High Performance Networking for Grid Applications

www.science.uva.nl/~delaat

Cees de Laat

Faculty of Science
High Performance Networking for Grid Applications

www.science.uva.nl/~deaat

Cees de Laat

EU

SURFnet

University of Amsterdam
Contents of this talk

This slide is intentionally left blank
I am not going to talk about this here, it feels like preaching to the pope.
A. Lightweight users, browsing, mailing, home use
   Need full Internet routing, one to many

B. Business applications, multicast, streaming, VPN’s, mostly LAN
   Need VPN services and full Internet routing, several to several + uplink

C. Special scientific applications, computing, data grids, virtual-presence
   Need very fat pipes, limited multiple Virtual Organizations, few to few
So what are the facts

- **Costs of fat pipes (fibers) are one/third of equipment to light them up**
  - Is what Lambda salesmen tell me

- **Costs of optical equipment 10% of switching 10 % of full routing equipment for same throughput**
  - 100 Byte packet @ 40 Gb/s -> 20 ns to look up in 140 kEntries routing table (light speed from me to you!)

- **Big sciences need fat pipes**

- **Bottom line: create a hybrid architecture which serves all users in one consistent cost effective way**
The only formula’s

\[ \# \lambda \approx \frac{200 \ast e^{(t-2002)}}{rtt} \]

Now, as having been a High Energy Physicist we set
\( c = 1 \)
\( e = 1 \)
\( \hbar = 1 \)
and the formula reduces to:

\[ \# \lambda \approx \frac{200 \ast e^{(t-2002)}}{rtt} \]
## Services

<table>
<thead>
<tr>
<th>SCALE</th>
<th>CLASS</th>
<th>2 Metro</th>
<th>20 National/regional</th>
<th>200 World</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Switching/routing</td>
<td>Routing</td>
<td>ROUTER$</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>VPN’s, (G)MPLS</td>
<td>VPN’s Routing</td>
<td>Routing</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>dark fiber Optical switching</td>
<td>Lambda switching</td>
<td>Sub-lambdas, ethernet-sdh</td>
<td></td>
</tr>
</tbody>
</table>

\[ \# \lambda \approx \frac{200 \times e^{(r-2002)}}{rtt} \]
SURFnet Lambda’s fibers (old already) See Erik’s talk
Current technology + (re)definition

- Current (to me) available technology consists of SONET/SDH switches, 10 gig ethernet and dark fiber environments
- Optical switch on the way (customs)!
- DWDM+switching included
- Starlight/NetherLight deploy VLAN’s on Ethernet switches to connect [exactly two] ports (but also routing)
- We want to understand routerless limited environments
- So redefine a $\lambda$ as:

  “a $\lambda$ is a pipe where you can inspect packets as they enter and when they exit, but principally not when in transit. In transit one only deals with the parameters of the pipe: number, color, bandwidth”
Architectures - L1 - L3

Long haul $\lambda$

Bring plumbing to the users, not just create sinks in the middle of nowhere
How low can you go?
• **lambda for high bandwidth applications**
  - Bypass of production network
  - Middleware may request (optical) pipe

• **RATIONALE:**
  - Lower the cost of transport per packet
Distributed L2
SURFnet backbone

Lambda’s to:
- Chicago,
- Geneve,
- Praha,
- NYC
- London

Dark fiber to Dwingeloo

NetherLight

15454

calien

10 Gbs

1 Gbs

DAS: 32*2cpu's IBM Myrinet

1 Gbs

100Mbs

4 HP servers

Fat pc

UvA/NikHEF/SARA

(14 of 15)
Problem Solving Environment

Applications and Supporting Tools

Application Development Support

Collective Grid Services
- Brokering
- Global Queuing
- Co-Scheduling
- Data Cataloguing
- Auditing
- Authorization
- Monitoring
- Fault Management

Common Grid Services
- Grid Information Service
- Uniform Resource Access
- Global Event Services
- Uniform Data Access
- Data Replication
- Communication Services

Grid Security Infrastructure (authentication, proxy, secure transport)

Communication

Fabric
- Grid task initiation

Local Resources
- Resource Manager
- CPUs
- Monitoring
- On-Line Storage
- Scientific Instruments
- Tertiary Storage
- Highspeed Data Transport

layers of increasing abstraction taxonomy

High performance computing and Processor memory co-allocation
Security and Generic AAA

Optical Networking

Researched in other program lines
Imported from the Globus toolkit
Starting point

Generic AAA server
Rule based engine

API

Application Specific Module

Service
Accounting Metering

Policy
Data

PEP

PDP

RFC 2903 - 2906, 3334, policy draft
(Future) Projects

• National:
  • NCF Grid project
  • VLE
  • GigaPort-NG
  • LOFAR

• European
  • DataGrid
  • DataTAG

• International
  • NetherLight
  • StarLight
  • AnyLight, LowLight, BackLight
  • Optiputer

Research:
Models of Lambda networking
Transport
AAA
Transport in the corners

BW*RTT

Needs more App & Middleware interaction

For what current Internet was designed

Full optical future

# FLOWS
The END

Thanks to
SURFnet: Kees Neggers
UIC&iCAIR: Tom DeFanti, Joel Mambretti
CANARIE: Bill St. Arnaud

This work is supported by:
SURFnet
EU-IST project DATATAG