DENIC‘s new DNS infrastructure

Christoph Strabel

<strabel@denic.de>

Prague 2010, OARC Meeting

Motivation

Set up a new environment that

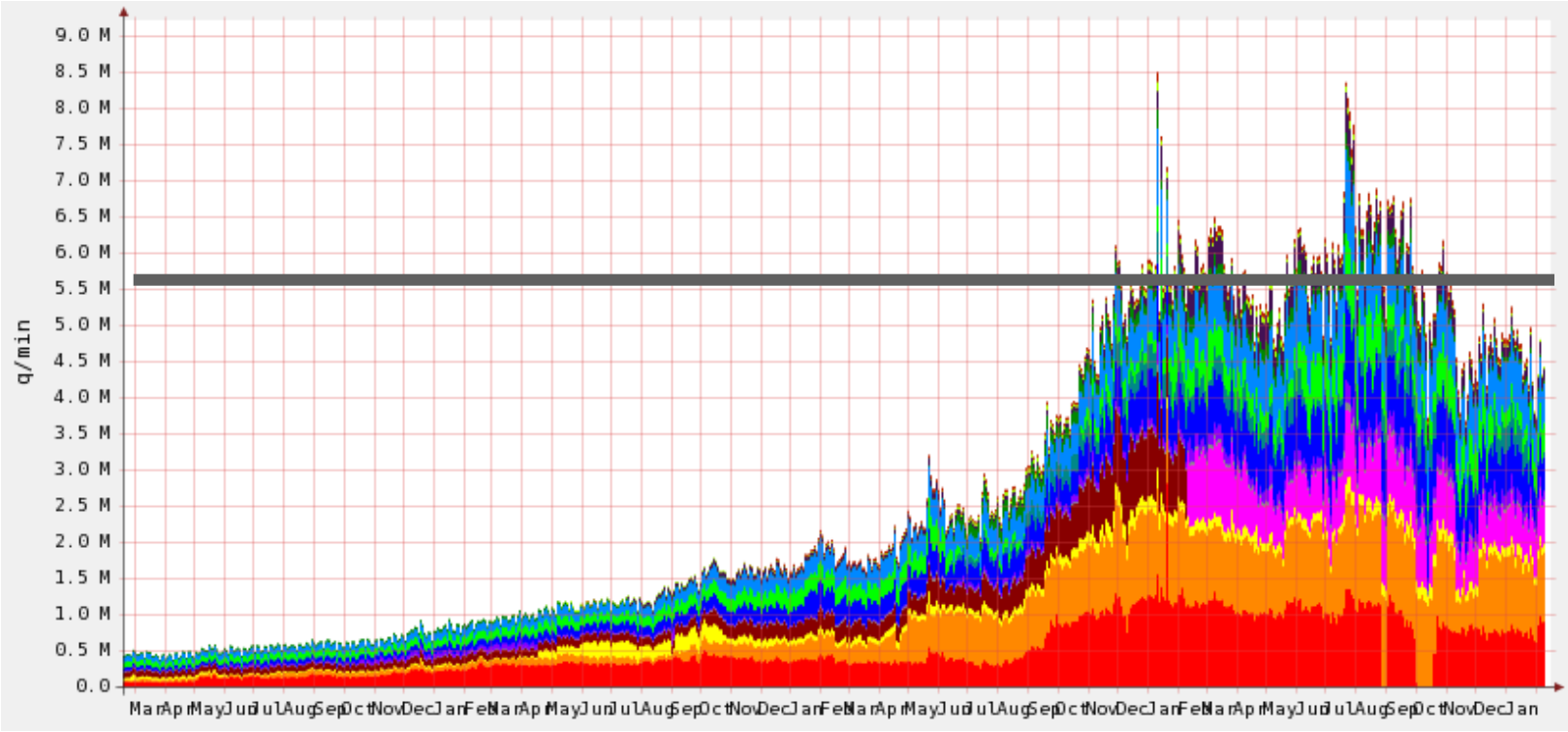
- **has an enhanced architecture**
- **will handle the query load growth for the next 5 years**
- **is scalable**
- **has optimized processes**
- **has a better cost efficiency**
- **can serve additional zones**

The old environment

- **Rollout in 2004/2005**
- **3 Nameserver and 1 Logserver per location**
- **Diversification of**
  - Server architecture (SPARC, AMD, PPC)
  - OS (Solaris, Fedora, Suse)
  - Nameserver software (BIND 8, BIND 9, NSD)
- **Reload every 2 hours**
  - Generating the zone and creating locally a diff
  - Copying zone/diff to each remote location (scp)
  - Recreating the actual zone in each location based on the diff
  - Distributing that zone to each nameserver within the location
  - Steppedly restarting each nameserver within the location

The old environment

Accumulated .de query load 2005 - 2009

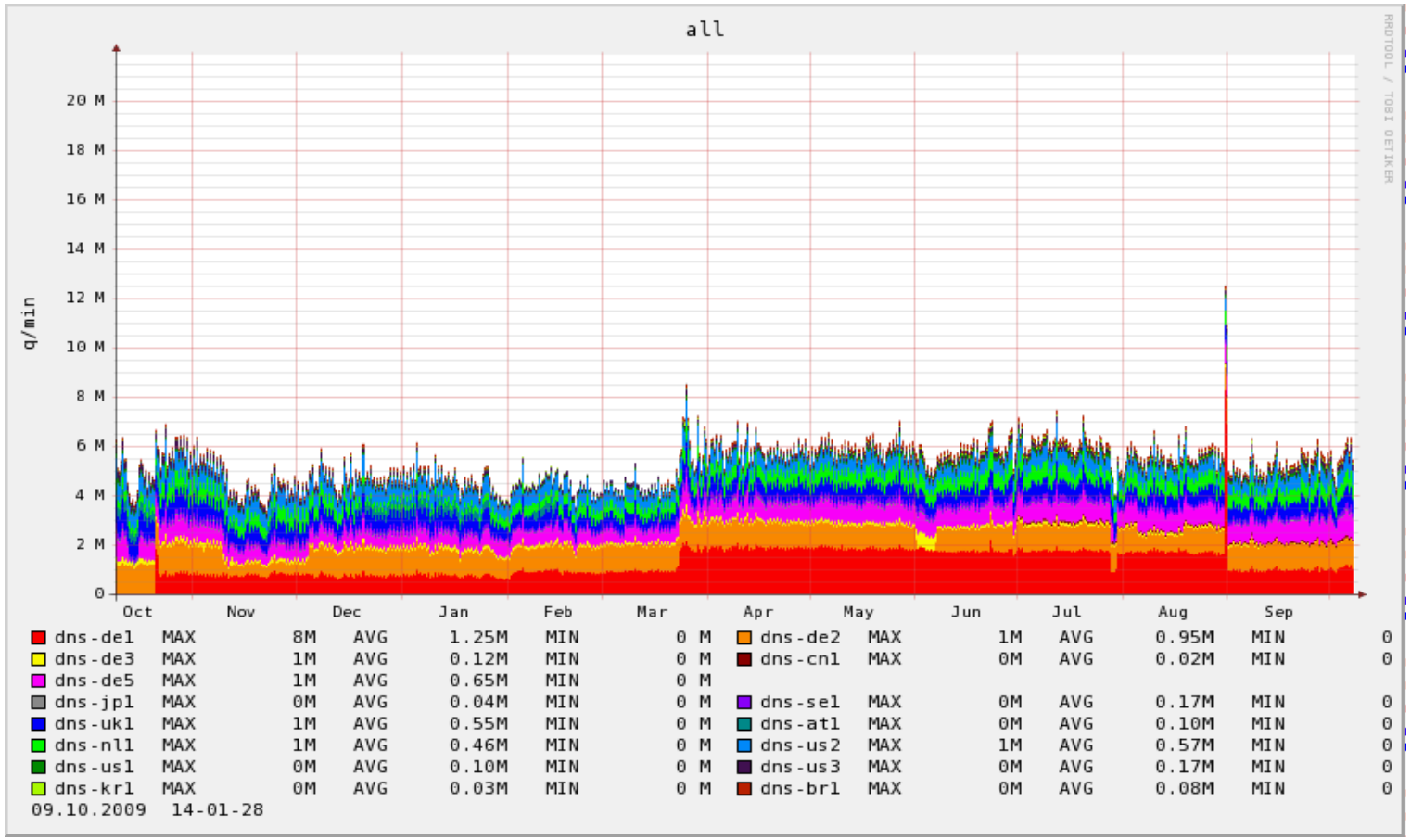


Jan./Aug 2008: Peaks up to 140k q/s

Avg. query load last year: 90k q/s

The old environment

.de query load 2009



Since March 2009 at about 100k q/s

The new setup

New concept:

- **No hardware diversification**
- **No OS diversification**
- **Running BIND 9 and NSD**
- **Using virtualisation**
- **Focusing on scalability**
- **Prepare for more frequent changes to zone data (based on IXFR)**

Initial scaling

- doubling of queries every 2-3 years
- estimated query load in 2011: 200k q/s
- targeted peak load ~15%
- load < 30%
- 16 locations
- 2-5 blades running nameservers per location

➔ max overall query load > 3 million q/sec, some 250 billion queries a day

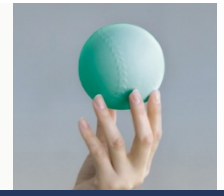
Initial setup

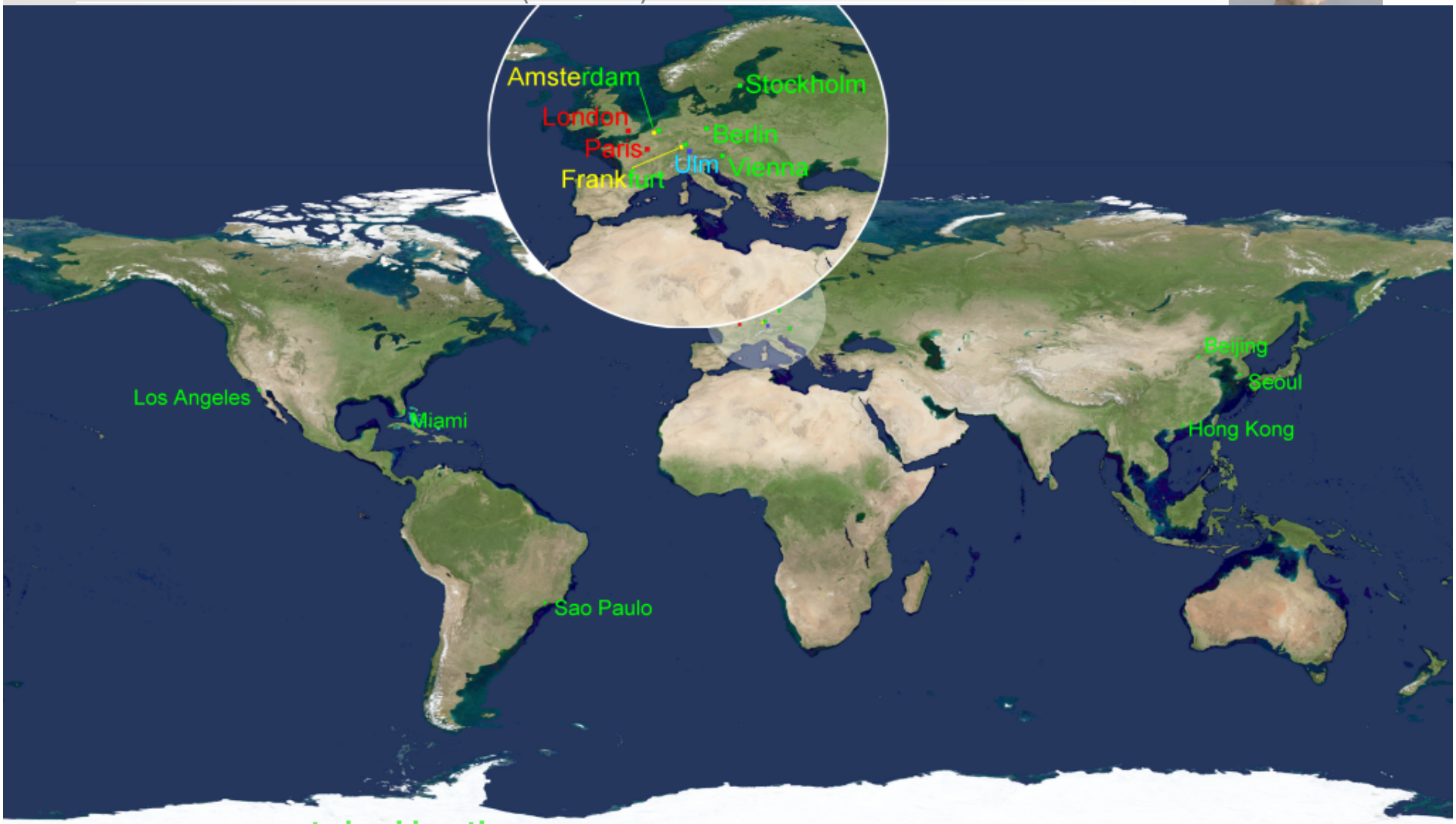
16 locations

- 2 IPv4 anycast clouds
- 1 IPv6 anycast cloud
- 2*2 locations with a provider internal anycast
- 1 unicast location
- additional test and spare setups

- 2 more locations to come this year

.de 2010: IPv4 - 16 locations at 10 exchanges ()
IPv6 - 7 locations ➔ 13 (June 2010)



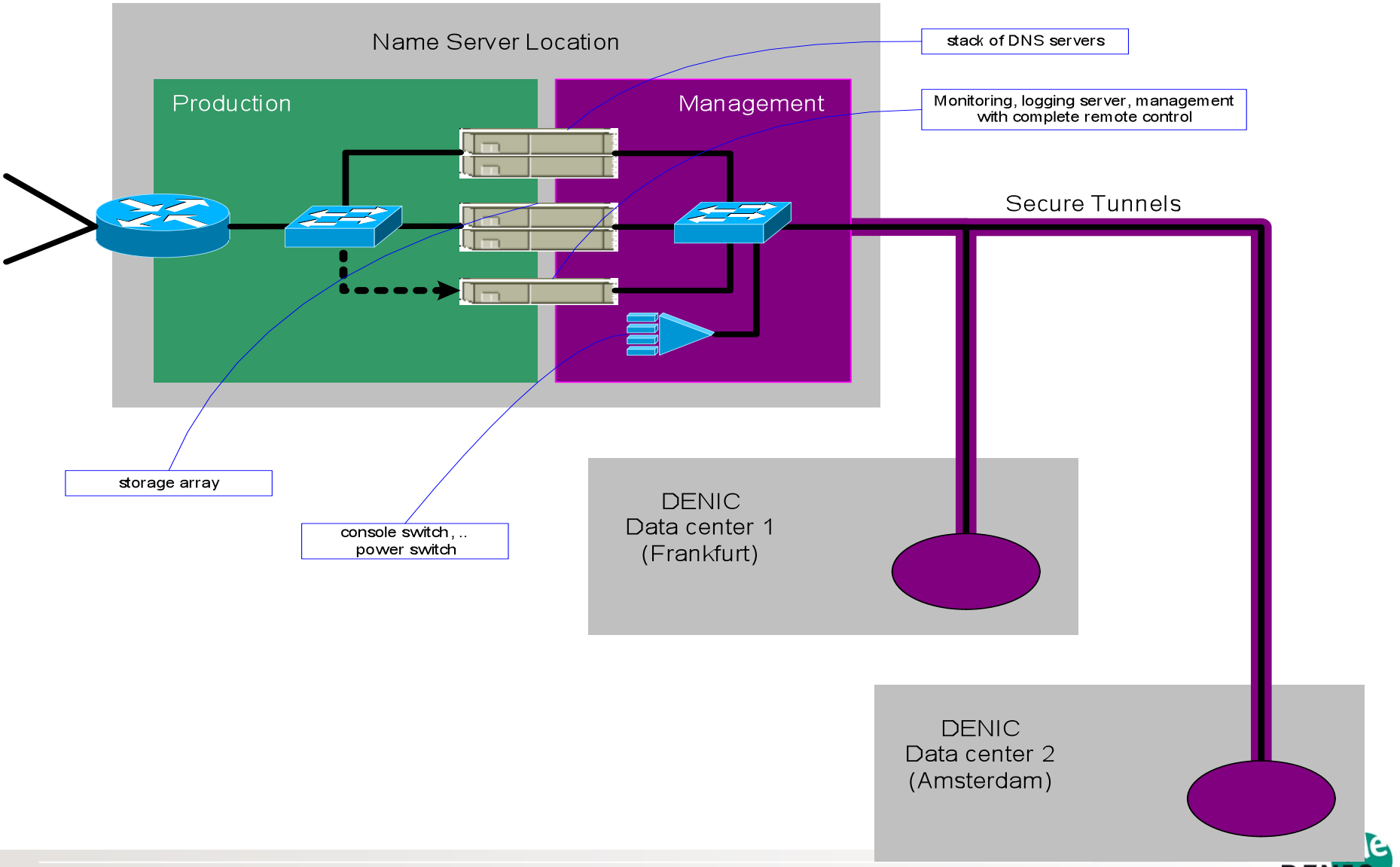


Legend:

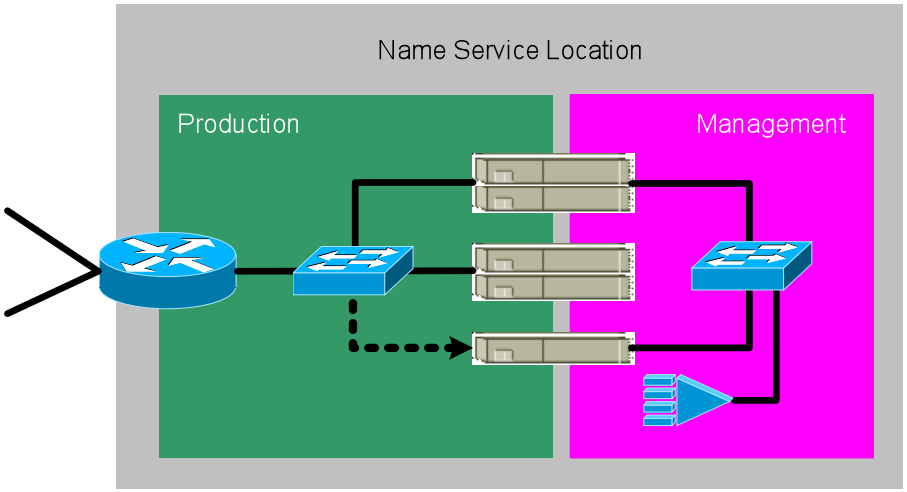
anycast cloud locations

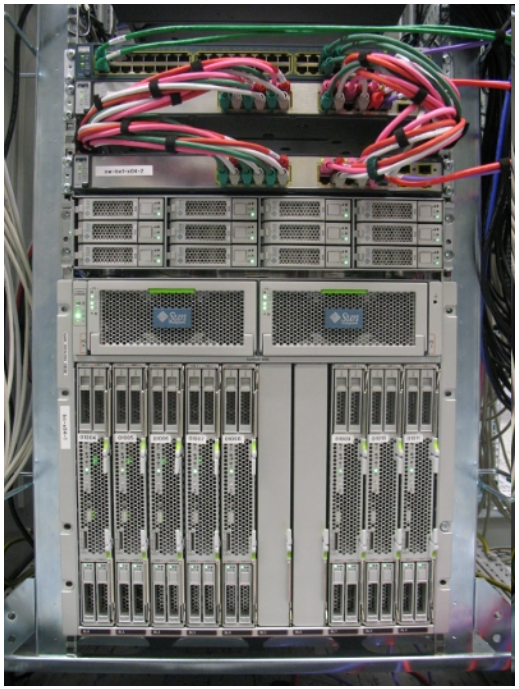
unicast locations, provider local anycast partners; dual locations uni-/anycast

Architecture (schematic)



Architecture (2)





Bladecenter: min. 4 Blades
(2x NS, 1x local Master/DSC/Logserver, 1x Spare)

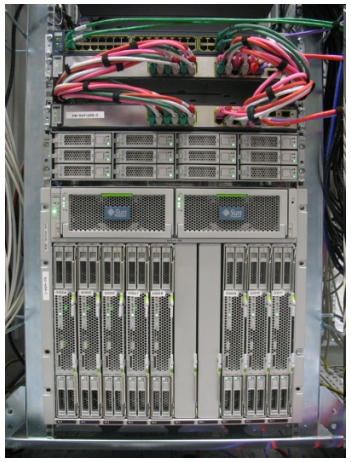
Blades: 2x QuadCore Opteron 2.5 GHz 32GB RAM

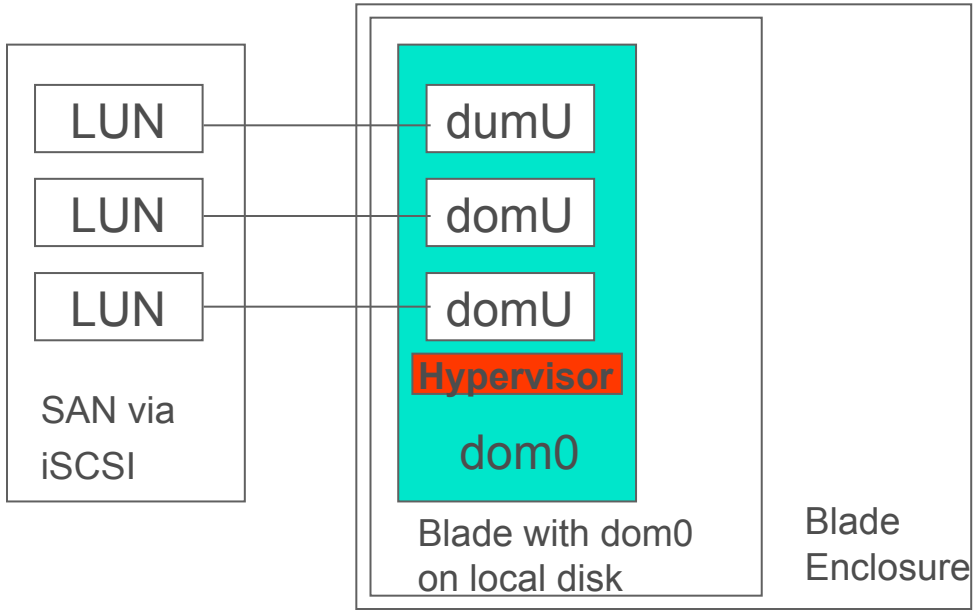
Storage: 1.5 TB iSCSI

Virtualisation: XEN

Nameserver Software: BIND, NSD

Architecture (3)





- Each application in its own virtual instance (domU)
- Highly scalable
- Easy rollout of new software
- Automatic installation and configuration
- Centralized administration

Deployment

Zone deployment now:

- **Creating a new zone every 2 hours**
- **Loading it to the central master**
- **Creating the update journals**
- **Automated distribution via local master to the nameserver using in-band mechanisms (Notify + IXFR)**

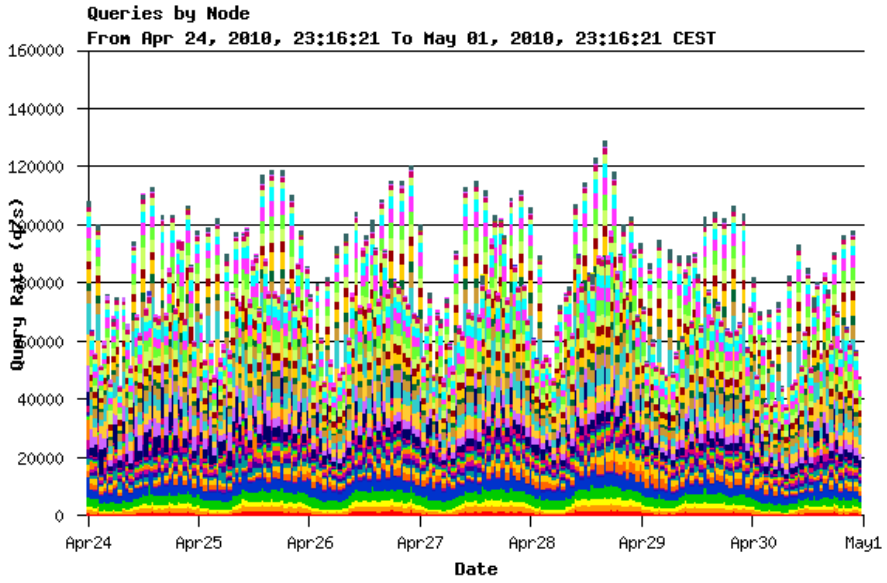
Plans for the future:

- **Dynamic zone generation and update**

Monitoring

Internal monitoring

- Nagios host- and service-checks
- DSC
- Centralized system logging
- Incoming traffic stored up to 36h



External monitoring

- Nagios service-checks
- DNSMON

Questions

?

Christoph Strabel
strabel@denic.de
+49 69 27235 -0