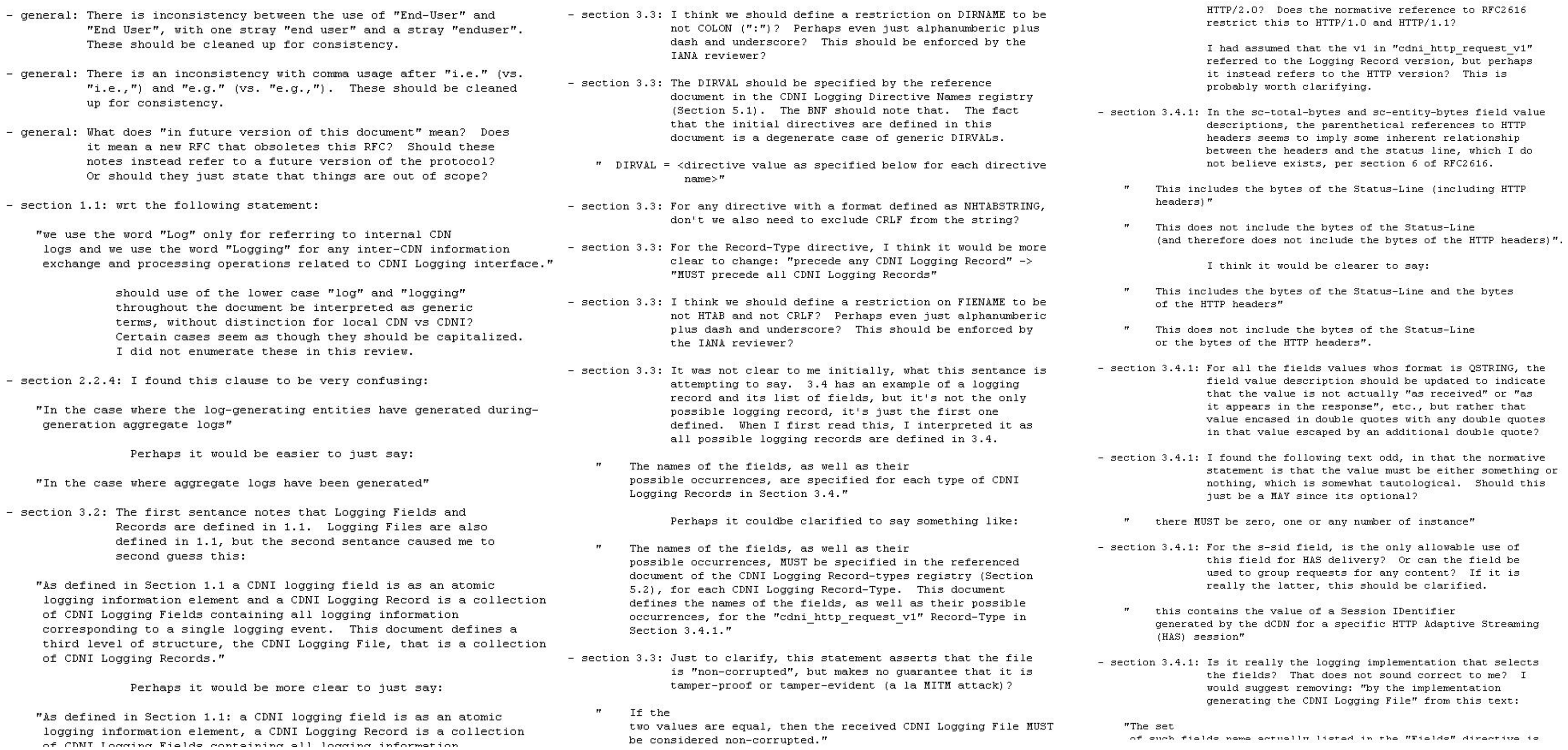
CDNI Logging
draft-ietf-cdni-logging-10

François Le Faucheur – Cisco
Gilles Bertrand - Orange
Iuniana Oprescu - Orange
Roy Peterkofsky – Skytide

IETF 89 London 2014

Changes between -08 and -09

- As agreed in Vancouver, Kevin Ma conducted a full review, before we move to WG Last Call
- A lot of comments given by Kevin Ma
- Fixes provided in 4 batches + follow up discussions
- Some items discussed on the list along the way



```
- general: There is inconsistency between the use of "End-User" and
           "End User", with one stray "end user" and a stray "enduser".
           These should be cleaned up for consistency.

- general: There is an inconsistency with comma usage after "i.e." (vs.
           "i.e.,") and "e.g." (vs. "e.g.,").  These should be cleaned
           up for consistency.

- general: What does "in future version of this document" mean?  Does
           it mean a new RFC that obsoletes this RFC?  Should these
           notes instead refer to a future version of the protocol?
           Or should they just state that things are out of scope?

- section 1.1: wrt the following statement:

    "we use the word "Log" only for referring to internal CDN
     logs and we use the word "Logging" for any inter-CDN information
     exchange and processing operations related to CDNI Logging interface."

               should use of the lower case "log" and "logging"
               throughout the document be interpreted as generic
               terms, without distinction for local CDN vs CDNI?
               Certain cases seem as though they should be capitalized.
               I did not enumerate these in this review.

- section 2.2.4: I found this clause to be very confusing:

    "In the case where the log-generating entities have generated during-
     generation aggregate logs"

                Perhaps it would be easier to just say:

    "In the case where aggregate logs have been generated"

- section 3.2: The first sentance notes that Logging Fields and
               Records are defined in 1.1.  Logging Files are also
               defined in 1.1, but the second sentance caused me to
               second guess this:

    "As defined in Section 1.1 a CDNI logging field is as an atomic
     logging information element and a CDNI Logging Record is a collection
     of CDNI Logging Fields containing all logging information
     corresponding to a single logging event.  This document defines a
     third level of structure, the CDNI Logging File, that is a collection
     of CDNI Logging Records."

                Perhaps it would be more clear to just say:

    "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
```

```
- section 3.3: I think we should define a restriction on DIRNAME to be
               not COLON (":")?  Perhaps even just alphanumberic plus
               dash and underscore?  This should be enforced by the
               IANA reviewer?

- section 3.3: The DIRVAL should be specified by the reference
               document in the CDNI Logging Directive Names registry
               (Section 5.1).  The BNF should note that.  The fact
               that the initial directives are defined in this
               document is a degenerate case of generic DIRVALs.

    "  DIRVAL = <directive value as specified below for each directive
                 name>"

- section 3.3: For any directive with a format defined as NHTABSTRING,
               don't we also need to exclude CRLF from the string?

- section 3.3: For the Record-Type directive, I think it would be more
               clear to change: "precede any CDNI Logging Record" ->
               "MUST precede all CDNI Logging Records"

- section 3.3: I think we should define a restriction on FIENAME to be
               not HTAB and not CRLF?  Perhaps even just alphanumberic
               plus dash and underscore?  This should be enforced by
               the IANA reviewer?

- section 3.3: It was not clear to me initially, what this sentance is
               attempting to say.  3.4 has an example of a logging
               record and its list of fields, but it's not the only
               possible logging record, it's just the first one
               defined.  When I first read this, I interpreted it as
               all possible logging records are defined in 3.4.

    "    The names of the fields, as well as their
         possible occurrences, are specified for each type of CDNI
         Logging Records in Section 3.4."

               Perhaps it couldbe clarified to say something like:

    "    The names of the fields, as well as their
         possible occurrences, MUST be specified in the referenced
         document of the CDNI Logging Record-types registry (Section
         5.2), for each CDNI Logging Record-Type.  This document
         defines the names of the fields, as well as their possible
         occurrences, for the "cdni_http_request_v1" Record-Type in
         Section 3.4.1."

- section 3.3: Just to clarify, this statement asserts that the file
               is "non-corrupted", but makes no guarantee that it is
               tamper-proof or tamper-evident (a la MITM attack)?

    "    If the
         two values are equal, then the received CDNI Logging File MUST
         be considered non-corrupted."
```

```
                 HTTP/2.0?  Does the normative reference to RFC2616
                 restrict this to HTTP/1.0 and HTTP/1.1?

                 I had assumed that the v1 in "cdni_http_request_v1"
                 referred to the Logging Record version, but perhaps
                 it instead refers to the HTTP version?  This is
                 probably worth clarifying.

- section 3.4.1: In the sc-total-bytes and sc-entity-bytes field value
                 descriptions, the parenthetical references to HTTP
                 headers seems to imply some inherent relationship
                 between the headers and the status line, which I do
                 not believe exists, per section 6 of RFC2616.

    "    This includes the bytes of the Status-Line (including HTTP
         headers)"

    "    This does not include the bytes of the Status-Line
         (and therefore does not include the bytes of the HTTP headers)".

                 I think it would be clearer to say:

    "    This includes the bytes of the Status-Line and the bytes
         of the HTTP headers"

    "    This does not include the bytes of the Status-Line
         or the bytes of the HTTP headers".

- section 3.4.1: For all the fields values whos format is QSTRING, the
                 field value description should be updated to indicate
                 that the value is not actually "as received" or "as
                 it appears in the response", etc., but rather that
                 value encased in double quotes with any double quotes
                 in that value escaped by an additional double quote?

- section 3.4.1: I found the following text odd, in that the normative
                 statement is that the value must be either something or
                 nothing, which is somewhat tautological.  Should this
                 just be a MAY since its optional?

    "    there MUST be zero, one or any number of instance"

- section 3.4.1: For the s-sid field, is the only allowable use of
                 this field for HAS delivery?  Or can the field be
                 used to group requests for any content?  If it is
                 really the latter, this should be clarified.

    "    this contains the value of a Session IDentifier
         generated by the dCDN for a specific HTTP Adaptive Streaming
         (HAS) session"

- section 3.4.1: Is it really the logging implementation that selects
                 the fields?  That does not sound correct to me?  I
                 would suggest removing: "by the implementation
                 generating the CDNI Logging File" from this text:

    "The set
     of such fields name actually listed in the "Fields" directive is
```

Changes between -08 and -09

- DIRNAME & DIRVAL format specifications
- NHTABSTRING excludes CRLF
- QSTRING format example for values “as received “ -> need to escape DQUOTES
- Extended *s*-sid to any HTTP session (previously used just for HAS)

Changes between -08 and -09

- Clarification on Mandatory-to-Implement fields and the associated valid values
- Explained that for the ATOM protocol server-side = dCDN, client-side = uCDN
- Field Names are valid beyond a given Record Type, but should have the same meaning

Changes between -08 and -09

- Added text to Privacy section
  - "anonymization of End Users IP address does not fully protect against deriving potentially sensitive information about traffic patterns"
  - the query string portion of the URL that may be conveyed inside the cs-uri and u-uri fields or the HTTP cookies may contain personnal information or information that can be exploited to derive personal information => better use u-uri that obfuscates the sensitive part of the request
- Editorial tweaks, clarifications, consistency…

Changes between -09 and -10

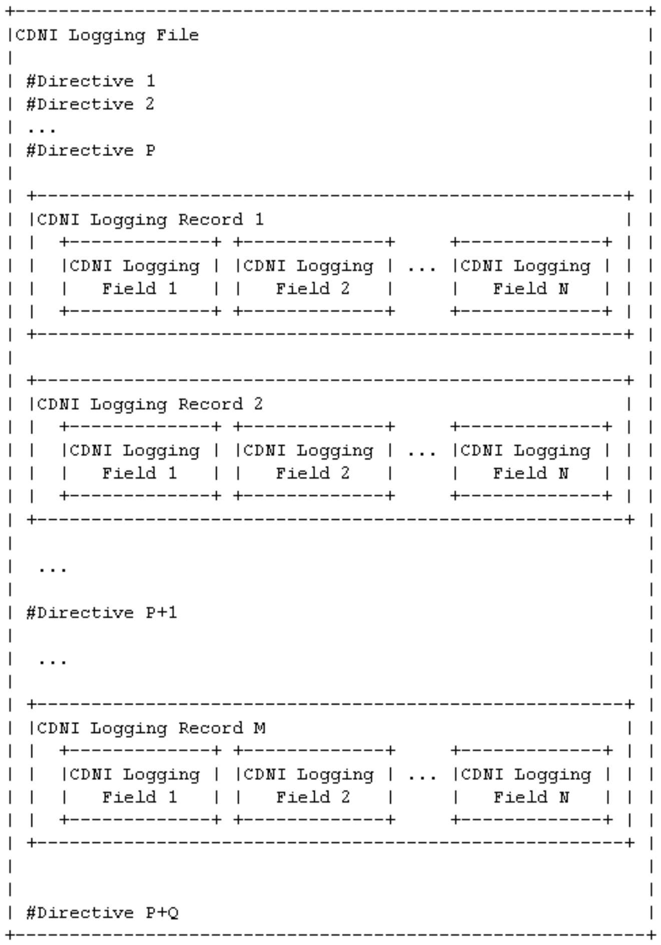
- Removed RFC2119 normative language from IANA section (see discussion on list)
- Duplicate <HTTP-header-name> unlikely to "usefully" happen => occurrence is zero or one
- In NHTAB definition, corrected "CRLF" into "CR" and "LF"
- More editorial tweaks

Next Steps

- **New rev to add statement about applicability of "cdni_http_request_v1" Logging Record**
  - To HTTP/1.0 and 1.1
  - To HTTP/2.0: same semantics as HTTP/1.1 so "cdni_http_request_v1" Logging Record is valid. From an operational point of view it is interesting to log additional information about a HTTP/2.0 delivery (mnot), so a future "cdni_http_request_v2" Logging Record may introduce additional fields for that
  - To HTTPS: "cdni_http_request_v1" Logging Record is valid. There might be an advantage to log more information, so a future "cdni_http_request_v2" Logging Record may introduce additional fields for that
- **Request Working Group Last Call**

CDNI Logging File Directives

- Version
- UUID
- Claimed-Origin
- Verified-Origin
- Record-Type
- Fields
- Integrity-Hash



```
+----------------------------------------------------------+
|CDNI Logging File                                         |
|                                                          |
| #Directive 1                                             |
| #Directive 2                                             |
| ...                                                      |
| #Directive P                                             |
|                                                          |
| +------------------------------------------------------+ |
| |CDNI Logging Record 1                                 | |
| |  +-------------+ +-------------+     +-------------+ | |
| |  |CDNI Logging | |CDNI Logging | ... |CDNI Logging | | |
| |  |   Field 1   | |   Field 2   |     |   Field N   | | |
| |  +-------------+ +-------------+     +-------------+ | |
| +------------------------------------------------------+ |
|                                                          |
| +------------------------------------------------------+ |
| |CDNI Logging Record 2                                 | |
| |  +-------------+ +-------------+     +-------------+ | |
| |  |CDNI Logging | |CDNI Logging | ... |CDNI Logging | | |
| |  |   Field 1   | |   Field 2   |     |   Field N   | | |
| |  +-------------+ +-------------+     +-------------+ | |
| +------------------------------------------------------+ |
|                                                          |
|  ...                                                     |
|                                                          |
| #Directive P+1                                           |
|                                                          |
|  ...                                                     |
|                                                          |
| +------------------------------------------------------+ |
| |CDNI Logging Record M                                 | |
| |  +-------------+ +-------------+     +-------------+ | |
| |  |CDNI Logging | |CDNI Logging | ... |CDNI Logging | | |
| |  |   Field 1   | |   Field 2   |     |   Field N   | | |
| |  +-------------+ +-------------+     +-------------+ | |
| +------------------------------------------------------+ |
|                                                          |
|                                                          |
| #Directive P+Q                                           |
+----------------------------------------------------------+
```

Current HTTP Logging Fields

- Date
- Time
- Time-taken
- C-IP
- C-IP-anonimizing
- 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-name>)
- SC(<HTTP-header-name>)
- S-CCID
- S-SID
- S-cached
- S-URI-signing

where C: client (User Agent), S: server (dCDN surrogate), U: uCDN, D: dCDN