Systems Engineering Update
OARC 31, Austin

Matthew Pounsett

2019/10/10

Contents

Contents **1**

1 Introduction **2**

2 OARC Services Overview **2**
2.1 Data Archiving 2
2.2 File Servers and Storage 3
2.3 Data Analysis Servers 4

3 System and Service Status **4**
3.1 General Condition 4
3.2 Web Services 4
3.3 File Servers 5
3.4 DNSVIZ 5
3.5 Analysis Servers 6

4 Conclusion **7**

1 Introduction

This summer has seen significantly slower progress in improvements to OARC's systems, owing primarily to ongoing hardware problems and resulting difficulties in managing data. The bulk of Systems Engineering time, over the last few months, has been spent on mitigating these issues. Yet, we are entering the fall in a much better place, with the DNSViz historical data on-boarding nearly complete, and our DITL collection stabilized while we make plans to improve the underlying infrastructure.

2 OARC Services Overview

2.1 Data Archiving

OARC maintains a large store of multiple data sets.

Day in the Life OARC coordinates annual and occasional ad-hoc Day in the Life (DITL) DNS traffic capture events. These involve many operators of significant DNS infrastructures—including root server operators, TLDs, and recursive operators—running packet captures of their traffic over the same 24 hour period. The data are uploaded to OARC where it is organized for use in research.

The DITL collections go back to 2009.

DNS Statistics Collector DSC is a data collection and statistics generation tool for DNS. Several members contribute their DSC data to OARC, and we make this available to all members to view on our centralized DSC installation.

There is more information on DSC at `https://www.dns-oarc.net/oarc/data/dsc`.

RSSAC 002 Statistics The Root Server System Advisory Committee's publication RSSAC 002 is the Advisory on Measurement of the Root Server System. It defines an initial set of statistics to be collected by root server operators from their systems. OARC collects the output of this reporting from each root server operator, daily, and maintains a history of these statistics available for analysis or review.

Zone File Repository OARC maintains an historical archive of zone files which includes daily updates of the root zone going back to 1993, and weekly updates of several TLDs beginning at various times between 2009 and 2018.

Other Data OARC also periodically accepts submissions of other data that may be relevant to researchers interested in the DNS:

- derivative data from research done on OARC's other datasets

- data collected from OARC testing tools, such as the DNS Entropy Tester
- DITL-like collections from outside regular DITL windows, such as ongoing contributions from AS112 server operators
- packet captures from OARC's Open DNSSEC Validating Resolver (ODVR) which includes forwarded queries from the DNS Privacy Testbed
- Case Western Reserve University's "Case Connection Zone" FTTH data
- other ad-hoc contributions of relevant data

The new implementation of the member portal does not include the joint presenter for the DSC data that the old portal had. In recent years the number of contributors has declined, and the number of people viewing this data has declined even more. We are considering options for the future of this data set. As discussed in the previous Systems Engineering report, one possible option is to stop accepting DSC data from members entirely, and archive the existing dataset. Another option is to continue to accept data, and set up a new presenter, possibly based on Grafana, as future work. We have not heard from any members since OARC 30 about their interest in this service; we are still interested in feedback from members about how useful they find this service, and what path we should take with it in the future.

2.2 File Servers and Storage

OARC's datasets are stored on six file servers. The first five file servers, located in Fremont, California, have 424.31TB used of their 532.64TB of capacity. Two of these have multiple filesystems, marked as A and B in the chart below. The sixth file server, located in Ottawa, Ontario, is an off-site copy of a selection of datasets from the first five servers.

File Server	Used	Capacity
FS1	118TB	121TB
FS2a	36TB	42TB
FS2b	25TB	125TB
FS3	34TB	42TB
FS4	72TB	84TB
FS5a	69TB	84TB
FS5b	33TB	42TB
FS6	117TB	121TB

Each file server uses either ZFS (RaidZ2) or XFS over software RAID for its filesystem to provide redundancy within the file server. Each dataset is stored on more than one file server in order to create cross-chassis redundancy of data; some datasets currently have copies on three systems. This means

that the total size of all unique datasets is slightly less than half of the 504TB indicated above.

All capacity numbers above are the filesystem capacity, rather than the raw size of the disks in service.

2.3 Data Analysis Servers

OARC maintains four UNIX shell servers with access to the above data sets. Three in Fremont, CA (an1, an2, an4) and one in Ottawa, ON (an3). Members and participants who have signed a Data Sharing Agreement and request access are given accounts on these analysis servers, which they can use to do research into the DNS using any of OARC's datasets.

> **Note Well**: No data, even derived data, may leave OARC analysis systems without express written authorization, in compliance with the Data Sharing Agreement. Contact admin@dns-oarc.net first, *always*.

3 System and Service Status

3.1 General Condition

With very few—but notable—exceptions, OARC's systems and network continue to be significantly more stable than they were this time last year. Unfortunately, the issues that we have experienced have been disruptive enough to delay much of the expected progress replacing or updating older systems and software.

However, some progress has still been made improving the stability of, and and reducing the distraction caused by, several smaller services and processes.

Some of the more visible and time-consuming work is disccused below.

3.2 Web Services

In the last few weeks we experienced an outage of the web server which provides many of our public-facing web sites, including OARC's main web site. The web server in question was the last remaining one using custom compiled software, rather than OS-packaged software, and had some known issues which were limited to some internal-use sites. The outage was extremely similar to of a similar outage experienced early this spring, mentioned in the previous Systems Engineering report.

The config for the failed web server was updated to use more current, OS-distributed software, and many of the more esoteric configuration choices of the older config were investigated and either cleaned up or removed. The

replacement web server has been performing fantastically since, and load on the host system has been reduced.

3.3 File Servers

As mentioned in the previous Systems Engineering report, OARC's file server infrastructure is aging and beginning to experience an increasing frequency of odd issues which appear to be hardware related. These include, but are not limited to, crashes and data errors when the systems are put under load.

The issues with our current file servers have consumed the majority of my time over the past five months, have delayed us in our ability to make available the fall 2018 root KSK revocation DITL collection the regular spring 2019 DITL collection, and have prevented us from accepting uploads of an additional root KSK removal collection done by the root operators.

The need for a significant hardware refresh, combined with the operational load required by the current architecture, has led us to the conclusion that we should replace the existing file servers and their eight separate NFS shares with a single unified clustered file system. This plan was discussed at a high level in the previous Systems Engineering report, and is covered in more detail in the separate File Server Infrastructure Proposal which will be distributed to members prior to OARC 31.

The decision to pursue a complete replacement of the existing infrastructure has led us to delay the annual summer replacement of aging file server drives. We expect to put those funds toward new servers (and storage) later in the year.

The problem of offsite replication of the data housed on our file servers is still constrained by a combination of budget and our data sharing agreement. We cannot afford to run a second, entirely duplicate file server infrastructure offsite, and our data sharing agreement prevents the duplication of that data onto systems not controlled by OARC, even for backup purposes. Possible changes to the data sharing agreement are still under review by the Board.

3.4 DNSVIZ

The recently acquired DNSViz service has been the other primary use of my time since OARC 30. As discussed in the previous Systems Engineering report, some early issues with the historical database have snowballed into an extremely time consuming cleanup job.

The restoration of this database is nearly complete, and we anticipate that the historical functions of the DNSViz service should be restored before OARC 31. This is thanks, in no small part, to the help of DNSViz's author, Casey Deccio. Casey has been instrumental in the effort to correct errors introduced into the data during our cleanup efforts.

OARC has already purchased a replacement database server which will be put into production within the next few weeks, very soon after OARC 31. We

will use the replication functions of the database software to populate a copy of the database on the new hardware, move production service from the old to the new database, and then reverse the replication direction. The old database server will remain as a hot backup until the budget allows us to replace it as well. Offsite backup of the data will be handled by a weekly export copied to cloud storage or our systems in Ottawa, Canada.

3.5 Analysis Servers

User Cleanup

Cleanup of the users on the analysis systems, and synchronization of those users' meta-data with the portal database, was a huge success. Many thanks to those individuals who helped with sorting out inconsistencies. We are now in a much better position to implement gradually increasing levels of automation in the management of member shell accounts.

Initial steps in that direction will be invisible to end users, but will significantly reduce the operational workload of supporting and maintaining member shell accounts.

OS Refresh and Additional Hardware

Planned upgrades to the analysis servers are likely to be delayed, as we may be required to re-purpose those funds to support the above improvements to the underlying file server infrastructure. However, if time and budget both allow we may go ahead with the planned addition of local "scratch" storage to the existing Fremont analysis servers late in the year.

Likewise, the OS refresh of the analysis servers will be put off to this coming winter. The limited amount of time available, due to the pursuit of other problems, has not allowed this to proceed as planned this summer. The analysis servers have also been surprisingly busy this summer, which would have made timing this work awkward, even if the the hours had been available.

Indexing of Data

Another possible enhancement to the analysis infrastructure which is under consideration is the addition of a searchable copy of our DITL datasets, and other DNS packet captures. I'm hoping to take advantage of the analysis work already done by Sinodun and ICANN into various options, and the work they have done importing ICANN's root server data into a column store database.

We feel that making the DITL datasets easily searchable would significantly improve the usefulness of OARC's data, and greatly simplify many of the research projects currently done on our analysis infrastructure.

However, if we were to pursue a project like this it would be above and beyond OARC's typical hardware budget, and would likely need to be presented to the community as a project needing additional funding.

4 Conclusion

Some significant challenges have consumed a large amount of my time and attention in recent months. However, I believe we're about to push past those distractions and regain much of the momentum we had last fall and winter. I expect we should have some very visible changes to report by OARC 33, in the spring.