Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: September 26, 2014

F. Le Faucheur, Ed.
Cisco Systems
G. Bertrand, Ed.
I. Oprescu, Ed.
Orange
R. Peterkofsky
Skytide, Inc.
March 25, 2014

CDNI Logging Interface
draft-ietf-cdni-logging-11

Abstract

This memo specifies the Logging interface between a downstream CDN (dCDN) and an upstream CDN (uCDN) that are interconnected as per the CDN Interconnection (CDNI) framework. First, it describes a reference model for CDNI logging. Then, it specifies the CDNI Logging File format and the actual protocol for exchange of CDNI Logging Files.

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 September 26, 2014.

Copyright Notice

Copyright (c) 2014 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. Terminology 3
1.2. Requirements Language 4
2. CDNI Logging Reference Model 5
2.1. CDNI Logging interactions 5
2.2. Overall Logging Chain 8
2.2.1. Logging Generation and During-Generation Aggregation 9
2.2.2. Logging Collection 10
2.2.3. Logging Filtering 10
2.2.4. Logging Rectification and Post-Generation Aggregation 11
2.2.5. Log-Consuming Applications 12
2.2.5.1. Maintenance/Debugging 12
2.2.5.2. Accounting 12
2.2.5.3. Analytics and Reporting 12
2.2.5.4. Security 13
2.2.5.5. Legal Logging Duties 13
2.2.5.6. Notions common to multiple Log Consuming Applications 13
3. CDNI Logging File 15
3.1. Rules 15
3.2. CDNI Logging File Structure 16
3.3. CDNI Logging File Directives 19
3.4. CDNI Logging Records 22
3.4.1. HTTP Request Logging Record 23
3.5. CDNI Logging File Example 32
4. CDNI Logging File Exchange Protocol 32
4.1. CDNI Logging Feed 33
4.1.1. Atom Formatting 33
4.1.2. Updates to Log Files and the Feed 34
4.1.3. Redundant Feeds 34
4.1.4. Example CDNI Logging Feed 35
4.2. CDNI Logging File Pull 37
5. IANA Considerations 38
5.1. CDNI Logging Directive Names Registry 38
5.2. CDNI Logging Record-Types Registry 39
5.3. CDNI Logging Field Names Registry 40
5.4. CDNI Logging MIME Media Type 41
6. Security Considerations 41
6.1. Authentication, Confidentiality, Integrity Protection . . 41
6.2. Denial of Service 42

6.3. Privacy 42
7. Acknowledgments 43
8. References 44
8.1. Normative References 44
8.2. Informative References 44
Authors' Addresses 46

1. Introduction

This memo specifies the CDNI Logging interface between a downstream CDN (dCDN) and an upstream CDN (uCDN). First, it describes a reference model for CDNI logging. Then, it specifies the CDNI Logging File format and the actual protocol for exchange of CDNI Logging Files.

The reader should be familiar with the following documents:

- CDNI problem statement [RFC6707] and framework [I-D.ietf-cdni-framework] identify a Logging interface,
- Section 8 of [I-D.ietf-cdni-requirements] specifies a set of requirements for Logging,
- [RFC6770] outlines real world use-cases for interconnecting CDNs. These use cases require the exchange of Logging information between the dCDN and the uCDN.

As stated in [RFC6707], "the CDNI Logging interface enables details of logs or events to be exchanged between interconnected CDNs".

The present document describes:

- The CDNI Logging reference model (Section 2),
- The CDNI Logging File format (Section 3),
- The CDNI Logging File Exchange protocol (Section 4).

1.1. Terminology

In this document, the first letter of each CDNI-specific term is capitalized. We adopt the terminology described in [RFC6707] and [I-D.ietf-cdni-framework], and extend it with the additional terms defined below.

Intra-CDN Logging information: logging information generated and collected within a CDN. The format of the Intra-CDN Logging

information may be different to the format of the CDNI Logging information.

CDNI Logging information: logging information exchanged across CDNs using the CDNI Logging Interface.

Logging information: logging information generated and collected within a CDN or obtained from another CDN using the CDNI Logging Interface.

CDNI Logging Field: an atomic element of information that can be included in a CDNI Logging Record. The time an event/task started, the IP address of an End User to whom content was delivered, and the URI of the content delivered are examples of CDNI Logging Fields.

CDNI Logging Record: an information record providing information about a specific event. This comprises a collection of CDNI Logging Fields.

CDNI Logging File: a file containing CDNI Logging Records, as well as additional information facilitating the processing of the CDNI Logging Records.

CDN Reporting: the process of providing the relevant information that will be used to create a formatted content delivery report provided to the CSP in deferred time. Such information typically includes aggregated data that can cover a large period of time (e.g., from hours to several months). Uses of Reporting include the collection of charging data related to CDN services and the computation of Key Performance Indicators (KPIs).

CDN Monitoring: the process of providing or displaying content delivery information in a timely fashion with respect to the corresponding deliveries. Monitoring typically includes visibility of the deliveries in progress for service operation purposes. It presents a view of the global health of the services as well as information on usage and performance, for network services supervision and operation management. In particular, monitoring data can be used to generate alarms.

1.2. 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 RFC 2119 [RFC2119].

2. CDNI Logging Reference Model

2.1. CDNI Logging interactions

The CDNI logging reference model between a given uCDN and a given dCDN involves the following interactions:

- customization by the uCDN of the CDNI Logging information to be provided by the dCDN to the uCDN (e.g., control of which CDNI Logging Fields are to be communicated to the uCDN for a given task performed by the dCDN or control of which types of events are to be logged). The dCDN takes into account this CDNI Logging customization information to determine what Logging information to provide to the uCDN, but it may, or may not, take into account this CDNI Logging customization information to influence what CDN logging information is to be generated and collected within the dCDN (e.g., even if the uCDN requests a restricted subset of the logging information, the dCDN may elect to generate a broader set of logging information). The mechanism to support the customization by the uCDN of CDNI Logging information is outside the scope of this document and left for further study. Until such a mechanism is available, the uCDN and dCDN are expected to agree off-line on what exact set of CDNI Logging information is to be provided by the dCDN to the uCDN, and to rely on management plane actions to configure the CDNI Logging functions in the dCDN to generate this information set and in the uCDN to expect this information set.

- generation and collection by the dCDN of the intra-CDN Logging information related to the completion of any task performed by the dCDN on behalf of the uCDN (e.g., delivery of the content to an End User) or related to events happening in the dCDN that are relevant to the uCDN (e.g., failures or unavailability in dCDN). This takes place within the dCDN and does not directly involve CDNI interfaces.

- communication by the dCDN to the uCDN of the Logging information collected by the dCDN relevant to the uCDN. This is supported by the CDNI Logging interface and in the scope of the present document. For example, the uCDN may use this Logging information to charge the CSP, to perform analytics and monitoring for operational reasons, to provide analytics and monitoring views on its content delivery to the CSP or to perform trouble-shooting. This document exclusively specifies non-real-time exchange of Logging information. Closer to real-time exchange of Logging information (say sub-minute or sub-second) is outside the scope of the present document and left for further study. This document exclusively specifies exchange of Logging information related to

content delivery. Exchange of Logging information related to operational events (e.g., dCDN request routing function unavailable, content acquisition failure by dCDN) for audit or operational reactive adjustments by uCDN is outside the scope of the present document and left for further study.

- customization by the dCDN of the CDNI Logging information to be provided by the uCDN on behalf of the dCDN. The mechanism to support the customization by the dCDN of CDNI Logging information is outside the scope of this document and left for further study.

- generation and collection by the uCDN of Intra-CDN Logging information related to the completion of any task performed by the uCDN on behalf of the dCDN (e.g., serving of content by uCDN to dCDN for acquisition purposes by dCDN) or related to events happening in the uCDN that are relevant to the dCDN. This takes place within the uCDN and does not directly involve CDNI interfaces.

- communication by the uCDN to the dCDN of the Logging information collected by the uCDN relevant to the dCDN. For example, the dCDN might potentially benefit from this information for security auditing or content acquisition troubleshooting. This is outside the scope of this document and left for further study.

Figure 1 provides an example of CDNI Logging interactions (focusing only on the interactions that are in the scope of this document) in a particular scenario where four CDNs are involved in the delivery of content from a given CSP: the uCDN has a CDNI interconnection with dCDN-1 and dCDN-2. In turn, dCDN2 has a CDNI interconnection with dCDN3. In this example, uCDN, dCDN-1, dCDN-2 and dCDN-3 all participate in the delivery of content for the CSP. In this example, the CDNI Logging interface enables the uCDN to obtain Logging information from all the dCDNs involved in the delivery. In the example, the uCDN uses the Logging information:

- to analyze the performance of the delivery performed by the dCDNs and to adjust its operations after the fact (e.g., request routing) as appropriate,

- to provide (non-real-time) reporting and monitoring information to the CSP.

For instance, the uCDN merges Logging information, extracts relevant KPIs, and presents a formatted report to the CSP, in addition to a bill for the content delivered by uCDN itself or by its dCDNs on the CSP's behalf. The uCDN may also provide Logging information as raw

log files to the CSP, so that the CSP can use its own logging analysis tools.

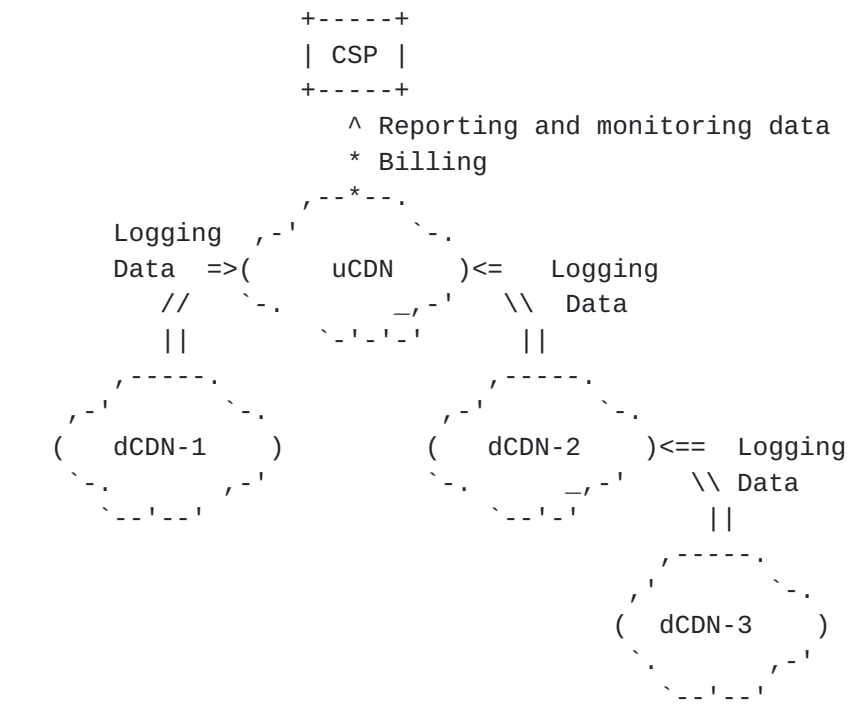


Figure 1: Interactions in CDNI Logging Reference Model

A dCDN (e.g., dCDN-2) integrates the relevant Logging information obtained from its dCDNs (i.e., dCDN-3) in the Logging information that it provides to the uCDN, so that the uCDN ultimately obtains all Logging information relevant to a CSP for which it acts as the authoritative CDN.

Note that the format of Logging information that a CDN provides over the CDNI interface might be different from the one that the CDN uses internally. In this case, the CDN needs to reformat the Logging information before it provides this information to the other CDN over the CDNI Logging interface. Similarly, a CDN might reformat the Logging information that it receives over the CDNI Logging interface before injecting it into its log-consuming applications or before providing some of this Logging information to the CSP. Such reformatting operations introduce latency in the logging distribution chain and introduce a processing burden. Therefore, there are benefits in specifying CDNI Logging formats that are suitable for use

inside CDNs and also are close to the intra-CDN Logging formats commonly used in CDNs today.

2.2. Overall Logging Chain

This section discusses the overall logging chain within and across CDNs to clarify how CDN Logging information is expected to fit in this overall chain. Figure 2 illustrates the overall logging chain within the dCDN, across CDNs using the CDNI Logging interface and within the uCDN. Note that the logging chain illustrated in the Figure is obviously only an example and varies depending on the specific environments. For example, there may be more or fewer instantiations of each entity (e.g., there may be 4 Log consuming applications in a given CDN). As another example, there may be one instance of Rectification process per Log Consuming Application instead of a shared one.

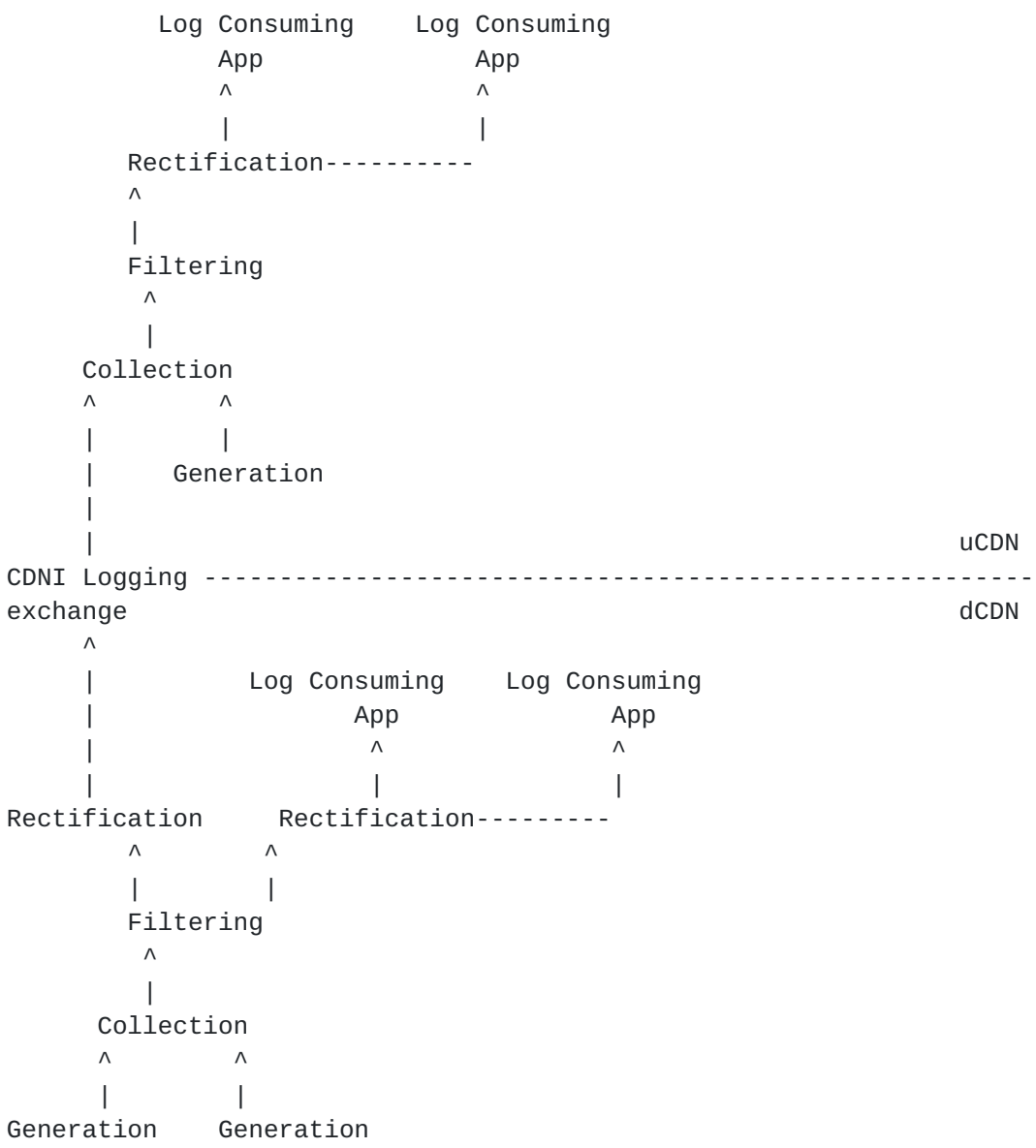


Figure 2: CDNI Logging in the overall Logging Chain

The following subsections describe each of the processes potentially involved in the logging chain of Figure 2.

2.2.1. Logging Generation and During-Generation Aggregation

CDNs typically generate Logging information for all significant task completions, events, and failures. Logging information is typically generated by many devices in the CDN including the surrogates, the request routing system, and the control system.

The amount of Logging information generated can be huge. Therefore, during contract negotiations, interconnected CDNs often agree on a retention duration for Logging information, and/or potentially on a maximum volume of Logging information that the dCDN must keep. If this volume is exceeded, the dCDN is expected to alert the uCDN but may not keep more Logging information for the considered time period. In addition, CDNs may aggregate Logging information and transmit only summaries for some categories of operations instead of the full Logging information. Note that such aggregation leads to an information loss, which may be problematic for some usages of the Logging information (e.g., debugging).

[RFC6983] discusses logging for HTTP Adaptive Streaming (HAS). In accordance with the recommendations articulated there, it is expected that a surrogate will generate separate Logging information for delivery of each chunk of HAS content. This ensures that separate Logging information can then be provided to interconnected CDNs over the CDNI Logging interface. Still in line with the recommendations of [RFC6983], the Logging information for per-chunck delivery may include some information (a Content Collection IDentifier and a Session IDentifier) intended to facilitate subsequent post-generation aggregation of per-chunk logs into per-session logs. Note that a CDN may also elect to generate aggregate per-session logs when performing HAS delivery, but this needs to be in addition to, and not instead of, the per-chunk delivery logs. We note that aggregate per-session logs for HAS delivery are for further study and outside the scope of this document.

2.2.2. Logging Collection

This is the process that continuously collects Logging information generated by the log-generating entities within a CDN.

In a CDNI environment, in addition to collecting Logging information from log-generating entities within the local CDN, the Collection process also collects Logging information provided by another CDN, or other CDNs, through the CDNI Logging interface. This is illustrated in Figure 2 where we see that the Collection process of the uCDN collects Logging information from log-generating entities within the uCDN as well as Logging information coming from the dCDNs through the CDNI Logging interface.

2.2.3. Logging Filtering

A CDN may be required to only present different subsets of the whole Logging information collected to various log-consuming applications. This is achieved by the Filtering process.

In particular, the Filtering process can also filter the right subset of Logging information that needs to be provided to a given interconnected CDN. For example, the filtering process in the dCDN can be used to ensure that only the Logging information related to tasks performed on behalf of a given uCDN are made available to that uCDN (thereby filtering out all the Logging information related to deliveries by the dCDN of content for its own CSPs). Similarly, the Filtering process may filter or partially mask some fields, for example, to protect End Users' privacy when communicating CDNI Logging information to another CDN. Filtering of Logging information prior to communication of this information to other CDNs via the CDNI Logging interface requires that the downstream CDN can recognize the subset of Logging information that relate to each interconnected CDN.

The CDN will also filter some internal scope information such as information related to its internal alarms (security, failures, load, etc).

In some use cases described in [RFC6770], the interconnected CDNs do not want to disclose details on their internal topology. The filtering process can then also filter confidential data on the dCDNs' topology (number of servers, location, etc.). In particular, information about the requests served by each Surrogate may be confidential. Therefore, the Logging information must be protected so that data such as Surrogates' hostnames are not disclosed to the uCDN. In the "Inter-Affiliates Interconnection" use case, this information may be disclosed to the uCDN because both the dCDN and the uCDN are operated by entities of the same group.

2.2.4. Logging Rectification and Post-Generation Aggregation

If Logging information is generated periodically, it is important that the sessions that start in one Logging period and end in another are correctly reported. If they are reported in the starting period, then the Logging information of this period will be available only after the end of the session, which delays the Logging information generation.

A Logging rectification/update mechanism could be useful to reach a good trade-off between the Logging information generation delay and the Logging information accuracy.

In the presence of HAS, some log-consuming applications can benefit from aggregate per-session logs. For example, for analytics, per-session logs allow display of session-related trends which are much more meaningful for some types of analysis than chunk-related trends. In the case where aggregate logs have been generated directly by the log-generating entities, those can be used by the applications. In

the case where aggregate logs have not been generated, the Rectification process can be extended with a Post-Generation Aggregation process that generates per-session logs from the per-chunk logs, possibly leveraging the information included in the per-chunk logs for that purpose (Content Collection IDentifier and a Session IDentifier). However, in accordance with [RFC6983], this document does not define exchange of such aggregate logs on the CDNI Logging interface. We note that this is for further study and outside the scope of this document..

2.2.5. Log-Consuming Applications

2.2.5.1. Maintenance/Debugging

Logging information is useful to permit the detection (and limit the risk) of content delivery failures. In particular, Logging information facilitates the detection of configuration issues.

To detect faults, Logging information needs to report success and failure of CDN delivery operations. The uCDN can summarize such information into KPIs. For instance, Logging information needs to allow the computation of the number of times, during a given time period, that content delivery related to a specific service succeeds/fails.

Logging information enables the CDN providers to identify and troubleshoot performance degradations. In particular, Logging information enables tracking of traffic data (e.g., the amount of traffic that has been forwarded by a dCDN on behalf of an uCDN over a given period of time), which is particularly useful for CDN and network planning operations.

2.2.5.2. Accounting

Logging information is essential for accounting, to permit inter-CDN billing and CSP billing by uCDNs. For instance, Logging information provided by dCDNs enables the uCDN to compute the total amount of traffic delivered by every dCDN for a particular Content Provider, as well as, the associated bandwidth usage (e.g., peak, 95th percentile), and the maximum number of simultaneous sessions over a given period of time.

2.2.5.3. Analytics and Reporting

The goal of analytics is to gather any relevant information to track audience, analyze user behavior, and monitor the performance and quality of content delivery. For instance, Logging information

enables the CDN providers to report on content consumption (e.g., delivered sessions per content) in a specific geographic area.

The goal of reporting is to gather any relevant information to monitor the performance and quality of content delivery and allow detection of delivery issues. For instance, reporting could track the average delivery throughput experienced by End Users in a given region for a specific CSP or content set over a period of time.

2.2.5.4. Security

The goal of security is to prevent and monitor unauthorized access, misuse, modification, and denial of access of a service. A set of information is logged for security purposes. In particular, a record of access to content is usually collected to permit the CSP to detect infringements of content delivery policies and other abnormal End User behaviors.

2.2.5.5. Legal Logging Duties

Depending on the country considered, the CDNs may have to retain specific Logging information during a legal retention period, to comply with judicial requisitions.

2.2.5.6. Notions common to multiple Log Consuming Applications

2.2.5.6.1. Logging Information Views

Within a given log-consuming application, different views may be provided to different users depending on privacy, business, and scalability constraints.

For example, an analytics tool run by the uCDN can provide one view to an uCDN operator that exploits all the Logging information available to the uCDN, while the tool may provide a different view to each CSP exploiting only the Logging information related to the content of the given CSP.

As another example, maintenance and debugging tools may provide different views to different CDN operators, based on their operational role.

2.2.5.6.2. Key Performance Indicators (KPIs)

This section presents, for explanatory purposes, a non-exhaustive list of Key Performance Indicators (KPIs) that can be extracted/ produced from logs.

Multiple log-consuming applications, such as analytics, monitoring, and maintenance applications, often compute and track such KPIs.

In a CDNI environment, depending on the situation, these KPIs may be computed by the uCDN or by the dCDN. But it is usually the uCDN that computes KPIs, because the uCDN and dCDN may have different definitions of the KPIs and the computation of some KPIs requires a vision of all the deliveries performed by the uCDN and all its dCDNs.

Here is a list of important examples of KPIs:

- Number of delivery requests received from End Users in a given region for each piece of content, during a given period of time (e.g., hour/day/week/month)
- Percentage of delivery successes/failures among the aforementioned requests
- Number of failures listed by failure type (e.g., HTTP error code) for requests received from End Users in a given region and for each piece of content, during a given period of time (e.g., hour/day/week/month)
- Number and cause of premature delivery termination for End Users in a given region and for each piece of content, during a given period of time (e.g., hour/day/week/month)
- Maximum and mean number of simultaneous sessions established by End Users in a given region, for a given Content Provider, and during a given period of time (e.g., hour/day/week/month)
- Volume of traffic delivered for sessions established by End Users in a given region, for a given Content Provider, and during a given period of time (e.g., hour/day/week/month)
- Maximum, mean, and minimum delivery throughput for sessions established by End Users in a given region, for a given Content Provider, and during a given period of time (e.g., hour/day/week/month)
- Cache-hit and byte-hit ratios for requests received from End Users in a given region for each piece of content, during a given period of time (e.g., hour/day/week/month)
- Top 10 most popularly requested contents (during a given day/week/month)

- Terminal type (mobile, PC, STB, if this information can be acquired from the browser type header, for example).

Additional KPIs can be computed from other sources of information than the Logging information, for instance, data collected by a content portal or by specific client-side application programming interfaces. Such KPIs are out of scope for the present document.

The KPIs used depend strongly on the considered log-consuming application -- the CDN operator may be interested in different metrics than the CSP is. In particular, CDN operators are often interested in delivery and acquisition performance KPIs, information related to Surrogates' performance, caching information to evaluate the cache-hit ratio, information about the delivered file size to compute the volume of content delivered during peak hour, etc.

Some of the KPIs, for instance those providing an instantaneous vision of the active sessions for a given CSP's content, are useful essentially if they are provided in a timely manner. By contrast, some other KPIs, such as those averaged on a long period of time, can be provided in non-real-time.

3. CDNI Logging File

3.1. Rules

This specification uses the Augmented Backus-Naur Form (ABNF) notation and core rules of [RFC5234]. In particular, the present document uses the following rules from [RFC5234]:

```
CR = %x0D ; carriage return

ALPHA = %x41-5A / %x61-7A ; A-Z / a-z

DIGIT = %x30-39 ; 0-9

DQUOTE = %x22 ; " (Double Quote)

CRLF = CR LF ; Internet standard newline

HEXDIG = DIGIT / "A" / "B" / "C" / "D" / "E" / "F"

HTAB = %x09 ; horizontal tab

LF = %x0A ; linefeed

OCTET = %x00-FF ; 8 bits of data
```

The present document also uses the following rules from [RFC3986]:

host = as specified in section 3.2.2 of [RFC3986].

IPv4address = as specified in section 3.2.2 of [RFC3986].

IPv6address = as specified in section 3.2.2 of [RFC3986].

The present document also defines the following additional rules:

ADDRESS = IPv4address / IPv6address

ALPHANUM = ALPHA / DIGIT

DATE = 4DIGIT "-" 2DIGIT "-" 2DIGIT

Dates are recorded in the format YYYY-MM-DD where YYYY, MM and DD stand for the numeric year, month and day respectively. All dates are specified in Universal Time Coordinated (UTC).

DEC = 1*DIGIT ["." *DIGIT]

NAMEFORMAT = ALPHANUM *(ALPHANUM / "_" / "-")

QSTRING = DQUOTE *NDQUOTE DQUOTE ; where

NDQUOTE = <any OCTET excluding DQUOTE> / 2DQUOTE ; whereby a DQUOTE is conveyed inside a QSTRING unambiguously by repeating it.

NHTABSTRING = *NHTAB ; where

NHTAB = <any OCTET excluding HTAB, CR and LF>

TIME = 2DIGIT ":" 2DIGIT ":" 2DIGIT ["." *DIGIT]

Times are recorded in the form HH:MM:SS or HH:MM:SS.S where HH is the hour in 24 hour format, MM is minutes and SS is seconds. All times are specified in Universal Time Coordinated (UTC).

3.2. CDNI Logging File Structure

As defined in Section 1.1: a CDNI Logging Field is as an atomic logging information element, a CDNI Logging Record is a collection of CDNI Logging Fields containing all logging information corresponding to a single logging event, and a CDNI Logging File contains a

collection of CDNI Logging Records. This structure is illustrated in Figure 3. The use of a file structure for transfer of CDNI Logging information is selected since this is the most common practise today for exchange of logging information within and across CDNs.

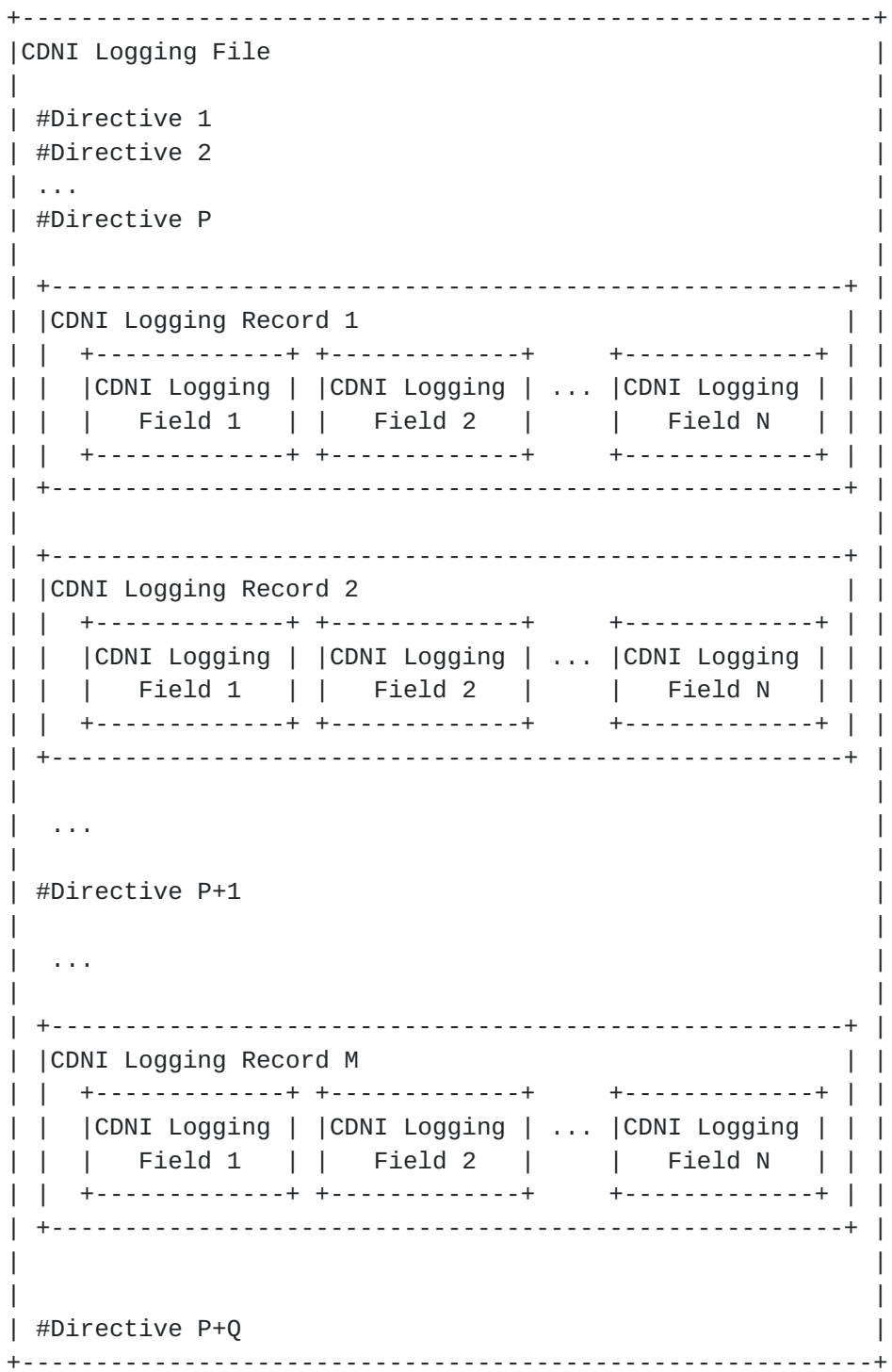


Figure 3: Structure of Logging Files

The CDNI Logging File format is inspired from the W3C Extended Log File Format [ELF]. However, it is fully specified by the present document. Where the present document differs from the W3C Extended

Log File Format, an implementation of CDNI Logging MUST comply with the present document.

Using a format that resembles the W3C Extended Log File Format is intended to keep CDNI logging format close to the intra-CDN Logging information format commonly used in CDNs today, thereby minimizing systematic translation at CDN/CDNI boundary.

A CDNI Logging File MUST contain a sequence of lines containing US-ASCII characters [CHAR_SET] terminated by CRLF.

Each line of a CDNI Logging File MUST contain either a directive or a CDNI Logging Record.

Directives record information about the CDNI Logging process itself. Lines containing directives MUST begin with the "#" character. Directives are specified in Section 3.3.

Logging Records provide actual details of the logged event. Logging Records are specified in Section 3.4.

The CDNI File structure is defined by the following rules:

```
DIRLINE = "#" directive CRLF

DIRGROUP = 1*DIRLINE

RECLINE = <CDNI Logging Record> CRLF

RECGROUP = *RECLINE

<CDNI Logging File> = 1*<DIRGROUP RECGROUP>
```

3.3. CDNI Logging File Directives

The CDNI Logging File directives are defined by the following rules:

```
directive = DIRNAME ":" HTAB DIRVAL

DIRNAME = <any CDNI Logging Directive name registered in the CDNI
Logging Directive Names registry (Section 5.1)>

DIRVAL = <directive value, as specified by the document identified
as Reference in the CDNI Logging Directive Names registry
(Section 5.1) for the corresponding DIRNAME>
```

An implementation of the CDNI Logging interface MUST support all of the following directives, listed below by their directive name:

- Version:
  - format: "CDNI" "/" 1*DIGIT "." 1*DIGIT
  - directive value: indicates the version of the CDNI Logging File format. The value MUST be "CDNI/1.0" for the version specified in the present document.
  - occurrence: there MUST be one and only one instance of this directive per CDNI Logging File. It MUST be the first line of the CDNI Logging File.
- UUID:
  - format: NHTABSTRING
  - directive value: this a Universally Unique IDentifier (UUID) from the UUID Uniform Resource Name (URN) namespace specified in [RFC4122]) for the CDNI Logging File.
  - occurrence: there MUST be one and only one instance of this directive per CDNI Logging File.
- Claimed-Origin:
  - format: host
  - directive value: this contains the claimed identification of the entity transmitting the CDNI Logging File (e.g., the host in a dCDN supporting the CDNI Logging interface) or the entity responsible for transmitting the CDNI Logging File (e.g., the dCDN).
  - occurrence: there MUST be zero or exactly one instance of this directive per CDNI Logging File. This directive MAY be included by the dCDN. It MUST NOT be included or modified by the uCDN.
- Verified-Origin:
  - format: host
  - directive value: this contains the identification, as established by the entity receiving the CDNI Logging File, of the entity transmitting the CDNI Logging File (e.g., the host in a dCDN supporting the CDNI Logging interface) or the entity responsible for transmitting the CDNI Logging File (e.g., the dCDN).

  * occurrence: there MUST be zero or exactly one instance of this directive per CDNI Logging File. This directive MAY be added by the uCDN (e.g., before storing the CDNI Logging File). It MUST NOT be included by the dCDN. The mechanisms used by the uCDN to establish and validate the entity responsible for the CDNI Logging File is outside the scope of the present document. We observe that, in particular, this may be achieved through authentication mechanisms that are part of the CDNI Logging File pull mechanism (Section 4.2).

- Record-Type:

  * format: NAMEFORMAT

  * directive value: indicates the type of the CDNI Logging Records that follow this directive, until another Record-Type directive (or the end of the CDNI Logging File). This can be any CDNI Logging Record type registered in the CDNI Logging Record-types registry (Section 5.2). "cdni_http_request_v1" MUST be indicated as the Record-Type directive value for CDNI Logging records corresponding to HTTP requests (e.g., a HTTP delivery request) as specified in Section 3.4.1.

  * occurrence: there MUST be at least one instance of this directive per CDNI Logging File. The first instance of this directive MUST precede a Fields directive and MUST precede all CDNI Logging Records.

- Fields:

  * format: FIENAME *<HTAB FIENAME> ; where FIENAME can take any CDNI Logging field name registered in the CDNI Logging Field Names registry (Section 5.3).

  * directive value: this lists the names of all the fields for which a value is to appear in the CDNI Logging Records that follow the instance of this directive (until another instance of this directive). The names of the fields, as well as their occurrences, MUST comply with the corresponding rules specified in the document referenced in the CDNI Logging Record-types registry (Section 5.2) for the corresponding CDNI Logging Record-Type.

  * occurrence: there MUST be at least one instance of this directive per Record-Type directive. The first instance of this directive for a given Record-Type MUST appear before any CDNI Logging Record for this Record-Type.

- Integrity-Hash:
  - format: 32HEXDIG
  - directive value: This directive permits the detection of a corrupted CDNI Logging File. This can be useful, for instance, if a problem occurs on the filesystem of the dCDN Logging system and leads to a truncation of a logging file. The valid Integrity-Hash value is included in this directive by the entity that transmits the CDNI Logging File. It is computed by applying the MD5 ([RFC1321]) cryptographic hash function on the CDNI Logging File, including all the directives and logging records, up to the Integrrity-Hash directive itself, excluding the Integrity-Hash directive itself. The Integrity-Hash value is represented as a US-ASCII encoded hexadecimal number, 32 digits long (representing a 128 bit hash value). The entity receiving the CDNI Logging File also computes in a similar way the MD5 hash on the received CDNI Logging File and compares this hash to the value of the Integrity-Hash directive. If the two values are equal, then the received CDNI Logging File MUST be considered non-corrupted. Note that this is not a guarantee that the file has not been tampered with by a third party. If the two values are different, the received CDNI Logging File MUST be considered corrupted. The behavior of the entity that received a corrupted CDNI Logging File is outside the scope of this specification; we note that the entity MAY attempt to pull again the same CDNI Logging File from the transmitting entity. If the entity receiving a non-corrupted CDNI Logging File adds a Verified-Origin directive, it MUST then recompute and update the Integrity-Hash directive so it also protects the added Verified-Origin directive.
  - occurrence: there MUST be zero or exactly one instance of this directive. There SHOULD be exactly one instance of this directive. One situation where that directive could be omitted is where integrity protection is already provided via another mechanism (for example if an integrity hash is associated to the CDNI Logging File out of band through the CDNI Logging Logging Feed Section 4.1 leveraging ATOM extensions such as those proposed in [I-D.snell-atompub-link-extensions]. When present, this field MUST be the last line of the CDNI Logging File.

3.4. CDNI Logging Records

A CDNI Logging Record consists of a sequence of CDNI Logging Fields relating to that single CDNI Logging Record.

CDNI Logging Fields MUST be separated by the "horizontal tabulation (HTAB)" character.

To facilitate readability, a prefix scheme is used for CDNI Logging field names in a similar way to the one used in W3C Extended Log File Format [ELF]. The semantics of the prefix in the present document is:

- c: refers to the User Agent that issues the request (corresponds to the "client" of W3C Extended Log Format)
- d: refers to the dCDN (relative to a given CDN acting as a uCDN)
- s: refers to the dCDN Surrogate that serves the request (corresponds to the "server" of W3C Extended Log Format)
- u: refers to the uCDN (relative to a given CDN acting as a dCDN)
- cs: refers to communication from the User Agent towards the dCDN Surrogate
- sc: refers to communication from the dCDN Surrogate towards the User Agent

An implementation of the CDNI Logging interface as per the present specification MUST support the CDNI HTTP Delivery Records as specified in Section 3.4.1.

A CDNI Logging Record is defined by the following rules:

FIEVAL = <CDNI Logging Field value>

<CDNI Logging Record> = FIEVAL *<HTAB FIEVAL> ; where FIEVAL contains the CDNI Logging field values corresponding to the CDNI Logging field names (FIENAME) listed is the last Fields directive predecing the present CDNI Logging Record.

3.4.1. HTTP Request Logging Record

This section defines the CDNI Logging Record of Record-Type "cdni_http_request_v1". It is applicable to content delivery performed by the dCDN using HTTP/1.0([RFC1945]), HTTP/1.1([RFC2616]) or HTTPS ([RFC2818]). We observe that, in the case of HTTPS delivery, there may be value in logging additional information specific to the operation of HTTP over TLS and we note that this is outside the scope of the present document and may be addressed in a future document defining another CDNI Logging Record or another version of the HTTP Request Logging Record.

The "cdni_http_request_v1" Record-Type is also expected to be applicable to HTTP/2.0 [I-D.ietf-httpbis-http2] (which is still under development at the time of writing the present document) since a fundamental design tenet of HTTP/2.0 is to preserve the HTTP/1.1 semantics. We observe that, in the case of HTTP/2.0 delivery, there may be value in logging additional information specific to the additional functionality of HTTP/2.0 (e.g. information related to connection identification, to stream identification, to stream priority and to flow control). We note that such additional information is outside the scope of the present document and may be addressed in a future document defining another CDNI Logging Record or another version of the HTTP Request Logging Record.

The "cdni_http_request_v1" Record-Type contains the following CDNI Logging Fields, listed by their field name:

- date:
  - format: DATE
  - field value: the date at which the processing of request completed on the Surrogate.
  - occurrence: there MUST be one and only one instance of this field.
- time:
  - format: TIME
  - field value: the time at which the processing of request completed on the Surrogate.
  - occurrence: there MUST be one and only one instance of this field.
- time-taken:
  - format: DEC
  - field value: decimal value of the duration, in seconds, between the start of the processing of the request and the completion of the request processing (e.g., completion of delivery) by the Surrogate.
  - occurrence: there MUST be one and only one instance of this field.

- c-ip:
  - format: ADDRESS
  - field value: the source IPv4 or IPv6 address (i.e., the "client" address) in the request received by the Surrogate.
  - occurrence: there MUST be one and only one instance of this field.
- c-ip-anonymizing:
  - format: 1*DIGIT
  - field value: the number of rightmost bits of the address in the c-ip field that are zeroed-out in order to anonymize the logging record. The mechanism by which the two ends of the CDNI Logging interface agree on whether anonymization is to be supported and the number of bits that need to be zeroed-out for this purpose are outside the scope of the present document.
  - occurrence: there MUST be zero or exactly one instance of this field.
- c-port:
  - format: 1*DIGIT
  - field value: the source TCP port (i.e., the "client" port) in the request received by the Surrogate.
  - occurrence: there MUST be zero or exactly one instance of this field.
- s-ip:
  - format: ADDRESS
  - field value: the IPv4 or IPv6 address of the Surrogate that served the request (i.e., the "server" address).
  - occurrence: there MUST be zero or exactly one instance of this field.
- s-hostname:
  - format: host

  * field value: the hostname of the Surrogate that served the request (i.e., the "server" hostname).

  * occurrence: there MUST be zero or exactly one instance of this field.

- s-port:

  * format: 1*DIGIT

  * field value: the destination TCP port (i.e., the "server" port) in the request received by the Surrogate.

  * occurrence: there MUST be zero or exactly one instance of this field.

- cs-method:

  * format: NHTABSTRING

  * field value: this is the HTTP method of the HTTP request received by the Surrogate.

  * occurrence: There MUST be one and only one instance of this field.

- cs-uri:

  * format: NHTABSTRING

  * field value: this is the complete URL of the request received by the Surrogate. It is exactly in the format of a http_URL specified in [RFC2616]) or, when the request was a HTTPS request ([RFC2818]), it is in the format of a http_URL but with the scheme part set to "https" instead of "http".

  * occurrence: there MUST be zero or exactly one instance of this field.

- u-uri:

  * format: NHTABSTRING

  * field value: this is a complete URL, derived from the complete URI of the request received by the Surrogate (i.e., the cs-uri) but transformed by the entity generating or transmitting the CDNI Logging Record, in a way that is agreed upon between the two ends of the CDNI Logging interface, so the transformed URI

is meaningful to the uCDN. For example, the two ends of the CDNI Logging interface could agree that the u-uri is constructed from the cs-uri by removing the part of the hostname that exposes which individual Surrogate actually performed the delivery. The details of modification performed to generate the u-uri, as well as the mechanism to agree on these modifications between the two sides of the CDNI Logging interface are outside the scope of the present document.

* occurrence: there MUST be one and only one instance of this field.

- protocol:
  * format: NHTABSTRING
  * field value: this is value of the HTTP-Version field as specified in [RFC2616] of the Request-Line of the request received by the Surrogate (e.g., "HTTP/1.1").
  * occurrence: there MUST be one and only one instance of this field.
- sc-status:
  * format: 3DIGIT
  * field value: this is the HTTP Status-Code in the HTTP response from the Surrogate.
  * occurrence: There MUST be one and only one instance of this field.
- sc-total-bytes:
  * format: 1*DIGIT
  * field value: this is the total number of bytes of the HTTP response sent by the Surrogate in response to the request. This includes the bytes of the Status-Line, the bytes of the HTTP headers and the bytes of the message-body.
  * occurrence: There MUST be one and only one instance of this field.
- sc-entity-bytes:
  * format: 1*DIGIT

  * field value: this is the number of bytes of the message-body in the HTTP response sent by the Surrogate in response to the request. This does not include the bytes of the Status-Line or the bytes of the HTTP headers.

  * occurrence: there MUST be zero or exactly one instance of this field.

- cs(<HTTP-header-name>):

  * format: QSTRING

  * field value: the value of the HTTP header (identified by the <HTTP-header-name> in the CDNI Logging field name) as it appears in the request processed by the Surrogate, but prepended by a DQUOTE and appended by a DQUOTE. For example, when the CDNI Logging field name (FIENAME) listed in the preceding Fields directive is cs(User-Agent), this CDNI Logging field value contains the value of the User-Agent HTTP header as received by the Surrogate in the request it processed, but prepended by a DQUOTE and appended by a DQUOTE. If the HTTP header as it appeared in the request processed by the Surrogate contains one or more DQUOTE, each DQUOTE MUST be escaped by an additional DQUOTE. For example, if the HTTP header contains My_Header"value", then the field value of the cs(<HTTP-header-name>) is "My_Header""value""".

  * occurrence: there MAY be zero, one or any number of instance of this field.

- sc(<HTTP-header-name>):

  * format: QSTRING

  * field value: the value of the HTTP header (identified by the <HTTP-header-name> in the CDNI Logging field name) as it appears in the response issued by the Surrogate to serve the request, but prepended by a DQUOTE and appended by a DQUOTE. If the HTTP header as it appeared in the request processed by the Surrogate contains one or more DQUOTE, each DQUOTE MUST be escaped by an additional DQUOTE. For example, if the HTTP header contains My_Header"value", then the field value of the cs(<HTTP-header-name>) is "My_Header""value""".

  * occurrence: there MAY be zero, one or any number of instances of this field. For a given <HTTP-header-name>, there MUST be zero or exactly one instance of this field.

- s-ccid:
  - format: QSTRING
  - field value: this contains the value of the Content Collection IDentifier (CCID) associated by the uCDN to the content served by the Surrogate via the CDNI Metadata interface ([I-D.ietf-cdni-metadata]), prepended by a DQUOTE and appended by a DQUOTE. If the CCID conveyed in the CDNI Metadata interface contains one or more DQUOTE, each DQUOTE MUST be escaped by an additional DQUOTE. For example, if the CCID conveyed in the CDNI Metadata interface is My_CCIDD"value", then the field value of the s-ccid is "My_CCID""value""".
  - occurrence: there MUST be zero or exactly one instance of this field. For a given <HTTP-header-name>, there MUST be zero or exactly one instance of this field.
- s-sid:
  - format: QSTRING
  - field value: this contains the value of a Session IDentifier (SID) generated by the dCDN for a specific HTTP session, prepended by a DQUOTE and appended by a DQUOTE. In particular, for HTTP Adaptive Streaming (HAS) session, the Session IDentifier value is included in the Logging record for every content chunk delivery of that session in view of facilitating the later correlation of all the per content chunk log records of a given HAS session. See section 3.4.2.2. of [RFC6983] for more discussion on the concept of Session IDentifier in the context of HAS. If the SID conveyed contains one or more DQUOTE, each DQUOTE MUST be escaped by an additional DQUOTE. For example, if the SID is My_SID"value", then the field value of the s-sid is "My_SID""value""".
  - occurrence: there MUST be zero or exactly one instance of this field.
- s-cached:
  - format: 1DIGIT
  - field value: this characterises whether the Surrogate served the request using content already stored on its local cache or not. The allowed values are "0" (for miss) and "1" (for hit). "1" MUST be used when the Surrogate did serve the request using exclusively content already stored on its local cache. "0" MUST

be used otherwise (including cases where the Surrogate served the request using some, but not all, content already stored on its local cache). Note that a "0" only means a cache miss in the Surrogate and does not provide any information on whether the content was already stored, or not, in another device of the dCDN, i.e., whether this was a "dCDN hit" or "dCDN miss".

* occurrence: there MUST be zero or exactly one instance of this field.

The "Fields" directive corresponding to a HTTP Request Logging Record MUST contain all the fields names whose occurrence is specified above as "There MUST be one and only one instance of this field". The corresponding fields value MUST be present in every HTTP Request Logging Record.

The "Fields" directive corresponding to a HTTP Request Logging Record MAY list all the fields value whose occurrence is specified above as "there MUST be zero or exactly one instance of this field" or "there MAY be zero, one or any number of instances of this field". The set of such field names actually listed in the "Fields" directive is selected by the CDN generating the CDNI Logging File based on agreements between the interconnected CDNs established through mechanisms outside the scope of this specification (e.g., contractual agreements). When such a field name is not listed in the "Fields" directive, the corresponding field value MUST NOT be included in the Logging Record. When such a field name is listed in the "Fields" directive, the corresponding field value MUST be included in the Logging Record; if the value for the field is not available, this MUST be conveyed via a dash character ("-").

The fields names listed in the "Fields" directive MAY be listed in the order in which they are listed in Section 3.4.1 or MAY be listed in any other order.

A dCDN-side implementation of the CDNI Logging interface MUST implement all the following Logging Fields in a CDNI Logging Record of Record-Type "cdni_http_request_v1", and MUST support the ability to include valid values for each of them:

o date

o time

o time-taken

o c-ip

- c-port
- s-ip
- s-hostname
- s-port
- cs-method
- cs-uri
- u-uri
- protocol
- sc-status
- sc-total-bytes
- sc-entity-bytes
- cs(<HTTP-header>)
- sc(<HTTP-header>)
- s-cached

A dCDN-side implementation of the CDNI Logging interface MAY support the following Logging Fields in a CDNI Logging Record of Record-Type "cdni_http_request_v1":

- c-ip-anonymizing
- s-ccid
- s-sid

If a dCDN-side implementation of the CDNI Logging interface supports these Fields, it MUST support the ability to include valid values for them.

An uCDN-side implementation of the CDNI Logging interface MUST be able to accept CDNI Logging Files with CDNI Logging Records of Record-Type "cdni_http_request_v1" containing any CDNI Logging Field defined in Section 3.4.1 as long as the CDNI Logging Record and the CDNI Logging File are compliant with the present document.

3.5. CDNI Logging File Example

```
#Version:<HTAB>CDNI/1.0<CRLF>

#UUID:<HTAB>urn:uuid:f81d4fae-7dec-11d0-a765-00a0c91e6bf6<CRLF>

#Claimed-Origin:<HTAB>cdni-logging-entity.dcdn.example.com<CRLF>

#Record-Type:<HTAB>cdni_http_request_v1<CRLF>

#Fields:<HTAB>date<HTAB>time<HTAB>time-taken<HTAB>c-ip<HTAB>cs-
method<HTAB>u-uri<HTAB>protocol<HTAB>sc-status<HTAB>sc-total-
bytes<HTAB>cs(User-Agent)<HTAB>cs(Referer)<HTAB>s-cached<CRLF>

2013-05-17<HTAB>00:38:06.825<HTAB>9.058<HTAB>10.5.7.1<HTAB>GET<HTAB>h
ttp://cdni-ucdn.dcdn.example.com/video/movie100.mp4<HTAB>HTTP/
1.1<HTAB>200<HTAB>6729891<HTAB>"Mozilla/5.0 (Windows; U; Windows NT
6.0; en-US) AppleWebKit/533.4 (KHTML, like Gecko) Chrome/5.0.375.127
Safari /533.4"<HTAB>"host1.example.com"<HTAB>1<CRLF>

2013-05-17<HTAB>00:39:09.145<HTAB>15.32<HTAB>10.5.10.5<HTAB>GET<HTAB>
http://cdni-ucdn.dcdn.example.com/video/movie118.mp4<HTAB>HTTP/
1.1<HTAB>200<HTAB>15799210<HTAB>"Mozilla/5.0 (Windows; U; Windows NT
6.0; en-US) AppleWebKit/533.4 (KHTML, like Gecko) Chrome/5.0.375.127
Safari /533.4"<HTAB>"host1.example.com"<HTAB>1<CRLF>

2013-05-17<HTAB>00:42:53.437<HTAB>52.879<HTAB>10.5.10.5<HTAB>GET<HTAB
>http://cdni-ucdn.dcdn.example.com/video/picture11.mp4<HTAB>HTTP/
1.0<HTAB>200<HTAB>97234724<HTAB>"Mozilla/5.0 (Windows; U; Windows NT
6.0; en-US) AppleWebKit/533.4 (KHTML, like Gecko) Chrome/5.0.375.127
Safari /533.4"<HTAB>"host5.example.com"<HTAB>0<CRLF>

#Integrity-Hash:<HTAB>fe113dfce8fec91323a4fc02261af26e<CRLF>
```

4. CDNI Logging File Exchange Protocol

This document specifies a protocol for the exchange of CDNI Logging Files as specified in Section 3.

This protocol comprises:

- a CDNI Logging feed, allowing the dCDN to notify the uCDN about the CDNI Logging Files that can be retrieved by that uCDN from the dCDN, as well as all the information necessary for retrieving each of these CDNI Logging Files. The CDNI Logging feed is specified in Section 4.1.

- a CDNI Logging File pull mechanism, allowing the uCDN to obtain from the dCDN a given CDNI Logging File at the uCDN's convenience. The CDNI Logging File pull mechanisms is specified in Section 4.2.

An implementation of the CDNI Logging interface on the dCDN side (the entity generating the CDNI Logging file) MUST support the server side of the CDNI Logging feed and the server side of the CDNI Logging pull mechanism.

An implementation of the CDNI Logging interface on the uCDN side (the entity consuming the CDNI Logging file) MUST support the client side of the CDNI Logging feed and the client side of the CDNI Logging pull mechanism.

We note that implementations of the CDNI Logging interface MAY also support other mechanisms to exchange CDNI Logging Files, for example in view of exchanging logging information with minimum time-lag (e.g., sub-minute or sub-second) between when the event occurred in the dCDN and when the corresponding Logging Record is made available to the uCDN (e.g., for log-consuming applications requiring extremely fresh logging information such as near-real-time content delivery monitoring). Such mechanisms are for further study and outside the scope of this document.

4.1. CDNI Logging Feed

The server-side implementation of the CDNI Logging feed MUST produce an Atom feed [RFC4287]. This feed is used to advertise log files that are available for the client-side to retrieve using the CDNI Logging pull mechanism.

4.1.1. Atom Formatting

A CDNI Logging feed MUST be structured as an Archived feed, as defined in [RFC5005], and MUST be formatted in Atom [RFC4287]. This means it consists of a subscription document that is regularly updated as new CDNI Logging Files become available, and information about older CDNI Logging files is moved into archive documents. Once created, archive documents are never modified.

Each CDNI Logging File listed in an Atom feed MUST be described in an atom:entry container element.

The atom:entry MUST contain an atom:content element whose "src" attribute is a link to the CDNI Logging File and whose "type" attribute is the MIME Media Type indicating that the entry is a CDNI Logging File. We define this MIME Media Type as "application/cdni.LoggingFile" (See Section 5.4).

For compatibility with some Atom feed readers the atom:entry MAY also contain an atom:link entry whose "href" attribute is a link to the CDNI Logging File and whose "type" attribute is the MIME Media Type indicating that the entry is a CDNI Logging File using the "application/cdni.LoggingFile" MIME Media Type (See Section 5.4).

The URI used in the atom:id of the atom:entry MUST contain the UUID of the CDNI Logging File.

The atom:updated in the atom:entry MUST indicate the time at which the CDNI Logging File was last updated.

4.1.2. Updates to Log Files and the Feed

CDNI Logging Files MUST NOT be modified by the dCDN once published in the CDNI Logging feed.

The frequency with which the subscription feed is updated, the period of time covered by each CDNI Logging File or each archive document, and timeliness of publishing of CDNI Logging Files are outside the scope of the present document and are expected to be agreed upon by uCDN and dCDN via other means (e.g., human agreement).

The server-side implementation SHOULD use HTTP cache control headers on the subscription feed to indicate the frequency at which the client-side is to poll for updates.

The potential retention limits (e.g., sliding time window) within which the dCDN is to retain and be ready to serve an archive document is outside the scope of the present document and is expected to be agreed upon by uCDN and dCDN via other means (e.g., human agreement). The server-side implementation MUST retain, and be ready to serve, any archive document within the agreed retention limits. Outside these agreed limits, the server-side implementation MAY indicate its inability to serve (e.g., with HTTP status code 404) an archive document or MAY refuse to serve it (e.g., with HTTP status code 403 or 410).

4.1.3. Redundant Feeds

The server-side implementation MAY present more than one CDNI Logging feed and for redundancy. CDNI Logging Files MAY be published in more than one feed.

A client-side implementation MAY support such redundant CDNI Logging feeds. If it supports redundant CDNI Logging feed, the client-side SHOULD use the UUID of the CDNI Logging File, presented in the

atom:id element of the Atom feed, to avoid unnecessarily pulling and storing each CDNI Logging File more than once.

4.1.4. Example CDNI Logging Feed

Figure 4 illustrates an example of the subscription document of a CDNI Logging feed.

```
<?xml version="1.0" encoding="utf-8"?>
<feed xmlns="http://www.w3.org/2005/Atom"
<http://www.w3.org/2005/Atom%22>>
  <title type="text">CDNI Logging Feed</title>
  <updated>2013-03-23T14:46:11Z</updated>
  <id>urn:uuid:663ae677-40fb-e99a-049d-c5642916b8ce</id>
  <link href="https://dcdn.example/logfeeds/ucdn1"
     rel="self" type="application/atom+xml" />
  <link href="https://dcdn.example/logfeeds/ucdn1"
     rel="current" type="application/atom+xml" />
  <link href="https://dcdn.example/logfeeds/ucdn1/201303231400"
     rel="prev-archive" type="application/atom+xml" />
  <generator version="example version 1">CDNI Log Feed
     Generator</generator>
  <author><name>dcdn.example</name></author>
  <entry>
    <title type="text">CDNI Logging File for uCDN at
      2013-03-23 14:15:00</title>
      <id>urn:uuid:12345678-1234-abcd-00aa-01234567abcd</id>
      <updated>2013-03-23T14:15:00Z</updated>
      <content src="https://dcdn.example/logs/ucdn/
         http-requests-20130323141500000000"
         type="application/cdni.LoggingFile" />
      <summary>CDNI Logging File for uCDN at
      2013-03-23 14:15:00</summary>
  </entry>
  <entry>
    <title type="text">CDNI Logging File for uCDN at
      2013-03-23 14:30:00</title>
      <id>urn:uuid:87654321-4321-dcba-aa00-dcba7654321</id>
      <updated>2013-03-23T14:30:00Z</updated>
      <content src="https://dcdn.example/logs/ucdn/
         http-requests-20130323143000000000"
         type="application/cdni.LoggingFile" />
      <summary>CDNI Logging File for uCDN at
      2013-03-23 14:30:00</summary>
  </entry>
  ...
  <entry>
    ...
  </entry>
</feed>
```

Figure 4: Example subscription document of a CDNI Logging Feed

4.2. CDNI Logging File Pull

A client-side implementation of the CDNI Logging interface MAY pull, at its convenience, a CDNI Logging File that is published by the server-side in the CDNI Logging Feed (in the subscription document or an archive document). To do so, the client-side:

- MUST use HTTP v1.1 [RFC2616];
- SHOULD use TLS (i.e., use what is loosely referred to as "HTTPS") as per [RFC2818] whenever protection of the CDNI Logging information is required (see Section 6.1);
- MUST use the URI that was associated to the CDNI Logging File (within the "src" attribute of the corresponding atom:content element) in the CDNI Logging Feed;
- MUST support exchange of CDNI Logging Files with no content encoding applied to the representation;
- SHOULD support exchange of CDNI Logging Files with "gzip" content encoding (as defined in [RFC2616]) applied to the representation.

Note that a client-side implementation of the CDNI Logging interface MAY pull a CDNI Logging File that it has already pulled.

The server-side implementation MUST respond to valid pull request by a client-side implementation for a CDNI Logging File published by the server-side in the CDNI Logging Feed (in the subscription document or an archive document). The server-side implementation:

- MUST handle the client-side request as per HTTP v1.1;
- MUST include the CDNI Logging File identified by the request URI inside the body of the HTTP response;
- MUST support exchange of CDNI Logging Files with no content encoding applied to the representation;
- SHOULD support exchange of CDNI Logging Files with "gzip" content encoding (as defined in [RFC2616]) applied to the representation.

Content negotiation approaches defined in [RFC2616] (e.g., using Accept-Encoding request-header field or Content-Encoding entity-header field) MAY be used by the client-side and server-side implementations to establish the content-coding to be used for a particular exchange of a CDNI Logging File.

Applying compression content encoding (such as "gzip") is expected to mitigate the impact of exchanging the large volumes of logging information expected across CDNs. This is expected to be particularly useful in the presence of HTTP Adaptive Streaming (HAS) which, as per the present version of the document, will result in a separate CDNI Log Record for each HAS segment delivery in the CDNI Logging File.

The potential retention limits (e.g., sliding time window, maximum aggregate file storage quotas) within which the dCDN is to retain and be ready to serve a CDNI Logging File previously advertised in the CDNI Logging Feed is outside the scope of the present document and is expected to be agreed upon by uCDN and dCDN via other means (e.g., human agreement). The server-side implementation MUST retain, and be ready to serve, any CDNI Logging File within the agreed retention limits. Outside these agreed limits, the server-side implementation MAY indicate its inability to serve (e.g., with HTTP status code 404) a CDNI Logging File or MAY refuse to serve it (e.g., with HTTP status code 403 or 410).

5. IANA Considerations

5.1. CDNI Logging Directive Names Registry

The IANA is requested to create a new registry, CDNI Logging Directive Names.

The initial contents of the CDNI Logging File Directives registry comprise the names of the directives specified in Section 3.3 of the present document, and are as follows:

Directive Name	Reference
Version	RFC xxxx
UUID	RFC xxxx
Claimed-Origin	RFC xxxx
Verified-Origin	RFC xxxx
Record-Type	RFC xxxx
Fields	RFC xxxx
Integrity-Hash	RFC xxxx

Figure 5

[Instructions to IANA: Replace "RFC xxxx" above by the RFC number of the present document]

Within the registry, names are to be allocated by IANA according to the "Specification Required" policy specified in [RFC5226]. Directive names are to be allocated by IANA with a format of NAMEFORMAT (see Section 3.1).

Each specification that defines a new CDNI Logging directive needs to contain a description for the new directive with the same set of information as provided in Section 3.3 (i.e., format, directive value and occurrence).

5.2. CDNI Logging Record-Types Registry

The IANA is requested to create a new registry, CDNI Logging Record-Types.

The initial contents of the CDNI Logging Record-Types registry comprise the names of the CDNI Logging Record types specified in Section 3.4 of the present document, and are as follows:

Record-Types	Reference
cdni_http_request_v1	RFC xxxx

Figure 6

[Instructions to IANA: Replace "RFC xxxx" above by the RFC number of the present document]

Within the registry, Record-Types are to be allocated by IANA according to the "Specification Required" policy specified in [RFC5226]. Record-Types are to be allocated by IANA with a format of NAMEFORMAT (see Section 3.1).

Each specification that defines a new Record-Type needs to contain a description for the new Record-Type with the same set of information as provided in Section 3.4.1. This includes:

- a list of all the CDNI Logging Fields that can appear in a CDNI Logging Record of the new Record-Type
- for all these Fields: a specification of the occurrence for each Field in the new Record-Type
- for every newly defined Field, i.e., for every Field that results in a registration in the CDNI Logging Field Names Registry

(Section 5.3): a specification of the field name, format and field value.

5.3. CDNI Logging Field Names Registry

The IANA is requested to create a new registry, CDNI Logging Field Names.

The initial contents of the CDNI Logging Fields Names registry comprise the names of the CDNI Logging fields specified in Section 3.4 of the present document, and are as follows:

Field Name	Reference
date	RFC xxxx
time	RFC xxxx
time-taken	RFC xxxx
c-ip	RFC xxxx
c-ip-anonymizing	RFC xxxx
c-port	RFC xxxx
s-ip	RFC xxxx
s-hostname	RFC xxxx
s-port	RFC xxxx
cs-method	RFC xxxx
cs-uri	RFC xxxx
u-uri	RFC xxxx
protocol	RFC xxxx
sc-status	RFC xxxx
sc-total-bytes	RFC xxxx
sc-entity-bytes	RFC xxxx
cs(<HTTP-header>)	RFC xxxx
sc(<HTTP-header>)	RFC xxxx
s-ccid	RFC xxxx
s-sid	RFC xxxx
s-cached	RFC xxxx

Figure 7

[Instructions to IANA: Replace "RFC xxxx" above by the RFC number of the present document]

Within the registry, names are to be allocated by IANA according to the "Specification Required" policy specified in [RFC5226]. Field names are to be allocated by IANA with a format of NHTABSTRING (see Section 3.1).

The above registry is intended to be shared across the currently defined Record-Type (i.e., cdni_http_request_v1) as well as potential other CDNI Logging Record-Types that may be defined in separate specifications. When a Field from this registry is used by another CDNI Logging Record-Type, it is to be used with the exact semantics and format specified in the document that registered this field and that is identified in the Reference column of the registry. If another CDNI Logging Record-Type requires a Field with semantics that are not strictly identical, or a format that is not strictly identical then this new Field is to be registered in the registry with a different Field name. When a Field from this registry is used by another CDNI Logging Record-Type, it can be used with different occurence rules.

5.4. CDNI Logging MIME Media Type

The IANA is requested to allocate the "application/cdni.LoggingFile" MIME Media Type (whose use is specified in Section 4.1.1 of the present document) in the MIME Media Types registry.

6. Security Considerations

6.1. Authentication, Confidentiality, Integrity Protection

A CDNI Logging implementation MUST support TLS transport of the CDNI Logging feed (Section 4.1) and of the CDNI Logging File pull (Section 4.2) as per [RFC2818].

The use of TLS for transport of the CDNI Logging feed and CDNI Logging File pull allows:

- the dCDN and uCDN to authenticate each other (to ensure they are transmitting/receiving CDNI Logging File from an authenticated CDN)
- the CDNI Logging information to be transmitted with confidentiality
- the integrity of the CDNI Logging information to be protected during the exchange

In an environment where any such protection is required, TLS SHOULD be used for transport of the CDNI Logging feed and the CDNI Logging File pull unless alternate methods are used for ensuring the confidentiality of the information in the logging files (such as setting up an IPsec tunnel between the two CDNs or using a physically secured internal network between two CDNs that are owned by the same

corporate entity). Both parties of the transaction (uCDN and dCDN) SHOULD use mutual authentication.

A CDNI Logging implementation MUST support the TLS_DHE_RSA_WITH_AES_128_GCM_SHA256 cipher suite ([RFC5288]). An implementation of the CDNI Logging Interface SHOULD prefer cipher suites which support perfect forward secrecy over cipher suites that don't.

The Integrity-Hash directive inside the CDNI Logging File provides additional integrity protection, this time targeting potential corruption of the CDNI logging information during the CDNI Logging File generation. This mechanism does not allow restoration of the corrupted CDNI Logging information, but it allows detection of such corruption and therefore triggering of appropriate correcting actions (e.g., discard of corrupted information, attempt to re-obtain the CDNI Logging information).

6.2. Denial of Service

This document does not define specific mechanism to protect against Denial of Service (DoS) attacks on the Logging Interface. However, the CDNI Logging feed and CDNI Logging pull endpoints can be protected against DoS attacks through the use of TLS transport and/or via mechanisms outside the scope of the CDNI Logging interface such as firewalling or use of Virtual Private Networks (VPNs).

Protection of dCDN Surrogates against spoofed delivery requests is outside the scope of the CDNI Logging interface.

6.3. Privacy

CDNs have the opportunity to collect detailed information about the downloads performed by End Users. The provision of this information to another CDN introduces potential End Users privacy protection concerns.

The use of TLS for transport of the CDNI Logging feed and CDNI Logging pull as discussed in Section 6.1 protects the confidentiality of logged information by preventing any other party than the authorised uCDN to gain access to the logging information.

We observe that when CDNI interconnection is realised as per [I-D.ietf-cdni-framework], the uCDN handles the initial End User requests (before it is redirected to the dCDN) so, regardless of which information is, or is not, communicated to the uCDN through the CDNI Logging interface, the uCDN has visibility on significant

information such as the IP address of the End User request and the URL of the request.

Nonetheless, if the dCDN and uCDN agree that anonymization is required to avoid making some detailed information available to the uCDN (such as how many bytes of the content have been watched by an End User and/or at what time) or is required to meet some legal obligations, then the uCDN and dCDN can agree to exchange anonymized End Users IP address in CDNI Logging Files and the c-ip-anonymization field can be used to convey the number of bits that have been anonymized so that the meaningful information can still be easily extracted from the anonymized addressses (e.g., for geolocation aware analytics).

We note that anonymization of End Users IP address does not fully protect against deriving potentially sensitive information about traffic patterns; in general, increasing the number of bits that are anonymized can mitigate the risks of deriving such sensitive traffic pattern information.

We also note that independently of IP addresses, the query string portion of the URL that may be conveyed inside the cs-uri and u-uri fields of CDNI Logging Files, or the HTTP cookies([RFC6265]) that may be conveyed inside the cs(<HTTP-header-name>) field of CDNI Logging Fields, may contain personnal information or information that can be exploited to derive personal information. Where this is a concern, the CDNI Logging interface specification allows the dCDN to not include the cs-uri and to include a u-uri that removes (or hides) the sensitive part of the query string and allows the dCDN to not include the cs(<HTTP-header-name>) fields corresponding to HTTP headers associated with cookies.

7. Acknowledgments

This document borrows from the W3C Extended Log Format [ELF].

Rob Murray significantly contributed into the text of Section 4.1.

The authors thank Ben Niven-Jenkins, Kevin Ma, David Mandelberg and Ray van Brandenburg for their ongoing input.

Finally, we also thank Sebastien Cubaud, Pawel Grochocki, Christian Jacquenet, Yannick Le Louedec, Anne Marrec , Emile Stephan, Fabio Costa, Sara Oueslati, Yvan Massot, Renaud Edel, Joel Favier and the contributors of the EU FP7 OCEAN project for their input in the early versions of this document.

8. References

8.1. Normative References

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

[RFC2616] Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P., and T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1", RFC 2616, June 1999.

[RFC3986] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", STD 66, RFC 3986, January 2005.

[RFC4122] Leach, P., Mealling, M., and R. Salz, "A Universally Unique IDentifier (UUID) URN Namespace", RFC 4122, July 2005.

[RFC4287] Nottingham, M., Ed. and R. Sayre, Ed., "The Atom Syndication Format", RFC 4287, December 2005.

[RFC5005] Nottingham, M., "Feed Paging and Archiving", RFC 5005, September 2007.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.

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

[RFC5288] Salowey, J., Choudhury, A., and D. McGrew, "AES Galois Counter Mode (GCM) Cipher Suites for TLS", RFC 5288, August 2008.

8.2. Informative References

[CHAR_SET] "IANA Character Sets registry", <http://www.iana.org/assignments/character-sets/character-sets.xml>.

[ELF] Phillip M. Hallam-Baker, and Brian Behlendorf, "Extended Log File Format, W3C (work in progress), WD-logfile-960323", <http://www.w3.org/TR/WD-logfile.html>.

[I-D.ietf-cdni-framework]
Peterson, L., Davie, B., and R. Brandenburg, "Framework for CDN Interconnection", draft-ietf-cdni-framework-10 (work in progress), March 2014.

[I-D.ietf-cdni-metadata]
Niven-Jenkins, B., Murray, R., Watson, G., Caulfield, M., Leung, K., and K. Ma, "CDN Interconnect Metadata", draft-ietf-cdni-metadata-06 (work in progress), February 2014.

[I-D.ietf-cdni-requirements]
Leung, K. and Y. Lee, "Content Distribution Network Interconnection (CDNI) Requirements", draft-ietf-cdni-requirements-17 (work in progress), January 2014.

[I-D.ietf-httpbis-http2]
Belshe, M., Peon, R., and M. Thomson, "Hypertext Transfer Protocol version 2", draft-ietf-httpbis-http2-10 (work in progress), February 2014.

[I-D.snell-atompub-link-extensions]
Snell, J., "Atom Link Extensions", draft-snell-atompub-link-extensions-09 (work in progress), June 2012.

[RFC1321] Rivest, R., "The MD5 Message-Digest Algorithm", RFC 1321, April 1992.

[RFC1945] Berners-Lee, T., Fielding, R., and H. Nielsen, "Hypertext Transfer Protocol -- HTTP/1.0", RFC 1945, May 1996.

[RFC2818] Rescorla, E., "HTTP Over TLS", RFC 2818, May 2000.

[RFC6265] Barth, A., "HTTP State Management Mechanism", RFC 6265, April 2011.

[RFC6707] Niven-Jenkins, B., Le Faucheur, F., and N. Bitar, "Content Distribution Network Interconnection (CDNI) Problem Statement", RFC 6707, September 2012.

[RFC6770] Bertrand, G., Stephan, E., Burbridge, T., Eardley, P., Ma, K., and G. Watson, "Use Cases for Content Delivery Network Interconnection", RFC 6770, November 2012.

[RFC6983] van Brandenburg, R., van Deventer, O., Le Faucheur, F., and K. Leung, "Models for HTTP-Adaptive-Streaming-Aware Content Distribution Network Interconnection (CDNI)", RFC 6983, July 2013.

Authors' Addresses

Francois Le Faucheur (editor)
Cisco Systems
E.Space Park - Batiment D
6254 Allee des Ormes - BP 1200
Mougins cedex 06254
FR

Phone: +33 4 97 23 26 19
Email: flefauch@cisco.com

Gilles Bertrand (editor)
Orange
38-40 rue du General Leclerc
Issy les Moulineaux 92130
FR

Phone: +33 1 45 29 89 46
Email: gilles.bertrand@orange.com

Iuniana Oprescu (editor)
Orange
38-40 rue du General Leclerc
Issy les Moulineaux 92130
FR

Phone: +33 6 89 06 92 72
Email: iuniana.oprescu@orange.com

Roy Peterkofsky
Skytide, Inc.
One Kaiser Plaza, Suite 785
Oakland CA 94612
USA

Phone: +01 510 250 4284
Email: roy@skytide.com