

OARC Systems Update

DNS-OARC Workshop
Warsaw, Poland
May 6, 2013



DNS-OARC

Domain Name System Operations Analysis and Research Center

DNS-OARC Services

- To recap, there are a number of services DNS-OARC provides
 - Access to data archives (mostly DITL captures)
 - Mailings lists, websites, collections, accounts, database servers, ODVR, ZFR, DSC, don't-probe, kitchen sinks, etc.,
 - Dedicated capture systems to receive and process imported data types (DSC, DITL data)
 - Various collections of support systems, like DNS, out-of-band management and network switches
 - Monitoring systems such as TLDMon and DSC
 - And in one corner, AS112 number resource registrations plus website



DNS-OARC

Domain Name System Operations Analysis and Research Center

System Status I: Migrations

- We're now down to just one aging server, the main one, in1
 - The mail system (user accounts, mailing lists, mail archives)
- ix1 now holds the majority of OARC public-facing and 'corporate services' save for the above
- ix2 holds anything else research and analysis (ODVR, etc)
- TLDMon has migrated from FreeBSD to Linux, and is revitalised
 - TLD holders may be interested in being notified directly, not just ICANN only
- ZFR has migrated to a dedicated Linux server



DNS-OARC

Domain Name System Operations Analysis and Research Center

System Status II

- Backend OARC data (i.e., everything not DITL, ODVR, etc) is an issue. Where does it go if the larger file servers take priority?
 - A recycled Sun Fire X4500 has been put into service as 'thump3' to house OARC internal data, freeing space on the file servers
- Another recycled Sun Fire x4500 was commissioned to act as a fourth NFS mount to the analysis servers, containing 100% DURZ data from RIPE-NCC



DNS-OARC

Domain Name System Operations Analysis and Research Center

System Status III

- The DNS Lab servers acting as DITL capture boxes also had their OS upgraded and are current
- DSC system is stable but a bit short on DSC upload data (more at about this at the end)
 - Shared to act as backup server with tape (using BackupPC)
- RIPE was kind enough to send us their DURZ data, 13.7TB transferred from Amsterdam in 4 weeks using GridFTP-Lite (thank you Iñigo Ortiz) after the October 2013 workshop



DNS-OARC

Domain Name System Operations Analysis and Research Center

System Status IV: Jabber

- Pain!!
 - Relief is spelled: 'rm -rf <whatever>'
- Despite the quirks found within various ejabberd versions, OARC Jabber was finally migrated to ix1, with a software refresh done on the venerable nagios/cc1 server
- Ejabberd was also reconfigured to be more resilient beyond that of just DNS round-robin and hoping for the best
 - Internal ejabberd 'mnesia' databases are now synchronizing so that if one of the pair of servers goes down, the other continues
 - However, we are asking a little bit too much of the current configuration
 - Ejabberd clustering should be within a data centre, not across continents or oceans according to an old post on the ejabberd mailing
 - Despite this, clustering works for OARC Jabber today as configured
- The same was done to the root-ops Jabber: No longer a jail on another machine, now full-blown dedicated hardware isolated from any OARC shenanigans.



DNS-OARC

Domain Name System Operations Analysis and Research Center

System Status V: Indico

- Indico is OARC's successor conference and workshop management system which replaced the version long used under Drupal and some custom mechanisms.
 - This is not to be confused with live conferencing systems such as Webex, etc.
- Peculiar incidents involving instabilities within Indico have caused a lot of headaches
- Fortunately, BackUpPC has rescued us a couple of times
- Triggering a 're-migration' seems to have solved some of the problems with the database backend.
- Indico seems to be stable at time of writing
 - But Keith may have other plans



DNS-OARC

Domain Name System Operations Analysis and Research Center

Other Migrations and Updates

- dns-tender and the DNS lab need some extensive documenting and updates to available OSs for the lab
 - In need of replacement SCSI disks as well
- The mail system
- The Drupal website which fronts as the face of OARC on the web
 - This one in particular needs special consideration
 - What's the intention, how to express it (fluff vs content), etc?
 - The current Drupal version and 'theme' used by OARC is un-supported by Drupal, but much OARC process, internal database organisation, auto-page updates, ODVR, ZFR graphs etc. have been built around it to fully take advantage of the features
 - This is not a trivial project to undertake
 - But was successfully migrated on the night of May 1, 2014, as-is
- Dontprobe service is working well, but would like to hear from anyone using this service for feedback, besides dns-surveyor
- Heartbleed: Despite only looking at version numbers for OpenSSL, all systems have been patched against this already



DNS-OARC

Domain Name System Operations Analysis and Research Center

“Collisions” boxes

- Donated hardware each consisting of
 - 64GB of RAM
 - 4.5TB disks
- Running as an3 and an4
- NFS mounts to fs2, fs3, fs4 and thump4 repositories of DITL data.



DNS-OARC

Domain Name System Operations Analysis and Research Center

Data Archives Status

- fs2, 6.5TB of 17TB used
- fs3, 20TB of 22TB used
- fs4, 42TB of 45TB used
- Thump4, 14TB of 15TB used
- Total: 82.5TB used, 99TB total possible
- Put another way, the file servers are now nearly full
 - (Each DITL run can consume 10TB when all is said and done depending on duration of contribution and how many contribute.)
- Access to that data available via an1, an3 and an4!



DNS-OARC

Domain Name System Operations Analysis and Research Center

More Data Archiving

- fs1, planned box for some time this year
 - To become a copy of all archived data giving us the choice to free fs2, fs3 and fs4 for any capture runs we require.
 - Partially acquired, still short of the full compliment of 4.0TB desktop SATA disk drives
- This is a StoragePOD, based on the famous Backblaze blog articles
 - See 45drives.com for more info



DNS-OARC

Domain Name System Operations Analysis and Research Center

Fs2, fs3, in, an2, tldmon, etc



DNS-OARC

Domain Name System Operations Analysis and Research Center

DSC

- DSC as it is implemented at OARC won't scale for very much longer, unless more hardware and CPU are dedicated to the effort
- An alternative is to do what some others are doing: stop sending data to OARC to process, and just send us the post-processed .dat files instead via rsync
 - Everyone is processing the data already
 - Why process it again if the outcome is simply to display pretty graphs in the presenter under the Portal?
- Long-term, this is the most scalable solution for this OARC service
- Comments appreciated



DNS-OARC

Domain Name System Operations Analysis and Research Center

Growth Rate of Data

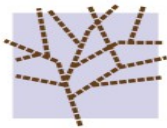
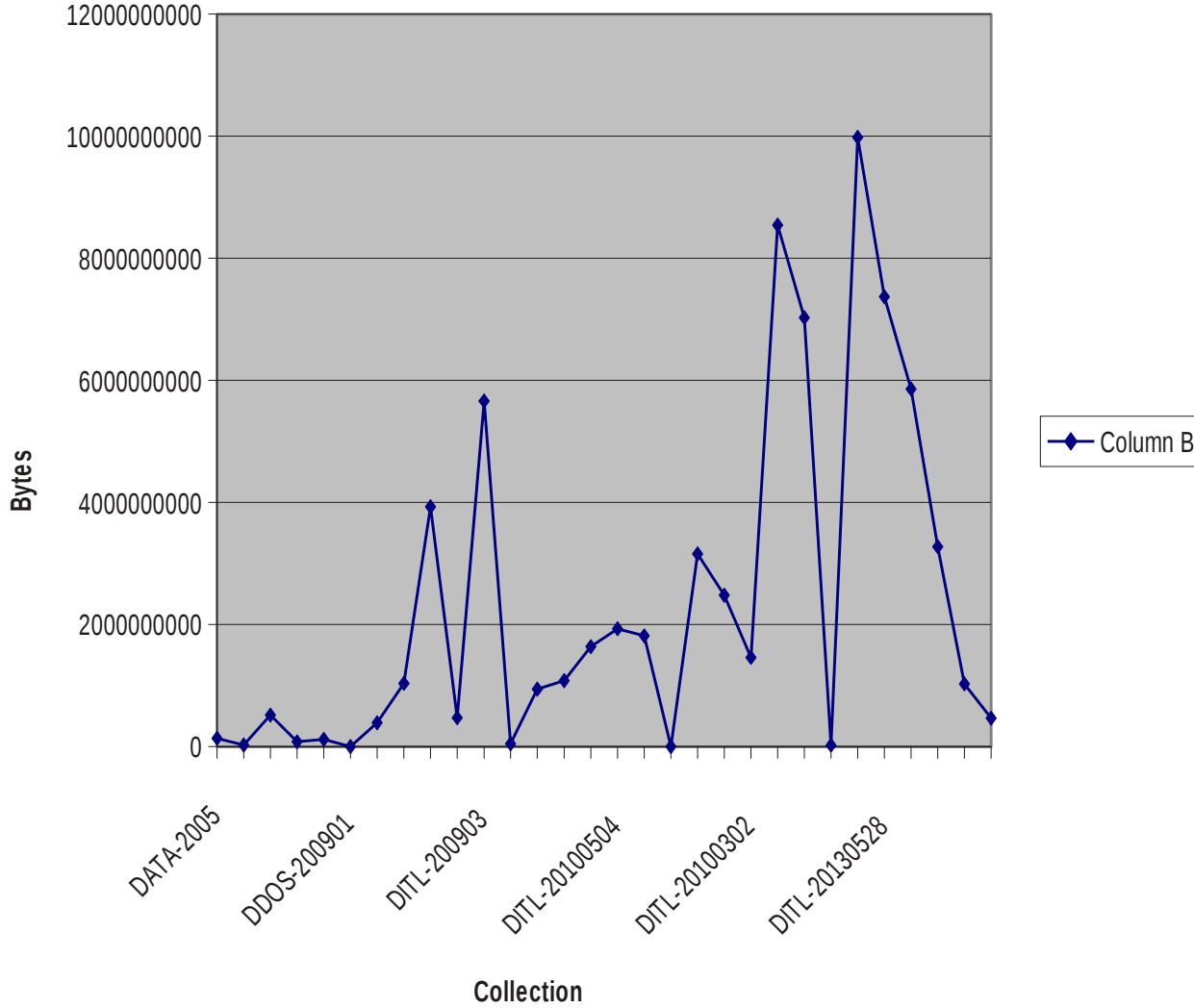
- Given those capacities, what is our data growth like for the foreseeable future?
- What is the impact of a special collection triggered by a predicted event or DDOS attack in progress?



DNS-OARC

Domain Name System Operations Analysis and Research Center

DNS-OARC Data Collection Sizes



DNS-OARC

Domain Name System Operations Analysis and Research Center

DITL Collection Interpretation

- In 6 years the amount of data collected looks to have more than doubled, from 4TB to 10TB
- This trend is however dictated by:
 - Different contributors for every event;
 - The amount of data varies with the DITL duration;
 - The type of additional data requested.
- Nonetheless, the next purchase of hardware needs to be adequate for the next 5 years, assuming no deletion of any data



DNS-OARC

Domain Name System Operations Analysis and Research Center

DITL 2014

- Currently at 2.9GB of RAW data, and counting
 - Contrast with 5TB for 2013, 3.7TB for 2012
- There is a tremendous amount of data still being uploaded, so expect this number to increase
- Watch for announcement of data availability on the OARC website
- Thanks to all that participated.
- A new collection for in.addr-arpa? Perhaps.



DNS-OARC

Domain Name System Operations Analysis and Research Center

Future Considerations

- Consideration must be given first of all to the data:
 - What to do with this data (old versus new, retention)?
 - Where else on the planet to store a copy?
 - What to do about where it resides today?
- Also, we need better monitoring of OARC itself
 - Network utilisations are chief on the list
 - OARC-controlled gateway paves the way
- Infrastructure plan is under development
 - Includes jumbo frames and 10Gb/s in the back-end
 - Being circulated internally



DNS-OARC

Domain Name System Operations Analysis and Research Center

Future Path

- From a certain perspective, OARC needs to realise that it is walking like a duck, talking like a duck and acting like a duck. But what kind of duck?
 - In the future, there will be a requirement to have a duplicate of this data set elsewhere, which means moving a lot of data around securely with updates to the archives
 - Analysis of this data will also grow as OARC becomes more well-known in the third-party research (commercial or academic) arena.
 - There will be more data collected feeding the appetite of the above
- It's looking a lot like a science/HPC duck with a comparatively lean budget
 - In other words, it is business as usual



DNS-OARC

Domain Name System Operations Analysis and Research Center

Questions ?



DNS-OARC

Domain Name System Operations Analysis and Research Center