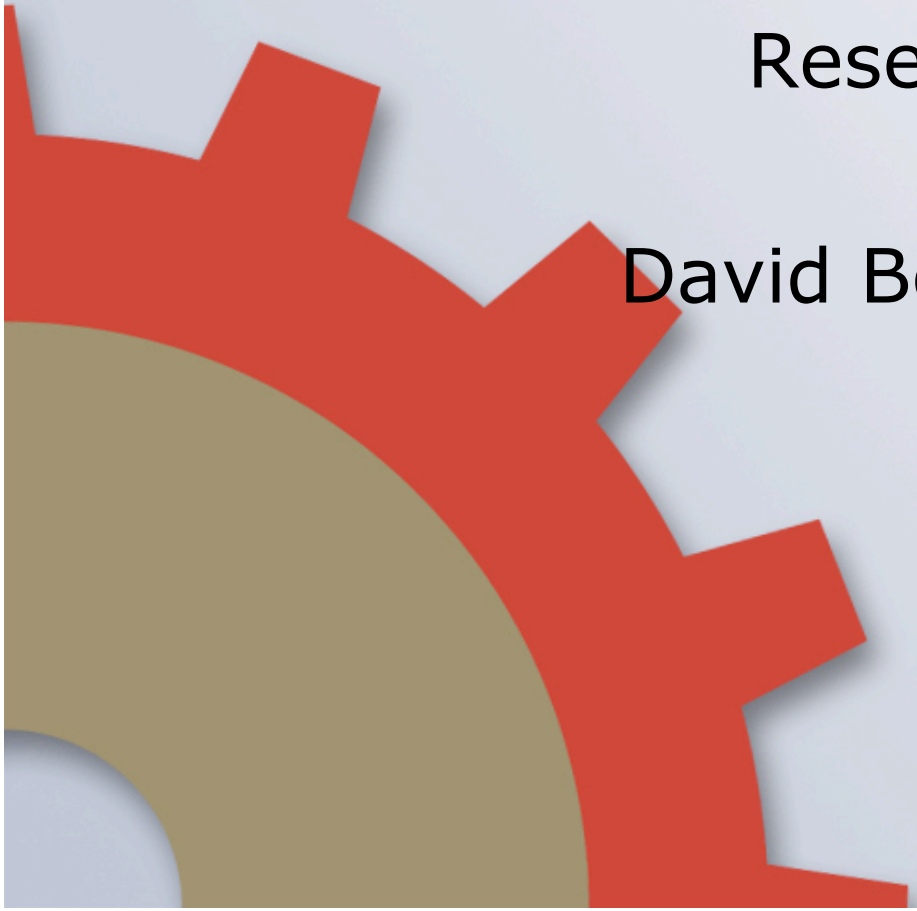
Testing DNS Performance limits



Research by ISC for CAIDA

Funded by NSF

David Boggs, lead investigator

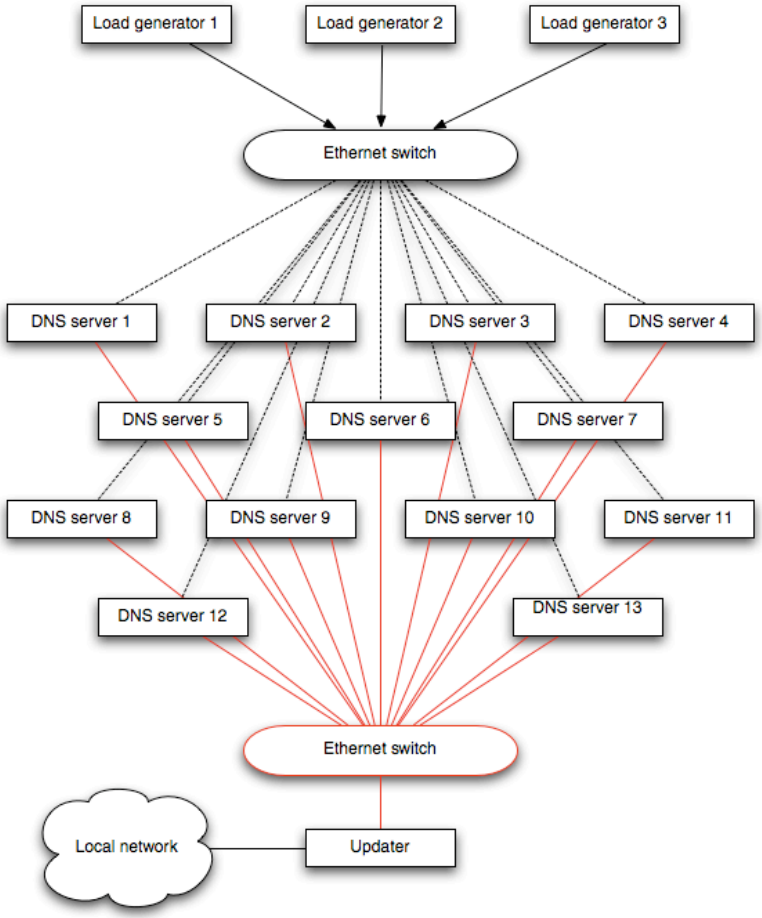
DNSPERF Project overview

- Build testbed big enough to test .COM, .NET TLD service
- Test its maximum capacity (query rate at server overload point)
- Reconfigure to use DDNS for updates, IXFR for distribution
- Test under load, find maximum

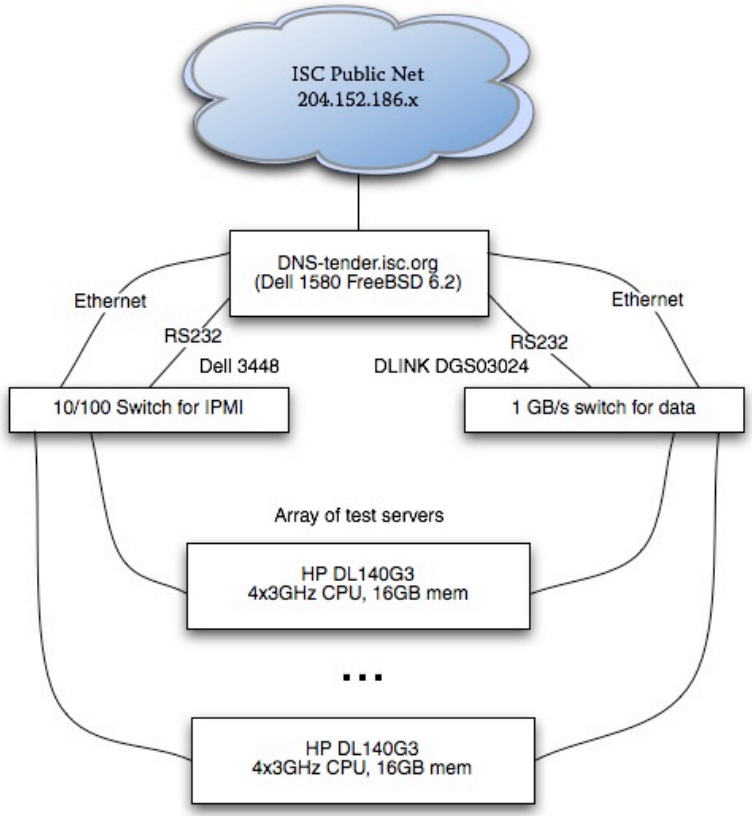
Physical testbed

- 13 affordable COTS servers
- 1 Stealth master for IXFR sourcing
- Non-blocking GBE connectivity
- Load generator
- Update generator
- Monitoring

Logical diagram



Physical diagram



What hardware?

- Affordable under limited budget ($100K available to buy 16 servers)
- Candidates: Sun X4200, HP DL140G3, Iron Systems I-class, M-class (Intel Xeon and AMD28x)
- Must run open-source OS
- Choose by memory performance

Hardware test results

	L1 Memtest MB/sec	LMbench Bandwidth MB/sec	L1 calibrator (NS for miss)	LMbench latency (NS)	STREAM Copy (MB/sec)	STREAM Add (MB/sec)	STREAM Triad (MB/sec)
HP DL140/G3	49058	2984	3.07	72	2586	2884	2890
Sun x4200 AMD254	22886	2316	3.48	83	1724	1896	1893
Sun x4200 AMD285	21251	2368	3.73	83	1816	1994	1958
Iron Systems M (AMD)	19717	-	4.08	-	-	-	-
Iron Systems I (Intel)	19607	2047	6.82	109	1309	1329	1524
HP Celestica	16331	1303	5.07	155	1122	1254	1138

Hardware decision

- HP DL140/G3
- Surprised that Intel processors outperformed AMD for these tests
- Able to afford 16GB RAM in each (8 pairs of matched 1GB parts)

What software?

- BIND 9.4
- OS: Test these, pick the fastest

  Linux (Gentoo, Fedora), FreeBSD (6, 7), Solaris 10, NetBSD 4, OpenBSD 4.1, Windows 2003 Server, Windows XP Pro64

What test?

- Loaded server with .PT zone
- Used queries from 48-hour F-Root capture, sent with queryperf
- Ramped query rate until server limit reached
- Ran test at server limit for 1 hour (1.13 millioin queries)

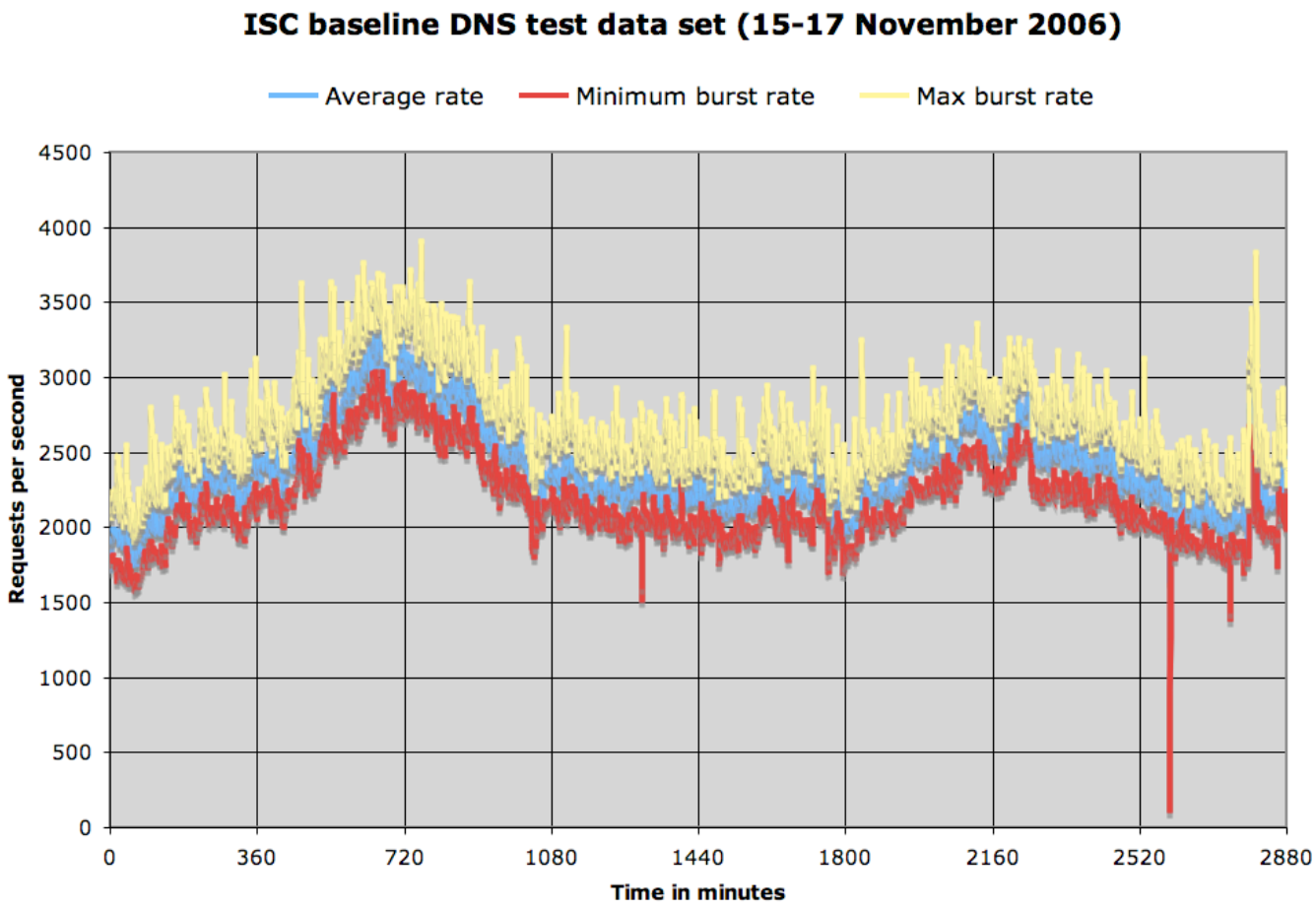
OS Performance queries/sec

Linux-Gentoo	Kernel 2.6.20.7	92327	Solaris-10	SunOS 5.11 snv-64a	41306
Linux-Fedora	Kernel 2.6.20.7	86732	NetBSD	4.0-beta2	36331
FreeBSD	7-current 200708	83089	OpenBSD	4.1-current 200705	35237
FreeBSD	6-stable 200708	54076	Windows 2003 Server	SP2 5.2.3790	22548
Solaris-10	SunOS 5.10 120012-14	53539	Windows XP Pro	SP2 5.2.3790	19888
FreeBSD	6.2-release	50611	Windows 2000 Pro	SP4 5.0.2195	18957

Test data stream

- 48-hour capture from F-Root
- 414931073 requests (38.8% failed)
- Avg rate (req/sec) = 2401.2
  95%ile burst = 3011.0
  Max burst = 3921.9

Test data stream



Testing with .COM

- Used COM zone from 5 Oct 2007
- 175,762,611 entries
- Raw zone file size 6GB
- BIND 9 RSS varied by OS from 9.2GB (FreeBSD) to 14GB (Linux)

Testing with .COM

OS	Queries/sec
Gentoo Linux	67900
Fedora Linux	65159
Solaris-10 (Proprietary edition)	BIND failed to start*
Solaris-10 (Open source edition)	BIND failed to start*
FreeBSD 7-CURRENT	56811
FreeBSD 6-STABLE	40512
FreeBSD 6.2-RELEASE	40239
NetBSD 4-CURRENT	BIND failed to start*
OpenBSD 4.1-CURRENT	BIND failed to start*
Windows XP Professional	BIND failed to start*
*BIND exited during initialization with an “Out of memory” error	

Next step

- One test remains: measure BIND performance during constant update
- Use nsupdate on “Stealth master”
- Use IXFR to update individual servers from Stealth master
- Feed generated nsupdate stream at controlled rate

Status

- Our hardware cannot cope with the full .COM zone (memory size)
- Truncating .COM zone at random until it fits
- Will re-generate nsupdate test stream not to generate update or delete of removed zone

For more information

- Website http://new.isc.org/proj/dnsperf
- Contact info@isc.org to inquire about research access to this testbed (it is available to other researchers)