• A showcase of applications that are “early adopters” of very-high-bandwidth national and international networks
  – What can you do with a 10Gbps network?
  – What applications have insatiable bandwidth appetites?
• Scientists and technologists to optimally utilize 10Gbps experimental networks, with special emphasis on e-Science, Grid and Virtual Laboratory applications
• Registration is open (www.igrid2002.org)
• iGrid is not just a conference/demonstration event, it is also a testbed!!
• Contact
  – maxine@startup.net or deLaat@science.uva.nl
The Road towards Optical Networking

www.science.uva.nl/~delaat

Cees de Laat

EU SURFnet
University of Amsterdam
What is this buzz about optical networking

- Networks are already optical for ages
- Users won’t see the light
- Almost all current projects are about SONET circuits and Ethernet (old wine in new bags?)
- Are we going back to the telecom world, do NRN’s want to become telco’s
- Does it scale
- Is it all about speed or is it integrated services
The term VLBI is easily capable of generating many Gb of data per day. The sensitivity of the VLBI array scales with the data-rate (3/15) and there is a strong push to increase rates of 8Gb/s or more are entirely feasible under development. It is expected that parallel correlator will remain the most efficient approach, with distributed processing may have an application. Multi-gigabit data streams will aggregate into larger computers and the capacity of the final link to the database.
Know the user

A -> Lightweight users, browsing, mailing, home use
B -> Business applications, multicast, streaming
C -> Special scientific applications, computing, data grids, virtual-presence
So, what’s up doc

Suppose:

- Optical components get cheaper and cheaper
- Dark (well, dark?) fibers abundant
- Number of available \(\lambda/\text{user} \rightarrow \infty\)
- Speeds of 10, 100, 1000 Gbit/s make electrical domain packet handling physically difficult
  - 150 bytes @ 40 Gbit/s = 30 ns = 15 meter fiber
  - QoS makes no sense at these speeds
- Cost per packet forwarding lower at L1 / L2

Then:

- **Long term view ---\(\rightarrow\) full optical**
- \(\lambda\) provisioning for grid applications
- How low can you go
Optical networking, 3 scenarios

• **Lambdas for internal ISP bandwidth provisioning**
  – An ISP uses a lambda switching network to make better use of its (suppliers) dark fibers and to provision to the POP's. In this case the optical network is just within one domain and as such is a relatively simple case.

• **Lambda switching as peering point technology**
  – In this use case a layer 1 Internet exchange is build. ISP's peer by instantiating lambdas to each other. Is a $N*(N-1)$ and multi domain management problem.

• **Lambda switching as grid application bandwidth provisioning**
  – This is by far the most difficult since it needs UNI and NNI protocols to provision the optical paths through different domains.
Current technology + (re)definition

- Current (to me) available technology consists of SONET/SDH switches
- DWDM+switching coming up
- Starlight uses for the time being VLAN’s on Ethernet switches to connect [exactly] two ports
- So redefine a $\lambda$ as:

  “$\lambda$ 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”
- Lambda for high bandwidth applications
  - Bypass of production network
  - Middleware may request (optical) pipe
Other architectures - L1 - 3
Other architectures - Distributed virtual IEX’es

Problem: vlan tag distribution ==> gmpls
Distributed L2

Layer 2 VPN

Univ A

Univ B

SN5
CHICAGO

Univ X

SN5
A'DAM

Univ Y

lambda

Univ

2.5 Gbps SONET/SDH “Lambda”
10/100/1000 Mbps Ethernet

Amsterdam

L2/3

SurfNet5

Amsterdam 2nd phase

Almere

VLAN SARA

 AMSIX

 AMSIX

R

GIGA cluster

DASII

AMSIX

AMSIX

VLAN SARA

SARA
First experiences with SURFnet pure for research Lambda

- 2.5 Gbit SONET λ delivered dec 2001
  - Took about 3 months, should be 300 ms
- First generation equipment delivered nov 2001
- Back to back tests => OC12 limit -> 560 Mbit/s
- 1 unit shipped to Chicago (literally, took 3 weeks)
- End to end now 80 Mbit/s
- So, what is going on?
- Second generation equipment just delivered
- 1 unit shipped to Chicago (yes, is going to take 3 weeks)
slope = 100 ms
Sun Throughput [Mbit/s]
TCP is bursty due to sliding window protocol and slow start algorithm. So pick from menu:

- Flow control
- Traffic Shaping
- RED (Random Early Discard)
- Self clocking in TCP
- Deep memory

Window = BandWidth * RTT & BW = slow

Memory-at-bottleneck = fast - slow

\[ \text{Memory-at-bottleneck} = \frac{\text{fast}}{\text{RTT}} \times \text{slow} \times \text{RTT} \]
5000 1 kByte UDP packets
Memory = \frac{fast - slow}{fast} \times slow \times RTT

For RTT = 100 ms
Layer - 2 requirements from 3/4

Window = BandWidth * RTT & BW == slow

Memory-at-bottleneck = \[ \frac{\text{fast - slow}}{\text{fast}} \] * slow * RTT

Given M and f, solve for slow ===> 

\[ 0 = s^2 - f * s + \frac{f * M}{\text{RTT}} \]

\[ s_1, s_2 = \frac{f}{2} \left( 1 \pm \sqrt{1 - 4 \frac{M}{f * \text{RTT}}} \right) \]
Forbidden area, solutions for $s$ when $f = 1 \text{ Gb/s}$, $M = 0.5 \text{ Mbyte}$
AND NOT USING FLOWCONTROL

$\text{rtt} = 158 \text{ ms} = \text{RTT Amsterdam - Vancouver or Berkeley}$
Self-clocking of TCP

WS

L2
fast->slow

high RTT

L2
slow->fast

WS

14 µsec

20 µsec

20 µsec

20 µsec
Revisiting the truck of tapes

Consider one fiber

• Current technology allows for 320 \( \lambda \) in one of the frequency bands

• Each \( \lambda \) has a bandwidth of 40 Gbit/s

• Transport: \( 320 \times 40 \times 10^9 / 8 = 1600 \) GByte/sec

• Take a 10 metric ton truck

• One tape contains 50 Gbyte, weights 100 gr

• Truck contains \( (10000 / 0.1) \times 50 \) Gbyte = 5 PByte

• Truck / fiber = 5 PByte / 1600 GByte/sec = 3125 s \( \approx \) one hour

• For distances further away than a truck drives in one hour (50 km) minus loading and handling 100000 tapes the fiber wins!!!
iGrid 2002
The International Virtual Laboratory

www.igrid2002.org

24-26 September 2002
Amsterdam Science and Technology Centre (WTCW)
The Netherlands

• A showcase of applications that are “early adopters” of very-high-bandwidth national and international networks
  – What can you do with a 10Gbps network?
  – What applications have insatiable bandwidth appetites?
• Scientists and technologists to optimally utilize 10Gbps experimental networks, with special emphasis on e-Science, Grid and Virtual Laboratory applications
• Registration is open (www.igrid2002.org)
• iGrid is not just a conference/demonstration event, it is also a testbed!!
• Contact
  – maxine@startap.net      or      deLaat@science.uva.nl