

Delay-Tolerant Networking
Internet-Draft
Intended status: Informational
Expires: May 3, 2018

E. Birrane
Johns Hopkins Applied Physics Laboratory
October 30, 2017

Asynchronous Management Architecture
draft-birrane-dtn-ama-06

Abstract

This document describes an asynchronous management architecture (AMA) suitable for providing application-level network management services in a challenged networking environment. Challenged networks are those that require fault protection, configuration, and performance reporting while unable to provide humans-in-the-loop with synchronous feedback or otherwise preserve transport-layer sessions. In such a context, networks must exhibit behavior that is both determinable and autonomous while maintaining compatibility with existing network management protocols and operational concepts.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on May 3, 2018.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect

to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
1.1. Purpose	3
1.2. Scope	3
1.3. Requirements Language	4
1.4. Organization	4
2. Terminology	5
3. Motivation	6
3.1. Challenged Networks	6
3.2. Current Approaches and Their Limitations	7
4. Service Definitions	8
4.1. Configuration	9
4.2. Reporting	9
4.3. Autonomous Parameterized Procedure Calls	10
4.4. Administration	10
5. Desirable Properties	11
5.1. Intelligent Push of Information	11
5.2. Minimize Message Size Not Node Processing	12
5.3. Absolute Data Identification	12
5.4. Custom Data Definition	12
5.5. Autonomous Operation	13
6. Roles and Responsibilities	13
6.1. Agent Responsibilities	14
6.2. Manager Responsibilities	15
7. Logical Data Model	15
7.1. EDDs, VARs, and Reporting	16
7.2. Controls and Macros	17
7.3. Rules	17
7.4. Operators and Literals	18
8. System Model	18
8.1. Control and Data Flows	18
8.2. Control Flow by Role	19
8.2.1. Notation	19
8.2.2. Serialized Management	20
8.2.3. Multiplexed Management	21
8.2.4. Data Fusion	23
9. IANA Considerations	24
10. Security Considerations	24
11. Informative References	24
Author's Address	25

1. Introduction

The Asynchronous Management Architecture (AMA) provides application-layer network management services over links where delivery delays prevent timely communications between a network operator and a managed device. These delays may be caused by long signal propagations or frequent link disruptions (such as described in [RFC4838]) or by non-environmental factors such as unavailability of network operators, administrative delays, or delays caused by quality-of-service prioritizations and service-level agreements.

1.1. Purpose

This document describes the motivation, service definitions, desirable properties, roles/responsibilities, system model, and logical data model that form the AMA. These descriptions should be of sufficient specificity that implementations conformant to this architecture will operate successfully in a challenged networking environment.

This document is not a prescriptive standardization of a physical data model or protocol. Instead, it serves as informative guidance to authors of such models and protocols.

An AMA is necessary as the assumptions inherent to the architecture and design of synchronous management tools and techniques are not valid in challenged network scenarios. In these scenarios, synchronous approaches either patiently wait for periods of bi-directional connectivity or require the investment of significant time and resources to evolve a challenged network into a well-connected, low-latency network. In some cases such evolution is merely a costly way to over-resource a network. In other cases, such evolution is impossible given physical limitations imposed by signal propagation delays, power, transmission technologies, and other phenomena. Asynchronous management of asynchronous networks enables large-scale deployments, distributed technical capabilities, and reduced deployment and operations costs.

The rationale and motivation for asynchronous management is captured in [BIRrane1], [BIRrane2],[BIRrane3]. The properties and feasibility of such a system are taken from prototyping work done in accordance with [I-D.irtf-dtnrg-dtnmp].

1.2. Scope

It is assumed that any challenged network where network management would be usefully applied supports basic services (where necessary) such as naming, addressing, integrity, confidentiality,

authentication, fragmentation, and traditional network/session layer functions. Therefore, these items are outside of the scope of the AMA and not covered in this document.

While possible that a challenged network may interface with an unchallenged network, this document does not address the concept of network management compatibility with synchronous approaches.

1.3. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

1.4. Organization

The remainder of this document is organized into seven sections that, together, describe an AMA suitable for enterprise management of asynchronous networks: terminology, motivation, service definitions, desirable properties, roles/responsibilities, logical data model, and system model. The description of each section is as follows.

- o Terminology - This section identifies those terms critical to understanding the proper operation of the AMA. Whenever possible, these terms align in both word selection and meaning with their analogs from other management protocols.
- o Motivation - This section provides an overall motivation for this work as providing a novel and useful alternative to current network management approaches. Specifically, this section describes common network functions and how synchronous mechanisms fail to provide these functions in an asynchronous environment.
- o Service Definitions - This section defines asynchronous network management services in terms of terminology, scope, and impact.
- o Desirable Properties - This section identifies the properties to which an asynchronous management system should adhere to effectively implement service definitions in an asynchronous environment. These properties guide the subsequent definition of the system and logical models that comprise the AMA.
- o Roles and Responsibilities - This section identifies the roles in the AMA and their associated responsibilities. It provides the terminology and context for discussing how network management services interact.

- o Logical Data Model - This section describes the kinds of data that should be represented in deployment asynchronous management system.
- o System Model - This section describes data flows amongst various defined Actor roles. These flows capture how the AMA system works to provide asynchronous network management services in accordance with defined desirable properties.

2. Terminology

- o Actor - A software service running on either managed or managing devices for the purpose of implementing management protocols between such devices. Actors may implement the "Manager" role, "Agent" role, or both.
- o Agent Role (or Agent) - The role associated with a managed device, responsible for reporting performance data, enforcing administrative policies, and accepting/performing actions. Agents exchange information with Managers operating either on the same device or on a remote managing device.
- o Externally Defined Data (EDD) - Information made available to an Agent by a managed device, but not computed directly by the Agent.
- o Variable (VAR) - Information that is computed by an Agent, typically as a function of EDD values and/or other Variables.
- o Controls (CTRLs) - Operations that may be undertaken by an Actor to change the behavior, configuration, or state of an application or protocol managed by an AMP.
- o Literals (LIT) - Constants, enumerations, and other immutable definitions.
- o Macros - A named, ordered collection of Controls.
- o Manager - A role associated with a managing device responsible for configuring the behavior of, and receiving information from, Agents. Managers interact with one or more Agents located on the same device and/or on remote devices in the network.
- o Operator (OP) - The enumeration and specification of a mathematical function used to calculate computed data definitions and construct expressions to calculate state.

- o Report (RPT) - A named, typed, ordered collection of data values gathered by one or more Agents and provided to one or more Managers.
- o Rule - A unit of autonomous specification that provides a stimulus-response relationship between time or state on an Agent and the Controls to be run as a result of that time or state.

3. Motivation

Challenged networks, to include networks challenged by administrative or policy delays, cannot guarantee capabilities required to enable synchronous management techniques. These capabilities include high-rate, highly-available data, round-trip data exchange, and operators "in-the-loop". The inability of current approaches to provide network management services in a challenged network motivates the need for a new network management architecture focused on asynchronous, open-loop, autonomous control of network components.

3.1. Challenged Networks

A growing variety of link-challenged networks support packetization to increase data communications reliability without otherwise guaranteeing a simultaneous end-to-end path. Examples of such networks include Mobile Ad-Hoc Networks (MANets), Vehicular Ad-Hoc Networks (VANets), Space-Terrestrial Internetworks (STINTs), and heterogeneous networking overlays. Links in such networks are often unavailable due to attenuations, propagation delays, occultation, and other limitations imposed by energy and mass considerations. Data communications in such networks rely on store-and-forward and other queuing strategies to wait for the connectivity necessary to usefully advance a packet along its route.

Similarly, there also exist well-resourced networks that incur high message delivery delays due to non-environmental limitations. For example, networks whose operations centers are understaffed or where data volume and management requirements exceed the real-time cognitive load of operators or the associated operations console software support. Also, networks where policy restricts user access to existing bandwidth creates situations functionally similar to link disruption and delay.

Independent of the reason, when a node experiences an inability to communicate it must rely on autonomous mechanisms to ensure its safe operation and ability to usefully re-join the network at a later time. In cases of sparsely-populated networks, there may never be a practical concept of "the connected network" as most nodes may be disconnected most of the time. In such environments, defining a

network in terms of instantaneous connectivity becomes impractical or impossible.

Specifically, challenged networks exhibit the following properties that may violate assumptions built into current approaches to synchronous network management.

- o Links may be uni-directional.
- o Bi-directional links may have asymmetric data rates.
- o No end-to-end path is guaranteed to exist at any given time between any two nodes.
- o Round-trip communications between any two nodes within any given time window may be impossible.

3.2. Current Approaches and Their Limitations

Network management tools in unchallenged networks provide mechanisms for communicating locally-collected data from Agents to Managers, typically using a "pull" mechanism where data must be explicitly requested by a Manager in order to be transmitted by an Agent.

Management approaches that rely on timely data exchange, such as those that rely on negotiated sessions or other synchronized acknowledgment, do not function in challenged network environments. Familiar examples of TCP/IP based management via closed-loop, synchronous messaging do not work when network disruptions increase in frequency and severity. While no protocol delivers data in the absence of a networking link, protocols that eliminate or drastically reduce overhead and end-point coordination require smaller transmission windows and continue to function when confronted with scaling delays and disruptions in the network.

A legacy method for management in unchallenged networks today is the Simple Network Management Protocol (SNMP) [RFC3416]. SNMP utilizes a request/response model to set and retrieve data values such as host identifiers, link utilizations, error rates, and counters between application software on Agents and Managers. Data may be directly sampled or consolidated into representative statistics. Additionally, SNMP supports a model for asynchronous notification messages, called traps, based on predefined triggering events. Thus, Managers can query Agents for status information, send new configurations, and be informed when specific events have occurred. Traps and queryable data are defined in one or more Managed Information Bases (MIBs) which define the information for a particular data standard, protocol, device, or application.

While there is a large installation base for SNMP there are several aspects of the protocol that make it inappropriate for use in a challenged networking environment. SNMP relies on sessions with low round-trip latency to support its "pull" model. The SNMP trap model provides some Agent-side processing, but with very low fidelity and traps are typically "fire and forget" requiring the underlying transport to support reliable, in-order message delivery. Adaptive modifications to SNMP to support challenged networks would alter the basic function of the protocol (data models, control flows, and syntax) so as to be functionally incompatible with existing SNMP installations. Therefore, this approach is not suitable for an asynchronous network management system.

The Network Configuration Protocol (NETCONF) provides device-level configuration capabilities [RFC6241] to replace vendor-specific command line interface (CLI) configuration software. The XML-based protocol provides a remote procedure call (RPC) syntax such that any exposed functionality on an Agent can be exercised via a software application interface. NETCONF places no specific functional requirements or constraints on the capabilities of the Agent, which makes it a very flexible tool for configuring a homogeneous network of devices.

NETCONF places specific constraints on any underlying transport protocol: a long-lived, reliable, low-latency sequenced data delivery session. This is a fundamental requirement given the RPC-nature of the operating concept, and it is unsustainable in a challenged network. Aspects of the data modeling associated with NETCONF may apply to an asynchronous network management system, such that some modeling tools may be used, even if the network control plane cannot.

Just as the concept of a loosely-confederated set of nodes changes the definition of a network, it also changes the operational concept of what it means to manage a network. When a network stops being a single entity exhibiting a single behavior, "network management" becomes large-scale "node management". Individual nodes must share the burden of implementing desirable behavior without reliance on a single oracle of configuration or other coordinating function such as an operator-in-the-loop.

4. Service Definitions

This section identifies the services that must exist between Managers and Agents within an AMA. These services include configuration, reporting, parameterized control, and administration.

4.1. Configuration

Configuration services update Agent data associated with managed applications and protocols. Some configuration data might be defined in the context of an application or protocol, such that any network using that application or protocol would understand that data. Other configuration data may be defined tactically for use in a specific network deployment and not available to other networks even if they use the same applications or protocols.

New configurations received by an Agent must be validated to ensure that they do not conflict with other configurations or would otherwise prevent the Agent from effectively working with other Actors in its region. With no guarantee of round-trip data exchange, Agents cannot rely on remote Managers to correct erroneous or stale configurations from harming the flow of data through a challenged network.

Examples of configuration service behavior include the following.

- o Creating a new datum as a function of other well-known data:
C = A + B.
- o Creating a new report as a unique, ordered collection of known data:
RPT = {A, B, C}.
- o Storing predefined, parameterized responses to potential future conditions:
IF (X > 3) THEN RUN CMD(PARM).

4.2. Reporting

Reporting services populate report templates with values collected or computed by an Agent. The resultant reports are sent to one or more Managers by the Agent. The term "reporting" is used in place of the term "monitoring", as monitoring implies a timeliness and regularity that cannot be guaranteed by a challenged network. Reports sent by an Agent provide best-effort information to receiving Managers.

Since a Manager is not actively "monitoring" an Agent, the Agent must make its own determination on when to send what Reports based on its own local time and state information. Agents should produce Reports of varying fidelity and with varying frequency based on thresholds and other information set as part of configuration services.

Examples of reporting service behavior include the following.

- o Generate Report R1 every hour (time-based production).
- o Generate Report R2 when $X > 3$ (state-based production).

4.3. Autonomous Parameterized Procedure Calls

Similar to an RPC call, some mechanism MUST exist to allow a procedure to be run on an Agent to effect behavior or otherwise change the Agent's internal state. Since there is no guarantee that a Manager will be in contact with an Agent at any given time, the decisions of whether and when a procedure should be run MUST be made locally and autonomously by the Agent. Two types of automation triggers are identified in the AMA: triggers based on the general state of the Agent and triggers based on an Agent's notion of time. As such, the autonomous execution of procedures can be viewed as a stimulus-response system, where the stimulus is the positive evaluation of a state or time based predicate and the response is the function to be executed.

The autonomous nature of procedure execution by an Agent implies that the full suite of information necessary to run a procedure may not be known by a Manager in advance. To address this situation, a parameterization mechanism MUST be available so that required data can be provided at the time of execution on the Agent rather than at the time of definition/configuration by the Manager.

Autonomous, parameterized procedure calls provide a powerful mechanism for Managers to "manage" an Agent asynchronously during periods of no communication by pre-configuring responses to events that may be encountered by the Agent at a future time.

Examples of potential behavior include the following.

- o Updating local routing information based on instantaneous link analysis.
- o Managing storage on the device to enforce quotas.
- o Applying or modifying local security policy.

4.4. Administration

Administration services enforce the potentially complex mapping of configuration, reporting, and control services amongst Agents and Managers in the network. Fine-grained access control specifying which Managers may apply which services to which Agents may be necessary in networks dealing with multiple administrative entities or overlay networks crossing multiple administrative boundaries.

Whitelists, blacklists, key-based infrastructures, or other schemes may be used for this purpose.

Examples of administration service behavior include the following.

- o Agent A1 only Sends reports for Protocol P1 to Manager M1.
- o Agent A2 only accepts a configurations for Application Y from Managers M2 and M3.
- o Agent A3 accepts services from any Manager providing the proper authentication token.

Note that the administrative enforcement of access control is different from security services provided by the networking stack carrying AMP messages.

5. Desirable Properties

This section describes those design properties that are desirable when defining an architecture that must operate across challenged links in a network. These properties ensure that network management capabilities are retained even as delays and disruptions in the network scale. Ultimately, these properties are the driving design principles for the AMA.

5.1. Intelligent Push of Information

Pull management mechanisms require that a Manager send a query to an Agent and then wait for the response to that query. This practice implies a control-session between entities and increases the overall message traffic in the network. Challenged networks cannot guarantee timely roundtrip data-exchange and, in extreme cases, are comprised solely of uni-directional links. Therefore, pull mechanisms must be avoided in favor of push mechanisms.

Push mechanisms, in this context, refer to Agents making their own determinations relating to the information that should be sent to Managers. Such mechanisms do not require round-trip communications as Managers do not request each reporting instance; Managers need only request once, in advance, that information be produced in accordance with a predetermined schedule or in response to a predefined state on the Agent. In this way information is "pushed" from Agents to Managers and the push is "intelligent" because it is based on some internal evaluation performed by the Agent.

5.2. Minimize Message Size Not Node Processing

Protocol designers must balance message size versus message processing time at sending and receiving nodes. Verbose representations of data simplify node processing whereas compact representations require additional activities to generate/parse the compacted message. There is no asynchronous management advantage to minimizing node processing time in a challenged network. However, there is a significant advantage to smaller message sizes in such networks. Compact messages require smaller periods of viable transmission for communication, incur less re-transmission cost, and consume less resources when persistently stored en-route in the network. AMPs should minimize PDUs whenever practical, to include packing and unpacking binary data, variable-length fields, and pre-configured data definitions.

5.3. Absolute Data Identification

Elements within the management system must be uniquely identifiable so that they can be individually manipulated. Identification schemes that are relative to system configuration make data exchange between Agents and Managers difficult as system configurations may change faster than nodes can communicate.

Consider the following common technique for approximating an associative array lookup. A manager wishing to do an associative lookup for some key K1 will (1) query a list of array keys from the agent, (2) find the key that matches K1 and infer the index of K1 from the returned key list, and (3) query the discovered index on the agent to retrieve the desired data.

Ignoring the inefficiency of two pull requests, this mechanism fails when the Agent changes its key-index mapping between the first and second query. Rather than constructing an artificial mapping from K1 to an index, an AMP must provide an absolute mechanism to lookup the value K1 without an abstraction between the Agent and Manager.

5.4. Custom Data Definition

Custom definition of new data from existing data (such as through data fusion, averaging, sampling, or other mechanisms) provides the ability to communicate desired information in as compact a form as possible. Specifically, an Agent should not be required to transmit a large data set for a Manager that only wishes to calculate a smaller, inferred data set. The Agent should calculate the smaller data set on its own and transmit that instead. Since the identification of custom data sets is likely to occur in the context

of a specific network deployment, AMPs must provide a mechanism for their definition.

5.5. Autonomous Operation

AMA network functions must be achievable using only knowledge local to the Agent. Rather than directly controlling an Agent, a Manager configures an engine of the Agent to take its own action under the appropriate conditions in accordance with the Agent's notion of local state and time.

Such an engine may be used for simple automation of predefined tasks or to support semi-autonomous behavior in determining when to run tasks and how to configure or parameterize tasks when they are run. Wholly autonomous operations MAY be supported where required. Generally, autonomous operations should provide the following benefits.

- o Distributed Operation - The concept of pre-configuration allows the Agent to operate without regular contact with Managers in the system. The initial configuration (and periodic update) of the system remains difficult in a challenged network, but an initial synchronization on stimuli and responses drastically reduces needs for centralized operations.
- o Deterministic Behavior - Such behavior is necessary in critical operational systems where the actions of a platform must be well understood even in the absence of an operator in the loop. Depending on the types of stimuli and responses, these systems may be considered simple automation or semi-autonomous behavior, both of which imply the ability of a frequently-out-of-contact Manager to better predict the state of an Agent than if controls were to be run by an independent, fully autonomous system.
- o Engine-Based Behavior - Several operational systems are unable to deploy "mobile code" based solutions due to network bandwidth, memory or processor loading, or security concerns. Engine-based approaches are preferred as they can be flexible without incurring a set of problematic requirements or concerns.

6. Roles and Responsibilities

By definition, Agents reside on managed devices and Managers reside on managing devices. This section describes how these roles participate in the network management functions outlined in the prior section.

6.1. Agent Responsibilities

Application Support

Agents MUST collect all data, execute all procedures, populate all reports and run operations required by each application which the Agent claims to manage. Agents MUST report supported applications so that Managers in a network understands what information is understood by what Agent.

Local Data Collection

Agents MUST collect from local firmware (or other on-board mechanisms) and report all data defined for the management of applications for which they have been configured.

Autonomous Control

Agents MUST determine, without Manager intervention, whether a procedure should be invoked. Agents MAY also invoke procedures on other devices for which they act as proxy.

User Data Definition

Agents MUST provide mechanisms for operators in the network to use configuration services to create customized data definitions in the context of a specific network or network use-case. Agents MUST allow for the creation, listing, and removal of such definitions in accordance with whatever security models are deployed within the particular network.

Where applicable, Agents MUST verify the validity of these definitions when they are configured and respond in a way consistent with the logging/error-handling policies of the Agent and the network.

Autonomous Reporting

Agents MUST determine, without real-time Manager intervention, whether and when to populate and transmit a given report targeted to one or more Managers in the network.

Consolidate Messages

Agents SHOULD produce as few messages as possible when sending information. For example, rather than sending multiple messages, each with one report to a Manager, an Agent SHOULD prefer to send a single message containing multiple reports.

Regional Proxy

Agents MAY perform any of their responsibilities on behalf of other network nodes that, themselves, do not have an Agent.

In such a configuration, the Agent acts as a proxy for these other network nodes.

6.2. Manager Responsibilities

Agent Capabilities Mapping

Managers MUST understand what applications are managed by the various Agents with which they communicate. Managers should not attempt to request, invoke, or refer to application information for applications not managed by an Agent.

Data Collection

Managers MUST receive information from Agents by asynchronously configuring the production of reports and then waiting for, and collecting, responses from Agents over time. Managers MAY try to detect conditions where Agent information has not been received within operationally relevant time spans and react in accordance with network policy.

Custom Definitions

Managers should provide the ability to define custom data definitions. Any custom definitions MUST be transmitted to appropriate Agents and these definitions MUST be remembered to interpret the reporting of these custom values from Agents in the future.

Data Translation

Managers should provide some interface to other network management protocols. Managers MAY accomplish this by accumulating a repository of push-data from high-latency parts of the network from which data may be pulled by low-latency parts of the network.

Data Fusion

Managers MAY support the fusion of data from multiple Agents with the purpose of transmitting fused data results to other Managers within the network. Managers MAY receive fused reports from other Managers pursuant to appropriate security and administrative configurations.

7. Logical Data Model

The AMA logical data model captures the types of information that should be collected and exchanged to implement necessary roles and responsibilities. The data model presented in this section does not presuppose a specific mapping to a physical data model or encoding technique; it is included to provide a way to logically reason about

the types of data that should be exchanged in an asynchronously managed network.

The elements of the AMA logical data model are described as follows.

7.1. EDDs, VARs, and Reporting

There are three fundamental representations of data in the AMA: (1) data which are sampled/calculated external to the network management system, (2) data which are calculated internal to the network management system, and (3) ordered collections of data items used for reporting.

Data that is sampled/calculated external to the network management system is defined as "externally defined data" (EDD). EDD values represent the most useful information in the management system as they are provided by the applications or protocol being managed on the Agent. It is RECOMMENDED that EDD values be strongly typed to avoid issues with interpreting the data value. It is also RECOMMENDED that the timeliness/staleness of the data value be considered when using the data in the context of autonomous action on the Agent.

Data that is calculated internal to the network management system is defined as a "variable" (VAR). VARs allow the creation of new data values for use in the network management system. New value definitions are useful for storing user-defined information, storing the results of complex calculations for easier re-use, and providing a mechanism for combining information from multiple external sources. It is RECOMMENDED that VARs be strongly typed to avoid issues with interpreting the data value. In cases where a VAR definition relies on other VAR definitions, mechanisms to prevent circular references MUST be included in any actual data model or implementation.

Ordered collections of EDD values and VARs should be produced by Agents and sent to Managers as a way of communicating Agent state. Such an ordered collection is called a "report" (RPT). It is RECOMMENDED that the structure of a RPT be given in a template that can be synchronized between an Agent and a Manager so that RPTs themselves do not need to be self-describing. A RPT may include EDD values, VARs, and also other RPTs. In cases where a RPT includes another RPT, mechanisms to prevent circular references MUST be included in any actual data model or implementation.

In cases where sets of data can comply to a tabular format, a Table structure should be used instead of a report structure. In these cases, table columns provide name and type information that does not need to be repeated in every data value. However, unlike

reports, the regularity of the table structure allows for the addition of rows without altering the format of the table structure. Attempting to add a new data set at the end of a report would require alterations to the report template. Additionally, the prescriptive nature of the Table structure allows for the possibility of advanced filtering which may reduce traffic between Agents and Managers.

7.2. Controls and Macros

Low-latency, high-availability approaches to network management use mechanisms such as (or similar to) remote procedure calls (RPCs) to cause some action to be performed on an Agent. The AMA requires similar capabilities, though without requiring that the Manager be in the processing loop of the Agent.

A "control" (CTRL) represents a parameterized, predefined procedure that can be run on an Agent. CTRLs do not have a return code as there is not the same concept of sequential execution in an asynchronous model. Parameters can be provided when running a command from a Manager, pre-configured as part of an autonomy response on the Agent, or auto-generated as needed on the Agent. The success or failure of a control MAY be inferred by reports generated for that purpose.

Often, a series of controls must be executed in concert to achieve a particular outcome. A "macro" (MACRO) represents an ordered collection of controls (or other macros). In cases where a MACRO includes another MACRO, mechanisms to prevent circular references and maximum nesting levels MUST be included in any actual data model or implementation.

7.3. Rules

The AMA data model contains EDD values and VARs that capture the state of applications on an Agent. The model also contains controls and macros to perform actions on an Agent. A mechanism is needed to relate these two capabilities; to perform an action on the Agent in response to the state of the Agent.

One way of mapping Agent state to Agent actions is via a stimulus-response system. A "rule" represents a stimulus-response pairing in the following form.

IF predicate THEN response

The predicate is a logical expression that evaluates to true if the rule stimulus is present and evaluates to false otherwise. The response may be any control or macro known to the Agent. An example of a time-based predicate is to perform some activity every 24 hours (e.g., `((CUR_TIME - START_TIME) % 24Hrs) == 0`). An example of a

state-based predicate is to perform some activity if a given EDD value exceeds a predefined threshold such as a measured temperature exceeds 80 degrees centigrade (e.g., (TEMP > 80.0))

Rules should be allowed to construct their stimuli from the full set of EDD values and VARs available to the network management system. Similarly, macro responses should be allowed to include controls from all applications known by the Agent. This enables an expressive capability to have multiple applications monitored and managed by the Agent.

7.4. Operators and Literals

Actions such as computing a VAR value or describing a rule predicate require calculating mathematical expressions. An element of an expression will be one of four types of data: an EDD value, a VAR value, a mathematical operations, and literal values.

An "operator" (OP) represents a mathematical operation in an expression. OPs should support multiple operands based on the operation supported. A common set of OPs SHOULD be defined for any Agent and systems MAY choose to allow individual applications to define new OPs to assist in the generation of new VAR values and predicates for managing that application. OPs may be simple binary operations such as "A + B" or more complex functions such as sin(A) or avg(A,B,C,D).

A "literal" (LIT) represents a constant value, such as simple numbers (e.g., 4), well-known mathematical numbers (e.g., PI, E), or other useful data such as Epoch times. LITs should be strongly typed to avoid any misinterpretation of their data value.

8. System Model

This section describes the notional data flows and control flows that illustrate how Managers and Agents within an AMA cooperate to perform network management services.

8.1. Control and Data Flows

The AMA identifies three significant data flows: control flows from Managers to Agents, reports flows from Agents to Managers, and fusion reports from Managers to other Managers. These data flows are illustrated in Figure 1.

AMA Control and Data Flows

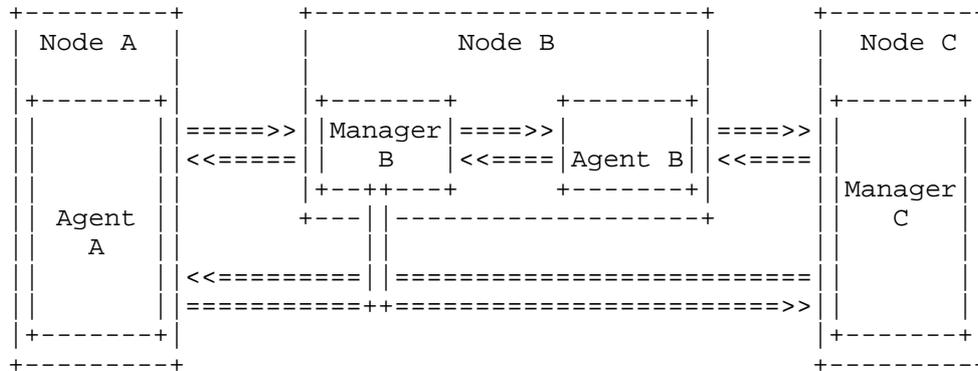


Figure 1

In this data flow, the Agent on node A receives Controls from Managers on nodes B and C, and replies with Report Entries back to these Managers. Similarly, the Agent on node B interacts with the local Manager on node B and the remote Manager on node C. Finally, the Manager on node B may fuse Report Entries received from Agents at nodes A and B and send these fused Report Entries back to the Manager on node C.

From this figure it is clear that there exist many-to-many relationships amongst Managers, amongst Agents, and between Agents and Managers. Note that Agents and Managers are roles, not necessarily differing software applications. Node A may represent a single software application fulfilling only the Agent role, whereas node B may have a single software application fulfilling both the Agent and Manager roles. The specifics of how these roles are realized is an implementation matter.

8.2. Control Flow by Role

This section describes three common configurations of Agents and Managers and the flow of messages between them. These configurations involve local and remote management and data fusion.

8.2.1. Notation

The notation outlined in Table 1 describes the types of control messages exchanged between Agents and Managers.

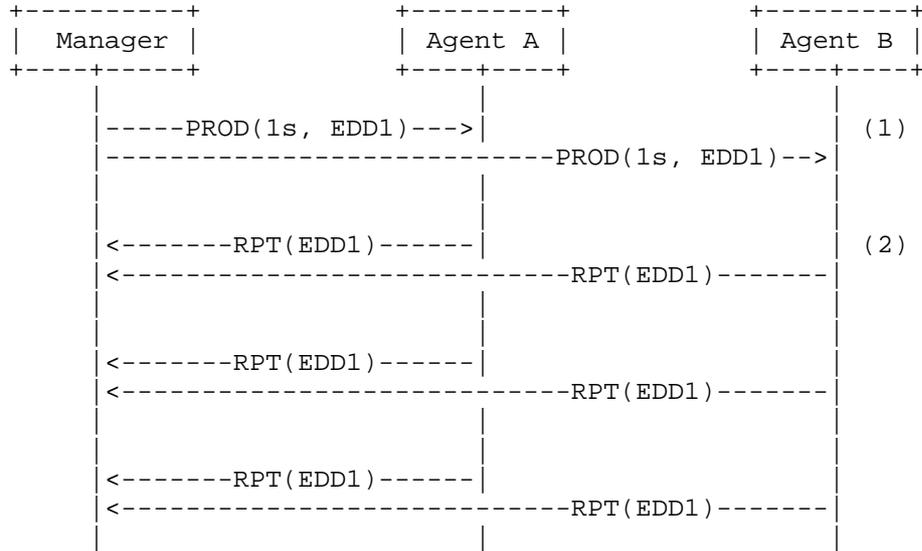
Term	Definition	Example
EDD#	EDD definition.	EDD1
V#	Custom data definition.	V1 = EDD1 + V0.
DEF([ACL], ID,EXPR)	Define id from expression. Allow managers in access control list (ACL) to request this id.	DEF([*], V1, EDD1 + EDD2)
PROD(P, ID)	Produce ID according to predicate P. P may be a time period (1s) or an expression (EDD1 > 10).	PROD(1s, EDD1)
RPT(ID)	A report identified by ID.	RPT(EDD1)

Table 1: Terminology

8.2.2. Serialized Management

This is a nominal configuration of network management where a Manager interacts with a set of Agents. The control flows for this are outlined in Figure 2.

Serialized Management Control Flow



In a simple network, a Manager interacts with multiple Agents.

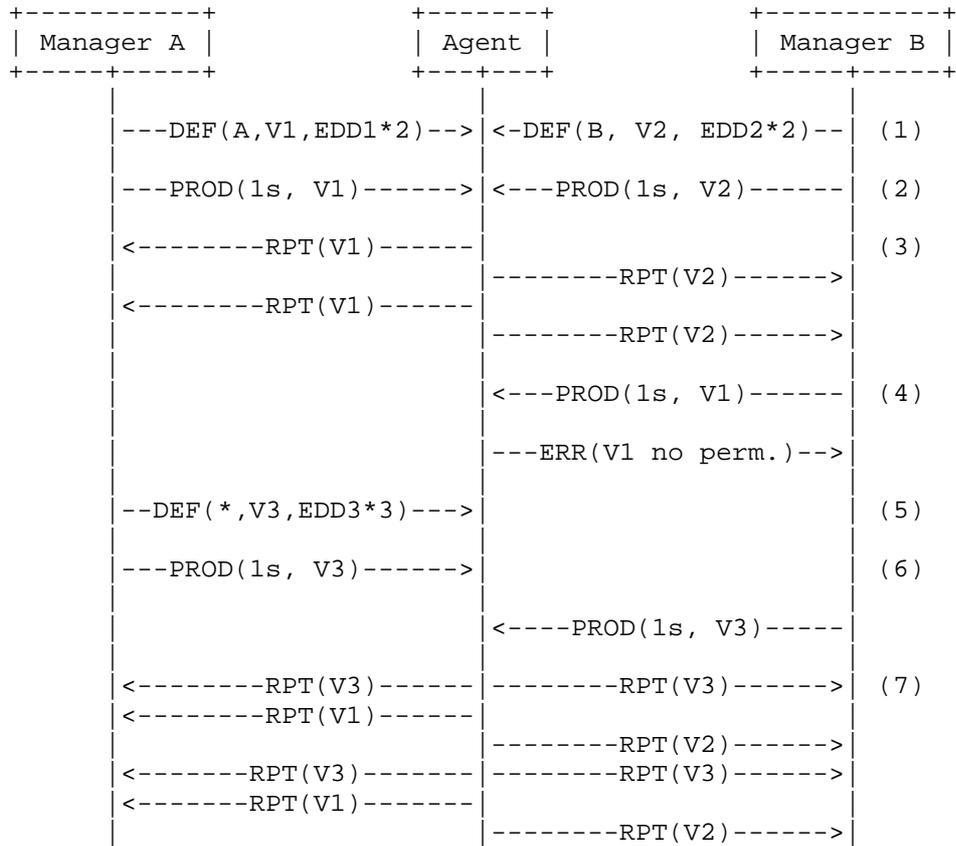
Figure 2

In this figure, the Manager configures Agents A and B to produce EDD1 every second in (1). At some point in the future, upon receiving and configuring this message, Agents A and B then build a Report Entry containing EDD1 and send those reports back to the Manager in (2).

8.2.3. Multiplexed Management

Networks spanning multiple administrative domains may require multiple Managers (for example, one per domain). When a Manager defines custom Reports/Variables to an Agent, that definition may be tagged with an access control list (ACL) to limit what other Managers will be privy to this information. Managers in such networks should synchronize with those other Managers granted access to their custom data definitions. When Agents generate messages, they MUST only send messages to Managers according to these ACLs, if present. The control flows in this scenario are outlined in Figure 3.

Multiplexed Management Control Flow



Complex networks require multiple Managers interfacing with Agents.

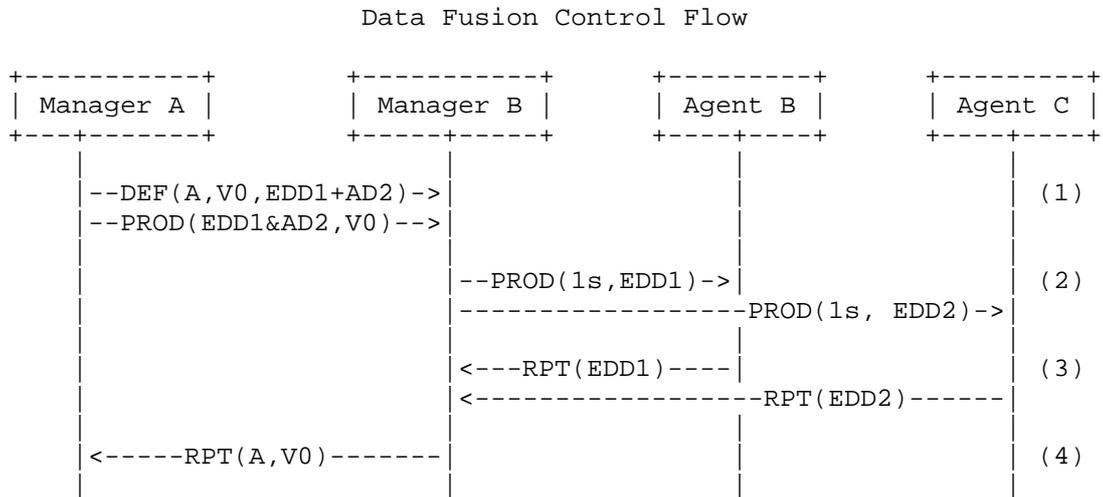
Figure 3

In more complex networks, any Manager may choose to define custom Reports and Variables, and Agents may need to accept such definitions from multiple Managers. Variable definitions may include an ACL that describes who may query and otherwise understand these definitions. In (1), Manager A defines V1 only for A while Manager B defines V2 only for B. Managers may, then, request the production of Report Entries containing these definitions, as shown in (2). Agents produce different data for different Managers in accordance with configured production rules, as shown in (3). If a Manager requests the production of a custom definition for which the Manager has no permissions, a response consistent with the configured logging policy on the Agent should be implemented, as shown in (4). Alternatively,

as shown in (5), a Manager may define custom data with no restrictions allowing all other Managers to request and use this definition. This allows all Managers to request the production of Report Entries containing this definition, shown in (6) and have all Managers receive this and other data going forward, as shown in (7).

8.2.4. Data Fusion

In some networks, Agents do not individually transmit their data to a Manager, preferring instead to fuse reporting data with local nodes prior to transmission. This approach reduces the number and size of messages in the network and reduces overall transmission energy expenditure. The AMA supports fusion of NM reports by co-locating Agents and Managers on nodes and offloading fusion activities to the Manager. This process is illustrated in Figure 4.



Data fusion occurs amongst Managers in the network.

Figure 4

In this example, Manager A requires the production of a Variable V0, from node B, as shown in (1). The Manager role understands what data is available from what agents in the subnetwork local to B, understanding that EDD1 is available locally and EDD2 is available remotely. Production messages are produced in (2) and data collected in (3). This allows the Manager at node B to fuse the collected Report Entries into V0 and return it in (4). While a trivial example, the mechanism of associating fusion with the Manager function rather than the Agent function scales with fusion complexity, though it is important to reiterate that Agent and

Manager designations are roles, not individual software components. There may be a single software application running on node B implementing both Manager B and Agent B roles.

9. IANA Considerations

At this time, this protocol has no fields registered by IANA.

10. Security Considerations

Security within an AMA MUST exist in two layers: transport layer security and access control.

Transport-layer security addresses the questions of authentication, integrity, and confidentiality associated with the transport of messages between and amongst Managers and Agents in the AMA. This security is applied before any particular Actor in the system receives data and, therefore, is outside of the scope of this document.

Finer grain application security is done via ACLs which are defined via configuration messages and implementation specific.

11. Informative References

[BIRrane1]

Birrane, E. and R. Cole, "Management of Disruption-Tolerant Networks: A Systems Engineering Approach", 2010.

[BIRrane2]

Birrane, E., Burleigh, S., and V. Cerf, "Defining Tolerance: Impacts of Delay and Disruption when Managing Challenged Networks", 2011.

[BIRrane3]

Birrane, E. and H. Kruse, "Delay-Tolerant Network Management: The Definition and Exchange of Infrastructure Information in High Delay Environments", 2011.

[I-D.irtf-dtnrg-dtnmp]

Birrane, E. and V. Ramachandran, "Delay Tolerant Network Management Protocol", draft-irtf-dtnrg-dtnmp-01 (work in progress), December 2014.

[RFC2119]

Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.

- [RFC3416] Presuhn, R., Ed., "Version 2 of the Protocol Operations for the Simple Network Management Protocol (SNMP)", STD 62, RFC 3416, DOI 10.17487/RFC3416, December 2002, <<https://www.rfc-editor.org/info/rfc3416>>.
- [RFC4838] Cerf, V., Burleigh, S., Hooke, A., Torgerson, L., Durst, R., Scott, K., Fall, K., and H. Weiss, "Delay-Tolerant Networking Architecture", RFC 4838, April 2007.
- [RFC6241] Enns, R., Ed., Bjorklund, M., Ed., Schoenwaelder, J., Ed., and A. Bierman, Ed., "Network Configuration Protocol (NETCONF)", RFC 6241, DOI 10.17487/RFC6241, June 2011, <<https://www.rfc-editor.org/info/rfc6241>>.

Author's Address

Edward J. Birrane
Johns Hopkins Applied Physics Laboratory

Email: Edward.Birrane@jhuapl.edu

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 8, 2018

S. Burleigh
D. Horres
JPL, Calif. Inst. Of Technology
K. Viswanathan
M. Benson
F. Templin
Boeing Research & Technology
October 05, 2017

Architecture for Delay-Tolerant Key Administration
draft-burleigh-dtnwg-dtka-00.txt

Abstract

Delay-Tolerant Key Administration (DTKA) is a system of public-key management protocols intended for use in Delay Tolerant Networking (DTN). This document outlines a DTKA proposal for space-based communications, which are characterized by long communication delays and planned communication contacts.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 8, 2018.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect

to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	2
1.1.	Motivation and Design Strategy	3
1.2.	Scope	3
1.3.	About This Document	3
1.4.	Related Documents	3
2.	Terminology	5
3.	High Level Architecture	5
3.1.	Application Domains	6
3.2.	System Entities	6
3.3.	System Interconnections	7
3.4.	Architectural Assumption on Communication	8
3.5.	System Security Configuration	8
4.	Detailed Design	9
4.1.	Message Formats	9
4.2.	Node Registration	11
4.3.	Key Revocation	13
4.4.	Key Roll-over	14
4.5.	Key Distribution	15
4.6.	Secure Communications	16
4.7.	Communication Stack View	17
5.	IANA Considerations	17
6.	Security Considerations	17
7.	References	17
7.1.	Normative References	17
7.2.	Informative References	18
	Authors' Addresses	18

1. Introduction

Delay-Tolerant Key Administration (DTKA) is a system of public-key management protocols intended for use in Delay Tolerant Networking (DTN) [RFC4838]. This document outlines a DTKA proposal for space-based communications, which are characterized by long communication delays and planned communication contacts. The proposal satisfies the requirements for DTN Security Key Management [I-D.templin-dtnskmreq].

1.1. Motivation and Design Strategy

In general, on-demand interactive communications, like client-server interactions, are not feasible in DTN's network model. Terrestrial public-key management protocols require on-demand interactions with remote computing nodes to distribute and validate public-keys. For example, terrestrial public-key management protocols require on-demand interactions with a remote trusted authority (Certificate Revocation List (CRL)) to determine if a given public-key certificate has been revoked or not. Therefore, such terrestrial public-key management protocols cannot be used in DTN.

Periodic and planned communications are an inherent property of space-based communication systems. Thus, the core principle of DTKA is to exploit this property of space-based communication systems in order to avoid the need for on-demand interactive communications for key management. Therefore, the design strategy for DTKA is to proactively distribute authenticated public-keys to all nodes in a given DTN instance in advance to ensure that keys will be available when needed even if there may be significant delays or disruptions. This design strategy is to be contrasted with protocols for terrestrial Public-Key Infrastructures, in which authenticated public-keys are exchanged interactively, just-in-time and on demand.

1.2. Scope

DTKA was originally designed for space-based DTN environments, but it could potentially be used in terrestrial DTN environments as well.

1.3. About This Document

This document describes the high-level architecture of DTKA and lists the architectural entities, their interactions, and system assumptions.

1.4. Related Documents

The following documents provide the necessary context for the high-level design described in this document.

RFC 4838 [RFC4838] describes the architecture for DTN and is titled, "Delay-Tolerant Networking Architecture." That document provides a high-level overview of DTN architecture and the decisions that underpin the DTN architecture.

RFC 5050 [RFC5050] describes the protocol and message formats for DTN and is titled, "Bundle Protocol Specification." That document provides the format of the network protocol message for DTN,

called a Bundle, along with descriptions of processes for generating, sending, forwarding, and receiving Bundles. It also specifies an encoding format called SDNV (Self-Delimiting Numeric Values) for use in DTN. Each bundle comprises a primary block, a payload block, and zero or more additional extension blocks. A node may receive and process a bundle even when the bundle contains one or more extension blocks that the node is not equipped to process.

RFC 6257 [RFC6257] is titled, "Bundle Security Protocol Specification." It specifies the message formats and processing rules for providing three types of security services to bundles, namely: confidentiality, integrity, and authentication. It does not specify mechanisms for key management. Rather, it assumes that cryptographic keys are somehow in place and then specifies how the keys shall be used to provide the security services. Additionally, it attempts to standardize a default cipher suite for DTN.

The revised Internet Draft [I-D.ietf-dtn-bpsec] for DTN communication security is titled, "Bundle Security Protocol Specification (bpsec)." When compared with RFC 6257, it is silent on concepts such as Security Regions, at-most-once-delivery option, and cipher suite specification. It deletes the Bundle Authentication Block and generalized the Payload Integrity and Payload Confidentiality Blocks to Block Integrity Block and Block Confidentiality Block. It provides more detailed specification for bundle canonicalization and rules for processing bundles received from other nodes. Like RFC 6257, the draft does not describe any key management mechanisms for DTN but assumes that a suitable key management mechanism shall be in place.

5050bis [I-D.ietf-dtn-bpbis] is an Internet Draft on standards track that intends to update RFC 5050. It introduces a new concept called "node ID" as distinguished from the existing concept of "endpoint ID": a single DTN endpoint may contain one or more nodes. It also migrates some primary block fields into extension blocks, making the primary block immutable. In the Security Considerations section, 5050bis explicitly describes end-to-end security using Block-Integrity-Block (BIB) and Block-Confidentiality-Block (BCB). It does not specify link-by-link security considerations to be part of the bundle protocol level using the Bundle-Authenticity-Block (BAB), which was described in RFC 6257. The convergence layers may provide link-by-link authentication instead of bundle protocol agent.

The Internet Draft for specifying requirements for DTN Key Management [I-D.templin-dtnskmreq] is titled, "DTN Security Key

Management - Requirements and Design." It sketches nine requirements and four design criteria for DTN Key Management system. The last two requirements are the need to support revocation in a delay tolerant manner. It also specifies the requirements for avoiding single points of failure and opportunities for the presence of multiple key management authorities.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119]. Lower case uses of these words are not to be interpreted as carrying RFC2119 significance.

3. High Level Architecture

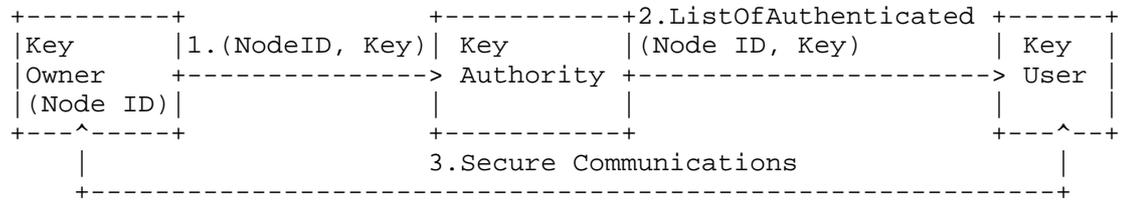


Figure 1: Abstract Data-Flow-Diagram for DTKA

The DTKA system includes Key Owners, Key Agents (which, in aggregate, constitute the Key Authority), and Key Users. For the sake of simplicity and to promote conceptual clarity, Figure 1 shows a single Key Agent. In order to avoid a single point-of-trust, DTKA provides mechanisms to distribute the Key Authority function among one or more DTKA Key Agents using an erasure-coding technique. This trust-distributing mechanism is discussed later in this document.

Each Key Owner has a unique DTN Node ID and chooses its own public-private key pair. In order to associate a public-key (Key) with its Node ID, a Key Owner sends an assertion of the form: (Node ID, Key) to the Key Authority. Key Owners need to authenticate their respective keys in one of two ways:

1. in the case of out-of-band bootstrapping, Key Authority shall rely on the physical security of the out-of-band channel to validate the integrity of the received message and the Key Owner needs to sign the assertion (Node ID, Key) using the private key corresponding to the Key in the assertion; or,

2. in the case of in-band authentication, the Key Owner needs to sign the assertion (Node ID, Key) using the private key corresponding to the previously authenticated and currently effective public-key for that NodeID.

Each Key User periodically receives a list of authenticated public-keys from the Key Authority and uses the authenticated public-keys as needed.

3.1. Application Domains

DTN can be used in various theatres such as space, airspace, on earth and at sea. There can be more than one installation of DTN in each of these theatres administered by different administrative entities, which may represent countries, companies and institutions. A particular installation of DTN with a single aggregate key authority is called an Application Domain.

3.2. System Entities

The architectural elements of DTKA, which shall henceforth be called DTKA Entities, are listed below.

DTKA Key Agent (DTKA-KA)

DTKA-KA is part of the root of trust for authenticated distribution of public-keys for a given application domain. All DTKA Entities must have physically authenticated public-keys of all DTKA Key Agents (DTKA-KAs), which together constitute the DTKA Key Authority for a given application domain.

DTKA Key Owner[Node ID] (DTKA-KO[Node ID])

DTKA-KO[Node ID] is a computing node that has possession of the private key corresponding to the public-key authenticated for a given Node Identity (Node ID) by the DTKA-KAs for the Key Owner's application domain.

DTKA Key User (DTKA-KU)

DTKA-KU is a computing node that receives authenticated public-keys from DTKA-KAs and distributes the same within a single computing machine through a suitable Interprocess Communication mechanism, which is outside the scope of this document.

DTKA Key Manager (DTKA-KM) and DTKA Key Manager Client (DTKA-KMC)

DTKA-KM is a DTKA Key User that receives authenticated public-keys from DTKA-KAs and distributes the same over a communication network to DTKA-KMCs, which are not DTKA Entities. DTKA-KMC can be a DTN node that can receive key distributions from DTKA KMs. The communication and security protocols for the interactions

run a reliable Time Synchronization Mechanism (TSM), like the Network Time Protocol (NTP) service. TSM shall ensure that time is synchronized between the DTKA-KAs that realize the DTKA Key Authority for a given application domain.

A potentially delayed and frequently disrupted communication link is assumed to interconnect DTKA-KAs, DTKA-KOs and DTKA-KUs. This delayed-and-disrupted communication link is used by the DTKA-KAs to multicast authenticated public-key assertions to DTKA-KUs. The DTKA-KUs are assumed to have access to autonomous clocks. Autonomous clocks keep time without external correction signals and with an allowed drift in the order of a few seconds. But, delay-tolerant mechanisms for clock agreement such as issuance of UTC offsets in network management messages may be present.

3.4. Architectural Assumption on Communication

In the subsequent sections, it shall be seen that DTKA-KAs shall dispatch updates to the list of authenticated public-keys in the system using erasure coding techniques. It is evident that at least a sub-set of such communications updates must reach each DTKA-KU. Therefore, the DTN upon which the DTKA operates must satisfy the following communication assumption before DTKA can function along expected lines: all addressed receivers MUST receive sufficient number of bundles from the DTKA-KAs before the earliest effective time among the effective times of all public-key assertions in the payloads of the bundles. Note that the underlying DTN will not be aware of the effective times of the public-key assertions in the payloads of the bundles.

The above assumption can be restated using DTKA protocol terminologies, which shall be seen in the subsequent sections, as follows: All addressed receivers MUST receive enough of the code blocks for a given bulletin to enable reassembly of that bulletin before the earliest effective-time among all assertions in the bulletin.

3.5. System Security Configuration

The current public-keys of all designated DTKA-KAs for a given application domain must be securely configured into every DTKA-KA and DTKA-KU that needs to participate in that application domain; this is a pre-condition for initializing those DTKA-Entities. This process will ensure that the DTKA Agents are established as the root of trust for that application domain.

4. Detailed Design

4.1. Message Formats

Every DTKA-KA in an application domain will receive requests for associating public-keys with Node IDs from the respective DTKA-KOs. After authenticating the requests, every DTKA-KA reaches consensus with all other DTKA-KAs, which constitute the Key Authority in its application domain, on some subset of the authenticated requests. Thereafter, each DTKA-KA must multicast to all participating DTKA Entities the subset of authenticated list of address-and-key assertions on which consensus was reached. The message format for this multicast, which is called a Bulletin, supports message authentication and redundancy. The goal of message authentication is to prevent DTKA Entities' acceptance of malicious multicast messages issued by hostile nodes. The goal of message redundancy is to ensure that a minimal set of collaborating DTKA-KAs in the application domain will be able to successfully send assertions or revocations for address-and-key associations to all DTKA Entities -- the DTKA Entities need not know which DTKA-KAs are not collaborating.

A bulletin is a collection of association blocks such that each association block represents a single association of a Node ID with a public-key as depicted in Figure 3. Each block issues either an assert or revoke or roll-over instruction to the receiving DTKA Entities, which use the key information message to execute the instruction locally. The block labelled "Bulletin Hash" contains the cryptographic hash computed over all association blocks in that bulletin.

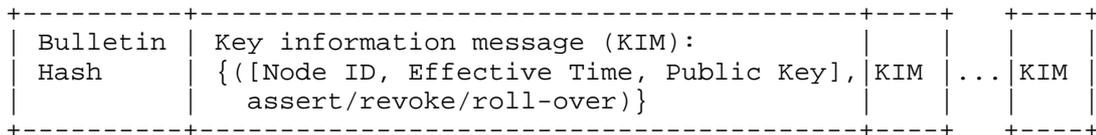


Figure 3: Bulletin

After forming a bulletin, a (Q+k)-erasure code algorithm is used to create an erasure code for the bulletin. Thus, receipt of any Q distinct code blocks will be sufficient to decode the bulletin. To ensure that the incapacity or compromise -- or veto (disagreement on bulletin content) -- of any single DTKA-KA will not result in malfunction of the key authority mechanism, each DTKA-KA is assigned primary responsibility for transmission of some limited subset of the bulletin's code blocks and backup responsibility for some other limited subset. The assigned code block subsets for the various DTKA-KAs are selected in such a way that every code block is to be

transmitted by two different DTKA-KAs. The combination of these two transmission redundancy mechanisms (parity code blocks and duplicate transmissions), together with reliable bundle transmission at the convergence layer under bundle multicast, minimizes the likelihood of any client node being unable to reconstruct the bulletin from the code blocks it receives.

During system initialization, the code-block assignments for each DTKA-KA need to be configured into every DTKA Entity. The code-block assignment for the example considered in this section is shown below in the table, in which an x-mark depicts the assignment of a code block to a DTKA-KA. It can be seen in the table that, in this example, code-blocks from at least five ($t=5$) DTKA-KAs must be received before the bulletin blocks can be decoded. Also, when all DTKA-KAs multicast their pre-defined code blocks, $n * m$ ($8 * 3 = 24$) code blocks are sent to all DTKA Entities. To further defend against a compromised DTKA-KA node introducing error into the key distribution system:

- o All nodes are informed of the code block subsets for which all DTKA-KA nodes are responsible. Any received code block that was transmitted by a DTKA-KA node which was not responsible for transmission of that code block is discarded by the receiving node.
- o Each bulletin issued by the aggregate KA is signed in a private key that is held in common by all DTKA-KA nodes. This signature functionally identifies the bulletin. Every transmitted code block is accompanied by the signature of the bulletin whose encoding DTKA-KA generated this code block. All - and only - code blocks tagged with a common signature are reassembled into the bulletin identified by that signature.
- o If the signature of a bulletin reassembled from a set of received code blocks is not verified then, for each of the DTKA-KA nodes that transmitted one or more of the constituent code blocks, all code blocks transmitted by that node are excluded from the reassembled bulletin and verification of the bulletin's signature is attempted again. Upon success, the node whose transmitted code blocks had been excluded from the reassembled bulletin may be presumed to be compromised.

Code Block Numbers (0 to (Q + k - 1))	0	1	2	3	4	5	6	7
KA 1	x	x	x					
KA 2		x	x	x				
KA 3			x	x	x			
KA 4				x	x	x		
KA 5					x	x	x	
KA 6						x	x	x
KA 7	x						x	x
KA 8	x	x						x

Table 1: Example: Code Blocks Assignments for Key Authorities

The message format for transmitting the assigned code-blocks by each DTKA-KA is shown in Figure 4. Note that each such message is the payload of a Bundle and that the authenticity of that payload is nominally protected by a Block Integrity Block containing a digital signature computed in the private key of the issuing Key Agent; the message itself contains no self-authentication material. Reading the figure left to right, the first item is the bulletin hash as defined in Figure 3, the second item identifies the pre-defined code blocks for that DTKA-KA as conceptualized in Table 1, and the last item is the value of the identified code blocks. The identity of the DTKA-KA (KAx) that generated the code blocks is given by the source node ID of the DTN bundle in which the message arrived. KAx is used to validate the signature in the bundle’s Block Integrity Block before the message is delivered to DTKA.

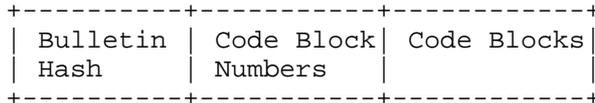


Figure 4: Message Format for Code Blocks

4.2. Node Registration

In order to register a new DTKA-KO in the system, DTKA requires the DTKA-KO with a Node ID (DTKA-KO[Node ID]) to generate a public-private key pair and preserve the secrecy of its private key. The DTKA-KO[Node ID] needs to generate an association message of the form (Node ID, effective-time, public-key), where effective-time specifies the start time after which the public-key is valid. That is, each bundle sent by this node is to be authenticated using the node’s most recently effective public key whose effective time is less than the

bundle's creation time. The DTKA-KO[Node ID] must send the association message, along with a signature on the message using its private key, to the DTKA-KA as depicted in Figure 5. Since DTKA-KA would not have seen the association of the public-key to that key owner previously, it cannot trust that the message indeed originated from DTKA-KO[Node ID]. Therefore, for registration purposes, this initial message from the DTKA-KO[Node ID] to the DTKA-KA MUST be protected by transmitting it over an independently (e.g., physically) authenticated channel. The independently authenticated channel can be realized by physically securing the access to the DTKA-KA server, using a physical communication medium, such as a USB dongle, and manually verifying the authenticity of the communication from the DTKA-KO. The manual verification is a one-time process for a given Key Owner. When an application domain has more than one DTKA-KA (KAx), the message from DTKA-KO[Node ID] must be sent to each DTKA-KA (KAx) in a similarly secure manner.

Although the messages to DTKA-KA (KAx) are independently authenticated, the DTKA-KO[Node ID] must sign the association message using its private key. The signature is not intended to cryptographically authenticate the message but only to prove to the DTKA-KA that the DTKA-KO[Node ID] is indeed in possession of the private key. This self-signed message by the DTKA-KO is useful to ensure that the physical courier, which is used to realize the physically authenticated channel, has not tampered the message sent by the DTKA-KO to the DTKA-KA. Additionally, the self-signed message is useful to audit the operations of the DTKA-KA.

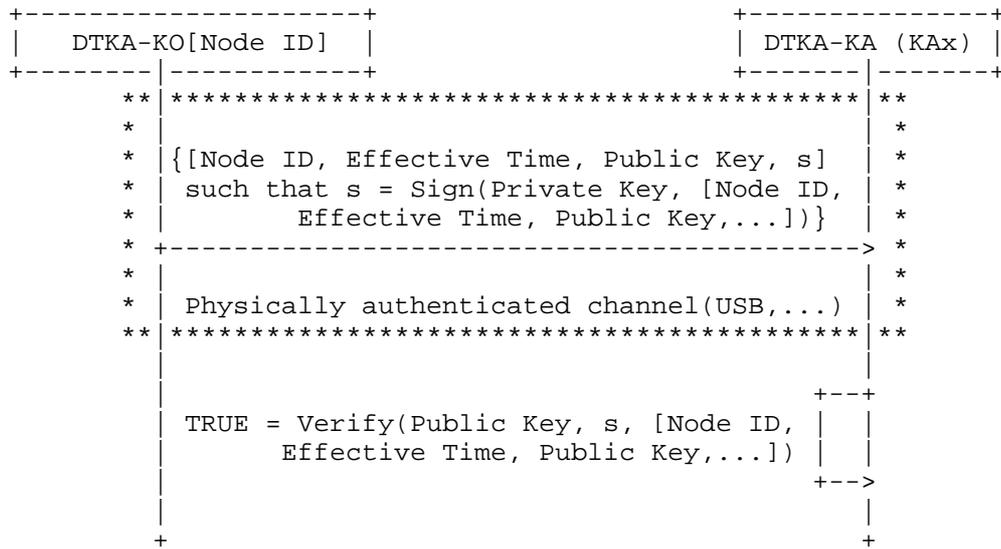


Figure 5: Interaction Diagram 1: Node Registration

As an alternative to the use of a physically authenticated channel, the registration association message might be sent by a trusted third-party node whose authenticated public key is known to all KAs, so that the message may be authenticated by verifying the digital signature (formed using the trusted third-party node’s current private key) in the BIB of the bundle containing the message.

The DTKA-KA will insert the received association message into its next bulletin (refer to Figure 3), for multicast, as an assertion. The bulletin will be multicast to all DTKA Entities using the protocol described in Section 4.5.

4.3. Key Revocation

Manual decisions trigger the key revocation procedure. Every DTKA-KA in the application domain is assumed to have a human operator who can trigger the revocation process. When a key is to be revoked, the human operator will need to authenticate to the respective DTKA-KA (KAx) server, identify the public-key and Node ID to be revoked, and instruct that DTKA-KA (KAx) revocation software to schedule a revocation message. The revocation software in DTKA-KA (KAx) will: (a) insert the revocation message in its next bulletin as described in Figure 3; and, (b) send a revocation notice to the other DTKA-KA (KAy) instances using the protocol described in Figure 6. After authenticating the revocation instruction and arriving at a consensus for a given bulletin, of which the revocation information is a part,

each DTKA-KA will multicast the revocation in their next bulletin code block as described in Figure 3.

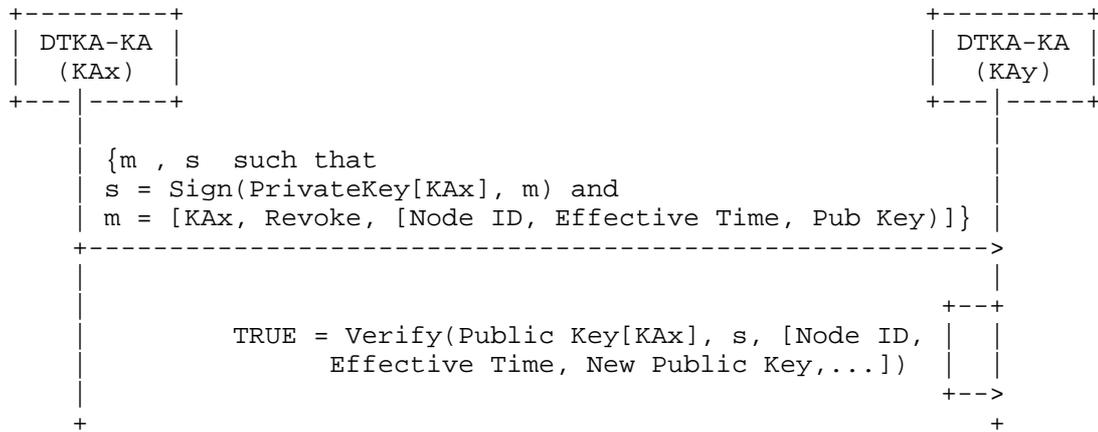


Figure 6: Interaction Diagram 1.1: Key Revocation

4.4. Key Roll-over

When a DTKA-KO[Node ID] has been registered by the DTKA-KA using the protocol described in Figure 5, the DTKA-KO[Node ID] can periodically roll-over to a new public-private key pair by following the key roll-over protocol described in Figure 7. The protocol for key roll-over is similar to the one for key registration except that: (a) the protocol can be executed using DTN bundles issued by the KO itself without requiring any independently secured out-of-band communication channels; and, (b) the old (current) public-key is used to authenticate the association of the new public-key with the Node ID for that DTKA KO. The DTKA-KO [Node ID] must send this message to every key agent in its application domain. Upon accepting the roll-over message from the DTKA-KO[Node ID], each key agent will schedule the roll-over instruction for identified Node ID and public-key in its next bulletin as described in Section 4.1. A DTKA-KO can schedule any number of future roll-overs but the number of such roll-over schedules may need to be limited to avoid Denial of Service attacks by registered nodes -- but this topic is beyond the scope of this document.

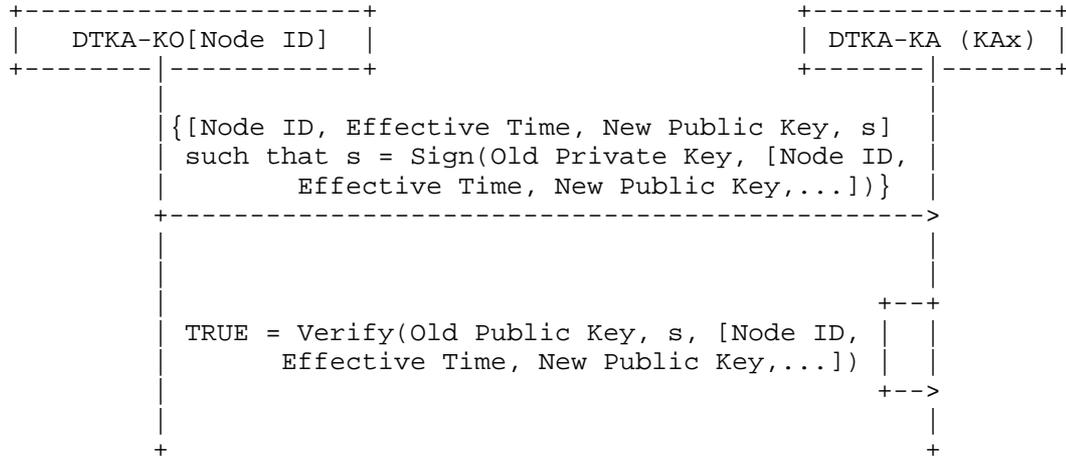


Figure 7: Interaction Diagram 1.2: Key Rollover

4.5. Key Distribution

Each DTKA-KA collects multiple assertion, revocation, and roll-over messages from different parties by following the protocols described in Section 4.2, Section 4.3, and Section 4.4. Then, each DTKA-KA forms and multicasts the code blocks for its bulletin to all DTKA Key Users as explained in Section 4.1. The DTKA-KUs verify the authenticity of each code block from all the DTKA-KAs before using the code blocks to decode the bulletin, which will contains key assertion, revocation, and roll-over instructions. The DTKA-KUs perform these instructions in their respective local key database. This interaction between the DTKA-KAs and the DTKA-KUs of an application domain is shown in Figure 8.

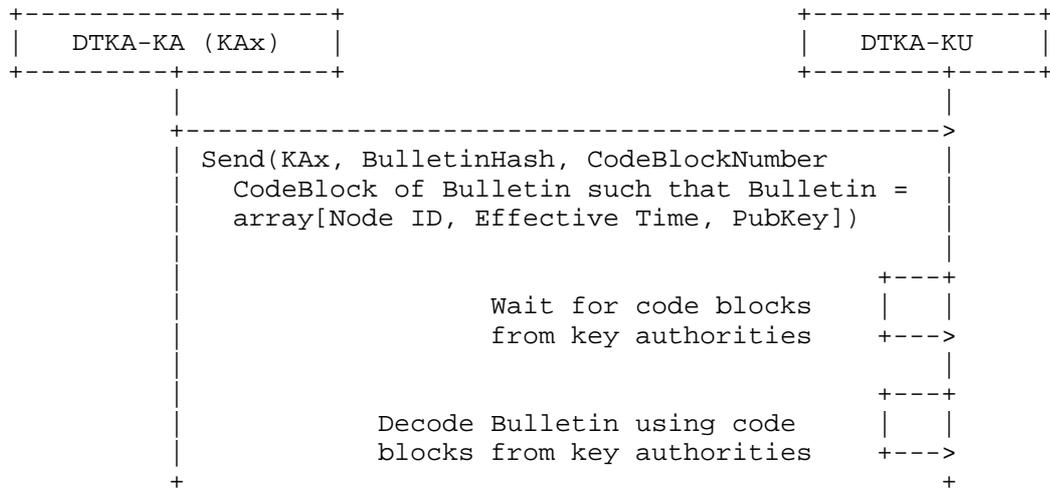


Figure 8: Interaction Diagram 2: Bulk Key Distribution

4.6. Secure Communications

After receiving key assertion and roll-over information, every DTKA-KU shall have authenticated public-keys for different Node IDs in its local database. These authenticated public-keys can be used to authenticate messages received from the DTKA-KO[Node ID] and to send confidential messages to the DTKA-KO[Node ID] after the specified effective-time for each Node ID and public-key pair. This interaction is specified in Figure 9.

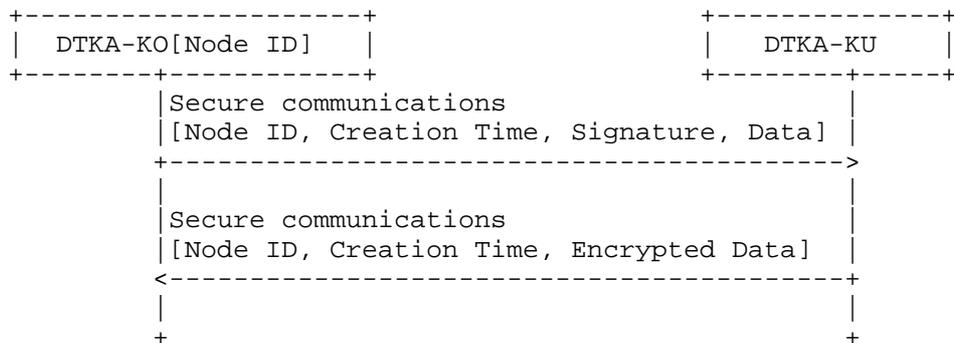


Figure 9: Interaction Diagram 3: Secure communication

4.7. Communication Stack View

DTKA is designed to be a special DTN application that shall perform key management operations using the services of the Bundle Protocol and BPsec. DTKA will use BP, which in turn will use BPsec to authenticate the messages containing the public-keys that are subsequently to be used by BPsec for securing future communications as shown in Figure 10.

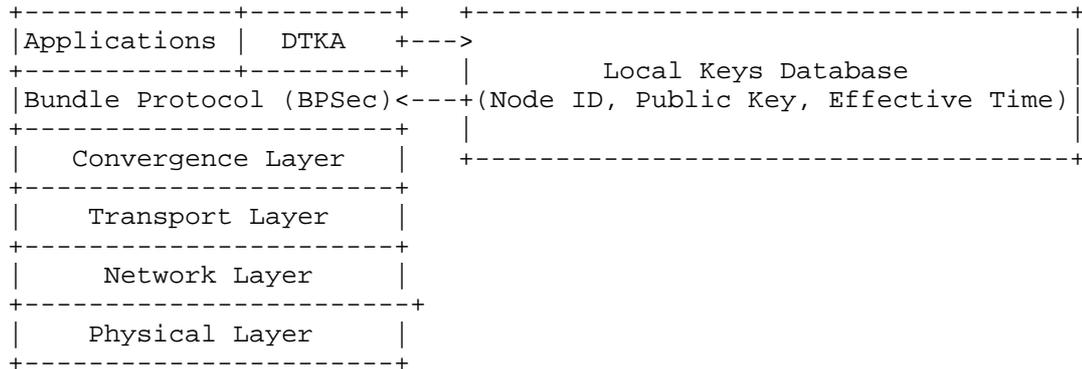


Figure 10: Interaction Diagram 3: Secure communication

5. IANA Considerations

This document potentially contains IANA considerations depending on the design choices adopted for future work. But, in its present form, there are no immediate IANA considerations.

6. Security Considerations

Security issues and considerations are discussed through out this document.

7. References

7.1. Normative References

[I-D.ietf-dtn-bpbis]
 Burleigh, S., Fall, K., and E. Birrane, "Bundle Protocol",
 draft-ietf-dtn-bpbis-08 (work in progress), August 2017.

- [I-D.ietf-dtn-bpsec]
Birrane, E. and K. McKeever, "Bundle Protocol Security Specification", draft-ietf-dtn-bpsec-05 (work in progress), July 2017.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.

7.2. Informative References

- [I-D.templin-dtnskmreq]
Templin, F. and S. Burleigh, "DTN Security Key Management - Requirements and Design", draft-templin-dtnskmreq-00 (work in progress), February 2015.
- [RFC4838] Cerf, V., Burleigh, S., Hooke, A., Torgerson, L., Durst, R., Scott, K., Fall, K., and H. Weiss, "Delay-Tolerant Networking Architecture", RFC 4838, DOI 10.17487/RFC4838, April 2007, <<https://www.rfc-editor.org/info/rfc4838>>.
- [RFC5050] Scott, K. and S. Burleigh, "Bundle Protocol Specification", RFC 5050, DOI 10.17487/RFC5050, November 2007, <<https://www.rfc-editor.org/info/rfc5050>>.
- [RFC6257] Symington, S., Farrell, S., Weiss, H., and P. Lovell, "Bundle Security Protocol Specification", RFC 6257, DOI 10.17487/RFC6257, May 2011, <<https://www.rfc-editor.org/info/rfc6257>>.

Authors' Addresses

Scott Burleigh
JPL, Calif. Inst. Of Technology
4800 Oak Grove Dr.
Pasadena, CA 91109-8099
USA

Email: Scott.Burleigh@jpl.nasa.gov

David Horres
JPL, Calif. Inst. Of Technology
4800 Oak Grove Dr.
Pasadena, CA 91109-8099
USA

Email: David.C.Horres@jpl.nasa.gov

Kapali Viswanathan
Boeing Research & Technology
Boeing International Corporation India Private Limited
A Block, 4th Floor, Lake View Building
Bagmane Tech Park, C.V. Raman Nagar
Bangalore, KA 560093
IN

Email: kapaleeswaran.viswanathan@boeing.com

Michael W. Benson
Boeing Research & Technology
The Boeing Company
499 Boeing Boulevard
Huntsville, AL 35824
USA

Email: michael.w.benson@boeing.com

Fred L. Templin
Boeing Research & Technology
P.O. Box 3707
Seattle, WA 98124
USA

Email: fltemplin@acm.org

Delay-Tolerant Networking Working Group
Internet Draft
Intended status: Standards Track
Expires: February 2018

S. Burleigh
JPL, Calif. Inst. Of Technology
K. Fall
Nefeli Networks, Inc.
E. Birrane
APL, Johns Hopkins University
August 8, 2017

Bundle Protocol
draft-ietf-dtn-bpbis-08.txt

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>

This Internet-Draft will expire on February 9, 2018.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in

Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Abstract

This Internet Draft presents a specification for Bundle Protocol, adapted from the experimental Bundle Protocol specification developed by the Delay-Tolerant Networking Research group of the Internet Research Task Force and documented in RFC 5050.

Table of Contents

1. Introduction.....	3
2. Conventions used in this document.....	5
3. Service Description.....	5
3.1. Definitions.....	5
3.2. Discussion of BP concepts.....	9
3.3. Services Offered by Bundle Protocol Agents.....	11
4. Bundle Format.....	12
4.1. BP Fundamental Data Structures.....	12
4.1.1. CRC Type.....	12
4.1.2. CRC.....	13
4.1.3. Bundle Processing Control Flags.....	13
4.1.4. Block Processing Control Flags.....	14
4.1.5. Identifiers.....	15
4.1.5.1. Endpoint ID.....	15
4.1.5.2. Node ID.....	16
4.1.6. DTN Time.....	17
4.1.7. Creation Timestamp.....	17
4.1.8. Block-type-specific Data.....	17
4.2. Bundle Representation.....	17
4.2.1. Bundle.....	17
4.2.2. Primary Bundle Block.....	18
4.2.3. Canonical Bundle Block Format.....	19
4.3. Extension Blocks.....	20
4.3.1. Previous Node.....	21
4.3.2. Bundle Age.....	21
4.3.3. Hop Count.....	22
5. Bundle Processing.....	22
5.1. Generation of Administrative Records.....	23
5.2. Bundle Transmission.....	23
5.3. Bundle Dispatching.....	24
5.4. Bundle Forwarding.....	24
5.4.1. Forwarding Contraindicated.....	26
5.4.2. Forwarding Failed.....	26
5.5. Bundle Expiration.....	26

5.6. Bundle Reception.....	27
5.7. Local Bundle Delivery.....	27
5.8. Bundle Fragmentation.....	28
5.9. Application Data Unit Reassembly.....	30
5.10. Bundle Deletion.....	30
5.11. Discarding a Bundle.....	30
5.12. Canceling a Transmission.....	30
6. Administrative Record Processing.....	31
6.1. Administrative Records.....	31
6.1.1. Bundle Status Reports.....	32
6.2. Generation of Administrative Records.....	34
7. Services Required of the Convergence Layer.....	34
7.1. The Convergence Layer.....	34
7.2. Summary of Convergence Layer Services.....	35
8. Implementation Status.....	35
9. Security Considerations.....	36
10. IANA Considerations.....	38
11. References.....	39
11.1. Normative References.....	39
11.2. Informative References.....	39
12. Acknowledgments.....	40
13. Significant Changes from RFC 5050.....	40
Appendix A. For More Information.....	41
Appendix B. CDDL expression.....	42

1. Introduction

Since the publication of the Bundle Protocol Specification (Experimental RFC 5050) in 2007, the Delay-Tolerant Networking Bundle Protocol has been implemented in multiple programming languages and deployed to a wide variety of computing platforms. This implementation and deployment experience has identified opportunities for making the protocol simpler, more capable, and easier to use. The present document, standardizing the Bundle Protocol (BP), is adapted from RFC 5050 in that context.

This document describes version 7 of BP.

Delay Tolerant Networking is a network architecture providing communications in and/or through highly stressed environments. Stressed networking environments include those with intermittent connectivity, large and/or variable delays, and high bit error rates. To provide its services, BP may be viewed as sitting at the application layer of some number of constituent networks, forming a store-carry-forward overlay network. Key capabilities of BP include:

- . Ability to use physical motility for the movement of data
- . Ability to move the responsibility for error control from one node to another
- . Ability to cope with intermittent connectivity, including cases where the sender and receiver are not concurrently present in the network
- . Ability to take advantage of scheduled, predicted, and opportunistic connectivity, whether bidirectional or unidirectional, in addition to continuous connectivity
- . Late binding of overlay network endpoint identifiers to underlying constituent network addresses

For descriptions of these capabilities and the rationale for the DTN architecture, see [ARCH] and [SIGC].

BP's location within the standard protocol stack is as shown in Figure 1. BP uses underlying "native" transport and/or network protocols for communications within a given constituent network.

The interface between the bundle protocol and a specific underlying protocol is termed a "convergence layer adapter".

Figure 1 shows three distinct transport and network protocols (denoted T1/N1, T2/N2, and T3/N3).

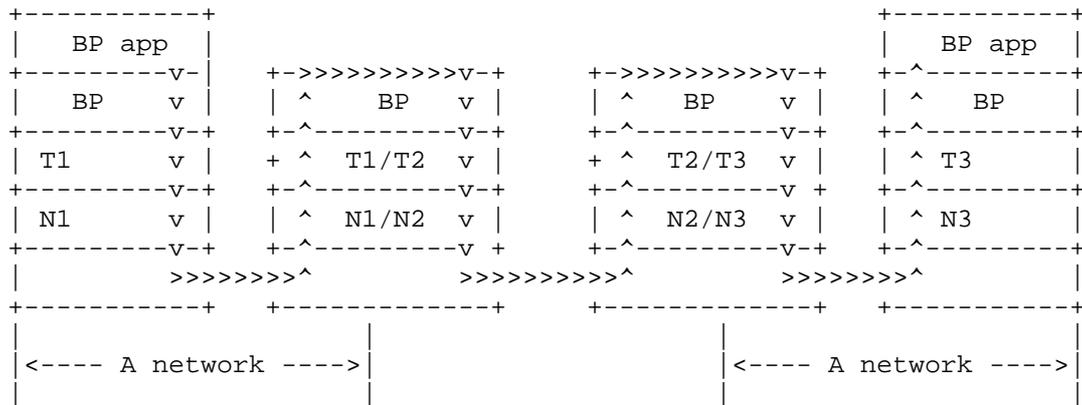


Figure 1: The Bundle Protocol in the Protocol Stack Model

This document describes the format of the protocol data units (called "bundles") passed between entities participating in BP communications.

The entities are referred to as "bundle nodes". This document does not address:

- . Operations in the convergence layer adapters that bundle nodes use to transport data through specific types of internets. (However, the document does discuss the services that must be provided by each adapter at the convergence layer.)
- . The bundle route computation algorithm.
- . Mechanisms for populating the routing or forwarding information bases of bundle nodes.
- . The mechanisms for securing bundles en route.
- . The mechanisms for managing bundle nodes.

Note that implementations of the specification presented in this document will not be interoperable with implementations of RFC 5050.

2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC-2119 [RFC2119].

In this document, these words will appear with that interpretation only when in ALL CAPS. Lower case uses of these words are not to be interpreted as carrying RFC-2119 significance.

3. Service Description

3.1. Definitions

Bundle - A bundle is a protocol data unit of BP, so named because negotiation of the parameters of a data exchange may be impractical in a delay-tolerant network: it is often better practice to "bundle" with a unit of data all metadata that might be needed in order to make the data immediately usable when delivered to applications. Each bundle comprises a sequence of two or more "blocks" of protocol data, which serve various purposes.

Block - A bundle protocol block is one of the protocol data structures that together constitute a well-formed bundle.

Bundle payload - A bundle payload (or simply "payload") is the application data whose conveyance to the bundle's destination is the purpose for the transmission of a given bundle; it is the content of the bundle's payload block. The terms "bundle content", "bundle payload", and "payload" are used interchangeably in this document.

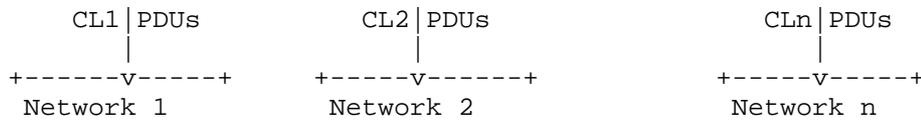


Figure 2: Components of a BP Node

Bundle protocol agent - The bundle protocol agent (BPA) of a node is the node component that offers the BP services and executes the procedures of the bundle protocol.

Convergence layer adapter - A convergence layer adapter (CLA) is a node component that sends and receives bundles on behalf of the BPA, utilizing the services of some 'native' protocol stack that is supported in one of the networks within which the node is functionally located.

Application agent - The application agent (AA) of a node is the node component that utilizes the BP services to effect communication for some user purpose. The application agent in turn has two elements, an administrative element and an application-specific element.

Application-specific element - The application-specific element of an AA is the node component that constructs, requests transmission of, accepts delivery of, and processes units of user application data.

Administrative element - The administrative element of an AA is the node component that constructs and requests transmission of administrative records (defined below), including status reports, and accepts delivery of and processes any administrative records that the node receives.

Administrative record - A BP administrative record is an application data unit that is exchanged between the administrative elements of nodes' application agents for some BP administrative purpose. The only administrative record defined in this specification is the status report, discussed later.

Bundle endpoint - A bundle endpoint (or simply "endpoint") is a set of zero or more bundle nodes that all identify themselves for BP purposes by some common identifier, called a "bundle endpoint ID" (or, in this document, simply "endpoint ID"; endpoint IDs are described in detail in Section 4.4.1 below).

Singleton endpoint - A singleton endpoint is an endpoint that always contains exactly one member.

Registration - A registration is the state machine characterizing a given node's membership in a given endpoint. Any single registration has an associated delivery failure action as defined below and must at any time be in one of two states: Active or Passive.

Delivery - A bundle is considered to have been delivered at a node subject to a registration as soon as the application data unit that is the payload of the bundle, together with any relevant metadata (an implementation matter), has been presented to the node's application agent in a manner consistent with the state of that registration.

Deliverability - A bundle is considered "deliverable" subject to a registration if and only if (a) the bundle's destination endpoint is the endpoint with which the registration is associated, (b) the bundle has not yet been delivered subject to this registration, and (c) the bundle has not yet been "abandoned" (as defined below) subject to this registration.

Abandonment - To abandon a bundle subject to some registration is to assert that the bundle is not deliverable subject to that registration.

Delivery failure action - The delivery failure action of a registration is the action that is to be taken when a bundle that is "deliverable" subject to that registration is received at a time when the registration is in the Passive state.

Destination - The destination of a bundle is the endpoint comprising the node(s) at which the bundle is to be delivered (as defined below).

Transmission - A transmission is an attempt by a node's BPA to cause copies of a bundle to be delivered to one or more of the nodes that are members of some endpoint (the bundle's destination) in response to a transmission request issued by the node's application agent.

Forwarding - To forward a bundle to a node is to invoke the services of one or more CLAs in a sustained effort to cause a copy of the bundle to be received by that node.

Discarding - To discard a bundle is to cease all operations on the bundle and functionally erase all references to it. The specific procedures by which this is accomplished are an implementation matter.

Retention constraint - A retention constraint is an element of the state of a bundle that prevents the bundle from being discarded. That is, a bundle cannot be discarded while it has any retention constraints.

Deletion - To delete a bundle is to remove unconditionally all of the bundle's retention constraints, enabling the bundle to be discarded.

3.2. Discussion of BP concepts

Multiple instances of the same bundle (the same unit of DTN protocol data) might exist concurrently in different parts of a network -- possibly differing in some blocks -- in the memory local to one or more bundle nodes and/or in transit between nodes. In the context of the operation of a bundle node, a bundle is an instance (copy), in that node's local memory, of some bundle that is in the network.

The payload for a bundle forwarded in response to a bundle transmission request is the application data unit whose location is provided as a parameter to that request. The payload for a bundle forwarded in response to reception of a bundle is the payload of the received bundle.

In the most familiar case, a bundle node is instantiated as a single process running on a general-purpose computer, but in general the definition is meant to be broader: a bundle node might alternatively be a thread, an object in an object-oriented operating system, a special-purpose hardware device, etc.

The manner in which the functions of the BPA are performed is wholly an implementation matter. For example, BPA functionality might be coded into each node individually; it might be implemented as a shared library that is used in common by any number of bundle nodes on a single computer; it might be implemented as a daemon whose services are invoked via inter-process or network communication by any number of bundle nodes on one or more computers; it might be implemented in hardware.

Every CLA implements its own thin layer of protocol, interposed between BP and the (usually "top") protocol(s) of the underlying native protocol stack; this "CL protocol" may only serve to multiplex and de-multiplex bundles to and from the underlying native protocol, or it may offer additional CL-specific functionality. The manner in which a CLA sends and receives bundles, as well as the definitions of CLAs and CL protocols, are beyond the scope of this specification.

Note that the administrative element of a node's application agent may itself, in some cases, function as a convergence-layer adapter. That is, outgoing bundles may be "tunneled" through encapsulating bundles:

- . An outgoing bundle constitutes a byte array. This byte array may, like any other, be presented to the bundle protocol agent as an application data unit that is to be transmitted to some endpoint.
- . The original bundle thus forms the payload of an encapsulating bundle that is forwarded using some other convergence-layer protocol(s).
- . When the encapsulating bundle is received, its payload is delivered to the peer application agent administrative element, which then instructs the bundle protocol agent to dispatch that original bundle in the usual way.

The purposes for which this technique may be useful (such as cross-domain security) are beyond the scope of this specification.

The only interface between the BPA and the application-specific element of the AA is the BP service interface. But between the BPA and the administrative element of the AA there is a (conceptual) private control interface in addition to the BP service interface. This private control interface enables the BPA and the administrative element of the AA to direct each other to take action under specific circumstances.

In the case of a node that serves simply as a BP "router", the AA may have no application-specific element at all. The application-specific elements of other nodes' AAs may perform arbitrarily complex application functions, perhaps even offering multiplexed DTN communication services to a number of other applications. As with the BPA, the manner in which the AA performs its functions is wholly an implementation matter.

Singletons are the most familiar sort of endpoint, but in general the endpoint notion is meant to be broader. For example, the nodes in a sensor network might constitute a set of bundle nodes that identify themselves by a single common endpoint ID and thus form a single bundle endpoint. *Note* too that a given bundle node might identify itself by multiple endpoint IDs and thus be a member of multiple bundle endpoints.

The destination of every bundle is an endpoint, which may or may not be singleton. The source of every bundle is a node, identified by the endpoint ID for some singleton endpoint that contains that node.

Any number of transmissions may be concurrently undertaken by the bundle protocol agent of a given node.

When the bundle protocol agent of a node determines that a bundle must be forwarded to a node (either to a node that is a member of the bundle's destination endpoint or to some intermediate forwarding node) in the course of completing the successful transmission of that bundle, it invokes the services of one or more CLAs in a sustained effort to cause a copy of the bundle to be received by that node.

Upon reception, the processing of a bundle that has been received by a given node depends on whether or not the receiving node is registered in the bundle's destination endpoint. If it is, and if the payload of the bundle is non-fragmentary (possibly as a result of successful payload reassembly from fragmentary payloads, including the original payload of the newly received bundle), then the bundle is normally delivered to the node's application agent subject to the registration characterizing the node's membership in the destination endpoint.

The bundle protocol does not natively ensure delivery of a bundle to its destination. Data loss along the path to the destination node can be minimized by utilizing reliable convergence-layer protocols between neighbors on all segments of the end-to-end path, but for end-to-end bundle delivery assurance it will be necessary to develop extensions to the bundle protocol and/or application-layer mechanisms.

The bundle protocol is designed for extensibility. Bundle protocol extensions, documented elsewhere, may extend this specification by:

- . defining additional blocks;
- . defining additional administrative records;
- . defining additional bundle processing flags;
- . defining additional block processing flags;
- . defining additional types of bundle status reports;
- . defining additional bundle status report reason codes;
- . defining additional mandates and constraints on processing that conformant bundle protocol agents must perform at specified points in the inbound and outbound bundle processing cycles.

3.3. Services Offered by Bundle Protocol Agents

The BPA of each node is expected to provide the following services to the node's application agent:

- . commencing a registration (registering the node in an endpoint);
- . terminating a registration;
- . switching a registration between Active and Passive states;
- . transmitting a bundle to an identified bundle endpoint;
- . canceling a transmission;
- . polling a registration that is in the Passive state;
- . delivering a received bundle.

4. Bundle Format

The format of bundles SHALL conform to the Concise Binary Object Representation (CBOR [RFC7049]).

Each bundle SHALL be a concatenated sequence of at least two blocks, represented as a CBOR indefinite-length array. The first block in the sequence (the first item of the array) MUST be a primary bundle block in CBOR representation as described below; the bundle MUST have exactly one primary bundle block. The primary block MUST be followed by one or more canonical bundle blocks (additional array items) in CBOR representation as described below. The last such block MUST be a payload block; the bundle MUST have exactly one payload block. The last item of the array, immediately following the payload block, SHALL be a CBOR "break" stop code.

An implementation of the Bundle Protocol MAY discard any sequence of bytes that does not conform to the Bundle Protocol specification.

An implementation of the Bundle Protocol MAY accept a sequence of bytes that does not conform to the Bundle Protocol specification (e.g., one that represents data elements in fixed-length arrays rather than indefinite-length arrays) and transform it into conformant BP structure before processing it. Procedures for accomplishing such a transformation are beyond the scope of this specification.

4.1. BP Fundamental Data Structures

4.1.1. CRC Type

CRC type is an unsigned integer type code for which the following values (and no others) are valid:

- . 0 indicates "no CRC is present."
- . 1 indicates "a standard X-25 CRC-16 is present." [CRC]
- . 2 indicates "a standard PKZIP CRC-32 is present." [CRC]

CRC type SHALL be represented as a CBOR unsigned integer.

4.1.2. CRC

CRC SHALL be omitted from a block if and only if the block's CRC type code is zero.

When not omitted, the CRC SHALL be represented as sequence of two bytes (if CRC type is 1) or as a sequence of four bytes (if CRC type is 2).

4.1.3. Bundle Processing Control Flags

Bundle processing control flags assert properties of the bundle as a whole rather than of any particular block of the bundle. They are conveyed in the primary block of the bundle.

The following properties are asserted by the bundle processing control flags:

- . The bundle is a fragment. (Boolean)
- . The bundle's payload is an administrative record. (Boolean)
- . The bundle must not be fragmented. (Boolean)
- . The bundle's destination endpoint is a singleton. (Boolean)
- . Acknowledgment by the user application is requested. (Boolean)
- . Status time is requested in all status reports. (Boolean)
- . The bundle contains a "manifest" extension block. (Boolean)
- . Flags requesting types of status reports (all Boolean):
 - o Request reporting of bundle reception.
 - o Request reporting of bundle forwarding.
 - o Request reporting of bundle delivery.
 - o Request reporting of bundle deletion.

If the bundle processing control flags indicate that the bundle's application data unit is an administrative record, then all status report request flag values must be zero.

If the bundle's source node is omitted (i.e., the source node ID is the ID of the null endpoint, which has no members as discussed below; this option enables anonymous bundle transmission), then the bundle is not uniquely identifiable and all bundle protocol features that rely on bundle identity must therefore be disabled: the "Bundle must not be fragmented" flag value must be 1 and all status report request flag values must be zero.

The bundle processing control flags SHALL be represented as a CBOR unsigned integer item containing a bit field of 16 bits indicating the control flag values as follows:

- . Bit 0 (the high-order bit, 0x8000): reserved.
- . Bit 1 (0x4000): reserved.
- . Bit 2 (0x2000): reserved.
- . Bit 3(0x1000): bundle deletion status reports are requested.
- . Bit 4(0x0800): bundle delivery status reports are requested.
- . Bit 5(0x0400): bundle forwarding status reports are requested.
- . Bit 6(0x0200): reserved.
- . Bit 7(0x0100): bundle reception status reports are requested.
- . Bit 8(0x0080): bundle contains a Manifest block.
- . Bit 9(0x0040): status time is requested in all status reports.
- . Bit 10(0x0020): user application acknowledgement is requested.
- . Bit 11(0x0010): destination is a singleton endpoint.
- . Bit 12(0x0008): reserved.
- . Bit 13(0x0004): bundle must not be fragmented.
- . Bit 14(0x0002): payload is an administrative record.
- . Bit 15 (the low-order bit, 0x0001: bundle is a fragment.

4.1.4. Block Processing Control Flags

The block processing control flags assert properties of canonical bundle blocks. They are conveyed in the header of the block to which they pertain.

The following properties are asserted by the block processing control flags:

- . This block must be replicated in every fragment. (Boolean)
- . Transmission of a status report is requested if this block can't be processed. (Boolean)
- . Block must be removed from the bundle if it can't be processed. (Boolean)

- . Bundle must be deleted if this block can't be processed.
(Boolean)

For each bundle whose bundle processing control flags indicate that the bundle's application data unit is an administrative record, or whose source node ID is the null endpoint ID as defined below, the value of the "Transmit status report if block can't be processed" flag in every canonical block of the bundle must be zero.

The block processing control flags SHALL be represented as a CBOR unsigned integer item containing a bit field of 8 bits indicating the control flag values as follows:

- . Bit 0 (the high-order bit, 0x80): reserved.
- . Bit 1 (0x40): reserved.
- . Bit 2(0x20): reserved.
- . Bit 3(0x10): reserved.
- . Bit 4(0x08): bundle must be deleted if block can't be processed.
- . Bit 5(0x04): transmission of a status report is requested if block can't be processed.
- . Bit 6(0x02): block must be removed from bundle if it can't be processed.
- . Bit 7(the low-order bit, 0x01): block must be replicated in every fragment.

4.1.5. Identifiers

4.1.5.1. Endpoint ID

The destinations of bundles are bundle endpoints, identified by text strings termed "endpoint IDs" (see Section 3.1). Each endpoint ID (EID) is a Uniform Resource Identifier (URI; [URI]). As such, each endpoint ID can be characterized as having this general structure:

< scheme name > : < scheme-specific part, or "SSP" >

The scheme identified by the < scheme name > in an endpoint ID is a set of syntactic and semantic rules that fully explain how to parse and interpret the SSP. The set of allowable schemes is effectively unlimited. Any scheme conforming to [URIREG] may be used in a bundle protocol endpoint ID.

Note that, although endpoint IDs are URIs, implementations of the BP service interface may support expression of endpoint IDs in some internationalized manner (e.g., Internationalized Resource Identifiers (IRIs); see [RFC3987]).

The endpoint ID "dtn:none" identifies the "null endpoint", the endpoint that by definition never has any members.

Each BP endpoint ID (EID) SHALL be represented as a CBOR array comprising a 2-tuple.

The first item of the array SHALL be the code number identifying the endpoint's URI scheme [URI], as defined in the registry of URI scheme code numbers for Bundle Protocol maintained by IANA as described in Section 10. [URIREG]. Each URI scheme code number SHALL be represented as a CBOR unsigned integer.

The second item of the array SHALL be the applicable CBOR representation of the scheme-specific part (SSP) of the EID, defined as follows:

- . If the EID's URI scheme is "dtn" then the SSP SHALL be represented as a CBOR text string unless the EID's SSP is "none", in which case the SSP SHALL be represented as a CBOR unsigned integer with the value zero.
- . If the EID's URI scheme is "ipn" then the SSP SHALL be represented as a CBOR array comprising a 2-tuple. The first item of this array SHALL be the EID's node number represented as a CBOR unsigned integer. The second item of this array SHALL be the EID's service number represented as a CBOR unsigned integer.
- . Definitions of the CBOR representations of the SSPs of EIDs encoded in other URI schemes are included in the specifications defining those schemes.

4.1.5.2. Node ID

For many purposes of the Bundle Protocol it is important to identify the node that is operative in some context.

As discussed in 3.1 above, nodes are distinct from endpoints; specifically, an endpoint is a set of zero or more nodes. But rather than define a separate namespace for node identifiers, we instead use endpoint identifiers to identify nodes, subject to the following restrictions:

- . Every node MUST be a member of at least one singleton endpoint.
- . The EID of any singleton endpoint of which a node is a member MAY be used to identify that node. A "node ID" is an EID that is used in this way.
- . A node's membership in a given singleton endpoint MUST be sustained at least until the nominal operation of the Bundle

Protocol no longer depends on the identification of that node using that endpoint's ID.

4.1.6. DTN Time

A DTN time is an unsigned integer indicating a count of seconds since the start of the year 2000 on the Coordinated Universal Time (UTC) scale [UTC]. Each DTN time SHALL be represented as a CBOR unsigned integer item.

4.1.7. Creation Timestamp

Each creation timestamp SHALL be represented as a CBOR array item comprising a 2-tuple.

The first item of the array SHALL be a DTN time.

The second item of the array SHALL be the creation timestamp's sequence number, represented as a CBOR unsigned integer.

4.1.8. Block-type-specific Data

Block-type-specific data in each block (other than the primary block) SHALL be the applicable CBOR representation of the content of the block. Details of this representation are included in the specification defining the block type.

4.2. Bundle Representation

This section describes the primary block in detail and non-primary blocks in general. Rules for processing these blocks appear in Section 5 of this document.

Note that supplementary DTN protocol specifications (including, but not restricted to, the Bundle Security Protocol [BPSEC]) may require that BP implementations conforming to those protocols construct and process additional blocks.

4.2.1. Bundle

Each bundle SHALL be represented as a CBOR indefinite-length array. The first item of this array SHALL be the CBOR representation of a Primary Block. Every other item of the array except the last SHALL be the CBOR representation of a Canonical Block. The last item of the array SHALL be a CBOR "break" stop code.

4.2.2. Primary Bundle Block

The primary bundle block contains the basic information needed to forward bundles to their destinations.

Each primary block SHALL be represented as a CBOR array; the number of elements in the array SHALL be 8 (if the bundle is not a fragment and CRC type is zero) or 9 (if the bundle is not a fragment and CRC type is non-zero) or 10 (if the bundle is a fragment and CRC type is zero) or 11 (if the bundle is a fragment and CRC-type is non-zero).

The fields of the primary bundle block SHALL be as follows, listed in the order in which they MUST appear:

Version: An unsigned integer value indicating the version of the bundle protocol that constructed this block. The present document describes version 7 of the bundle protocol. Version number SHALL be represented as a CBOR unsigned integer item.

Bundle Processing Control Flags: The Bundle Processing Control Flags are discussed in Section 4.1.3. above.

CRC Type: CRC Type codes are discussed in Section 4.1.1. above.

Destination EID: The Destination EID field identifies the bundle endpoint that is the bundle's destination, i.e., the endpoint that contains the node(s) at which the bundle is to be delivered.

Source node ID: The Source node ID field identifies the bundle node at which the bundle was initially transmitted, except that Source node ID may be the null endpoint ID in the event that the bundle's source chooses to remain anonymous.

Report-to EID: The Report-to EID field identifies the bundle endpoint to which status reports pertaining to the forwarding and delivery of this bundle are to be transmitted.

Creation Timestamp: The creation timestamp is a pair of unsigned integers that, together with the source node ID and (if the bundle is a fragment) the fragment offset and payload length, serve to identify the bundle. The first of these integers is the bundle's creation time, while the second is the bundle's creation timestamp sequence number. Bundle creation time shall be the time - expressed in seconds since the start of the year 2000, on the Coordinated Universal Time (UTC) scale [UTC] - at which the transmission request was received that resulted in the creation of the bundle. Sequence count shall be the latest value (as of the time at which that

transmission request was received) of a monotonically increasing positive integer counter managed by the source node's bundle protocol agent that may be reset to zero whenever the current time advances by one second. For nodes that lack accurate clocks, it is recommended that bundle creation time be set to zero and that the counter used as the source of the bundle sequence count never be reset to zero. Note that, in general, the creation of two distinct bundles with the same source node ID and bundle creation timestamp may result in unexpected network behavior and/or suboptimal performance. The combination of source node ID and bundle creation timestamp serves to identify a single transmission request, enabling it to be acknowledged by the receiving application (provided the source node ID is not the null endpoint ID).

Lifetime: The lifetime field is an unsigned integer that indicates the time at which the bundle's payload will no longer be useful, encoded as a number of microseconds past the creation time. (For high-rate deployments with very brief disruptions, fine-grained expression of bundle lifetime may be useful.) When a bundle's age exceeds its lifetime, bundle nodes need no longer retain or forward the bundle; the bundle SHOULD be deleted from the network. Bundle lifetime SHALL be represented as a CBOR unsigned integer item.

Fragment offset: If and only if the Bundle Processing Control Flags of this Primary block indicate that the bundle is a fragment, fragment offset SHALL be present in the primary block. Fragment offset SHALL be represented as a CBOR unsigned integer indicating the offset from the start of the original application data unit at which the bytes comprising the payload of this bundle were located.

CRC: If and only if the value of the CRC type field of this Primary block is non-zero, a CRC SHALL be present in the primary block. The length and nature of the CRC SHALL be as indicated by the CRC type. The CRC SHALL be computed over the concatenation of all bytes (including CBOR "break" characters) of the primary block including the CRC field itself, which for this purpose SHALL be temporarily populated with the value zero.

4.2.3. Canonical Bundle Block Format

Every block other than the primary block (which blocks are termed "canonical" blocks) SHALL be represented as a CBOR array; the number of elements in the array SHALL be 6 (if CRC type is zero) or 7 (otherwise).

The fields of every canonical block SHALL be as follows, listed in the order in which they MUST appear:

- . Block type code, an unsigned integer. Bundle block type code 1 indicates that the block is a bundle payload block. Block type codes 2 through 9 are explicitly reserved as noted later in this specification. Block type codes 192 through 255 are not reserved and are available for private and/or experimental use. All other block type code values are reserved for future use.
- . Block number, an unsigned integer. The block number uniquely identifies the block within the bundle, enabling blocks (notably bundle security protocol blocks) to explicitly reference other blocks in the same bundle. Block numbers need not be in continuous sequence, and blocks need not appear in block number sequence in the bundle. The block number of the payload block is always zero.
- . Block processing control flags as discussed in Section 4.1.4 above.
- . CRC type as discussed in Section 4.1.1 above.
- . Block data length, an unsigned integer. The block data length field SHALL contain the aggregate length of all remaining fields of the block, i.e., the block-type-specific data fields. Block data length SHALL be represented as a CBOR unsigned integer item.
- . Block-type-specific data fields, whose nature and order are type-specific and whose aggregate length in octets is the value of the block data length field. For the Payload Block in particular (block type 1), there SHALL be exactly one block-type-specific data field, termed the "payload", which SHALL be an application data unit, or some contiguous extent thereof, represented as a CBOR byte string.
- . If and only if the value of the CRC type field of this block is non-zero, a CRC. If present, the length and nature of the CRC SHALL be as indicated by the CRC type and the CRC SHALL be computed over the concatenation of all bytes of the block (including CBOR "break" characters) including the CRC field itself, which for this purpose SHALL be temporarily populated with the value zero.

4.3. Extension Blocks

"Extension blocks" are all blocks other than the primary and payload blocks. Because not all extension blocks are defined in the Bundle Protocol specification (the present document), not all nodes conforming to this specification will necessarily instantiate Bundle Protocol implementations that include procedures for processing (that is, recognizing, parsing, acting on, and/or producing) all extension blocks. It is therefore possible for a node to receive a bundle that includes extension blocks that the node cannot process.

The values of the block processing control flags indicate the action to be taken by the bundle protocol agent when this is the case.

(Note that, while CBOR permits considerable flexibility in the encoding of bundles, this flexibility must not be interpreted as inviting increased complexity in protocol data unit structure.)

The following extension blocks are defined in other DTN protocol specification documents as noted:

- . Block Integrity Block (block type 2) and Block Confidentiality Block (block type 3) are defined in the Bundle Security Protocol specification (work in progress).
- . Manifest Block (block type 4) is defined in the Manifest Extension Block specification (work in progress). The manifest block identifies the blocks that were present in the bundle at the time it was created. The bundle MUST contain one (1) occurrence of this type of block if the value of the "manifest" flag in the bundle processing control flags is 1; otherwise the bundle MUST NOT contain any Manifest block.
- . The Flow Label Block (block type 6) is defined in the Flow Label Extension Block specification (work in progress). The flow label block is intended to govern transmission of the bundle by convergence-layer adapters.

The following extension blocks are defined in the current document.

4.3.1. Previous Node

The Previous Node block, block type 7, identifies the node that forwarded this bundle to the local node (i.e., to the node at which the bundle currently resides); its block-type-specific data is the node ID of that forwarder node which SHALL take the form of a node ID represented as described in Section 4.1.5.2. above. If the local node is the source of the bundle, then the bundle MUST NOT contain any previous node block. Otherwise the bundle SHOULD contain one (1) occurrence of this type of block.

4.3.2. Bundle Age

The Bundle Age block, block type 8, contains the number of seconds that have elapsed between the time the bundle was created and time at which it was most recently forwarded. It is intended for use by nodes lacking access to an accurate clock, to aid in determining the time at which a bundle's lifetime expires. The block-type-specific data of this block is an unsigned integer containing the age of the bundle in seconds, which SHALL be represented as a CBOR unsigned

integer item. (The age of the bundle is the sum of all known intervals of the bundle's residence at forwarding nodes, up to the time at which the bundle was most recently forwarded, plus the summation of signal propagation time over all episodes of transmission between forwarding nodes. Determination of these values is an implementation matter.) If the bundle's creation time is zero, then the bundle MUST contain exactly one (1) occurrence of this type of block; otherwise, the bundle MAY contain at most one (1) occurrence of this type of block. A bundle MUST NOT contain multiple occurrences of the bundle age block, as this could result in processing anomalies.

4.3.3. Hop Count

The Hop Count block, block type 9, contains two unsigned integers, hop limit and hop count. A "hop" is here defined as an occasion on which a bundle was forwarded from one node to another node. The hop limit value SHOULD NOT be changed at any time after creation of the Hop Count block; the hop count value SHOULD initially be zero and SHOULD be increased by 1 on each hop.

The hop count block is mainly intended as a safety mechanism, a means of identifying bundles for removal from the network that can never be delivered due to a persistent forwarding error. When a bundle's hop count exceeds its hop limit, the bundle SHOULD be deleted for the reason "hop limit exceeded", following the bundle deletion procedure defined in Section 5.10. . Procedures for determining the appropriate hop limit for a block are beyond the scope of this specification. The block-type-specific data in a hop count block SHALL be represented as a CBOR array comprising a 2-tuple. The first item of this array SHALL be the bundle's hop limit, represented as a CBOR unsigned integer. The second item of this array SHALL be the bundle's hop count, represented as a CBOR unsigned integer. A bundle MAY contain at most one (1) occurrence of this type of block.

5. Bundle Processing

The bundle processing procedures mandated in this section and in Section 6 govern the operation of the Bundle Protocol Agent and the Application Agent administrative element of each bundle node. They are neither exhaustive nor exclusive. Supplementary DTN protocol specifications (including, but not restricted to, the Bundle Security Protocol [BPSEC]) may augment, override, or supersede the mandates of this document.

5.1. Generation of Administrative Records

All transmission of bundles is in response to bundle transmission requests presented by nodes' application agents. When required to "generate" an administrative record (such as a bundle status report), the bundle protocol agent itself is responsible for causing a new bundle to be transmitted, conveying that record. In concept, the bundle protocol agent discharges this responsibility by directing the administrative element of the node's application agent to construct the record and request its transmission as detailed in Section 6 below. In practice, the manner in which administrative record generation is accomplished is an implementation matter, provided the constraints noted in Section 6 are observed.

Under some circumstances, the requesting of status reports could result in an unacceptable increase in the bundle traffic in the network. For this reason, the generation of status reports **MUST** be disabled by default and enabled only when the risk of excessive network traffic is deemed acceptable.

When the generation of status reports is enabled, the decision on whether or not to generate a requested status report is left to the discretion of the bundle protocol agent. Mechanisms that could assist in making such decisions, such as pre-placed agreements authorizing the generation of status reports under specified circumstances, are beyond the scope of this specification.

Notes on administrative record terminology:

- . A "bundle reception status report" is a bundle status report with the "reporting node received bundle" flag set to 1.
- . A "bundle forwarding status report" is a bundle status report with the "reporting node forwarded the bundle" flag set to 1.
- . A "bundle delivery status report" is a bundle status report with the "reporting node delivered the bundle" flag set to 1.
- . A "bundle deletion status report" is a bundle status report with the "reporting node deleted the bundle" flag set to 1.

5.2. Bundle Transmission

The steps in processing a bundle transmission request are:

Step 1: Transmission of the bundle is initiated. An outbound bundle **MUST** be created per the parameters of the bundle transmission request, with the retention constraint "Dispatch pending". The source node ID of the bundle **MUST** be either the null endpoint ID, indicating that the source of the bundle is anonymous, or else the

EID of a singleton endpoint whose only member is the node of which the BPA is a component.

Step 2: Processing proceeds from Step 1 of Section 5.4.

5.3. Bundle Dispatching

The steps in dispatching a bundle are:

Step 1: If the bundle's destination endpoint is an endpoint of which the node is a member, the bundle delivery procedure defined in Section 5.7 MUST be followed.

Step 2: Processing proceeds from Step 1 of Section 5.4.

5.4. Bundle Forwarding

The steps in forwarding a bundle are:

Step 1: The retention constraint "Forward pending" MUST be added to the bundle, and the bundle's "Dispatch pending" retention constraint MUST be removed.

Step 2: The bundle protocol agent MUST determine whether or not forwarding is contraindicated for any of the reasons listed in Figure 4. In particular:

- . The bundle protocol agent MAY choose either to forward the bundle directly to its destination node(s) (if possible) or to forward the bundle to some other node(s) for further forwarding. The manner in which this decision is made may depend on the scheme name in the destination endpoint ID and/or on other state but in any case is beyond the scope of this document. If the BPA elects to forward the bundle to some other node(s) for further forwarding but finds it impossible to select any node(s) to forward the bundle to, then forwarding is contraindicated.
- . Provided the bundle protocol agent succeeded in selecting the node(s) to forward the bundle to, the bundle protocol agent MUST select the convergence layer adapter(s) whose services will enable the node to send the bundle to those nodes. The manner in which specific appropriate convergence layer adapters are selected is beyond the scope of this document. If the agent finds it impossible to select any appropriate convergence layer adapter(s) to use in forwarding this bundle, then forwarding is contraindicated.

Step 3: If forwarding of the bundle is determined to be contraindicated for any of the reasons listed in Figure 4, then the Forwarding Contraindicated procedure defined in Section 5.4.1 MUST be followed; the remaining steps of Section 5 are skipped at this time.

Step 4: For each node selected for forwarding, the bundle protocol agent MUST invoke the services of the selected convergence layer adapter(s) in order to effect the sending of the bundle to that node. Determining the time at which the bundle protocol agent invokes convergence layer adapter services is a BPA implementation matter. Determining the time at which each convergence layer adapter subsequently responds to this service invocation by sending the bundle is a convergence-layer adapter implementation matter. Note that:

- . If the bundle contains a flow label extension block (to be defined in a future document) then that flow label value MAY identify procedures for determining the order in which convergence layer adapters must send bundles, e.g., considering bundle source when determining the order in which bundles are sent. The definition of such procedures is beyond the scope of this specification.
- . If the bundle has a bundle age block, as defined in 4.3.2. above, then at the last possible moment before the CLA initiates conveyance of the bundle node via the CL protocol the bundle age value MUST be increased by the difference between the current time and the time at which the bundle was received (or, if the local node is the source of the bundle, created).

Step 5: When all selected convergence layer adapters have informed the bundle protocol agent that they have concluded their data sending procedures with regard to this bundle:

- . If the "request reporting of bundle forwarding" flag in the bundle's status report request field is set to 1, and status reporting is enabled, then a bundle forwarding status report SHOULD be generated, destined for the bundle's report-to endpoint ID. The reason code on this bundle forwarding status report MUST be "no additional information".
- . If any applicable bundle protocol extensions mandate generation of status reports upon conclusion of convergence-layer data sending procedures, all such status reports SHOULD be generated with extension-mandated reason codes.
- . The bundle's "Forward pending" retention constraint MUST be removed.

5.4.1. Forwarding Contraindicated

The steps in responding to contraindication of forwarding are:

Step 1: The bundle protocol agent MUST determine whether or not to declare failure in forwarding the bundle. Note: this decision is likely to be influenced by the reason for which forwarding is contraindicated.

Step 2: If forwarding failure is declared, then the Forwarding Failed procedure defined in Section 5.4.2 MUST be followed.

Otherwise, when -- at some future time - the forwarding of this bundle ceases to be contraindicated, processing proceeds from Step 4 of Section 5.4.

5.4.2. Forwarding Failed

The steps in responding to a declaration of forwarding failure are:

Step 1: The bundle protocol agent MAY forward the bundle back to the node that sent it, as identified by the Previous Node block, if present.

Step 2: If the bundle's destination endpoint is an endpoint of which the node is a member, then the bundle's "Forward pending" retention constraint MUST be removed. Otherwise, the bundle MUST be deleted: the bundle deletion procedure defined in Section 5.14 MUST be followed, citing the reason for which forwarding was determined to be contraindicated.

5.5. Bundle Expiration

A bundle expires when the bundle's age exceeds its lifetime as specified in the primary bundle block. Bundle age MAY be determined by subtracting the bundle's creation timestamp time from the current time if (a) that timestamp time is not zero and (b) the local node's clock is known to be accurate (as discussed in section 4.5.1 above); otherwise bundle age MUST be obtained from the Bundle Age extension block. Bundle expiration MAY occur at any point in the processing of a bundle. When a bundle expires, the bundle protocol agent MUST delete the bundle for the reason "lifetime expired": the bundle deletion procedure defined in Section 5.14 MUST be followed.

5.6. Bundle Reception

The steps in processing a bundle that has been received from another node are:

Step 1: The retention constraint "Dispatch pending" MUST be added to the bundle.

Step 2: If the "request reporting of bundle reception" flag in the bundle's status report request field is set to 1, and status reporting is enabled, then a bundle reception status report with reason code "No additional information" SHOULD be generated, destined for the bundle's report-to endpoint ID.

Step 3: For each block in the bundle that is an extension block that the bundle protocol agent cannot process:

- . If the block processing flags in that block indicate that a status report is requested in this event, and status reporting is enabled, then a bundle reception status report with reason code "Block unintelligible" SHOULD be generated, destined for the bundle's report-to endpoint ID.
- . If the block processing flags in that block indicate that the bundle must be deleted in this event, then the bundle protocol agent MUST delete the bundle for the reason "Block unintelligible"; the bundle deletion procedure defined in Section 5.14 MUST be followed and all remaining steps of the bundle reception procedure MUST be skipped.
- . If the block processing flags in that block do NOT indicate that the bundle must be deleted in this event but do indicate that the block must be discarded, then the bundle protocol agent MUST remove this block from the bundle.
- . If the block processing flags in that block indicate neither that the bundle must be deleted nor that that the block must be discarded, then processing continues with the next extension block that the bundle protocol agent cannot process, if any; otherwise, processing proceeds from step 4.

Step 4: Processing proceeds from Step 1 of Section 5.3.

5.7. Local Bundle Delivery

The steps in processing a bundle that is destined for an endpoint of which this node is a member are:

Step 1: If the received bundle is a fragment, the application data unit reassembly procedure described in Section 5.9 MUST be followed.

If this procedure results in reassembly of the entire original application data unit, processing of this bundle (whose fragmentary payload has been replaced by the reassembled application data unit) proceeds from Step 2; otherwise, the retention constraint "Reassembly pending" MUST be added to the bundle and all remaining steps of this procedure MUST be skipped.

Step 2: Delivery depends on the state of the registration whose endpoint ID matches that of the destination of the bundle:

- . An additional implementation-specific delivery deferral procedure MAY optionally be associated with the registration.
- . If the registration is in the Active state, then the bundle MUST be delivered automatically as soon as it is the next bundle that is due for delivery according to the BPA's bundle delivery scheduling policy, an implementation matter.
- . If the registration is in the Passive state, or if delivery of the bundle fails for some implementation-specific reason, then the registration's delivery failure action MUST be taken. Delivery failure action MUST be one of the following:
 - o defer delivery of the bundle subject to this registration until (a) this bundle is the least recently received of all bundles currently deliverable subject to this registration and (b) either the registration is polled or else the registration is in the Active state, and also perform any additional delivery deferral procedure associated with the registration; or
 - o abandon delivery of the bundle subject to this registration (as defined in 3.1.).

Step 3: As soon as the bundle has been delivered, if the "request reporting of bundle delivery" flag in the bundle's status report request field is set to 1 and bundle status reporting is enabled, then a bundle delivery status report SHOULD be generated, destined for the bundle's report-to endpoint ID. Note that this status report only states that the payload has been delivered to the application agent, not that the application agent has processed that payload.

5.8. Bundle Fragmentation

It may at times be advantageous for bundle protocol agents to reduce the sizes of bundles in order to forward them. This might be the case, for example, if a node to which a bundle is to be forwarded is accessible only via intermittent contacts and no upcoming contact is long enough to enable the forwarding of the entire bundle.

The size of a bundle can be reduced by "fragmenting" the bundle. To fragment a bundle whose payload is of size M is to replace it with two "fragments" -- new bundles with the same source node ID and creation timestamp as the original bundle -- whose payloads are the first N and the last $(M - N)$ bytes of the original bundle's payload, where $0 < N < M$. Note that fragments may themselves be fragmented, so fragmentation may in effect replace the original bundle with more than two fragments. (However, there is only one 'level' of fragmentation, as in IP fragmentation.)

Any bundle whose primary block's bundle processing flags do NOT indicate that it must not be fragmented MAY be fragmented at any time, for any purpose, at the discretion of the bundle protocol agent. NOTE, however, that some combinations of bundle fragmentation, replication, and routing might result in unexpected traffic patterns.

Fragmentation SHALL be constrained as follows:

- . The concatenation of the payloads of all fragments produced by fragmentation MUST always be identical to the payload of the fragmented bundle (that is, the bundle that is being fragmented). Note that the payloads of fragments resulting from different fragmentation episodes, in different parts of the network, may be overlapping subsets of the fragmented bundle's payload.
- . The primary block of each fragment MUST differ from that of the fragmented bundle, in that the bundle processing flags of the fragment MUST indicate that the bundle is a fragment and both fragment offset and total application data unit length must be provided. Additionally, the CRC of the primary block of the fragmented bundle, if any, MUST be replaced in each fragment by a new CRC computed for the primary block of that fragment.
- . The payload blocks of fragments will differ from that of the fragmented bundle as noted above.
- . If the fragmented bundle is not a fragment or is the fragment with offset zero, then all extension blocks of the fragmented bundle MUST be replicated in the fragment whose offset is zero.
- . Each of the fragmented bundle's extension blocks whose "Block must be replicated in every fragment" flag is set to 1 MUST be replicated in every fragment.
- . Beyond these rules, replication of extension blocks in the fragments is an implementation matter.

5.9. Application Data Unit Reassembly

If the concatenation -- as informed by fragment offsets and payload lengths -- of the payloads of all previously received fragments with the same source node ID and creation timestamp as this fragment, together with the payload of this fragment, forms a byte array whose length is equal to the total application data unit length in the fragment's primary block, then:

- . This byte array -- the reassembled application data unit -- MUST replace the payload of this fragment.
- . The "Reassembly pending" retention constraint MUST be removed from every other fragment whose payload is a subset of the reassembled application data unit.

Note: reassembly of application data units from fragments occurs at the nodes that are members of destination endpoints as necessary; an application data unit MAY also be reassembled at some other node on the path to the destination.

5.10. Bundle Deletion

The steps in deleting a bundle are:

Step 1: If the "request reporting of bundle deletion" flag in the bundle's status report request field is set to 1, and if status reporting is enabled, then a bundle deletion status report citing the reason for deletion SHOULD be generated, destined for the bundle's report-to endpoint ID.

Step 2: All of the bundle's retention constraints MUST be removed.

5.11. Discarding a Bundle

As soon as a bundle has no remaining retention constraints it MAY be discarded, thereby releasing any persistent storage that may have been allocated to it.

5.12. Canceling a Transmission

When requested to cancel a specified transmission, where the bundle created upon initiation of the indicated transmission has not yet been discarded, the bundle protocol agent MUST delete that bundle for the reason "transmission cancelled". For this purpose, the procedure defined in Section 5.14 MUST be followed.

6. Administrative Record Processing

6.1. Administrative Records

Administrative records are standard application data units that are used in providing some of the features of the Bundle Protocol. One type of administrative record has been defined to date: bundle status reports. Note that additional types of administrative records may be defined by supplementary DTN protocol specification documents.

Every administrative record consists of:

- . Record type code (an unsigned integer for which valid values are as defined below).
- . Record content in type-specific format.

Valid administrative record type codes are defined as follows:

Value	Meaning
1	Bundle status report.
(other)	Reserved for future use.

Figure 3: Administrative Record Type Codes

Each BP administrative record SHALL be represented as a CBOR array comprising a 2-tuple.

The first item of the array SHALL be a record type code, which SHALL be represented as a CBOR unsigned integer.

The second element of this array SHALL be the applicable CBOR representation of the content of the record. Details of the CBOR representation of administrative record type 1 are provided below. Details of the CBOR representation of other types of administrative record type are included in the specifications defining those records.

6.1.1.1. Bundle Status Reports

The transmission of "bundle status reports" under specified conditions is an option that can be invoked when transmission of a bundle is requested. These reports are intended to provide information about how bundles are progressing through the system, including notices of receipt, forwarding, final delivery, and deletion. They are transmitted to the Report-to endpoints of bundles.

Each bundle status report SHALL be represented as a CBOR array. The number of elements in the array SHALL be either 6 (if the subject bundle is a fragment) or 4 (otherwise).

The first item of the bundle status report array SHALL be bundle status information represented as a CBOR array of at least 4 elements. The first four items of the bundle status information array shall provide information on the following four status assertions, in this order:

- . Reporting node received bundle.
- . Reporting node forwarded the bundle.
- . Reporting node delivered the bundle.
- . Reporting node deleted the bundle.

Each item of the bundle status information array SHALL be a bundle status item represented as a CBOR array; the number of elements in each such array SHALL be either 2 (if the value of the first item of this bundle status item is 1 AND the "Report status time" flag was set to 1 in the bundle processing flags of the bundle whose status is being reported) or 1 (otherwise). The first item of the bundle status item array SHALL be a status indicator, a Boolean value indicating whether or not the corresponding bundle status is asserted, represented as a CBOR Boolean value. The second item of the bundle status item array, if present, SHALL indicate the time (as reported by the local system clock, an implementation matter) at which the indicated status was asserted for this bundle, represented as a DTN time as described in Section 4.1.6. above.

The second item of the bundle status report array SHALL be the bundle status report reason code explaining the value of the status indicator, represented as a CBOR unsigned integer. Valid status report reason codes are defined in Figure 4 below but the list of status report reason codes provided here is neither exhaustive nor exclusive; supplementary DTN protocol specifications (including, but not restricted to, the Bundle Security Protocol [BPSEC]) may define additional reason codes.

Value	Meaning
0	No additional information.
1	Lifetime expired.
2	Forwarded over unidirectional link.
3	Transmission canceled.
4	Depleted storage.
5	Destination endpoint ID unintelligible.
6	No known route to destination from here.
7	No timely contact with next node on route.
8	Block unintelligible.
9	Hop limit exceeded.
(other)	Reserved for future use.

defined minimal set of services to the bundle protocol agent. This convergence layer service specification enumerates those services.

7.2. Summary of Convergence Layer Services

Each convergence layer protocol adapter is expected to provide the following services to the bundle protocol agent:

- . sending a bundle to a bundle node that is reachable via the convergence layer protocol;
- . delivering to the bundle protocol agent a bundle that was sent by a bundle node via the convergence layer protocol.

The convergence layer service interface specified here is neither exhaustive nor exclusive. That is, supplementary DTN protocol specifications (including, but not restricted to, the Bundle Security Protocol [BPSEC]) may expect convergence layer adapters that serve BP implementations conforming to those protocols to provide additional services such as reporting on the transmission and/or reception progress of individual bundles (at completion and/or incrementally), retransmitting data that were lost in transit, discarding bundle-conveying data units that the convergence layer protocol determines are corrupt or inauthentic, or reporting on the integrity and/or authenticity of delivered bundles.

8. Implementation Status

[NOTE to the RFC Editor: please remove this section before publication, as well as the reference to RFC 7942.]

This section records the status of known implementations of the protocol defined by this specification at the time of posting of this Internet-Draft, and is based on a proposal described in RFC 7942. The description of implementations in this section is intended to assist the IETF in its decision processes in progressing drafts to RFCs. Please note that the listing of any individual implementation here does not imply endorsement by the IETF. Furthermore, no effort has been spent to verify the information presented here that was supplied by IETF contributors. This is not intended as, and must not be construed to be, a catalog of available implementations or their features. Readers are advised to note that other implementations may exist.

According to RFC 7942, "this will allow reviewers and working groups to assign due consideration to documents that have the benefit of running code, which may serve as evidence of valuable experimentation and feedback that have made the implemented

protocols more mature. It is up to the individual working groups to use this information as they see fit".

At the time of this writing, the only known implementation of the current document is microPCN (<https://upcn.eu/>). According to the developers:

The Micro Planetary Communication Network (uPCN) is a free software project intended to offer an implementation of Delay-tolerant Networking protocols for POSIX operating systems (well, and for Linux) plus for the ARM Cortex STM32F4 microcontroller series. More precisely it currently provides an implementation of

- . the Bundle Protocol (BP, RFC 5050),
- . the Bundle Protocol version 7 specification draft (version 6),
- . the DTN IP Neighbor Discovery (IPND) protocol, and
- . a routing approach optimized for message-ferry micro LEO satellites.

uPCN is written in C and is built upon the real-time operating system FreeRTOS. The source code of uPCN is released under the "BSD 3-Clause License".

The project depends on an execution environment offering link layer protocols such as AX.25. The source code uses the USB subsystem to interact with the environment.

9. Security Considerations

The bundle protocol security architecture and the available security services are specified in an accompanying document, the Bundle Security Protocol specification [BPSEC].

The bpsec extensions to Bundle Protocol enable each block of a bundle (other than a bpsec extension block) to be individually authenticated by a signature block (Block Integrity Block, or BIB) and also enable each block of a bundle other than the primary block (and the bpsec extension blocks themselves) to be individually encrypted by a BCB.

Because the security mechanisms are extension blocks that are themselves inserted into the bundle, the integrity and confidentiality of bundle blocks are protected while the bundle is at rest, awaiting transmission at the next forwarding opportunity, as well as in transit.

Additionally, convergence-layer protocols that ensure authenticity of communication between adjacent nodes in BP network topology SHOULD be used where available, to minimize the ability of unauthenticated nodes to introduce inauthentic traffic into the network.

Note that, while the primary block must remain in the clear for routing purposes, the Bundle Protocol can be protected against traffic analysis to some extent by using bundle-in-bundle encapsulation to tunnel bundles to a safe forward distribution point: the encapsulated bundle forms the payload of an encapsulating bundle, and that payload block may be encrypted by a BCB.

Note that the generation of bundle status reports is disabled by default because malicious initiation of bundle status reporting could result in the transmission of extremely large numbers of bundle, effecting a denial of service attack.

The bpsec extensions accommodate an open-ended range of ciphersuites; different ciphersuites may be utilized to protect different blocks. One possible variation is to sign and/or encrypt blocks in symmetric keys securely formed by Diffie-Hellman procedures (such as EKDH) using the public and private keys of the sending and receiving nodes. For this purpose, the key distribution problem reduces to the problem of trustworthy delay-tolerant distribution of public keys, a current research topic.

Bundle security MUST NOT be invalidated by forwarding nodes even though they themselves might not use the Bundle Security Protocol.

In particular, while blocks MAY be added to bundles transiting intermediate nodes, removal of blocks with the "Discard block if it can't be processed" flag set in the block processing control flags may cause security to fail.

Inclusion of the Bundle Security Protocol in any Bundle Protocol implementation is RECOMMENDED. Use of the Bundle Security Protocol in Bundle Protocol operations is OPTIONAL, subject to the following guidelines:

- . Every block (that is not a bpsec extension block) of every bundle SHOULD be authenticated by a BIB citing the ID of the node that inserted that block. (Note that a single BIB may authenticate multiple "target" blocks.) BIB authentication MAY be omitted on (and only on) any initial end-to-end path segments on which it would impose unacceptable overhead, provided that satisfactory authentication is ensured at the

convergence layer and that BIB authentication is asserted on the first path segment on which the resulting overhead is acceptable and on all subsequent path segments.

- . If any segment of the end-to-end path of a bundle will traverse the Internet or any other potentially insecure communication environment, then the payload block SHOULD be encrypted by a BCB on this path segment and all subsequent segments of the end-to-end path.

10. IANA Considerations

This document defines the following additional Bundle Protocol block types, for which values are to be assigned from the Bundle Administrative Record Types namespace [RFC6255]:

Value	Name	Meaning	Reference
7	Previous node	Identifies sender	This document
8	Bundle age	Bundle age in seconds	This document
9	Hop count	#prior transmission attempts	This document

This document also defines a new URI scheme type field - an unsigned integer of undefined length - for which IANA is to create and maintain a new registry named "URI scheme type values". Initial values for the Bundle Protocol URI scheme type registry are given below; future assignments are to be made through Expert Review. Each assignment consists of a URI scheme type name and its associated value.

Value	URI Scheme Type Name	Reference
0	Reserved	
1	dtn	RFC5050, Section 4.4
2	ipn	RFC6260, Section 4
3-254	Unassigned	
255	Reserved	

11. References

11.1. Normative References

[CRC] International Telecommunication Union, "Error-correcting procedures for DCEs using asynchronous-to-synchronous conversion", ITU-T Recommendation V.42, March 2002.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC7049] Borman, C. and P. Hoffman, "Concise Binary Object Representation (CBOR)", RFC 7049, October 2013.

[URI] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", RFC 3986, STD 66, January 2005.

[URIREG] Thaler, D., Hansen, T., and T. Hardie, "Guidelines and Registration Procedures for URI Schemes", RFC 7595, BCP 35, June 2015.

11.2. Informative References

[ARCH] V. Cerf et al., "Delay-Tolerant Network Architecture", RFC 4838, April 2007.

[BPSEC] Birrane, E., "Bundle Security Protocol Specification", Work In Progress, October 2015.

[RFC3987] Duerst, M. and M. Suignard, "Internationalized Resource Identifiers (IRIs)", RFC 3987, January 2005.

[RFC6255] Blanchet, M., "Delay-Tolerant Networking Bundle Protocol IANA Registries", RFC 6255, May 2011.

[SIGC] Fall, K., "A Delay-Tolerant Network Architecture for Challenged Internets", SIGCOMM 2003.

[UTC] Arias, E. and B. Guinot, "Coordinated universal time UTC: historical background and perspectives" in "Journées systemes de reference spatio-temporels", 2004.

12. Acknowledgments

This work is freely adapted from RFC 5050, which was an effort of the Delay Tolerant Networking Research Group. The following DTNRG participants contributed significant technical material and/or inputs to that document: Dr. Vinton Cerf of Google, Scott Burleigh, Adrian Hooke, and Leigh Torgerson of the Jet Propulsion Laboratory, Michael Demmer of the University of California at Berkeley, Robert Durst, Keith Scott, and Susan Symington of The MITRE Corporation, Kevin Fall of Carnegie Mellon University, Stephen Farrell of Trinity College Dublin, Peter Lovell of SPARTA, Inc., Manikantan Ramadas of Ohio University, and Howard Weiss of SPARTA, Inc.

This document was prepared using 2-Word-v2.0.template.dot.

13. Significant Changes from RFC 5050

Points on which this draft significantly differs from RFC 5050 include the following:

- . Clarify the difference between transmission and forwarding.
- . Migrate custody transfer to the bundle-in-bundle encapsulation specification.
- . Introduce the concept of "node ID" as functionally distinct from endpoint ID, while having the same syntax.
- . Restructure primary block, making it immutable. Add optional CRC.
- . Add optional CRCs to non-primary blocks.
- . Add block ID number to canonical block format (to support streamlined BSP).
- . Add bundle age extension block, defined in this specification.
- . Add previous node extension block, defined in this specification.
- . Add flow label extension block, *not* defined in this specification.
- . Add manifest extension block, *not* defined in this specification.
- . Add hop count extension block, defined in this specification.
- . Migrate Quality of Service markings to a new QoS extension block, *not* defined in this specification.

Appendix A.

For More Information

Please refer comments to dtn@ietf.org. The Delay Tolerant Networking Research Group (DTNRG) Web site is located at <http://www.dtnrg.org>.

Copyright (c) 2017 IETF Trust and the persons identified as authors of the code. All rights reserved.

Redistribution and use in source and binary forms, with or without modification, is permitted pursuant to, and subject to the license terms contained in, the Simplified BSD License set forth in Section 4.c of the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>).

Appendix B.

CDDL expression

For informational purposes, Carsten Bormann has kindly provided an expression of the Bundle Protocol specification in the Concise Data Definition Language (CDDL). That CDDL expression is presented below, somewhat edited by the authors. Note that wherever the CDDL expression is in disagreement with the textual representation of the BP specification presented in the earlier sections of this document, the textual representation rules.

```
start = bundle

dtn-time = uint

creation-timestamp = [dtn-time, sequence: uint]

eid-generic = [uri-code, SSP: any]

uri-code = uint

eid = eid-choice .within eid-generic

eid-choice /= [dtn-code, SSP: (text / 0)]

dtn-code = 1 ; TBD

eid-choice /= [ipn-code, SSP: [nodenum: uint, servicenum: uint]]

ipn-code = 2 ; TBD

bundle-control-flags = uint .bits bundleflagbits

bundleflagbits = &(
    reserved: 15
    reserved: 14
    reserved: 13
    bundle-deletion-status-reports-are-requested: 12
    bundle-delivery-status-reports-are-requested: 11
    bundle-forwarding-status-reports-are-requested: 10
    reserved: 9
```

```
bundle-reception-status-reports-are-requested: 8
bundle-contains-a-Manifest-block: 7
status-time-is-requested-in-all-status-reports: 6
user-application-acknowledgement-is-requested: 5
destination-is-a-singleton-endpoint: 4
reserved: 3
bundle-must-not-be-fragmented: 2
payload-is-an-administrative-record: 1
bundle-is-a-fragment: 0
)
crc = bytes
block-control-flags = uint .bits blockflagbits
blockflagbits = &(amp;
    reserved: 7
    reserved: 6
    reserved: 5
    reserved: 4
    bundle-must-be-deleted-if-block-cannot-be-processed: 3
    status-report-must-be-transmitted-if-block-cannot-be-processed: 2
    block-must-be-removed-from-bundle-if-it-cannot-be-processed: 1
    block-must-be-replicated-in-every-fragment: 0
)
bundle = [primary-block, *extension-block, payload-block]
primary-block = [
```

```
    version: 7,  
    bundle-control-flags,  
    crc-type: uint,  
    destination: eid,  
    source-node: eid,  
    report-to: eid,  
    creation-timestamp,  
    lifetime: uint,  
    ? fragment-offset: uint,  
    ? total-application-data-length: uint,  
    ? crc,  
]  
canonical-block-generic = [  
    block-type-code: uint,  
    canonical-block-common,  
    content: any,  
    ? crc  
]  
canonical-block-common = (  
    block-number: uint,  
    block-control-flags,  
    crc-type: uint,  
    block-data-length: uint  
)
```

```
canonical-block = canonical-block-choice .within canonical-block-
generic
canonical-block-choice /= payload-block
payload-block = [1, canonical-block-common, adu-extent: payload]
payload = bytes / bytes .cbor admin-record
canonical-block-choice /= extension-block
extension-block = extension-block-choice .within canonical-block
extension-block-choice /= previous-node-block
previous-node-block = [7, canonical-block-common, eid]
extension-block-choice /= bundle-age-block
bundle-age-block = [8, canonical-block-common, bundle-age: uint]
extension-block-choice /= hop-count-block
hop-count-block = [9, canonical-block-common,
    [hop-limit: uint,
    hop-count: uint]
]
admin-record-generic = [record-type: uint, any]
admin-record = admin-record-choice .within admin-record-generic
admin-record-choice /= bundle-status-report
bundle-status-report = [1, [bundle-status-information,
    bundle-status-reason: uint,
    admin-common]
]
```

```
admin-common = (  
    source-node: eid,  
    creation-timestamp,  
    ? fragment-offset: uint,  
    ? payload-length: uint  
)  
  
bundle-status-information = [  
    reporting-node-received-bundle: bundle-status-item,  
    reporting-node-forwarded-the-bundle: bundle-status-item,  
    reporting-node-delivered-the-bundle: bundle-status-item,  
    reporting-node-deleted-the-bundle: bundle-status-item,  
]  
  
bundle-status-item = [  
    asserted: bool,  
    ? time-of-assertion: dtn-time  
]  
]
```

Authors' Addresses

Scott Burleigh
Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Dr.
Pasadena, CA 91109-8099
US
Phone: +1 818 393 3353
Email: Scott.Burleigh@jpl.nasa.gov

Kevin Fall
Nefeli Networks, Inc.
2150 Shattuck Ave.
Berkeley, CA 94704
US
Email: kfall@kfall.com

Edward J. Birrane
Johns Hopkins University Applied Physics Laboratory
11100 Johns Hopkins Rd
Laurel, MD 20723
US
Phone: +1 443 778 7423
Email: Edward.Birrane@jhuapl.edu

Delay-Tolerant Networking Working Group
Internet Draft
Obsoletes: 5050 and 6255 (if approved)
Intended status: Standards Track
Expires: February 5, 2020

S. Burleigh
JPL, Calif. Inst. Of Technology
K. Fall
Roland Computing Services
E. Birrane
APL, Johns Hopkins University
August 4, 2019

Bundle Protocol Version 7
draft-ietf-dtn-bpbis-14.txt

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>

This Internet-Draft will expire on February 5, 2020.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in

Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Abstract

This Internet Draft presents a specification for Bundle Protocol, adapted from the experimental Bundle Protocol specification developed by the Delay-Tolerant Networking Research group of the Internet Research Task Force and documented in RFC 5050.

It obsoletes RFC 5050 and RFC 6255.

Table of Contents

1. Introduction.....	3
2. Conventions used in this document.....	5
3. Service Description.....	6
3.1. Definitions.....	6
3.2. Discussion of BP concepts.....	9
3.3. Services Offered by Bundle Protocol Agents.....	12
4. Bundle Format.....	13
4.1. BP Fundamental Data Structures.....	13
4.1.1. CRC Type.....	13
4.1.2. CRC.....	14
4.1.3. Bundle Processing Control Flags.....	14
4.1.4. Block Processing Control Flags.....	15
4.1.5. Identifiers.....	16
4.1.5.1. Endpoint ID.....	16
4.1.5.2. Node ID.....	17
4.1.6. DTN Time.....	18
4.1.7. Creation Timestamp.....	20
4.1.8. Block-type-specific Data.....	20
4.2. Bundle Representation.....	20
4.2.1. Bundle.....	20
4.2.2. Primary Bundle Block.....	21
4.2.3. Canonical Bundle Block Format.....	23
4.3. Extension Blocks.....	23
4.3.1. Previous Node.....	24
4.3.2. Bundle Age.....	25
4.3.3. Hop Count.....	25
5. Bundle Processing.....	26
5.1. Generation of Administrative Records.....	26
5.2. Bundle Transmission.....	27
5.3. Bundle Dispatching.....	27
5.4. Bundle Forwarding.....	27
5.4.1. Forwarding Contraindicated.....	29

5.4.2. Forwarding Failed.....	29
5.5. Bundle Expiration.....	30
5.6. Bundle Reception.....	30
5.7. Local Bundle Delivery.....	31
5.8. Bundle Fragmentation.....	32
5.9. Application Data Unit Reassembly.....	33
5.10. Bundle Deletion.....	34
5.11. Discarding a Bundle.....	34
5.12. Canceling a Transmission.....	34
6. Administrative Record Processing.....	34
6.1. Administrative Records.....	34
6.1.1. Bundle Status Reports.....	35
6.2. Generation of Administrative Records.....	38
7. Services Required of the Convergence Layer.....	38
7.1. The Convergence Layer.....	38
7.2. Summary of Convergence Layer Services.....	38
8. Implementation Status.....	39
9. Security Considerations.....	40
10. IANA Considerations.....	42
10.1. Bundle Block Types.....	42
10.2. Primary Bundle Protocol Version.....	43
10.3. Bundle Processing Control Flags.....	44
10.4. Block Processing Control Flags.....	45
10.5. Bundle Status Report Reason Codes.....	46
10.6. URI scheme types.....	47
10.7. New URI scheme "dtn".....	48
11. References.....	50
11.1. Normative References.....	50
11.2. Informative References.....	51
12. Acknowledgments.....	51
13. Significant Changes from RFC 5050.....	52
Appendix A. For More Information.....	53
Appendix B. CDDL expression.....	54

1. Introduction

Since the publication of the Bundle Protocol Specification (Experimental RFC 5050) in 2007, the Delay-Tolerant Networking (DTN) Bundle Protocol has been implemented in multiple programming languages and deployed to a wide variety of computing platforms. This implementation and deployment experience has identified opportunities for making the protocol simpler, more capable, and easier to use. The present document, standardizing the Bundle Protocol (BP), is adapted from RFC 5050 in that context and obsoletes RFC 5050 for that reason. Significant changes from the Bundle Protocol specification defined in RFC 5050 are listed in section 13. In addition, those registry rules defined for RFC 5050

in RFC 6255 that are relevant to the current document have been transcribed into section 10, with modifications as necessary; therefore the rules defined in RFC 6255 are now relevant only to RFC 5050, so in obsoleting RFC 5050 we also obsolete RFC 6255.

This document describes version 7 of BP.

Delay Tolerant Networking is a network architecture providing communications in and/or through highly stressed environments. Stressed networking environments include those with intermittent connectivity, large and/or variable delays, and high bit error rates. To provide its services, BP may be viewed as sitting at the application layer of some number of constituent networks, forming a store-carry-forward overlay network. Key capabilities of BP include:

- . Ability to use physical motility for the movement of data
- . Ability to move the responsibility for error control from one node to another
- . Ability to cope with intermittent connectivity, including cases where the sender and receiver are not concurrently present in the network
- . Ability to take advantage of scheduled, predicted, and opportunistic connectivity, whether bidirectional or unidirectional, in addition to continuous connectivity
- . Late binding of overlay network endpoint identifiers to underlying constituent network addresses

For descriptions of these capabilities and the rationale for the DTN architecture, see [ARCH] and [SIGC].

BP's location within the standard protocol stack is as shown in Figure 1. BP uses underlying "native" transport and/or network protocols for communications within a given constituent network.

The interface between the bundle protocol and a specific underlying protocol is termed a "convergence layer adapter".

Figure 1 shows three distinct transport and network protocols (denoted T1/N1, T2/N2, and T3/N3).

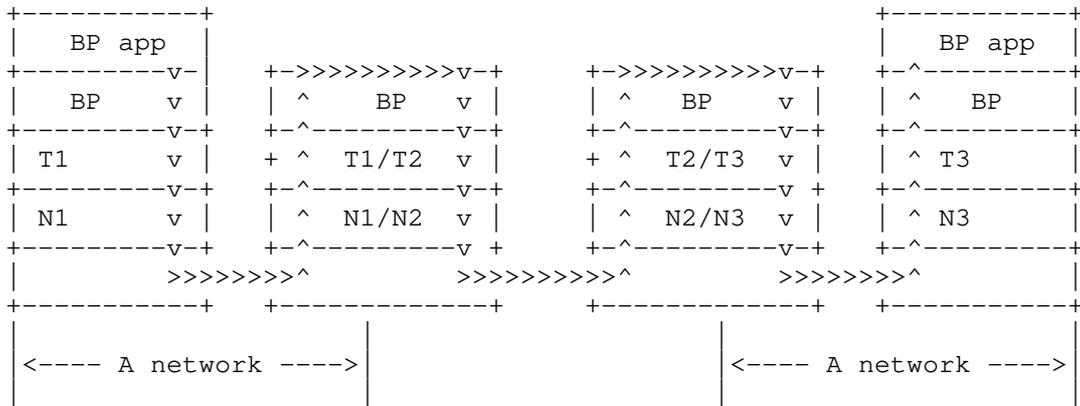


Figure 1: The Bundle Protocol in the Protocol Stack Model

This document describes the format of the protocol data units (called "bundles") passed between entities participating in BP communications.

The entities are referred to as "bundle nodes". This document does not address:

- . Operations in the convergence layer adapters that bundle nodes use to transport data through specific types of internets. (However, the document does discuss the services that must be provided by each adapter at the convergence layer.)
- . The bundle route computation algorithm.
- . Mechanisms for populating the routing or forwarding information bases of bundle nodes.
- . The mechanisms for securing bundles en route.
- . The mechanisms for managing bundle nodes.

Note that implementations of the specification presented in this document will not be interoperable with implementations of RFC 5050.

2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Service Description

3.1. Definitions

Bundle - A bundle is a protocol data unit of BP, so named because negotiation of the parameters of a data exchange may be impractical in a delay-tolerant network: it is often better practice to "bundle" with a unit of application data all metadata that might be needed in order to make the data immediately usable when delivered to the application. Each bundle comprises a sequence of two or more "blocks" of protocol data, which serve various purposes.

Block - A bundle protocol block is one of the protocol data structures that together constitute a well-formed bundle.

Application Data Unit (ADU) - An application data unit is the unit of data whose conveyance to the bundle's destination is the purpose for the transmission of some bundle that is not a fragment (as defined below).

Bundle payload - A bundle payload (or simply "payload") is the content of the bundle's payload block. The terms "bundle content", "bundle payload", and "payload" are used interchangeably in this document. For a bundle that is not a fragment (as defined below), the payload is an application data unit.

Partial payload - A partial payload is a payload that comprises either the first N bytes or the last N bytes of some other payload of length M, such that $0 < N < M$. Note that every partial payload is a payload and therefore can be further subdivided into partial payloads.

Fragment - A fragment is a bundle whose payload block contains a partial payload.

Bundle node - A bundle node (or, in the context of this document, simply a "node") is any entity that can send and/or receive bundles. Each bundle node has three conceptual components, defined below, as shown in Figure 2: a "bundle protocol agent", a set of zero or more "convergence layer adapters", and an "application agent".

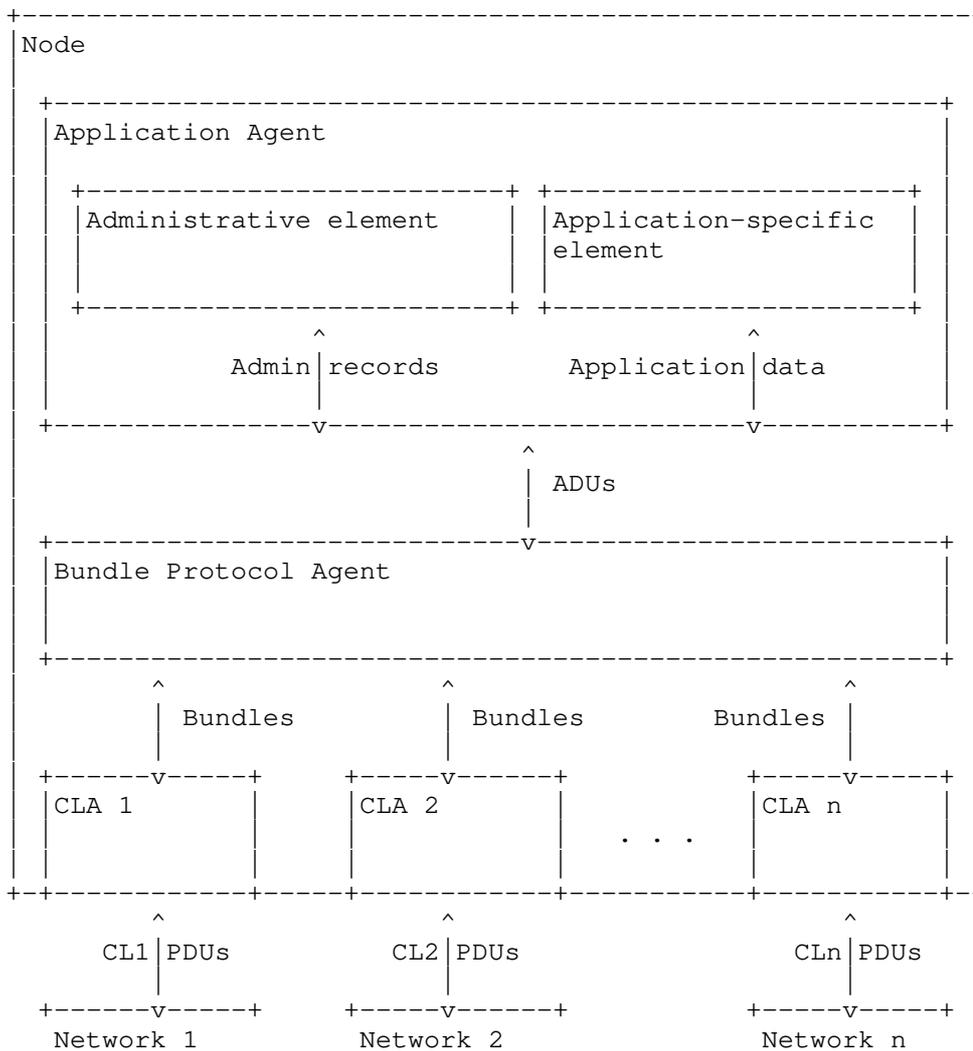


Figure 2: Components of a BP Node

Bundle protocol agent - The bundle protocol agent (BPA) of a node is the node component that offers the BP services and executes the procedures of the bundle protocol.

Convergence layer adapter - A convergence layer adapter (CLA) is a node component that sends and receives bundles on behalf of the BPA, utilizing the services of some 'native' protocol stack that is

supported in one of the networks within which the node is functionally located.

Application agent - The application agent (AA) of a node is the node component that utilizes the BP services to effect communication for some user purpose. The application agent in turn has two elements, an administrative element and an application-specific element.

Application-specific element - The application-specific element of an AA is the node component that constructs, requests transmission of, accepts delivery of, and processes units of user application data.

Administrative element - The administrative element of an AA is the node component that constructs and requests transmission of administrative records (defined below), including status reports, and accepts delivery of and processes any administrative records that the node receives.

Administrative record - A BP administrative record is an application data unit that is exchanged between the administrative elements of nodes' application agents for some BP administrative purpose. The only administrative record defined in this specification is the status report, discussed later.

Bundle endpoint - A bundle endpoint (or simply "endpoint") is a set of zero or more bundle nodes that all identify themselves for BP purposes by some common identifier, called a "bundle endpoint ID" (or, in this document, simply "endpoint ID"; endpoint IDs are described in detail in Section 4.4.1 below).

Singleton endpoint - A singleton endpoint is an endpoint that always contains exactly one member.

Registration - A registration is the state machine characterizing a given node's membership in a given endpoint. Any single registration has an associated delivery failure action as defined below and must at any time be in one of two states: Active or Passive.

Delivery - A bundle is considered to have been delivered at a node subject to a registration as soon as the application data unit that is the payload of the bundle, together with any relevant metadata (an implementation matter), has been presented to the node's application agent in a manner consistent with the state of that registration.

Deliverability - A bundle is considered "deliverable" subject to a registration if and only if (a) the bundle's destination endpoint is the endpoint with which the registration is associated, (b) the bundle has not yet been delivered subject to this registration, and (c) the bundle has not yet been "abandoned" (as defined below) subject to this registration.

Abandonment - To abandon a bundle subject to some registration is to assert that the bundle is not deliverable subject to that registration.

Delivery failure action - The delivery failure action of a registration is the action that is to be taken when a bundle that is "deliverable" subject to that registration is received at a time when the registration is in the Passive state.

Destination - The destination of a bundle is the endpoint comprising the node(s) at which the bundle is to be delivered (as defined below).

Transmission - A transmission is an attempt by a node's BPA to cause copies of a bundle to be delivered to one or more of the nodes that are members of some endpoint (the bundle's destination) in response to a transmission request issued by the node's application agent.

Forwarding - To forward a bundle to a node is to invoke the services of one or more CLAs in a sustained effort to cause a copy of the bundle to be received by that node.

Discarding - To discard a bundle is to cease all operations on the bundle and functionally erase all references to it. The specific procedures by which this is accomplished are an implementation matter.

Retention constraint - A retention constraint is an element of the state of a bundle that prevents the bundle from being discarded. That is, a bundle cannot be discarded while it has any retention constraints.

Deletion - To delete a bundle is to remove unconditionally all of the bundle's retention constraints, enabling the bundle to be discarded.

3.2. Discussion of BP concepts

Multiple instances of the same bundle (the same unit of DTN protocol data) might exist concurrently in different parts of a network --

possibly differing in some blocks -- in the memory local to one or more bundle nodes and/or in transit between nodes. In the context of the operation of a bundle node, a bundle is an instance (copy), in that node's local memory, of some bundle that is in the network.

The payload for a bundle forwarded in response to a bundle transmission request is the application data unit whose location is provided as a parameter to that request. The payload for a bundle forwarded in response to reception of a bundle is the payload of the received bundle.

In the most familiar case, a bundle node is instantiated as a single process running on a general-purpose computer, but in general the definition is meant to be broader: a bundle node might alternatively be a thread, an object in an object-oriented operating system, a special-purpose hardware device, etc.

The manner in which the functions of the BPA are performed is wholly an implementation matter. For example, BPA functionality might be coded into each node individually; it might be implemented as a shared library that is used in common by any number of bundle nodes on a single computer; it might be implemented as a daemon whose services are invoked via inter-process or network communication by any number of bundle nodes on one or more computers; it might be implemented in hardware.

Every CLA implements its own thin layer of protocol, interposed between BP and the (usually "top") protocol(s) of the underlying native protocol stack; this "CL protocol" may only serve to multiplex and de-multiplex bundles to and from the underlying native protocol, or it may offer additional CL-specific functionality. The manner in which a CLA sends and receives bundles, as well as the definitions of CLAs and CL protocols, are beyond the scope of this specification.

Note that the administrative element of a node's application agent may itself, in some cases, function as a convergence-layer adapter. That is, outgoing bundles may be "tunneled" through encapsulating bundles:

- . An outgoing bundle constitutes a byte array. This byte array may, like any other, be presented to the bundle protocol agent as an application data unit that is to be transmitted to some endpoint.
- . The original bundle thus forms the payload of an encapsulating bundle that is forwarded using some other convergence-layer protocol(s).

- . When the encapsulating bundle is received, its payload is delivered to the peer application agent administrative element, which then instructs the bundle protocol agent to dispatch that original bundle in the usual way.

The purposes for which this technique may be useful (such as cross-domain security) are beyond the scope of this specification.

The only interface between the BPA and the application-specific element of the AA is the BP service interface. But between the BPA and the administrative element of the AA there is a (conceptual) private control interface in addition to the BP service interface. This private control interface enables the BPA and the administrative element of the AA to direct each other to take action under specific circumstances.

In the case of a node that serves simply as a BP "router", the AA may have no application-specific element at all. The application-specific elements of other nodes' AAs may perform arbitrarily complex application functions, perhaps even offering multiplexed DTN communication services to a number of other applications. As with the BPA, the manner in which the AA performs its functions is wholly an implementation matter.

Singletons are the most familiar sort of endpoint, but in general the endpoint notion is meant to be broader. For example, the nodes in a sensor network might constitute a set of bundle nodes that identify themselves by a single common endpoint ID and thus form a single bundle endpoint. *Note* too that a given bundle node might identify itself by multiple endpoint IDs and thus be a member of multiple bundle endpoints.

The destination of every bundle is an endpoint, which may or may not be singleton. The source of every bundle is a node, identified by the endpoint ID for some singleton endpoint that contains that node. Note, though, that the source node ID asserted in a given bundle may be the null endpoint ID (as described later) rather than the endpoint ID of the actual source node; bundles for which the asserted source node ID is the null endpoint ID are termed "anonymous" bundles.

Any number of transmissions may be concurrently undertaken by the bundle protocol agent of a given node.

When the bundle protocol agent of a node determines that a bundle must be forwarded to a node (either to a node that is a member of the bundle's destination endpoint or to some intermediate forwarding

node) in the course of completing the successful transmission of that bundle, the bundle protocol agent invokes the services of one or more CLAs in a sustained effort to cause a copy of the bundle to be received by that node.

Upon reception, the processing of a bundle that has been received by a given node depends on whether or not the receiving node is registered in the bundle's destination endpoint. If it is, and if the payload of the bundle is non-fragmentary (possibly as a result of successful payload reassembly from fragmentary payloads, including the original payload of the newly received bundle), then the bundle is normally delivered to the node's application agent subject to the registration characterizing the node's membership in the destination endpoint.

The bundle protocol does not natively ensure delivery of a bundle to its destination. Data loss along the path to the destination node can be minimized by utilizing reliable convergence-layer protocols between neighbors on all segments of the end-to-end path, but for end-to-end bundle delivery assurance it will be necessary to develop extensions to the bundle protocol and/or application-layer mechanisms.

The bundle protocol is designed for extensibility. Bundle protocol extensions, documented elsewhere, may extend this specification by:

- . defining additional blocks;
- . defining additional administrative records;
- . defining additional bundle processing flags;
- . defining additional block processing flags;
- . defining additional types of bundle status reports;
- . defining additional bundle status report reason codes;
- . defining additional mandates and constraints on processing that conformant bundle protocol agents must perform at specified points in the inbound and outbound bundle processing cycles.

3.3. Services Offered by Bundle Protocol Agents

The BPA of each node is expected to provide the following services to the node's application agent:

- . commencing a registration (registering the node in an endpoint);
- . terminating a registration;
- . switching a registration between Active and Passive states;
- . transmitting a bundle to an identified bundle endpoint;

- . canceling a transmission;
- . polling a registration that is in the Passive state;
- . delivering a received bundle.

4. Bundle Format

The format of bundles SHALL conform to the Concise Binary Object Representation (CBOR [RFC7049]).

Each bundle SHALL be a concatenated sequence of at least two blocks, represented as a CBOR indefinite-length array. The first block in the sequence (the first item of the array) MUST be a primary bundle block in CBOR representation as described below; the bundle MUST have exactly one primary bundle block. The primary block MUST be followed by one or more canonical bundle blocks (additional array items) in CBOR representation as described below. The last such block MUST be a payload block; the bundle MUST have exactly one payload block. The last item of the array, immediately following the payload block, SHALL be a CBOR "break" stop code.

(Note that, while CBOR permits considerable flexibility in the encoding of bundles, this flexibility must not be interpreted as inviting increased complexity in protocol data unit structure.)

An implementation of the Bundle Protocol MAY discard any sequence of bytes that does not conform to the Bundle Protocol specification.

An implementation of the Bundle Protocol MAY accept a sequence of bytes that does not conform to the Bundle Protocol specification (e.g., one that represents data elements in fixed-length arrays rather than indefinite-length arrays) and transform it into conformant BP structure before processing it. Procedures for accomplishing such a transformation are beyond the scope of this specification.

4.1. BP Fundamental Data Structures

4.1.1. CRC Type

CRC type is an unsigned integer type code for which the following values (and no others) are valid:

- . 0 indicates "no CRC is present."
- . 1 indicates "a standard X-25 CRC-16 is present." [CRC16]
- . 2 indicates "a standard CRC32C (Castagnoli) CRC-32 is present." [CRC32C]

CRC type SHALL be represented as a CBOR unsigned integer.

For examples of CRC32C CRCs, see Appendix A.4 of [RFC7143].

4.1.2. CRC

CRC SHALL be omitted from a block if and only if the block's CRC type code is zero.

When not omitted, the CRC SHALL be represented as sequence of two bytes (if CRC type is 1) or as a sequence of four bytes (if CRC type is 2); in each case the sequence of bytes SHALL constitute an unsigned integer value (of 16 or 32 bits, respectively) in network byte order.

4.1.3. Bundle Processing Control Flags

Bundle processing control flags assert properties of the bundle as a whole rather than of any particular block of the bundle. They are conveyed in the primary block of the bundle.

The following properties are asserted by the bundle processing control flags:

- . The bundle is a fragment. (Boolean)
- . The bundle's payload is an administrative record. (Boolean)
- . The bundle must not be fragmented. (Boolean)
- . Acknowledgment by the user application is requested. (Boolean)
- . Status time is requested in all status reports. (Boolean)
- . The bundle contains a "manifest" extension block. (Boolean)
- . Flags requesting types of status reports (all Boolean):
 - o Request reporting of bundle reception.
 - o Request reporting of bundle forwarding.
 - o Request reporting of bundle delivery.
 - o Request reporting of bundle deletion.

If the bundle processing control flags indicate that the bundle's application data unit is an administrative record, then all status report request flag values must be zero.

If the bundle's source node is omitted (i.e., the source node ID is the ID of the null endpoint, which has no members as discussed below; this option enables anonymous bundle transmission), then the bundle is not uniquely identifiable and all bundle protocol features that rely on bundle identity must therefore be disabled: the "Bundle must not be fragmented" flag value must be 1 and all status report request flag values must be zero.

The bundle processing control flags SHALL be represented as a CBOR unsigned integer item containing a bit field of 16 bits indicating the control flag values as follows:

- . Bit 0 (the high-order bit, 0x8000): reserved.
- . Bit 1 (0x4000): reserved.
- . Bit 2 (0x2000): reserved.
- . Bit 3 (0x1000): bundle deletion status reports are requested.
- . Bit 4 (0x0800): bundle delivery status reports are requested.
- . Bit 5 (0x0400): bundle forwarding status reports are requested.
- . Bit 6 (0x0200): reserved.
- . Bit 7 (0x0100): bundle reception status reports are requested.
- . Bit 8 (0x0080): reserved.
- . Bit 9 (0x0040): status time is requested in all status reports.
- . Bit 10 (0x0020): user application acknowledgement is requested.
- . Bit 11 (0x0010): reserved.
- . Bit 12 (0x0008): reserved.
- . Bit 13 (0x0004): bundle must not be fragmented.
- . Bit 14 (0x0002): payload is an administrative record.
- . Bit 15 (the low-order bit, 0x0001): bundle is a fragment.

Note: bit 8 is reserved with the intention of using it to indicate the presence of a Manifest extension block, not yet defined.

4.1.4. Block Processing Control Flags

The block processing control flags assert properties of canonical bundle blocks. They are conveyed in the header of the block to which they pertain.

The following properties are asserted by the block processing control flags:

- . This block must be replicated in every fragment. (Boolean)

- . Transmission of a status report is requested if this block can't be processed. (Boolean)
- . Block must be removed from the bundle if it can't be processed. (Boolean)
- . Bundle must be deleted if this block can't be processed. (Boolean)

For each bundle whose bundle processing control flags indicate that the bundle's application data unit is an administrative record, or whose source node ID is the null endpoint ID as defined below, the value of the "Transmit status report if block can't be processed" flag in every canonical block of the bundle must be zero.

The block processing control flags SHALL be represented as a CBOR unsigned integer item containing a bit field of 8 bits indicating the control flag values as follows:

- . Bit 0 (the high-order bit, 0x80): reserved.
- . Bit 1 (0x40): reserved.
- . Bit 2(0x20): reserved.
- . Bit 3(0x10): reserved.
- . Bit 4(0x08): bundle must be deleted if block can't be processed.
- . Bit 5(0x04): transmission of a status report is requested if block can't be processed.
- . Bit 6(0x02): block must be removed from bundle if it can't be processed.
- . Bit 7(the low-order bit, 0x01): block must be replicated in every fragment.

4.1.5. Identifiers

4.1.5.1. Endpoint ID

The destinations of bundles are bundle endpoints, identified by text strings termed "endpoint IDs" (see Section 3.1). Each endpoint ID (EID) is a Uniform Resource Identifier (URI; [URI]). As such, each endpoint ID can be characterized as having this general structure:

< scheme name > : < scheme-specific part, or "SSP" >

The scheme identified by the < scheme name > in an endpoint ID is a set of syntactic and semantic rules that fully explain how to parse and interpret the SSP. The set of allowable schemes is effectively

unlimited. Any scheme conforming to [URIREG] may be used in a bundle protocol endpoint ID.

Note that, although endpoint IDs are URIs, implementations of the BP service interface may support expression of endpoint IDs in some internationalized manner (e.g., Internationalized Resource Identifiers (IRIs); see [RFC3987]).

The endpoint ID "dtn:none" identifies the "null endpoint", the endpoint that by definition never has any members.

Each BP endpoint ID (EID) SHALL be represented as a CBOR array comprising a 2-tuple.

The first item of the array SHALL be the code number identifying the endpoint's URI scheme [URI], as defined in the registry of URI scheme code numbers for Bundle Protocol maintained by IANA as described in Section 10. [URIREG]. Each URI scheme code number SHALL be represented as a CBOR unsigned integer.

The second item of the array SHALL be the applicable CBOR representation of the scheme-specific part (SSP) of the EID, defined as follows:

- . If the EID's URI scheme is "dtn" then the SSP SHALL be represented as a CBOR text string unless the EID's SSP is "none", in which case the SSP SHALL be represented as a CBOR unsigned integer with the value zero.
- . If the EID's URI scheme is "ipn" then the SSP SHALL be represented as a CBOR array comprising a 2-tuple. The first item of this array SHALL be the EID's node number represented as a CBOR unsigned integer. The second item of this array SHALL be the EID's service number represented as a CBOR unsigned integer.
- . Definitions of the CBOR representations of the SSPs of EIDs encoded in other URI schemes are included in the specifications defining those schemes.

4.1.5.2. Node ID

For many purposes of the Bundle Protocol it is important to identify the node that is operative in some context.

As discussed in 3.1 above, nodes are distinct from endpoints; specifically, an endpoint is a set of zero or more nodes. But rather than define a separate namespace for node identifiers, we

instead use endpoint identifiers to identify nodes, subject to the following restrictions:

- . Every node MUST be a member of at least one singleton endpoint.
- . The EID of any singleton endpoint of which a node is a member MAY be used to identify that node. A "node ID" is an EID that is used in this way.
- . A node's membership in a given singleton endpoint MUST be sustained at least until the nominal operation of the Bundle Protocol no longer depends on the identification of that node using that endpoint's ID.

4.1.6. DTN Time

A DTN time is an unsigned integer indicating an interval of Unix epoch time [EPOCH] that has elapsed since the start of the year 2000 on the Coordinated Universal Time (UTC) scale [UTC], which is Unix epoch timestamp 946684800. (Note that the DTN time that equates to the current time as reported by the UNIX time() function can be derived by subtracting 946684800 from that reported time value.) Each DTN time SHALL be represented as a CBOR unsigned integer item.

Note: The choice of Unix epoch time as the scale on which time values in DTN are expressed may need some explanation.

The computation of time intervals is integral to several DTN protocol procedures. Inconsistency in the results of these computations would result in inconsistent performance of those procedures and would compromise the operation of the protocol.

So the key qualities sought in selecting the time scale to be used for expressing DTN times were these: (a) the broadest possible access to the value of the current time on the selected time scale, enabling all nodes of the network to perform protocol procedures in the same way using the same information, and (b) ease of time interval computation.

UTC was an obvious candidate but fell short on both counts. First, millions of devices can readily query the current UTC time, thanks to NTP, but spacecraft operating beyond Earth orbit cannot. There is currently no adaptation of NTP that operates over the long and variable signal propagation delays between vehicles in deep space.

Moreover, computing the number of actual elapsed seconds between two UTC times is non-trivial because UTC times include leap seconds. As an illustration of the issue, consider the passage of UTC and TAI time at a ground station antenna that began transmitting data at

8Kbps around midnight December 31, 2016 (UTC), when a leap second was added (*):

	UTC	TAI	Total bytes sent
t1	2016-12-31 23:59:58	2017-01-01 00:00:34	0
t2	2016-12-31 23:59:59	2017-01-01 00:00:35	1000
t3	2016-12-31 23:59:60*	2017-01-01 00:00:36	2000
t4	2017-01-01 00:00:00	2017-01-01 00:00:37	3000
t5	2017-01-01 00:00:01	2017-01-01 00:00:38	4000

Suppose we must compute the volume of data transmitted in the interval between t1 and t5. If we use TAI time values, the elapsed time interval is 4 seconds (00:00:38 minus 00:00:34); at 8Kbps, the computed transmission volume is 4000 bytes, which is correct. If we instead use UTC time values as stated, without special compensation for the insertion of the leap second, the elapsed time interval is 3 seconds (00:00:01 minus 23:59:58); the computed transmission volume is then 3000 bytes, which is incorrect.

TAI, then, would be an ideal time scale for DTN, as the interval in seconds between two TAI times can be computed by simply subtracting one from the other; there is no need to consult a table of leap seconds each time a time interval is computed. Unfortunately the current value of TAI, as tracked by atomic clocks on Earth and carefully managed by the International Bureau of Weights and Measures, is likewise not directly accessible to spacecraft.

Unix epoch time is the next best option. Like TAI, Unix epoch time is simply a count of seconds elapsed since a standard epoch. Unlike TAI, the current value of Unix epoch time is provided by virtually all operating systems on which BP is likely to run.

Implementers of Bundle Protocol need to be aware that the difference between DTN time and UTC time will increase with the passing years as additional leap seconds are inserted into UTC. Converting DTN time to the correct corresponding UTC time, in the event that such conversion is needed, will require an understanding of the leap second adjustments made to UTC over time; for software written in C, the widely supported `gmtime()` function provides this service.

Implementers also need to be aware that DTN time values conveyed in CBOR representation in bundles can conceivably exceed $(2^{32} - 1)$.

4.1.7. Creation Timestamp

Each creation timestamp SHALL be represented as a CBOR array item comprising a 2-tuple.

The first item of the array SHALL be a DTN time.

The second item of the array SHALL be the creation timestamp's sequence number, represented as a CBOR unsigned integer.

4.1.8. Block-type-specific Data

Block-type-specific data in each block (other than the primary block) SHALL be the applicable CBOR representation of the content of the block. Details of this representation are included in the specification defining the block type.

4.2. Bundle Representation

This section describes the primary block in detail and non-primary blocks in general. Rules for processing these blocks appear in Section 5 of this document.

Note that supplementary DTN protocol specifications (including, but not restricted to, the Bundle Security Protocol [BPSEC]) may require that BP implementations conforming to those protocols construct and process additional blocks.

4.2.1. Bundle

Each bundle SHALL be represented as a CBOR indefinite-length array. The first item of this array SHALL be the CBOR representation of a Primary Block. Every other item of the array except the last SHALL be the CBOR representation of a Canonical Block. The last item of the array SHALL be a CBOR "break" stop code.

Associated with each block of a bundle is a block number. The block number uniquely identifies the block within the bundle, enabling blocks (notably bundle security protocol blocks) to reference other blocks in the same bundle without ambiguity. The block number of the primary block is implicitly zero; the block numbers of all other blocks are explicitly stated in block headers as noted below. Block numbering is unrelated to the order in which blocks are sequenced in the bundle. The block number of the payload block is always 1.

4.2.2. Primary Bundle Block

The primary bundle block contains the basic information needed to forward bundles to their destinations.

Each primary block SHALL be represented as a CBOR array; the number of elements in the array SHALL be 9 (if the bundle is not a fragment) or 11 (if the bundle is a fragment).

The primary block of each bundle SHALL be immutable. The values of all fields in the primary block must remain unchanged from the time the block is created to the time it is delivered.

The fields of the primary bundle block SHALL be as follows, listed in the order in which they MUST appear:

Version: An unsigned integer value indicating the version of the bundle protocol that constructed this block. The present document describes version 7 of the bundle protocol. Version number SHALL be represented as a CBOR unsigned integer item.

Bundle Processing Control Flags: The Bundle Processing Control Flags are discussed in Section 4.1.3. above.

CRC Type: CRC Type codes are discussed in Section 4.1.1. above.

Destination EID: The Destination EID field identifies the bundle endpoint that is the bundle's destination, i.e., the endpoint that contains the node(s) at which the bundle is to be delivered.

Source node ID: The Source node ID field identifies the bundle node at which the bundle was initially transmitted, except that Source node ID may be the null endpoint ID in the event that the bundle's source chooses to remain anonymous.

Report-to EID: The Report-to EID field identifies the bundle endpoint to which status reports pertaining to the forwarding and delivery of this bundle are to be transmitted.

Creation Timestamp: The creation timestamp is a pair of unsigned integers that, together with the source node ID and (if the bundle is a fragment) the fragment offset and payload length, serve to identify the bundle. The first of these integers is the bundle's creation time, while the second is the bundle's creation timestamp sequence number. Bundle creation time shall be the DTN time at which the transmission request was received that resulted in the creation of the bundle. Sequence count shall be the latest value (as of the

time at which that transmission request was received) of a monotonically increasing positive integer counter managed by the source node's bundle protocol agent that may be reset to zero whenever the current time advances by one second. For nodes that lack accurate clocks, it is recommended that bundle creation time be set to zero and that the counter used as the source of the bundle sequence count never be reset to zero. Note that, in general, the creation of two distinct bundles with the same source node ID and bundle creation timestamp may result in unexpected network behavior and/or suboptimal performance. The combination of source node ID and bundle creation timestamp serves to identify a single transmission request, enabling it to be acknowledged by the receiving application (provided the source node ID is not the null endpoint ID).

Lifetime: The lifetime field is an unsigned integer that indicates the time at which the bundle's payload will no longer be useful, encoded as a number of microseconds past the creation time. (For high-rate deployments with very brief disruptions, fine-grained expression of bundle lifetime may be useful.) When a bundle's age exceeds its lifetime, bundle nodes need no longer retain or forward the bundle; the bundle SHOULD be deleted from the network. For bundles originating at nodes that lack accurate clocks, it is recommended that bundle age be obtained from the Bundle Age extension block (see 4.3.2 below) rather than from the difference between current time and bundle creation time. Bundle lifetime SHALL be represented as a CBOR unsigned integer item.

Fragment offset: If and only if the Bundle Processing Control Flags of this Primary block indicate that the bundle is a fragment, fragment offset SHALL be present in the primary block. Fragment offset SHALL be represented as a CBOR unsigned integer indicating the offset from the start of the original application data unit at which the bytes comprising the payload of this bundle were located.

Total Application Data Unit Length: If and only if the Bundle Processing Control Flags of this Primary block indicate that the bundle is a fragment, total application data unit length SHALL be present in the primary block. Total application data unit length SHALL be represented as a CBOR unsigned integer indicating the total length of the original application data unit of which this bundle's payload is a part.

CRC: A CRC SHALL be present in the primary block. The length and nature of the CRC SHALL be as indicated by the CRC type. The CRC SHALL be computed over the concatenation of all bytes (including CBOR "break" characters) of the primary block including the CRC

field itself, which for this purpose SHALL be temporarily populated with the value zero.

4.2.3. Canonical Bundle Block Format

Every block other than the primary block (all such blocks are termed "canonical" blocks) SHALL be represented as a CBOR array; the number of elements in the array SHALL be 5 (if CRC type is zero) or 6 (otherwise).

The fields of every canonical block SHALL be as follows, listed in the order in which they MUST appear:

- . Block type code, an unsigned integer. Bundle block type code 1 indicates that the block is a bundle payload block. Block type codes 2 through 9 are explicitly reserved as noted later in this specification. Block type codes 192 through 255 are not reserved and are available for private and/or experimental use. All other block type code values are reserved for future use.
- . Block number, an unsigned integer as discussed above.
- . Block processing control flags as discussed in Section 4.1.4 above.
- . CRC type as discussed in Section 4.1.1 above.
- . Block-type-specific data represented as a single definite-length CBOR byte string, i.e., a CBOR byte string that is not of indefinite length. For each type of block, the block-type-specific data byte string is the serialization, in a block-type-specific manner, of the data conveyed by that type of block; definitions of blocks are required to define the manner in which block-type-specific data are serialized within the block-type-specific data field. For the Payload Block in particular (block type 1), the block-type-specific data field, termed the "payload", SHALL be an application data unit, or some contiguous extent thereof, represented as a definite-length CBOR byte string.
- . If and only if the value of the CRC type field of this block is non-zero, a CRC. If present, the length and nature of the CRC SHALL be as indicated by the CRC type and the CRC SHALL be computed over the concatenation of all bytes of the block (including CBOR "break" characters) including the CRC field itself, which for this purpose SHALL be temporarily populated with the value zero.

4.3. Extension Blocks

"Extension blocks" are all blocks other than the primary and payload blocks. Because not all extension blocks are defined in the Bundle

Protocol specification (the present document), not all nodes conforming to this specification will necessarily instantiate Bundle Protocol implementations that include procedures for processing (that is, recognizing, parsing, acting on, and/or producing) all extension blocks. It is therefore possible for a node to receive a bundle that includes extension blocks that the node cannot process. The values of the block processing control flags indicate the action to be taken by the bundle protocol agent when this is the case.

Extension block types 2 and 3 are reserved for the Block Integrity Block and Block Confidentiality Block as defined in the Bundle Security Protocol specification [BPSEC].

The following extension block types are reserved for extension blocks for which a need is anticipated but for which no definitions yet exist:

- . Block type 4 is reserved for the anticipated Manifest Block.
Note: it is anticipated that the manifest block will identify the blocks that were present in the bundle at the time it was created, implying that the bundle **MUST** contain one (1) occurrence of this type of block if the value of the "manifest" flag in the bundle processing control flags is 1, but otherwise the bundle **MUST NOT** contain any Manifest block.
- . Block type 5 is reserved for the anticipated Metadata Block.
Note: the structure and function of the anticipated Metadata Block are currently undefined.
- . Block type 6 is reserved for the anticipated Data Label Block.
Note: it is anticipated that the data label block will provide additional information that can assist nodes in making forwarding decisions.

The following extension blocks are defined in the current document.

4.3.1. Previous Node

The Previous Node block, block type 7, identifies the node that forwarded this bundle to the local node (i.e., to the node at which the bundle currently resides); its block-type-specific data is the node ID of that forwarder node which **SHALL** take the form of a node ID represented as described in Section 4.1.5.2. above. If the local node is the source of the bundle, then the bundle **MUST NOT** contain any previous node block. Otherwise the bundle **SHOULD** contain one (1) occurrence of this type of block.

4.3.2. Bundle Age

The Bundle Age block, block type 8, contains the number of microseconds that have elapsed between the time the bundle was created and time at which it was most recently forwarded. It is intended for use by nodes lacking access to an accurate clock, to aid in determining the time at which a bundle's lifetime expires. The block-type-specific data of this block is an unsigned integer containing the age of the bundle in microseconds, which SHALL be represented as a CBOR unsigned integer item. (The age of the bundle is the sum of all known intervals of the bundle's residence at forwarding nodes, up to the time at which the bundle was most recently forwarded, plus the summation of signal propagation time over all episodes of transmission between forwarding nodes. Determination of these values is an implementation matter.) If the bundle's creation time is zero, then the bundle MUST contain exactly one (1) occurrence of this type of block; otherwise, the bundle MAY contain at most one (1) occurrence of this type of block. A bundle MUST NOT contain multiple occurrences of the bundle age block, as this could result in processing anomalies.

4.3.3. Hop Count

The Hop Count block, block type 9, contains two unsigned integers, hop limit and hop count. A "hop" is here defined as an occasion on which a bundle was forwarded from one node to another node. Hop limit MUST be in the range 1 through 255. The hop limit value SHOULD NOT be changed at any time after creation of the Hop Count block; the hop count value SHOULD initially be zero and SHOULD be increased by 1 on each hop.

The hop count block is mainly intended as a safety mechanism, a means of identifying bundles for removal from the network that can never be delivered due to a persistent forwarding error. When a bundle's hop count exceeds its hop limit, the bundle SHOULD be deleted for the reason "hop limit exceeded", following the bundle deletion procedure defined in Section 5.10. . Procedures for determining the appropriate hop limit for a block are beyond the scope of this specification. The block-type-specific data in a hop count block SHALL be represented as a CBOR array comprising a 2-tuple. The first item of this array SHALL be the bundle's hop limit, represented as a CBOR unsigned integer. The second item of this array SHALL be the bundle's hop count, represented as a CBOR unsigned integer. A bundle MAY contain at most one (1) occurrence of this type of block.

5. Bundle Processing

The bundle processing procedures mandated in this section and in Section 6 govern the operation of the Bundle Protocol Agent and the Application Agent administrative element of each bundle node. They are neither exhaustive nor exclusive. Supplementary DTN protocol specifications (including, but not restricted to, the Bundle Security Protocol [BPSEC]) may augment, override, or supersede the mandates of this document.

5.1. Generation of Administrative Records

All transmission of bundles is in response to bundle transmission requests presented by nodes' application agents. When required to "generate" an administrative record (such as a bundle status report), the bundle protocol agent itself is responsible for causing a new bundle to be transmitted, conveying that record. In concept, the bundle protocol agent discharges this responsibility by directing the administrative element of the node's application agent to construct the record and request its transmission as detailed in Section 6 below. In practice, the manner in which administrative record generation is accomplished is an implementation matter, provided the constraints noted in Section 6 are observed.

Note that requesting status reports for any single bundle might easily result in the generation of $(1 + (2 * (N-1)))$ status report bundles, where N is the number of nodes on the path from the bundle's source to its destination, inclusive. That is, the requesting of status reports for large numbers of bundles could result in an unacceptable increase in the bundle traffic in the network. For this reason, the generation of status reports MUST be disabled by default and enabled only when the risk of excessive network traffic is deemed acceptable.

When the generation of status reports is enabled, the decision on whether or not to generate a requested status report is left to the discretion of the bundle protocol agent. Mechanisms that could assist in making such decisions, such as pre-placed agreements authorizing the generation of status reports under specified circumstances, are beyond the scope of this specification.

Notes on administrative record terminology:

- . A "bundle reception status report" is a bundle status report with the "reporting node received bundle" flag set to 1.
- . A "bundle forwarding status report" is a bundle status report with the "reporting node forwarded the bundle" flag set to 1.

- . A "bundle delivery status report" is a bundle status report with the "reporting node delivered the bundle" flag set to 1.
- . A "bundle deletion status report" is a bundle status report with the "reporting node deleted the bundle" flag set to 1.

5.2. Bundle Transmission

The steps in processing a bundle transmission request are:

Step 1: Transmission of the bundle is initiated. An outbound bundle MUST be created per the parameters of the bundle transmission request, with the retention constraint "Dispatch pending". The source node ID of the bundle MUST be either the null endpoint ID, indicating that the source of the bundle is anonymous, or else the EID of a singleton endpoint whose only member is the node of which the BPA is a component.

Step 2: Processing proceeds from Step 1 of Section 5.4.

5.3. Bundle Dispatching

The steps in dispatching a bundle are:

Step 1: If the bundle's destination endpoint is an endpoint of which the node is a member, the bundle delivery procedure defined in Section 5.7 MUST be followed and for the purposes of all subsequent processing of this bundle at this node the node's membership in the bundle's destination endpoint SHALL be disavowed; specifically, even though the node is a member of the bundle's destination endpoint, the node SHALL NOT undertake to forward the bundle to itself in the course of performing the procedure described in Section 5.4.

Step 2: Processing proceeds from Step 1 of Section 5.4.

5.4. Bundle Forwarding

The steps in forwarding a bundle are:

Step 1: The retention constraint "Forward pending" MUST be added to the bundle, and the bundle's "Dispatch pending" retention constraint MUST be removed.

Step 2: The bundle protocol agent MUST determine whether or not forwarding is contraindicated for any of the reasons listed in Figure 4. In particular:

- . The bundle protocol agent MAY choose either to forward the bundle directly to its destination node(s) (if possible) or to forward the bundle to some other node(s) for further forwarding. The manner in which this decision is made may depend on the scheme name in the destination endpoint ID and/or on other state but in any case is beyond the scope of this document. If the BPA elects to forward the bundle to some other node(s) for further forwarding but finds it impossible to select any node(s) to forward the bundle to, then forwarding is contraindicated.
- . Provided the bundle protocol agent succeeded in selecting the node(s) to forward the bundle to, the bundle protocol agent MUST select the convergence layer adapter(s) whose services will enable the node to send the bundle to those nodes. The manner in which specific appropriate convergence layer adapters are selected is beyond the scope of this document. If the agent finds it impossible to select any appropriate convergence layer adapter(s) to use in forwarding this bundle, then forwarding is contraindicated.

Step 3: If forwarding of the bundle is determined to be contraindicated for any of the reasons listed in Figure 4, then the Forwarding Contraindicated procedure defined in Section 5.4.1 MUST be followed; the remaining steps of Section 5.4 are skipped at this time.

Step 4: For each node selected for forwarding, the bundle protocol agent MUST invoke the services of the selected convergence layer adapter(s) in order to effect the sending of the bundle to that node. Determining the time at which the bundle protocol agent invokes convergence layer adapter services is a BPA implementation matter. Determining the time at which each convergence layer adapter subsequently responds to this service invocation by sending the bundle is a convergence-layer adapter implementation matter. Note that:

- . If the bundle contains a data label extension block (to be defined in a future document) then that data label value MAY identify procedures for determining the order in which convergence layer adapters must send bundles, e.g., considering bundle source when determining the order in which bundles are sent. The definition of such procedures is beyond the scope of this specification.
- . If the bundle has a bundle age block, as defined in 4.3.2. above, then at the last possible moment before the CLA initiates conveyance of the bundle node via the CL protocol the bundle age value MUST be increased by the difference between

the current time and the time at which the bundle was received (or, if the local node is the source of the bundle, created).

Step 5: When all selected convergence layer adapters have informed the bundle protocol agent that they have concluded their data sending procedures with regard to this bundle:

- . If the "request reporting of bundle forwarding" flag in the bundle's status report request field is set to 1, and status reporting is enabled, then a bundle forwarding status report SHOULD be generated, destined for the bundle's report-to endpoint ID. The reason code on this bundle forwarding status report MUST be "no additional information".
- . If any applicable bundle protocol extensions mandate generation of status reports upon conclusion of convergence-layer data sending procedures, all such status reports SHOULD be generated with extension-mandated reason codes.
- . The bundle's "Forward pending" retention constraint MUST be removed.

5.4.1. Forwarding Contraindicated

The steps in responding to contraindication of forwarding are:

Step 1: The bundle protocol agent MUST determine whether or not to declare failure in forwarding the bundle. Note: this decision is likely to be influenced by the reason for which forwarding is contraindicated.

Step 2: If forwarding failure is declared, then the Forwarding Failed procedure defined in Section 5.4.2 MUST be followed.

Otherwise, when -- at some future time - the forwarding of this bundle ceases to be contraindicated, processing proceeds from Step 4 of Section 5.4.

5.4.2. Forwarding Failed

The steps in responding to a declaration of forwarding failure are:

Step 1: The bundle protocol agent MAY forward the bundle back to the node that sent it, as identified by the Previous Node block, if present. This forwarding, if performed, SHALL be accomplished by performing Step 4 and Step 5 of section 5.4 where the sole node selected for forwarding SHALL be the node that sent the bundle.

Step 2: If the bundle's destination endpoint is an endpoint of which the node is a member, then the bundle's "Forward pending" retention constraint MUST be removed. Otherwise, the bundle MUST be deleted: the bundle deletion procedure defined in Section 5.10 MUST be followed, citing the reason for which forwarding was determined to be contraindicated.

5.5. Bundle Expiration

A bundle expires when the bundle's age exceeds its lifetime as specified in the primary bundle block. Bundle age MAY be determined by subtracting the bundle's creation timestamp time from the current time if (a) that timestamp time is not zero and (b) the local node's clock is known to be accurate; otherwise bundle age MUST be obtained from the Bundle Age extension block. Bundle expiration MAY occur at any point in the processing of a bundle. When a bundle expires, the bundle protocol agent MUST delete the bundle for the reason "lifetime expired": the bundle deletion procedure defined in Section 5.10 MUST be followed.

5.6. Bundle Reception

The steps in processing a bundle that has been received from another node are:

Step 1: The retention constraint "Dispatch pending" MUST be added to the bundle.

Step 2: If the "request reporting of bundle reception" flag in the bundle's status report request field is set to 1, and status reporting is enabled, then a bundle reception status report with reason code "No additional information" SHOULD be generated, destined for the bundle's report-to endpoint ID.

Step 3: If any block of the bundle is malformed according to this specification, or if any block has an attached CRC and the CRC computed for this block upon reception differs from that attached CRC, then the bundle protocol agent MUST delete the bundle for the reason "Block unintelligible". The bundle deletion procedure defined in Section 5.10 MUST be followed and all remaining steps of the bundle reception procedure MUST be skipped.

Step 4: For each block in the bundle that is an extension block that the bundle protocol agent cannot process:

- . If the block processing flags in that block indicate that a status report is requested in this event, and status reporting

is enabled, then a bundle reception status report with reason code "Block unintelligible" SHOULD be generated, destined for the bundle's report-to endpoint ID.

- . If the block processing flags in that block indicate that the bundle must be deleted in this event, then the bundle protocol agent MUST delete the bundle for the reason "Block unintelligible"; the bundle deletion procedure defined in Section 5.10 MUST be followed and all remaining steps of the bundle reception procedure MUST be skipped.
- . If the block processing flags in that block do NOT indicate that the bundle must be deleted in this event but do indicate that the block must be discarded, then the bundle protocol agent MUST remove this block from the bundle.
- . If the block processing flags in that block indicate neither that the bundle must be deleted nor that that the block must be discarded, then processing continues with the next extension block that the bundle protocol agent cannot process, if any; otherwise, processing proceeds from step 5.

Step 5: Processing proceeds from Step 1 of Section 5.3.

5.7. Local Bundle Delivery

The steps in processing a bundle that is destined for an endpoint of which this node is a member are:

Step 1: If the received bundle is a fragment, the application data unit reassembly procedure described in Section 5.9 MUST be followed. If this procedure results in reassembly of the entire original application data unit, processing of this bundle (whose fragmentary payload has been replaced by the reassembled application data unit) proceeds from Step 2; otherwise, the retention constraint "Reassembly pending" MUST be added to the bundle and all remaining steps of this procedure MUST be skipped.

Step 2: Delivery depends on the state of the registration whose endpoint ID matches that of the destination of the bundle:

- . An additional implementation-specific delivery deferral procedure MAY optionally be associated with the registration.
- . If the registration is in the Active state, then the bundle MUST be delivered automatically as soon as it is the next bundle that is due for delivery according to the BPA's bundle delivery scheduling policy, an implementation matter.
- . If the registration is in the Passive state, or if delivery of the bundle fails for some implementation-specific reason, then

the registration's delivery failure action MUST be taken.
Delivery failure action MUST be one of the following:

- o defer delivery of the bundle subject to this registration until (a) this bundle is the least recently received of all bundles currently deliverable subject to this registration and (b) either the registration is polled or else the registration is in the Active state, and also perform any additional delivery deferral procedure associated with the registration; or
- o abandon delivery of the bundle subject to this registration (as defined in 3.1.).

Step 3: As soon as the bundle has been delivered, if the "request reporting of bundle delivery" flag in the bundle's status report request field is set to 1 and bundle status reporting is enabled, then a bundle delivery status report SHOULD be generated, destined for the bundle's report-to endpoint ID. Note that this status report only states that the payload has been delivered to the application agent, not that the application agent has processed that payload.

5.8. Bundle Fragmentation

It may at times be advantageous for bundle protocol agents to reduce the sizes of bundles in order to forward them. This might be the case, for example, if a node to which a bundle is to be forwarded is accessible only via intermittent contacts and no upcoming contact is long enough to enable the forwarding of the entire bundle.

The size of a bundle can be reduced by "fragmenting" the bundle. To fragment a bundle whose payload is of size M is to replace it with two "fragments" -- new bundles with the same source node ID and creation timestamp as the original bundle -- whose payloads are the first N and the last $(M - N)$ bytes of the original bundle's payload, where $0 < N < M$. Note that fragments may themselves be fragmented, so fragmentation may in effect replace the original bundle with more than two fragments. (However, there is only one 'level' of fragmentation, as in IP fragmentation.)

Any bundle whose primary block's bundle processing flags do NOT indicate that it must not be fragmented MAY be fragmented at any time, for any purpose, at the discretion of the bundle protocol agent. NOTE, however, that some combinations of bundle fragmentation, replication, and routing might result in unexpected traffic patterns.

Fragmentation SHALL be constrained as follows:

- . The concatenation of the payloads of all fragments produced by fragmentation MUST always be identical to the payload of the fragmented bundle (that is, the bundle that is being fragmented). Note that the payloads of fragments resulting from different fragmentation episodes, in different parts of the network, may be overlapping subsets of the fragmented bundle's payload.
- . The primary block of each fragment MUST differ from that of the fragmented bundle, in that the bundle processing flags of the fragment MUST indicate that the bundle is a fragment and both fragment offset and total application data unit length must be provided. Additionally, the CRC of the primary block of the fragmented bundle, if any, MUST be replaced in each fragment by a new CRC computed for the primary block of that fragment.
- . The payload blocks of fragments will differ from that of the fragmented bundle as noted above.
- . If the fragmented bundle is not a fragment or is the fragment with offset zero, then all extension blocks of the fragmented bundle MUST be replicated in the fragment whose offset is zero.
- . Each of the fragmented bundle's extension blocks whose "Block must be replicated in every fragment" flag is set to 1 MUST be replicated in every fragment.
- . Beyond these rules, replication of extension blocks in the fragments is an implementation matter.

5.9. Application Data Unit Reassembly

If the concatenation -- as informed by fragment offsets and payload lengths -- of the payloads of all previously received fragments with the same source node ID and creation timestamp as this fragment, together with the payload of this fragment, forms a byte array whose length is equal to the total application data unit length in the fragment's primary block, then:

- . This byte array -- the reassembled application data unit -- MUST replace the payload of this fragment.
- . The "Reassembly pending" retention constraint MUST be removed from every other fragment whose payload is a subset of the reassembled application data unit.

Note: reassembly of application data units from fragments occurs at the nodes that are members of destination endpoints as necessary; an application data unit MAY also be reassembled at some other node on the path to the destination.

5.10. Bundle Deletion

The steps in deleting a bundle are:

Step 1: If the "request reporting of bundle deletion" flag in the bundle's status report request field is set to 1, and if status reporting is enabled, then a bundle deletion status report citing the reason for deletion SHOULD be generated, destined for the bundle's report-to endpoint ID.

Step 2: All of the bundle's retention constraints MUST be removed.

5.11. Discarding a Bundle

As soon as a bundle has no remaining retention constraints it MAY be discarded, thereby releasing any persistent storage that may have been allocated to it.

5.12. Canceling a Transmission

When requested to cancel a specified transmission, where the bundle created upon initiation of the indicated transmission has not yet been discarded, the bundle protocol agent MUST delete that bundle for the reason "transmission cancelled". For this purpose, the procedure defined in Section 5.10 MUST be followed.

6. Administrative Record Processing

6.1. Administrative Records

Administrative records are standard application data units that are used in providing some of the features of the Bundle Protocol. One type of administrative record has been defined to date: bundle status reports. Note that additional types of administrative records may be defined by supplementary DTN protocol specification documents.

Every administrative record consists of:

- . Record type code (an unsigned integer for which valid values are as defined below).
- . Record content in type-specific format.

Valid administrative record type codes are defined as follows:

Value	Meaning
1	Bundle status report.
(other)	Reserved for future use.

Figure 3: Administrative Record Type Codes

Each BP administrative record SHALL be represented as a CBOR array comprising a 2-tuple.

The first item of the array SHALL be a record type code, which SHALL be represented as a CBOR unsigned integer.

The second element of this array SHALL be the applicable CBOR representation of the content of the record. Details of the CBOR representation of administrative record type 1 are provided below. Details of the CBOR representation of other types of administrative record type are included in the specifications defining those records.

6.1.1. Bundle Status Reports

The transmission of "bundle status reports" under specified conditions is an option that can be invoked when transmission of a bundle is requested. These reports are intended to provide information about how bundles are progressing through the system, including notices of receipt, forwarding, final delivery, and deletion. They are transmitted to the Report-to endpoints of bundles.

Each bundle status report SHALL be represented as a CBOR array. The number of elements in the array SHALL be either 6 (if the subject bundle is a fragment) or 4 (otherwise).

The first item of the bundle status report array SHALL be bundle status information represented as a CBOR array of at least 4

elements. The first four items of the bundle status information array shall provide information on the following four status assertions, in this order:

- . Reporting node received bundle.
- . Reporting node forwarded the bundle.
- . Reporting node delivered the bundle.
- . Reporting node deleted the bundle.

Each item of the bundle status information array SHALL be a bundle status item represented as a CBOR array; the number of elements in each such array SHALL be either 2 (if the value of the first item of this bundle status item is 1 AND the "Report status time" flag was set to 1 in the bundle processing flags of the bundle whose status is being reported) or 1 (otherwise). The first item of the bundle status item array SHALL be a status indicator, a Boolean value indicating whether or not the corresponding bundle status is asserted, represented as a CBOR Boolean value. The second item of the bundle status item array, if present, SHALL indicate the time (as reported by the local system clock, an implementation matter) at which the indicated status was asserted for this bundle, represented as a DTN time as described in Section 4.1.6. above.

The second item of the bundle status report array SHALL be the bundle status report reason code explaining the value of the status indicator, represented as a CBOR unsigned integer. Valid status report reason codes are defined in Figure 4 below but the list of status report reason codes provided here is neither exhaustive nor exclusive; supplementary DTN protocol specifications (including, but not restricted to, the Bundle Security Protocol [BPSEC]) may define additional reason codes.

Value	Meaning
0	No additional information.
1	Lifetime expired.
2	Forwarded over unidirectional link.

3	Transmission canceled.
4	Depleted storage.
5	Destination endpoint ID unintelligible.
6	No known route to destination from here.
7	No timely contact with next node on route.
8	Block unintelligible.
9	Hop limit exceeded.
10	Traffic pared (e.g., status reports).
(other)	Reserved for future use.

Figure 4: Status Report Reason Codes

The third item of the bundle status report array SHALL be the source node ID identifying the source of the bundle whose status is being reported, represented as described in Section 4.1.5.2. above.

The fourth item of the bundle status report array SHALL be the creation timestamp of the bundle whose status is being reported, represented as described in Section 4.1.7. above.

The fifth item of the bundle status report array SHALL be present if and only if the bundle whose status is being reported contained a fragment offset. If present, it SHALL be the subject bundle's fragment offset represented as a CBOR unsigned integer item.

The sixth item of the bundle status report array SHALL be present if and only if the bundle whose status is being reported contained a fragment offset. If present, it SHALL be the length of the subject bundle's payload represented as a CBOR unsigned integer item.

6.2. Generation of Administrative Records

Whenever the application agent's administrative element is directed by the bundle protocol agent to generate an administrative record with reference to some bundle, the following procedure must be followed:

Step 1: The administrative record must be constructed. If the administrative record references a bundle and the referenced bundle is a fragment, the administrative record MUST contain the fragment offset and fragment length.

Step 2: A request for transmission of a bundle whose payload is this administrative record MUST be presented to the bundle protocol agent.

7. Services Required of the Convergence Layer

7.1. The Convergence Layer

The successful operation of the end-to-end bundle protocol depends on the operation of underlying protocols at what is termed the "convergence layer"; these protocols accomplish communication between nodes. A wide variety of protocols may serve this purpose, so long as each convergence layer protocol adapter provides a defined minimal set of services to the bundle protocol agent. This convergence layer service specification enumerates those services.

7.2. Summary of Convergence Layer Services

Each convergence layer protocol adapter is expected to provide the following services to the bundle protocol agent:

- . sending a bundle to a bundle node that is reachable via the convergence layer protocol;
- . notifying the bundle protocol agent when it has concluded its data sending procedures with regard to a bundle;

- . delivering to the bundle protocol agent a bundle that was sent by a bundle node via the convergence layer protocol.

The convergence layer service interface specified here is neither exhaustive nor exclusive. That is, supplementary DTN protocol specifications (including, but not restricted to, the Bundle Security Protocol [BPSEC]) may expect convergence layer adapters that serve BP implementations conforming to those protocols to provide additional services such as reporting on the transmission and/or reception progress of individual bundles (at completion and/or incrementally), retransmitting data that were lost in transit, discarding bundle-conveying data units that the convergence layer protocol determines are corrupt or inauthentic, or reporting on the integrity and/or authenticity of delivered bundles.

8. Implementation Status

[NOTE to the RFC Editor: please remove this section before publication, as well as the reference to RFC 7942.]

This section records the status of known implementations of the protocol defined by this specification at the time of posting of this Internet-Draft, and is based on a proposal described in RFC 7942. The description of implementations in this section is intended to assist the IETF in its decision processes in progressing drafts to RFCs. Please note that the listing of any individual implementation here does not imply endorsement by the IETF. Furthermore, no effort has been spent to verify the information presented here that was supplied by IETF contributors. This is not intended as, and must not be construed to be, a catalog of available implementations or their features. Readers are advised to note that other implementations may exist.

According to RFC 7942, "this will allow reviewers and working groups to assign due consideration to documents that have the benefit of running code, which may serve as evidence of valuable experimentation and feedback that have made the implemented protocols more mature. It is up to the individual working groups to use this information as they see fit".

At the time of this writing, there are three known implementations of the current document.

The first known implementation is microPCN (<https://upcn.eu/>). According to the developers:

The Micro Planetary Communication Network (uPCN) is a free software project intended to offer an implementation of Delay-tolerant Networking protocols for POSIX operating systems (well, and for Linux) plus for the ARM Cortex STM32F4 microcontroller series. More precisely it currently provides an implementation of

- . the Bundle Protocol (BP, RFC 5050),
- . the Bundle Protocol version 7 specification draft (version 6),
- . the DTN IP Neighbor Discovery (IPND) protocol, and
- . a routing approach optimized for message-ferry micro LEO satellites.

uPCN is written in C and is built upon the real-time operating system FreeRTOS. The source code of uPCN is released under the "BSD 3-Clause License".

The project depends on an execution environment offering link layer protocols such as AX.25. The source code uses the USB subsystem to interact with the environment.

The second known implementation is PyDTN, developed by X-works, s.r.o (<https://x-works.sk/>). The final third of the implementation was developed during the IETF 101 Hackathon. According to the developers, PyDTN implements bundle coding/decoding and neighbor discovery. PyDTN is written in Python and has been shown to be interoperable with uPCN.

The third known implementation is "Terra" (<https://github.com/RightMesh/Terra/>), a Java implementation developed in the context of terrestrial DTN. It includes an implementation of a "minimal TCP" convergence layer adapter.

9. Security Considerations

The bundle protocol security architecture and the available security services are specified in an accompanying document, the Bundle Security Protocol specification [BPSEC].

The bpsec extensions to Bundle Protocol enable each block of a bundle (other than a bpsec extension block) to be individually authenticated by a signature block (Block Integrity Block, or BIB) and also enable each block of a bundle other than the primary block (and the bpsec extension blocks themselves) to be individually encrypted by a BCB.

Because the security mechanisms are extension blocks that are themselves inserted into the bundle, the integrity and

confidentiality of bundle blocks are protected while the bundle is at rest, awaiting transmission at the next forwarding opportunity, as well as in transit.

Additionally, convergence-layer protocols that ensure authenticity of communication between adjacent nodes in BP network topology SHOULD be used where available, to minimize the ability of unauthenticated nodes to introduce inauthentic traffic into the network. Convergence-layer protocols that ensure confidentiality of communication between adjacent nodes in BP network topology SHOULD also be used where available, to minimize exposure of the bundle's primary block and other clear-text blocks, thereby offering some defense against traffic analysis.

Note that, while the primary block must remain in the clear for routing purposes, the Bundle Protocol can be protected against traffic analysis to some extent by using bundle-in-bundle encapsulation to tunnel bundles to a safe forward distribution point: the encapsulated bundle forms the payload of an encapsulating bundle, and that payload block may be encrypted by a BCB.

Note that the generation of bundle status reports is disabled by default because malicious initiation of bundle status reporting could result in the transmission of extremely large numbers of bundles, effecting a denial of service attack.

The bpsec extensions accommodate an open-ended range of ciphersuites; different ciphersuites may be utilized to protect different blocks. One possible variation is to sign and/or encrypt blocks using symmetric keys securely formed by Diffie-Hellman procedures (such as EKDH) using the public and private keys of the sending and receiving nodes. For this purpose, the key distribution problem reduces to the problem of trustworthy delay-tolerant distribution of public keys, a current research topic.

Bundle security MUST NOT be invalidated by forwarding nodes even though they themselves might not use the Bundle Security Protocol.

In particular, while blocks MAY be added to bundles transiting intermediate nodes, removal of blocks with the "Discard block if it can't be processed" flag set in the block processing control flags may cause security to fail.

Inclusion of the Bundle Security Protocol in any Bundle Protocol implementation is RECOMMENDED. Use of the Bundle Security Protocol in Bundle Protocol operations is OPTIONAL, subject to the following guidelines:

- . Every block (that is not a bpsec extension block) of every bundle SHOULD be authenticated by a BIB citing the ID of the node that inserted that block. (Note that a single BIB may authenticate multiple "target" blocks.) BIB authentication MAY be omitted on (and only on) any initial end-to-end path segments on which it would impose unacceptable overhead, provided that satisfactory authentication is ensured at the convergence layer and that BIB authentication is asserted on the first path segment on which the resulting overhead is acceptable and on all subsequent path segments.
- . If any segment of the end-to-end path of a bundle will traverse the Internet or any other potentially insecure communication environment, then the payload block SHOULD be encrypted by a BCB on this path segment and all subsequent segments of the end-to-end path.

10. IANA Considerations

The Bundle Protocol includes fields requiring registries managed by IANA.

10.1. Bundle Block Types

The Bundle Protocol has a Bundle Block Type code field (Section 4.2.3). An IANA registry has been set up as follows.

The registration policy for this registry is:

- 0-191: Specification Required.
- 192-255: Private or experimental use. No assignment by IANA.

The Value range is: unsigned 8-bit integer.

Bundle Block Type Registry

Value	Description	Reference
0	Reserved	This document
1	Bundle Payload Block	section 4.2.3
2	Block Integrity Block	[BPSEC]

3	Block Confidentiality Block	[BPSEC]
4-6	Reserved	section 4.3
7	Previous node (proximate sender)	section 4.3.1
8	Bundle age (in seconds)	section 4.3.2
9	Hop count (#prior xmit attempts)	section 4.3.3
10-191	Unassigned	
192-255	Private and/or Experimental Use	This document

IANA is requested to add values 2-9, as noted above, to the existing registry.

The value "0" was not defined in any document or in the ad hoc registry. As per consensus by the DTNRG research group, it is reserved per this document.

10.2. Primary Bundle Protocol Version

The Bundle Protocol has a version field (Section 4.2.2). An IANA registry has been set up as follows.

The registration policy for this registry is: RFC Required.

The value range is: unsigned 8-bit integer.

Primary Bundle Protocol Version Registry

Value	Description	Reference
0-5	Reserved	This document
6	Assigned	[RFC5050]
7	Assigned	section 4.2.2
8-255	Unassigned	

+-----+-----+-----+

The value "0-5" was not defined in any document or in the ad hoc registry. As per consensus by the DTNRG research group, it is reserved per this document.

10.3. Bundle Processing Control Flags

The Bundle Protocol has a Bundle Processing Control Flags field (Section 4.1.3) for which IANA is requested to create and maintain a new registry named "BPv7 Bundle Processing Control Flags". Initial values for this registry are given below.

The registration policy for this registry is: Specification Required.

The value range is: variable length. Maximum number of flag bit positions: 16.

Bundle Processing Control Flags Registry

Bit Position (right to left)	Description	Reference
0	Bundle is a fragment	4.1.3
1	Application data unit is an administrative record	4.1.3
2	Bundle must not be fragmented	4.1.3
3	reserved	4.1.3
4	reserved	4.1.3
5	Acknowledgement by application is requested	4.1.3
6	Status time requested in reports	4.1.3

	7	Reserved	4.1.3	
	8	Request reporting of bundle	4.1.3	
		reception		
	9	Reserved	4.1.3	
	10	Request reporting of bundle	4.1.3	
		forwarding		
	11	Request reporting of bundle	4.1.3	
		delivery		
	12	Request reporting of bundle	4.1.3	
		deletion		
	13-15	Unassigned		

10.4. Block Processing Control Flags

The Bundle Protocol has a Block Processing Control Flags field (Section 4.1.4) for which IANA is requested to create and maintain a new registry named "BPv7 Block Processing Control Flags". Initial values for this registry are given below.

The registration policy for this registry is: Specification Required.

The value range is: variable length. Maximum number of flag bit positions: 8.

Block Processing Control Flags Registry

Bit Position	Description	Reference
(right to left)		

0	Block must be replicated in every fragment	4.1.4
1	Discard block if it can't be processed	4.1.4
2	Transmit status report if block can't be processed	4.1.4
3	Delete bundle if block can't be processed	4.1.4
4-7	Reserved	

10.5. Bundle Status Report Reason Codes

The Bundle Protocol has a Bundle Status Report Reason Codes field (Section 6.1.1) for which IANA is requested to create and maintain a new registry named "BPv7 Bundle Status Report Reason Codes". Initial values for this registry are given below.

The registration policy for this registry is: Specification Required.

The value range is: unsigned 8-bit integer.

Bundle Status Report Reason Codes Registry

Value	Description	Reference
0	No additional information	6.1.1
1	Lifetime expired	6.1.1
2	Forwarded over unidirectional link	6.1.1
3	Transmission canceled	6.1.1

4	Depleted storage	6.1.1
5	Destination endpoint ID unintelligible	6.1.1
6	No known route to destination from here	6.1.1
7	No timely contact with next node on route	6.1.1
8	Block unintelligible	6.1.1
9	Hop limit exceeded	6.1.1
8	Traffic pared	6.1.1
9-254	Unassigned	
255	Reserved	

10.6. URI scheme types

The Bundle Protocol has a URI scheme type field - an unsigned integer of undefined length - for which IANA is requested to create and maintain a new registry named "URI scheme type values". Initial values for the Bundle Protocol URI scheme type registry are given below.

The registration policy for this registry is: RFC Required.

The value range is: unsigned 8-bit integer.

Each assignment consists of a URI scheme type name and its associated value.

Bundle Block Type Registry

Value	Description	Reference
0	Reserved	
1	dtn	section 10.7

2	ipn	RFC6260, Section 4
3-254	Unassigned	
255	reserved	
+-----+-----+-----+		

10.7. New URI scheme "dtn"

IANA is requested to register a URI scheme with the string "dtn" as the scheme name, as follows:

URI scheme name: "dtn"

Status: provisional

URI scheme syntax:

This specification uses the Augmented Backus-Naur Form (ABNF) notation of [RFC5234].

dtn-uri = "dtn:" dtn-hier-part

dtn-hier-part = "//" node-name name-delim demux ; a path-rootless

node-name = 1*VCHAR

name-delim = "/"

demux = *VCHAR

None of the reserved characters defined in the generic URI syntax are used as delimiters within URIs of the DTN scheme.

URI scheme semantics: URIs of the DTN scheme are used as endpoint identifiers in the Delay-Tolerant Networking (DTN) Bundle Protocol (BP) as described in Section 4.1.5.1.

Encoding considerations: URIs of the DTN scheme are encoded exclusively in US-ASCII characters.

Applications and/or protocols that use this URI scheme name: the Delay-Tolerant Networking (DTN) Bundle Protocol (BP).

Interoperability considerations: as noted above, URIs of the DTN scheme are encoded exclusively in US-ASCII characters.

Security considerations:

- . Reliability and consistency: none of the BP endpoints identified by the URIs of the DTN scheme are guaranteed to be reachable at any time, and the identity of the processing entities operating on those endpoints is never guaranteed by the Bundle Protocol itself. Bundle authentication as defined by the Bundle Security Protocol is required for this purpose.
- . Malicious construction: malicious construction of a conformant DTN-scheme URI is limited to the malicious selection of node names and the malicious selection of demux strings. That is, a maliciously constructed DTN-scheme URI could be used to direct a bundle to an endpoint that might be damaged by the arrival of that bundle or, alternatively, to declare a false source for a bundle and thereby cause incorrect processing at a node that receives the bundle. In both cases (and indeed in all bundle processing), the node that receives a bundle should verify its authenticity and validity before operating on it in any way.
- . Back-end transcoding: the limited expressiveness of URIs of the DTN scheme effectively eliminates the possibility of threat due to errors in back-end transcoding.
- . Rare IP address formats: not relevant, as IP addresses do not appear anywhere in conformant DTN-scheme URIs.
- . Sensitive information: because DTN-scheme URIs are used only to represent the identities of Bundle Protocol endpoints, the risk of disclosure of sensitive information due to interception of these URIs is minimal. Examination of DTN-scheme URIs could be used to support traffic analysis; where traffic analysis is a plausible danger, bundles should be conveyed by secure convergence-layer protocols that do not expose endpoint IDs.
- . Semantic attacks: the simplicity of DTN-scheme URI syntax minimizes the possibility of misinterpretation of a URI by a human user.

Contact:

Scott Burleigh
Jet Propulsion Laboratory,
California Institute of Technology
scott.c.burleigh@jpl.nasa.gov

+1 (800) 393-3353

Author/Change controller:

Scott Burleigh

Jet Propulsion Laboratory,

California Institute of Technology

scott.c.burleigh@jpl.nasa.gov

11. References

11.1. Normative References

[BPSEC] Birrane, E., "Bundle Security Protocol Specification", Work In Progress, October 2015.

[CRC16] ITU-T Recommendation X.25, p. 9, section 2.2.7.4, International Telecommunications Union, October 1996.

[CRC32C] Castagnoli, G., Brauer, S., and M. Herrmann, "Optimization of Cyclic Redundancy-Check Codes with 24 and 32 Parity Bits", IEEE Transact. on Communications, Vol. 41, No. 6, June 1993.

[EPOCH] Thompson, K. and D. M. Ritchie, "UNIX Programmer's Manual, Fifth Edition", Bell Telephone Laboratories Inc., June 1974.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008.

[RFC7049] Borman, C. and P. Hoffman, "Concise Binary Object Representation (CBOR)", RFC 7049, October 2013.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, May 2017.

[URI] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", RFC 3986, STD 66, January 2005.

[URIREG] Thaler, D., Hansen, T., and T. Hardie, "Guidelines and Registration Procedures for URI Schemes", RFC 7595, BCP 35, June 2015.

11.2. Informative References

[ARCH] V. Cerf et al., "Delay-Tolerant Network Architecture", RFC 4838, April 2007.

[BIBE] Burleigh, S., "Bundle-in-Bundle Encapsulation", Work In Progress, June 2017.

[RFC3987] Duerst, M. and M. Suignard, "Internationalized Resource Identifiers (IRIs)", RFC 3987, January 2005.

[RFC5050] Scott, K. and S. Burleigh, "Bundle Protocol Specification", RFC 5050, November 2007.

[RFC7143] Chadalapaka, M., Satran, J., Meth, K., and D. Black, "Internet Small Computer System Interface (iSCSI) Protocol (Consolidated)", RFC 7143, April 2014.

[SIGC] Fall, K., "A Delay-Tolerant Network Architecture for Challenged Internets", SIGCOMM 2003.

[UTC] Arias, E. and B. Guinot, "Coordinated universal time UTC: historical background and perspectives" in "Journées systèmes de référence spatio-temporels", 2004.

12. Acknowledgments

This work is freely adapted from RFC 5050, which was an effort of the Delay Tolerant Networking Research Group. The following DTNRG participants contributed significant technical material and/or inputs to that document: Dr. Vinton Cerf of Google, Scott Burleigh, Adrian Hooke, and Leigh Torgerson of the Jet Propulsion Laboratory, Michael Demmer of the University of California at Berkeley, Robert Durst, Keith Scott, and Susan Symington of The MITRE Corporation, Kevin Fall of Carnegie Mellon University, Stephen Farrell of Trinity College Dublin, Howard Weiss and Peter Lovell of SPARTA, Inc., and Manikantan Ramadas of Ohio University.

This document was prepared using 2-Word-v2.0.template.dot.

13. Significant Changes from RFC 5050

Points on which this draft significantly differs from RFC 5050 include the following:

- . Clarify the difference between transmission and forwarding.
- . Migrate custody transfer to the bundle-in-bundle encapsulation specification [BIBE].
- . Introduce the concept of "node ID" as functionally distinct from endpoint ID, while having the same syntax.
- . Restructure primary block, making it immutable. Add optional CRC.
- . Add optional CRCs to non-primary blocks.
- . Add block ID number to canonical block format (to support BPSEC).
- . Add definition of bundle age extension block.
- . Add definition of previous node extension block.
- . Add definition of hop count extension block.
- . Remove Quality of Service markings.
- . Change from SDNVs to CBOR representation.

Appendix A.

For More Information

Please refer comments to dtm@ietf.org. DTN Working Group documents are located at <https://datatracker.ietf.org/wg/dtn/documents>. The original Delay Tolerant Networking Research Group (DTNRG) Web site is located at <https://irtf.org/concluded/dtnrg>.

Copyright (c) 2019 IETF Trust and the persons identified as authors of the code. All rights reserved.

Redistribution and use in source and binary forms, with or without modification, is permitted pursuant to, and subject to the license terms contained in, the Simplified BSD License set forth in Section 4.c of the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>).

Appendix B.

CDDL expression

For informational purposes, Carsten Bormann and Brian Sipos have kindly provided an expression of the Bundle Protocol specification in the Concise Data Definition Language (CDDL). That CDDL expression is presented below. Note that wherever the CDDL expression is in disagreement with the textual representation of the BP specification presented in the earlier sections of this document, the textual representation rules.

```
start = bundle / #6.55799(bundle)

; Times before 2000 are invalid

dtn-time = uint

; CRC enumerated type

crc-type = &(
    crc-none: 0,
    crc-16bit: 1,
    crc-32bit: 2
)

; Either 16-bit or 32-bit

crc-value = (bstr .size 2) / (bstr .size 4)

creation-timestamp = [
    dtn-time, ; absolute time of creation
    sequence: uint ; sequence within the time
]

eid = $eid .within eid-structure

eid-structure = [
    uri-code: uint,
```

```
    SSP: any
  ]
  $eid /= [
    uri-code: 1,
    SSP: (tstr / 0)
  ]
  $eid /= [
    uri-code: 2,
    SSP: [
      nodenum: uint,
      servicenum: uint
    ]
  ]
]
; The root bundle array
bundle = [primary-block, *extension-block, payload-block]
primary-block = [
  version: 7,
  bundle-control-flags,
  crc-type,
  destination: eid,
  source-node: eid,
  report-to: eid,
  creation-timestamp,
  lifetime: uint,
```

```
? (  
    fragment-offset: uint,  
    total-application-data-length: uint  
)  
? crc-value,  
]  
bundle-control-flags = uint .bits bundleflagbits  
bundleflagbits = &(  
    reserved: 15,  
    reserved: 14,  
    reserved: 13,  
    bundle-deletion-status-reports-are-requested: 12,  
    bundle-delivery-status-reports-are-requested: 11,  
    bundle-forwarding-status-reports-are-requested: 10,  
    reserved: 9,  
    bundle-reception-status-reports-are-requested: 8,  
    bundle-contains-a-Manifest-block: 7,  
    status-time-is-requested-in-all-status-reports: 6,  
    user-application-acknowledgement-is-requested: 5,  
    reserved: 4,  
    reserved: 3,  
    bundle-must-not-be-fragmented: 2,  
    payload-is-an-administrative-record: 1,  
    bundle-is-a-fragment: 0
```

```
)  
; Abstract shared structure of all non-primary blocks  
canonical-block-structure = [  
    block-type-code: uint,  
    block-number: uint,  
    block-control-flags,  
    crc-type,  
    ; Each block type defines the content within the bytestring  
    block-type-specific-data,  
    ? crc-value  
]  
block-control-flags = uint .bits blockflagbits  
blockflagbits = &(br/>    reserved: 7,  
    reserved: 6,  
    reserved: 5,  
    reserved: 4,  
    bundle-must-be-deleted-if-block-cannot-be-processed: 3,  
    status-report-must-be-transmitted-if-block-cannot-be-processed: 2,  
    block-must-be-removed-from-bundle-if-it-cannot-be-processed: 1,  
    block-must-be-replicated-in-every-fragment: 0  
)  
block-type-specific-data = bstr / #6.24(bstr)
```

```
; Actual CBOR data embedded in a bytestring, with optional tag to
indicate so
```

```
embedded-cbor<Item> = (bstr .cbor Item) / #6.24(bstr .cbor Item)
```

```
; Extension block type, which does not specialize other than the
code/number
```

```
extension-block = $extension-block-structure .within canonical-
block-structure
```

```
; Generic shared structure of all non-primary blocks
```

```
extension-block-use<CodeValue, BlockData> = [
```

```
  block-type-code: CodeValue,
```

```
  block-number: (uint .gt 1),
```

```
  block-control-flags,
```

```
  crc-type,
```

```
  BlockData,
```

```
  ? crc-value
```

```
]
```

```
; Payload block type
```

```
payload-block = payload-block-structure .within canonical-block-
structure
```

```
payload-block-structure = [
```

```
  block-type-code: 1,
```

```
  block-number: 1,
```

```
  block-control-flags,
```

```
  crc-type,
```

```
  $payload-block-data,
```

```
  ? crc-value
```

```
]
; Arbitrary payload data, including non-CBOR bytestring
$payload-block-data /= block-type-specific-data
; Administrative record as a payload data specialization
$payload-block-data /= embedded-cbor<admin-record>
admin-record = $admin-record .within admin-record-structure
admin-record-structure = [
  record-type-code: uint,
  record-content: any
]
; Only one defined record type
$admin-record /= [1, status-record-content]
status-record-content = [
  bundle-status-information,
  status-report-reason-code: uint,
  source-node-eid: eid,
  subject-creation-timestamp: creation-timestamp,
  ? (
    subject-payload-offset: uint,
    subject-payload-length: uint
  )
]
bundle-status-information = [
  reporting-node-received-bundle: status-info-content,
```

```
    reporting-node-forwarded-bundle: status-info-content,  
    reporting-node-delivered-bundle: status-info-content,  
    reporting-node-deleted-bundle: status-info-content  
]  
status-info-content = [  
    status-indicator: bool,  
    ? timestamp: dtn-time  
]  
; Previous Node extension block  
$extension-block-structure /=  
    extension-block-use<7, embedded-cbor<ext-data-previous-node>>  
ext-data-previous-node = eid  
; Bundle Age extension block  
$extension-block-structure /=  
    extension-block-use<8, embedded-cbor<ext-data-bundle-age>>  
ext-data-bundle-age = uint  
; Hop Count extension block  
$extension-block-structure /=  
    extension-block-use<9, embedded-cbor<ext-data-hop-count>>  
ext-data-hop-count = [  
    hop-limit: uint,  
    hop-count: uint  
]  
]
```

Authors' Addresses

Scott Burleigh
Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Dr.
Pasadena, CA 91109-8099
US
Phone: +1 818 393 3353
Email: Scott.C.Burleigh@jpl.nasa.gov

Kevin Fall
Roland Computing Services
3871 Piedmont Ave. Suite 8
Oakland, CA 94611
US
Email: kfall+rsc@kfall.com

Edward J. Birrane
Johns Hopkins University Applied Physics Laboratory
11100 Johns Hopkins Rd
Laurel, MD 20723
US
Phone: +1 443 778 7423
Email: Edward.Birrane@jhuapl.edu

Delay-Tolerant Networking
Internet-Draft
Intended status: Standards Track
Expires: May 3, 2018

E. Birrane
K. McKeever
JHU/APL
October 30, 2017

Bundle Protocol Security Specification
draft-ietf-dtn-bpsec-06

Abstract

This document defines a security protocol providing end to end data integrity and confidentiality services for the Bundle Protocol.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on May 3, 2018.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

- 1. Introduction 3
 - 1.1. Supported Security Services 3
 - 1.2. Specification Scope 4
 - 1.3. Related Documents 5
 - 1.4. Terminology 5
- 2. Design Decisions 6
 - 2.1. Block-Level Granularity 6
 - 2.2. Multiple Security Sources 7
 - 2.3. Mixed Security Policy 7
 - 2.4. User-Selected Cipher Suites 8
 - 2.5. Deterministic Processing 8
- 3. Security Blocks 8
 - 3.1. Block Definitions 8
 - 3.2. Uniqueness 9
 - 3.3. Target Multiplicity 9
 - 3.4. Target Identification 10
 - 3.5. Block Representation 10
 - 3.6. Abstract Security Block 11
 - 3.7. Block Integrity Block 14
 - 3.8. Block Confidentiality Block 15
 - 3.9. Block Interactions 16
 - 3.10. Cipher Suite Parameter and Result Identification 17
 - 3.11. BSP Block Example 18
- 4. Canonical Forms 19
- 5. Security Processing 20
 - 5.1. Bundles Received from Other Nodes 20
 - 5.1.1. Receiving BCB Blocks 20
 - 5.1.2. Receiving BIB Blocks 21
 - 5.2. Bundle Fragmentation and Reassembly 22
- 6. Key Management 22
- 7. Security Policy Considerations 23
- 8. Security Considerations 24
 - 8.1. Attacker Capabilities and Objectives 24
 - 8.2. Attacker Behaviors and BPSec Mitigations 25
 - 8.2.1. Eavesdropping Attacks 25
 - 8.2.2. Modification Attacks 26
 - 8.2.3. Topology Attacks 27
 - 8.2.4. Message Injection 28
- 9. Cipher Suite Authorship Considerations 28
- 10. Defining Other Security Blocks 29
- 11. IANA Considerations 30
 - 11.1. Bundle Block Types 30
- 12. References 30
 - 12.1. Normative References 31
 - 12.2. Informative References 31
- Appendix A. Acknowledgements 31

Authors' Addresses 31

1. Introduction

This document defines security features for the Bundle Protocol (BP) [BPBIS] and is intended for use in Delay Tolerant Networks (DTNs) to provide end-to-end security services.

The Bundle Protocol specification [BPBIS] defines DTN as referring to "a networking architecture providing communications in and/or through highly stressed environments" where "BP may be viewed as sitting at the application layer of some number of constituent networks, forming a store-carry-forward overlay network". The term "stressed" environment refers to multiple challenging conditions including intermittent connectivity, large and/or variable delays, asymmetric data rates, and high bit error rates.

The BP might be deployed such that portions of the network cannot be trusted, posing the usual security challenges related to confidentiality and integrity. However, the stressed nature of the BP operating environment imposes unique conditions where usual transport security mechanisms may not be sufficient. For example, the store-carry-forward nature of the network may require protecting data at rest, preventing unauthorized consumption of critical resources such as storage space, and operating without regular contact with a centralized security oracle (such as a certificate authority).

An end-to-end security service is needed that operates in all of the environments where the BP operates.

1.1. Supported Security Services

BPsec provides end-to-end integrity and confidentiality services for BP bundles.

Integrity services ensure that protected data within a bundle are not changed from the time they are provided to the network to the time they are delivered at their destination. Data changes may be caused by processing errors, environmental conditions, or intentional manipulation.

Confidentiality services ensure that protected data is unintelligible to nodes in the DTN, except for authorized nodes possessing special information. Confidentiality, in this context, applies to the contents of protected data and does not extend to hiding the fact that protected data exist in the bundle.

NOTE: Hop-by-hop authentication is NOT a supported security service in this specification, for three reasons.

1. The term "hop-by-hop" is ambiguous in a BP overlay, as nodes that are adjacent in the overlay may not be adjacent in physical connectivity. This condition is difficult or impossible to detect and therefore hop-by-hop authentication is difficult or impossible to enforce.
2. Networks in which BPSec may be deployed may have a mixture of security-aware and not-security-aware nodes. Hop-by-hop authentication cannot be deployed in a network if adjacent nodes in the network have different security capabilities.
3. Hop-by-hop authentication is a special case of data integrity and can be achieved with the integrity mechanisms defined in this specification. Therefore, a separate authentication service is not necessary.

1.2. Specification Scope

This document defines the security services provided by the BPSec. This includes the data specification for representing these services as BP extension blocks, and the rules for adding, removing, and processing these blocks at various points during the bundle's traversal of the DTN.

BPSec applies only to those nodes that implement it, known as "security-aware" nodes. There might be other nodes in the DTN that do not implement BPSec. While all nodes in a BP overlay can exchange bundles, BPSec security operations can only happen at BPSec security-aware nodes.

This specification does not address individual cipher suite implementations. Different networking conditions and operational considerations require varying strengths of security mechanism such that mandating a cipher suite in this specification may result in too much security for some networks and too little security in others. It is expected that separate documents will be standardized to define cipher suites compatible with BPSec, to include operational cipher suites and interoperability cipher suites.

This specification does not address the implementation of security policy and does not provide a security policy for the BPSec. Similar to cipher suites, security policies are based on the nature and capabilities of individual networks and network operational concepts. This specification does provide policy considerations when building a security policy.

This specification does not address how to combine the BPsec security blocks with other protocols, other BP extension blocks, or other best practices to achieve security in any particular network implementation.

1.3. Related Documents

This document is best read and understood within the context of the following other DTN documents:

"Delay-Tolerant Networking Architecture" [RFC4838] defines the architecture for DTNs and identifies certain security assumptions made by existing Internet protocols that are not valid in a DTN.

The Bundle Protocol [BPBIS] defines the format and processing of bundles, defines the extension block format used to represent BPsec security blocks, and defines the canonicalization algorithms used by this specification.

The Bundle Security Protocol [RFC6257] and Streamlined Bundle Security Protocol [SBSP] documents introduced the concepts of using BP extension blocks for security services in a DTN. The BPsec is a continuation and refinement of these documents.

1.4. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

This section defines terminology either unique to the BPsec or otherwise necessary for understanding the concepts defined in this specification.

- o Bundle Source - the node which originates a bundle. The Node ID of the BPA originating the bundle.
- o Forwarder - any node that transmits a bundle in the DTN. The Node ID of the Bundle Protocol Agent (BPA) that sent the bundle on its most recent hop.
- o Intermediate Receiver, Waypoint, or "Next Hop" - any node that receives a bundle from a Forwarder that is not the Destination. The Node ID of the BPA at any such node.

- o Path - the ordered sequence of nodes through which a bundle passes on its way from Source to Destination. The path is not necessarily known in advance by the bundle or any BPAs in the DTN.
- o Security Block - a BPSec extension block in a bundle.
- o Security Operation - the application of a security service to a security target, notated as OP(security service, security target). For example, OP(confidentiality, payload). Every security operation in a bundle MUST be unique, meaning that a security service can only be applied to a security target once in a bundle. A security operation is implemented by a security block.
- o Security Service - the security features supported by this specification: integrity and confidentiality.
- o Security Source - a bundle node that adds a security block to a bundle. The Node ID of that node.
- o Security Target - the block within a bundle that receives a security-service as part of a security-operation.

2. Design Decisions

The application of security services in a DTN is a complex endeavor that must consider physical properties of the network, policies at each node, and various application security requirements. This section identifies those desirable properties that guide design decisions for this specification and are necessary for understanding the format and behavior of the BPSec protocol.

2.1. Block-Level Granularity

Security services within this specification must allow different blocks within a bundle to have different security services applied to them.

Blocks within a bundle represent different types of information. The primary block contains identification and routing information. The payload block carries application data. Extension blocks carry a variety of data that may augment or annotate the payload, or otherwise provide information necessary for the proper processing of a bundle along a path. Therefore, applying a single level and type of security across an entire bundle fails to recognize that blocks in a bundle may represent different types of information with different security needs.

For example, a payload block might be encrypted to protect its contents and an extension block containing summary information related to the payload might be integrity signed but unencrypted to provide waypoints access to payload-related data without providing access to the payload.

2.2. Multiple Security Sources

A bundle MAY have multiple security blocks and these blocks MAY have different security sources.

The Bundle Protocol allows extension blocks to be added to a bundle at any time during its existence in the DTN. When a waypoint adds a new extension block to a bundle, that extension block may have security services applied to it by that waypoint. Similarly, a waypoint may add a security service to an existing extension block, consistent with its security policy. For example, a node representing a boundary between a trusted part of the network and an untrusted part of the network may wish to apply payload encryption for bundles leaving the trusted portion of the network.

When a waypoint adds a security service to the bundle, the waypoint is the security source for that service. The security block(s) which represent that service in the bundle may need to record this security source as the bundle destination might need this information for processing. For example, a destination node might interpret policy as it related to security blocks as a function of the security source for that block.

2.3. Mixed Security Policy

The security policy enforced by nodes in the DTN MAY differ.

Some waypoints may not be security aware and will not be able to process security blocks. Therefore, security blocks must have their processing flags set such that the block will be treated appropriately by non-security-aware waypoints

Some waypoints will have security policies that require evaluating security services even if they are not the bundle destination or the final intended destination of the service. For example, a waypoint may choose to verify an integrity service even though the waypoint is not the bundle destination and the integrity service will be needed by other node along the bundle's path.

Some waypoints will determine, through policy, that they are the intended recipient of the security service and terminate the security service in the bundle. For example, a gateway node may determine

that, even though it is not the destination of the bundle, it should verify and remove a particular integrity service or attempt to decrypt a confidentiality service, before forwarding the bundle along its path.

Some waypoints may understand security blocks but refuse to process them unless they are the bundle destination.

2.4. User-Selected Cipher Suites

The security services defined in this specification rely on a variety of cipher suites providing integrity signatures, cipher-text, and other information necessary to populate security blocks. Users MAY select different cipher suites to implement security services. For example, some users might prefer a SHA2 hash function for integrity whereas other users may prefer a SHA3 hash function instead. The security services defined in this specification must provide a mechanism for identifying what cipher suite has been used to populate a security block.

2.5. Deterministic Processing

Whenever a node determines that it must process more than one security block in a received bundle (either because the policy at a waypoint states that it should process security blocks or because the node is the bundle destination) the order in which security blocks are processed must be deterministic. All nodes must impose this same deterministic processing order for all security blocks. This specification provides determinism in the application and evaluation of security services, even when doing so results in a loss of flexibility.

3. Security Blocks

3.1. Block Definitions

This specification defines two types of security block: the Block Integrity Block (BIB) and the Block Confidentiality Block (BCB).

The BIB is used to ensure the integrity of its security target(s). The integrity information in the BIB MAY be verified by any node in between the BIB security source and the bundle destination. Security-aware waypoints may add or remove BIBs from bundles in accordance with their security policy.

The BCB indicates that the security target(s) have been encrypted at the BCB security source in order to protect its content while in transit. The BCB may be decrypted by security-aware nodes in

the network, up to and including the bundle destination, as a matter of security policy.

3.2. Uniqueness

Security operations in a bundle MUST be unique - the same security service MUST NOT be applied to a security target more than once in a bundle. Since a security operation is represented as a security block, this limits what security blocks may be added to a bundle: if adding a security block to a bundle would cause some other security block to no longer represent a unique security operation then the new block MUST NOT be added.

If multiple security blocks representing the same security operation were allowed in a bundle at the same time, there would exist ambiguity regarding block processing order and the property of deterministic processing blocks would be lost.

Using the notation `OP(service,target)`, several examples illustrate this uniqueness requirement.

- o Signing the payload twice: The two operations `OP(integrity, payload)` and `OP(integrity, payload)` are redundant and MUST NOT both be present in the same bundle at the same time.
- o Signing different blocks: The two operations `OP(integrity, payload)` and `OP(integrity, extension_block_1)` are not redundant and both may be present in the same bundle at the same time. Similarly, the two operations `OP(integrity, extension_block_1)` and `OP(integrity, extension_block_2)` are also not redundant and may both be present in the bundle at the same time.
- o Different Services on same block: The two operations `OP(integrity, payload)` and `OP(confidentiality, payload)` are not inherently redundant and may both be present in the bundle at the same time, pursuant to other processing rules in this specification.

3.3. Target Multiplicity

Under special circumstances, a single security block may represent multiple security operations as a way of reducing the overall number of security blocks present in a bundle. In these circumstances, reducing the number of security blocks in the bundle reduces the amount of redundant information in the bundle.

A set of security operations may be represented by a single security block if and only if the following conditions are true.

- o The security operations apply the same security service. For example, they are all integrity operations or all confidentiality operations.
- o The cipher suite parameters and key information for the security operations are identical.
- o The security source for the security operations is the same. Meaning the set of operations are being added/removed by the same node.
- o No security operations have the same security target, as that would violate the need for security operations to be unique.
- o None of the security operations conflict with security operations already present in the bundle.

When representing multiple security operations in a single security block, the information that is common across all operations is represented once in the security block, and the information which is different (e.g., the security targets) are represented individually. When the security block is processed all security operations represented by the security block MUST be applied/evaluated at that time.

3.4. Target Identification

A security target is a block in the bundle to which a security service applies. This target must be uniquely and unambiguously identifiable when processing a security block. The definition of the extension block header from [BPBIS] provides a "Block Number" field suitable for this purpose. Therefore, a security target in a security block MUST be represented as the Block Number of the target block.

3.5. Block Representation

Each security block uses the Canonical Bundle Block Format as defined in [BPBIS]. That is, each security block is comprised of the following elements:

- o Block Type Code
- o Block Number
- o Block Processing Control Flags
- o CRC Type and CRC Field (if present)

- o Block Data Length
- o Block Type Specific Data Fields

Security-specific information for a security block is captured in the "Block Type Specific Data Fields".

3.6. Abstract Security Block

The structure of the security-specific portions of a security block is identical for both the BIB and BCB Block Types. Therefore, this section defines an Abstract Security Block (ASB) data structure and discusses the definition, processing, and other constraints for using this structure. An ASB is never directly instantiated within a bundle, it is only a mechanism for discussing the common aspects of BIB and BCB security blocks.

The fields of the ASB SHALL be as follows, listed in the order in which they must appear.

Security Targets:

This field identifies the block(s) targeted by the security operation(s) represented by this security block. Each target block is represented by its unique Block Number. This field SHALL be represented by a CBOR array of data items. Each target within this CBOR array SHALL be represented by a CBOR unsigned integer. This array MUST have at least 1 entry and each entry MUST represent the Block Number of a block that exists in the bundle. There MUST NOT be duplicate entries in this array.

Cipher Suite Id:

This field identifies the cipher suite used to implement the security service represented by this block and applied to each security target. This field SHALL be represented by a CBOR unsigned integer.

Cipher Suite Flags:

This field identifies which optional fields are present in the security block. This field SHALL be represented as a CBOR unsigned integer containing a bit field of 5 bits indicating the presence or absence of other security block fields, as follows.

Bit 1 (the most-significant bit, 0x10): reserved.

Bit 2 (0x08): reserved.

Bit 3 (0x04): reserved.

Bit 4 (0x02): Security Source Present Flag.

Bit 5 (the least-significant bit, 0x01): Cipher Suite Parameters Present Flag.

In this field, a value of 1 indicates that the associated security block field MUST be included in the security block. A value of 0 indicates that the associated security block field MUST NOT be in the security block.

Security Source (Optional Field):

This field identifies the Endpoint that inserted the security block in the bundle. If the security source field is not present then the source MAY be inferred from other information, such as the bundle source or the previous hop, as defined by security policy. This field SHALL be represented by a CBOR array in accordance with [BPBIS] rules for representing Endpoint Identifiers (EIDs).

Cipher Suite Parameters (Optional Field):

This field captures one or more cipher suite parameters that should be provided to security-aware nodes when processing the security service described by this security block. This field SHALL be represented by a CBOR array. Each entry in this array is a single cipher suite parameter. A single cipher suite parameter SHALL also be represented as a CBOR array comprising a 2-tuple of the id and value of the parameter, as follows.

- * Parameter Id. This field identifies which cipher suite parameter is being specified. This field SHALL be represented as a CBOR unsigned integer. Parameter ids are selected as described in Section 3.10.
- * Parameter Value. This field captures the value associated with this parameter. This field SHALL be represented by the applicable CBOR representation of the parameter, in accordance with Section 3.10.

The logical layout of the cipher suite parameters array is illustrated in Figure 1.

Parameter 1		Parameter 2		...	Parameter N	
Id	Value	Id	Value		Id	Value

Figure 1: Cipher Suite Parameters

Security Results:

This field captures the results of applying a security service to the security targets of the security block. This field SHALL be represented as a CBOR array of target results. Each entry in this array represents the set of security results for a specific security target. The target results MUST be ordered identically to the Security Targets field of the security block. This means that the first set of target results in this array corresponds to the first entry in the Security Targets field of the security block, and so on. There MUST be one entry in this array for each entry in the Security Targets field of the security block.

The set of security results for a target is also represented as a CBOR array of individual results. An individual result is represented as a 2-tuple of a result id and a result value, defined as follows.

- * Result Id. This field identifies which security result is being specified. Some security results capture the primary output of a cipher suite. Other security results contain additional annotative information from cipher suite processing. This field SHALL be represented as a CBOR unsigned integer. Security result ids will be as specified in Section 3.10.
- * Result Value. This field captures the value associated with the result. This field SHALL be represented by the applicable CBOR representation of the result value, in accordance with Section 3.10.

The logical layout of the security results array is illustrated in Figure 2. In this figure there are N security targets for this security block. The first security target contains M results and the Nth security target contains K results.

Target 1				Target N			
Result 1		Result M	...	Result 1		Result K	
Id	Value	Id	Value	Id	Value	Id	Value

Figure 2: Security Results

3.7. Block Integrity Block

A BIB is a bundle extension block with the following characteristics.

- o The Block Type Code value is as specified in Section 11.1.
- o The Block Type Specific Data Fields follow the structure of the ASB.
- o A security target listed in the Security Targets field MUST NOT reference a security block defined in this specification (e.g., a BIB or a BCB).
- o The Cipher Suite Id MUST be documented as an end-to-end authentication-cipher suite or as an end-to-end error-detection-cipher suite.
- o An EID-reference to the security source MAY be present. If this field is not present, then the security source of the block SHOULD be inferred according to security policy and MAY default to the bundle source. The security source may also be specified as part of key information described in Section 3.10.

Notes:

- o It is RECOMMENDED that cipher suite designers carefully consider the effect of setting flags that either discard the block or delete the bundle in the event that this block cannot be processed.
- o Since OP(integrity, target) is allowed only once in a bundle per target, it is RECOMMENDED that users wishing to support multiple integrity signatures for the same target define a multi-signature cipher suite.
- o For some cipher suites, (e.g., those using asymmetric keying to produce signatures or those using symmetric keying with a group key), the security information MAY be checked at any hop on the

way to the destination that has access to the required keying information, in accordance with Section 3.9.

- o The use of a generally available key is RECOMMENDED if custodial transfer is employed and all nodes SHOULD verify the bundle before accepting custody.

3.8. Block Confidentiality Block

A BCB is a bundle extension block with the following characteristics.

The Block Type Code value is as specified in Section 11.1.

The Block Processing Control flags value can be set to whatever values are required by local policy, except that this block MUST have the "replicate in every fragment" flag set if the target of the BCB is the Payload Block. Having that BCB in each fragment indicates to a receiving node that the payload portion of each fragment represents cipher-text.

The Block Type Specific Data Fields follow the structure of the ASB.

A security target listed in the Security Targets field MAY reference the payload block, a non-security extension block, or a BIB block. A BCB MUST NOT include another BCB as a security target. A BCB MUST NOT target the primary block.

The Cipher Suite Id MUST be documented as a confidentiality cipher suite.

Any additional bytes generated from applying the cipher suite to a security target (such as additional authenticated text) MAY be placed in an appropriate security result (e.g., an Integrity Check Value) in accordance with cipher suite and security policy.

An EID-reference to the security source MAY be present. If this field is not present, then the security source of the block SHOULD be inferred according to security policy and MAY default to the bundle source. The security source may also be specified as part of key information described in Section 3.10.

The BCB modifies the contents of its security target(s). When a BCB is applied, the security target body data are encrypted "in-place". Following encryption, the security target Block Type Specific Data Fields contains cipher-text, not plain-text. Other block fields remain unmodified, with the exception of the Block Data Length field,

which may be changed if the BCB is allowed to change the length of the block (see below).

Fragmentation, reassembly, and custody transfer are adversely affected by a change in size of the payload block due to ambiguity about what byte range of the block is actually in any particular fragment. Therefore, when the security target of a BCB is the bundle payload, the BCB MUST NOT alter the size of the payload block body data. This "in-place" encryption allows fragmentation, reassembly, and custody transfer to operate without knowledge of whether or not encryption has occurred.

If a BCB cannot alter the size of the security target (e.g., the security target is the payload block or block length modifications are disallowed by policy) then differences in the size of the cipher-text and plain-text must be handled in the following way. If the cipher-text is shorter in length than the plain-text, padding MUST be used in accordance with the cipher suite policy. If the cipher-text is larger than the plain-text, overflow bytes MUST be placed in overflow parameters in the Security Result field.

Notes:

- o It is RECOMMENDED that cipher suite designers carefully consider the effect of setting flags that either discard the block or delete the bundle in the event that this block cannot be processed.
- o The BCB block processing control flags MAY be set independently from the processing control flags of the security target(s). The setting of such flags SHOULD be an implementation/policy decision for the encrypting node.
- o A BCB MAY include information as part of additional authenticated data to address parts of the target block that are not converted to cipher-text.

3.9. Block Interactions

The security block types defined in this specification are designed to be as independent as possible. However, there are some cases where security blocks may share a security target creating processing dependencies.

If confidentiality is being applied to a target that already has integrity applied to it, then an undesirable condition occurs where a security aware waypoint would be unable to check the integrity result of a block because the block contents have been encrypted after the

integrity signature was generated. To address this concern, the following processing rules must be followed.

- o If confidentiality is to be applied to a target, it MUST also be applied to any integrity operation already defined for that target. This means that if a BCB is added to encrypt a block, another BCB MUST also be added to encrypt a BIB also targeting that block.
- o An integrity operation MUST NOT be applied to a security target if a BCB in the bundle shares the same security target. This prevents ambiguity in the order of evaluation when receiving a BIB and a BCB for a given security target.
- o An integrity value MUST NOT be evaluated if the BIB providing the integrity value is the security target of an existing BCB block in the bundle. In such a case, the BIB data contains cipher-text as it has been encrypted.
- o An integrity value MUST NOT be evaluated if the security target of the BIB is also the security target of a BCB in the bundle. In such a case, the security target data contains cipher-text as it has been encrypted.
- o As mentioned in Section 3.7, a BIB MUST NOT have a BCB as its security target. BCBs may embed integrity results as part of security results.

These restrictions on block interactions impose a necessary ordering when applying security operations within a bundle. Specifically, for a given security target, BIBs MUST be added before BCBs. This ordering MUST be preserved in cases where the current BPA is adding all of the security blocks for the bundle or whether the BPA is a waypoint adding new security blocks to a bundle that already contains security blocks.

3.10. Cipher Suite Parameter and Result Identification

Cipher suite parameters and security results each represent multiple distinct pieces of information in a security block. Each piece of information is assigned an identifier and a CBOR encoding. Identifiers MUST be unique for a given cipher suite but do not need to be unique across all cipher suites. Therefore, parameter ids and security result ids are specified in the context of a cipher suite definition.

Individual BPSec cipher suites SHOULD use existing registries of identifiers and CBOR encodings, such as those defined in [COSE],

whenever possible. Cipher suites MAY define their own identifiers and CBOR encodings when necessary.

A cipher suite MAY include multiple instances of the same identifier for a parameter or result in a security block. Parameters and results are represented using CBOR, and any identification of a new parameter or result must include how the value will be represented using the CBOR specification. Ids themselves are always represented as a CBOR unsigned integer.

3.11. BSP Block Example

An example of BPsec blocks applied to a bundle is illustrated in Figure 3. In this figure the first column represents blocks within a bundle and the second column represents the Block Number for the block, using the terminology B1...Bn for the purpose of illustration.

Block in Bundle	ID
Primary Block	B1
BIB OP(integrity, target=B1)	B2
BCB OP(confidentiality, target=B4)	B3
Extension Block	B4
BIB OP(integrity, target=B6)	B5
Extension Block	B6
BCB OP(confidentiality, targets=B8, B9)	B7
BIB (encrypted by B7) OP(integrity, target=B9)	B8
Payload Block	B9

Figure 3: Sample Use of BPsec Blocks

In this example a bundle has four non-security-related blocks: the primary block (B1), two extension blocks (B4,B6), and a payload block

(B9). The following security applications are applied to this bundle.

- o An integrity signature applied to the canonicalized primary block. This is accomplished by a single BIB (B2).
- o Confidentiality for the first extension block (B4). This is accomplished by a BCB block (B3).
- o Integrity for the second extension block (B6). This is accomplished by a BIB block (B5). NOTE: If the extension block B6 contains a representation of the serialized bundle (such as a hash over all blocks in the bundle at the time of its last transmission) then the BIB block is also providing an authentication service.
- o An integrity signature on the payload (B10). This is accomplished by a BIB block (B8).
- o Confidentiality for the payload block and its integrity signature. This is accomplished by a BCB block, B7, encrypting B8 and B9. In this case, the security source, key parameters, and service are identical, so a single security block MAY be used for this purpose, rather than requiring two BCBs one to encrypt B8 and one to encrypt B9.

4. Canonical Forms

Security services require consistency and determinism in how information is presented to cipher suites at the security source and at a receiving node. For example, integrity services require that the same target information (e.g., the same bits in the same order) is provided to the cipher suite when generating an original signature and when generating a comparison signature. Canonicalization algorithms are used to construct a stable, end-to-end bit representation of a target block.

Canonical forms are not transmitted, they are used to generate input to a cipher suite for security processing at a security-aware node.

The canonicalization of the primary block is as specified in [BPBIS].

All non-primary blocks share the same block structure and are canonicalized as specified in [BPBIS] with the following exception.

- o If the service being applied is a confidentiality service, then the Block Type Code, Block Number, Block Processing Control Flags, CRC Type and CRC Field (if present), and Block Data Length fields

MUST NOT be included in the canonicalization. Confidentiality services are used solely to convert the Block Type Specific Data Fields from plain-text to cipher-text.

- o Reserved flags MUST NOT be included in any canonicalization as it is not known if those flags will change in transit.

These canonicalization algorithms assume that Endpoint IDs do not change from the time at which a security source adds a security block to a bundle and the time at which a node processes that security block.

Cipher suites MAY define their own canonicalization algorithms and require the use of those algorithms over the ones provided in this specification. In the event of conflicting canonicalization algorithms, cipher suite algorithms take precedence over this specification.

5. Security Processing

This section describes the security aspects of bundle processing.

5.1. Bundles Received from Other Nodes

Security blocks must be processed in a specific order when received by a security-aware node. The processing order is as follows.

- o All BCB blocks in the bundle MUST be evaluated prior to evaluating any BIBs in the bundle. When BIBs and BCBS share a security target, BCBS MUST be evaluated first and BIBs second.

5.1.1. Receiving BCB Blocks

If a received bundle contains a BCB, the receiving node must determine whether it has the responsibility of decrypting the BCB security target and removing the BCB prior to delivering data to an application at the node or forwarding the bundle.

If the receiving node is the destination of the bundle, the node MUST decrypt any BCBS remaining in the bundle. If the receiving node is not the destination of the bundle, the node MAY decrypt the BCB if directed to do so as a matter of security policy.

If the security policy of a security-aware node specifies that a bundle should have applied confidentiality to a specific security target and no such BCB is present in the bundle, then the node MUST process this security target in accordance with the security policy. This MAY involve removing the security target from the bundle. If

the removed security target is the payload block, the bundle MAY be discarded.

If an encrypted payload block cannot be decrypted (i.e., the decryption key cannot be deduced or decryption fails), then the bundle MUST be discarded and processed no further. If an encrypted security target other than the payload block cannot be decrypted then the associated security target and all security blocks associated with that target MUST be discarded and processed no further. In both cases, requested status reports (see [BPBIS]) MAY be generated to reflect bundle or block deletion.

When a BCB is decrypted, the recovered plain-text MUST replace the cipher-text in the security target Block Type Specific Data Fields. If the Block Data Length field was modified at the time of encryption it MUST be updated to reflect the decrypted block length.

If a BCB contains multiple security targets, all security targets MUST be processed when the BCB is processed. Errors and other processing steps SHALL be made as if each security target had been represented by an individual BCB with a single security target.

5.1.2. Receiving BIB Blocks

If a received bundle contains a BIB, the receiving node MUST determine whether it has the final responsibility of verifying the BIB security target and removing it prior to delivering data to an application at the node or forwarding the bundle. If a BIB check fails, the security target has failed to authenticate and the security target SHALL be processed according to the security policy. A bundle status report indicating the failure MAY be generated. Otherwise, if the BIB verifies, the security target is ready to be processed for delivery.

A BIB MUST NOT be processed if the security target of the BIB is also the security target of a BCB in the bundle. Given the order of operations mandated by this specification, when both a BIB and a BCB share a security target, it means that the security target must have been encrypted after it was integrity signed and, therefore, the BIB cannot be verified until the security target has been decrypted by processing the BCB.

If the security policy of a security-aware node specifies that a bundle should have applied integrity to a specific security target and no such BIB is present in the bundle, then the node MUST process this security target in accordance with the security policy. This MAY involve removing the security target from the bundle. If the removed security target is the payload or primary block, the bundle

MAY be discarded. This action may occur at any node that has the ability to verify an integrity signature, not just the bundle destination.

If a receiving node does not have the final responsibility of verifying the BIB it MAY still attempt to verify the BIB to prevent the needless forwarding of corrupt data. If the check fails, the node SHALL process the security target in accordance to local security policy. It is RECOMMENDED that if a payload integrity check fails at a waypoint that it is processed in the same way as if the check fails at the destination. If the check passes, the node MUST NOT remove the BIB prior to forwarding.

If a BIB contains multiple security targets, all security targets MUST be processed if the BIB is processed by the Node. Errors and other processing steps SHALL be made as if each security target had been represented by an individual BIB with a single security target.

5.2. Bundle Fragmentation and Reassembly

If it is necessary for a node to fragment a bundle payload, and security services have been applied to that bundle, the fragmentation rules described in [BPBIS] MUST be followed. As defined there and summarized here for completeness, only the payload block may be fragmented; security blocks, like all extension blocks, can never be fragmented.

Due to the complexity of payload block fragmentation, including the possibility of fragmenting payload block fragments, integrity and confidentiality operations are not to be applied to a bundle representing a fragment. Specifically, a BCB or BIB MUST NOT be added to a bundle if the "Bundle is a Fragment" flag is set in the Bundle Processing Control Flags field.

Security processing in the presence of payload block fragmentation MAY be handled by other mechanisms outside of the BPsec protocol or by applying BPsec blocks in coordination with an encapsulation mechanism.

6. Key Management

There exist a myriad of ways to establish, communicate, and otherwise manage key information in a DTN. Certain DTN deployments might follow established protocols for key management whereas other DTN deployments might require new and novel approaches. BPsec assumes that key management is handled as a separate part of network management and this specification neither defines nor requires a specific key management strategy.

7. Security Policy Considerations

When implementing BPSec, several policy decisions must be considered. This section describes key policies that affect the generation, forwarding, and receipt of bundles that are secured using this specification. No single set of policy decisions is envisioned to work for all secure DTN deployments.

- o If a bundle is received that contains more than one security operation, in violation of BPSec, then the BPA must determine how to handle this bundle. The bundle may be discarded, the block affected by the security operation may be discarded, or one security operation may be favored over another.
- o BPAs in the network must understand what security operations they should apply to bundles. This decision may be based on the source of the bundle, the destination of the bundle, or some other information related to the bundle.
- o If a waypoint has been configured to add a security operation to a bundle, and the received bundle already has the security operation applied, then the receiver must understand what to do. The receiver may discard the bundle, discard the security target and associated BPSec blocks, replace the security operation, or some other action.
- o It is recommended that security operations only be applied to the blocks that absolutely need them. If a BPA were to apply security operations such as integrity or confidentiality to every block in the bundle, regardless of need, there could be downstream errors processing blocks whose contents must be inspected or changed at every hop along the path.
- o Adding a BIB to a security target that has already been encrypted by a BCB is not allowed. If this condition is likely to be encountered, there are (at least) three possible policies that could handle this situation.
 1. At the time of encryption, an integrity signature may be generated and added to the BCB for the security target as additional information in the security result field.
 2. The encrypted block may be replicated as a new block and integrity signed.
 3. An encapsulation scheme may be applied to encapsulate the security target (or the entire bundle) such that the encapsulating structure is, itself, no longer the security

target of a BCB and may therefore be the security target of a BIB.

8. Security Considerations

Given the nature of DTN applications, it is expected that bundles may traverse a variety of environments and devices which each pose unique security risks and requirements on the implementation of security within BPSec. For these reasons, it is important to introduce key threat models and describe the roles and responsibilities of the BPSec protocol in protecting the confidentiality and integrity of the data against those threats. This section provides additional discussion on security threats that BPSec will face and describes how BPSec security mechanisms operate to mitigate these threats.

It should be noted that BPSec addresses only the security of data traveling over the DTN, not the underlying DTN itself. Additionally, BPSec addresses neither the fitness of externally-defined cryptographic methods nor the security of their implementation. It is the responsibility of the BPSec implementer that appropriate algorithms and methods are chosen. Furthermore, the BPSec protocol does not address threats which share computing resources with the DTN and/or BPSec software implementations. These threats may be malicious software or compromised libraries which intend to intercept data or recover cryptographic material. Here, it is the responsibility of the BPSec implementer to ensure that any cryptographic material, including shared secret or private keys, is protected against access within both memory and storage devices.

The threat model described here is assumed to have a set of capabilities identical to those described by the Internet Threat Model in [RFC3552], but the BPSec threat model is scoped to illustrate threats specific to BPSec operating within DTN environments and therefore focuses on man-in-the-middle (MITM) attackers.

8.1. Attacker Capabilities and Objectives

BPSec was designed to protect against MITM threats which may have access to a bundle during transit from its source, Alice, to its destination, Bob. A MITM node, Mallory, is a non-cooperative node operating on the DTN between Alice and Bob that has the ability to receive bundles, examine bundles, modify bundles, forward bundles, and generate bundles at will in order to compromise the confidentiality or integrity of data within the DTN. For the purposes of this section, any MITM node is assumed to effectively be security-aware even if it does not implement the BPSec protocol.

There are three classes of MITM nodes which are differentiated based on their access to cryptographic material:

- o Unprivileged Node: Mallory has not been provisioned within the secure environment and only has access to cryptographic material which has been publicly-shared.
- o Legitimate Node: Mallory is within the secure environment and therefore has access to cryptographic material which has been provisioned to Mallory (i.e., K_M) as well as material which has been publicly-shared.
- o Privileged Node: Mallory is a privileged node within the secure environment and therefore has access to cryptographic material which has been provisioned to Mallory, Alice and/or Bob (i.e. K_M , K_A , and/or K_B) as well as material which has been publicly-shared.

If Mallory is operating as a privileged node, this is tantamount to compromise; BPSec does not provide mechanisms to detect or remove Mallory from the DTN or BPSec secure environment. It is up to the BPSec implementer or the underlying cryptographic mechanisms to provide appropriate capabilities if they are needed. It should also be noted that if the implementation of BPSec uses a single set of shared cryptographic material for all nodes, a legitimate node is equivalent to a privileged node because $K_M == K_A == K_B$.

A special case of the legitimate node is when Mallory is either Alice or Bob (i.e., $K_M == K_A$ or $K_M == K_B$). In this case, Mallory is able to impersonate traffic as either Alice or Bob, which means that traffic to and from that node can be decrypted and encrypted, respectively. Additionally, messages may be signed as originating from one of the endpoints.

8.2. Attacker Behaviors and BPSec Mitigations

8.2.1. Eavesdropping Attacks

Once Mallory has received a bundle, she is able to examine the contents of that bundle and attempt to recover any protected data or cryptographic keying material from the blocks contained within. The protection mechanism that BPSec provides against this action is the BCB, which encrypts the contents of its security target, providing confidentiality of the data. Of course, it should be assumed that Mallory is able to attempt offline recovery of encrypted data, so the cryptographic mechanisms selected to protect the data should provide a suitable level of protection.

When evaluating the risk of eavesdropping attacks, it is important to consider the lifetime of bundles on a DTN. Depending on the network, bundles may persist for days or even years. Long-lived bundles imply that the data exists in the network for a longer period of time and, thus, there may be more opportunities to capture those bundles. Additionally, bundles that are long-lived imply that the information stored within them may remain relevant and sensitive for long enough that, once captured, there is sufficient time to crack encryption associated with the bundle. If a bundle does persist on the network for years and the cipher suite used for a BCB provides inadequate protection, Mallory may be able to recover the protected data either before that bundle reaches its intended destination or before the information in the bundle is no longer considered sensitive.

8.2.2. Modification Attacks

As a node participating in the DTN between Alice and Bob, Mallory will also be able to modify the received bundle, including non-BPsec data such as the primary block, payload blocks, or block processing control flags as defined in [BPBIS]. Mallory will be able to undertake activities which include modification of data within the blocks, replacement of blocks, addition of blocks, or removal of blocks. Within BPsec, both the BIB and BCB provide integrity protection mechanisms to detect or prevent data manipulation attempts by Mallory.

The BIB provides that protection to another block which is its security target. The cryptographic mechanisms used to generate the BIB should be strong against collision attacks and Mallory should not have access to the cryptographic material used by the originating node to generate the BIB (e.g., K_A). If both of these conditions are true, Mallory will be unable to modify the security target or the BIB and lead Bob to validate the security target as originating from Alice.

Since BPsec security operations are implemented by placing blocks in a bundle, there is no in-band mechanism for detecting or correcting certain cases where Mallory removes blocks from a bundle. If Mallory removes a BCB block, but keeps the security target, the security target remains encrypted and there is a possibility that there may no longer be sufficient information to decrypt the block at its destination. If Mallory removes both a BCB (or BIB) and its security target there is no evidence left in the bundle of the security operation. Similarly, if Mallory removes the BIB but not the security target there is no evidence left in the bundle of the security operation. In each of these cases, the implementation of BPsec must be combined with policy configuration at endpoints in the network which describe the expected and required security operations

that must be applied on transmission and are expected to be present on receipt. This or other similar out-of-band information is required to correct for removal of security information in the bundle.

A limitation of the BIB may exist within the implementation of BIB validation at the destination node. If Mallory is a legitimate node within the DTN, the BIB generated by Alice with K_A can be replaced with a new BIB generated with K_M and forwarded to Bob. If Bob is only validating that the BIB was generated by a legitimate user, Bob will acknowledge the message as originating from Mallory instead of Alice. In order to provide verifiable integrity checks, both a BIB and BCB should be used and the BCB should require an IND-CCA2 encryption scheme. Such an encryption scheme will guard against signature substitution attempts by Mallory. In this case, Alice creates a BIB with the protected data block as the security target and then creates a BCB with both the BIB and protected data block as its security targets.

8.2.3. Topology Attacks

If Mallory is in a MITM position within the DTN, she is able to influence how any bundles that come to her may pass through the network. Upon receiving and processing a bundle that must be routed elsewhere in the network, Mallory has three options as to how to proceed: not forward the bundle, forward the bundle as intended, or forward the bundle to one or more specific nodes within the network.

Attacks that involve re-routing the packets throughout the network are essentially a special case of the modification attacks described in this section where the attacker is modifying fields within the primary block of the bundle. Given that BPSec cannot encrypt the contents of the primary block, alternate methods must be used to prevent this situation. These methods MAY include requiring BIBs for primary blocks, using encapsulation, or otherwise strategically manipulating primary block data. The specifics of any such mitigation technique are specific to the implementation of the deploying network and outside of the scope of this document.

Furthermore, routing rules and policies may be useful in enforcing particular traffic flows to prevent topology attacks. While these rules and policies may utilize some features provided by BPSec, their definition is beyond the scope of this specification.

8.2.4. Message Injection

Mallory is also able to generate new bundles and transmit them into the DTN at will. These bundles may either be copies or slight modifications of previously-observed bundles (i.e., a replay attack) or entirely new bundles generated based on the Bundle Protocol, BPSec, or other bundle-related protocols. With these attacks Mallory's objectives may vary, but may be targeting either the bundle protocol or application-layer protocols conveyed by the bundle protocol.

BPSec relies on cipher suite capabilities to prevent replay or forged message attacks. A BCB used with appropriate cryptographic mechanisms (e.g., a counter-based cipher mode) may provide replay protection under certain circumstances. Alternatively, application data itself may be augmented to include mechanisms to assert data uniqueness and then protected with a BIB, a BCB, or both along with other block data. In such a case, the receiving node would be able to validate the uniqueness of the data.

9. Cipher Suite Authorship Considerations

Cipher suite developers or implementers should consider the diverse performance and conditions of networks on which the Bundle Protocol (and therefore BPSec) will operate. Specifically, the delay and capacity of delay-tolerant networks can vary substantially. Cipher suite developers should consider these conditions to better describe the conditions when those suites will operate or exhibit vulnerability, and selection of these suites for implementation should be made with consideration to the reality. There are key differences that may limit the opportunity to leverage existing cipher suites and technologies that have been developed for use in traditional, more reliable networks:

- o Data Lifetime: Depending on the application environment, bundles may persist on the network for extended periods of time, perhaps even years. Cryptographic algorithms should be selected to ensure protection of data against attacks for a length of time reasonable for the application.
- o One-Way Traffic: Depending on the application environment, it is possible that only a one-way connection may exist between two endpoints, or if a two-way connection does exist, the round-trip time may be extremely large. This may limit the utility of session key generation mechanisms, such as Diffie-Hellman, as a two-way handshake may not be feasible or reliable.

- o Opportunistic Access: Depending on the application environment, a given endpoint may not be guaranteed to be accessible within a certain amount of time. This may make asymmetric cryptographic architectures which rely on a key distribution center or other trust center impractical under certain conditions.

When developing new cipher suites for use with BPSec, the following information SHOULD be considered for inclusion in these specifications.

- o Cipher Suite Parameters. Cipher suites MUST define their parameter ids, the data types of those parameters, and their CBOR encoding.
- o Security Results. Cipher suites MUST define their security result ids, the data types of those results, and their CBOR encoding.
- o New Canonicalizations. Cipher suites MAY define new canonicalization algorithms as necessary.

10. Defining Other Security Blocks

Other security blocks (OSBs) may be defined and used in addition to the security blocks identified in this specification. Both the usage of BIB, BCB, and any future OSBs MAY co-exist within a bundle and MAY be considered in conformance with BPSec if each of the following requirements are met by any future identified security blocks.

- o Other security blocks (OSBs) MUST NOT reuse any enumerations identified in this specification, to include the block type codes for BIB and BCB.
- o An OSB definition MUST state whether it can be the target of a BIB or a BCB. The definition MUST also state whether the OSB can target a BIB or a BCB.
- o An OSB definition MUST provide a deterministic processing order in the event that a bundle is received containing BIBs, BCBs, and OSBs. This processing order MUST NOT alter the BIB and BCB processing orders identified in this specification.
- o An OSB definition MUST provide a canonicalization algorithm if the default non-primary-block canonicalization algorithm cannot be used to generate a deterministic input for a cipher suite. This requirement MAY be waived if the OSB is defined so as to never be the security target of a BIB or a BCB.

- o An OSB definition MAY NOT require any behavior of a BPSEC-BPA that is in conflict with the behavior identified in this specification. In particular, the security processing requirements imposed by this specification must be consistent across all BPSEC-BPAs in a network.
- o The behavior of an OSB when dealing with fragmentation must be specified and MUST NOT lead to ambiguous processing states. In particular, an OSB definition should address how to receive and process an OSB in a bundle fragment that may or may not also contain its security target. An OSB definition should also address whether an OSB may be added to a bundle marked as a fragment.

Additionally, policy considerations for the management, monitoring, and configuration associated with blocks SHOULD be included in any OSB definition.

NOTE: The burden of showing compliance with processing rules is placed upon the standards defining new security blocks and the identification of such blocks shall not, alone, require maintenance of this specification.

11. IANA Considerations

A registry of cipher suite identifiers will be required.

11.1. Bundle Block Types

This specification allocates two block types from the existing "Bundle Block Types" registry defined in [RFC6255] .

Additional Entries for the Bundle Block-Type Codes Registry:

Value	Description	Reference
TBD	Block Integrity Block	This document
TBD	Block Confidentiality Block	This document

Table 1

12. References

12.1. Normative References

- [BPBIS] Burleigh, S., Fall, K., and E. Birrane, "Bundle Protocol", draft-ietf-dtn-bpbis-06 (work in progress), July 2016.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3552] Rescorla, E. and B. Korver, "Guidelines for Writing RFC Text on Security Considerations", BCP 72, RFC 3552, DOI 10.17487/RFC3552, July 2003, <<https://www.rfc-editor.org/info/rfc3552>>.
- [RFC6255] Blanchet, M., "Delay-Tolerant Networking Bundle Protocol IANA Registries", RFC 6255, May 2011.

12.2. Informative References

- [COSE] Schaad, J., "CBOR Object Signing and Encryption (COSE)", draft-ietf-cose-msg-24 (work in progress), November 2016.
- [RFC4838] Cerf, V., Burleigh, S., Hooke, A., Torgerson, L., Durst, R., Scott, K., Fall, K., and H. Weiss, "Delay-Tolerant Networking Architecture", RFC 4838, April 2007.
- [RFC6257] Symington, S., Farrell, S., Weiss, H., and P. Lovell, "Bundle Security Protocol Specification", RFC 6257, May 2011.
- [SBSP] Birrane, E., "Streamlined Bundle Security Protocol", draft-birrane-dtn-sbsp-01 (work in progress), October 2015.

Appendix A. Acknowledgements

The following participants contributed technical material, use cases, and useful thoughts on the overall approach to this security specification: Scott Burleigh of the Jet Propulsion Laboratory, Amy Alford and Angela Hennessy of the Laboratory for Telecommunications Sciences, and Angela Dalton and Cherita Corbett of the Johns Hopkins University Applied Physics Laboratory.

Authors' Addresses

Edward J. Birrane, III
The Johns Hopkins University Applied Physics Laboratory
11100 Johns Hopkins Rd.
Laurel, MD 20723
US

Phone: +1 443 778 7423
Email: Edward.Birrane@jhuapl.edu

Kenneth McKeever
The Johns Hopkins University Applied Physics Laboratory
11100 Johns Hopkins Rd.
Laurel, MD 20723
US

Phone: +1 443 778 2237
Email: Ken.McKeever@jhuapl.edu

Delay-Tolerant Networking
Internet-Draft
Intended status: Standards Track
Expires: March 12, 2020

E. Birrane
K. McKeever
JHU/APL
September 9, 2019

Bundle Protocol Security Specification
draft-ietf-dtn-bpsec-11

Abstract

This document defines a security protocol providing end to end data integrity and confidentiality services for the Bundle Protocol.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on March 12, 2020.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

- 1. Introduction 3
 - 1.1. Supported Security Services 3
 - 1.2. Specification Scope 4
 - 1.3. Related Documents 5
 - 1.4. Terminology 6
- 2. Design Decisions 7
 - 2.1. Block-Level Granularity 7
 - 2.2. Multiple Security Sources 8
 - 2.3. Mixed Security Policy 8
 - 2.4. User-Defined Security Contexts 9
 - 2.5. Deterministic Processing 9
- 3. Security Blocks 9
 - 3.1. Block Definitions 9
 - 3.2. Uniqueness 10
 - 3.3. Target Multiplicity 11
 - 3.4. Target Identification 11
 - 3.5. Block Representation 12
 - 3.6. Abstract Security Block 12
 - 3.7. Block Integrity Block 15
 - 3.8. Block Confidentiality Block 16
 - 3.9. Block Interactions 17
 - 3.10. Parameter and Result Identification 18
 - 3.11. BSP Block Examples 19
 - 3.11.1. Example 1: Constructing a Bundle with Security . . . 19
 - 3.11.2. Example 2: Adding More Security At A New Node . . . 20
- 4. Canonical Forms 22
- 5. Security Processing 22
 - 5.1. Bundles Received from Other Nodes 23
 - 5.1.1. Receiving BCBs 23
 - 5.1.2. Receiving BIBs 24
 - 5.2. Bundle Fragmentation and Reassembly 25
- 6. Key Management 25
- 7. Security Policy Considerations 25
- 8. Security Considerations 27
 - 8.1. Attacker Capabilities and Objectives 27
 - 8.2. Attacker Behaviors and BPsec Mitigations 28
 - 8.2.1. Eavesdropping Attacks 28
 - 8.2.2. Modification Attacks 29
 - 8.2.3. Topology Attacks 30
 - 8.2.4. Message Injection 30
- 9. Security Context Considerations 31
 - 9.1. Identification and Configuration 31
 - 9.2. Authorship 31
- 10. Defining Other Security Blocks 33
- 11. IANA Considerations 34
 - 11.1. Bundle Block Types 34

11.2. Security Context Identifiers 34
12. References 35
12.1. Normative References 35
12.2. Informative References 35
Appendix A. Acknowledgements 36
Authors' Addresses 36

1. Introduction

This document defines security features for the Bundle Protocol (BP) [I-D.ietf-dtn-bpbis] and is intended for use in Delay Tolerant Networks (DTNs) to provide end-to-end security services.

The Bundle Protocol specification [I-D.ietf-dtn-bpbis] defines DTN as referring to "a networking architecture providing communications in and/or through highly stressed environments" where "BP may be viewed as sitting at the application layer of some number of constituent networks, forming a store-carry-forward overlay network". The term "stressed" environment refers to multiple challenging conditions including intermittent connectivity, large and/or variable delays, asymmetric data rates, and high bit error rates.

The BP might be deployed such that portions of the network cannot be trusted, posing the usual security challenges related to confidentiality and integrity. However, the stressed nature of the BP operating environment imposes unique conditions where usual transport security mechanisms may not be sufficient. For example, the store-carry-forward nature of the network may require protecting data at rest, preventing unauthorized consumption of critical resources such as storage space, and operating without regular contact with a centralized security oracle (such as a certificate authority).

An end-to-end security service is needed that operates in all of the environments where the BP operates.

1.1. Supported Security Services

BPSec provides end-to-end integrity and confidentiality services for BP bundles, as defined in this section.

Integrity services ensure that changes to target data within a bundle, if any, can be discovered. Data changes may be caused by processing errors, environmental conditions, or intentional manipulation. In the context of BPSec, integrity services apply to plain-text in the bundle.

Confidentiality services ensure that target data is unintelligible to nodes in the DTN, except for authorized nodes possessing special information. This generally means producing cipher-text from plain-text and generating authentication information for that cipher-text. Confidentiality, in this context, applies to the contents of target data and does not extend to hiding the fact that confidentiality exists in the bundle.

NOTE: Hop-by-hop authentication is NOT a supported security service in this specification, for three reasons.

1. The term "hop-by-hop" is ambiguous in a BP overlay, as nodes that are adjacent in the overlay may not be adjacent in physical connectivity. This condition is difficult or impossible to detect and therefore hop-by-hop authentication is difficult or impossible to enforce.
2. Networks in which BPsec may be deployed may have a mixture of security-aware and not-security-aware nodes. Hop-by-hop authentication cannot be deployed in a network if adjacent nodes in the network have different security capabilities.
3. Hop-by-hop authentication is a special case of data integrity and can be achieved with the integrity mechanisms defined in this specification. Therefore, a separate authentication service is not necessary.

1.2. Specification Scope

This document defines the security services provided by the BPsec. This includes the data specification for representing these services as BP extension blocks, and the rules for adding, removing, and processing these blocks at various points during the bundle's traversal of the DTN.

BPsec applies only to those nodes that implement it, known as "security-aware" nodes. There might be other nodes in the DTN that do not implement BPsec. While all nodes in a BP overlay can exchange bundles, BPsec security operations can only happen at BPsec security-aware nodes.

BPsec addresses only the security of data traveling over the DTN, not the underlying DTN itself. Furthermore, while the BPsec protocol can provide security-at-rest in a store-carry-forward network, it does not address threats which share computing resources with the DTN and/or BPsec software implementations. These threats may be malicious software or compromised libraries which intend to intercept data or recover cryptographic material. Here, it is the responsibility of

the BPSec implementer to ensure that any cryptographic material, including shared secret or private keys, is protected against access within both memory and storage devices.

This specification addresses neither the fitness of externally-defined cryptographic methods nor the security of their implementation. Different networking conditions and operational considerations require varying strengths of security mechanism such that mandating a cipher suite in this specification may result in too much security for some networks and too little security in others. It is expected that separate documents will be standardized to define security contexts and cipher suites compatible with BPSec, to include those that should be used to assess interoperability and those fit for operational use in various network scenarios. A sample security context has been defined ([I-D.ietf-dtn-bpsec-interop-sc]) to support interoperability testing and serve as an exemplar for how security contexts should be defined for this specification.

This specification does not address the implementation of security policy and does not provide a security policy for the BPSec. Similar to cipher suites, security policies are based on the nature and capabilities of individual networks and network operational concepts. This specification does provide policy considerations when building a security policy.

With the exception of the Bundle Protocol, this specification does not address how to combine the BPSec security blocks with other protocols, other BP extension blocks, or other best practices to achieve security in any particular network implementation.

1.3. Related Documents

This document is best read and understood within the context of the following other DTN documents:

"Delay-Tolerant Networking Architecture" [RFC4838] defines the architecture for DTNs and identifies certain security assumptions made by existing Internet protocols that are not valid in a DTN.

The Bundle Protocol [I-D.ietf-dtn-bpbis] defines the format and processing of bundles, defines the extension block format used to represent BPSec security blocks, and defines the canonicalization algorithms used by this specification.

The Concise Binary Object Representation (CBOR) format [RFC7049] defines a data format that allows for small code size, fairly small message size, and extensibility without version negotiation. The

block-specific data associated with BPsec security blocks are encoded in this data format.

The Bundle Security Protocol [RFC6257] and Streamlined Bundle Security Protocol [I-D.birrane-dtn-sbsp] documents introduced the concepts of using BP extension blocks for security services in a DTN. The BPsec is a continuation and refinement of these documents.

1.4. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

This section defines terminology either unique to the BPsec or otherwise necessary for understanding the concepts defined in this specification.

- o Bundle Destination - the node which receives a bundle and delivers the payload of the bundle to an application. Also, the Node ID of the BPA receiving the bundle. The bundle destination acts as the security destination for every security target in every security block in every bundle it receives.
- o Bundle Source - the node which originates a bundle. Also, the Node ID of the BPA originating the bundle.
- o Cipher Suite - a set of one or more algorithms providing integrity and confidentiality services. Cipher suites may define necessary parameters but do not provide values for those parameters.
- o Forwarder - any node that transmits a bundle in the DTN. Also, the Node ID of the Bundle Protocol Agent (BPA) that sent the bundle on its most recent hop.
- o Intermediate Receiver, Waypoint, or Next Hop - any node that receives a bundle from a Forwarder that is not the Destination. Also, the Node ID of the BPA at any such node.
- o Path - the ordered sequence of nodes through which a bundle passes on its way from Source to Destination. The path is not necessarily known in advance by the bundle or any BPAs in the DTN.
- o Security Block - a BPsec extension block in a bundle.
- o Security Context - the set of assumptions, algorithms, configurations and policies used to implement security services.

- o Security Destination - a bundle node that processes one or more security blocks in a bundle. Also, the Node ID of that node.
- o Security Operation - the application of a security service to a security target, notated as OP(security service, security target). For example, OP(confidentiality, payload). Every security operation in a bundle MUST be unique, meaning that a security service can only be applied to a security target once in a bundle. A security operation is implemented by a security block.
- o Security Service - the security features supported by this specification: either integrity or confidentiality.
- o Security Source - a bundle node that adds a security block to a bundle. Also, the Node ID of that node.
- o Security Target - the block within a bundle that receives a security-service as part of a security-operation.

2. Design Decisions

The application of security services in a DTN is a complex endeavor that must consider physical properties of the network, policies at each node, and various application security requirements. This section identifies those desirable properties that guide design decisions for this specification and are necessary for understanding the format and behavior of the BPsec protocol.

2.1. Block-Level Granularity

Security services within this specification must allow different blocks within a bundle to have different security services applied to them.

Blocks within a bundle represent different types of information. The primary block contains identification and routing information. The payload block carries application data. Extension blocks carry a variety of data that may augment or annotate the payload, or otherwise provide information necessary for the proper processing of a bundle along a path. Therefore, applying a single level and type of security across an entire bundle fails to recognize that blocks in a bundle represent different types of information with different security needs.

For example, a payload block might be encrypted to protect its contents and an extension block containing summary information related to the payload might be integrity signed but unencrypted to

provide waypoints access to payload-related data without providing access to the payload.

2.2. Multiple Security Sources

A bundle can have multiple security blocks and these blocks can have different security sources. BPsec implementations MUST NOT assume that all blocks in a bundle have the same security operations and/or security sources.

The Bundle Protocol allows extension blocks to be added to a bundle at any time during its existence in the DTN. When a waypoint adds a new extension block to a bundle, that extension block MAY have security services applied to it by that waypoint. Similarly, a waypoint MAY add a security service to an existing extension block, consistent with its security policy.

When a waypoint adds a security service to the bundle, the waypoint is the security source for that service. The security block(s) which represent that service in the bundle may need to record this security source as the bundle destination might need this information for processing.

For example, a bundle source may choose to apply an integrity service to its plain-text payload. Later a waypoint node, representing a gateway to an insecure portion of the DTN, may receive the bundle and choose to apply a confidentiality service. In this case, the integrity security source is the bundle source and the confidentiality security source is the waypoint node.

2.3. Mixed Security Policy

The security policy enforced by nodes in the DTN may differ.

Some waypoints might not be security aware and will not be able to process security blocks. Therefore, security blocks must have their processing flags set such that the block will be treated appropriately by non-security-aware waypoints.

Some waypoints will have security policies that require evaluating security services even if they are not the bundle destination or the final intended destination of the service. For example, a waypoint could choose to verify an integrity service even though the waypoint is not the bundle destination and the integrity service will be needed by other nodes along the bundle's path.

Some waypoints will determine, through policy, that they are the intended recipient of the security service and terminate the security

service in the bundle. For example, a gateway node could determine that, even though it is not the destination of the bundle, it should verify and remove a particular integrity service or attempt to decrypt a confidentiality service, before forwarding the bundle along its path.

Some waypoints could understand security blocks but refuse to process them unless they are the bundle destination.

2.4. User-Defined Security Contexts

A security context is the union of security algorithms (cipher suites), policies associated with the use of those algorithms, and configuration values. Different contexts may specify different algorithms, different policies, or different configuration values used in the implementation of their security services. BPsec must provide a mechanism for users to define their own security contexts.

For example, some users might prefer a SHA2 hash function for integrity whereas other users might prefer a SHA3 hash function. The security services defined in this specification must provide a mechanism for determining what cipher suite, policy, and configuration has been used to populate a security block.

2.5. Deterministic Processing

Whenever a node determines that it must process more than one security block in a received bundle (either because the policy at a waypoint states that it should process security blocks or because the node is the bundle destination) the order in which security blocks are processed must be deterministic. All nodes must impose this same deterministic processing order for all security blocks. This specification provides determinism in the application and evaluation of security services, even when doing so results in a loss of flexibility.

3. Security Blocks

3.1. Block Definitions

This specification defines two types of security block: the Block Integrity Block (BIB) and the Block Confidentiality Block (BCB).

The BIB is used to ensure the integrity of its plain-text security target(s). The integrity information in the BIB MAY be verified by any node along the bundle path from the BIB security source to the bundle destination. Security-aware waypoints add or remove

BIBs from bundles in accordance with their security policy. BIBs are never used to sign the cipher-text provided by a BCB.

The BCB indicates that the security target(s) have been encrypted at the BCB security source in order to protect their content while in transit. The BCB is decrypted by security-aware nodes in the network, up to and including the bundle destination, as a matter of security policy. BCBs additionally provide authentication mechanisms for the cipher-text they generate.

3.2. Uniqueness

Security operations in a bundle MUST be unique; the same security service MUST NOT be applied to a security target more than once in a bundle. Since a security operation is represented as a security block, this limits what security blocks may be added to a bundle: if adding a security block to a bundle would cause some other security block to no longer represent a unique security operation then the new block MUST NOT be added. It is important to note that any cipher-text integrity mechanism supplied by the BCB is considered part of the confidentiality service and, therefore, unique from the plain-text integrity service provided by the BIB.

If multiple security blocks representing the same security operation were allowed in a bundle at the same time, there would exist ambiguity regarding block processing order and the property of deterministic processing blocks would be lost.

Using the notation $OP(\text{service}, \text{target})$, several examples illustrate this uniqueness requirement.

- o Signing the payload twice: The two operations $OP(\text{integrity}, \text{payload})$ and $OP(\text{integrity}, \text{payload})$ are redundant and MUST NOT both be present in the same bundle at the same time.
- o Signing different blocks: The two operations $OP(\text{integrity}, \text{payload})$ and $OP(\text{integrity}, \text{extension_block_1})$ are not redundant and both may be present in the same bundle at the same time. Similarly, the two operations $OP(\text{integrity}, \text{extension_block_1})$ and $OP(\text{integrity}, \text{extension_block_2})$ are also not redundant and may both be present in the bundle at the same time.
- o Different Services on same block: The two operations $OP(\text{integrity}, \text{payload})$ and $OP(\text{confidentiality}, \text{payload})$ are not inherently redundant and may both be present in the bundle at the same time, pursuant to other processing rules in this specification.

3.3. Target Multiplicity

Under special circumstances, a single security block MAY represent multiple security operations as a way of reducing the overall number of security blocks present in a bundle. In these circumstances, reducing the number of security blocks in the bundle reduces the amount of redundant information in the bundle.

A set of security operations can be represented by a single security block when all of the following conditions are true.

- o The security operations apply the same security service. For example, they are all integrity operations or all confidentiality operations.
- o The security context parameters and key information for the security operations are identical.
- o The security source for the security operations is the same. Meaning the set of operations are being added by the same node.
- o No security operations have the same security target, as that would violate the need for security operations to be unique.
- o None of the security operations conflict with security operations already present in the bundle.

When representing multiple security operations in a single security block, the information that is common across all operations is represented once in the security block, and the information which is different (e.g., the security targets) are represented individually.

It is RECOMMENDED that if a node processes any security operation in a security block that it process all security operations in the security block. This allows security sources to assert that the set of security operations in a security block are expected to be processed by the same security destination. However, the determination of whether a node actually is a security destination or not is a matter of the policy of the node itself. In cases where a receiving node determines that it is the security destination of only a subset of the security operations in a security block, the node may choose to only process that subset of security operations.

3.4. Target Identification

A security target is a block in the bundle to which a security service applies. This target must be uniquely and unambiguously identifiable when processing a security block. The definition of the

extension block header from [I-D.ietf-dtn-bpbis] provides a "Block Number" field suitable for this purpose. Therefore, a security target in a security block MUST be represented as the Block Number of the target block.

3.5. Block Representation

Each security block uses the Canonical Bundle Block Format as defined in [I-D.ietf-dtn-bpbis]. That is, each security block is comprised of the following elements:

- o Block Type Code
- o Block Number
- o Block Processing Control Flags
- o CRC Type and CRC Field (if present)
- o Block Data Length
- o Block Type Specific Data Fields

Security-specific information for a security block is captured in the "Block Type Specific Data Fields".

3.6. Abstract Security Block

The structure of the security-specific portions of a security block is identical for both the BIB and BCB Block Types. Therefore, this section defines an Abstract Security Block (ASB) data structure and discusses the definition, processing, and other constraints for using this structure. An ASB is never directly instantiated within a bundle, it is only a mechanism for discussing the common aspects of BIB and BCB security blocks.

The fields of the ASB SHALL be as follows, listed in the order in which they must appear.

Security Targets:

This field identifies the block(s) targeted by the security operation(s) represented by this security block. Each target block is represented by its unique Block Number. This field SHALL be represented by a CBOR array of data items. Each target within this CBOR array SHALL be represented by a CBOR unsigned integer. This array MUST have at least 1 entry and each entry MUST represent the Block Number of a block that

exists in the bundle. There MUST NOT be duplicate entries in this array.

Security Context Id:

This field identifies the security context used to implement the security service represented by this block and applied to each security target. This field SHALL be represented by a CBOR unsigned integer.

Security Context Flags:

This field identifies which optional fields are present in the security block. This field SHALL be represented as a CBOR unsigned integer whose contents shall be interpreted as a bit field. Each bit in this bit field indicates the presence (bit set to 1) or absence (bit set to 0) of optional data in the security block. The association of bits to security block data is defined as follows.

Bit 1 (the least-significant bit, 0x01): Security Context Parameters Present Flag.

Bit 2 (0x02): Security Source Present Flag.

In this field, a value of 1 indicates that the associated security block field MUST be included in the security block. A value of 0 indicates that the associated security block field MUST NOT be in the security block.

Security Source (Optional):

This field identifies the Endpoint that inserted the security block in the bundle. If the security source field is not present then the source MUST be inferred from other information, such as the bundle source, previous hop, or other values defined by security policy. This field SHALL be represented by a CBOR array in accordance with [I-D.ietf-dtn-bpbis] rules for representing Endpoint Identifiers (EIDs).

Security Context Parameters (Optional):

This field captures one or more security context parameters that should be provided to security-aware nodes when processing the security service described by this security block. This field SHALL be represented by a CBOR array. Each entry in this array is a single security context parameter. A single parameter SHALL also be represented as a CBOR array comprising a 2-tuple of the id and value of the parameter, as follows.

- * Parameter Id. This field identifies which parameter is being specified. This field SHALL be represented as a CBOR unsigned integer. Parameter Ids are selected as described in Section 3.10.
- * Parameter Value. This field captures the value associated with this parameter. This field SHALL be represented by the applicable CBOR representation of the parameter, in accordance with Section 3.10.

The logical layout of the parameters array is illustrated in Figure 1.

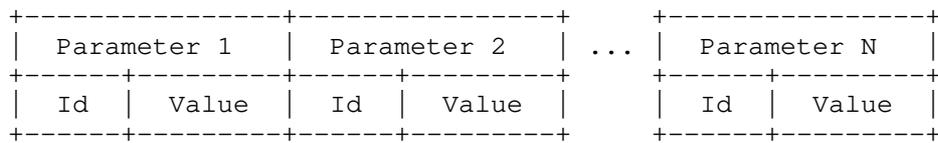


Figure 1: Security Context Parameters

Security Results:

This field captures the results of applying a security service to the security targets of the security block. This field SHALL be represented as a CBOR array of target results. Each entry in this array represents the set of security results for a specific security target. The target results MUST be ordered identically to the Security Targets field of the security block. This means that the first set of target results in this array corresponds to the first entry in the Security Targets field of the security block, and so on. There MUST be one entry in this array for each entry in the Security Targets field of the security block.

The set of security results for a target is also represented as a CBOR array of individual results. An individual result is represented as a 2-tuple of a result id and a result value, defined as follows.

- * Result Id. This field identifies which security result is being specified. Some security results capture the primary output of a cipher suite. Other security results contain additional annotative information from cipher suite processing. This field SHALL be represented as a CBOR unsigned integer. Security result Ids will be as specified in Section 3.10.

- * Result Value. This field captures the value associated with the result. This field SHALL be represented by the applicable CBOR representation of the result value, in accordance with Section 3.10.

The logical layout of the security results array is illustrated in Figure 2. In this figure there are N security targets for this security block. The first security target contains M results and the Nth security target contains K results.

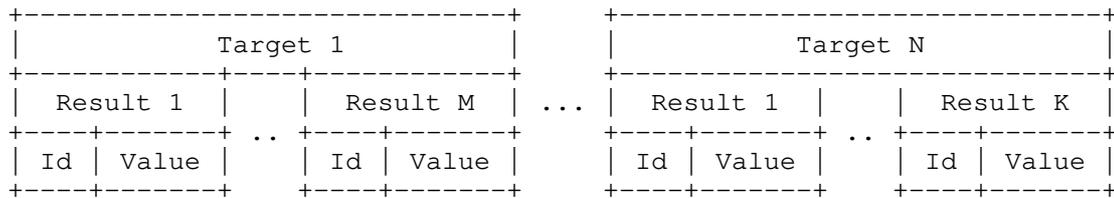


Figure 2: Security Results

3.7. Block Integrity Block

A BIB is a bundle extension block with the following characteristics.

- o The Block Type Code value is as specified in Section 11.1.
- o The Block Type Specific Data Fields follow the structure of the ASB.
- o A security target listed in the Security Targets field MUST NOT reference a security block defined in this specification (e.g., a BIB or a BCB).
- o The Security Context Id MUST utilize an end-to-end authentication cipher or an end-to-end error detection cipher.
- o An EID-reference to the security source MAY be present. If this field is not present, then the security source of the block SHOULD be inferred according to security policy and MAY default to the bundle source. The security source MAY be specified as part of key information described in Section 3.10.

Notes:

- o It is RECOMMENDED that cipher suite designers carefully consider the effect of setting flags that either discard the block or delete the bundle in the event that this block cannot be processed.

- o Since OP(integrity, target) is allowed only once in a bundle per target, it is RECOMMENDED that users wishing to support multiple integrity signatures for the same target define a multi-signature cipher suite.
- o For some cipher suites, (e.g., those using asymmetric keying to produce signatures or those using symmetric keying with a group key), the security information MAY be checked at any hop on the way to the destination that has access to the required keying information, in accordance with Section 3.9.
- o The use of a generally available key is RECOMMENDED if custodial transfer is employed and all nodes SHOULD verify the bundle before accepting custody.

3.8. Block Confidentiality Block

A BCB is a bundle extension block with the following characteristics.

The Block Type Code value is as specified in Section 11.1.

The Block Processing Control flags value can be set to whatever values are required by local policy, except that this block MUST have the "replicate in every fragment" flag set if the target of the BCB is the Payload Block. Having that BCB in each fragment indicates to a receiving node that the payload portion of each fragment represents cipher-text.

The Block Type Specific Data Fields follow the structure of the ASB.

A security target listed in the Security Targets field can reference the payload block, a non-security extension block, or a BIB. A BCB MUST NOT include another BCB as a security target. A BCB MUST NOT target the primary block.

The Security Context Id MUST utilize a confidentiality cipher that provides authenticated encryption with associated data (AEAD).

Additional information created by a cipher suite (such as additional authenticated data) can be placed either in a security result field or in the generated cipher-text. The determination of where to place these data is a function of the cipher suite and security context used.

An EID-reference to the security source MAY be present. If this field is not present, then the security source of the block SHOULD be inferred according to security policy and MAY default to the

bundle source. The security source MAY be specified as part of the key information described in Section 3.10.

The BCB modifies the contents of its security target(s). When a BCB is applied, the security target body data are encrypted "in-place". Following encryption, the security target Block Type Specific Data field contains cipher-text, not plain-text. Other block fields remain unmodified, with the exception of the Block Data Length field, which MUST be updated to reflect the new length of the Block Type Specific Data field.

Notes:

- o It is RECOMMENDED that cipher suite designers carefully consider the effect of setting flags that either discard the block or delete the bundle in the event that this block cannot be processed.
- o The BCB block processing control flags can be set independently from the processing control flags of the security target(s). The setting of such flags SHOULD be an implementation/policy decision for the encrypting node.

3.9. Block Interactions

The security block types defined in this specification are designed to be as independent as possible. However, there are some cases where security blocks may share a security target creating processing dependencies.

If a security target of a BCB is also a security target of a BIB, an undesirable condition occurs where a security aware waypoint would be unable to validate the BIB because one of its security target's contents have been encrypted by a BCB. To address this situation the following processing rules MUST be followed.

- o When adding a BCB to a bundle, if some (or all) of the security targets of the BCB also match all of the security targets of an existing BIB, then the existing BIB MUST also be encrypted. This can be accomplished by either adding a new BCB that targets the existing BIB, or by adding the BIB to the list of security targets for the BCB. Deciding which way to represent this situation is a matter of security policy.
- o When adding a BCB to a bundle, if some (or all) of the security targets of the BCB match some (but not all) of the security targets of a BIB then that BIB MUST be altered in the following way. Any security results in the BIB associated with the BCB

security targets MUST be removed from the BIB and placed in a new BIB. This newly created BIB MUST then be encrypted. The encryption of the new BIB can be accomplished by either adding a new BCB that targets the new BIB, or by adding the new BIB to the list of security targets for the BCB. Deciding which way to represent this situation is a matter of security policy.

- o A BIB MUST NOT be added for a security target that is already the security target of a BCB. In this instance, the BCB is already providing authentication and integrity of the security target and the BIB would be redundant, insecure, and cause ambiguity in block processing order.
- o A BIB integrity value MUST NOT be evaluated if the BIB is the security target of an existing BCB. In this case, the BIB data is encrypted.
- o A BIB integrity value MUST NOT be evaluated if the security target of the BIB is also the security target of a BCB. In such a case, the security target data contains cipher-text as it has been encrypted.
- o As mentioned in Section 3.7, a BIB MUST NOT have a BCB as its security target.

These restrictions on block interactions impose a necessary ordering when applying security operations within a bundle. Specifically, for a given security target, BIBs MUST be added before BCBs. This ordering MUST be preserved in cases where the current BPA is adding all of the security blocks for the bundle or whether the BPA is a waypoint adding new security blocks to a bundle that already contains security blocks.

NOTE: Since any cipher suite used with a BCB MUST be an AEAD cipher suite, it is inefficient and possibly insecure for a single security source to add both a BIB and a BCB for the same security target. In cases where a security source wishes to calculate both a plain-text integrity mechanism and encrypt a security target, a BCB with a cipher suite that generates such signatures as additional security results SHOULD be used instead.

3.10. Parameter and Result Identification

Each security context MUST define its own context parameters and security results. Each defined parameter and result is represented as the tuple of an identifier and a value. Identifiers are always represented as a CBOR unsigned integer. The CBOR encoding of values is as defined by the security context specification.

Identifiers MUST be unique for a given security context but do not need to be unique amongst all security contexts.

3.11. BSP Block Examples

This section provides two examples of BPsec blocks applied to a bundle. In the first example, a single node adds several security operations to a bundle. In the second example, a waypoint node received the bundle created in the first example and adds additional security operations. In both examples, the first column represents blocks within a bundle and the second column represents the Block Number for the block, using the terminology B1...Bn for the purpose of illustration.

3.11.1. Example 1: Constructing a Bundle with Security

In this example a bundle has four non-security-related blocks: the primary block (B1), two extension blocks (B4,B5), and a payload block (B6). The bundle source wishes to provide an integrity signature of the plain-text associated with the primary block, one of the extension blocks, and the payload. The resultant bundle is illustrated in Figure 3 and the security actions are described below.

Block in Bundle	ID
Primary Block	B1
BIB OP(integrity, targets=B1, B5, B6)	B2
BCB OP(confidentiality, target=B4)	B3
Extension Block (encrypted)	B4
Extension Block	B5
Payload Block	B6

Figure 3: Security at Bundle Creation

The following security actions were applied to this bundle at its time of creation.

- o An integrity signature applied to the canonicalized primary block (B1), the second extension block (B5) and the payload block (B6). This is accomplished by a single BIB (B2) with multiple targets.

A single BIB is used in this case because all three targets share a security source, security context, and security context parameters. Had this not been the case, multiple BIBs could have been added instead.

- o Confidentiality for the first extension block (B4). This is accomplished by a BCB (B3). Once applied, the contents of extension block B4 are encrypted. The BCB MUST hold an authentication signature for the cipher-text either in the cipher-text that now populated the first extension block or as a security result in the BCB itself, depending on which cipher suite is used to form the BCB. A plain-text integrity signature may also exist as a security result in the BCB if one is provided by the selected confidentiality cipher suite.

3.11.2. Example 2: Adding More Security At A New Node

Consider that the bundle as it is illustrated in Figure 3 is now received by a waypoint node that wishes to encrypt the first extension block and the bundle payload. The waypoint security policy is to allow existing BIBs for these blocks to persist, as they may be required as part of the security policy at the bundle destination.

The resultant bundle is illustrated in Figure 4 and the security actions are described below. Note that block IDs provided here are ordered solely for the purpose of this example and not meant to impose an ordering for block creation. The ordering of blocks added to a bundle MUST always be in compliance with [I-D.ietf-dtn-bpbis].

Block in Bundle	ID
Primary Block	B1
BIB OP(integrity, targets=B1)	B2
BIB (encrypted) OP(integrity, targets=B5, B6)	B7
BCB OP(confidentiality, target=B4,B6,B7)	B8
BCB OP(confidentiality, target=B4)	B3
Extension Block (encrypted)	B4
Extension Block (encrypted)	B5
Payload Block (encrypted)	B6

Figure 4: Security At Bundle Forwarding

The following security actions were applied to this bundle prior to its forwarding from the waypoint node.

- o Since the waypoint node wishes to encrypt blocks B5 and B6, it MUST also encrypt the BIBs providing plain-text integrity over those blocks. However, BIB B2 could not be encrypted in its entirety because it also held a signature for the primary block (B1). Therefore, a new BIB (B7) is created and security results associated with B5 and B6 are moved out of BIB B2 and into BIB B7.
- o Now that there is no longer confusion of which plain-text integrity signatures must be encrypted, a BCB is added to the bundle with the security targets being the second extension block (B5) and the payload (B6) as well as the newly created BIB holding their plain-text integrity signatures (B7). A single new BCB is used in this case because all three targets share a security source, security context, and security context parameters. Had this not been the case, multiple BCBs could have been added instead.

4. Canonical Forms

Security services require consistency and determinism in how information is presented to cipher suites at the security source and at a receiving node. For example, integrity services require that the same target information (e.g., the same bits in the same order) is provided to the cipher suite when generating an original signature and when generating a comparison signature. Canonicalization algorithms are used to construct a stable, end-to-end bit representation of a target block.

Canonical forms are not transmitted, they are used to generate input to a cipher suite for security processing at a security-aware node.

The canonicalization of the primary block is as specified in [I-D.ietf-dtn-bpbis].

All non-primary blocks share the same block structure and are canonicalized as specified in [I-D.ietf-dtn-bpbis] with the following exceptions.

- o If the service being applied is a confidentiality service, then the Block Type Code, Block Number, Block Processing Control Flags, CRC Type and CRC Field (if present), and Block Data Length fields MUST NOT be included in the canonicalization. Confidentiality services are used solely to convert the Block Type Specific Data Fields from plain-text to cipher-text.
- o Reserved flags MUST NOT be included in any canonicalization as it is not known if those flags will change in transit.

These canonicalization algorithms assume that Endpoint IDs do not change from the time at which a security source adds a security block to a bundle and the time at which a node processes that security block.

Cipher suites and security contexts MAY define their own canonicalization algorithms and require the use of those algorithms over the ones provided in this specification. In the event of conflicting canonicalization algorithms, those algorithms take precedence over this specification.

5. Security Processing

This section describes the security aspects of bundle processing.

5.1. Bundles Received from Other Nodes

Security blocks must be processed in a specific order when received by a security-aware node. The processing order is as follows.

- o When BIBs and BCBs share a security target, BCBs MUST be evaluated first and BIBs second.

5.1.1. Receiving BCBs

If a received bundle contains a BCB, the receiving node MUST determine whether it is the security destination for any of the security operations in the BCB. If so, the node MUST process those operations and remove any operation-specific information from the BCB prior to delivering data to an application at the node or forwarding the bundle. If processing a security operation fails, the target SHALL be processed according to the security policy. A bundle status report indicating the failure MAY be generated. When all security operations for a BCB have been removed from the BCB, the BCB MUST be removed from the bundle.

If the receiving node is the destination of the bundle, the node MUST decrypt any BCBs remaining in the bundle. If the receiving node is not the destination of the bundle, the node MUST process the BCB if directed to do so as a matter of security policy.

If the security policy of a security-aware node specifies that a bundle should have applied confidentiality to a specific security target and no such BCB is present in the bundle, then the node MUST process this security target in accordance with the security policy. This may involve removing the security target from the bundle. If the removed security target is the payload block, the bundle MUST be discarded.

If an encrypted payload block cannot be decrypted (i.e., the cipher-text cannot be authenticated), then the bundle MUST be discarded and processed no further. If an encrypted security target other than the payload block cannot be decrypted then the associated security target and all security blocks associated with that target MUST be discarded and processed no further. In both cases, requested status reports (see [I-D.ietf-dtn-bpbis]) MAY be generated to reflect bundle or block deletion.

When a BCB is decrypted, the recovered plain-text MUST replace the cipher-text in the security target Block Type Specific Data Fields. If the Block Data Length field was modified at the time of encryption it MUST be updated to reflect the decrypted block length.

If a BCB contains multiple security operations, each operation processed by the node MUST be treated as if the security operation has been represented by a single BCB with a single security operation.

5.1.2. Receiving BIBs

If a received bundle contains a BIB, the receiving node MUST determine whether it is the security destination for any of the security operations in the BIB. If so, the node MUST process those operations and remove any operation-specific information from the BIB prior to delivering data to an application at the node or forwarding the bundle. If processing a security operation fails, the target SHALL be processed according to the security policy. A bundle status report indicating the failure MAY be generated. When all security operations for a BIB have been removed from the BIB, the BIB MUST be removed from the bundle.

A BIB MUST NOT be processed if the security target of the BIB is also the security target of a BCB in the bundle. Given the order of operations mandated by this specification, when both a BIB and a BCB share a security target, it means that the security target must have been encrypted after it was integrity signed and, therefore, the BIB cannot be verified until the security target has been decrypted by processing the BCB.

If the security policy of a security-aware node specifies that a bundle should have applied integrity to a specific security target and no such BIB is present in the bundle, then the node MUST process this security target in accordance with the security policy. This may involve removing the security target from the bundle. If the removed security target is the payload or primary block, the bundle MAY be discarded. This action can occur at any node that has the ability to verify an integrity signature, not just the bundle destination.

If a receiving node is not the security destination of a security operation in a BIB it MAY attempt to verify the security operation anyway to prevent forwarding corrupt data. If the verification fails, the node SHALL process the security target in accordance to local security policy. It is RECOMMENDED that if a payload integrity check fails at a waypoint that it is processed in the same way as if the check fails at the bundle destination. If the check passes, the node MUST NOT remove the security operation from the BIB prior to forwarding.

If a BIB contains multiple security operations, each operation processed by the node MUST be treated as if the security operation

has been represented by a single BIB with a single security operation.

5.2. Bundle Fragmentation and Reassembly

If it is necessary for a node to fragment a bundle payload, and security services have been applied to that bundle, the fragmentation rules described in [I-D.ietf-dtn-bpbis] MUST be followed. As defined there and summarized here for completeness, only the payload block can be fragmented; security blocks, like all extension blocks, can never be fragmented.

Due to the complexity of payload block fragmentation, including the possibility of fragmenting payload block fragments, integrity and confidentiality operations are not to be applied to a bundle representing a fragment. Specifically, a BCB or BIB MUST NOT be added to a bundle if the "Bundle is a Fragment" flag is set in the Bundle Processing Control Flags field.

Security processing in the presence of payload block fragmentation may be handled by other mechanisms outside of the BPsec protocol or by applying BPsec blocks in coordination with an encapsulation mechanism.

6. Key Management

There exist a myriad of ways to establish, communicate, and otherwise manage key information in a DTN. Certain DTN deployments might follow established protocols for key management whereas other DTN deployments might require new and novel approaches. BPsec assumes that key management is handled as a separate part of network management and this specification neither defines nor requires a specific key management strategy.

7. Security Policy Considerations

When implementing BPsec, several policy decisions must be considered. This section describes key policies that affect the generation, forwarding, and receipt of bundles that are secured using this specification. No single set of policy decisions is envisioned to work for all secure DTN deployments.

- o If a bundle is received that contains more than one security operation, in violation of BPsec, then the BPA must determine how to handle this bundle. The bundle may be discarded, the block affected by the security operation may be discarded, or one security operation may be favored over another.

- o BPAs in the network must understand what security operations they should apply to bundles. This decision may be based on the source of the bundle, the destination of the bundle, or some other information related to the bundle.
- o If a waypoint has been configured to add a security operation to a bundle, and the received bundle already has the security operation applied, then the receiver must understand what to do. The receiver may discard the bundle, discard the security target and associated BPsec blocks, replace the security operation, or some other action.
- o It is recommended that security operations only be applied to the blocks that absolutely need them. If a BPA were to apply security operations such as integrity or confidentiality to every block in the bundle, regardless of need, there could be downstream errors processing blocks whose contents must be inspected or changed at every hop along the path.
- o It is recommended that BCBs be allowed to alter the size of extension blocks and the payload block. However, care must be taken to ensure that changing the size of the payload block while the bundle is in transit do not negatively affect bundle processing (e.g., calculating storage needs, scheduling transmission times, caching block byte offsets).
- o Adding a BIB to a security target that has already been encrypted by a BCB is not allowed. If this condition is likely to be encountered, there are (at least) three possible policies that could handle this situation.
 1. At the time of encryption, a plain-text integrity signature may be generated and added to the BCB for the security target as additional information in the security result field.
 2. The encrypted block may be replicated as a new block and integrity signed.
 3. An encapsulation scheme may be applied to encapsulate the security target (or the entire bundle) such that the encapsulating structure is, itself, no longer the security target of a BCB and may therefore be the security target of a BIB.
- o It is recommended that security policy address whether cipher suites whose cipher-text is larger (or smaller) than the initial plain-text are permitted and, if so, for what types of blocks. Changing the size of a block may cause processing difficulties for

networks that calculate block offsets into bundles or predict transmission times or storage availability as a function of bundle size. In other cases, changing the size of a payload as part of encryption has no significant impact.

8. Security Considerations

Given the nature of DTN applications, it is expected that bundles may traverse a variety of environments and devices which each pose unique security risks and requirements on the implementation of security within BPsec. For these reasons, it is important to introduce key threat models and describe the roles and responsibilities of the BPsec protocol in protecting the confidentiality and integrity of the data against those threats. This section provides additional discussion on security threats that BPsec will face and describes how BPsec security mechanisms operate to mitigate these threats.

The threat model described here is assumed to have a set of capabilities identical to those described by the Internet Threat Model in [RFC3552], but the BPsec threat model is scoped to illustrate threats specific to BPsec operating within DTN environments and therefore focuses on man-in-the-middle (MITM) attackers. In doing so, it is assumed that the DTN (or significant portions of the DTN) are completely under the control of an attacker.

8.1. Attacker Capabilities and Objectives

BPsec was designed to protect against MITM threats which may have access to a bundle during transit from its source, Alice, to its destination, Bob. A MITM node, Mallory, is a non-cooperative node operating on the DTN between Alice and Bob that has the ability to receive bundles, examine bundles, modify bundles, forward bundles, and generate bundles at will in order to compromise the confidentiality or integrity of data within the DTN. For the purposes of this section, any MITM node is assumed to effectively be security-aware even if it does not implement the BPsec protocol. There are three classes of MITM nodes which are differentiated based on their access to cryptographic material:

- o Unprivileged Node: Mallory has not been provisioned within the secure environment and only has access to cryptographic material which has been publicly-shared.
- o Legitimate Node: Mallory is within the secure environment and therefore has access to cryptographic material which has been provisioned to Mallory (i.e., K_M) as well as material which has been publicly-shared.

- o **Privileged Node:** Mallory is a privileged node within the secure environment and therefore has access to cryptographic material which has been provisioned to Mallory, Alice and/or Bob (i.e. K_M , K_A , and/or K_B) as well as material which has been publicly-shared.

If Mallory is operating as a privileged node, this is tantamount to compromise; BPsec does not provide mechanisms to detect or remove Mallory from the DTN or BPsec secure environment. It is up to the BPsec implementer or the underlying cryptographic mechanisms to provide appropriate capabilities if they are needed. It should also be noted that if the implementation of BPsec uses a single set of shared cryptographic material for all nodes, a legitimate node is equivalent to a privileged node because $K_M == K_A == K_B$.

A special case of the legitimate node is when Mallory is either Alice or Bob (i.e., $K_M == K_A$ or $K_M == K_B$). In this case, Mallory is able to impersonate traffic as either Alice or Bob, which means that traffic to and from that node can be decrypted and encrypted, respectively. Additionally, messages may be signed as originating from one of the endpoints.

8.2. Attacker Behaviors and BPsec Mitigations

8.2.1. Eavesdropping Attacks

Once Mallory has received a bundle, she is able to examine the contents of that bundle and attempt to recover any protected data or cryptographic keying material from the blocks contained within. The protection mechanism that BPsec provides against this action is the BCB, which encrypts the contents of its security target, providing confidentiality of the data. Of course, it should be assumed that Mallory is able to attempt offline recovery of encrypted data, so the cryptographic mechanisms selected to protect the data should provide a suitable level of protection.

When evaluating the risk of eavesdropping attacks, it is important to consider the lifetime of bundles on a DTN. Depending on the network, bundles may persist for days or even years. Long-lived bundles imply that the data exists in the network for a longer period of time and, thus, there may be more opportunities to capture those bundles. Additionally, bundles that are long-lived imply that the information stored within them may remain relevant and sensitive for long enough that, once captured, there is sufficient time to crack encryption associated with the bundle. If a bundle does persist on the network for years and the cipher suite used for a BCB provides inadequate protection, Mallory may be able to recover the protected data either

before that bundle reaches its intended destination or before the information in the bundle is no longer considered sensitive.

8.2.2. Modification Attacks

As a node participating in the DTN between Alice and Bob, Mallory will also be able to modify the received bundle, including non-BPsec data such as the primary block, payload blocks, or block processing control flags as defined in [I-D.ietf-dtn-bpbis]. Mallory will be able to undertake activities which include modification of data within the blocks, replacement of blocks, addition of blocks, or removal of blocks. Within BPsec, both the BIB and BCB provide integrity protection mechanisms to detect or prevent data manipulation attempts by Mallory.

The BIB provides that protection to another block which is its security target. The cryptographic mechanisms used to generate the BIB should be strong against collision attacks and Mallory should not have access to the cryptographic material used by the originating node to generate the BIB (e.g., K_A). If both of these conditions are true, Mallory will be unable to modify the security target or the BIB and lead Bob to validate the security target as originating from Alice.

Since BPsec security operations are implemented by placing blocks in a bundle, there is no in-band mechanism for detecting or correcting certain cases where Mallory removes blocks from a bundle. If Mallory removes a BCB, but keeps the security target, the security target remains encrypted and there is a possibility that there may no longer be sufficient information to decrypt the block at its destination. If Mallory removes both a BCB (or BIB) and its security target there is no evidence left in the bundle of the security operation. Similarly, if Mallory removes the BIB but not the security target there is no evidence left in the bundle of the security operation. In each of these cases, the implementation of BPsec must be combined with policy configuration at endpoints in the network which describe the expected and required security operations that must be applied on transmission and are expected to be present on receipt. This or other similar out-of-band information is required to correct for removal of security information in the bundle.

A limitation of the BIB may exist within the implementation of BIB validation at the destination node. If Mallory is a legitimate node within the DTN, the BIB generated by Alice with K_A can be replaced with a new BIB generated with K_M and forwarded to Bob. If Bob is only validating that the BIB was generated by a legitimate user, Bob will acknowledge the message as originating from Mallory instead of Alice. In order to provide verifiable integrity checks, both a BIB

and BCB should be used and the BCB should require an IND-CCA2 encryption scheme. Such an encryption scheme will guard against signature substitution attempts by Mallory. In this case, Alice creates a BIB with the protected data block as the security target and then creates a BCB with both the BIB and protected data block as its security targets.

8.2.3. Topology Attacks

If Mallory is in a MITM position within the DTN, she is able to influence how any bundles that come to her may pass through the network. Upon receiving and processing a bundle that must be routed elsewhere in the network, Mallory has three options as to how to proceed: not forward the bundle, forward the bundle as intended, or forward the bundle to one or more specific nodes within the network.

Attacks that involve re-routing the packets throughout the network are essentially a special case of the modification attacks described in this section where the attacker is modifying fields within the primary block of the bundle. Given that BPsec cannot encrypt the contents of the primary block, alternate methods must be used to prevent this situation. These methods may include requiring BIBs for primary blocks, using encapsulation, or otherwise strategically manipulating primary block data. The specifics of any such mitigation technique are specific to the implementation of the deploying network and outside of the scope of this document.

Furthermore, routing rules and policies may be useful in enforcing particular traffic flows to prevent topology attacks. While these rules and policies may utilize some features provided by BPsec, their definition is beyond the scope of this specification.

8.2.4. Message Injection

Mallory is also able to generate new bundles and transmit them into the DTN at will. These bundles may either be copies or slight modifications of previously-observed bundles (i.e., a replay attack) or entirely new bundles generated based on the Bundle Protocol, BPsec, or other bundle-related protocols. With these attacks Mallory's objectives may vary, but may be targeting either the bundle protocol or application-layer protocols conveyed by the bundle protocol.

BPsec relies on cipher suite capabilities to prevent replay or forged message attacks. A BCB used with appropriate cryptographic mechanisms (e.g., a counter-based cipher mode) may provide replay protection under certain circumstances. Alternatively, application data itself may be augmented to include mechanisms to assert data

uniqueness and then protected with a BIB, a BCB, or both along with other block data. In such a case, the receiving node would be able to validate the uniqueness of the data.

9. Security Context Considerations

9.1. Identification and Configuration

Security blocks must uniquely define the security context for their services. This context MUST be uniquely identifiable and MAY use parameters for customization. Where policy and configuration decisions can be captured as parameters, the security context identifier may identify a cipher suite. In cases where the same cipher suites are used with differing predetermined configurations and policies, users can define multiple security contexts.

Network operators must determine the number, type, and configuration of security contexts in a system. Networks with rapidly changing configurations may define relatively few security contexts with each context customized with multiple parameters. For networks with more stability, or an increased need for confidentiality, a larger number of contexts can be defined with each context supporting few, if any, parameters.

Security Context Examples

Context Id	Parameters	Definition
1	Key, IV	AES-GCM-256 cipher suite with provided ephemeral key and initialization vector.
2	IV	AES-GCM-256 cipher suite with predetermined key and predetermined key rotation policy.
3	Nil	AES-GCM-256 cipher suite with all info predetermined.

Table 1

9.2. Authorship

Cipher suite developers or implementers should consider the diverse performance and conditions of networks on which the Bundle Protocol (and therefore BPSec) will operate. Specifically, the delay and capacity of delay-tolerant networks can vary substantially. Cipher suite developers should consider these conditions to better describe

the conditions when those suites will operate or exhibit vulnerability, and selection of these suites for implementation should be made with consideration to the reality. There are key differences that may limit the opportunity to leverage existing cipher suites and technologies that have been developed for use in traditional, more reliable networks:

- o Data Lifetime: Depending on the application environment, bundles may persist on the network for extended periods of time, perhaps even years. Cryptographic algorithms should be selected to ensure protection of data against attacks for a length of time reasonable for the application.
- o One-Way Traffic: Depending on the application environment, it is possible that only a one-way connection may exist between two endpoints, or if a two-way connection does exist, the round-trip time may be extremely large. This may limit the utility of session key generation mechanisms, such as Diffie-Hellman, as a two-way handshake may not be feasible or reliable.
- o Opportunistic Access: Depending on the application environment, a given endpoint may not be guaranteed to be accessible within a certain amount of time. This may make asymmetric cryptographic architectures which rely on a key distribution center or other trust center impractical under certain conditions.

When developing new security contexts for use with BPsec, the following information SHOULD be considered for inclusion in these specifications.

- o Security Context Parameters. Security contexts MUST define their parameter Ids, the data types of those parameters, and their CBOR encoding.
- o Security Results. Security contexts MUST define their security result Ids, the data types of those results, and their CBOR encoding.
- o New Canonicalizations. Security contexts may define new canonicalization algorithms as necessary.
- o Cipher-Text Size. Security contexts MUST state whether their associated cipher suites generate cipher-text (to include any authentication information) that is of a different size than the input plain-text.

If a security context does not wish to alter the size of the plain-text, it should consider defining the following policy.

- * Place overflow bytes, authentication signatures, and any additional authenticated data in security result fields rather than in the cipher-text itself.
- * Pad the cipher-text in cases where the cipher-text is smaller than the plain-text.

10. Defining Other Security Blocks

Other security blocks (OSBs) may be defined and used in addition to the security blocks identified in this specification. Both the usage of BIB, BCB, and any future OSBs can co-exist within a bundle and can be considered in conformance with BPsec if each of the following requirements are met by any future identified security blocks.

- o Other security blocks (OSBs) MUST NOT reuse any enumerations identified in this specification, to include the block type codes for BIB and BCB.
- o An OSB definition MUST state whether it can be the target of a BIB or a BCB. The definition MUST also state whether the OSB can target a BIB or a BCB.
- o An OSB definition MUST provide a deterministic processing order in the event that a bundle is received containing BIBs, BCBs, and OSBs. This processing order MUST NOT alter the BIB and BCB processing orders identified in this specification.
- o An OSB definition MUST provide a canonicalization algorithm if the default non-primary-block canonicalization algorithm cannot be used to generate a deterministic input for a cipher suite. This requirement can be waived if the OSB is defined so as to never be the security target of a BIB or a BCB.
- o An OSB definition MUST NOT require any behavior of a BPSEC-BPA that is in conflict with the behavior identified in this specification. In particular, the security processing requirements imposed by this specification must be consistent across all BPSEC-BPAs in a network.
- o The behavior of an OSB when dealing with fragmentation must be specified and MUST NOT lead to ambiguous processing states. In particular, an OSB definition should address how to receive and process an OSB in a bundle fragment that may or may not also contain its security target. An OSB definition should also address whether an OSB may be added to a bundle marked as a fragment.

Additionally, policy considerations for the management, monitoring, and configuration associated with blocks SHOULD be included in any OSB definition.

NOTE: The burden of showing compliance with processing rules is placed upon the standards defining new security blocks and the identification of such blocks shall not, alone, require maintenance of this specification.

11. IANA Considerations

This specification includes fields requiring registries managed by IANA.

11.1. Bundle Block Types

This specification allocates two block types from the existing "Bundle Block Types" registry defined in [I-D.ietf-dtn-bpbis].

Additional Entries for the Bundle Block-Type Codes Registry:

Value	Description	Reference
2	Block Integrity Block	This document
3	Block Confidentiality Block	This document

Table 2

11.2. Security Context Identifiers

BPsec has a Security Context Identifier field () for which IANA is requested to create and maintain a new registry named "BPsec Security Context Identifiers". Initial values for this registry are given below.

The registration policy for this registry is: Specification Required.

The value range is: unsigned 16-bit integer.

BPSec Security Context Identifier Registry

Value	Description	Reference
0	Reserved	This document
1	BIB-HMAC256-SHA256	[I-D.ietf-dtn-bpsec-interop-sc]
2	BCB-AES-GCM-256	[I-D.ietf-dtn-bpsec-interop-sc]

Table 3

12. References

12.1. Normative References

- [I-D.ietf-dtn-bpbis] Burleigh, S., Fall, K., and E. Birrane, "Bundle Protocol Version 7", draft-ietf-dtn-bpbis-14 (work in progress), August 2019.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC3552] Rescorla, E. and B. Korver, "Guidelines for Writing RFC Text on Security Considerations", BCP 72, RFC 3552, DOI 10.17487/RFC3552, July 2003, <<https://www.rfc-editor.org/info/rfc3552>>.
- [RFC7049] Bormann, C. and P. Hoffman, "Concise Binary Object Representation (CBOR)", RFC 7049, DOI 10.17487/RFC7049, October 2013, <<https://www.rfc-editor.org/info/rfc7049>>.

12.2. Informative References

- [I-D.birrane-dtn-sbsp] Birrane, E., Pierce-Mayer, J., and D. Iannicca, "Streamlined Bundle Security Protocol Specification", draft-birrane-dtn-sbsp-01 (work in progress), October 2015.
- [I-D.ietf-dtn-bpsec-interop-sc] Birrane, E., "BPSec Interoperability Security Contexts", draft-ietf-dtn-bpsec-interop-sc-00 (work in progress), March 2019.

[RFC4838] Cerf, V., Burleigh, S., Hooke, A., Torgerson, L., Durst, R., Scott, K., Fall, K., and H. Weiss, "Delay-Tolerant Networking Architecture", RFC 4838, DOI 10.17487/RFC4838, April 2007, <<https://www.rfc-editor.org/info/rfc4838>>.

[RFC6257] Symington, S., Farrell, S., Weiss, H., and P. Lovell, "Bundle Security Protocol Specification", RFC 6257, DOI 10.17487/RFC6257, May 2011, <<https://www.rfc-editor.org/info/rfc6257>>.

Appendix A. Acknowledgements

The following participants contributed technical material, use cases, and useful thoughts on the overall approach to this security specification: Scott Burleigh of the Jet Propulsion Laboratory, Amy Alford and Angela Hennessy of the Laboratory for Telecommunications Sciences, and Angela Dalton and Cherita Corbett of the Johns Hopkins University Applied Physics Laboratory.

Authors' Addresses

Edward J. Birrane, III
The Johns Hopkins University Applied Physics Laboratory
11100 Johns Hopkins Rd.
Laurel, MD 20723
US

Phone: +1 443 778 7423
Email: Edward.Birrane@jhuapl.edu

Kenneth McKeever
The Johns Hopkins University Applied Physics Laboratory
11100 Johns Hopkins Rd.
Laurel, MD 20723
US

Phone: +1 443 778 2237
Email: Ken.McKeever@jhuapl.edu

Delay Tolerant Networking
Internet-Draft
Obsoletes: 7242 (if approved)
Intended status: Standards Track
Expires: February 21, 2020

B. Sipos
RKF Engineering
M. Demmer
UC Berkeley
J. Ott
Aalto University
S. Perreault
August 20, 2019

Delay-Tolerant Networking TCP Convergence Layer Protocol Version 4
draft-ietf-dtn-tcpclv4-13

Abstract

This document describes a revised protocol for the TCP-based convergence layer (TCPCL) for Delay-Tolerant Networking (DTN). The protocol revision is based on implementation issues in the original TCPCL Version 3 of RFC7242 and updates to the Bundle Protocol contents, encodings, and convergence layer requirements in Bundle Protocol Version 7. Specifically, the TCPCLv4 uses CBOR-encoded BPv7 bundles as its service data unit being transported and provides a reliable transport of such bundles. Several new IANA registries are defined for TCPCLv4 which define some behaviors inherited from TCPCLv3 but with updated encodings and/or semantics.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on February 21, 2020.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	3
1.1.	Scope	4
2.	Requirements Language	5
2.1.	Definitions Specific to the TCPCL Protocol	5
3.	General Protocol Description	8
3.1.	Convergence Layer Services	8
3.2.	TCPCL Session Overview	10
3.3.	TCPCL States and Transitions	12
3.4.	Transfer Segmentation Policies	18
3.5.	Example Message Exchange	19
4.	Session Establishment	21
4.1.	TCP Connection	21
4.2.	Contact Header	22
4.3.	Contact Validation and Negotiation	23
4.4.	Session Security	24
4.4.1.	TLS Handshake	24
4.4.2.	TLS Authentication	25
4.4.3.	Example TLS Initiation	26
4.5.	Message Header	27
4.6.	Session Initialization Message (SESS_INIT)	28
4.7.	Session Parameter Negotiation	30
4.8.	Session Extension Items	31
5.	Established Session Operation	32
5.1.	Upkeep and Status Messages	32
5.1.1.	Session Upkeep (KEEPALIVE)	32
5.1.2.	Message Rejection (MSG_REJECT)	33
5.2.	Bundle Transfer	34
5.2.1.	Bundle Transfer ID	35
5.2.2.	Data Transmission (XFER_SEGMENT)	35
5.2.3.	Data Acknowledgments (XFER_ACK)	37
5.2.4.	Transfer Refusal (XFER_REFUSE)	38
5.2.5.	Transfer Extension Items	41
6.	Session Termination	43
6.1.	Session Termination Message (SESS_TERM)	43
6.2.	Idle Session Shutdown	45
7.	Implementation Status	45

8. Security Considerations	46
9. IANA Considerations	47
9.1. Port Number	47
9.2. Protocol Versions	48
9.3. Session Extension Types	48
9.4. Transfer Extension Types	49
9.5. Message Types	50
9.6. XFER_REFUSE Reason Codes	50
9.7. SESS_TERM Reason Codes	51
9.8. MSG_REJECT Reason Codes	52
10. Acknowledgments	53
11. References	53
11.1. Normative References	53
11.2. Informative References	54
Appendix A. Significant changes from RFC7242	55
Authors' Addresses	56

1. Introduction

This document describes the TCP-based convergence-layer protocol for Delay-Tolerant Networking. Delay-Tolerant Networking is an end-to-end architecture providing communications in and/or through highly stressed environments, including those with intermittent connectivity, long and/or variable delays, and high bit error rates. More detailed descriptions of the rationale and capabilities of these networks can be found in "Delay-Tolerant Network Architecture" [RFC4838].

An important goal of the DTN architecture is to accommodate a wide range of networking technologies and environments. The protocol used for DTN communications is the Bundle Protocol Version 7 (BPv7) [I-D.ietf-dtn-bpbis], an application-layer protocol that is used to construct a store-and-forward overlay network. BPv7 requires the services of a "convergence-layer adapter" (CLA) to send and receive bundles using the service of some "native" link, network, or Internet protocol. This document describes one such convergence-layer adapter that uses the well-known Transmission Control Protocol (TCP). This convergence layer is referred to as TCP Convergence Layer Version 4 (TCPCLv4). For the remainder of this document, the abbreviation "BP" without the version suffix refers to BPv7. For the remainder of this document, the abbreviation "TCPCL" without the version suffix refers to TCPCLv4.

The locations of the TCPCL and the BP in the Internet model protocol stack (described in [RFC1122]) are shown in Figure 1. In particular, when BP is using TCP as its bearer with TCPCL as its convergence layer, both BP and TCPCL reside at the application layer of the Internet model.

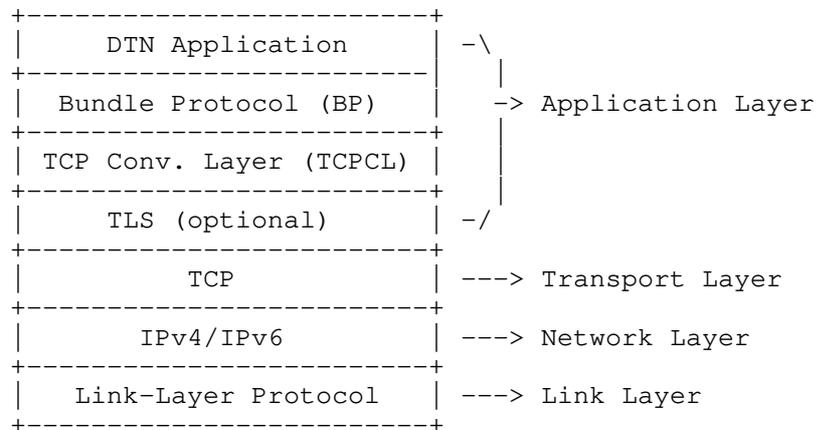


Figure 1: The Locations of the Bundle Protocol and the TCP Convergence-Layer Protocol above the Internet Protocol Stack

1.1. Scope

This document describes the format of the protocol data units passed between entities participating in TCPCL communications. This document does not address:

- o The format of protocol data units of the Bundle Protocol, as those are defined elsewhere in [I-D.ietf-dtn-bpbis]. This includes the concept of bundle fragmentation or bundle encapsulation. The TCPCL transfers bundles as opaque data blocks.
- o Mechanisms for locating or identifying other bundle entities (peers) within a network or across an internet. The mapping of Node ID to potential CL protocol and network address is left to implementation and configuration of the BP Agent and its various potential routing strategies.
- o Logic for routing bundles along a path toward a bundle's endpoint. This CL protocol is involved only in transporting bundles between adjacent nodes in a routing sequence.
- o Policies or mechanisms for assigning X.509 certificates, provisioning or deploying certificates and private keys, or configuring security parameters on an individual BP node or across a network.

Any TCPCL implementation requires a BP agent to perform those above listed functions in order to perform end-to-end bundle delivery.

2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALLNOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2.1. Definitions Specific to the TCPCL Protocol

This section contains definitions specific to the TCPCL protocol.

TCPCL Entity: This is the notional TCPCL application that initiates TCPCL sessions. This design, implementation, configuration, and specific behavior of such an entity is outside of the scope of this document. However, the concept of an entity has utility within the scope of this document as the container and initiator of TCPCL sessions. The relationship between a TCPCL entity and TCPCL sessions is defined as follows:

A TCPCL Entity MAY actively initiate any number of TCPCL Sessions and should do so whenever the entity is the initial transmitter of information to another entity in the network.

A TCPCL Entity MAY support zero or more passive listening elements that listen for connection requests from other TCPCL Entities operating on other entities in the network.

A TCPCL Entity MAY passively initiate any number of TCPCL Sessions from requests received by its passive listening element(s) if the entity uses such elements.

These relationships are illustrated in Figure 2. For most TCPCL behavior within a session, the two entities are symmetric and there is no protocol distinction between them. Some specific behavior, particularly during session establishment, distinguishes between the active entity and the passive entity. For the remainder of this document, the term "entity" without the prefix "TCPCL" refers to a TCPCL entity.

TCP Connection: The term Connection in this specification exclusively refers to a TCP connection and any and all behaviors, sessions, and other states association with that TCP connection.

TCPCL Session: A TCPCL session (as opposed to a TCP connection) is a TCPCL communication relationship between two TCPCL entities. Within a single TCPCL session there are two possible transfer streams; one in each direction, with one stream from each entity

being the outbound stream and the other being the inbound stream. The lifetime of a TCPCL session is bound to the lifetime of an underlying TCP connection. A TCPCL session is terminated when the TCP connection ends, due either to one or both entities actively terminating the TCP connection or due to network errors causing a failure of the TCP connection. For the remainder of this document, the term "session" without the prefix "TCPCL" refers to a TCPCL session.

Session parameters: These are a set of values used to affect the operation of the TCPCL for a given session. The manner in which these parameters are conveyed to the bundle entity and thereby to the TCPCL is implementation dependent. However, the mechanism by which two entities exchange and negotiate the values to be used for a given session is described in Section 4.3.

Transfer Stream: A Transfer stream is a uni-directional user-data path within a TCPCL Session. Messages sent over a transfer stream are serialized, meaning that one set of user data must complete its transmission prior to another set of user data being transmitted over the same transfer stream. Each uni-directional stream has a single sender entity and a single receiver entity.

Transfer: This refers to the procedures and mechanisms for conveyance of an individual bundle from one node to another. Each transfer within TCPCL is identified by a Transfer ID number which is unique only to a single direction within a single Session.

Transfer Segment: A subset of a transfer of user data being communicated over a transfer stream.

Idle Session: A TCPCL session is idle while the only messages being transmitted or received are KEEPALIVE messages.

Live Session: A TCPCL session is live while any messages are being transmitted or received.

Reason Codes: The TCPCL uses numeric codes to encode specific reasons for individual failure/error message types.

The relationship between connections, sessions, and streams is shown in Figure 3.

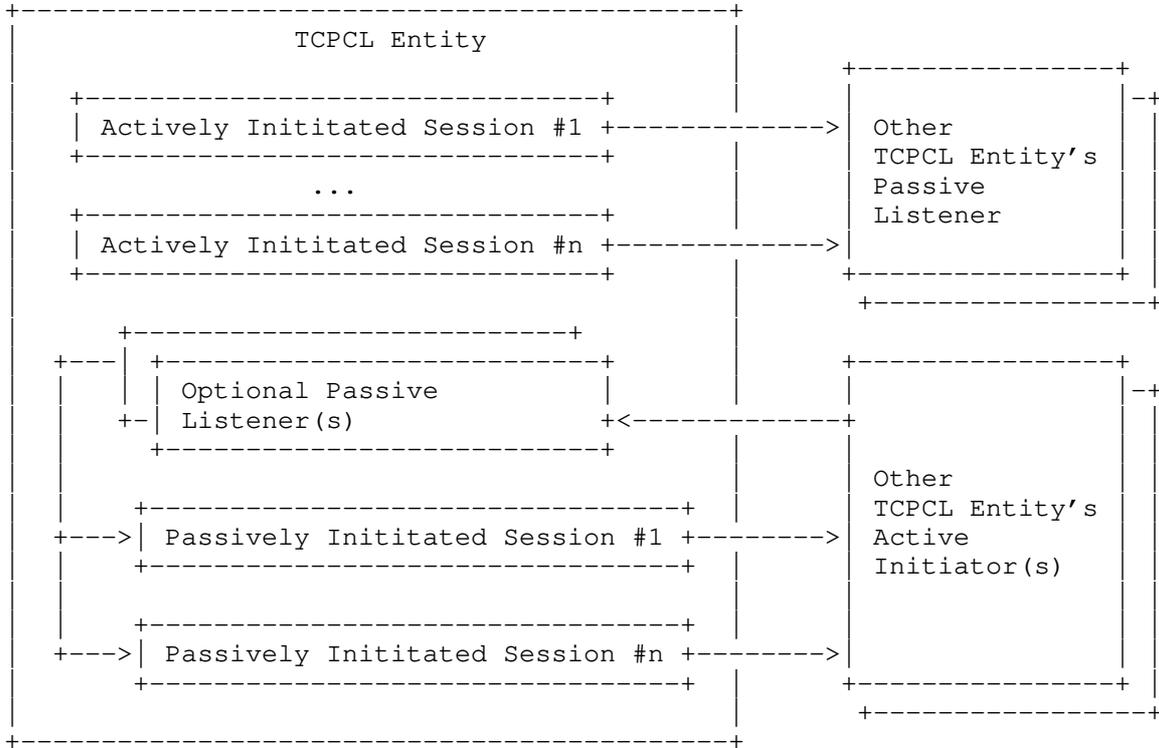


Figure 2: The relationships between TCPCL entities

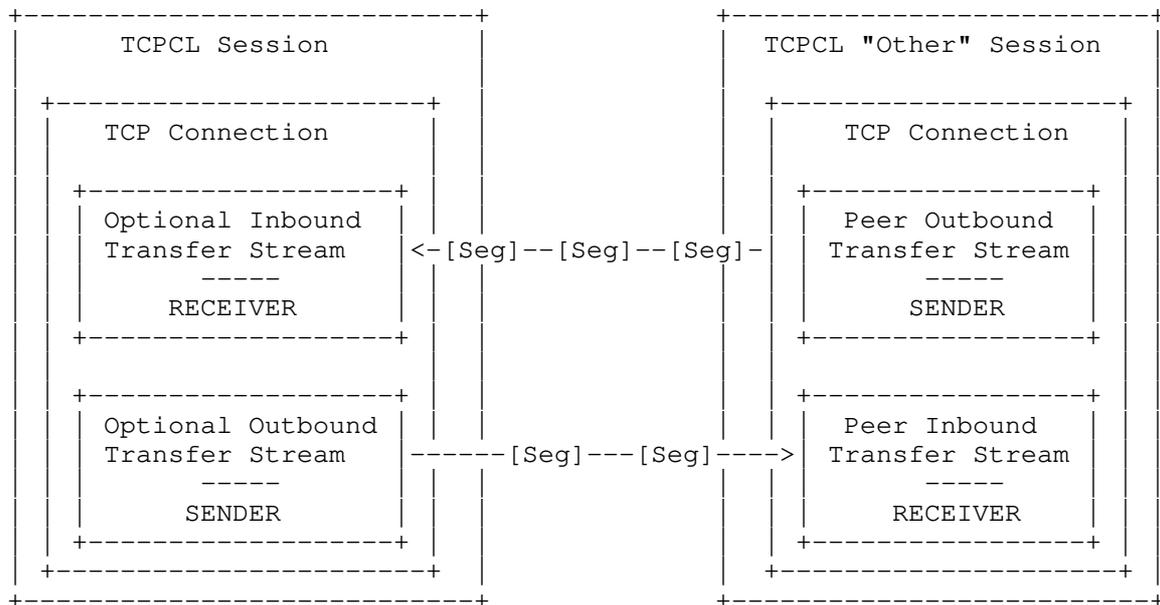


Figure 3: The relationship within a TCPCL Session of its two streams

3. General Protocol Description

The service of this protocol is the transmission of DTN bundles via the Transmission Control Protocol (TCP). This document specifies the encapsulation of bundles, procedures for TCP setup and teardown, and a set of messages and node requirements. The general operation of the protocol is as follows.

3.1. Convergence Layer Services

This version of the TCPCL provides the following services to support the overlaying Bundle Protocol agent. In all cases, this is not an API definition but a logical description of how the CL can interact with the BP agent. Each of these interactions can be associated with any number of additional metadata items as necessary to support the operation of the CL or BP agent.

Attempt Session: The TCPCL allows a BP agent to pre-emptively attempt to establish a TCPCL session with a peer entity. Each session attempt can send a different set of session negotiation parameters as directed by the BP agent.

Terminate Session: The TCPCL allows a BP agent to pre-emptively terminate an established TCPCL session with a peer entity. The terminate request is on a per-session basis.

Session State Changed: The TCPCL supports indication when the session state changes. The top-level session states indicated are:

Connecting: A TCP connection is being established. This state only applies to the active entity.

Contact Negotiating: A TCP connection has been made (as either active or passive entity) and contact negotiation has begun.

Session Negotiating: Contact negotiation has been completed (including possible TLS use) and session negotiation has begun.

Established: The session has been fully established and is ready for its first transfer.

Ending: The entity received a SESS_TERM message and is in the ending state.

Terminated: The session has finished normal termination sequencing.

Failed: The session ended without normal termination sequencing.

Session Idle Changed: The TCPCL supports indication when the live/idle sub-state of the session changes. This occurs only when the top-level session state is "Established". The session transitions from Idle to Live at the start of a transfer in either transfer stream; the session transitions from Live to Idle at the end of a transfer when the other transfer stream does not have an ongoing transfer. Because TCPCL transmits serially over a TCP connection, it suffers from "head of queue blocking" this indication provides information about when a session is available for immediate transfer start.

Begin Transmission: The principal purpose of the TCPCL is to allow a BP agent to transmit bundle data over an established TCPCL session. Transmission request is on a per-session basis, the CL does not necessarily perform any per-session or inter-session queueing. Any queueing of transmissions is the obligation of the BP agent.

Transmission Success: The TCPCL supports positive indication when a bundle has been fully transferred to a peer entity.

Transmission Intermediate Progress: The TCPCL supports positive indication of intermediate progress of transferr to a peer entity. This intermediate progress is at the granularity of each transferred segment.

Transmission Failure: The TCPCL supports positive indication of certain reasons for bundle transmission failure, notably when the peer entity rejects the bundle or when a TCPCL session ends before transferr success. The TCPCL itself does not have a notion of transfer timeout.

Reception Initialized: The TCPCL supports indication to the reciver just before any transmssion data is sent. This corresponds to reception of the XFER_SEGMENT message with the START flag set.

Interrupt Reception: The TCPCL allows a BP agent to interrupt an individual transfer before it has fully completed (successfully or not). Interruption can occur any time after the reception is initialized.

Reception Success: The TCPCL supports positive indication when a bundle has been fully transferred from a peer entity.

Reception Intermediate Progress: The TCPCL supports positive indication of intermediate progress of transfer from the peer entity. This intermediate progress is at the granularity of each transferred segment. Intermediate reception indication allows a BP agent the chance to inspect bundle header contents before the entire bundle is available, and thus supports the "Reception Interruption" capability.

Reception Failure: The TCPCL supports positive indication of certain reasons for reception failure, notably when the local entity rejects an attempted transfer for some local policy reason or when a TCPCL session ends before transfer success. The TCPCL itself does not have a notion of transfer timeout.

3.2. TCPCL Session Overview

First, one node establishes a TCPCL session to the other by initiating a TCP connection in accordance with [RFC0793]. After setup of the TCP connection is complete, an initial contact header is exchanged in both directions to establish a shared TCPCL version and possibly initiate TLS security. Once contact negotiation is complete, TCPCL messaging is available and the session negotiation is used to set parameters of the TCPCL session. One of these parameters is a Node ID of each TCPCL Entity. This is used to assist in routing

and forwarding messages by the BP Agent and is part of the authentication capability provided by TLS.

Once negotiated, the parameters of a TCPCL session cannot change and if there is a desire by either peer to transfer data under different parameters then a new session must be established. This makes CL logic simpler but relies on the assumption that establishing a TCP connection is lightweight enough that TCP connection overhead is negligible compared to TCPCL data sizes.

Once the TCPCL session is established and configured in this way, bundles can be transferred in either direction. Each transfer is performed by an sequence of logical segments of data within XFER_SEGMENT messages. Multiple bundles can be transmitted consecutively in a single direction on a single TCPCL connection. Segments from different bundles are never interleaved. Bundle interleaving can be accomplished by fragmentation at the BP layer or by establishing multiple TCPCL sessions between the same peers. There is no fundamental limit on the number of TCPCL sessions which a single node can establish beyond the limit imposed by the number of available (ephemeral) TCP ports of the passive peer.

A feature of this protocol is for the receiving node to send acknowledgment (XFER_ACK) messages as bundle data segments arrive. The rationale behind these acknowledgments is to enable the sender node to determine how much of the bundle has been received, so that in case the session is interrupted, it can perform reactive fragmentation to avoid re-sending the already transmitted part of the bundle. In addition, there is no explicit flow control on the TCPCL layer.

A TCPCL receiver can interrupt the transmission of a bundle at any point in time by replying with a XFER_REFUSE message, which causes the sender to stop transmission of the associated bundle (if it hasn't already finished transmission) Note: This enables a cross-layer optimization in that it allows a receiver that detects that it already has received a certain bundle to interrupt transmission as early as possible and thus save transmission capacity for other bundles.

For sessions that are idle, a KEEPALIVE message is sent at a negotiated interval. This is used to convey node live-ness information during otherwise message-less time intervals.

A SESS_TERM message is used to start the ending of a TCPCL session (see Section 6.1). During shutdown sequencing, in-progress transfers can be completed but no new transfers can be initiated. A SESS_TERM message can also be used to refuse a session setup by a peer (see

Section 4.3). It is an implementation matter to determine whether or not to close a TCPCL session while there are no transfers queued or in-progress.

Once a session is established, TCPCL is a symmetric protocol between the peers. Both sides can start sending data segments in a session, and one side's bundle transfer does not have to complete before the other side can start sending data segments on its own. Hence, the protocol allows for a bi-directional mode of communication. Note that in the case of concurrent bidirectional transmission, acknowledgment segments MAY be interleaved with data segments.

3.3. TCPCL States and Transitions

The states of a nominal TCPCL session (i.e. without session failures) are indicated in Figure 4.

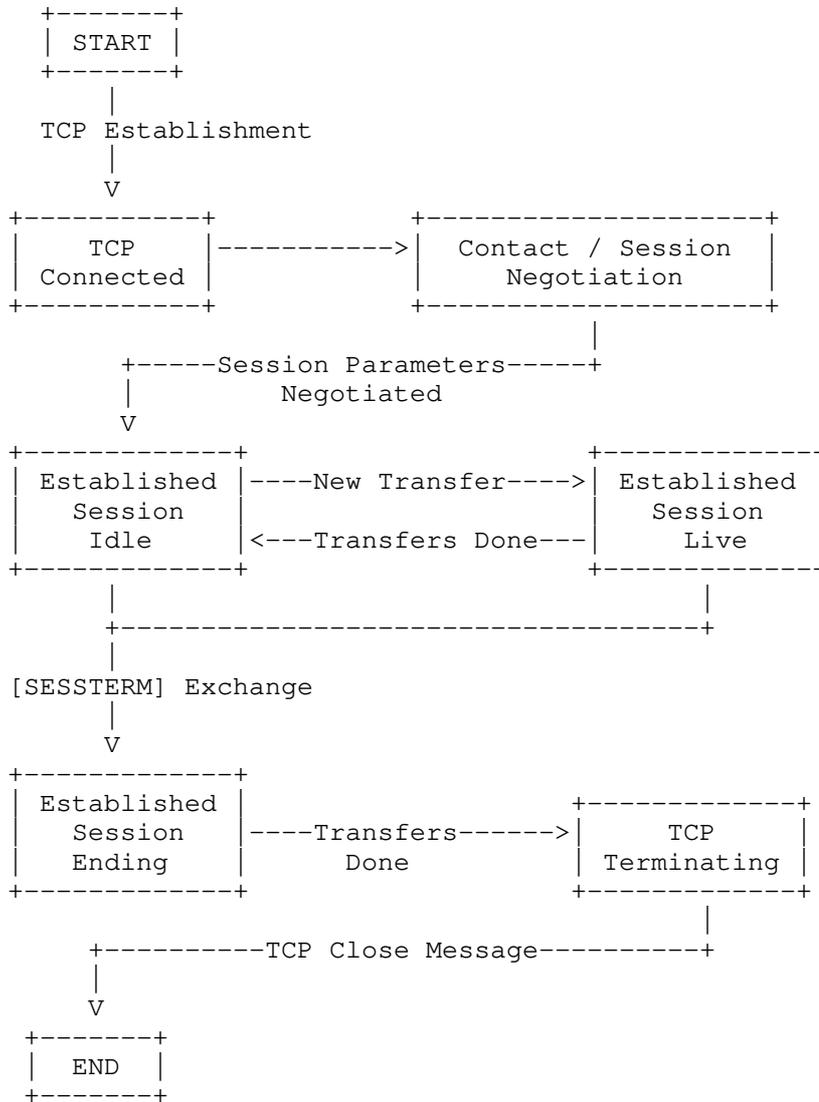


Figure 4: Top-level states of a TCPCL session

Notes on Established Session states:

Session "Live" means transmitting or receiving over a transfer stream.

Session "Idle" means no transmission/reception over a transfer stream.

Session "Ending" means no new transfers will be allowed.

Contact negotiation involves exchanging a Contact Header (CH) in both directions and deriving a negotiated state from the two headers. The contact negotiation sequencing is performed either as the active or passive peer, and is illustrated in Figure 5 and Figure 6 respectively which both share the data validation and analyze final states of the "[PCH]" activity of Figure 7 and the "[TCPCLOSE]" activity which indicates TCP connection close. Successful negotiation results in one of the Session Initiation "[SI]" activities being performed.

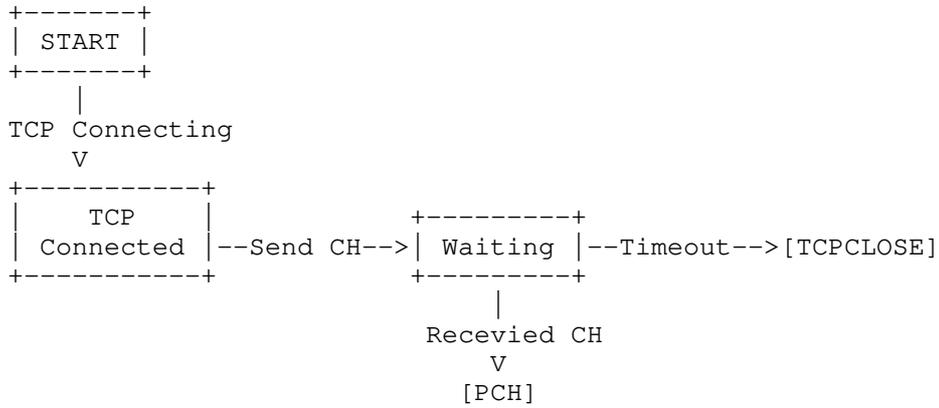


Figure 5: Contact Initiation as Active peer

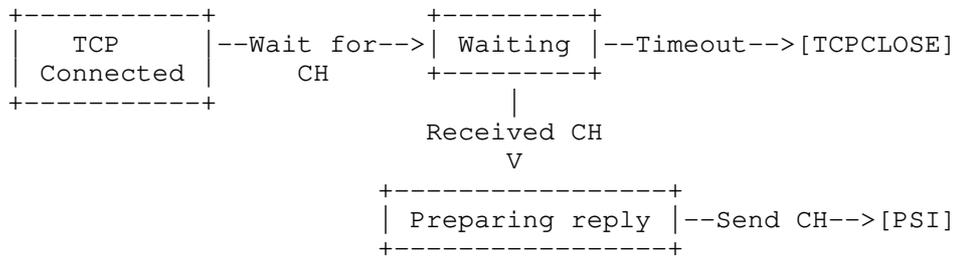


Figure 6: Contact Initiation as Passive peer

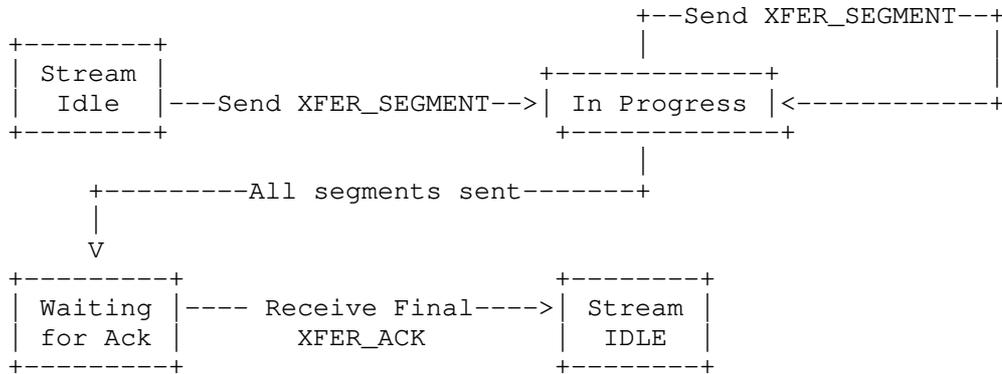


Figure 11: Transfer sender states

Notes on transfer sending:

Pipelining of transfers can occur when the sending entity begins a new transfer while in the "Waiting for Ack" state.

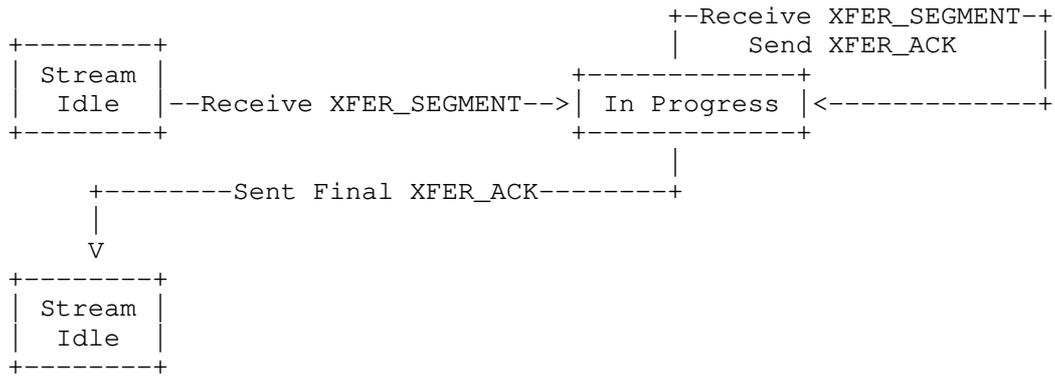


Figure 12: Transfer receiver states

Session termination involves one entity initiating the termination of the session and the other entity acknowledging the termination. For either entity, it is the sending of the SESS_TERM message which transitions the session to the ending substate. While a session is being terminated only in-progress transfers can be completed and no new transfers can be started.

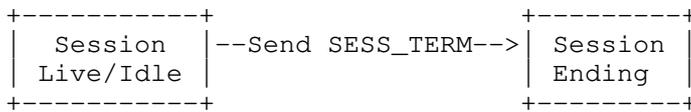


Figure 13: Session Termination [SESSTERM] from the Initiator

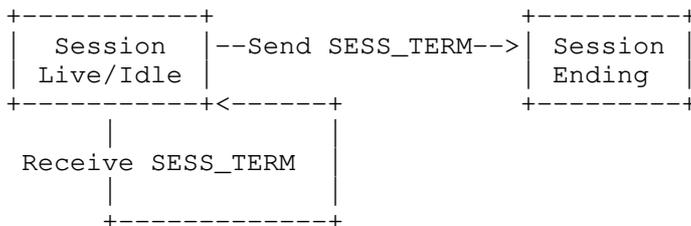


Figure 14: Session Termination [SESSTERM] from the Responder

3.4. Transfer Segmentation Policies

Each TCPCL session allows a negotiated transfer segmentation policy to be applied in each transfer direction. A receiving node can set the Segment MRU in its contact header to determine the largest acceptable segment size, and a transmitting node can segment a transfer into any sizes smaller than the receiver’s Segment MRU. It is a network administration matter to determine an appropriate segmentation policy for entities operating TCPCL, but guidance given here can be used to steer policy toward performance goals. It is also advised to consider the Segment MRU in relation to chunking/packetization performed by TLS, TCP, and any intermediate network-layer nodes.

Minimum Overhead: For a simple network expected to exchange relatively small bundles, the Segment MRU can be set to be identical to the Transfer MRU which indicates that all transfers can be sent with a single data segment (i.e. no actual segmentation). If the network is closed and all transmitters are known to follow a single-segment transfer policy, then receivers can avoid the necessity of segment reassembly. Because this CL operates over a TCP stream, which suffers from a form of head-of-queue blocking between messages, while one node is transmitting a single XFER_SEGMENT message it is not able to transmit any XFER_ACK or XFER_REFUSE for any associated received transfers.

Predictable Message Sizing: In situations where the maximum message size is desired to be well-controlled, the Segment MRU can be set to the largest acceptable size (the message size less XFER_SEGMENT header size) and transmitters can always segment a transfer into maximum-size chunks no larger than the Segment MRU. This

guarantees that any single XFER_SEGMENT will not monopolize the TCP stream for too long, which would prevent outgoing XFER_ACK and XFER_REFUSE associated with received transfers.

Dynamic Segmentation: Even after negotiation of a Segment MRU for each receiving node, the actual transfer segmentation only needs to guarantee that any individual segment is no larger than that MRU. In a situation where network "goodput" is dynamic, the transfer segmentation size can also be dynamic in order to control message transmission duration.

Many other policies can be established in a TCPCL network between the two extremes of minimum overhead (large MRU, single-segment) and predictable message sizing (small MRU, highly segmented). Different policies can be applied to each transfer stream to and from any particular node. Additionally, future header and transfer extension types can apply further nuance to transfer policies and policy negotiation.

3.5. Example Message Exchange

The following figure depicts the protocol exchange for a simple session, showing the session establishment and the transmission of a single bundle split into three data segments (of lengths "L1", "L2", and "L3") from Entity A to Entity B.

Note that the sending node can transmit multiple XFER_SEGMENT messages without waiting for the corresponding XFER_ACK responses. This enables pipelining of messages on a transfer stream. Although this example only demonstrates a single bundle transmission, it is also possible to pipeline multiple XFER_SEGMENT messages for different bundles without necessarily waiting for XFER_ACK messages to be returned for each one. However, interleaving data segments from different bundles is not allowed.

No errors or rejections are shown in this example.

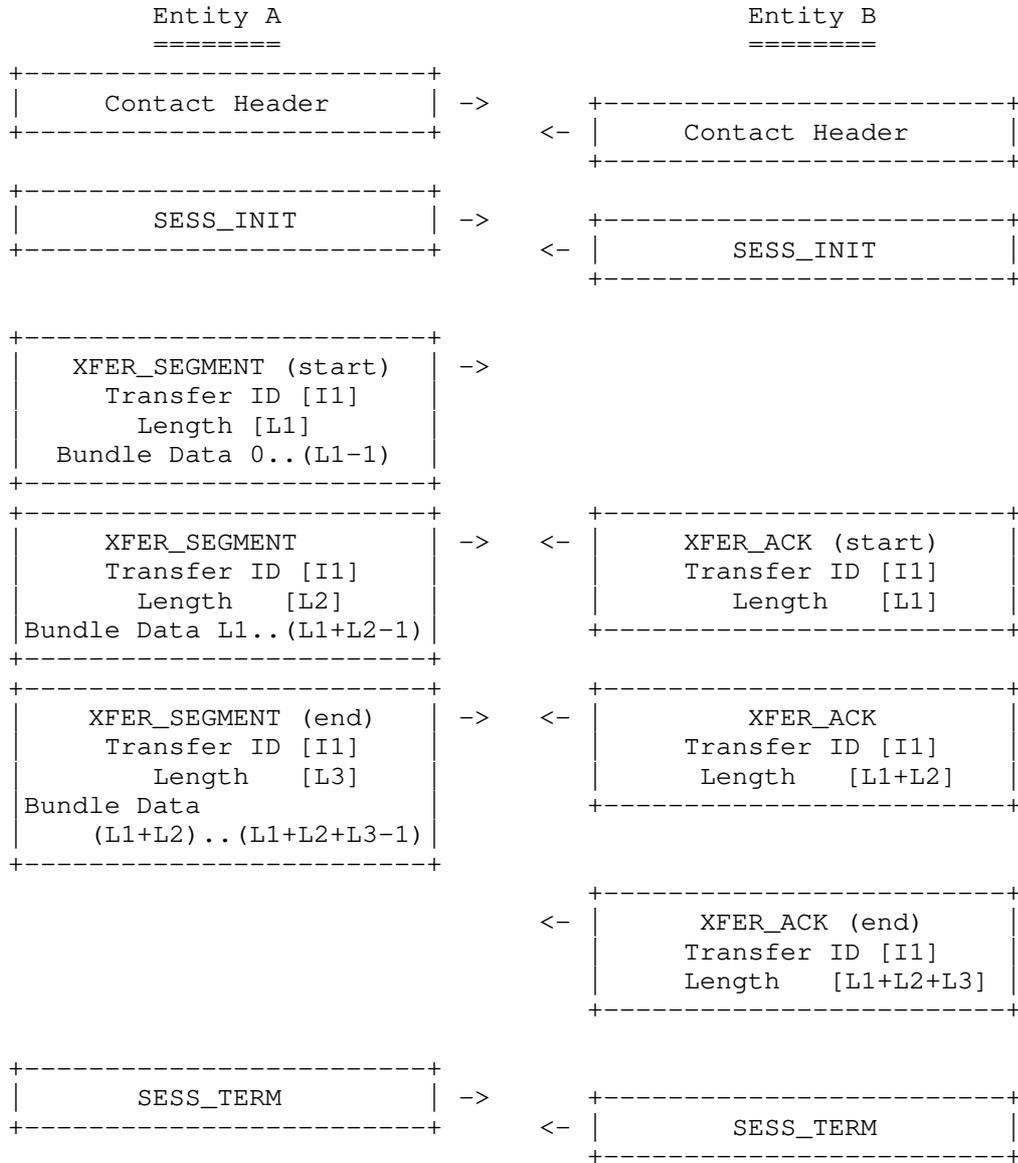


Figure 15: An example of the flow of protocol messages on a single TCP Session between two entities

4. Session Establishment

For bundle transmissions to occur using the TCPCL, a TCPCL session MUST first be established between communicating entities. It is up to the implementation to decide how and when session setup is triggered. For example, some sessions MAY be opened proactively and maintained for as long as is possible given the network conditions, while other sessions MAY be opened only when there is a bundle that is queued for transmission and the routing algorithm selects a certain next-hop node.

4.1. TCP Connection

To establish a TCPCL session, an entity MUST first establish a TCP connection with the intended peer entity, typically by using the services provided by the operating system. Destination port number 4556 has been assigned by IANA as the Registered Port number for the TCP convergence layer. Other destination port numbers MAY be used per local configuration. Determining a peer's destination port number (if different from the registered TCPCL port number) is up to the implementation. Any source port number MAY be used for TCPCL sessions. Typically an operating system assigned number in the TCP Ephemeral range (49152-65535) is used.

If the entity is unable to establish a TCP connection for any reason, then it is an implementation matter to determine how to handle the connection failure. An entity MAY decide to re-attempt to establish the connection. If it does so, it MUST NOT overwhelm its target with repeated connection attempts. Therefore, the entity MUST retry the connection setup no earlier than some delay time from the last attempt, and it SHOULD use a (binary) exponential back-off mechanism to increase this delay in case of repeated failures. The upper limit on a re-attempt back-off is implementation defined but SHOULD be no longer than one minute before signaling to the BP agent that a connection cannot be made.

Once a TCP connection is established, each entity MUST immediately transmit a contact header over the TCP connection. The format of the contact header is described in Section 4.2. Because the TCPCL protocol version in use is part of the initial contact header, nodes using TCPCL version 4 can coexist on a network with nodes using earlier TCPCL versions (with some negotiation needed for interoperation as described in Section 4.3).

4.2. Contact Header

Once a TCP connection is established, both parties exchange a contact header. This section describes the format of the contact header and the meaning of its fields.

Upon receipt of the contact header, both entities perform the validation and negotiation procedures defined in Section 4.3. After receiving the contact header from the other entity, either entity MAY refuse the session by sending a SESS_TERM message with an appropriate reason code.

The format for the Contact Header is as follows:

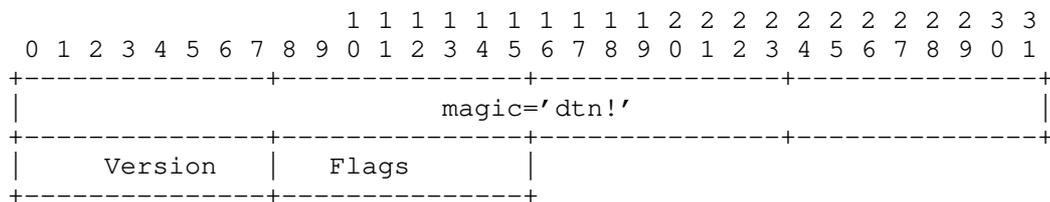


Figure 16: Contact Header Format

See Section 4.3 for details on the use of each of these contact header fields.

The fields of the contact header are:

magic: A four-octet field that always contains the octet sequence 0x64 0x74 0x6E 0x21, i.e., the text string "dtn!" in US-ASCII (and UTF-8).

Version: A one-octet field value containing the value 4 (current version of the TCPCL).

Flags: A one-octet field of single-bit flags, interpreted according to the descriptions in Table 1. All reserved header flag bits SHALL be not set by the sender. All reserved header flag bits SHALL be ignored by the receiver.

Name	Code	Description
CAN_TLS	0x01	If bit is set, indicates that the sending peer is capable of TLS security.
Reserved	others	

Table 1: Contact Header Flags

4.3. Contact Validation and Negotiation

Upon reception of the contact header, each node follows the following procedures to ensure the validity of the TCPCL session and to negotiate values for the session parameters.

If the magic string is not present or is not valid, the connection MUST be terminated. The intent of the magic string is to provide some protection against an inadvertent TCP connection by a different protocol than the one described in this document. To prevent a flood of repeated connections from a misconfigured application, an entity MAY elect to hold an invalid connection open and idle for some time before ending it.

The first negotiation is on the TCPCL protocol version to use. The active node always sends its Contact Header first and waits for a response from the passive node. The active node can repeatedly attempt different protocol versions in descending order until the passive node accepts one with a corresponding Contact Header reply. Only upon response of a Contact Header from the passive node is the TCPCL protocol version established and parameter negotiation begun.

During contact initiation, the active TCPCL node SHALL send the highest TCPCL protocol version on a first session attempt for a TCPCL peer. If the active node receives a Contact Header with a different protocol version than the one sent earlier on the TCP connection, the TCP connection SHALL be terminated. If the active node receives a SESS_TERM message with reason of "Version Mismatch", that node MAY attempt further TCPCL sessions with the peer using earlier protocol version numbers in decreasing order. Managing multi-TCPCL-session state such as this is an implementation matter.

If the passive node receives a contact header containing a version that is greater than the current version of the TCPCL that the node implements, then the node SHALL shutdown the session with a reason code of "Version mismatch". If the passive node receives a contact header with a version that is lower than the version of the protocol

that the node implements, the node MAY either terminate the session (with a reason code of "Version mismatch") or the node MAY adapt its operation to conform to the older version of the protocol. The decision of version fall-back is an implementation matter.

4.4. Session Security

This version of the TCPCL supports establishing a Transport Layer Security (TLS) session within an existing TCP connection. When TLS is used within the TCPCL it affects the entire session. Once established, there is no mechanism available to downgrade a TCPCL session to non-TLS operation. If this is desired, the entire TCPCL session MUST be terminated and a new non-TLS-negotiated session established.

4.4.1. TLS Handshake

The use of TLS is negotiated using the Contact Header as described in Section 4.3. After negotiating an Enable TLS parameter of true, and before any other TCPCL messages are sent within the session, the session entities SHALL begin a TLS handshake in accordance with TLS 1.2 [RFC5246] or any successors that are compatible with TLS 1.2. By convention, this protocol uses the node which initiated the underlying TCP connection as the "client" role of the TLS handshake request.

The TLS handshake, if it occurs, is considered to be part of the contact negotiation before the TCPCL session itself is established. Specifics about sensitive data exposure are discussed in Section 8.

The parameters within each TLS negotiation are implementation dependent but any TCPCL node SHALL follow all recommended practices of BCP 195 [RFC7525], or any updates or successors that become part of BCP 195. The TLS handshake SHOULD include a Server Name Indication from the active peer. The TLS handshake SHALL request a client-side certificate to allow authentication of the active peer. The passive peer SHOULD supply a certificate within the TLS handshake to allow authentication of its side of the session. The active peer SHOULD supply a certificate within the TLS handshake to allow authentication of its side of the session. All certificates supplied during TLS handshake SHALL conform with the profile of [RFC5280]. When a certificate is supplied during TLS handshake, the full certification chain SHOULD be included unless security policy indicates that is unnecessary.

If a TLS handshake cannot negotiate a TLS session, both entities of the TCPCL session SHALL close the TCP connection. At this point the TCPCL session has not yet been established so there is no TCPCL

session to terminate. This also avoids any potential security issues associated with further TCP communication with an untrusted peer.

After a TLS session is successfully established, the active peer SHALL send a SESS_INIT message to begin session negotiation. This session negotiation and all subsequent messaging are secured.

4.4.2. TLS Authentication

Using X.509 certificates exchanged during the TLS handshake, each of the entities can attempt to authenticate its peer at the network layer (host name and address) and at the application layer (BP Node ID). The Node ID exchanged in the Session Initialization is likely to be used by the BP agent for making transfer and routing decisions, so attempting host name validation is optional while attempting Node ID validation is required. The logic for attempting validation is separate from the logic for handling the result of validation, which is based on local security policy.

Any certificate received during TLS handshake SHALL be validated up to one or more trusted certificate authority (CA) certificates. If certificate validation fails or if security policy disallows a certificate for any reason, the entity SHOULD terminate the session (with a reason code of "Contact Failure").

Immediately after the TLS handshake, each side of the TCP connection SHOULD perform host name validation of its peer in accordance with [RFC6125] unless it is not needed by security policy. The active peer SHALL attempt to authenticate the host name (of the passive peer) used to initiate the TCP connection. The active peer MAY attempt to authenticate the IP address of the other side of the TCP connection. The passive peer SHALL attempt to authenticate the IP address of the other side of the TCP connection. The passive peer MAY use the IP address to resolve one or more host names of the active peer and attempt to authenticate those. If host name validation fails (including failure because the certificate does not contain any DNS-ID), the entity SHOULD terminate the session (with a reason code of "Contact Failure") unless security policy allows an unauthenticated host.

Immediately before Session Parameter Negotiation, each side of the session SHALL perform Node ID validation of its peer as described below. Node ID validation SHALL succeed if the associated certificate contains a subjectAltName entry of type uniformResourceIdentifier whose value matches the Node ID of the TCPCL entity. URI matching of Node IDs SHALL use the URI comparison logic of [RFC3986] and scheme-based normalization of those schemes specified in [I-D.ietf-dtn-bpbis]. This is similar to the URI-ID of

[RFC6125] but does not require any structure to the scheme-specific-part of the URI. If Node ID validation fails (including failure because the certificate does not contain any subjectAltName URI), the entity SHOULD terminate the session (with a reason code of "Contact Failure") unless security policy allows an unauthenticated node.

4.4.3. Example TLS Initiation

A summary of a typical TLS use is shown in the sequence in Figure 17 below.

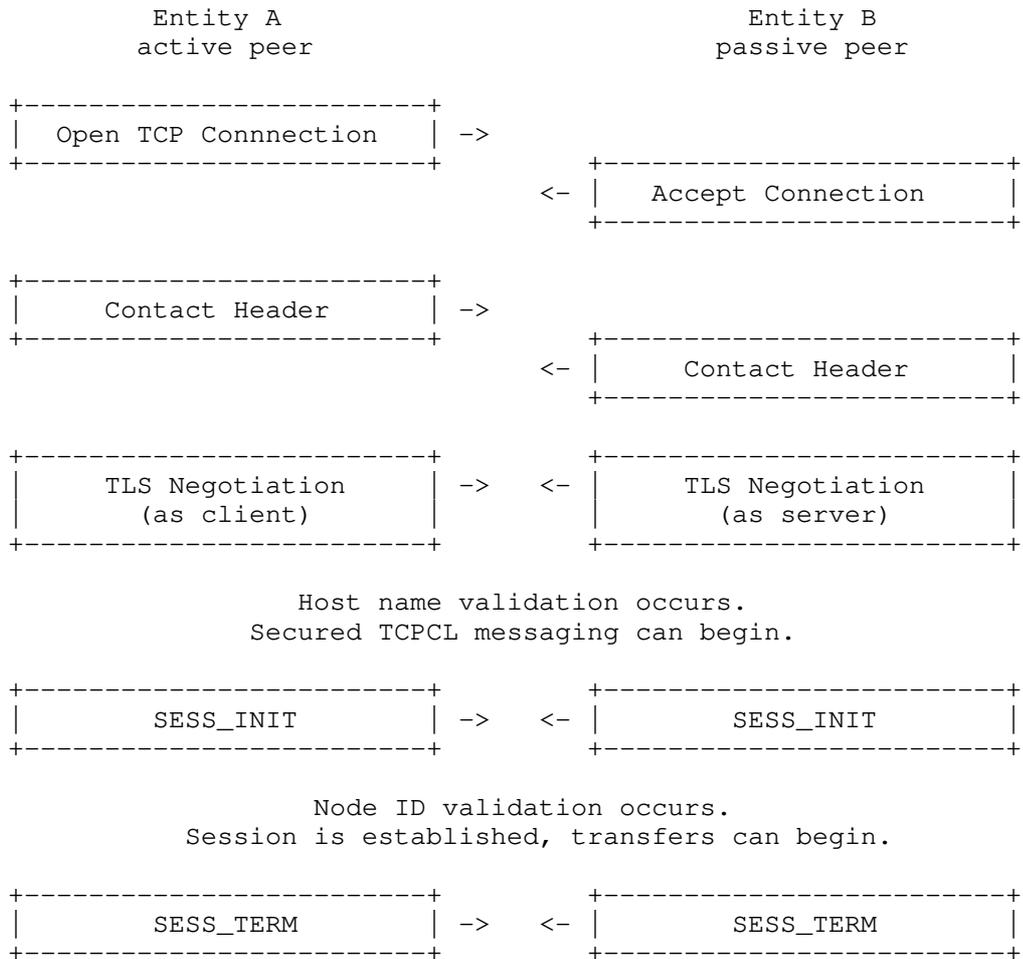


Figure 17: A simple visual example of TCPCL TLS Establishment between two entities

4.5. Message Header

After the initial exchange of a contact header, all messages transmitted over the session are identified by a one-octet header with the following structure:

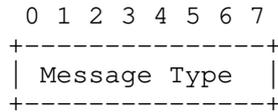


Figure 18: Format of the Message Header

The message header fields are as follows:

Message Type: Indicates the type of the message as per Table 2 below. Encoded values are listed in Section 9.5.

Name	Code	Description
SESS_INIT	0x07	Contains the session parameter inputs from one of the entities, as described in Section 4.6.
SESS_TERM	0x05	Indicates that one of the entities participating in the session wishes to cleanly terminate the session, as described in Section 6.
XFER_SEGMENT	0x01	Indicates the transmission of a segment of bundle data, as described in Section 5.2.2.
XFER_ACK	0x02	Acknowledges reception of a data segment, as described in Section 5.2.3.
XFER_REFUSE	0x03	Indicates that the transmission of the current bundle SHALL be stopped, as described in Section 5.2.4.
KEEPALIVE	0x04	Used to keep TCPCL session active, as described in Section 5.1.1.
MSG_REJECT	0x06	Contains a TCPCL message rejection, as described in Section 5.1.2.

Table 2: TCPCL Message Types

4.6. Session Initialization Message (SESS_INIT)

Before a session is established and ready to transfer bundles, the session parameters are negotiated between the connected entities. The SESS_INIT message is used to convey the per-entity parameters which are used together to negotiate the per-session parameters as described in Section 4.7.

The format of a SESS_INIT message is as follows in Figure 19.

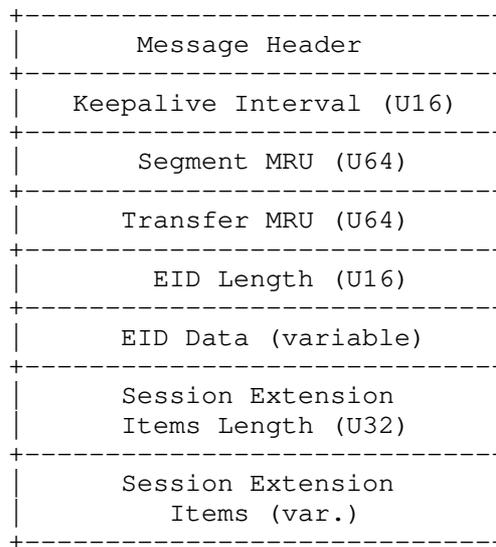


Figure 19: SESS_INIT Format

The fields of the SESS_INIT message are:

Keepalive Interval: A 16-bit unsigned integer indicating the interval, in seconds, between any subsequent messages being transmitted by the peer. The peer receiving this contact header uses this interval to determine how long to wait after any last-message transmission and a necessary subsequent KEEPALIVE message transmission.

Segment MRU: A 64-bit unsigned integer indicating the largest allowable single-segment data payload size to be received in this session. Any XFER_SEGMENT sent to this peer SHALL have a data payload no longer than the peer's Segment MRU. The two entities of a single session MAY have different Segment MRUs, and no relation between the two is required.

Transfer MRU: A 64-bit unsigned integer indicating the largest allowable total-bundle data size to be received in this session. Any bundle transfer sent to this peer SHALL have a Total Bundle Length payload no longer than the peer's Transfer MRU. This value can be used to perform proactive bundle fragmentation. The two entities of a single session MAY have different Transfer MRUs, and no relation between the two is required.

Node ID Length and Node ID Data: Together these fields represent a variable-length text string. The Node ID Length is a 16-bit

unsigned integer indicating the number of octets of Node ID Data to follow. A zero-length Node ID SHALL be used to indicate the lack of Node ID rather than a truly empty Node ID. This case allows an entity to avoid exposing Node ID information on an untrusted network. A non-zero-length Node ID Data SHALL contain the UTF-8 encoded Node ID of the Entity which sent the SESS_INIT message. Every Node ID SHALL be a URI consistent with the requirements of [RFC3986] and the URI schemes of [I-D.ietf-dtn-bpbis]. The Node ID itself can be authenticated as described in Section 4.4.2.

Session Extension Length and Session Extension Items: Together these fields represent protocol extension data not defined by this specification. The Session Extension Length is the total number of octets to follow which are used to encode the Session Extension Item list. The encoding of each Session Extension Item is within a consistent data container as described in Section 4.8. The full set of Session Extension Items apply for the duration of the TCPCL session to follow. The order and multiplicity of these Session Extension Items MAY be significant, as defined in the associated type specification(s).

4.7. Session Parameter Negotiation

An entity calculates the parameters for a TCPCL session by negotiating the values from its own preferences (conveyed by the contact header it sent to the peer) with the preferences of the peer node (expressed in the contact header that it received from the peer). The negotiated parameters defined by this specification are described in the following paragraphs.

Transfer MTU and Segment MTU: The maximum transmit unit (MTU) for whole transfers and individual segments are identical to the Transfer MRU and Segment MRU, respectively, of the received contact header. A transmitting peer can send individual segments with any size smaller than the Segment MTU, depending on local policy, dynamic network conditions, etc. Determining the size of each transmitted segment is an implementation matter.

Session Keepalive: Negotiation of the Session Keepalive parameter is performed by taking the minimum of this two contact headers' Keepalive Interval. The Session Keepalive interval is a parameter for the behavior described in Section 5.1.1.

Enable TLS: Negotiation of the Enable TLS parameter is performed by taking the logical AND of the two contact headers' CAN_TLS flags. A local security policy is then applied to determine if the negotiated value of Enable TLS is acceptable. It can be a

reasonable security policy to both require or disallow the use of TLS depending upon the desired network flows. Because this state is negotiated over an unsecured medium, there is a risk of a TLS Stripping as described in Section 8. If the Enable TLS state is unacceptable, the node SHALL terminate the session with a reason code of "Contact Failure". Note that this contact failure reason is different than a failure of TLS handshake or TLS authentication after an agreed-upon and acceptable Enable TLS state. If the negotiated Enable TLS value is true and acceptable then TLS negotiation feature (described in Section 4.4) begins immediately following the contact header exchange.

Once this process of parameter negotiation is completed (which includes a possible completed TLS handshake of the connection to use TLS), this protocol defines no additional mechanism to change the parameters of an established session; to effect such a change, the TCPCL session MUST be terminated and a new session established.

4.8. Session Extension Items

Each of the Session Extension Items SHALL be encoded in an identical Type-Length-Value (TLV) container form as indicated in Figure 20.

The fields of the Session Extension Item are:

Flags: A one-octet field containing generic bit flags about the Item, which are listed in Table 3. All reserved header flag bits SHALL be not set by the sender. All reserved header flag bits SHALL be ignored by the receiver. If a TCPCL entity receives a Session Extension Item with an unknown Item Type and the CRITICAL flag set, the entity SHALL close the TCPCL session with SESS_TERM reason code of "Contact Failure". If the CRITICAL flag is not set, an entity SHALL skip over and ignore any item with an unknown Item Type.

Item Type: A 16-bit unsigned integer field containing the type of the extension item. This specification does not define any extension types directly, but does allocate an IANA registry for such codes (see Section 9.3).

Item Length: A 16-bit unsigned integer field containing the number of Item Value octets to follow.

Item Value: A variable-length data field which is interpreted according to the associated Item Type. This specification places no restrictions on an extension's use of available Item Value data. Extension specifications SHOULD avoid the use of large data

lengths, as no bundle transfers can begin until the full extension data is sent.

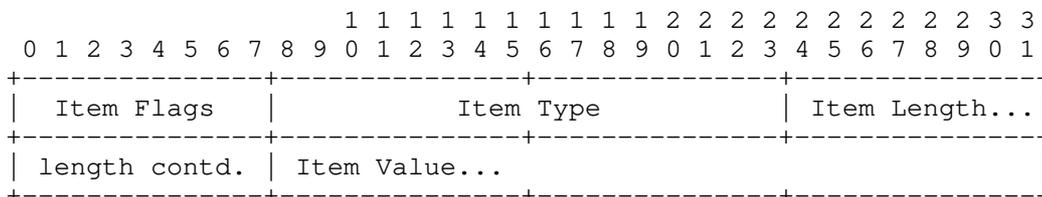


Figure 20: Session Extension Item Format

Name	Code	Description
CRITICAL	0x01	If bit is set, indicates that the receiving peer must handle the extension item.
Reserved	others	

Table 3: Session Extension Item Flags

5. Established Session Operation

This section describes the protocol operation for the duration of an established session, including the mechanism for transmitting bundles over the session.

5.1. Upkeep and Status Messages

5.1.1. Session Upkeep (KEEPALIVE)

The protocol includes a provision for transmission of KEEPALIVE messages over the TCPCL session to help determine if the underlying TCP connection has been disrupted.

As described in Section 4.3, a negotiated parameter of each session is the Session Keepalive interval. If the negotiated Session Keepalive is zero (i.e. one or both contact headers contains a zero Keepalive Interval), then the keepalive feature is disabled. There is no logical minimum value for the keepalive interval, but when used for many sessions on an open, shared network a short interval could lead to excessive traffic. For shared network use, entities SHOULD choose a keepalive interval no shorter than 30 seconds. There is no logical maximum value for the keepalive interval, but an idle TCP connection is liable for closure by the host operating system if the

keepalive time is longer than tens-of-minutes. Entities SHOULD choose a keepalive interval no longer than 10 minutes (600 seconds).

Note: The Keepalive Interval SHOULD NOT be chosen too short as TCP retransmissions MAY occur in case of packet loss. Those will have to be triggered by a timeout (TCP retransmission timeout (RTO)), which is dependent on the measured RTT for the TCP connection so that KEEPALIVE messages MAY experience noticeable latency.

The format of a KEEPALIVE message is a one-octet message type code of KEEPALIVE (as described in Table 2) with no additional data. Both sides SHALL send a KEEPALIVE message whenever the negotiated interval has elapsed with no transmission of any message (KEEPALIVE or other).

If no message (KEEPALIVE or other) has been received in a session after some implementation-defined time duration, then the node SHALL terminate the session by transmitting a SESS_TERM message (as described in Section 6.1) with reason code "Idle Timeout". If configurable, the idle timeout duration SHOULD be no shorter than twice the keepalive interval. If not configurable, the idle timeout duration SHOULD be exactly twice the keepalive interval.

5.1.2. Message Rejection (MSG_REJECT)

If a TCPCL node receives a message which is unknown to it (possibly due to an unhandled protocol mismatch) or is inappropriate for the current session state (e.g. a KEEPALIVE message received after contact header negotiation has disabled that feature), there is a protocol-level message to signal this condition in the form of a MSG_REJECT reply.

The format of a MSG_REJECT message is as follows in Figure 21.

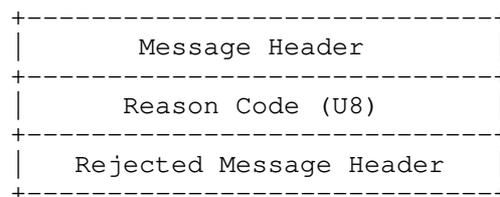


Figure 21: Format of MSG_REJECT Messages

The fields of the MSG_REJECT message are:

Reason Code: A one-octet refusal reason code interpreted according to the descriptions in Table 4.

Rejected Message Header: The Rejected Message Header is a copy of the Message Header to which the MSG_REJECT message is sent as a response.

Name	Code	Description
Message Type Unknown	0x01	A message was received with a Message Type code unknown to the TCPCL node.
Message Unsupported	0x02	A message was received but the TCPCL node cannot comply with the message contents.
Message Unexpected	0x03	A message was received while the session is in a state in which the message is not expected.

Table 4: MSG_REJECT Reason Codes

5.2. Bundle Transfer

All of the messages in this section are directly associated with transferring a bundle between TCPCL entities.

A single TCPCL transfer results in a bundle (handled by the convergence layer as opaque data) being exchanged from one node to the other. In TCPCL a transfer is accomplished by dividing a single bundle up into "segments" based on the receiving-side Segment MRU (see Section 4.2). The choice of the length to use for segments is an implementation matter, but each segment MUST be no larger than the receiving node's maximum receive unit (MRU) (see the field "Segment MRU" of Section 4.2). The first segment for a bundle MUST set the 'START' flag, and the last one MUST set the 'end' flag in the XFER_SEGMENT message flags.

A single transfer (and by extension a single segment) SHALL NOT contain data of more than a single bundle. This requirement is imposed on the agent using the TCPCL rather than TCPCL itself.

If multiple bundles are transmitted on a single TCPCL connection, they MUST be transmitted consecutively without interleaving of segments from multiple bundles.

5.2.1. Bundle Transfer ID

Each of the bundle transfer messages contains a Transfer ID which is used to correlate messages (from both sides of a transfer) for each bundle. A Transfer ID does not attempt to address uniqueness of the bundle data itself and has no relation to concepts such as bundle fragmentation. Each invocation of TCPCL by the bundle protocol agent, requesting transmission of a bundle (fragmentary or otherwise), results in the initiation of a single TCPCL transfer. Each transfer entails the sending of a sequence of some number of XFER_SEGMENT and XFER_ACK messages; all are correlated by the same Transfer ID.

Transfer IDs from each node SHALL be unique within a single TCPCL session. The initial Transfer ID from each node SHALL have value zero. Subsequent Transfer ID values SHALL be incremented from the prior Transfer ID value by one. Upon exhaustion of the entire 64-bit Transfer ID space, the sending node SHALL terminate the session with SESS_TERM reason code "Resource Exhaustion".

For bidirectional bundle transfers, a TCPCL node SHOULD NOT rely on any relation between Transfer IDs originating from each side of the TCPCL session.

5.2.2. Data Transmission (XFER_SEGMENT)

Each bundle is transmitted in one or more data segments. The format of a XFER_SEGMENT message follows in Figure 22.

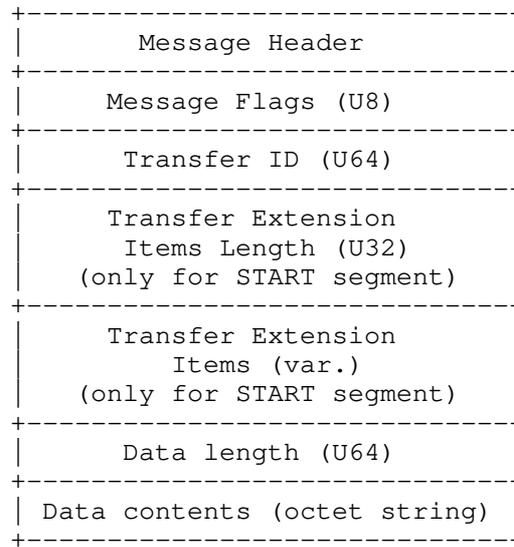


Figure 22: Format of XFER_SEGMENT Messages

The fields of the XFER_SEGMENT message are:

Message Flags: A one-octet field of single-bit flags, interpreted according to the descriptions in Table 5. All reserved header flag bits SHALL be not set by the sender. All reserved header flag bits SHALL be ignored by the receiver.

Transfer ID: A 64-bit unsigned integer identifying the transfer being made.

Transfer Extension Length and Transfer Extension Items: Together these fields represent protocol extension data for this specification. The Transfer Extension Length and Transfer Extension Item fields SHALL only be present when the 'START' flag is set on the message. The Transfer Extension Length is the total number of octets to follow which are used to encode the Transfer Extension Item list. The encoding of each Transfer Extension Item is within a consistent data container as described in Section 5.2.5. The full set of transfer extension items apply only to the associated single transfer. The order and multiplicity of these transfer extension items MAY be significant, as defined in the associated type specification(s).

Data length: A 64-bit unsigned integer indicating the number of octets in the Data contents to follow.

Data contents: The variable-length data payload of the message.

Name	Code	Description
END	0x01	If bit is set, indicates that this is the last segment of the transfer.
START	0x02	If bit is set, indicates that this is the first segment of the transfer.
Reserved	others	

Table 5: XFER_SEGMENT Flags

The flags portion of the message contains two optional values in the two low-order bits, denoted 'START' and 'END' in Table 5. The 'START' bit MUST be set to one if it precedes the transmission of the first segment of a transfer. The 'END' bit MUST be set to one when transmitting the last segment of a transfer. In the case where an entire transfer is accomplished in a single segment, both the 'START' and 'END' bits MUST be set to one.

Once a transfer of a bundle has commenced, the node MUST only send segments containing sequential portions of that bundle until it sends a segment with the 'END' bit set. No interleaving of multiple transfers from the same node is possible within a single TCPCL session. Simultaneous transfers between two entities MAY be achieved using multiple TCPCL sessions.

5.2.3. Data Acknowledgments (XFER_ACK)

Although the TCP transport provides reliable transfer of data between transport peers, the typical BSD sockets interface provides no means to inform a sending application of when the receiving application has processed some amount of transmitted data. Thus, after transmitting some data, the TCPCL needs an additional mechanism to determine whether the receiving agent has successfully received the segment. To this end, the TCPCL protocol provides feedback messaging whereby a receiving node transmits acknowledgments of reception of data segments.

The format of an XFER_ACK message follows in Figure 23.

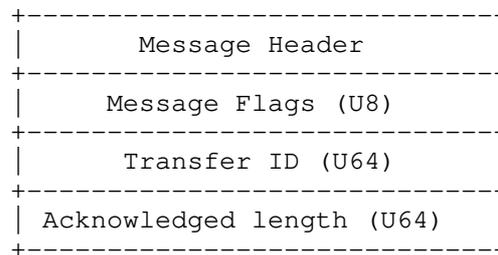


Figure 23: Format of XFER_ACK Messages

The fields of the XFER_ACK message are:

Message Flags: A one-octet field of single-bit flags, interpreted according to the descriptions in Table 5. All reserved header flag bits SHALL be not set by the sender. All reserved header flag bits SHALL be ignored by the receiver.

Transfer ID: A 64-bit unsigned integer identifying the transfer being acknowledged.

Acknowledged length: A 64-bit unsigned integer indicating the total number of octets in the transfer which are being acknowledged.

A receiving TCPCL node SHALL send an XFER_ACK message in response to each received XFER_SEGMENT message. The flags portion of the XFER_ACK header SHALL be set to match the corresponding DATA_SEGMENT message being acknowledged. The acknowledged length of each XFER_ACK contains the sum of the data length fields of all XFER_SEGMENT messages received so far in the course of the indicated transfer. The sending node SHOULD transmit multiple XFER_SEGMENT messages without waiting for the corresponding XFER_ACK responses. This enables pipelining of messages on a transfer stream.

For example, suppose the sending node transmits four segments of bundle data with lengths 100, 200, 500, and 1000, respectively. After receiving the first segment, the node sends an acknowledgment of length 100. After the second segment is received, the node sends an acknowledgment of length 300. The third and fourth acknowledgments are of length 800 and 1800, respectively.

5.2.4. Transfer Refusal (XFER_REFUSE)

The TCPCL supports a mechanism by which a receiving node can indicate to the sender that it does not want to receive the corresponding bundle. To do so, upon receiving an XFER_SEGMENT message, the node MAY transmit a XFER_REFUSE message. As data segments and

acknowledgments MAY cross on the wire, the bundle that is being refused SHALL be identified by the Transfer ID of the refusal.

There is no required relation between the Transfer MRU of a TCPCL node (which is supposed to represent a firm limitation of what the node will accept) and sending of a XFER_REFUSE message. A XFER_REFUSE can be used in cases where the agent's bundle storage is temporarily depleted or somehow constrained. A XFER_REFUSE can also be used after the bundle header or any bundle data is inspected by an agent and determined to be unacceptable.

A receiver MAY send an XFER_REFUSE message as soon as it receives any XFER_SEGMENT message. The sender MUST be prepared for this and MUST associate the refusal with the correct bundle via the Transfer ID fields.

The format of the XFER_REFUSE message is as follows in Figure 24.

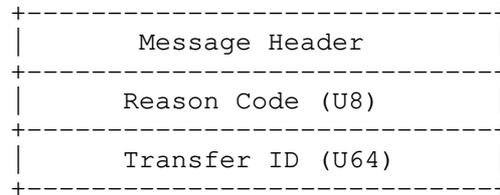


Figure 24: Format of XFER_REFUSE Messages

The fields of the XFER_REFUSE message are:

Reason Code: A one-octet refusal reason code interpreted according to the descriptions in Table 6.

Transfer ID: A 64-bit unsigned integer identifying the transfer being refused.

Name	Code	Description
Unknown	0x00	Reason for refusal is unknown or not specified.
Extension Failure	0x01	A failure processing the Transfer Extension Items has occurred.
Completed	0x02	The receiver already has the complete bundle. The sender MAY consider the bundle as completely received.
No Resources	0x03	The receiver's resources are exhausted. The sender SHOULD apply reactive bundle fragmentation before retrying.
Retransmit	0x04	The receiver has encountered a problem that requires the bundle to be retransmitted in its entirety.

Table 6: XFER_REFUSE Reason Codes

The receiver MUST, for each transfer preceding the one to be refused, have either acknowledged all XFER_SEGMENTS or refused the bundle transfer.

The bundle transfer refusal MAY be sent before an entire data segment is received. If a sender receives a XFER_REFUSE message, the sender MUST complete the transmission of any partially sent XFER_SEGMENT message. There is no way to interrupt an individual TCPCL message partway through sending it. The sender MUST NOT commence transmission of any further segments of the refused bundle subsequently. Note, however, that this requirement does not ensure that an entity will not receive another XFER_SEGMENT for the same bundle after transmitting a XFER_REFUSE message since messages MAY cross on the wire; if this happens, subsequent segments of the bundle SHALL also be refused with a XFER_REFUSE message.

Note: If a bundle transmission is aborted in this way, the receiver MAY not receive a segment with the 'END' flag set to '1' for the aborted bundle. The beginning of the next bundle is identified by the 'START' bit set to '1', indicating the start of a new transfer, and with a distinct Transfer ID value.

5.2.5. Transfer Extension Items

Each of the Transfer Extension Items SHALL be encoded in an identical Type-Length-Value (TLV) container form as indicated in Figure 25.

The fields of the Transfer Extension Item are:

Flags: A one-octet field containing generic bit flags about the Item, which are listed in Table 7. All reserved header flag bits SHALL be not set by the sender. All reserved header flag bits SHALL be ignored by the receiver. If a TCPCL node receives a Transfer Extension Item with an unknown Item Type and the CRITICAL flag set, the node SHALL refuse the transfer with an XFER_REFUSE reason code of "Extension Failure". If the CRITICAL flag is not set, an entity SHALL skip over and ignore any item with an unknown Item Type.

Item Type: A 16-bit unsigned integer field containing the type of the extension item. This specification allocates an IANA registry for such codes (see Section 9.4).

Item Length: A 16-bit unsigned integer field containing the number of Item Value octets to follow.

Item Value: A variable-length data field which is interpreted according to the associated Item Type. This specification places no restrictions on an extension's use of available Item Value data. Extension specifications SHOULD avoid the use of large data lengths, as the associated transfer cannot begin until the full extension data is sent.

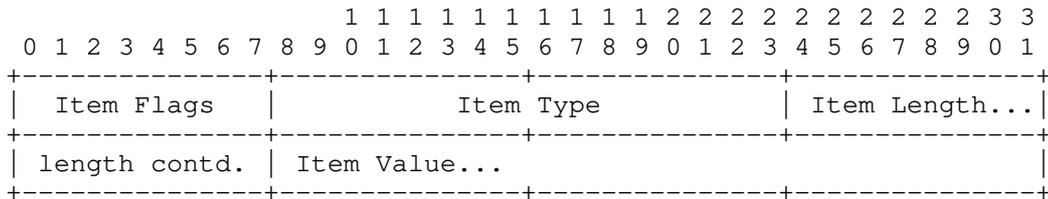


Figure 25: Transfer Extension Item Format

Name	Code	Description
CRITICAL	0x01	If bit is set, indicates that the receiving peer must handle the extension item.
Reserved	others	

Table 7: Transfer Extension Item Flags

5.2.5.1. Transfer Length Extension

The purpose of the Transfer Length extension is to allow entities to preemptively refuse bundles that would exceed their resources or to prepare storage on the receiving node for the upcoming bundle data.

Multiple Transfer Length extension items SHALL NOT occur within the same transfer. The lack of a Transfer Length extension item in any transfer SHALL NOT imply anything about the potential length of the transfer. The Transfer Length extension SHALL be assigned transfer extension type ID 0x0001.

If a transfer occupies exactly one segment (i.e. both START and END bits are set) the Transfer Length extension SHOULD NOT be present. The extension does not provide any additional information for single-segment transfers.

The format of the Transfer Length data is as follows in Figure 26.

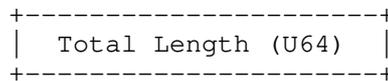


Figure 26: Format of Transfer Length data

The fields of the Transfer Length extension are:

Total Length: A 64-bit unsigned integer indicating the size of the data-to-be-transferred. The Total Length field SHALL be treated as authoritative by the receiver. If, for whatever reason, the actual total length of bundle data received differs from the value indicated by the Total Length value, the receiver SHALL treat the transmitted data as invalid.

6. Session Termination

This section describes the procedures for ending a TCPCL session.

6.1. Session Termination Message (SESS_TERM)

To cleanly shut down a session, a SESS_TERM message SHALL be transmitted by either node at any point following complete transmission of any other message. When sent to initiate a termination, the REPLY bit of a SESS_TERM message SHALL NOT be set. Upon receiving a SESS_TERM message after not sending a SESS_TERM message in the same session, an entity SHALL send an acknowledging SESS_TERM message. When sent to acknowledge a termination, a SESS_TERM message SHALL have identical data content from the message being acknowledged except for the REPLY bit, which is set to indicate acknowledgement.

After sending a SESS_TERM message, an entity MAY continue a possible in-progress transfer in either direction. After sending a SESS_TERM message, an entity SHALL NOT begin any new outgoing transfer for the remainder of the session. After receiving a SESS_TERM message, an entity SHALL NOT accept any new incoming transfer for the remainder of the session.

Instead of following a clean shutdown sequence, after transmitting a SESS_TERM message an entity MAY immediately close the associated TCP connection. When performing an unclean shutdown, a receiving node SHOULD acknowledge all received data segments before closing the TCP connection. Not acknowledging received segments can result in unnecessary retransmission. When performing an unclean shutdown, a transmitting node SHALL treat either sending or receiving a SESS_TERM message (i.e. before the final acknowledgment) as a failure of the transfer. Any delay between request to terminate the TCP connection and actual closing of the connection (a "half-closed" state) MAY be ignored by the TCPCL node.

The format of the SESS_TERM message is as follows in Figure 27.

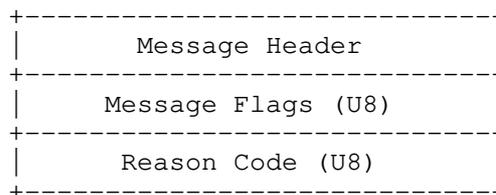


Figure 27: Format of SESS_TERM Messages

The fields of the SESS_TERM message are:

Message Flags: A one-octet field of single-bit flags, interpreted according to the descriptions in Table 8. All reserved header flag bits SHALL be not set by the sender. All reserved header flag bits SHALL be ignored by the receiver.

Reason Code: A one-octet refusal reason code interpreted according to the descriptions in Table 9.

Name	Code	Description
REPLY	0x01	If bit is set, indicates that this message is an acknowledgement of an earlier SESS_TERM message.
Reserved	others	

Table 8: SESS_TERM Flags

Name	Code	Description
Unknown	0x00	A termination reason is not available.
Idle timeout	0x01	The session is being closed due to idleness.
Version mismatch	0x02	The node cannot conform to the specified TCPCL protocol version.
Busy	0x03	The node is too busy to handle the current session.
Contact Failure	0x04	The node cannot interpret or negotiate contact header option.
Resource Exhaustion	0x05	The node has run into some resource limit and cannot continue the session.

Table 9: SESS_TERM Reason Codes

A session shutdown MAY occur immediately after transmission of a contact header (and prior to any further message transmit). This MAY, for example, be used to notify that the node is currently not

able or willing to communicate. However, an entity MUST always send the contact header to its peer before sending a SESS_TERM message.

If reception of the contact header itself somehow fails (e.g. an invalid "magic string" is received), an entity SHALL close the TCP connection without sending a SESS_TERM message. If the content of the Session Extension Items data disagrees with the Session Extension Length (i.e. the last Item claims to use more octets than are present in the Session Extension Length), the reception of the contact header is considered to have failed.

If a session is to be terminated before a protocol message has completed being sent, then the node MUST NOT transmit the SESS_TERM message but still SHALL close the TCP connection. Each TCPCL message is contiguous in the octet stream and has no ability to be cut short and/or preempted by an other message. This is particularly important when large segment sizes are being transmitted; either entire XFER_SEGMENT is sent before a SESS_TERM message or the connection is simply terminated mid-XFER_SEGMENT.

6.2. Idle Session Shutdown

The protocol includes a provision for clean shutdown of idle sessions. Determining the length of time to wait before ending idle sessions, if they are to be closed at all, is an implementation and configuration matter.

If there is a configured time to close idle links and if no TCPCL messages (other than KEEPALIVE messages) has been received for at least that amount of time, then either node MAY terminate the session by transmitting a SESS_TERM message indicating the reason code of "Idle timeout" (as described in Table 9).

7. Implementation Status

[NOTE to the RFC Editor: please remove this section before publication, as well as the reference to [RFC7942] and [github-dtn-bpbis-tcpcl].]

This section records the status of known implementations of the protocol defined by this specification at the time of posting of this Internet-Draft, and is based on a proposal described in [RFC7942]. The description of implementations in this section is intended to assist the IETF in its decision processes in progressing drafts to RFCs. Please note that the listing of any individual implementation here does not imply endorsement by the IETF. Furthermore, no effort has been spent to verify the information presented here that was supplied by IETF contributors. This is not intended as, and must not

be construed to be, a catalog of available implementations or their features. Readers are advised to note that other implementations can exist.

An example implementation of the this draft of TCPCLv4 has been created as a GitHub project [github-dtn-bpbis-tcpcl] and is intended to use as a proof-of-concept and as a possible source of interoperability testing. This example implementation uses D-Bus as the CL-BP Agent interface, so it only runs on hosts which provide the Python "dbus" library.

8. Security Considerations

TCPCL can be used to provide point-to-point transport security, but does not provide security of data-at-rest and does not guarantee end-to-end bundle security. The bundle security mechanisms defined in [I-D.ietf-dtn-bpsec] are to be used instead.

When negotiating whether to use TLS security as part of the contact header exchange, it is possible for a man-in-the-middle attacker to unset the CAN_TLS flag on either side of the exchange. This leads to the "SSL Stripping" attack described in [RFC7457]. If TLS is desired for use on any TCPCL network, it is strongly encouraged that the security policy disallow use of TCPCL when "Enable TLS" is negotiated to false. This requires that the TLS handshake occurs, regardless of the policy-driven parameters of the handshake and policy-driven handling of the handshake outcome.

Even when using TLS to secure the TCPCL session, the actual ciphersuite negotiated between the TLS peers MAY be insecure. TLS can be used to perform authentication without data confidentiality, for example. It is up to security policies within each TCPCL node to ensure that the negotiated TLS ciphersuite meets transport security requirements. This is identical behavior to STARTTLS use in [RFC2595].

The certificates exchanged by TLS enable authentication of peer host name and Node ID, but it is possible that a peer either not provide a valid certificate or that the certificate does not validate either the host name or Node ID of the peer. Having a CA-validated certificate does not alone guarantee the identity of the network host or BP node from which the certificate is provided; additional validation procedures bind the host name or node ID based on the contents of the certificate. The host name validation is a weaker form of authentication, because even if a peer is operating on an authenticated network host name it can provide an invalid Node ID and cause bundles to be "leaked" to an invalid node. Especially in DTN environments, network names and addresses of nodes can be time-

variable so binding a certificate to a Node ID is a more stable identity. Node ID validation ensures that the peer to which a bundle is transferred is in fact the node which the BP Agent expects it to be. It is a reasonable policy to skip host name validation if certificates can be guaranteed to validate the peer's Node ID. In circumstances where certificates can only be issued to network host names, Node ID validation is not possible but it could be reasonable to assume that a trusted host is not going to present an invalid Node ID. Trusting an authenticated host name can be feasible on a network secured by a private CA but is not advisable on the Internet when using a variety of public CAs.

Another consideration for this protocol relates to denial-of-service attacks. An entity MAY send a large amount of data over a TCPCL session, requiring the receiving entity to handle the data, attempt to stop the flood of data by sending a XFER_REFUSE message, or forcibly terminate the session. This burden could cause denial of service on other, well-behaving sessions. There is also nothing to prevent a malicious entity from continually establishing sessions and repeatedly trying to send copious amounts of bundle data. A listening entity MAY take countermeasures such as ignoring TCP SYN messages, closing TCP connections as soon as they are established, waiting before sending the contact header, sending a SESS_TERM message quickly or with a delay, etc.

9. IANA Considerations

Registration procedures referred to in this section are defined in [RFC8126].

Some of the registries have been defined as version specific to TCPCLv4, and imports some or all codepoints from TCPCLv3. This was done to disambiguate the use of these codepoints between TCPCLv3 and TCPCLv4 while preserving the semantics of some of the codepoints.

9.1. Port Number

Port number 4556 has been previously assigned as the default port for the TCP convergence layer in [RFC7242]. This assignment is unchanged by protocol version 4. Each TCPCL entity identifies its TCPCL protocol version in its initial contact (see Section 9.2), so there is no ambiguity about what protocol is being used.

Parameter	Value
Service Name:	dtn-bundle
Transport Protocol(s):	TCP
Assignee:	Simon Perreault <simon@per.reau.lt>
Contact:	Simon Perreault <simon@per.reau.lt>
Description:	DTN Bundle TCP CL Protocol
Reference:	[RFC7242]
Port Number:	4556

9.2. Protocol Versions

IANA has created, under the "Bundle Protocol" registry, a sub-registry titled "Bundle Protocol TCP Convergence-Layer Version Numbers" and initialize it with the following table. The registration procedure is RFC Required.

Value	Description	Reference
0	Reserved	[RFC7242]
1	Reserved	[RFC7242]
2	Reserved	[RFC7242]
3	TCPCL	[RFC7242]
4	TCPCLv4	This specification.
5-255	Unassigned	

9.3. Session Extension Types

EDITOR NOTE: sub-registry to-be-created upon publication of this specification.

IANA will create, under the "Bundle Protocol" registry, a sub-registry titled "Bundle Protocol TCP Convergence-Layer Version 4

Session Extension Types" and initialize it with the contents of Table 10. The registration procedure is Expert Review within the lower range 0x0001--0x7FFF. Values in the range 0x8000--0xFFFF are reserved for use on private networks for functions not published to the IANA.

Code	Session Extension Type
0x0000	Reserved
0x0001--0x7FFF	Unassigned
0x8000--0xFFFF	Private/Experimental Use

Table 10: Session Extension Type Codes

9.4. Transfer Extension Types

EDITOR NOTE: sub-registry to-be-created upon publication of this specification.

IANA will create, under the "Bundle Protocol" registry, a sub-registry titled "Bundle Protocol TCP Convergence-Layer Version 4 Transfer Extension Types" and initialize it with the contents of Table 11. The registration procedure is Expert Review within the lower range 0x0001--0x7FFF. Values in the range 0x8000--0xFFFF are reserved for use on private networks for functions not published to the IANA.

Code	Transfer Extension Type
0x0000	Reserved
0x0001	Transfer Length Extension
0x0002--0x7FFF	Unassigned
0x8000--0xFFFF	Private/Experimental Use

Table 11: Transfer Extension Type Codes

9.5. Message Types

EDITOR NOTE: sub-registry to-be-created upon publication of this specification.

IANA will create, under the "Bundle Protocol" registry, a sub-registry titled "Bundle Protocol TCP Convergence-Layer Version 4 Message Types" and initialize it with the contents of Table 12. The registration procedure is RFC Required within the lower range 0x01--0xEF. Values in the range 0xF0--0xFF are reserved for use on private networks for functions not published to the IANA.

Code	Message Type
0x00	Reserved
0x01	XFER_SEGMENT
0x02	XFER_ACK
0x03	XFER_REFUSE
0x04	KEEPALIVE
0x05	SESS_TERM
0x06	MSG_REJECT
0x07	SESS_INIT
0x08--0xEF	Unassigned
0xF0--0xFF	Private/Experimental Use

Table 12: Message Type Codes

9.6. XFER_REFUSE Reason Codes

EDITOR NOTE: sub-registry to-be-created upon publication of this specification.

IANA will create, under the "Bundle Protocol" registry, a sub-registry titled "Bundle Protocol TCP Convergence-Layer Version 4 XFER_REFUSE Reason Codes" and initialize it with the contents of Table 13. The registration procedure is Specification Required within the lower range 0x00--0xEF. Values in the range 0xF0--0xFF

are reserved for use on private networks for functions not published to the IANA.

Code	Refusal Reason
0x00	Unknown
0x01	Extension Failure
0x02	Completed
0x03	No Resources
0x04	Retransmit
0x05--0xEF	Unassigned
0xF0--0xFF	Private/Experimental Use

Table 13: XFER_REFUSE Reason Codes

9.7. SESS_TERM Reason Codes

EDITOR NOTE: sub-registry to-be-created upon publication of this specification.

IANA will create, under the "Bundle Protocol" registry, a sub-registry titled "Bundle Protocol TCP Convergence-Layer Version 4 SESS_TERM Reason Codes" and initialize it with the contents of Table 14. The registration procedure is Specification Required within the lower range 0x00--0xEF. Values in the range 0xF0--0xFF are reserved for use on private networks for functions not published to the IANA.

Code	Termination Reason
0x00	Unknown
0x01	Idle timeout
0x02	Version mismatch
0x03	Busy
0x04	Contact Failure
0x05	Resource Exhaustion
0x06--0xEF	Unassigned
0xF0--0xFF	Private/Experimental Use

Table 14: SESS_TERM Reason Codes

9.8. MSG_REJECT Reason Codes

EDITOR NOTE: sub-registry to-be-created upon publication of this specification.

IANA will create, under the "Bundle Protocol" registry, a sub-registry titled "Bundle Protocol TCP Convergence-Layer Version 4 MSG_REJECT Reason Codes" and initialize it with the contents of Table 15. The registration procedure is Specification Required within the lower range 0x01--0xEF. Values in the range 0xF0--0xFF are reserved for use on private networks for functions not published to the IANA.

Code	Rejection Reason
0x00	reserved
0x01	Message Type Unknown
0x02	Message Unsupported
0x03	Message Unexpected
0x04--0xEF	Unassigned
0xF0--0xFF	Private/Experimental Use

Table 15: MSG_REJECT Reason Codes

10. Acknowledgments

This specification is based on comments on implementation of [RFC7242] provided from Scott Burleigh.

11. References

11.1. Normative References

- [I-D.ietf-dtn-bpbis]
 Burleigh, S., Fall, K., and E. Birrane, "Bundle Protocol Version 7", draft-ietf-dtn-bpbis-14 (work in progress), August 2019.
- [RFC0793] Postel, J., "Transmission Control Protocol", STD 7, RFC 793, DOI 10.17487/RFC0793, September 1981, <<https://www.rfc-editor.org/info/rfc793>>.
- [RFC1122] Braden, R., Ed., "Requirements for Internet Hosts - Communication Layers", STD 3, RFC 1122, DOI 10.17487/RFC1122, October 1989, <<https://www.rfc-editor.org/info/rfc1122>>.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.

- [RFC3986] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", STD 66, RFC 3986, DOI 10.17487/RFC3986, January 2005, <<https://www.rfc-editor.org/info/rfc3986>>.
- [RFC5246] Dierks, T. and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2", RFC 5246, DOI 10.17487/RFC5246, August 2008, <<https://www.rfc-editor.org/info/rfc5246>>.
- [RFC5280] Cooper, D., Santesson, S., Farrell, S., Boeyen, S., Housley, R., and W. Polk, "Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile", RFC 5280, DOI 10.17487/RFC5280, May 2008, <<https://www.rfc-editor.org/info/rfc5280>>.
- [RFC6125] Saint-Andre, P. and J. Hodges, "Representation and Verification of Domain-Based Application Service Identity within Internet Public Key Infrastructure Using X.509 (PKIX) Certificates in the Context of Transport Layer Security (TLS)", RFC 6125, DOI 10.17487/RFC6125, March 2011, <<https://www.rfc-editor.org/info/rfc6125>>.
- [RFC7525] Sheffer, Y., Holz, R., and P. Saint-Andre, "Recommendations for Secure Use of Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS)", BCP 195, RFC 7525, DOI 10.17487/RFC7525, May 2015, <<https://www.rfc-editor.org/info/rfc7525>>.
- [RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <<https://www.rfc-editor.org/info/rfc8126>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

11.2. Informative References

- [github-dtn-bpbis-tcpcl]
Sipos, B., "TCPCL Example Implementation", <<https://github.com/BSipos-RKF/dtn-bpbis-tcpcl/tree/develop>>.

- [I-D.ietf-dtn-bpsec]
Birrane, E. and K. McKeever, "Bundle Protocol Security Specification", draft-ietf-dtn-bpsec-10 (work in progress), April 2019.
- [RFC2595] Newman, C., "Using TLS with IMAP, POP3 and ACAP", RFC 2595, DOI 10.17487/RFC2595, June 1999, <<https://www.rfc-editor.org/info/rfc2595>>.
- [RFC4838] Cerf, V., Burleigh, S., Hooke, A., Torgerson, L., Durst, R., Scott, K., Fall, K., and H. Weiss, "Delay-Tolerant Networking Architecture", RFC 4838, DOI 10.17487/RFC4838, April 2007, <<https://www.rfc-editor.org/info/rfc4838>>.
- [RFC7242] Demmer, M., Ott, J., and S. Perreault, "Delay-Tolerant Networking TCP Convergence-Layer Protocol", RFC 7242, DOI 10.17487/RFC7242, June 2014, <<https://www.rfc-editor.org/info/rfc7242>>.
- [RFC7457] Sheffer, Y., Holz, R., and P. Saint-Andre, "Summarizing Known Attacks on Transport Layer Security (TLS) and Datagram TLS (DTLS)", RFC 7457, DOI 10.17487/RFC7457, February 2015, <<https://www.rfc-editor.org/info/rfc7457>>.
- [RFC7942] Sheffer, Y. and A. Farrel, "Improving Awareness of Running Code: The Implementation Status Section", BCP 205, RFC 7942, DOI 10.17487/RFC7942, July 2016, <<https://www.rfc-editor.org/info/rfc7942>>.

Appendix A. Significant changes from RFC7242

The areas in which changes from [RFC7242] have been made to existing headers and messages are:

- o Split contact header into pre-TLS protocol negotiation and SESS_INIT parameter negotiation. The contact header is now fixed-length.
- o Changed contact header content to limit number of negotiated options.
- o Added session option to negotiate maximum segment size (per each direction).
- o Renamed "Endpoint ID" to "Node ID" to conform with BPv7 terminology.
- o Added session extension capability.

- o Added transfer extension capability. Moved transfer total length into an extension item.
- o Defined new IANA registries for message / type / reason codes to allow renaming some codes for clarity.
- o Segments of all new IANA registries are reserved for private/experimental use.
- o Expanded Message Header to octet-aligned fields instead of bit-packing.
- o Added a bundle transfer identification number to all bundle-related messages (XFER_SEGMENT, XFER_ACK, XFER_REFUSE).
- o Use flags in XFER_ACK to mirror flags from XFER_SEGMENT.
- o Removed all uses of SDNV fields and replaced with fixed-bit-length fields.
- o Renamed SHUTDOWN to SESS_TERM to deconflict term "shutdown" related to TCP connections.
- o Removed the notion of a re-connection delay parameter.

The areas in which extensions from [RFC7242] have been made as new messages and codes are:

- o Added contact negotiation failure SESS_TERM reason code.
- o Added MSG_REJECT message to indicate an unknown or unhandled message was received.
- o Added TLS session security mechanism.
- o Added Resource Exhaustion SESS_TERM reason code.

Authors' Addresses

Brian Sipos
RKF Engineering Solutions, LLC
7500 Old Georgetown Road
Suite 1275
Bethesda, MD 20814-6198
United States of America

Email: BSipos@rkf-eng.com

Michael Demmer
University of California, Berkeley
Computer Science Division
445 Soda Hall
Berkeley, CA 94720-1776
United States of America

Email: demmer@cs.berkeley.edu

Joerg Ott
Aalto University
Department of Communications and Networking
PO Box 13000
Aalto 02015
Finland

Email: ott@in.tum.de

Simon Perreault
Quebec, QC
Canada

Email: simon@per.reau.lt

Delay Tolerant Networking
Internet Draft
Intended status: Informational
Expires: January 2018

Taixin Li
Huachun Zhou
Qi Xu
Guanwen Li
Guanglei Li
Beijing Jiaotong University
May 31, 2017

software defined dtn-based satellite networks
draft-li-dtn-sd-dtn-sat-net-02.txt

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>

This Internet-Draft will expire on January 1, 2018.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Abstract

Delay/Disruption Tolerant Networking (DTN) is designed for a severe environment where communication quality is not guaranteed. It works as an overlay network associated with Bundle Protocol (BP) and some convergence layer protocols like Licklider Transmission Protocol (LTP). DTN is suitable for satellite networks. Because communication delay is long and peer-to-peer communication is not guaranteed in satellite networks. We implement SDN to solve the problems of controllable, manageable, and flexible in satellite networks. In this document, we propose a framework of Software Defined DTN-based satellite networks, using Bundle tunnel and protocol translation gateway.

Table of Contents

1. Introduction	3
2. Conventions used in this document	3
3. Key points of the design	3
3.1. Separated control plane and forwarding plane	4
3.2. Bundle tunnel	5
3.3. Satellite gateway	6
4. Use case	7
5. Security Considerations	8
6. IANA Considerations	8
7. Conclusions	8
8. References	9
8.1. Normative References	9
8.2. Informative References	9

9. Acknowledgments	9
--------------------------	---

1. Introduction

Delay/Disruption Tolerant Networking (DTN) [RFC4838] is designed for a severe environment where connectivity of network is intermittent and communication quality is not guaranteed. It works as an overlay network associated with Bundle Protocol (BP) [RFC5050] and convergence layer protocols like Licklider Transmission Protocol (LTP) [RFC5325] [RFC5326].

We implement DTN in the satellite networks to meet the need of high transmission delay with the help of Interplanetary Overlay Network (ION) [BURLEIGH07]. ION is an implementation of DTN architecture and is designed to enable inexpensive insertion of DTN functionality into embedded systems.

SDN [NUNES14] is a state-of-the-art network concept, introducing new possibilities for network management and configuration methods by decoupling the control decisions from forwarding hardware. A controller communicates with the switches by southbound interface, such as OpenFlow [LARA14], which is the core technology of SDN. We apply the idea of SDN to satellite network by separating control plane and forwarding plane in satellite network control structure and taking advantage of the global view of a controller.

In this document, we propose a framework of Software Defined DTN-based satellite networks, using Bundle tunnel to deploy OpenFlow over DTN and protocol translation gateway to achieve protocol translation between Bundle packets and IP packets.

2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

3. Key points of the design

The idea of SDN is applied in the proposed framework. The control link between the control plane and the forwarding plane is Bundle tunnel. Because we use DTN protocol stack in space network and the protocol stack of ground network is TCP/IP. There should be a

protocol translation gateway to achieve protocol translation between Bundle packets and IP packets.

3.1. Separated control plane and forwarding plane

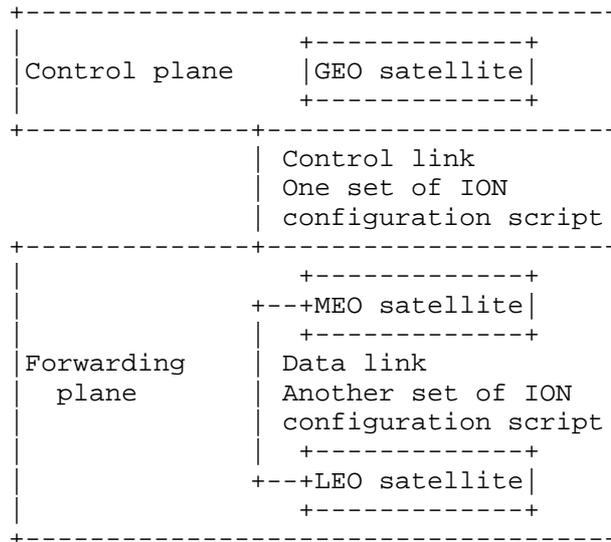


Figure 1 Illustration of the two planes

We apply the idea of SDN to satellite network by separating control plane and forwarding plane in satellite network control structure and taking advantage of the global view of a controller. The whole space network is divided into two parts, control plane and forwarding plane. The control plane contains Geosynchronous Earth Orbit (GEO) satellites, on which SDN controllers are deployed. The forwarding plane contains Medium Earth Orbit (MEO) satellites and Low Earth Orbit (LEO) satellites, and OpenFlow enabled switches are deployed on them.

We implement DTN with the help of ION in space network. The topology configuration mode of ION is reading the configuration scripts (.rc file), which contains the information of connections and nodes. As is shown in Figure 1, to achieve the goal of separating the control plane and forwarding plane in the space network, we adopt two set of unrelated ION configuration scripts when creating the topology. One is the script of Bundle tunnel (or we can say the control link). The

other one is the script of data link. Two ION processes run in the MEO/LEO satellite nodes without affecting each other.

3.2. Bundle tunnel

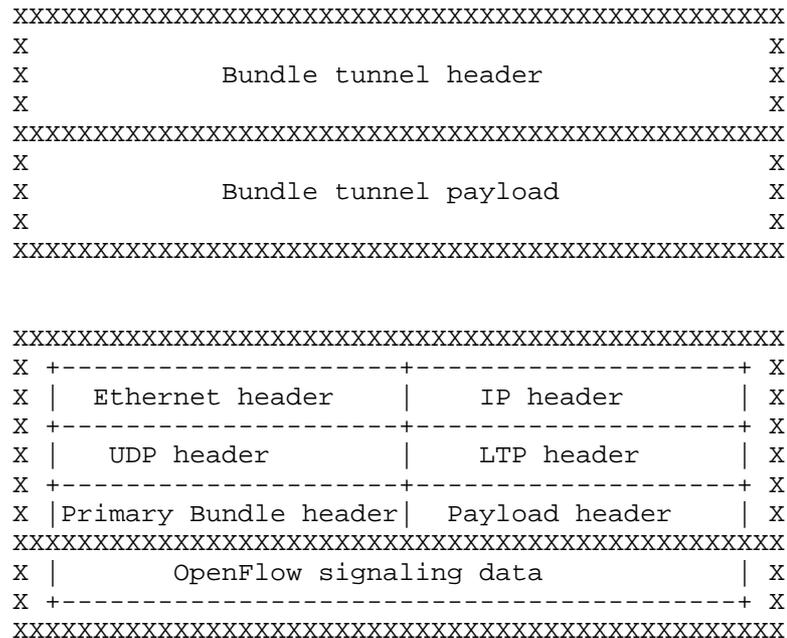


Figure 2 Illustration of bundle tunnel

We deploy OpenFlow over DTN by a method of tunnel. That is, signaling packets are transmitted in bundle tunnel when controller (GEO satellite) sets up connection to switches (MEO satellites) and when controllers send instructions to switches. The encapsulation format of the bundle tunnel is shown in Figure 2. Because DTN is implemented in ION in an overlay way, the first half of the Bundle tunnel header is the same as normal IP packets. The difference is that there are a 4-byte LTP header, a 14-byte Primary Bundle header, and a 5-byte Payload header before the payload data field due to the protocol stack of DTN. The link layer field is removed from the OpenFlow signaling packets between controller and switches and then the remaining fields are encapsulated in payload data.

The design of the Bundle tunnel adopts a dual process approach. One process is responsible for receiving OpenFlow signaling packets from

hierarchical, modular, and multi-process protocol translation function.

4. Use case

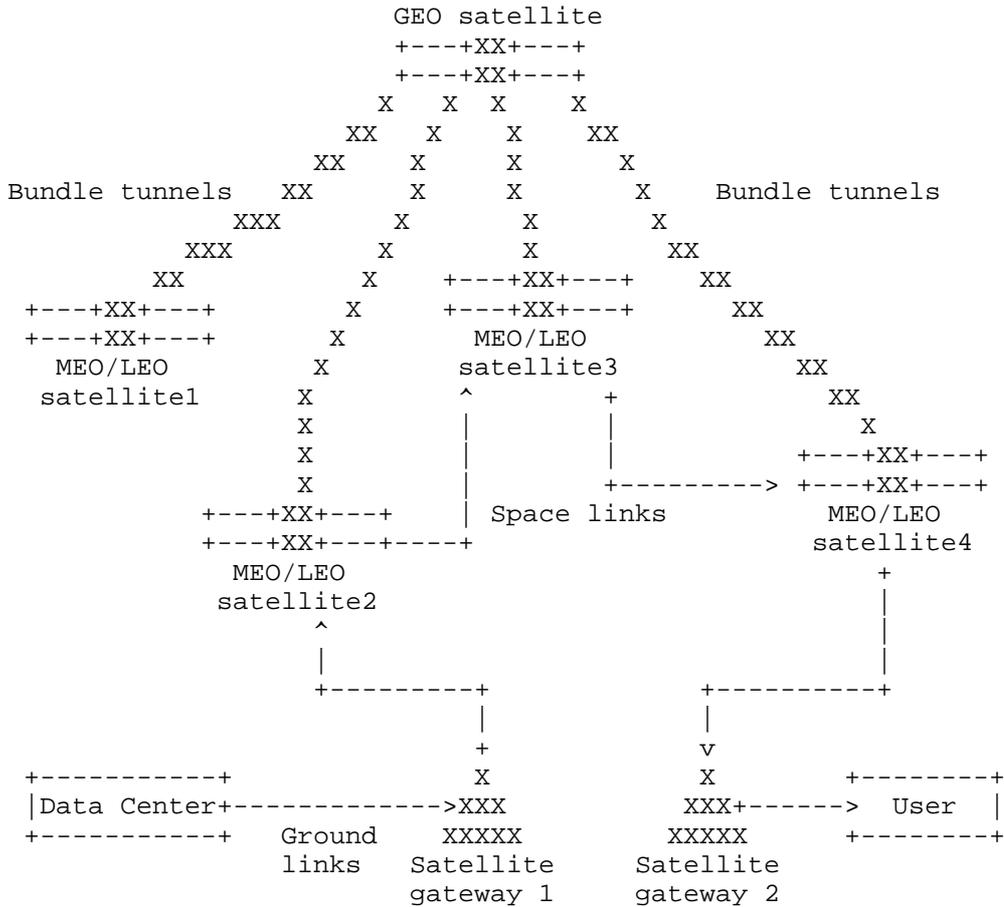


Figure 4 Use case of the proposed framework

The use case of proposed framework is shown in Figure 4. The GEO satellite set up control link to the four MEO/LEO satellites. The data center data are transmitted among the four MEO/LEO satellites. The ION configuration script of the control plane is about the connections of one GEO satellite to four MEO/LEO satellites. The ION configuration script of the forwarding plane is about the

connections among the four MEO/LEO satellites. That is to say, two sets of unrelated ION processes are running in the four MEO/LEO satellites.

A user applies for data from the data center via satellite networks. The traffic is sent to satellite gateway 1 and converted from IP packets to Bundle packets. The controller in GEO satellite send instructions to the MEO/LEO satellites and configure the flow tables of the switches in MEO/LEO satellites. Then the traffic is forwarded via the path: satellite2-->satellite3-->satellite4 under control of GEO satellite. Then, the traffic is sent to satellite gateway 2 and converted from Bundle packets to IP packets. Finally, the data are sent to the user.

5. Security Considerations

Introducing SDN in DTN-based space network can bring in some problems that any SDN-based frameworks have. The proposed framework adopts a centralized control architecture. So if GEO satellite is attacked (by viruses or physical attack), security problem should be considered. The possible solution may be reserving spare GEO satellite. When the GEO satellite in use breaks down, the spare one will take on the responsibility.

6. IANA Considerations

This document does not update or create any IANA registries.

7. Conclusions

This document describes the key points of the design of the proposed Software Defined DTN-based satellite networks framework: Separated control plane and forwarding plane in space network, Bundle tunnel, and satellite protocol translation gateway. And we describe the use case of the proposed framework in this document.

8. References

8.1. Normative References

- [RFC4838] Cerf, V., Burleigh, S., Hooke, A., Torgerson, L., Durst, R., Scott, K., Fall, K., and Weiss, H., "Delay-Tolerant Networking Architecture", RFC 4838, April 2007.
- [RFC5050] Scott, K., and Burleigh, S., "Bundle Protocol Specification", RFC 5050, RFC5050, November 2007.
- [RFC5325] Burleigh, S., Ramadas, M., and Farrell, S., "Licklider Transmission Protocol - Motivation", RFC 5325, September 2008.
- [RFC5326] Ramadas, M., Burleigh, S., and Farrell, S., "Licklider Transmission Protocol - Specification", RFC 5326, September 2008.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

8.2. Informative References

- [BURLEIGH07] Burleigh S. "Interplanetary Overlay Network: An Implementation of the DTN Bundle Protocol" Consumer Communications and NETWORKING Conference (CCNC), 2007, vol. 45(2). pp. 147-153, 2007.
- [NUNES14] B. Nunes, M. Mendonca, X. Nguyen, K. Obraczka, & T, Turletti, "A survey of software-defined networking: Past, present, and future of programmable networks," Communications Surveys & Tutorials, IEEE, vol. 16 (3), pp. 1617-1634, Feb. 2014.
- [LARA14] A. Lara, A. Kolasani, & B. Ramamurthy, "Network Innovation using Openflow: A Survey," Communications Surveys & Tutorials, IEEE, vol. 16(1), pp. 493-512, Feb. 2014.

9. Acknowledgments

This work in this document was supported by National High Technology of China ("863 program") under Grant No.2015AA015702.

Authors' Addresses

Taixin Li
Beijing Jiaotong University
Beijing, 100044, P.R. China

Email: 14111040@bjtu.edu.cn

Huachun Zhou
Beijing Jiaotong University
Beijing 100044, P.R. China

Email: hchzhou@bjtu.edu.cn

Qi Xu
Beijing Jiaotong University
Beijing, 100044, P.R. China

Email: 15111046@bjtu.edu.cn

Guanwen Li
Beijing Jiaotong University
Beijing, 100044, P.R. China

Email: 14120079@bjtu.edu.cn

Guanglei Li
Beijing Jiaotong University
Beijing, 100044, P.R. China

Email: 15111035@bjtu.edu.cn

