

Craig:

Good afternoon everyone and welcome to the Zone File Access Advisory Group workshop. My name is Craig Schwartz and I'm one of the ICANN staff supporting this particular piece of work. We would like to welcome you to today's work.

I would like to run through a few administrative items before introducing the speakers on today's call. They will lead us into a discussion on what we'll be covering today. I do want to remind folks the call is being recorded and that there will be an MP3 available of the meeting up within the next 24 hours, posted to the ZFA information web page on ICANN's website.

I would to acknowledge the 2 advisory group co-chairs Rod Rasmussen and Vladimir (last name). Also to acknowledge the ICANN staff supporting the work. Dave Piscatello, Mark McFadden and myself who are on the call.

For the folks signed into the Adobe link I'll be monitoring that so if you do have a question and would like to speak up you can raise your hand in the Adobe room and we'll start a queue and take questions that way. And throughout the presentation today I encourage folks to ask questions as they arise rather than waiting until the end. We will have a dedicated Q & A session following the presentation but we do encourage questions throughout the talk today.

I would also like to remind folks there is currently a public comment forum running on the Zone File Access concept paper and that comment is posted and supposed to run until April 8th and I'll presume we're just going to keep it open the extra 7 days.

I would also like to go ahead and acknowledge the folks that are, the advisory group folks here in the room in Reston. We have Mark McFadden who is with ICANN, we have Susan Prossor who is an advisory group member, Jonathan Frakes who is an advisory group member, Ching Chow also a group member, Mikey O'Connor and then we have a whole bunch of observers sitting here.

So with that said I would like to turn the session over to Vladimir who is with you in Nairobi to kick off the workshop today.

Vladimir:

Thank you Craig. I see here Brian (last name – 5:19) present, Tatiana from (location) and I'm not sure if there are other advisory group members in the room.

Barry:

Barry Cobb from the BC.

Vladimir:

Hello Barry. Okay so this workshop we will talk about zone files. First of all, let's talk about what TLD zone file is because our initial reach to the various communities indicated that there might be some misunderstanding as to what constitutes a zone file or what it is and mostly importantly what it is not.

So the zone file is the collection of data that is provisioned to the DNS servers of a particular top level domain. It generally contains a list of all delegated domain names in a TLD and a corresponding authoritative name server for each domain.

Besides that it usually contains glue records and that is the IP addresses of the name servers in case they belong to the same TLD as the main domain. DNS SEC information is also usually included in the zone file. And depending on the operational model of a TLD there may be other resource records included in the zone file.

Now what the zone file is not, the zone file doesn't contain any information related to the domain registration itself, meaning creation date, expiration date, sponsoring Registrar, domain status in the shared Registry system. It does not also contain any contact information for Registrants. So whoever gets access to the zone file that doesn't necessarily mean they are able to contact the Registrant.

The history of zone file access, zone files were available since pre-ICANN times. At that time, there was only one Registrar operating solutions and zone files were available via FTP to the signatures of the terms of service. When ICANN came to life there was still a single GTLD operator and the zone file access agreement was part of the agreement between ICANN and the event Registry operator.

During 2001 and 2009, 13 new GTLD's were introduced and the zone file access agreement was copied into these new GTLD's. So the system kind of worked in an environment where there were just a few GTLD operators. Scalability was not a priority when the original system was designed and while the new GTLD's were being added the scalability issues became noticeable but still the community was able to cope with these issues and still tolerable.

At the moment, GTLD Registries are required by their contracts with ICANN to provide free of charge zone file access to any party that signs an agreement with the Registry. Access is provided at no charge and as I already mentioned the zone file access agreements are almost identical across all the GTLD registries.

Registries are required to post up to date information at least every 24 hours. The access methods vary but quite often access is provided through FTP, through user password authentication without encryption. Some providers do maintain access control lists meaning that consumers need to provide their IP address from which the requests originate. These addresses are in the firewall access list.

The access is provided free of charge and as a result no SLA's, service level agreements, are applicable usually to the Registries. On the screen you can see the overview of number of zone file access accounts and the recent changes in the recent months over various GTLD's. You can see that the number of people accessing the zone file data varies from dozens to hundreds.

We also have to know that country code TLD's are generally not under contract with ICANN and they do not make their zone files available to the public. With that said we think that if the work of this advisory group proceeds to the point where some improved zone file system may be introduced, we will be happy to accept country code TLD's into that system so that they can benefit from any improvements that may come up.

Then I would like to turn the mike to Rod Rasmussen, the co-chair of this advisory group who will speak about possible uses of the zone file access, Rod.

Rod:

Thank you Vladimir. This is Rod Rasmussen and thanks everybody for your participation. I would like to thank your group members, especially for all the work we've been doing. I think it shows here in what we put out, so hopefully you can hear me. I do not have control by the way Vladimir with the presentation so I'll let you know when I'm done with the slides.

One of the things we discussed as part of the working group is uses for the presentation itself, the uses of the zone files themselves by different consumers of the data and got into various reasons why people are trying to get that information to better explore what kind of needs there would be in looking at this question around this.

So we have 3 different major areas we divided it into. The first is intellectual property and trademark protection uses. As you might imagine, people are taking a look at new domains as they're entering the zones as a place for potential trademark violations. It is also a great way to research new trademark registrations you might want to make and make sure you're not infringing on someone else's established domain name. Obviously, you can use it for competitive analysis to what new terms might be being used out there.

You can also by taking and looking at it from a zone level be able to look for systemic behavior where broad numbers of domains are registered and perhaps you can tie them through main servers, you can take a look at large numbers of domains containing a mark and put together a better case for infringement around that.

So the second area here we were talking about and came across or eluded from the people who were participating in the group was use for e-crime both in preventing and investigating incidences. So one of the ways you do that as new or different domains come into the zone you can look for a particular targets that get phished or some malware attacks and things like that

either by brand or terms. So if you have something like BigBank-account.net that would be a suspicious term for example.

Often time's bad guys will grab a bunch of names ahead of time and then use them later on. So you can find some of these things prior to disbanding them. And you can see this when you're researching different network structures that they'll use often the same name servers. For example, so if you have a bunch of new domains tied to a name server you can investigate those right away.

Then also as you're looking across different Registries you can see common patterns of abuse there.

Another important use of this for people who like to, who run black lists, they want to be able to white list, in other words, instead of making sure something gets blocked they make sure something doesn't get blocked when it's a legitimate domain name.

The third area we went through is kind of a big bucket of other uses in businesses, etc. and this isn't even a comprehensive list. I think there are many ways people are using zone files beyond this. But there are certain people who are just doing research into the internet itself, the reach of it and you can take a look at that from the hosting level or various analyses of things that people are doing on the internet.

Clearly there is an area of the marketplace that is using it for evaluating potential value of investments into domain names, whether that's looking for domains who expire or the history, keeping track of the history domains or looking up labels that are registered in 2nd level names registered in one Registry and trying to see if they're available in another one. And there are applications being used for keywords, knowing what trends are in terms various geographic things where you might want to do ad targeting other types of things where you're looking at enhancing your navigation behavior things like that.

So that covers the usage areas we identified and put up here. Going on to talking about looking at potential issues as we expand the GTLD space with the new GTLD program, we first were taking a look at some of the current issues we've identified with the current program and how those would or would not continue and those include things like the security around access, a lack of any service levels for provision of the zone files and that there is, it's fairly sub optimal as far as the way the costs are distributed.

Then as we take a look at overall expansion in the zone files, we have it becomes pretty challenging if the different Registries have different processes for creating and executing an agreement to access it. We have different ways already and we can imagine lots of different ways for converting the access to get the files from each Registry. And customer support and all the things you have to do in order to take care of problems, change management and things like that. If everybody has a different system for doing it and different expectations for when things

get done. We've seen that a bit with the current system already and as you expand something like this to an order of magnitude beyond where we currently are it becomes very...you start putting together the numbers and it looks kind of scary.

So there is also some concerns about DNS SEC as far as pure size of the zone files themselves as that gets implemented.

So I have a couple of graphs here that represent the current system. The first one is from the Registry point of view, so if you're a Registry providing access to dozens or hundreds of different types of consumers and there are a few labeled here, you have this one too many relationship between all the consumers of data and that Registry.

The next slide I have shows from the consumer point of view but as a collective if you will, where each consumer has a few Registries that they consume data from right now. Those in today's world a single consumer has connections to anywhere from 1 to a dozen Registries. Then that's repeated over and over and over again, as a collective marketplace. Now if you can imagine and I'm going to back up a slide here for a second, in the future this would be what a consumers view would look like, if you have hundreds of Registries out there. So you can imagine that this next slide would have that same massive scale issue. There we have the marketplace where you have lots and lots of data connections. So visually you can start to see some of the challenges that this represents.

I'm going to pass it back to Vladimir for the next set of slides here.

Vladimir:

Thank you Rod. So it was acknowledged through the ICANN reach out to the constituency that there are certain issues with the current zone file access system. When the applicant, draft applicant guidebook version 2 was published and when there was a comment period for it after the Sydney meeting, this issue was raised by the community.

In version 3 of the guidebook that was published before the Seoul meeting. The DAG contained positions for supporting a centralized zone file access. We talked about this issue in Seoul and it was decided that further consultation with various stakeholders will be required and therefore the decision was made to convene an advisory group in order to analyze the issues. So the advisory group was convened in December 2009.

The tasks before us were as follows, we need to collect the information to see what is the current status, what are the current uses of the zone files, who are the producers, who are the consumers? What are the issues and if there are indeed any issues with the current zone files access system. We had to come up with one of several solutions on how to improve that system and how to optimize costs for consumers and producers of data.

We had to summarize our findings into a concept paper which is currently available in the ICANN's website, which is posted for public comment a couple of weeks ago. The public comment will last until April 8th.

So what are we heading to? Our goal is to develop an improved zone file access system, which provides for scalability which is secure and provides consistency between the various GTLD operators. We think there are opportunities to improve costs not only for the consumers of data but also for the providers of data. We think there are opportunities to make the system more resilient and more...you understand what I'm saying right?

The important aspect is that the new system must not infringe on the Registries abilities to innovate because we already see Registries providing various improved services in regards to the zone file access such as, for example, rapid, VeriSign provides some very rapid updates to the zone files and there are probably (26:35 – tape skips and goes out).

The advisory group was co-chaired by one representative from the consumers side of the zone file access system which was Rod Rasmussen, the CEO of the internet security company called Internet Identity and the second co-chair was myself and my name is Vladimir (last name) and I'm a Director of Policy for Telnik the Registry for .TelTLD.

The problem which we, the way we express the problem that was before our group was as follows: we acknowledged that consumers have issues with the current system of the zone file access. We think that there are possibilities to improve that model in a way that will provide scaling it in an environment where there are numbers of parties, both from the consumer and the producer sides. The problem statement is published on the zone file access group web page.

Now I would like to turn the presentation to Dave Piscatello.

Dave:

Thank you Vlad. I'm going to talk a little bit about the burden that we perceive for both the consumers of data and the providers of data. The providers of data are the Registries and I'll begin with them. The first is obviously a business arrangement issue and has to do with managing contractual relationships and agreements between each of the parties that the Registry provides access to.

The process was a paper process in one sense because you have an application by a party, there has to be a review of that application to determine whether or not a party has an appropriate need. Then once the Registry approves that party there is an ongoing maintenance of that contract, keep the contact information ready and accurate and to essentially make certain that in some sense there are no stale contracts, people who stop using the service and no longer need to be contractually bound and the like.

The second is an infrastructure burden and this is maintaining some sort of bulk transfer service for the approved applicants. Typically this is using file transfer protocol. Within this burden you have the maintenance of accounts for each of the applicants, the creation of credentials using a name or password typically or some other authentication method. Maintaining security of the connection meaning providing some assurance that only the authorized users can authenticate and connect to the file server that the Registry provides. Providing them with connectivity sufficient enough to allow both parties to potentially very large files and providing sufficient bandwidth where those files can be delivered in a reasonable timeframe.

Obviously, the extreme example of that is the com zone with is 90 million com, domains is a very, very large file. You can't offer that service using a very, very small bandwidth. In addition to those burdens there is an ongoing real time need to check the service is not under attack by some form of denial service or some other kind of abuse. For example, scanning, trying to do a root force password attack and mitigating those kinds of abuses or potential attacks.

Lastly and one that actually has a pronounced effect on consumers is change management. If the Registry needs to renumber its IP's for the file server or needs to do some other troubleshooting to understand whether the routing to the file transfer service is correct or some other issue, all that is something the Registry has to do as a service provider. Much like anything anyone would do if they were running a public file transfer service or web.

And then lastly there is customer care. When something goes wrong the Registry does have to have someone to call for the consumers to understand what the problem is because the data are extremely important for each of the consumers.

Looking at it from the zone data consumer perspective, again you have 2 relationships. You have one which is contractual and you have one which is infrastructure. The contractual process here is relatively the same for the data consumer except the data consumer has different contracts that it would possibly sign with GTLD Registries. Now they are all pretty much proforma but there are some nuances that some of the Registries might consider. There are some issues that invariably arise when anyone is looking at a contract. So all that requires some sort of legal staff to process not only the applications but also do all the contractual maintenance for each of the data consumers.

Here again, you have the FTP client side of the connection for the data consumers and for each of the GTLD Registries the data consumer has to manage the account credentials, has to know the IP addressing and other access controls and other security mechanisms that the individual GTLD Registry (34:50 tape goes out)...

Mark:

File access advisory group will be doing a lot of work in this area in the coming months. Go to the next slide. Let me and that's a nice segeway for me. Let me talk a little bit about what is

going to happen for these zone file access advisory group. Vladimir and Rod talked about the public comment period that is currently open on the concept paper. That concept paper public comment period is open I think until the 8th of April and that is relatively easy to find online.

The next step here is really to get the advisory group back together to refine and do more work there, especially to incorporate more information from the provider community on zone file access class. To talk at much more length about the funding model and how that would work. Then those 2 things taken together will help inform a decision about what a candidate model would look like. I put and I think a couple of members of the advisory group advised me to do this, I put model and then in parentheses the letter “S” here with the idea that it might not be a single model.

What you would like to do is simplify as much as you can sort of ala Akum’s Razor but it might be a situation where certain consumers were better served by one model than another. So one thing you can imagine here is that there would be multiple models of operation. On the other hand, multiple models for operation opens the question of consistency which is one of the goals here that Dave talked about in the early part of the presentation.

Clearly what is intended here is to have some concrete specifications laid out in the next version of the applicant guidebook. Really the key thing here is for the advisory group to do some work here in March and in April so when the next version of the applicant guidebook is produced that in it is at least a fairly comprehensive view of how zone file access should work for the new round of GTLD’s.

At the bottom of the slide and Craig talked about this at the beginning of the call, is the URL for all the zone file access material that includes not only a link to the concept paper and a link to the cost model but also a link to all the notes from meetings and all the MP3 recordings. Again, if you’re interested in further details on the work of the advisory group that URL is pretty useful.

I think now I’ll turn this back to Vlad to finish up and start with questions and answers.

Vladimir:

Thank you Mark. So we already have one question in the chat, Pat King asks how there are heavy users when the current (38:33 – inaudible) download. May I answer that question Pat? There are users who just download the files occasionally or maybe even once or twice since they signed the agreement. There are patterns where users do that on a daily basis. Also there are users who just access a single zone file for a single GTLD and let’s say professional users who download multiple TLD’s. So these are the amount of load that different users put on the zone file access system where it really varies.

Pat:

The primary reason I was asking is because the way Mark was talking I kind of interpreted that as being an ongoing act over the course of the day. And one of the things I'm looking at from how we update the zone file is if there is an expectation that the data be near real time with the zone file access that there is simplification as to how the Registry manages providing that data. That is why I said that.

Mark:

And let me be clear because I misled you, unintentionally of course. Right now the zone files are actually provided very, very often, once every 24 hours and one of the things we want to be very careful to preserve is that for other uses where people have near real time or quasi real time, say 5 minute requirements, allow Registries to innovate in that space, allow them to create broader services in that space and not get in the way of that. So don't do anything to interfere with if Registries want to produce that and invest in that kind of service let them go ahead.

But the other kind of zone file access when you think about for instance like in the intellectual property community, probably don't need that immediate real time access. So what they're looking for is the sort of continued provision, the old style and I don't even want to call it old style, but the sort of provision a zone file that we currently have.

Pat:

Right and so just in discussions I've had with security analysts, basically getting a once a day download of the zone file you could actually get criminals come in and exit the zone file and never be discovered because they may have a very short window with which they come in and come out. So I've talked to Rod in the past about things like this and so the real advantage of something like this is going to be for the researchers and to your point not the trademark people because you want persistence and trademark is more of a persistence issue than the security stuff we're talking about with near zero time detection.

Mark:

Exactly.

Rod:

Vladimir we have a question here or a comment from Cathy Kleinman. Kathy you want to step on up.

Cathy:

Hi all it's Cathy Kleinman, new Director of Policy for PIR. I was wondering if there are any other types of questions we should be asking as we look at these other models. A disclaimer I'm new to this area. But are there other implications or ramifications of the new mode? Specifically are there any other, are there some dangers we should be looking at when we look at some of

these models with the consolidation and the access? Are there any dangers or issues about privacy? Thank you.

Dave:

Cathy certainly we'll take a look at that. I think we're actually having a conversation on the zone file access list now at the moment that is relevant to your question about privacy. The domain system is a public query response service. There is usually very little in a TLD zone file that discloses anything in a bulk download that wouldn't be disclosed in a single query. There are some mechanisms you could use to filter the information that are accessed to a query response but generally that's not the case. The zone files that the Registries operate don't have information that is in any way personally identifying.

So from that perspective of privacy, I think we're not changing things too much. If you're concerned about the question of the disclosure of who the data consumers are by the 3rd party or exposure of parties in the Registry, we would expect that whoever might eventually operate this service as a mutual operator would be someone who would be vetted for good business practices and would practice all the data protection best practices that are expected of a legitimate business partner.

So I don't think that that is something we would overlook once we got to the point where we would have an RFP to outsource this. And certainly today there are lots of security service providers or managed network providers who fall into that category. So I don't think we would spin this off to a brand new organization that has no experience in this field.

Did that answer your questions?

Cathy:

Yeah I think that answered it but let me do a follow up. Are the ccTLD's also part of the process being evaluated here? Do they use centralized zone files? If not, why not?

Dave:

I'll take a shot at that too unless someone else would like to. So the program is for GTLD Registries. If there were a ccTLD Registry that would want to participate in the program, I don't personally and I'm not speaking for ICANN, I personally don't see there would be any obstacle other than whatever privacy laws or other governance over the data and the operation of that TLD guided the Registry. So if someone wants to join the party so to speak, and they're not obstructed in any way legally, then I imagine that we can entertain having them.

There are some countries where there is, I'll make my own interpretation there is inflation of preceding the information a zone file and concluding it is relatively the same as the information that is available through WHOIS and essentially saying we don't want to allow bulk access in

either case. We're actually trying to tease that out as an issue among the, on the zone file access mailing list. In fact, I'm composing an email right now on a thread.

So the only thing that really prevents a ccTLD from participating is its own business decision and any sort of regulation that the country imposes on it.

Craig:

Dave thanks for that response and we have a couple of people in the queue now. We've got Jonathan Frakes standing here; we have Brian Cute who I think is in the room with you. Do you have folks in the queue on your end Vladimir?

Vladimir:

Just Brian.

Craig:

Let's go with Jonathan first and then we'll go to Brian.

Jonathan:

I'll go ahead and let Brian go first.

Brian:

Wow thanks Jonathan. Cathy to your question a couple of points in addition, in the existing zone file access agreement that Registries execute with users there are provisions under Section 4 which prohibit the users from among other things using the data for marketing purposes, obligating them to comply with all applicable laws and regulations governing the use of the data. So those constraints do exist in the existing agreement. And there has been Registry sentiment expressed to your point in the sense that if there were a 3rd party that in the repository model, for example, were to have access to all the zone file data across all the TLD's a recognition that that entity would be in a position itself where its use of the data would have to be appropriate and bounded much as the users would be.

So that is one point that has come up in the discussion and that is in the contract for you to review. Then secondly, a sentiment from the Registries that no matter what model is adopted or remains in place going forward that Registries want to have the ability to control access to their zone files. And if there are abusive situations or uses of the data, potentially including the privacy question you raised that whatever the model is we want to be able to have quick ability to control and shut down that abusive access.

Craig:

Rod it looks like you're next in the queue. Do you want to make a comment Rod?

Rod:

I just wanted to make a quick point on the previous question on the ccTLD's and providing access. There are some ccTLD's that provide access and they're typically very small and it is a direct transfer, zone transfer via DNS transfer commands. So you can get their zone files but very, the larger ccTLD's do not in general provide access although there have been some talking about doing so and trying to enhance security. Again, in the anti e-crime area trying to assist in those efforts but not on a standardized large scale basis.

Craig:

Thank Rod. I don't see any other hands raised here, actually we do we have Pat King coming to the mike.

Pat:

I put a couple of questions on chat that had to do with a consolidated model and how the 3rd party operator would operate. One of the questions I would have is around compliance around the usage items within the agreement today. I know what VeriSign does when we actually bring in a person for the zone file access; we actually go through the same process and authenticate people for certificates. So we go through an authentication process for that.

It is much harder to end up doing compliance on usage of the zone file. It's harder to track but by authenticating people on the front end that is a process where I know there is costs we incur but the 3rd party would also have to do that. But when you also get to a liability of the accuracy of the data, one of the things that I struggle with is when somebody else is publishing the data you provide to them where does that liability for accuracy fall?

Mikey:

This is Mikey and let me take a turn. One of the things we're still working on is sort of contractual framework of all this and we have a couple of threads going on that. One of the issues that came up fairly early in the conversation was who would the contract be with? There would still be bilateral contracts, what was the impact of the introduction of a 3rd party be on those? That is another part that is not quite spaced yet.

Pat:

So your suggestion would be that you would still have contract maintenance with the Registry and you would have a single serving up through this 3rd party that would control, would be a benefit to the consumer not necessarily off loading the work of the Registry.

Mikey:

The thought would be that you probably would get a similar benefit to the Registry. So if you imagine sort of today's situation on either end of the transaction, if you're a consumer or a provider. You look out into the world and you see the universe of the other end of the transaction. In this model, on either end of the transaction you would now look to one place and then you would have a bunch of choices that you can pick from. And the premise of all this is by going from looking at many things to looking at one thing that it will be easier to manage, cheaper to do, reduce the cost, etc. and give you a place to focus on for service.

That is the theory and we haven't quite baked it yet. One of the things that is embedded in all that is how does the contracting work in that situation? One of the things we've concluded is no way are we getting rid of the direct contractual relationship between producer and consumer. That is not on the table at all.

Pat:

Okay.

Mikey:

But facilitating the management of those consumer produced repairs is what we're after. It's 4 in the morning and probably better to do this when we both had a little more sleep.

Craig:

Anybody else want to get in the queue from Nairobi or on the phone? Oh Jonathan is coming up.

Jonathan:

I made a point in the chat and for the benefit of the folks listening, some of the things we were looking at is we want to make sure this data is accessible to emerging nations and looking at different methods of access. We talked about potentially offering like Delta, so a person can download just one main file and then the difference each day so it would reduce the bandwidth in places where bandwidth is an expense.

We looked at different things like what the frequency of these zone files were, the format of them and just basically things to make this more efficient and streamlined as part of what we're doing. So it's not just aggregating it and things like that.

Vladimir:

There are questions on the chat about the technical aspects of the zone file access in the future models. We really had some discussions about various improvements in the data protocols such as P2P networks, bit torrent or a number of other options. But I would say this is a task for the next stage of the group work because first of all we need to find out what is the structure of the

system that works best for everyone. Only then we can think of what technical solution will be the best here.

Craig:

Do we have any other questions?

Rod:

We don't have any further hands here in the Reston hub. What about you in Nairobi?

Vladimir:

I don't see anyone.

Craig:

I think with seeing no additional hands and/or questions on behalf of the zone file access advisory group I would like to thank everyone for their participation today. I would like to remind everyone that the public comment forum on the zone file access concept paper runs through the 8th of April and that we will capture all the information that was discussed in the chat room online today that may not have been raised on the phone bridge and be sure to address that in the work we take on going forward.

A reminder to about the zone file access information page that has a copy of all the MP3 recordings of all the teleconference calls that the group has had as well as a host of other information, including who the advisory group members are and a copy of the original concept paper, as well as a public archive of the email list that the group uses to conduct business.

So with all that shared and seeing no further comments I would like to thank the advisory group co-chairs and the advisory group members who participated in the call today and to thank all those who showed up in Nairobi for the workshop and those who dialed in from wherever in the world you are. And to say we look forward to your continued participation in our work.