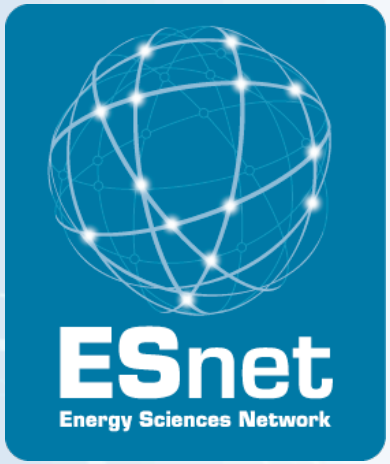


ESCC Evening Discussion: High Performance Data Transfer

Eli Dart, Network Engineer

ESnet Network Engineering Group

Summer ESCC/Joint Techs

Columbus, OH

July 14, 2010

High Performance Data Transfer

- This is a broad topic with several components
- Last discussion – perfSONAR
- Tonight's discussion – possible next steps to increase the performance achieved by scientists in their use of the network
  - What is "the network" from a user's perspective?
  - Deployment of dedicated systems
  - Need for cooperation
  - Do the easy stuff first

Data Transfer Is A Tractable Problem

- Data sets are, in a great many cases, less than 10TB in size, and many are less than 1TB.
- A well-configured infrastructure can do this quickly and easily
  - I can transfer a 10GB file from lbl-diskpt1 to the Data Transfer Nodes at NERSC and ORNL in less than a minute (1.6Gbps disk to disk at last check to NERSC)
  - 1TB/day means 100Mbps for 24 hours, or 300Mbps for 8 hours (see http://fasterdata.es.net/BandwidthRequirements.html)
  - This means that a 1Gbps data transfer host should be able to meet the needs of a great many scientists

Many Users Have Difficulty

- Data transfer is tractable…right…OK…then why is it so hard?

- One reason it is hard for many scientists is that the scientists are asked to integrate their own data transfer infrastructure
  - They don't have the funding for systems support
  - They don't have funding for infrastructure
  - "Why would I buy a new switch? $50k is another postdoc!"
- There is also no guarantee that the people "at the other end" have a well-configured infrastructure – why build my own?
- If the barriers to use of the network are too high, it doesn't matter if ESnet or the ESnet sites build cutting edge network infrastructures

What Is “The Network” Anyway?

- From the perspective of a user, “The Network” is not a bunch of routers, switches, fiber, and so on
- “The Network” is the thing that is broken when remote data transfers are hard (It’s cute and all, but what does this really say?)
- The primary user interface to “the network” is a data transfer tool or other user-agent application
- ***This means that without well-configured end systems, the fast networks at the sites and the ESnet backbone do not exist as effective scientific tools***
- Therefore, the utility of the network depends on the existence and availability of well-configured end systems

Next Steps: Pre-Deploy Dedicated Systems

- Several sites have deployed dedicated data transfer systems
- Large physics experiments (BaBar, LHC, RHIC, Tevatron, etc) already do this
- Recent success story – Fusion
  - Two systems – one at GA and one at EAST in China
  - Data transfers now keep up with instrument duty cycle
- Additional success stories – Data Transfer Nodes
  - ALCF, NERSC, OLCF
  - Dedicated hosts with access to shared global filesystems

Where To Deploy Dedicated Systems?

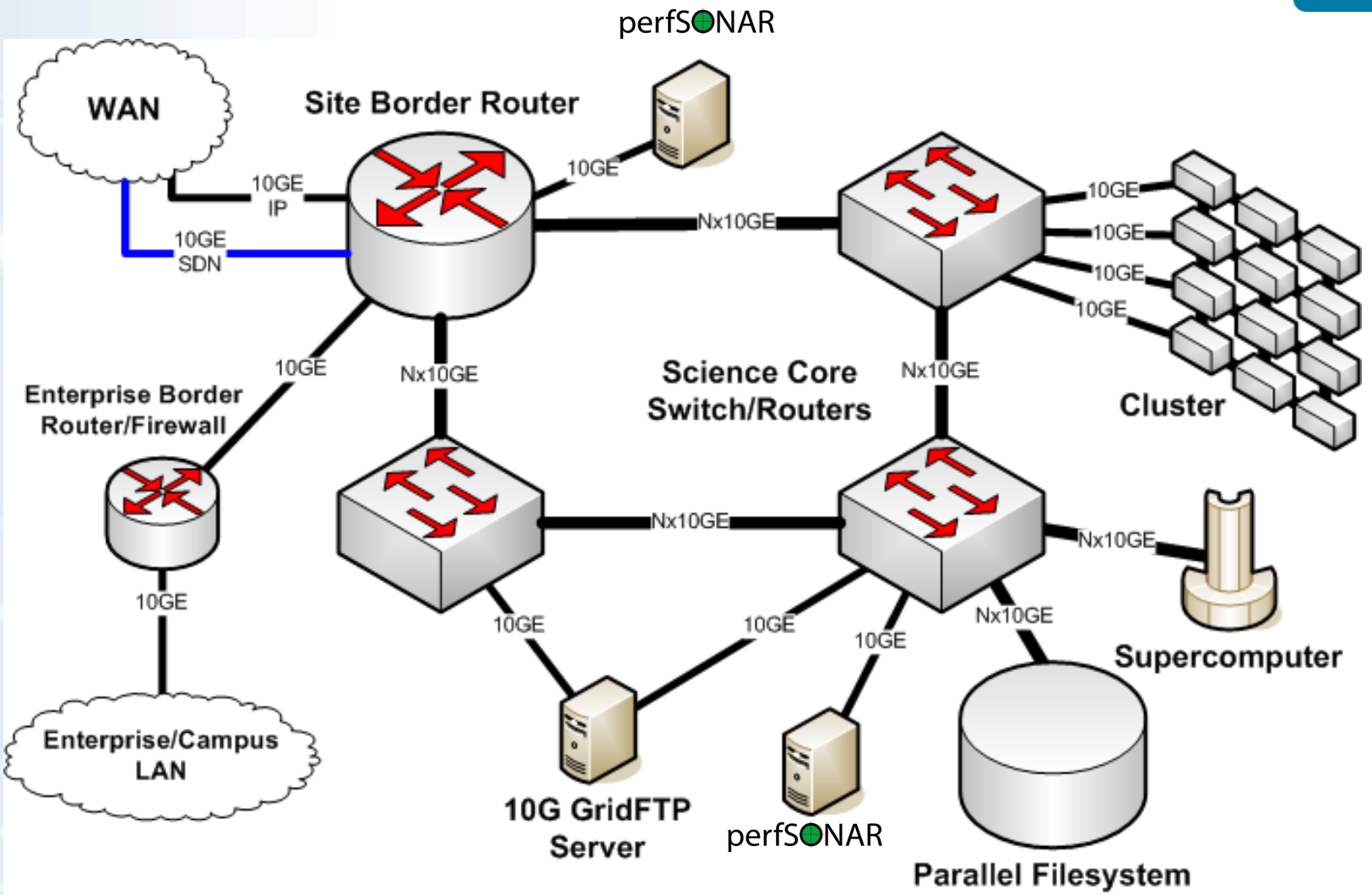
- Clearly dependent on network architecture – Know Your Network
- However, we have seen significant performance benefits when data transfer systems are moved near the site perimeter
- A DMZ network holds the external-facing servers that provide service to the Internet (e.g. DNS, Mail, Web)
- A “Science DMZ” could attach high-performance data servers to the site border router
  - This can be done with dark fiber if you’ve got the fiber – no need to move the machines to a different building, etc.
  - No need to drag large wide area data flows through the site network or the site firewall

We're All In This Together

- It is our collective job to support science
- Science is increasingly data-intensive
- Scope of collaboration is regional to global
- Therefore, science requires data movement, now and into the future
- Our customers cannot succeed unless we work together
  - Well-configured end systems and high performance networks are both necessary
  - Neither is a solution in itself
- ESnet can help with design, troubleshooting, etc., both for sites and for scientists

Discussion Topics

- Are there obvious places to put dedicated systems?
- Pick the low-hanging fruit first
- Science DMZs – issues for deployment
  - Firewalls (do we have to have firewalls if there are no windows clients?)
  - If the site security policy treats the boxes on the Science DMZ as external, does that help?
  - Funding silos, territory/influence concerns, etc
- http://fasterdata.es.net/


perfSONAR
WAN
Site Border Router
10GE
IP
10GE
SDN
10GE
Nx10GE
10GE
10GE
10GE
10GE
Enterprise Border
Router/Firewall
10GE
Nx10GE
Science Core
Switch/Routers
Nx10GE
Cluster
Nx10GE
10GE
Nx10GE
Supercomputer
10GE
10GE
10GE
Nx10GE
Enterprise/Campus
LAN
10G GridFTP
Server
perfSONAR
Parallel Filesystem

Questions?

Thanks!

Questions?

Thanks!