

Network Working Group
Internet-Draft
Intended status: Informational
Expires: April 30, 2017

B. Trammell, Ed.
M. Kuehlewind, Ed.
ETH Zurich
October 27, 2016

mPlane Protocol Specification
draft-trammell-mplane-protocol-02

Abstract

This document defines the mPlane protocol for coordination of heterogeneous network measurement Components: probes and repositories that measure, analyze, and store network measurements, data derived from measurements, and other ancillary data about elements of the network. The mPlane architecture is defined in terms of relationships between Components and Clients which communicate using the mPlane protocol defined in this document.

This revision of the document describes Version 2 of the mPlane 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 <http://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 30, 2017.

Copyright Notice

Copyright (c) 2016 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.

Table of Contents

1.	Introduction	3
1.1.	Changes from Revision -00 (Protocol version 1)	4
1.2.	Changes from Revision -01	4
2.	Terminology	4
3.	Overview of the mPlane Architecture	5
3.1.	Key Architectural Principles and Features	6
3.1.1.	Flexibility and Extensibility	6
3.1.2.	Schema-centric Measurement Definition	7
3.1.3.	Iterative Measurement Support	7
3.1.4.	Weak Imperativeness	7
3.2.	Entities and Relationships	8
3.2.1.	Components and Clients	9
3.2.2.	Probes and Repositories	9
3.3.	Message Types and Message Exchange Sequences	9
3.4.	Integrating Measurement Tools into mPlane	11
3.5.	From Architecture to Protocol Specification	11
4.	Protocol Information Model	11
4.1.	Element Registry	12
4.1.1.	Structured Element Names	13
4.1.2.	Primitive Types	14
4.1.3.	Augmented Registry Information	15
4.2.	Message Types	15
4.2.1.	Capability and Withdrawal	16
4.2.2.	Specification and Interrupt	16
4.2.3.	Result	16
4.2.4.	Receipt and Redemption	16
4.2.5.	Exception	17
4.2.6.	Envelope	17
4.3.	Message Sections	18
4.3.1.	Message Type and Verb	19
4.3.2.	Version	20
4.3.3.	Registry	20
4.3.4.	Label	20
4.3.5.	Temporal Scope (When)	21
4.3.6.	Parameters	22
4.3.7.	Metadata	23
4.3.8.	Result Columns and Values	24
4.3.9.	Export	24
4.3.10.	Link	25
4.3.11.	Token	25

4.3.12.	Contents	26
4.4.	Message Uniqueness and Idempotence	26
4.4.1.	Message Schema	26
4.4.2.	Message Identity	27
4.5.	Designing Measurement and Repository Schemas	27
5.	Representations and Session Protocols	28
5.1.	JSON representation	28
5.1.1.	Textual representations of element values	29
5.1.2.	Example mPlane Capabilities and Specifications in JSON	30
5.2.	mPlane over WebSockets over TLS	36
5.2.1.	mPlane PKI for WebSockets	37
5.2.2.	Capability Advertisement for WebSockets	37
5.2.3.	mPlane Link and Export URLs for WebSockets	37
6.	Deployment Considerations	38
6.1.	Supervisors and Federation	38
6.1.1.	Component Registration	40
6.1.2.	Client Authentication	40
6.1.3.	Capability Composition and Specification Decomposition	40
6.2.	Indirect Export	40
6.3.	Error Handling in mPlane Workflows	41
7.	Security Considerations	42
8.	IANA Considerations	42
9.	Contributors	42
10.	Acknowledgments	43
11.	Informative References	43
	Authors' Addresses	44

1. Introduction

This document describes the mPlane architecture and protocol, which is designed to provide control and coordination of heterogeneous network measurement tools. It is based heavily on the mPlane project's deliverable 1.4 [D14], and is submitted for the information of the Internet engineering community. [Section 3](#) gives an overview of the mPlane architecture, [Section 4](#) defines the protocol information model, and [Section 5](#) defines the representations of this data model and session protocols over which mPlane is supported.

Present work is focused on mPlane represented in JSON using WebSockets [RFC6455] as a session protocol.

This revision of the document describes Version 2 of the mPlane protocol.

1.1. Changes from Revision -00 (Protocol version 1)

- o WebSockets has been added as a session protocol for the mPlane protocol. Message flow has been modified to fit.
- o Redemptions have been expanded to allow temporally-scoped partial redemption of long-running measurements.
- o The "object" primitive type has been added to allow structured objects to be represented directly in mPlane messages using the underlying representation.
- o A number of obsolete features have been deprecated based on implementation and pilot deployment experience, and removed from the protocol description:
 - * HTTPS as a session protocol, as it is a poor fit for mPlane's exchanges
 - * Callback control, as it is no longer needed without the limitations of HTTPS as a session protocol
 - * The Indirection message type, as the same functionality is available for all message types using the link section.
 - * Repeating measurements, as their functionality has been replaced by partial redemption
- o References to the "core" registry have been removed, as all element registries must in any case be explicitly defined in mPlane Capabilities, Specifications, and Results. An eventual proposed standard mPlane protocol would refer to an IANA-managed core registry.

1.2. Changes from Revision -01

- o Editorial changes only.

2. Terminology

[EDITOR'S NOTE: these terms are not necessarily capitalized within the document at this time. Fix this.]

Client: An entity which implements the mPlane protocol, receives Capabilities published by one or more Components, and sends Specifications to those Component(s) to perform measurements and analysis. See [Section 3.2.1](#).

Component: An entity which implements the mPlane protocol specified within this document, advertises its Capabilities and accepts Specifications which request the use of those Capabilities. The measurements, analyses, storage facilities and other services provided by a Component are completely defined by its Capabilities. See [Section 3.2.1](#).

mPlane Message: The atomic unit of exchange in the mPlane protocol. The treatment of a message at a Client or Component receiving it is based upon its type; see [Section 4.2](#).

Capability: An mPlane message that contains a statement of a Component's ability and willingness to perform a specific operation, conveyed from a Component to a Client. A Capability does not represent a guarantee that the specific operation can or will be performed at a specific point in time. See [Section 4.2.1](#)

Specification: An mPlane message that contains a statement of a Client's desire that a Component should perform a specific operation, conveyed from a Client to a Component. It can be conceptually viewed as a Capability whose parameters have been filled in with values. See [Section 4.2.2](#).

Result: An mPlane message containing a statement produced by a Component that a particular measurement was taken and the given values were observed, or that a particular operation or analysis was performed and the given values were produced. It can be conceptually viewed as a Specification whose result columns have been filled in with values. See [Section 4.2.3](#).

Element: An identifier for a parameter or result column in a Capability, Specification, or Result, binding a name to a primitive type. Elements are contained in registries that contain the vocabulary from which mPlane Capabilities, Specifications, and Results can be built. See [Section 4.1](#).

3. Overview of the mPlane Architecture

mPlane is built around an architecture in which Components provide network measurement services and access to stored measurement data which they advertise via Capabilities completely describing these services and data. A Client makes use of these Capabilities by sending Specifications that respond to them back to the Components. Components may then either return Results directly to the Clients or sent to some third party via indirect export using an external protocol. The Capabilities, Specifications, and Results are carried over the mPlane protocol, defined in detail in this document. An

mPlane measurement infrastructure is built up from these basic blocks.

Components can be roughly classified into probes which generate measurement data and repositories which store and analyze measurement data, though the difference between a probe and a repository in the architecture is merely a matter of the Capabilities it provides. Components can be pulled together into an infrastructure by a supervisor (see [Section 6.1](#)), which presents a Client interface to subordinate Components and a Component interface to superordinate Clients, aggregating Capabilities into higher-level measurements and distributing Specifications to perform them.

The mPlane protocol is, in essence, a self-describing, error- and delay- tolerant remote procedure call (RPC) protocol between Clients and Components: each Capability exposes an entry point in the API provided by the Component; each Specification embodies an API call; and each Result returns the results of an API call. This arrangement is shown in Figure 1 below.

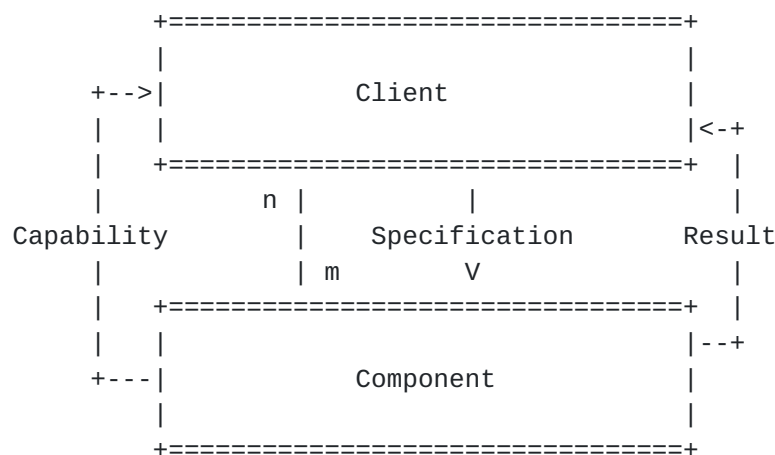


Figure 1: The simplest form of the mPlane architecture

[3.1.](#) Key Architectural Principles and Features

mPlane differs from a simple RPC facility in several important ways, detailed in the subsections below.

[3.1.1.](#) Flexibility and Extensibility

First, given the heterogeneity of the measurement tools and techniques applied, it is necessary for the protocol to be as flexible and extensible as possible. Therefore, the architecture in its simplest form consists of only two entities and one relationship, as shown in the diagram below: n Clients connect to m Components via

the mPlane protocol. Anything which can speak the mPlane protocol and exposes Capabilities thereby is a Component; anything which can understand these Capabilities and send Specifications to invoke them is a Client. Everything a Component can do, from the point of view of mPlane, is entirely described by its Capabilities. Capabilities are even used to expose optional internal features of the protocol itself, and provide a method for built-in protocol extensibility.

3.1.2. Schema-centric Measurement Definition

Second, given the flexibility required above, the key to measurement interoperability is the comparison of data types. Each Capability, Specification, and Result contains a schema, comprising the set of parameters required to execute a measurement or query and the columns in the data set that results. From the point of view of mPlane, the schema completely describes the measurement. This implies that when exposing a measurement using mPlane, the developer of a Component must build each Capabilities it advertises such that the semantics of the measurement are captured by the set of columns in its schema. The elements from which schemas can be built are captured in a type registry. The mPlane platform provides a core registry for common measurement use cases within the project, and the registry facility is itself fully extensible as well, for supporting new applications without requiring central coordination beyond the domain or set of domains running the application.

3.1.3. Iterative Measurement Support

Third, the exchange of messages in the protocol was chosen to support iterative measurement in which the aggregated, high-level results of a measurement are used as input to a decision process to select the next measurement. Specifically, the protocol blends control messages (Capabilities and Specifications) and data messages (Results) into a single workflow.

3.1.4. Weak Imperativeness

Fourth, the mPlane protocol is weakly imperative. A Capability represents a willingness and an ability to perform a given measurement or execute a query, but not a guarantee or a reservation to do so. Likewise, a Specification contains a set of parameters and a temporal scope for a measurement a Client wishes a Component to perform on its behalf, but execution of Specifications is best-effort. A Specification is not an instruction which must Result either in data or in an error. This property arises from our requirement to support large-scale measurement infrastructures with thousands of similar Components, including resource- and connectivity-limited probes such as smartphones and customer-premises

equipment (CPE) like home routers. These may be connected to a supervisor only intermittently. In this environment, the operability and conditions in which the probes find themselves may change more rapidly than can be practicably synchronized with a central supervisor; requiring reliable operation would compromise scalability of the architecture.

To support weak imperativeness, each message in the mPlane protocol is self-contained, and contains all the information required to understand the message. For instance, a Specification contains the complete information from the Capability which it responds to, and a Result contains its Specification. In essence, this distributes the state of the measurements running in an infrastructure across all Components, and any state resynchronization that is necessary after a disconnect happens implicitly as part of message exchange. The failure of a Component during a large-scale measurement can be detected and corrected after the fact, by examining the totality of the generated data.

This distribution of state throughout the measurement infrastructure carries with it a distribution of responsibility: a Component holding a Specification is responsible for ensuring that the measurement or query that the Specification describes is carried out, because the Client or supervisor which has sent the Specification does not necessarily keep any state for it.

Error handling in a weakly imperative environment is different to that in traditional RPC protocols. The exception facility provided by mPlane is designed only to report on failures of the handling of the protocol itself. Each Component and Client makes its best effort to interpret and process any authorized, well-formed mPlane protocol message it receives, ignoring those messages which are spurious or no longer relevant. This is in contrast with traditional RPC protocols, where even common exceptional conditions are signaled, and information about missing or otherwise defective data must be correlated from logs about measurement control. This traditional design pattern is not applicable in infrastructures where the supervisor has no control over the functionality and availability of its associated probes.

3.2. Entities and Relationships

The entities in the mPlane protocol and the relationships among them are described in more detail in the subsections below.

3.2.1. Components and Clients

Specifically, a Component is any entity which implements the mPlane protocol specified within this document, advertises its Capabilities and accepts Specifications which request the use of those Capabilities. The measurements, analyses, storage facilities and other services provided by a Component are completely defined by its Capabilities.

Conversely, a Client is any entity which implements the mPlane protocol, receives Capabilities published by one or more Components, and sends Specifications to those Component(s) to perform measurements and analysis.

Every interaction in the mPlane protocol takes place between a Component and a Client. Indeed, the simplest instantiation of the mPlane architecture consists of one or more Clients taking Capabilities from one or more Components, and sending Specifications to invoke those Capabilities, as shown in Figure 1. An mPlane domain may consist of as little as a single Client and a single Component. In this arrangement, mPlane provides a measurement-oriented RPC mechanism.

3.2.2. Probes and Repositories

Measurement Components can be roughly divided into two categories: probes and repositories. Probes perform measurements, and repositories provide access to stored measurements, analysis of stored measurements, or other access to related external data sources. External databases and data sources (e.g., routing looking glasses, WHOIS services, DNS, etc.) can be made available to mPlane Clients through repositories acting as gateways to these external sources, as well.

Note that this categorization is very rough: what a Component can do is completely described by its Capabilities, and some Components may combine properties of both probes and repositories.

3.3. Message Types and Message Exchange Sequences

The basic messages in the mPlane protocol are Capabilities, Specifications, and Results, as described above. The full protocol contains other message types as well. Withdrawals cancel Capabilities (i.e., indicate that the Component is no longer capable or willing to perform a given measurement) and interrupts cancel Specifications (i.e., indicate that the Component should stop performing the measurement). Receipts can be given in lieu of Results for not-yet completed measurements or queries, and

redemptions can be used to retrieve Results referred to by a receipt. Exceptions can be sent by Clients or Components at any time to signal protocol-level errors to their peers.

The following sequences of messages are possible within the mPlane protocol:



Figure 2: Possible sequences of messages in the mPlane protocol

In the nominal sequence, a Capability leads to a Specification leads to a Result, where Results may be transmitted by some other protocol. All the paths through the sequence of messages are shown in the diagram above; message types are described in detail in [Section 4.2](#).

Separate from the sequence of messages, the mPlane protocol supports two connection establishment patterns:

- o Client-initiated, in which Clients connect directly to Components at known, stable URLs. Client-initiated workflows are intended for use between Clients and supervisors, for access to repositories, and for access to probes embedded within a network infrastructure.
- o Component-initiated, in which Components initiate connections to Clients. Component-initiated workflows are intended for use between Components without stable routable addresses and

supervisors, e.g. for small probes on embedded devices, mobile devices, or software probes embedded in browsers on personal computers behind network-address translators (NATs) or firewalls which prevent a Client from establishing a connection to them.

Within a given mPlane domain, these patterns can be combined (along with indirect export and direct access) to facilitate complex interactions among Clients and Components according to the requirements imposed by the application and the deployment of Components in the network.

3.4. Integrating Measurement Tools into mPlane

mPlane's flexibility and the self-description of measurements provided by the Capability-Specification-Result cycle was designed to allow a wide variety of existing measurement tools, both probes and repositories, to be integrated into an mPlane domain. In both cases, the key to integration is to define a Capability for each of the measurements the tool can perform or the queries the repository needs to make available within an mPlane domain. Each Capability has a set of parameters - information required to run the measurement or the query - and a set of result columns - information which the measurement or query returns.

The parameters and result columns make up the measurement's schema, and are chosen from an extensible registry of elements. Practical details are given in [Section 4.5](#).

3.5. From Architecture to Protocol Specification

The remainder of this document builds the protocol Specification based on this architecture from the bottom up. First, we define the protocol's information model from the element registry through the types of mPlane messages and the sections they are composed of. We then define a concrete representation of this information model using Javascript Object Notation (JSON, [[RFC7159](#)]), and define bindings to WebSockets [[RFC6455](#)] over TLS as a session protocol.

4. Protocol Information Model

The mPlane protocol is message-oriented, built on the representation- and session-protocol-independent exchange of messages between Clients and Components. This section describes the information model, starting from the element registry which defines the elements from which Capabilities can be built, then detailing each type of message, and the sections that make these messages up. It then provides advice on using the information model to model measurements and queries.

4.1. Element Registry

An element registry makes up the vocabulary by which mPlane Components and Clients can express the meaning of parameters, metadata, and result columns for mPlane statements. A registry is represented as a JSON [[RFC7159](#)] object with the following keys:

- o registry-format: currently "mplane-0", determines the supported features of the registry format.
- o registry-uri: the URI identifying the registry. The URI must be dereferencable to retrieve the canonical version of this registry.
- o registry-revision: a serial number starting with 0 and incremented with each revision to the content of the registry.
- o includes: a list of URLs to retrieve additional registries from. Included registries will be evaluated in depth-first order, and elements with identical names will be replaced by registries parsed later.
- o elements: a list of objects, each of which has the following three keys:
 - * name: The name of the element.
- o prim: The name of the primitive type of the element, from the list of primitives in [Section 4.1.2](#).
 - * desc: An English-language description of the meaning of the element.

Since the element names will be used as keys in mPlane messages, mPlane binds to JSON, and JSON mandates lowercase key names, element names must use only lowercase letters.

An example registry with two elements and no includes follows:


```
{ "registry-format": "mplane-0",
  "registry-uri", "https://example.com/mplane/registry/core",
  "registry-revision": 0,
  "includes": [],
  "elements": [
    { "name": "full.structured.name",
      "prim": "string",
      "desc": "A representation of foo..."
    },
    { "name": "another.structured.name",
      "prim": "string",
      "desc": "A representation of bar..."
    },
  ],
}
```

Fully qualified element names consist of the element's name as an anchor after the URI from which the element came, e.g. "https://example.com/mplane/registry/core#full.structured.name". Elements within the type registry are considered globally equal based on their fully qualified names. However, within a given mPlane message, elements are considered equal based on unqualified names.

4.1.1.1. Structured Element Names

To ease understanding of mPlane type registries, element names are structured by convention; that is, an element name is made up of the following structural parts in order, separated by the dot ('.') character:

- o **basename:** exactly one, the name of the property the element specifies or measures. All elements with the same basename describe the same basic property. For example, all elements with basename "source" relate to the source of a packet, flow, active measurement, etc.; and elements with basename "delay" relate to the measured delay of an operation.
- o **modifier:** zero or more, additional information differentiating elements with the same basename from each other. Modifiers may associate the element with a protocol layer, or a particular variety of the property named in the basename. All elements with the same basename and modifiers refer to exactly the same property. Examples for the "delay" basename include "oneway" and "twoway", differentiating whether a delay refers to the path from the source to the destination or from the source to the source via the destination; and "icmp" and "tcp", describing the protocol used to measure the delay.

- o units: zero or one, present if the quantity can be measured in different units.
- o aggregation: zero or one, if the property is a metric derived from multiple singleton measurements. Supported aggregations are:
 - * "min": minimum value
 - * "max": maximum value
 - * "mean": mean value
 - * "sum": sum of values
 - * "NNpct" (where NN is a two-digit number 01-99): percentile
 - * "median": shorthand for and equivalent to "50pct".
 - * "count": count of values aggregated

When mapping mPlane structured names into contexts in which dots have special meaning (e.g. SQL column names or variable names in many programming languages), the dots may be replaced by underscores ('_'). When using external type registries (e.g. the IPFIX Information Element Registry), element names are not necessarily structured.

4.1.2. Primitive Types

The mPlane protocol supports the following primitive types for elements in the type registry:

- o string: a sequence of Unicode characters
- o natural: an unsigned integer
- o real: a real (floating-point) number
- o bool: a true or false (boolean) value
- o time: a timestamp, expressed in terms of UTC. The precision of the timestamp is taken to be unambiguous based on its representation.
- o address: an identifier of a network-level entity, including an address family. The address family is presumed to be implicit in the format of the message, or explicitly stored. Addresses may represent specific endpoints or entire networks.

- o url: a uniform resource locator
- o object: a structured object, serialized according to the serialization rules of the underlying representation.

4.1.3. Augmented Registry Information

Additional keys beyond prim, desc, and name may appear in an mPlane registry to augment information about each element. The following additional registry keys have been found useful by some implementors of the protocol:

- o units: If applicable, string describing the units in which the element is expressed; equal to the units part of a structured name if present.
- o ipfix-eid: The element ID of the equivalent IPFIX [[RFC7011](#)] Information Element.
- o ipfix-pen: The SMI Private Enterprise Number of the equivalent IPFIX [[RFC7011](#)] Information Element, if any.

4.2. Message Types

Workflows in mPlane are built around the Capability - Specification - Result cycle. Capabilities, Specifications, and Results are kinds of statements: a Capability is a statement that a Component can perform some action (generally a measurement); a Specification is a statement that a Client would like a Component to perform the action advertised in a Capability; and a Result is a statement that a Component measured a given set of values at a given point in time according to a Specification.

Other types of messages outside this nominal cycle are referred to as notifications. Types of notifications include Withdrawals, Interrupts, Receipts, Redemptions, and Exceptions. These notify Clients or Components of conditions within the measurement infrastructure itself, as opposed to directly containing information about measurements or observations.

Messages may also be grouped together into a single envelope message. Envelopes allow multiple messages to be represented within a single message, for example multiple Results pertaining to the same Receipt; and multiple Capabilities or Specifications to be transferred in a single transaction in the underlying session protocol.

The following types of messages are supported by the mPlane protocol:

4.2.1. Capability and Withdrawal

A Capability is a statement of a Component's ability and willingness to perform a specific operation, conveyed from a Component to a Client. It does not represent a guarantee that the specific operation can or will be performed at a specific point in time.

A withdrawal is a notification of a Component's inability or unwillingness to perform a specific operation. It cancels a previously advertised Capability. A withdrawal can also be sent in reply to a Specification which attempts to invoke a Capability no longer offered.

4.2.2. Specification and Interrupt

A Specification is a statement that a Component should perform a specific operation, conveyed from a Client to a Component. It can be conceptually viewed as a Capability whose parameters have been filled in with values.

An interrupt is a notification that a Component should stop performing a specific operation, conveyed from Client to Component. It terminates a previously sent Specification. If the Specification uses indirect export, the indirect export will simply stop running. If the Specification has pending Results, those Results are returned in response to the interrupt.

4.2.3. Result

A Result is a statement produced by a Component that a particular measurement was taken and the given values were observed, or that a particular operation or analysis was performed and the given values were produced. It can be conceptually viewed as a Specification whose Result columns have been filled in with values. Note that, in keeping with the stateless nature of the mPlane protocol, a Result contains the full set of parameters from which it was derived.

Note that not every Specification will lead to a Result being returned; for example, in case of indirect export, only a receipt which can be used for future interruption will be returned, as the results will be conveyed to a third Component using an external protocol.

4.2.4. Receipt and Redemption

A receipt is returned instead of a Result by a Component in response to a Specification which either:

- o will never return results, as it initiated an indirect export, or
- o will not return results immediately, as the operation producing the results will have a long run time.

Receipts have the same content Specification they are returned for. A Component may optionally add a token section, which can be used in future redemptions or interruptions by the Client. The content of the token is an opaque string generated by the Component.

A redemption is sent from a Client to a Component for a previously received receipt to attempt to retrieve delayed results. It may contain only the token section, the token and temporal scope, or all sections of the received receipt.

When the temporal scope of a redemption for a running measurement is different than the temporal scope of the original Specification, it is treated by the Component as a partial redemption: all rows resulting from the measurement within the specified temporal scope are returned as a Result. Otherwise, a Component responds with a Result only when the measurement is complete; otherwise, another receipt is returned.

Note that redemptions are optional: when a Component has a Result available for a Client after some period of time, it may send it immediately, without waiting for a redemption to retrieve it.

[4.2.5.](#) Exception

An exception is sent from a Client to a Component or from a Component to a Client to signal an exceptional condition within the infrastructure itself. They are not meant to signal exceptional conditions within a measurement performed by a Component; see [Section 6.3](#) for more. An exception contains only two sections: an optional token referring back to the message to which the exception is related (if any), and a message section containing free-form, preferably human readable information about the exception.

[4.2.6.](#) Envelope

An envelope is used to contain other messages. Message containment is necessary in contexts in which multiple mPlane messages must be grouped into a single transaction in the underlying session protocol. It is legal to group any kind of message, and to mix messages of different types, in an envelope. However, in the current revision of the protocol, envelopes are primarily intended to be used for three distinct purposes:

- o To group multiple Capabilities together within a single message (e.g., all the Capabilities a given Component has).
- o To return multiple Results for a single receipt or Specification.
- o To group multiple Specifications into a single message.

4.3. Message Sections

Each message is made up of sections, as described in the subsection below. The following table shows the presence of each of these sections in each of the message types supported by mPlane: "req" means the section is required, "opt" means it is optional, and "tok" means it may be replaced by reference via a token when available (i.e., in withdrawals and redemptions); see the subsection on each message section for details.

Section	Cap.	Spec.	Result	Receipt	Envelope
Verb	req	req	req.	req.	
Content Type					req
"version"	req	req	req	req	req
"registry"	req	req	req	opt	
"label"	opt	opt	opt	opt	opt
"when"	req	req	req	req	
"parameters"	req/tok	req	req	opt/tok	
"metadata"	opt/tok	opt	opt	opt/tok	
"results"	req/tok	req	req	opt/tok	
"resultvalues"			req		
"export"	opt	opt	opt	opt	
"link"	opt	opt			
"token"	opt	opt	opt	opt	opt
"contents"					req

Table 1: Message Sections for Each Message Type

Withdrawals take the same sections as Capabilities; and redemptions and interrupts take the same sections as receipts. Exceptions are not shown in this table.

[4.3.1.](#) Message Type and Verb

The verb is the action to be performed by the Component. The following verbs are supported by the base mPlane protocol, but arbitrary verbs may be specified by applications:

- o "measure": Perform a measurement
- o "query": Query a database about a past measurement

- o "collect": Receive Results via indirect export
- o "callback": Used for callback control in Component-initiated workflows

In the JSON representation of mPlane messages, the verb is the value of the key corresponding to the message's type, represented as a lowercase string (e.g. "Capability", "Specification", "result" and so on).

Roughly speaking, probes implement "measure" Capabilities, and repositories implement "query" and "collect" Capabilities. Of course, any single Component can implement Capabilities with any number of different verbs.

Within the SDK, the primary difference between "measure" and "query" is that the temporal scope of a "measure" Specification is taken to refer to when the measurement should be scheduled, while the temporal scope of a "query" Specification is taken to refer to the time window (in the past) of a query.

Envelopes have no verb; instead, the value of the "envelope" key is the kind of messages the envelope contains, or "message" if the envelope contains a mixture of different unspecified kinds of messages.

4.3.2. Version

The "version" section contains the version of the mPlane protocol to which the message conforms, as an integer serially incremented with each new protocol revision. This section is required in all messages. This document describes version 2 of the protocol.

4.3.3. Registry

The "registry" section contains the URL identifying the element registry used by this message, and from which the registry can be retrieved. This section is required in all messages containing element names (statements, and receipts/redemptions/interrupts not using tokens for identification; see the "token" section).

4.3.4. Label

The "label" section of a statement contains a human-readable label identifying it, intended solely for use when displaying information about messages in user interfaces. Results, receipts, redemptions, and interrupts inherit their label from the Specification from which they follow; otherwise, Client and Component software can arbitrarily

assign labels . The use of labels is optional in all messages, but as labels do greatly ease human-readability of arbitrary messages within user interfaces, their use is recommended.

mPlane Clients and Components should never use the label as a unique identifier for a message, or assume any semantic meaning in the label - the test of message equality and relatedness is always based upon the schema and values as in [Section 4.4](#).

4.3.5. Temporal Scope (When)

The "when" section of a statement contains its temporal scope.

A temporal scope refers to when a measurement can be run (in a Capability), when it should be run (in a Specification), or when it was run (in a Result). Temporal scopes can be either absolute or relative, and may have an optional period, referring to how often single measurements should be taken.

The general form of a temporal scope (in BNF-like syntax) is as follows:

```
when = <singleton> |                                # A single point in time
      <range> |                                       # A range in time
      <range> ' / ' <duration> # A range with a period

singleton = <iso8601> | # absolute singleton
            'now'      # relative singleton

range = <iso8601> ' ... ' <iso8601> | # absolute range
        <iso8601> ' + ' <duration> | # relative range
        'now' ' ... ' <iso8601> |   # definite future
        'now' ' + ' <duration> |   # relative future
        <iso8601> ' ... ' 'now' |   # definite past
        'past ... now' |           # indefinite past
        'now ... future' |         # indefinite future
        <iso8601> ' ... ' 'future' | # absolute indefinite future
        'past ... future' |       # forever

duration = [ <n> 'd' ] # days
           [ <n> 'h' ] # hours
           [ <n> 'm' ] # minute
           [ <n> 's' ] # seconds
```

```
iso8601 = <n> '-' <n> '-' <n> [ ' ' <n> ':' <n> ':' <n> [ '.' <n> ] ]
```

All absolute times are always given in UTC and expressed in ISO8601 format with variable precision.

In Capabilities, if a period is given it represents the minimum period supported by the measurement; this is done to allow rate limiting. If no period is given, the measurement is not periodic. A Capability with a period can only be fulfilled by a Specification with period greater than or equal to the period in the Capability. Conversely, a Capability without a period can only be fulfilled by a Specification without a period.

Within a Result, only absolute ranges are allowed within the temporal scope, and refers to the time range of the measurements contributing to the Result. Note that the use of absolute times here implies that the Components and Clients within a domain should have relatively well-synchronized clocks, e.g., to be synchronized using the Network Time Protocol [[RFC5905](#)] in order for Results to be temporally meaningful.

So, for example, an absolute range in time might be expressed as:

```
"when: 2009-02-20 13:02:15 ... 2014-04-04 04:27:19"
```

A relative range covering three and a half days might be:

```
"when: 2009-04-04 04:00:00 + 3d12h"
```

In a Specification for running an immediate measurement for three hours every seven and a half minutes:

```
"when: now + 3h / 7m30s"
```

In a Capability noting that a Repository can answer questions about the past:

```
"when: past ... now".
```

In a Specification requesting that a measurement run from a specified point in time until interrupted:

```
"when: 2017-11-23 18:30:00 ... future"
```

4.3.6. Parameters

The "parameters" section of a message contains an ordered list of the parameters for a given measurement: values which must be provided by a Client to a Component in a Specification to convey the specifics of the measurement to perform. Each parameter in an mPlane message is a key-value pair, where the key is the name of an element from the element registry. In Specifications and Results, the value is the value of the parameter. In Capabilities, the value is a constraint

on the possible values the Component will accept for the parameter in a subsequent Specification.

Four kinds of constraints are currently supported for mPlane parameters:

- o No constraint: all values are allowed. This is signified by the special constraint string `""`.
- o Single value constraint: only a single value is allowed. This is intended for use for Capabilities which are conceivably configurable, but for which a given Component only supports a single value for a given parameter due to its own out-of-band configuration or the permissions of the Client for which the Capability is valid. For example, the source address of an active measurement of a single-homed probe might be given as `"source.ip4: 192.0.2.19"`.
- o Set constraint: multiple values are allowed, and are explicitly listed, separated by the `","` character. For example, a multi-homed probe allowing two potential source addresses on two different networks might be given as `"source.ip4: 192.0.2.19, 192.0.3.21"`.
- o Range constraint: multiple values are allowed, between two ordered values, separated by the special string `"..."`. Range constraints are inclusive. A measurement allowing a restricted range of source ports might be expressed as `"source.port: 32768 ... 65535"`.
- o Prefix constraint: multiple values are allowed within a single network, as specified by a network address and a prefix. A prefix constraint may be satisfied by any network of host address completely contained within the prefix. An example allowing probing of any host within a given /24 might be `"destination.ip4: 192.0.2.0/24"`.

Parameter and constraint values must be a representation of an instance of the primitive type of the associated element.

[4.3.7](#). Metadata

The "metadata" section contains measurement metadata: key-value pairs associated with a Capability inherited by its Specification and Results. Metadata can also be thought of as immutable parameters. This is intended to represent information which can be used to make decisions at the Client as to the applicability of a given Capability (e.g. details of algorithms used or implementation-specific

information) as well as to make adjustments at post-measurement analysis time when contained within Results.

An example metadata element might be `"measurement.identifier: qof"`, which identifies the underlying tool taking measurements, such that later analysis can correct for known peculiarities in the implementation of the tool. Another example might be `"location.longitude = 8.55272"`, which while not particularly useful for analysis purposes, can be used to draw maps of measurements.

4.3.8. Result Columns and Values

Results are represented using two sections: "results" which identify the elements to be returned by the measurement, and "resultvalues" which contains the actual values. "results" appear in all statements, while "resultvalues" appear only in Result messages.

The "results" section contains an ordered list of result columns for a given measurement: names of elements which will be returned by the measurement. The result columns are identified by the names of the elements from the element registry.

The "resultvalues" section contains an ordered list of ordered lists (or, rather, a two dimensional array) of values of results for a given measurement, in row-major order. The columns in the result values appear in the same order as the columns in the "results" section.

Values for each column must be a representation of an instance of the primitive type of the associated result column element.

4.3.9. Export

The "export" section contains a URL or partial URL for indirect export. Its meaning depends on the kind and verb of the message:

- o For Capabilities with the "collect" verb, the "export" section contains the URL of the collector which can accept indirect export for the schema defined by the "parameters" and "results" sections of the Capability, using the protocol identified by the URL's schema.
- o For Capabilities with any verb other than "collect", the "export" section contains either the URL of a collector to which the Component can indirectly export results, or a URL schema identifying a protocol over which the Component can export to arbitrary collectors.

- o For Specifications with any verb other than "collect", the "export" section contains the URL of a collector to which the Component should indirectly export results. A receipt will be returned for such Specifications.

If a Component can indirectly export or indirectly collect using multiple protocols, each of those protocols must be identified by its own Capability; Capabilities with an "export" section can only be used by Specifications with a matching "export" section.

4.3.10. Link

The "link" section contains the URL to which messages in the next step in the workflow (i.e. a Specification for a Capability, a Result or receipt for a Specification) can be sent, providing indirection. The link URL must currently have the schema "wss", and refers to the URL to which to initiate a connection.

If present in a Capability, the Client must send Specifications for the given Capability to the Component at the URL given in order to use the Capability, connecting to the URL if no connection is currently established. If present in a Specification, the Component must send Results for the given Specification back to the Client at the URL given, connecting to the URL if no connection is currently established.

4.3.11. Token

The "token" section contains an arbitrary string by which a message may be identified in subsequent communications in an abbreviated fashion. Unlike labels, tokens are not necessarily intended to be human-readable; instead, they provide a way to reduce redundancy on the wire by replacing the parameters, metadata, and results sections in messages within a workflow, at the expense of requiring more state at Clients and Components. Their use is optional.

Tokens are scoped to the association between the Component and Client in which they are first created; i.e., at a Component, the token will be associated with the Client's identity, and vice-versa at a Client. Tokens should be created with sufficient entropy to avoid collision from independent processes at the same Client or token reuse in the case of Client or Component state loss at restart.

If a Capability contains a token, it may be subsequently withdrawn by the same Component using a withdrawal containing the token instead of the parameters, metadata, and results sections.

If a Specification contains a token, it may be answered by the Component with a receipt containing the token instead of the parameters, metadata, and results sections. A Specification containing a token may likewise be interrupted by the Client with an interrupt containing the token. A Component must not answer a Specification with a token with a receipt or Result containing a different token, but the token may be omitted in subsequent receipts and Results.

If a receipt contains a token, it may be redeemed by the same Client using a redemption containing the token instead of the parameters, metadata, and results sections.

4.3.12. Contents

The "contents" section appears only in envelopes, and is an ordered list of messages. If the envelope's kind identifies a message kind, the contents may contain only messages of the specified kind, otherwise if the kind is "message", the contents may contain a mix of any kind of message.

4.4. Message Uniqueness and Idempotence

Messages in the mPlane protocol are intended to support state distribution: Capabilities, Specifications, and Results are meant to be complete declarations of the state of a given measurement. In order for this to hold, it must be possible for messages to be uniquely identifiable, such that duplicate messages can be recognized. With one important exception (i.e., Specifications with relative temporal scopes), messages are idempotent: the receipt of a duplicate message at a Client or Component is a null operation.

4.4.1. Message Schema

The combination of elements in the "parameters" and "results" sections, together with the registry from which these elements are drawn, is referred to as a message's schema. The schema of a measurement can be loosely thought of as the definition of the table, rows of which the message represents.

The schema contributes not only to the identity of a message, but also to the semantic interpretation of the parameter and result values. The meanings of element values in mPlane are dependent on the other elements present in the message; in other words, the key to interpreting an mPlane message is that the unit of semantic identity is a message. For example, the element "destination.ip4" as a parameter means "the target of a given active measurement" when together with elements describing an active metric (e.g.

"delay.twoway.icmp.us")), but "the destination of the packets in a flow" when together with other elements in result columns describing a passively-observed flow.

The interpretation of the semantics of an entire message is application-specific. The protocol does not forbid the transmission of messages representing semantically meaningless or ambiguous schemas.

4.4.2. Message Identity

A message's identity is composed of its schema, together with its temporal scope, metadata, parameter values, and indirect export properties. Concretely, the full content of the "registry", "when", "parameters", "metadata", "results", and "export" sections taken together comprise the message's identity.

One convenience feature complicates this somewhat: when the temporal scope is not absolute, multiple Specifications may have the same literal temporal scope but refer to different measurements. In this case, the current time at the Client or Component when a message is invoked must be taken as part of the message's identity as well. Implementations may use hashes over the values of the message's identity sections to uniquely identify messages; e.g. to generate message tokens.

4.5. Designing Measurement and Repository Schemas

As noted, the key to integrating a measurement tool into an mPlane infrastructure is properly defining the schemas for the measurements and queries it performs, then defining those schemas in terms of mPlane Capabilities. Specifications and Results follow naturally from Capabilities, and the mPlane SDK allows Python methods to be bound to Capabilities in order to execute them. A schema should be defined such that the set of parameters, the set of result columns, and the verb together naturally and uniquely define the measurement or the query being performed. For simple metrics, this is achieved by encoding the entire meaning of the metric in its name. For example, "delay.twoway.icmp.us" as a result column together with "source.ip4" and "destination.ip4" as parameters uniquely defines a single ping measurement, measured via ICMP, expressed in microseconds.

Aggregate measurements are defined by returning metrics with aggregations: "delay.twoway.icmp.min.us", "delay.twoway.icmp.max.us", "delay.twoway.icmp.mean.us", and "delay.twoway.icmp.count.us" as result columns represent aggregate ping measurements with multiple samples.

Note that mPlane Results may contain multiple rows. In this case, the parameter values in the Result, taken from the Specification, apply to all rows. In this case, the rows are generally differentiated by the values in one or more result columns; for example, the "time" element can be used to represent time series, or the "hops.ip" different elements along a path between source and destination, as in a traceroute measurement.

For measurements taken instantaneously, the verb "measure" should be used; for direct queries from repositories, the verb "query" should be used. Other actions that cannot be differentiated by schema alone should be differentiated by a custom verb.

When integrating a repository into an mPlane infrastructure, only a subset of the queries the repository can perform will generally be exposed via the mPlane interface. Consider a generic repository which provides an SQL interface for querying data; wrapping the entire set of possible queries in specific Capabilities would be impossible, while providing direct access to the underlying SQL (for instance, by creating a custom registry with a "query.sql" string element to be used as a parameter) would make it impossible to differentiate Capabilities by schema (thereby making the interoperability benefits of mPlane integration pointless). Instead, specific queries to be used by Clients in concert with Capabilities provided by other Components are each wrapped within a separate Capability, analogous to stored procedure programming in many database engines. Of course, Clients which do speak the precise dialect of SQL necessary can integrate directly with the repository separate from the Capabilities provided over mPlane.

5. Representations and Session Protocols

The mPlane protocol is built atop an abstract data model in order to support multiple representations and session protocols. The canonical representation supported by the present SDK involves JSON [\[RFC7159\]](#) objects transported via Websockets [\[RFC6455\]](#) over TLS [\[RFC5246\]](#) (known by the "wss" URL schema).

5.1. JSON representation

In the JSON representation, an mPlane message is a JSON object, mapping sections by name to their contents. The name of the message type is a special section key, which maps to the message's verb, or to the message's content type in the case of an envelope.

Each section name key in the object has a value represented in JSON as follows:

- o "version" : an integer identifying the mPlane protocol version used by the message.
- o "registry" : a URL identifying the registry from which element names are taken.
- o "label" : an arbitrary string.
- o "when" : a string containing a temporal scope, as described in [Section 4.3.5](#).
- o "parameters" : a JSON object mapping (non-qualified) element names, either to constraints or to parameter values, as appropriate, and as described in [Section 4.3.6](#).
- o "metadata" : a JSON object mapping (non-qualified) element names to metadata values.
- o "results" : an array of element names.
- o "resultvalues" : an array of arrays of element values in row major order, where each row array contains values in the same order as the element names in the "results" section.
- o "export" : a URL for indirect export.
- o "link" : a URL for message indirection.
- o "token" : an arbitrary string.
- o "contents" : an array of objects containing messages.

[5.1.1.1](#). Textual representations of element values

Each primitive type is represented as a value in JSON as follows, following the Textual Representation of IPFIX Abstract Data Types defined in [\[RFC7373\]](#).

Natural and real values are represented in JSON using native JSON representation for numbers.

Booleans are represented by the reserved words "true" and "false".

Strings and URLs are represented as JSON strings, subject to JSON escaping rules.

Addresses are represented as dotted quads for IPv4 addresses as they would be in URLs, and canonical IPv6 textual addresses as in section

2.2 of [\[RFC4291\]](#) as updated by [section 4 of \[RFC5952\]](#). When representing networks, addresses may be suffixed as in CIDR notation, with a '/' character followed by the mask length in bits *n*, provided that the least significant 32 - *n* or 128 - *n* bits of the address are zero, for IPv4 and IPv6 respectively.

Timestamps are represented in [\[RFC3339\]](#) and ISO 8601, with two important differences. First, all mPlane timestamps are expressed in terms of UTC, so time zone offsets are neither required nor supported, and are always taken to be 0. Second, fractional seconds are represented with a variable number of digits after an optional decimal point after the fraction.

Objects are represented as JSON objects.

5.1.2. Example mPlane Capabilities and Specifications in JSON

To illustrate how mPlane messages are encoded, we consider first two Capabilities for a very simple application - ping - as mPlane JSON Capabilities. The following Capability states that the Component can measure ICMP two-way delay from 192.0.2.19 to anywhere on the IPv4 Internet, with a minimum delay between individual pings of 1 second, returning aggregate statistics:

```
{
  "capability": "measure",
  "version":    0,
  "registry":   "https://example.com/mplane/registry/core",
  "label":      "ping-aggregate",
  "when":       "now ... future / 1s",
  "parameters": {"source.ip4":      "192.0.2.19",
                 "destination.ip4": "*"},
  "results":    ["delay.twoway.icmp.us.min",
                 "delay.twoway.icmp.us.mean",
                 "delay.twoway.icmp.us.50pct",
                 "delay.twoway.icmp.us.max",
                 "delay.twoway.icmp.count"]
}
```

In contrast, the following Capability would return timestamped singleton delay measurements given the same parameters:


```
{
  "capability": "measure",
  "version":    0,
  "registry":   "https://example.com/mplane/registry/core",
  "label":      "ping-singletons",
  "when":       "now ... future / 1s",
  "parameters": {"source.ip4":      "192.0.2.19",
                 "destination.ip4": "*"},
  "results":    ["time",
                 "delay.twoway.icmp.us"]
}
```

A Specification is merely a Capability with filled-in parameters, e.g.:

```
{
  "specification": "measure",
  "version":       0,
  "registry":      "https://example.com/mplane/registry/core",
  "label":         "ping-aggregate-three-thirtythree",
  "token":         "0f31c9033f8fce0c9be41d4942c276e4",
  "when":          "now + 30s / 1s",
  "parameters": {"source.ip4":      "192.0.2.19",
                 "destination.ip4": "192.0.3.33"},
  "results":       ["delay.twoway.icmp.us.min",
                    "delay.twoway.icmp.us.mean",
                    "delay.twoway.icmp.us.50pct",
                    "delay.twoway.icmp.us.max",
                    "delay.twoway.icmp.count"]
}
```

Results are merely Specifications with result values filled in and an absolute temporal scope:


```
{
  "result": "measure",
  "version": 0,
  "registry": "https://example.com/mplane/registry/core",
  "label": "ping-aggregate-three-thirtythree",
  "token": "0f31c9033f8fce0c9be41d4942c276e4",
  "when": "2014-08-25 14:51:02.623 ... 2014-08-25 14:51:32.701 / 1s",
  "parameters": {"source.ip4": "192.0.2.19",
                 "destination.ip4": "192.0.3.33"},
  "results": ["delay.twoway.icmp.us.min",
              "delay.twoway.icmp.us.mean",
              "delay.twoway.icmp.us.50pct",
              "delay.twoway.icmp.us.max",
              "delay.twoway.icmp.count"],
  "resultvalues": [ [ 23901,
                      29833,
                      27619,
                      66002,
                      30] ]
}
```

More complex measurements can be modeled by mapping them back to tables with multiple rows. For example, a traceroute Capability would be defined as follows:

```
{
  "capability": "measure",
  "version": 0,
  "registry": "https://example.com/mplane/registry/core",
  "label": "traceroute",
  "when": "now ... future / 1s",
  "parameters": {"source.ip4": "192.0.2.19",
                 "destination.ip4": "*",
                 "hops.ip.max": "0..32"},
  "results": ["time",
              "intermediate.ip4",
              "hops.ip",
              "delay.twoway.icmp.us"]
}
```

with a corresponding Specification:


```
{
  "specification": "measure",
  "version":      0,
  "registry":     "https://example.com/mplane/registry/core",
  "label":        "traceroute-three-thirtythree",
  "token":        "2f4123588b276470b3641297ae85376a",
  "when":         "now",
  "parameters":  {"source.ip4":      "192.0.2.19",
                  "destination.ip4": "192.0.3.33",
                  "hops.ip.max": 32},
  "results":      ["time",
                  "intermediate.ip4",
                  "hops.ip",
                  "delay.twoway.icmp.us"]
}
```

and an example result:

```
{
  "result": "measure",
  "version": 0,
  "registry": "https://example.com/mplane/registry/core",
  "label":    "traceroute-three-thirtythree",
  "token":    "2f4123588b276470b3641297ae85376a",
  "when":     "2014-08-25 14:53:11.019 ... 2014-08-25 14:53:12.765",
  "parameters": {"source.ip4":      "192.0.2.19",
                 "destination.ip4": "192.0.3.33",
                 "hops.ip.max": 32},
  "results":     ["time",
                 "intermediate.ip4",
                 "hops.ip",
                 "delay.twoway.icmp.us"],
  "resultvalues": [ [ "2014-08-25 14:53:11.019", "192.0.2.1",
                      1, 162 ],
                    [ "2014-08-25 14:53:11.220", "217.147.223.101",
                      2, 15074 ],
                    [ "2014-08-25 14:53:11.570", "77.109.135.193",
                      3, 30093 ],
                    [ "2014-08-25 14:53:12.091", "77.109.135.34",
                      4, 34979 ],
                    [ "2014-08-25 14:53:12.310", "192.0.3.1",
                      5, 36120 ],
                    [ "2014-08-25 14:53:12.765", "192.0.3.33",
                      6, 36202 ]
                  ]
}
```


Indirect export to a repository with subsequent query requires three Capabilities: one in which the repository advertises its ability to accept data over a given external protocol, one in which the probe advertises its ability to export data of the same type using that protocol, and one in which the repository advertises its ability to answer queries about the stored data. Returning to the aggregate ping measurement, first let's consider a repository which can accept these measurements via direct POST of mPlane Result messages:

```
{
  "capability": "collect",
  "version": 0,
  "registry": "https://example.com/mplane/registry/core",
  "label": "ping-aggregate-collect",
  "when": "past ... future",
  "export": "wss://repository.example.com:4343/",
  "parameters": {"source.ip4": "*",
                 "destination.ip4": "*"},
  "results": ["delay.twoway.icmp.us.min",
              "delay.twoway.icmp.us.mean",
              "delay.twoway.icmp.us.50pct",
              "delay.twoway.icmp.us.max",
              "delay.twoway.icmp.count"]
}
```

This Capability states that the repository at "wss://repository.example.com:4343/" will accept mPlane Result messages matching the specified schema, without any limitations on time. Note that this schema matches that of the export Capability provided by the probe:

```
{
  "capability": "measure",
  "version": 0,
  "registry": "https://example.com/mplane/registry/core",
  "label": "ping-aggregate-export",
  "when": "now ... future / 1s",
  "export": "wss",
  "parameters": {"source.ip4": "192.0.2.19",
                 "destination.ip4": "*"},
  "results": ["delay.twoway.icmp.us.min",
              "delay.twoway.icmp.us.mean",
              "delay.twoway.icmp.us.50pct",
              "delay.twoway.icmp.us.max",
              "delay.twoway.icmp.count"]
}
```


which differs only from the previous probe Capability in that it states that Results can be exported via the "wss" protocol. Subsequent queries can be sent to the repository in response to the query Capability:

```
{
  "capability": "query",
  "version": 0,
  "registry": "https://example.com/mplane/registry/core",
  "label": "ping-aggregate-query",
  "when": "past ... future",
  "link": "wss://repo.example.com:4343/",
  "parameters": {"source.ip4": "*",
                 "destination.ip4": "*"},
  "results": ["delay.twoway.icmp.us.min",
              "delay.twoway.icmp.us.mean",
              "delay.twoway.icmp.us.50pct",
              "delay.twoway.icmp.us.max",
              "delay.twoway.icmp.count"]
}
```

The examples in this section use the following registry:

```
{ "registry-format": "mplane-0",
  "registry-uri": "https://example.com/mplane/registry/core",
  "registry-revision": 0,
  "includes": [],
  "elements": [
    { "name": "time",
      "prim": "time",
      "desc": "Time at which a single observation was taken"
    },
    { "name": "source.ip4",
      "prim": "address",
      "desc": "Source (or probe) IPv4 address"
    },
    { "name": "destination.ip4",
      "prim": "address",
      "desc": "Destination (or target) IPv4 address"
    },
    { "name": "intermediate.ip4",
      "prim": "address",
      "desc": "IPv4 address of intermediate node on a path"
    },
    { "name": "hops.ip",
      "prim": "natural",
      "desc": "IP-layer hops to identified node"
    }
  ],
}
```



```
{ "name": "hops.ip.max",
  "prim": "natural",
  "desc": "Maximum number of IP-layer hops to measure"
},
{ "name": "delay.twoway.icmp.us",
  "prim": "natural",
  "desc": "Singleton two-way delay as measured via ICMP Echo"
},
{ "name": "delay.twoway.icmp.us.min",
  "prim": "natural",
  "desc": "Minimum two-way delay as measured via ICMP Echo"
},
{ "name": "delay.twoway.icmp.us.50pct",
  "prim": "natural",
  "desc": "Median two-way delay as measured via ICMP Echo"
},
{ "name": "delay.twoway.icmp.us.mean",
  "prim": "natural",
  "desc": "Mean two-way delay as measured via ICMP Echo"
},
{ "name": "delay.twoway.icmp.us.max",
  "prim": "natural",
  "desc": "Maximum two-way delay as measured via ICMP Echo"
},
{ "name": "delay.twoway.icmp.us.count",
  "prim": "natural",
  "desc": "Count of two-way delay samples ... "
},
]
}
```

5.2. mPlane over WebSockets over TLS

The default session protocol for mPlane is WebSockets [[RFC6455](#)]. Once a WebSockets handshake between Client and Component is complete, messages are simply exchanged as outlined in [Section 4.2](#) as JSON objects in WebSockets text frames over the channel.

When WebSockets is used as a session protocol for mPlane, it MUST be used over TLS for mPlane message exchanges. Both TLS Clients and servers MUST present certificates for TLS mutual authentication. mPlane Components MUST use the certificate presented by the mPlane Client to determine the Client's identity, and therefore the Capabilities which it is authorized to invoke. mPlane Clients MUST use the certificate presented by the mPlane Component to authenticate the Results received. mPlane Clients and Components MUST NOT use network-layer address or name (e.g. derived from DNS PTR queries) information to identify peers.

mPlane Components may either act as WebSockets servers, for Client-initiated connection establishment, or as WebSockets Clients, for Component-initiated connection establishment. In either case, it is the responsibility of the connection initiator to re-establish connection in case it is lost. In this case, the entity acting as WebSockets server SHOULD maintain a queue of pending mPlane messages to identified peers to be sent on reconnection.

5.2.1. mPlane PKI for WebSockets

The Clients and Components within an mPlane domain generally share a single certificate issuer, specific to a single mPlane domain. Issuing a certificate to a Client or Component then grants it membership within the domain. Any Client or Component within the domain can then communicate with Components and Clients within that domain. In a domain containing one or more supervisors, all Clients and Components within the domain can connect to a supervisor. This deployment pattern can be used to scale mPlane domains to large numbers of Clients and Components without needing to specifically configure each Client and Component identity at the supervisor.

In the case of interdomain federation, where supervisors connect to each other, each supervisor will have its own issuer. In this case, each supervisor must be configured to trust each remote domain's issuer, but only to identify that domain's supervisor. This compartmentalization is necessary to keep one domain from authorizing Clients within another domain.

5.2.2. Capability Advertisement for WebSockets

When a connection is first established between a Client and a Component, regardless of whether the Client or the Component initiates the connection, the Component sends an envelope containing all the Capabilities it wishes to advertise to the Client, based on the Client's identity.

5.2.3. mPlane Link and Export URLs for WebSockets

Components acting as WebSockets servers (for Client-initiated connection establishment) are identified in the Link sections of Capabilities and receipts by URLs with the "wss:" schema. Similarly, Clients acting as WebSockets servers (for Component-initiated connection establishment) are identified in the Link sections of Specifications by URLs with the "wss:" schema.

When the "wss:" schema appears in the export section of the Capability, this represents the Component's willingness to establish a WebSockets connection over which mPlane Result messages will be

pushed. A "wss:" schema URL in a Specification export section, similarly, directs the Component to the WebSockets server to push Results to.

Path information in WebSockets URLs is presently unused by the mPlane protocol, but path information **MUST** be conserved. mPlane Clients and Components acting as WebSockets servers can use path information as they see fit, for example to separate Client and Component workflows on the same server (as on a supervisor), to run mPlane and other protocols over WebSockets on the same server, and/or to pass cryptographic tokens for additional context separation or authorization purposes. Future versions of the mPlane protocol may use path information in WebSockets URLs, but this path information will be relative to this conserved "base" URL, as opposed to relative to the root.

6. Deployment Considerations

This section outlines considerations for building larger-scale infrastructures from the building blocks defined in this document.

6.1. Supervisors and Federation

For simple infrastructures, a set of Components may be controlled directly by a Client. However, in more complex infrastructures providing support for multiple Clients, a supervisor can mediate between Clients and Components. This supervisor is responsible for collecting Capabilities from a set of Components, and providing Capabilities based on these to its Clients. From the point of view of the mPlane protocol, a supervisor is merely a combined Component and Client. The logic binding Client and Component interfaces within the supervisor is application-specific, as it involves the following operations according to the semantics of each application:

- o translating lower-level Capabilities from subordinate Components into higher-level (composed) Capabilities, according to the application's semantics
- o translating higher-level Specifications from subordinate Components into lower-level (decomposed) Specifications
- o relaying or aggregating Results from subordinate Components to supervisor Clients

This arrangement is shown in Figure 3.

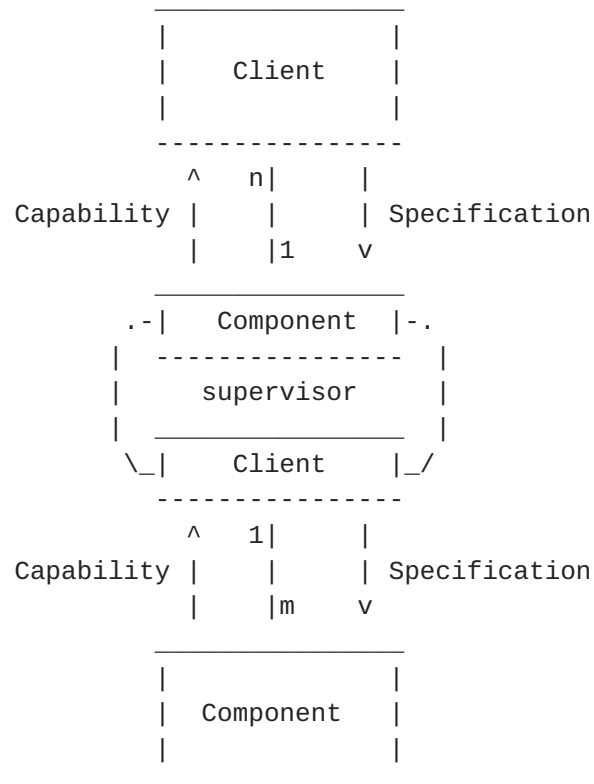


Figure 3: Simple mPlane architecture with a supervisor

The set of Components which respond to Specifications from a single supervisor is referred to as an mPlane domain. Domain membership is also determined by the issuer of the certificates identifying the Clients, Components, and supervisor. Within a given domain, each Client and Component connects to only one supervisor. Underlying measurement Components and Clients may indeed participate in multiple domains, but these are separate entities from the point of view of the architecture. Interdomain measurement is supported by federation among supervisors: a local supervisor delegates measurements in a remote domain to that domain's supervisor.

In addition to Capability composition and Specification decomposition, supervisors are responsible for Client and Component registration and authentication, as well as access control based on identity information provided by the session protocol (WebSockets) in the general case.

The general responsibilities of a supervisor are elaborated in the subsections below:

6.1.1. Component Registration

In order to be able to use Components to perform measurements, the supervisor must register the Components associated with it. For Client-initiated workflows - large repositories and the address of the Components is often a configuration parameter of the supervisor. Capabilities describing the available measurements and queries at large-scale Components can even be part of the supervisor's externally managed static configuration, or can be dynamically retrieved and updated from the Components or from a Capability discovery server.

For Component-initiated workflows, Components connect to the supervisor and send Capabilities and withdrawals, which requires the supervisor to maintain a set of Capabilities associated with a set of Components currently part of the mPlane infrastructure it supervises.

6.1.2. Client Authentication

For many Components - probes and simple repositories - very simple authentication often suffices, such that any Client with a certificate with an issuer recognized as valid is acceptable, and all Capabilities are available to. Larger repositories often need finer grained control, mapping specific peer certificates to identities internal to the repository's access control system (e.g. database users).

In an mPlane infrastructure, it is therefore the supervisor's responsibility to map Client identities to the set of Capabilities each Client is authorized to access. This mapping is part of the supervisor's configuration.

6.1.3. Capability Composition and Specification Decomposition

The most dominant responsibility of the supervisor is composing Capabilities from its subordinate Components into aggregate Capabilities, and decomposing Specifications from Clients to more-specific Specifications to pass to each Component. This operation is always application-specific, as the semantics of the composition and decomposition operations depend on the Capabilities available from the Components, the granularity of the Capabilities to be provided to the Clients.

6.2. Indirect Export

Many common measurement infrastructures involve a large number of probes exporting large volumes of data to a (much) smaller number of repositories, where data is reduced and analyzed. Since (1) the

mPlane protocol is not particularly well-suited to the bulk transfer of data and (2) fidelity is better ensured when minimizing translations between representations, the channel between the probes and the repositories is in this case external to mPlane. This indirect export channel runs either a standard export protocol such as IPFIX, or a proprietary protocol unique to the probe/repository pair. It coordinates an exporter which will produce and export data with a collector which will receive it. All that is necessary is that (1) the Client, exporter, and collector agree on a schema to define the data to be transferred and (2) the exporter and collector share a common protocol for export.

Here, we consider a Client speaking to an exporter and a collector. The Client first receives an export Capability from the exporter (with verb "measure" and with a protocol identified in the "export" section) and a collection Capability from the collector (with the verb "collect" and with a URL in the "export" section describing where the exporter should export), either via a Client-initiated workflow or a Capability discovery server. The Client then sends a Specification to the exporter, which matches the schema and parameter constraints of both the export and collection Capabilities, with the collector's URL in the "export" section.

The exporter initiates export to the collector using the specified protocol, and replies with a receipt that can be used to interrupt the export, should it have an indefinite temporal scope. In the meantime, it sends data matching the Capability's schema directly to the collector.

This data, or data derived from the analysis thereof, can then be subsequently retrieved by a Client using a Client-initiated workflow to the collector.

6.3. Error Handling in mPlane Workflows

Any Component may signal an error to its Client or supervisor at any time by sending an exception message. While the taxonomy of error messages is at this time left up to each individual Component, given the weakly imperative nature of the mPlane protocol, exceptions should be used sparingly, and only to notify Components and Clients of errors with the mPlane infrastructure itself.

It is generally presumed that diagnostic information about errors which may require external human intervention to correct will be logged at each Component; the mPlane exception facility is not intended as a replacement for logging facilities (such as syslog).

Specifically, Components using Component-initiated connection establishment should not use the exception mechanism for common error conditions (e.g., device losing connectivity for small network-edge probes) - Specifications sent to such Components are expected to be best-effort. Exceptions should also not be returned for Specifications which would normally not be delayed but are due to high load - receipts should be used in this case, instead. Likewise, Specifications which cannot be fulfilled because they request the use of Capabilities that were once available but are no longer should be answered with withdrawals.

Exceptions should always be sent in reply to messages sent to Components or Clients which cannot be handled due to a syntactic or semantic error in the message itself.

7. Security Considerations

The mPlane protocol allows the control of network measurement devices. The protocol itself uses WebSockets using TLS as a session layer. TLS mutual authentication must be used for the exchange of mPlane messages, as access control decisions about which Clients and Components are trusted for which Capabilities take identity information from the certificates TLS Clients and servers use to identify themselves. Current operational best security practices for the deployment of TLS-secured protocols must be followed for the deployment of mPlane.

Indirect export, as a design feature, presents a potential for information leakage, as indirectly exported data is necessarily related to measurement data and control transported with the mPlane protocol. Though out of scope for this document, indirect export protocols used within an mPlane domain must be secured at least as well as the mPlane protocol itself.

8. IANA Considerations

This document has no actions for IANA. Future versions of this document, should it become a standards-track Specification, may specify the initial contents of a core mPlane registry to be managed by IANA.

9. Contributors

This document is based on Deliverable 1.4, the architecture and protocol Specification document produced by the mPlane project [[D14](#)], which is the work of the mPlane consortium, specifically Brian Trammell and Mirja Kuehlewind (the editors of this document), as well as Marco Mellia, Alessandro Finamore, Stefano Pentassuglia, Gianni De

Rosa, Fabrizio Invernizzi, Marco Milanese, Dario Rossi, Saverio Niccolini, Ilias Leontiadis, Tivadar Szemethy, Balas Szabo, Rolf Winter, Michael Faath, Benoit Donnet, and Dimitri Papadimitriou.

10. Acknowledgments

Thanks to Lingli Deng and Robert Kisteleki for feedback leading to the improvement of this document and the protocol.

This work is supported by the European Commission under grant agreement FP7-318627 mPlane and H2020-688421 MAMI, and by the Swiss State Secretariat for Education, Research, and Innovation under contract no. 15.0268. This support does not imply endorsement of the contents of this document.

11. Informative References

- [RFC3205] Moore, K., "On the use of HTTP as a Substrate", [BCP 56](#), [RFC 3205](#), DOI 10.17487/RFC3205, February 2002, <<http://www.rfc-editor.org/info/rfc3205>>.
- [RFC3339] Klyne, G. and C. Newman, "Date and Time on the Internet: Timestamps", [RFC 3339](#), DOI 10.17487/RFC3339, July 2002, <<http://www.rfc-editor.org/info/rfc3339>>.
- [RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", [RFC 4291](#), DOI 10.17487/RFC4291, February 2006, <<http://www.rfc-editor.org/info/rfc4291>>.
- [RFC5246] Dierks, T. and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2", [RFC 5246](#), DOI 10.17487/RFC5246, August 2008, <<http://www.rfc-editor.org/info/rfc5246>>.
- [RFC5905] Mills, D., Martin, J., Ed., Burbank, J., and W. Kasch, "Network Time Protocol Version 4: Protocol and Algorithms Specification", [RFC 5905](#), DOI 10.17487/RFC5905, June 2010, <<http://www.rfc-editor.org/info/rfc5905>>.
- [RFC5952] Kawamura, S. and M. Kawashima, "A Recommendation for IPv6 Address Text Representation", [RFC 5952](#), DOI 10.17487/RFC5952, August 2010, <<http://www.rfc-editor.org/info/rfc5952>>.
- [RFC6455] Fette, I. and A. Melnikov, "The WebSocket Protocol", [RFC 6455](#), DOI 10.17487/RFC6455, December 2011, <<http://www.rfc-editor.org/info/rfc6455>>.

- [RFC7011] Claise, B., Ed., Trammell, B., Ed., and P. Aitken, "Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information", STD 77, [RFC 7011](#), DOI 10.17487/RFC7011, September 2013, <<http://www.rfc-editor.org/info/rfc7011>>.
- [RFC7159] Bray, T., Ed., "The JavaScript Object Notation (JSON) Data Interchange Format", [RFC 7159](#), DOI 10.17487/RFC7159, March 2014, <<http://www.rfc-editor.org/info/rfc7159>>.
- [RFC7230] Fielding, R., Ed. and J. Reschke, Ed., "Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing", [RFC 7230](#), DOI 10.17487/RFC7230, June 2014, <<http://www.rfc-editor.org/info/rfc7230>>.
- [RFC7373] Trammell, B., "Textual Representation of IP Flow Information Export (IPFIX) Abstract Data Types", [RFC 7373](#), DOI 10.17487/RFC7373, September 2014, <<http://www.rfc-editor.org/info/rfc7373>>.
- [D14] Trammell, B., "mPlane Architecture Specification", April 2015, <<https://www.ict-mplane.eu/sites/default/files//public/public-page/public-deliverables//1095mplane-d14.pdf>>.

Authors' Addresses

Brian Trammell (editor)
ETH Zurich
Gloriastrasse 35
8092 Zurich
Switzerland

Email: ietf@trammell.ch

Mirja Kuehlewind (editor)
ETH Zurich
Gloriastrasse 35
8092 Zurich
Switzerland

Email: mirja.kuehlewind@tik.ee.ethz.ch

