STRATEGY PROPOSAL

gTLD ZONE FILE ACCESS FOR THE FUTURE

STATUS OF THIS DOCUMENT

This is the Strategy Proposal (the "Proposal") as developed by the community members participating in the Zone File Access Advisory Group (the "Group"). The Proposal presents the Group's recommendation of a scalable, uniform implementation model for zone file access in new gTLDs and encourages existing TLD operators to adopt the model. The implementation recommended by the Group incorporates elements of two models proposed in the original ZFA Concept Paper (the "Paper") and thus referred to as the Hybrid Model (the "Model"), will be incorporated to the next draft of the new gTLD Applicant Guidebook for discussion at the ICANN meeting in Brussels, Belgium.

Zone File Access - Strategy Proposal
Prepared by: Zone File Access Advisory Group
12 May 2010

Table of Contents

1.0 SUMMARY

The Group was formed in December 2009 to discuss benefits and explore methods to effectively and efficiently enhance and make scalable access to zone file information in an environment with many gTLDs. Enhanced access to zone file information was identified in Applicant Guidebook v3 as a means to mitigate potential malicious conduct in new gTLDs. As reported in the Group's Paper posted on 22 February 2010, several stakeholder groups including anti-abuse and trademark protection organizations have described access to zone data as an effective and necessary tool for combating Domain Name System (DNS) abuse.

The Group's Paper explored and presented four potential solutions that included (i) an enhanced bi-lateral model where the essential elements of relationships between registries and consumers become standardized; (ii) a repository model where a third-party collects zone files from registries and distributes them to zone file consumers; (iii) a proxy model where a third-party acts as an intermediary for standardized authorization and where data is delivered via secure proxy connections to the registry; and, (iv) a clearinghouse model where credentials and authorization are maintained by an intermediary but where data is delivered over secure connections between the consumer and the registry.

The Group convened weekly calls that commencing in December 2009 and on 29 April 2010 reached consensus that the most effective and efficient means to enhance the current zone file access system is via the proposed Model – a combination of the enhanced bi-lateral and clearinghouse models described in this Proposal. The Model offers a single point of contact for applicants seeking zone file access and largely preserves existing roles and operational functions of data providers. The Model introduces two changes to the current zone file access system. First, it standardizes the relationships between zone file data providers (i.e., registry operators) and consumers (e.g., anti-abuse and trademark protection organizations, researchers, academia, etc.) in three main categories: application standards, access standards, and file/record format standards. Second, it introduces a lightweight clearinghouse for identity management in the zone file access system that is intended to provide a single point of contact for consumers who seek zone file access. These modifications are presented in detail in Sections 5.1 and 5.2 in this Proposal.

With the submission of this Proposal, the Group's work is complete. What is to follow is the development of an implementation plan, including incorporation of the Model in to the Registry Agreement for new gTLDs. The implementation plan will be developed by ICANN Staff with input via public comment from the global, stakeholder community.

2.0 WHAT IS THE PROBLEM TO BE SOLVED?

The current model for Zone File Access is built on individual, bi-lateral agreements and operational relationships between consumers and gTLD registry operators (providers of gTLD zone data). Scaling this model into an environment where there are many consumers and providers may create

both operational and cost inefficiencies for both consumers and providers of gTLD zone data. A new model for Zone File Access that can scale in both the current environment, and in an environment including new gTLDs and larger numbers of consumers, is required to address these concerns.

3.0 WHAT IS THE SOLUTION?

The solution proposed by the Group is a combination of two models proposed in the ZFA Concept Paper:

1. standardization of the essential relationships between consumers of gTLD Zone File data and the providers of that data; and,
2. a lightweight clearinghouse for contract processing and identity management in the ZFA system.

Throughout this paper this combination of models will be referred to as the Model because it is the result of simplifying and refining a pair of approaches first presented in the Paper, to reduce business and operational overheads while leaving TLD zone data under the control of the registry operators. The Model emerged from the Group as it deliberated a set of four individual approaches to solving the ZFA problem. During the course of deliberations, the Group considered how the approaches might fit together to provide the greatest value to both providers and consumers. The consensus reached by the Group is that the combination of standardization and clearinghouse approaches best meets the needs of the majority of zone file consumers and providers.

3.1 WHAT IS MEANT BY STANDARDIZATION?

Section 5.1 of this paper presents details of the standardization that all new gTLD registries will be required to comply with per their Registry Agreement. Essentially, standardization can be seen as:

> *changing from a system where there is a large variety of formats for the zone file data and large number of transport options to a system where the variety in both cases is reduced -- in fact, the optimum circumstance is a single, "standard" format and single access method for zone file data.*

Standardization refers not just to the format of the zone file data itself, but also to the operational mechanisms and technology for accessing that data.

3.2 WHAT IS A LIGHTWEIGHT CLEARINGHOUSE?

In the current model for gTLD zone file access, each consumer of data engages in a separate relationship with each registry where they have an interest in zone file data. This means that, for those people and organizations that have an interest in many gTLD zone files, they are managing many relationships with many providers of data. As the number of registries increases, the complexity and cost of managing those relationships increases.

The lightweight clearinghouse approach attempts to address this scaling problem by providing a single point of contact for administrative coordination of legal and technical identity information. Essentially, the clearinghouse can be seen as:

> *a lightweight management function that, 1) coordinates a standardized application process for gTLD zone file access on behalf of consumers and gTLD registry operators; 2) generates and manages credentials for consumers that gTLD operators have approved for access; and 3) distributes consumer credentials to approving gTLD registry operators.*

The relationship between consumers and providers of the gTLD zone data is a direct relationship. The zone file access agreement is a legal arrangement between the consumer and provider. The clearinghouse provides consumers a single contact point for zone file access agreement processing, credential issuance and customer care/support. The clearinghouse is not a party to the legal agreement between registry and consumer, and it is not involved in the transfer of gTLD zone data (there is, for instance, no proxy or repository service between the consumer and the provider as certain models in the ZFA Paper suggested). The implementation of the Model would not prohibit a provider (i.e., registry) and consumer from executing a separate agreement with a modified Zone File Access Agreement to for example permit the aggregation and distribution of zone file data.

Section 5.3 provides significant technical detail for the design of a clearinghouse for identity management in ZFA relationships.

3.3 HOW DOES THIS SOLVE THE PROBLEM?

The combination of standardization and clearinghouse function are not the only possible solutions to the problem outlined in section 2.0 of this paper. However, the Group determined that the Model solution offered the greatest benefit to zone file consumers and providers while also addressing the issues of scaling and multiple registry access.

First, the commitment to standardization in both access methods and zone file data formats means that the addition of new registries will not add to the number of access methods and data formats that have to be managed at the consumer end. For professional data consumers, the addition of new registries as data suppliers will involve invocation of well documented and easily programmable procedures. This means that consumers will not have to build special purpose code or processes in order to access new gTLD registry zone data. This addresses one of the key features of the scaling problem: for consumers, the addition of new gTLD zone files meant the possibility of coping with diverse access and data formats and standardization significantly limits this problem.

In addition, standardization provides a significant benefit to providers of the data: no longer will they have to create custom (per registry) processes for provisioning the zone file access scheme. Instead, a set of standard access procedures, naming conventions, and data formats will already be in place -- and possibly, publicly available code and technology to support those procedures.

Second, the clearinghouse component of the Model ensures that zone file access can scale as interested parties apply for large numbers of new gTLD zone files. Individual consumers will not have to hunt for the appropriate contacts and paperwork for each registry in which they have an interest; instead, the consumer will only have to access a single contact. That contact -- the clearinghouse -- is the provider for administrative, informational, and access details for all the zone files for which the consumer has an interest.

While standardization provides an effective increase in scalability in access to the data, the clearinghouse model provides increased scalability in access to the service. This two-pronged approach -- improving access to the administrative and technical components of zone file access -- is fundamental to the Model. Standardization was recognized by the Group as particularly necessary to address scaling issues no matter what model from the Group's Paper was eventually chosen.

This Hybrid approach makes use of the Clearinghouse Model described in the Paper by having a third party facilitate the necessary overhead of contract and user management. It enhances the existing bi-lateral model by identifying and adopting conventions that simplify the configuration and zone data processing overhead. It leaves enhancement opportunities as green fields for individual registries, or as enhancements that could be incorporated as a convention along with the present set being studied by ZFA standardization subgroup.

3.4 DOES THIS APPROACH APPLY TO EXISTING CUSTOMER/PROVIDER ZFA RELATIONSHIPS?

The Group believes that the Model should apply to all new gTLD registries. The Group also encourages existing gTLDs to adopt the Model as soon as is practical for the supporting registry.

4.0 HYBRID ZFA MODEL - ROLES

4.1 The ZFA Consumer

The ZFA Consumer is defined as the person or organization who applies for zone access, agrees to a standardized zone file access agreement and uses standard protocols to retrieve the zone files. This definition does not differentiate between those consumers who use the files for research, commercial activities or other applications approved by gTLD registries. In the Model, all consumers of the zone files are treated equally.

4.2 The ZFA Data Provider

The ZFA Data Provider is the TLD registry that participates in the ZFA system. While this program is being defined for gTLDs, we note that ZFA Data Providers do not necessarily need to be limited to gTLDs. In the case where zone files are managed by a third-party on behalf of the registry, the ZFA Data Provider is still the registry responsible for the TLD. This definition of does not differentiate between ZFA Data Providers of different sizes or between ZFA Data Providers who have different organizational models for their TLDs.

4.3 The ZFA Clearinghouse

The ZFA Clearinghouse is a third-party, independent entity that provides a single point of contact for information and services for the Hybrid ZFA system. It provides consumers a single point of contact for zone file access, an application submission and processing service, and relays necessary notifications between the ZFA Consumer and ZFA Data Providers (e.g., requests for clarification or additional information) as required by specific implementation. In cases where gTLD registries approve the ZFA Consumer for zone file access, the ZFA Clearinghouse manages the ZFA consumer's identities and credential-related materials (including issuance, management and revocation, and customer care related to contracts and credentialing). The ZFA Clearinghouse is specifically limited to these functions.

5.0 WHAT ARE THE KEY FEATURES OF THE HYBRID ZFA MODEL?

The Model is effectively the combination of:

- standardization of ***access methodology supported by ZFA data providers***;
- standardization of ***data formats supported by ZFA data providers***; and
- simplification of the administration for ZFA and identity management, *supported by the ZFA Clearinghouse*.

5.1 STANDARDIZATION

The technical recommendations on standardization for Zone File Access fall into three main categories:

- **Application Standards** is the process for consumers to apply for access;
- **Access Standards** which includes methods of access, directory and file structures, compression, and publishing frequency; and,
- **File and Record Format Standards** recommendations to help normalize and standardize formats for providers and consumers.

Application standards are part of the clearinghouse approach and are detailed in section 3.3 of this document. The Group expects the application standards to be very similar to the requirements commonly imposed on ZFA consumers by gTLD registry operators today.

5.1.1 Access Standards -- Method of Access

Access credentials are consistent for each consumer. The clearinghouse is responsible for generating and managing a single credential assigned to each user for all TLDs. The credentials may take the form of a triple composed of user, password and IP address, or simply a user/password (if the provider does not limit by the IP address or subnet(s) of the consumer). User and password are defined as ASCII strings. IP address is defined as either an IPv4 dotted quad format address, and IPv6 canonical format address, or the null string. Passwords generated by the clearinghouse are to be strong and random, as they are under the current zone file access arrangement.

5.1.2 Access Standards -- Access Protocol

There is a single access protocol supported by the Model for Zone File Access. The standard access methodology is FTP and the access scenario is simply for the consumer to connect to the registry's FTP server, from the IP identified in the credentials supplied by the clearinghouse, and log in using the credentials from the clearinghouse. FTP is specified as a baseline technology because it is widely available, easily implementable, and well-understood. However, the Group recommends that the implementation team should investigate and deploy more modern data transport and access protocols for distribution of the Zone Files (see Section 6.2).

5.1.3 Access Standards -- Server Naming

For every zone file access server, an alias is provided by ICANN so that finding the proper server is easy. ICANN will provide aliases in the DNS as <tld>.zfa.icann.org, analogous to the <tld>.whois-servers.net. All ftp servers will be named in the ICANN.org namespace -- this is an extension to the convention already applied for whois servers. ICANN will implement this with a CNAME resource record to ensure that registries have maximum flexibility to manage their services while ensuring routine and stabile access for Zone File Consumers.

5.1.4 Access Standards -- Paths to Zone Files

For every zone file access server, the zone files can be found in a consistent directory at the server. For the Model, the zone files are in the top level directory called <tld>.zone.gz, with <tld>.zone.gz.md5 and <tld>.zone.gz.sig to verify downloads. If the registry also provides historical data, it uses the naming pattern <tld>-yyyymmdd.zone.gz, etc.

Note that the undated file (e.g. <tld>.zone.gz) must still exist as an alias for the current version of the zone file.

5.1.5 Access Standards -- Compression

The Model for zone file access incorporates a single mode of mandatory compression. While the Group noted that bzip2 is slightly more efficient than gzip, gzip is in wider use and is well supported. gzip is the mandatory compression mechanism for zone files.

5.1.6 Access Standards -- Frequency and Timing of Updates

Each registry must establish a frequency of updating its zone file for publication using the Model. The frequency must be often enough to ensure that the delta between current and previous versions is reasonable. The registry establishes the frequency and documents this in a README file available on its FTP server.

Each registry will also establish the time when the updated zone file will become available in the update window. The time must be predictable and reliable. The registry establishes the timing for the publication of updates and communicates this to the consumer through the clearinghouse.

5.1.7 File and Format Standards

Zone files are to be provided as BIND-compatible zone master files. New registries must use the following format for zone files:

1. Each record must include all fields in one line as: <domain-name> <TTL> <class> <type> <RDATA>
2. Class and Type must use the standard mnemonics and must be in upper case.
3. TTL must be present as a decimal integer.
4. Use of /X and /DDD inside domain names is allowed.
5. All domain names must be in upper case.
6. Must use exactly one tab as separator of fields inside a record.
7. All domain names must be fully qualified.
8. No $ORIGIN directives.
9. No use of "@" to denote current origin.
10. No use of "blank domain names" at the beginning of a record to continue the use of the domain name in the previous record.
11. No $INCLUDE directives.
12. No $TTL directives.
13. No use of parentheses, e.g., to continue the list of fields in a record across a line boundary.
14. No use of comments.
15. No blank lines.
16. Exactly one SOA record should be present at the top of the zone.
17. With the exception of the SOA record, all the records in a file must be in alphabetical order.
18. One zone per file. If a TLD divides its DNS data into multiple zones, each goes into a separate file named as above, with all the files combined using tar into a file called <tld>.zone.tar. (There's no point in compressing a tar file of already compressed files.)

The following practices are specifically not permitted:

1. The use of .NAME is in mixed case with many $ORIGIN lines.
2. Defaulted blank names that mean to use the same name as the previous non-blank name.

The Group recognizes that three major formats are already in use by registries. There should be no further expansion of varieties of legacy zone file formats beyond the three listed here:

1. use relative names for names within the current TLD, absolute otherwise, omit TTL and class, fields separated by one space, records are in arbitrary order but consistent from one version of the zone file to the next;

2. formats all names as absolute names, NS records are alphabetized, other records follow the relevant NS records; or,

3. use relative names, explicit IN fields, put A records first in alphabetical order, then NS records in alphabetical order.

Just as in the case where the registry documents timings of its postings in a README file, the registry must also indicate which of the legacy file formats are in use in the README file.

5.2 THE ZFA CLEARINGHOUSE

The ZFA Clearinghouse is one part of the hybrid solution based on the Clearinghouse and Enhanced Bi-lateral Models in the ZFA Paper.

The purpose of the ZFA clearinghouse is to reduce contractual overhead and to manage the issuance, recovery or replacement of credentials. To make the ZFA clearinghouse approach work, three well-defined sets of participants must exist:

1. the ZFA consumer;
2. a ZFA consumer registration authority (the ZFA Clearinghouse), credential issuing agent and credential distribution agent which is trusted by all participating registries and consumers, and
3. a ZFA Data Provider, which obtains consumer authenticating material from the registration agent.

The ZFA Data Provider is operated by each registry whose zone file transfer services are enhanced by agreements and conventions (i.e., standardization in section 5.1 above) that reduce the configuration and complexity for consumers.

5.2.1 Clearinghouse Role -- Part 1: Application Processing Services

The Clearinghouse provides a single, clearly defined URL to access ZFA request documents, so that a potential consumer can easily locate and access the necessary legal agreements. In the actions below, a standard contract is envisioned for ZFA. In particular, it is expected that there will be a consistent content and structure to ZFA agreements, so that users can apply for access to several or all gTLD registry zone files but would need to review and sign a single contract.

The role of the ZFA Clearinghouse would include the following actions:

- Process ZFA applications from a consumer, by providing them with a standard contract and some means of indicating which zones the consumer wishes to access;
- Forward the ZFA application to individual registries for approval; and,
- Assist consumer or registry in reconciling processing difficulties (i.e., should a registry require a clarification from consumer or additional contact information)

All parties who agree to abide by a standard agreement are eligible for ZFA. Today, registries take measures to verify the contact information an applicant submits. These measures include confirmation that the applicant can be reached at the postal and email address, telephone and fax numbers

the applicant submits. Existing process of entering into ZFA agreements is normally paper-based. It is worth studying whether with introduction of the Clearinghouse an electronic process can be implemented.

5.2.2 Clearinghouse Role -- Part 2: Approved Applicants

- For approved applicants,
  - Create credentials in the form of a triple (consumer identity, consumer password, additional factor which is currently an IP address for certain registries) [Note: this would be the IP address of the consumer's ZFA client machine, provided by the consumer in the application process, and if the registry is not interested in using this factor, the value could be null.] See Section 5.1.1 above.
  - Propagate credentials to the individual registries that have approved the consumer for ZFA. This could be a single file that is periodically uploaded to or downloaded by individual registries using authenticated FTP. Changes to this file are likely to be infrequent, so a daily update may be sufficient.
  - Provide credentials and instructions to consumers on how to access zone data from selected registries.
  - Provide configuration information for server access for each registry.
  - Manage identity and credential related problems (password recovery/reset, change of IP), synchronize authentication problem resolution outcomes with the individual registries that have approved the consumer for ZFA. This is an outcome of updating the approved consumer's authentication data file, which is either downloaded to or uploaded by the registries.
  - Serve as a single point of contact for customer care, problem resolution, and notification of revocation. The registry determines if the consumer has violated terms of service, and would coordinate its suspension of service with the ZFA registration agent.
  - Manage the configuration information required for ZFA registration FTP service with participating registries. In the "push" variant, the ZFA registration agent must have account information for each registry whereas in the "pull" variant, the registries must have the account information for the ZFA registration agent.

5.3 ZFA ACCESS MECHANICS

Zone file access proceeds rather similarly to the existing process. A consumer's client machine connects to a registry ZFA server. This is a gTLD registry operated host from which consumers will download zone files, see Section 5.1.1 for naming convention and access method. The registry ZFA server challenges the consumer's client for credentials. The consumer's client machine submits the credentials the consumer obtained from the ZFA registration agent. The registry ZFA server consults its local copy of the approved consumers authentication data file to see if the consumer has submitted the expected identity/password information (from the expected IP address). If the consumer's credentials

(and origin IP address) result in a positive match to an entry in the approved consumers authentication data file, the registry's ZFA server accepts the connection. The consumer's client machine then submits requests to download zone files (current or historical). See Sections 5.1.2 - 5.7 for details regarding naming and format conventions for zone files.

5.4 IMPLICATIONS FOR EXISTING/LEGACY TLDs AND THEIR CUSTOMERS

Except for the inclusion of the Clearinghouse, the Model is not significantly different from the existing approach for providing access to zone files. As a result, the Group proposes that existing registries adopt the model quickly. The intent is to eventually move all zone file access to the Model.

It is important to note that the Model does not eliminate the possibility that a registry might want to execute a bi-lateral agreement between the registry and a zone file customer for example to provide value-added services. While it is intended that all registries participate in the Model, nothing prevents them from additionally providing service on an individual basis.

Another essential feature of the Model is that registries retain the freedom to create and market advanced services for zone file access. Some registries already provide value-added services related to zone file access -- nothing in this proposal is intended to limit any registry's ability to provide existing advanced services or marketing new services.

Finally, the Model ensures that the access mechanism (see Section 5.3 above) is between the ZFA Data Provider and the ZFA Consumer. In the event that the ZFA Data Provider detects potential malicious conduct by the ZFA Consumer, the Data Provider can terminate provisioning of data immediately -- without the intervention of the Clearinghouse.

6.0 NEXT STEPS

6.1 INCLUSION IN THE APPLICANT GUIDEBOOK

The Group was formed to develop and recommend a method of enhanced zone file access for new gTLDs for inclusion in the Applicant Guidebook. The purpose of this strategy paper is to provide a source for Zone File Access details in the Applicant Guidebook. The timing of the drafting of the strategy paper is done with this goal in mind. The strategy paper is not intended to be the final document related to zone file access: for instance, detailed implementation planning needs to take place.

6.2 IMPLEMENTATION PLANNING

Once the strategy proposal has been used as a source for information for the Applicant Guidebook, the next step is to do detailed implementation planning.

The detailed implementation planning will include the development of precise flows of work between the ZFA Data Consumer, Clearinghouse, and ZFA Data Provider. In addition, it will provide a clear design for funding, staffing and implementation of the Clearinghouse role. Detailed implementation planning is expected to be completed by staff in ICANN.

The implementation planning will include an examination of the potential for third-party aggregation of zone file data and the creation of new by-products of the access to the zone files. For example, to allow for the possibility of a market-created aggregation and distribution service provider the Group proposes that the Zone File Access Agreement be amended to permit this function provided the consumer and registry execute a separate agreement for this stated activity.

Finally, there has been a discussion within the Group as to what entity could act as the Clearinghouse and whether ICANN can run or manage one. A reasonable rationale in making this choice would be minimization of costs. Zone files are currently available for consumers free of charge and it is important that they remain free under the Model. Ideally introduction of the additional service should not cause significant increase of the expenses for data providers.

APPENDIX A: NOTES ON A COST MODEL FOR ZONE FILE ACCESS

The ZFA **Cost Model** [XLS, 175 KB] is designed to thoroughly explore the potential costs of providing zone file access under four different approaches explored by the Group. The approaches are as follows: Protocol Standardization, Clearinghouse, Repository, and Proxy. The model accounts for various variable and fixed costs of service provisioning under the various models, and provides an examination of the resulting costs to individual registry operators, any "centralized" provider described in the model, and to the overall ecosystem of providers. The primary assumptions and results are combined on a single base worksheet, allowing the user to modify assumptions to look at their impacts. Additional worksheets present the calculations necessary to drive the overall analysis and may be examined and modified by the user if desired.

The assumptions are broken up into three primary areas: market sizing (market make-up and growth), cost structures (fixed and variable costs necessary to provide ZFA services as an entity), and zone file data characteristics (size of files and download frequencies). Changes to any of the assumptions will lead to an immediate update of the overall costs presented per model at the top of the worksheet. Note that an important assumption whether one is reviewing the current "known" universe of registries, or the projected model that will be with us in the future, and this assumption can be found under "market sizing". Users wishing to apply their own assumptions to how the market may grow, their own knowledge of cost structures, and alternative assumptions on zone file growth are encouraged to use this model to drive their own analysis of the various potential methods for providing zone file access the working group identified.

APPENDIX B: ZFA ADVISORY GROUP PARTICIPANTS

Mike O'Connor	O'Connor Company of St Paul
John Levine	Taughannock Networks
Nacho Amadoz	Fundacio puntCAT (.CAT)
Adam Palmer	Symantec
David Maher	Public Interest Registry (.ORG)
Vladimir Shadrunov	Telnic (.TEL)
Brian Cute	Afilias (.INFO)
Tom Barrett	EnCirca - ICANN-accredited Registrar
Paul Stahura	
Rick Wilhelm	Network Solutions - ICANN-accredited Registrar
Tatyana Khramtsova	RU-Center, Registrar
Rod Rasmussen	President/CTO, Internet Identity
Rod Dixon	Attorney
Joe St Sauver	Internet2 and the University of Oregon
Wang Wei	CNNIC - ccTLD Registry
John Kristoff	Research Analyst, Team Cymru
Jothan Frakes	Minds + Machines
Ken Stubbs	Afilias (.INFO)
Asif Kabani	International Sustainable Development – Resource Centre (ISD-RC)
Berry Cobb	Infinity Portals, LLC
Susan Prosser	DomainTools
John McCormac	Hosterstats.com
Rick Wesson	Support Intelligence, Inc.

ICANN

Francisco Arias
Mark McFadden
Dave Piscitello
Kurt Pritz
Craig Schwartz