Individual Contribution
Internet-Draft
Expires: April 28, 2003

E. Brunner-Williams
Wampumpeag

October 28, 2002

EPP Data Considerations

<draft-brunner-epp-data-considerations-00.txt>

Status of this Memo

Abstract

This memo discusses data collection considerations and EPP. It is far from complete, but deadlines press.

Distribution of this document is unlimited.

Dedication

This draft is dedicated to Paul Wellstone, Senator from Minisota.

Table of Contents

1. Introduction 3
2. Background 4
3. The PROVREG WG 5
4. The Problem (or IESG Considerations) 7
5. The EPP Data Collection Element <dcp> 9
6. W3C P3P Overview 12
7. Discussion 13
References 15
Author's Address 15
A. Acknowledgements 16

1. Introduction

After Working Group Last Call (WGLC) on the set of memos that form the core specification [EPP-P],[EPP-C],[EPP-D],[EPP-H] of a provisioning protocol for domain names meeting the requirements contained in [RFC3XXX], an issue relating to the access model for data was raised.

Provisioning protocols for domain names [RFC2832],[SRS],[EPP-P] minimally transport data necessary for name servers implementing [RFC1034][RFC1035] et seq., aka "DNS", to provide mappings between some addresse(s) and some name(s). This is generally referred to as "publication" via a zone file. The usage of the PROVREG Working Group is to refer to this data as "technical data". The prevailing policy at the time [RFC881] was written, which defined the modern form of namespaces for doman name (see [RFC606],[RFC608]), required additional data to be provisioned. This policy was originally articulated in [RFC742], and subsequently revised in [RFC812], and ultmately [RFC954]. The usage of the PROVREG Working Group is to refer to this data as "social data". The primary distinguishing characteristic between "technical" and "social" data is that absence of or error in "social" data has no effect on the DNS.

[Discussion of Thick vs Thin.]

Fee-for-service operator of domain name operators, "pay-registries" hereafter, and fee-for-service regstrars, "pay-registries" hereafter, generally require sufficient additional data to support payment mechanisms. One policy for the management of domain name spaces imposes an expiry property on domain names. This policy is used by fee-for-service operators for service subscription, and a common form of this is the annual renewal model. For historical reasons the prevailing practice by fee-for-service registry and registrar operators is to require [RFC954] "social data", in addition to the data sufficient to support payment mechanisms, and to publish the "social data", following the practices set down in [RFC954].

2. Background

Several memos have appeared that attempt to reduce the impact of trafficing in network endpoint identfiers, e.g.,

1. Anti-Spam Recommendations for SMTP MTAs [RFC2505],

2. DON'T SPEW A Set of Guidelines for Mass Unsolicited Mailings and Postings (spam*) [RFC2635],

3. How to Advertise Responsibly Using E-Mail and Newsgroups or - how NOT to $$$$$ MAKE ENEMIES FAST! $$$$$ [RFC3098],

Additionally, several memos have appeared that directly or incidently reduce the accessibility of network endpoint identifiers to trafficers in identfiers, e.g.,

1. The IP Network Address Translator (NAT) [RFC1631]

2. Network Address Translator (NAT)-Friendly Application Design Guidelines [RFC3235

None has been published which obsoletes or updates NICNAME/WHOIS [RFC954], which mandates whois operators to acquire, and publish without any access controls, the names, postal, telephonic, and internet endpoint identfiers, of domain name infrastructure providors (aka "registrants").

3. The PROVREG WG

Some contributors to PROVREG are EPP implementors, and some are registry (or registrar) operators, or both. The operational practices of each operator may be limited to a single controlling authority, however, operators may operate multiple instances, each with a distinct operational practice. Implementors may seek to support "lazy evaluation" of operational practices.

In particular, collection practices of registries, registrars are not assumed to be uniform. This lead to the the requirement:

8.4 Data Collection Requirements [1], GRRQ

> The protocol MUST provide services to identify data collection policies.

Note that the requirement is for collector policies, not originator preferences.

The PROVREG WG considered the mechanism of the W3C's P3P activity, and adopted the following basic elements:

1. the mechanism would allow description of data collection practice, necessarily by the data collector
2. portions of the P3P XML Schema, specifically:
    1. <access> element
    2. <statement> element containing:
        1. <purpose> element
        2. <recipient> element
        3. <retention> element
        4. <expiry> element (optional)
3. [anything else I think of]
4. the optional character of P3P.

These allow for the registry data-collector to announce to the registrar the registry's data collection practices, and for registrar

data to be provisioned to a registry under a specific policy.

4. The Problem (or IESG Considerations)

The problem or issue raised upon IESG review of the work product of the PROVREG contributors is two-fold, some IESG commentors sought finer granularity for data protection, some to add originator preference.

The PROVREG contributors considered these possibilities in mid-2001. A semantic could be attached to an individual XML element, or to an EPP object, or to an operation on an object, or to a group of EPP commands and responses, aka, an EPP "session". The semantic could originate from a registry, or from a regstrar, or from a reseller, or from a registrant.

Because EPP is protocol between regstrar and registry peers, allowing resellers or users mechanisms to affect the state of an EPP transaction was eventually considered "out-of-scope".

Because the initial focus of the PROVREG contributors was on a connection-oriented, session-maintaining transport (TCP), and operational practice of PROVREG contributing registry operators and registrars did not suggest a use case for multiple data collection practices within a single namespace, reflected within the protocol as a single logical instance of an EPP server, the "EPP session" was associated with the <dcp> element.

The proposal from the IESG to the Working Group was to add a new attribute to every object element in the domain, contact, and host mappings. The new attribute would be a Boolean type, named "private", and will be used to note that the value of an element SHOULD NOT (emphasis added) be disclosed to third parties.

This proposal would change the EPP schema, which of necessity, is a change to the protocol, as the protocol is defined in XML Schema. It is not an extension, hence optional to implement. Ironically, it is optional to evaluate. Conformant EPP+XML parsers could ignore the value of the attribute, or simply treat as white-space anything following the closure of each element in the Working Group Schema.

The proposal additionally held that the default value should be "false", meaning that the element should be disclosed to third parties.

This proposal would be consistent with the "opt-out" legal regime of the United States, but inconsistent with the "opt-in" regime of the OECD States, and inconflict with the "opt-in" regime of the EU States.

Examples of the IESG proposal:

```
<contact:email private="true">jdoe@example.com</contact:email>

<domain:name private="false">example.com</domain:name>

<host:addr ip="v4" private="true">192.1.2.3</host:addr>

<extension>
  <noWhois><contact:email></noWhois>
</extension>
```

Note: This extension example isn't syntactically valid, it is shown for expository purposes only.

5. The EPP Data Collection Element <dcp>

From the -07 xsd.

```
<!--
A greeting is sent by a server in response to a client connection
or <hello>.
-->
  <complexType name="greetingType">
    <sequence>
      ...
      <element name="dcp" type="epp:dcpType" minOccurs="0"/>
      ...
    </sequence>
  </complexType>

<!--
Data Collection Policy types.
-->
  <complexType name="dcpType">
    <sequence>
      <element name="access" type="epp:dcpAccessType"/>
      <element name="statement" type="epp:dcpStatementType"
       maxOccurs="unbounded"/>
      <element name="expiry" type="epp:dcpExpiryType"
       minOccurs="0"/>
    </sequence>
  </complexType>

  <complexType name="dcpAccessType">
    <choice>
      <element name="all"/>
      <element name="none"/>
      <element name="null"/>
      <element name="other"/>
      <element name="personal"/>
      <element name="personalAndOther"/>
    </choice>
  </complexType>

  <complexType name="dcpStatementType">
    <sequence>
      <element name="purpose" type="epp:dcpPurposeType"/>
      <element name="recipient" type="epp:dcpRecipientType"/>
      <element name="retention" type="epp:dcpRetentionType"/>
    </sequence>
```

```
    </complexType>

    <complexType name="dcpPurposeType">
      <sequence>
        <element name="admin"
         minOccurs="0"/>
        <element name="contact"
         minOccurs="0"/>
        <element name="other"
         minOccurs="0"/>
        <element name="prov"
         minOccurs="0"/>
      </sequence>
    </complexType>

    <complexType name="dcpRecipientType">
      <sequence>
        <element name="other"
         minOccurs="0"/>
        <element name="ours" type="epp:dcpOursType"
         minOccurs="0" maxOccurs="unbounded"/>
        <element name="public"
         minOccurs="0"/>
        <element name="same"
         minOccurs="0"/>
        <element name="unrelated"
         minOccurs="0"/>
      </sequence>
    </complexType>

    <complexType name="dcpOursType">
      <sequence>
        <element name="recDesc" type="epp:dcpRecDescType"
         minOccurs="0"/>
      </sequence>
    </complexType>

    <simpleType name="dcpRecDescType">
      <restriction base="token">
        <minLength value="1"/>
        <maxLength value="255"/>
      </restriction>
    </simpleType>

    <complexType name="dcpRetentionType">
      <choice>
        <element name="business"/>
        <element name="indefinite"/>
```

```
        <element name="legal"/>
        <element name="none"/>
        <element name="stated"/>
      </choice>
    </complexType>

    <complexType name="dcpExpiryType">
      <choice>
        <element name="absolute" type="dateTime"/>
        <element name="relative" type="duration"/>
      </choice>
    </complexType>
```

6. W3C P3P Overview

This section left mostly blank for the time being.

Some intro blurb for those lacking clue.

P3P requires that policy reference files MUST be encoded using UTF-8, (2.3.2), and all policies MUST be encoded using UTF-8 (3.2).

The exception to this is the p3p-link-tag, which has the same character encoding will be the same as that of the HTML document it is embedded in.

This is not required by XML, see draft-brunner-epp-smtp-00 for details.

7. Discussion

There are 5 areas of significant difference between the PROVREG and IESG approaches:

1. optional element vs pervasive attribute

2. command (or "session") scope vs element scope

3. extensible vocabulary vs binary toggle

4. collector policy vs user preference

5. undisclosed policy default vs opt-out preference default

What follows is a brief discussion of these.

The impact of the IESG's proposal is that every implementor would have to implement the attribute. While evaluation is "MAY", discussed later, the schema-change is "MUST". Since implemention of the <dcp> element is optional, and may be absent from a server's <greeting> response, it is possible that the IESG's proposal would have the effect of making the PROVREG approach moot.

The IESG's proposal is that third parties are not recipients of data, ignoring the difference between "preference" and (IESG) and "policy" (PROVREG). There is no benefit from narrowing the scope to a leaf elements, except subdivision "social" data into discretionary categories with combinatorial scaling issues. Indicating that third parties are not recipients of data is trivially accomplished with the <dcp>.

The IESG's proposal is restricted to a binary value set, or possibly an enumerated set. If enumerated, changes to it would require a protocol version identifier change, as it is a proposed pervasive property of the EPP schema. The PROVREG approach is extensible, and uses a well-known vocabulary covering registrant access to the registry-resident data, recipients of the registrant data, repurposing of registrant data, and retention of registrant data, with problem-domain specific modification.

The IESG's proposal is to express the third-party disclosure preference of registrants. EPP clients MUST emit "privacy" attributes on each <element>, which EPP servers MAY ignore. The PROVREG approach allows EPP clients to REQUIRE EPP servers to emit a <dcp> element, and evalute the policy, before sending data to the EPP server. The net of the IESG proposal is that registries may accept registrant social data with this preference, and may then ignore it,

except for faithfully returning the attribue's value in <info> responses, but treating the data as unencumbered. The user's PREFERENCE is not binding. The PROVREG approach is the collector states its policy, allowing registrant selection of operators based upon policy.

The IESG's proposal is that the registrant must opt-out of third-party disclosure. The PROVREG approach is that the registrars data collection policy is unknown unless stated.

One issue remains in the author's mind as a sub-text in the IESG's proposal that is worth pursuing. Charitably restated, the IESG proposes element-wise access policy for the EPP schema. Mechanism left to the imagination. This can be accomplished by element-wise encryption, which has the disadvantage of adding key management to the problem doamin, and the advantages of allowing a richer access policy. The W3C XML Encryption Activity has published in this area.

References

[1] Bradner, S., "The Internet Standards Process -- Revision 3", RFC 2026, BCP 9, October 1996.

[2] Hollenbeck, S., "Extensible Provisioning Protocol", , August 2002.

[3] Hollenbeck, S., "Extensible Provisioning Protocol Contact Mapping", , August 2002.

[4] Hollenbeck, S., "Extensible Provisioning Protocol Domain Name Mapping", , August 2002.

[5] Hollenbeck, S., "Extensible Provisioning Protocol Host Mapping", , August 2002.

Author's Address

Eric Brunner-Williams
Wampumpeag, LLC
1415 Forest Avenue
Portland, Maine 04103
United States

EMail: brunner@nic-naa.net

Appendix A. Acknowledgements

The author gratefully acknowledges the contributions over the years of the Chair, Staff, Member Company contributing representitives, and Invited Experts of the W3C's P3P Specification Working Group. Many thanks go to the contributors to the IETF EPP drafts.

In addition the author thanks Marshall Rose who provided the extremely useful [RFC2629] document type description and xml2rfc tool used to edit this specification. The author may yet learn how to use this tool.