Proposed Format For Presence Information
draft-hudson-impp-presence-01.txt

1. Status of this Memo

This document is an Internet-Draft and is in full conformance with all
provisions of Section 10 of RFC2026.

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

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

The list of current Internet-Drafts can be accessed at
http://www.ietf.org/ietf/1id-abstracts.txt

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

Please send comments to the IMPP working group at impp@iastate.edu.

2. Abstract

This document proposes a syntax and initial tag set for presence
information to be used in the IMPP protocol suite. The encoding is a
subset of well-formed but not valid [XML] documents, such that it can
be parsed either by a simple hand-written parser or by an XML
implementation.

3. Terminology

The following terms are defined in [Model] and are used with those
definitions in this document:

PRESENTITY
PRINCIPAL
WATCHER USER AGENT

However, those terms are used in lowercase for improved readability,
since they are relatively distinctive.

The terms MUST, SHOULD, and MAY are used in uppercase with the meaning
defined in [RFC 2119].

[4](#). **Syntax**

[POINT OF CONTENTION: Some have argued that we should define our
syntax by referring to XML and adding restrictions so that we don't
accidentally introduce variations. My view is that this would force
implementors to consult the full XML spec; having a self-contained,
reduced grammar seems more conducive to implementations.]

[POINT OF CONTENTION: Several people think we should use the MIME type
application/presence-xml or try to get presence/xml to allow for
future media types which encode presence information differently.
Precedents like application/pip suggest to me that
"application/presence" is more along the lines of common practice than
creating a whole new hierarchy of different encodings.]

Presence information is a MIME [RFC 2045-2049] object of type
application/presence. The contents of the MIME object is a presence
document. The underlying character set for a presence document is
[Unicode], which will be represented in UTF-8 or as determined
otherwise by a charset parameter in the media type of the MIME object.
Following is an ABNF [RFC 2234] grammar describing the syntax for
presence information:

```
        presence-doc    = "<presence>" content "</presence>"
        content         = *(element / char-data / reference)
        element         = empty-tag / start-tag content end-tag
                          ; end-tag name must match start-tag name.
        empty-tag       = "<" name "/>"
        start-tag       = "<" name ">"
        end-tag         = "</" name ">"
        name            = (Letter / "_" / ":") *NameChar
        char-data       = 1*DataChar
                          ; "]]>" may not appear, for compatibility
                          ; with SGML.
        reference       = char-ref / entity-ref
        char-ref        = "&#" 1*ASCIIDigit ';' /
                          "&#x" 1*ASCIIHexDigit
                          ; Must refer to a valid Char
        entity-ref      = "&lt;" / "&gt;" / "&amp;" / "&apos;" /
                          "&quot;"
```

The character classes Letter, Digit, CombiningChar, and Extender are
defined in [XML] Appendix B. The other character classes are defined
as follows:

```
        NameChar        = Letter / Digit / "." / "-" / "_" / ":" /
                          CombiningChar / Extender
        DataChar        = %x9 / %xA / %xD / %x20-25 / %x27-3B /
                          %x3D-D7FF / %xE000-FFFD / %x10000-310FFFF
                          ; Most valid Unicode characters
        Char            = DataChar / "&" / "<"
```

```
        ASCIIDigit      = %x30-39
                          ; [0-9]
        ASCIIHexDigit   = %x30-39 / %x41-46 / %x61-66
                          ; [0-9A-Fa-f]
```

A char-ref refers to a Unicode character by number, either in decimal
("&#" prefix) or in hexadecimal ("&#x" prefix). An entity-ref refers
to a specific Unicode character by name, as follows:

```
        entity-ref      Character
        ----------      ---------
        &lt;            <
        &gt;            >
        &amp;           &
        &apos;          '
        &quot;          "
```

5. Syntactic interpretation

After parsing, a presence document consists of a tree of elements,
where each element consists of a name (or "tag"), text (the
concatenation of all char-data and reference productions in the
element's content, but not char-data and reference productions inside
sub-elements), and an ordered list of child elements. For example,
the presence document:

```
        <presence>a&lt;a<foo/>bbb<bar>ccc</bar>ddd</presence>
```

parses into a tree of three elements named "presence", "foo", and
"bar", and which can be viewed pictorially as:

```
         presence
        "a<abbbddd"
             |
         ---------
         |       |
        foo     bar
        ""      "ccc"
```

A watcher user agent MUST discard an element, including all text and
sub-elements inside that element, if it does not recognize the
element's tag in that element's context. For instance, if a watcher
user agent recognizes the tag "foo" in the context of a "presence"
element but does not recognize the tag "bar", it MUST treat the
presence document from the previous example as equivalent to:

```
        <presence>a&lt;a<foo/>bbbddd</presence>
```

6. Tag set

Some tag definitions include a list of constraints on that element's
children. If an element's children do not meet the specified

constraints, the watcher user agent MUST discard that element.

```
Tag:           presence
Context:       (top level)
Sub-elements:  date presentity location status contact
               [XXX Do we want sub-elements here for personal
               information, or is that out of scope for presence?]
Constraints:   time and presentity must appear exactly once.
               location and status must appear at most once.
Description:   This tag introduces the presence document.  Any text
               in the element is discarded.


Tag:           date
Context:       presence
Description:   The text of this element gives the date and time for
               which the presence information is being reported.
               [XXX Obviously we need to pick a standard format, but
               the details are unimportant at this stage.]


Tag:           presentity
Context:       presence
Description:   The text of this element specifies the identifier of the
               presentity whose presence is being reported.


Tag:           location
Context:       presence
Description:   The text of this element specifies the location of the
               principal as a human-readable description.
               [XXX Open issue: is it useful to define a
               human-readable field like this and restrict it to flat
               text?  Or is it only useful if it can also be a video
               clip or HTML or whatnot?]


Tag:           status
Context:       presence
Description:   The text of this element specifies the current status
               of the presentity.  It must be one of "available",
               "busy", and "idle".  [XXX Should we be more precise
               about what those values mean, or is it good enough
               just to make sure programs use one of those three
               words?  Should we allow for more values in the future,
               or is it better for interoperability not to make this
               particular field extensible?]


Tag:           contact
Context:       presence
Sub-elements:  address capabilities preference
Constraints:   address must appear exactly once.  capabilities and
               preference must appear at most once.
Description:   This tag introduces a means of communicating with the
               principal.  Any text in the element is discarded.
```

There may be multiple contact elements within the
                    presence document.

Tag: address
Context: contact
Description: The text of this element gives the communications
                address as a URL [RFC 1738]. The URL type must
                correspond to a communication means and not a document
                type. [XXX How can we be more precise about this
                distinction? Obviously we don't want HTTP URLs here
                to be considered valid.]

Tag: capabilities
Context: contact
Description: The text of this element specifies the media features
                which can be processed by a means of communication,
                using the filter syntax defined in [RFC 2533].
                [XXX RFC 2533 filters are probably not all we need.
                More delving into the CONNEG framework is required.]

Tag: preference
Context: contact
Description: The text of this element is an unsigned integer giving
                the preference of a contact relative to other
                contacts. When selecting between contact addresses to
                use to contact a principal, addresses with lower
                priorities should be considered more desirable than
                addresses with higher priorities. If no preference
                element appears in a contact address, it should be
                considered less desirable than any contact address
                with a preference element.

7. Examples

The following presence document might be given as presence information
for a presentity which might be identified as joe@example.com. Note
that clarifying whitespace in the presence document must be used with
some care; it is fine to have extra whitespace directly within a
"presence" or "contact" element where it will be ignored, but it
should not be included in elements such as "address" in which text is
significant and extra whitespace is not specifically allowed.

```
        <presence>
        <presentity>joe@example.com</presentity>
        <date>2000-02-11 17:34:12</date>
        <status>idle</status>
        <location>Out to lunch at Mel's Diner</location>
        <contact>
                <address>im:joe@example.com</address>
                <capabilities>
                        (&amp; (pix-x&lt;=1024) (pix-y&lt;=768)
```

```
                              (color&lt;=256))
               </capabilities>
               <preference>1</preference>
        </contact>
        <contact>
               <address>mailto:joe@example.com</address>
        </contact>
        </presence>
```

8. Extensions

New element tags can only be standardized in the form of a
standards-track RFC. Element names beginning with "x-" may be used
for experimental purposes for all three kinds of extensions. New
element names should avoid the use of the ":" character, since it may
be used in the future for XML namespaces.

9. Security considerations

Watcher user agents should be careful to present communications
addresses to users when users choose to send a message to a
principal, so that users cannot be easily fooled into sending
authenticated messages to their work supervisors other unintended
parties.

10. IANA considerations

The current extensions proposal does not place any load on the IANA.

11. References

[Model]
M. Day, J. Rosenberg, H. Sugano. "A Model for Presence." Work in
progress, draft-ietf-impp-model-03.txt.

[Reqts]
M. Day, S. Aggarwal, G. Mohr, J. Vincent. "Instant Message / Presence
Protocol Requirements." Work in progress,
draft-ietf-impp-reqts-03.txt.

[Type-feature]
G. Klyne. "MIME content types in media feature expressions." Work in
progress, draft-ietf-conneg-feature-type-01.txt.

[RFC 1738]
T. Berners-Lee, L. Masinter, M. McCahill. "Uniform Resource Locators
(URL)." RFC 1738, December 1994.

[RFC 2045-2049]
N. Freed, N. Borenstein. "Multipurpose Internet Mail Extensions
(MIME)." RFC 2045-2049, November 1996.

[RFC 2119]
S. **Bradner. "Key Words for Use in RFCs to Indicate Requirement**
Levels." RFC 2119, March 1997.

[RFC 2234]
D. **Crocker, Ed., P. Overell. "Augmented BNF for Syntax**
Specifications: ABNF." RFC 2234, November 1997.

[RFC 2533]
G. **Klyne. "A Syntax for Describing Media Feature Sets." RFC 2533**,
March 1999.

[Unicode]
ISO (International Organization for Standardization). "ISO/IEC
10646-1993 (E). Information technology -- Universal Multiple-Octet
Coded Character Set (UCS) -- Part 1: Architecture and Basic
Multilingual Plane." [Geneva]: International Organization for
Standardization, 1993 (plus amendments AM 1 through AM 7).

[XML]
T. **Bray, J. Paoli, C. M. Sperberg-McQueen. "Extensible Markup**
Language (XML) 1.0." W3C Recommendation REC-xml-19980210, February
1998.