Internet Innovation to support Science

Cees de Laat
Internet
From a network experiment that never ended (Vint Cerf)

- 1974: for the first time the word internet (RFC 675 - Specification of Internet Transmission Control Program) [note -> Open process!]
- 1981: the TCP/IP standard was ready to be adopted (RFC 791,792,793)

To a network for society

- 1989: WWW was born

- Jan 2011 → IANA IPv4 address space depleted! → Ipv6day.nl
# Internet is a Billion - Business!

<table>
<thead>
<tr>
<th>Company</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google</td>
<td>197</td>
</tr>
<tr>
<td>Amazon</td>
<td>83</td>
</tr>
<tr>
<td>Facebook</td>
<td>50</td>
</tr>
<tr>
<td>BAIDU</td>
<td>37</td>
</tr>
<tr>
<td>eBay</td>
<td>36</td>
</tr>
<tr>
<td>Yahoo</td>
<td>22</td>
</tr>
<tr>
<td>PriceLine</td>
<td>21</td>
</tr>
<tr>
<td>SalesForce</td>
<td>18</td>
</tr>
<tr>
<td>F5 Networks</td>
<td>11</td>
</tr>
<tr>
<td>CheckPoint</td>
<td>9</td>
</tr>
<tr>
<td>NetFlix</td>
<td>9</td>
</tr>
<tr>
<td>Expedia</td>
<td>7</td>
</tr>
</tbody>
</table>

- e.g.: Exxon Mobil, 368
- Apple Inc., 333

Monday 3 January 2011

1 miljard in 100$ biljetten
Internet developments

... more users!
... more data!
... more realtime!
GPU cards are disruptive!

- Fastest supercomputer in the world
- Nr. 500 supercomputer in the world
- 1 single Graphics Processing Unit

20,000,000$ | 7 year | 500$

Top 500

#1

#500
Data storage: doubling every 1.5 year!
Multiple colors / Fiber

Per fiber: ~ 80-100 colors * 50 GHz
Per color: 10 – 40 – 100 Gbit/s
BW * Distance ~ 2*10^{17} bm/s

New: Hollow Fiber!
⇒ less RTT!
Next Generation Wireless LAN Technology

802.11ac 1 Gbps throughput with WiFi

WiFi is one of the most preferred communication protocol LAN due to the easy comparison and convenience in the digital home. While consumer PC products has just started to migrate to a much higher bandwidth of 802.11n wireless LAN now working on next-generation standard definition is already in progress.
protocol LAN due to the easy comparison and convenience in the digital home. While consumer PC products has just started to migrate to a much higher bandwidth of 802.11n wireless LAN now working on next-generation standard definition is already in progress.
<table>
<thead>
<tr>
<th>Category</th>
<th>Speed</th>
<th>Volume</th>
<th>Deterministic</th>
<th>Real-time</th>
<th>Scalable</th>
<th>Secure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green-IT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Privacy/Trust</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authorization/policy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Programmable networks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-100Gig/TCP/WF/QoS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topology/Architecture</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optical Photonic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- IJkdijk/Urban Flood: X
- Medical: X
- LifeWatch: X
- CosmoGrid/eVLBI: X
- CineGrid: X
- EU-GN3/NOVI/Geyser: X
- SURFnet/GLIF/Cloud: X
<table>
<thead>
<tr>
<th></th>
<th>Jkdijk/Urban Flood</th>
<th>Medical</th>
<th>LifeWatch</th>
<th>EU-GN3/NOVI/Geyser</th>
<th>SURFnet/GLIF/Cloud</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green-IT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Privacy/Trust</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authorization/policy</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Programmable networks</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-100Gig/TCP/WF/QoS</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topology/Architecture</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Optical Photonic</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>
LOFAR as a Sensor Network

- LOFAR is a large distributed research infrastructure:
  - Astronomy:
    - >100 phased array stations
    - Combined in aperture synthesis array
    - 13,000 small “LF” antennas
    - 13,000 small “HF” tiles
  - Geophysics:
    - 18 vibration sensors per station
    - Infrasound detector per station
  - >20 Tbit/s generated digitally
  - >40 Tflop/s supercomputer
  - innovative software systems
    - new calibration approaches
    - full distributed control
    - VO and Grid integration
    - datamining and visualisation

20 flops/byte
2 Tflops/s
e - Very Large Base Interferometer
**Deadline for submitting observing proposals**

Program committee:
- rates proposals
- allocates observing time

VLBI Observing Session

- Disks shipped to JIVE
- Correlation at JIVE

Data shipped

Data arrives at scientist's desk!
eEVN: European VLBI Network

Data processing centre: Dwingeloo
16 Gbps (2005)

China
USA
South Africa
Russia

1-30 Gbps

star topology

Slide courtesy of Richard Schilizzi <schilizzi@jive.nl>
eEVN+: European VLBI Network

Dec 4

Deadline for submitting eVLBI observing proposals

Dec 5

Program committee decides if eVLBI science can be justified

eVLBI Observing Run

Correlation at JIVE

Dec 6

Scientist downloads data from www.jive.nl

Slide courtesy of Richard Schilizzi <schilizzi@jive.nl>
**The SCARIlE project**

**SCARIlE**: a research project to create a Software Correlator for e-VLBI.  
**VLBI Correlation**: signal processing technique to get high precision image from spatially distributed radio-telescope.

---

**Research:**

![Diagram](image.png)

- **Telescopes**
- **Input nodes**
- **Correlator nodes**
- **Output node**

---

**Figure 2.** Grid architecture that includes programmable network services.
CosmoGrid

- Motivation: previous simulations found >100 times more substructure than is observed!
- Simulate large structure formation in the Universe
  - Dark Energy (cosmological constant)
  - Dark Matter (particles)
- Method: Cosmological $N$-body code
- Computation: Intercontinental SuperComputer Grid
The hardware setup

- 2 supercomputers:
  - 1 in Amsterdam (60Tflops Power6 @ SARA)
  - 1 in Tokyo (30Tflops Cray XD0-4 @ CFCA)

- Both computers are connected via an intercontinental optical 10 Gbit/s network
<table>
<thead>
<tr>
<th>Topic</th>
<th>IJkdijk/Urban Flood</th>
<th>Medical</th>
<th>LifeWatch</th>
<th>CosmoGrid/eVLBI</th>
<th>EU-GN3/NOVI/Geyers</th>
<th>SURFnet/GLIF/Cloud</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green-IT</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Privacy/Trust</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authorization/policy</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Programmable networks</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-100Gig/TCP/WF/QoS</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topology/Architecture</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Optical Photonic</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

Where will it happen?
IJKDIJK

Sensors: 15000km* 800 bps/m ->12 Gbit/s to cover all Dutch dikes
Sensor grid: instrument the dikes
First controlled breach occurred on sept 27th ‘08:

Many small flows $\rightarrow$ 12 Gb/s
Tera-Thinking

- What constitutes a Tb/s network?
- think back to teraflop computing!
  - MPI turns a room full of pc’s in a teraflop machine
- massive parallel channels in hosts, NIC’s
- TeraApps programming model supported by
  - TFlops -> MPI / Globus / Cloud
  - TBytes -> DAIS / MONETdb …
  - TPixels -> SAGE
  - TSensors -> LOFAR, LHC, LOOKING, CineGrid, ...
  - Tbit/s -> ?
- ? -> Programmable Networks
User Programmable Virtualized Networks.

- The network is virtualized as a collection of resources.
- UPVNs enable network resources to be programmed as part of the application.
- Mathematica interacts with virtualized networks using UPVNs and optimize network + computation.

ref: Robert J. Meijer, Rudolf J. Strijkers, Leon Gommans, Cees de Laat, User Programmable Virtualized Networks, accepted for publication to the IEEE e-Science 2006 conference Amsterdam.
TouchTable Demonstration @ SC08
<table>
<thead>
<tr>
<th>Category</th>
<th>IJkdijk/Urban Flood</th>
<th>Medical LifeWatch</th>
<th>EU-GN3/NOVI/Geyser</th>
<th>SURFnet/GLIF/Cloud</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green-IT</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Privacy/Trust</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Authorization/policy</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Programmable networks</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>40-100Gig/TCP/WF/QoS</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Topology/Architecture</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Optical Photonic</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
ATLAS detector @ CERN Geneve
ATLAS detector @ CERN Geneve
One Heavy Ion Collision in Atlas!

Snapshot of a heavy ion collision directly from the ATLAS experiment.
LHC Data Grid Hierarchy
CMS as example, Atlas is similar

100000 flops/byte
10 Pflops/s

Status 2002!

Tier 0+1

Online System

~100 MBytes/sec

event reconstruction

~100000 flops/byte

10 Pflops/s

event simulation

CMS detector: 15m X 15m X 22m
12,500 tons, $700M.

Tier 1

Italian Regional Center

German Regional Center

NIKHEF Dutch Regional Center

FermiLab, USA Regional Center

Tier 2

~0.6-2.5 Gbps

Tier 3

~0.6-2.5 Gbps

Institute

Institute

Institute

Institute

Tier 2 Center

Tier 2 Center

Tier 2 Center

Tier 2 Center

Tier 4

Physics data cache

100 - 1000 Mbits/sec

Workstations

CERN/CMS data goes to 6-8 Tier 1 regional centers, and from each of these to 6-10 Tier 2 centers.

Physicists work on analysis “channels” at 135 institutes. Each institute has ~10 physicists working on one or more channels.

2000 physicists in 31 countries are involved in this 20-year experiment in which DOE is a major player.
Status 2011!

4 X 100G Trans-Atlantic
+ NY-CHI
+ AMS-GVA

Using 100G Links
+ Next-Gen.
Optical Muxes
Big and small flows don’t go well together on the same wire! 😞
A. Lightweight users, browsing, mailing, home use
   Need full Internet routing, one to all

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

C. E-Science applications, distributed data processing, all sorts of grids
   Need very fat pipes, limited multiple Virtual Organizations, P2P, few to few

For the Netherlands 2011
\[ \Sigma A = \Sigma B = \Sigma C \approx 1 \text{ Tb/s} \]
However:
- A -> all connects
- B -> on several
- C -> just a few (SP, LHC, LOFAR)

Ref: Cees de Laat, Erik Radius, Steven Wallace, "The Rationale of the Current Optical Networking Initiatives"
A. Lightweight users, browsing, mailing, home use
   Need full Internet routing, one to all

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

C. E-Science applications, distributed data processing, all sorts of grids
   Need very fat pipes, limited multiple Virtual Organizations, P2P, few to few

For the Netherlands 2011
\[ \Sigma A = \Sigma B = \Sigma C \approx 1 \text{ Tb/s} \]

However:
• A -> all connects
• B -> on several
• C -> just a few (SP, LHC, LOFAR)

Ref: Cees de Laat, Erik Radius, Steven Wallace, "The Rationale of the Current Optical Networking Initiatives"
Towards Hybrid Networking!

- Costs of photonic equipment 10% of switching 10% of full routing
  - for same throughput!
  - Photonic vs Optical (optical used for SONET, etc, 10-50 k$/port)
  - DWDM lasers for long reach expensive, 10-50 k$

- Bottom line: look for a hybrid architecture which serves all classes in a cost effective way
  - map A -> L3, B -> L2, C -> L1 and L2

- Give each packet in the network the service it needs, but no more!

L1 ≈ 2-3 k$/port

L2 ≈ 5-8 k$/port

L3 ≈ 75+ k$/port
How low can you go?

Application Endpoint A

Router
Ethernet
SONET
DWDM
Fiber

Local Ethernet
MEMS
Regional dark fiber
POS
15454
6500
HDXc
Trans-Oceanic

Application Endpoint B

CERN
NetherLight
GLIF
StarLight

10454
6500
HDXc
Trans-Oceanic
In The Netherlands SURFnet connects between 180:
- universities;
- academic hospitals;
- most polytechnics;
- research centers.
with an indirect ~750K user base

~ 8860 km scale comparable to railway system
Diagram for SAGE video streaming to ATS

Lab 10, Nortel

SAGE Display
SAGE Servers
MERS

Netherlight Canarie

Internet
Content Choice

User
Regular Browser

UvA, Amsterdam

comp clusters

Traffic Generators

Content Portal
Streaming Server
100 TB Storage

Content Request
PBT is **SIMPLE** and **EFFECTIVE** technology to build a shared Media-Ready Network
Alien light
From idea to realisation!

40Gb/s alien wavelength transmission via a multi-vendor 10Gb/s DWDM infrastructure

Alien wavelength advantages
- Direct connection of customer equipment\(^1\)
  \(\rightarrow\) cost savings
- Avoid OEO regeneration \(\rightarrow\) power savings
- Faster time to service\(^2\) \(\rightarrow\) time savings
- Support of different modulation formats\(^3\)
  \(\rightarrow\) extend network lifetime

Alien wavelength challenges
- Complex end-to-end optical path engineering in terms of linear (i.e. OSNR, dispersion) and non-linear (FWM, SPM, XPM, Raman) transmission effects for different modulation formats.
- Complex interoperability testing.
- End-to-end monitoring, fault isolation and resolution.
- End-to-end service activation.

In this demonstration we will investigate the performance of a 40Gb/s PM-QPSK alien wavelength installed on a 10Gb/s DWDM infrastructure.

New method to present fiber link quality, FoM (Figure of Merit)
In order to quantify optical link grade, we propose a new method of representing system quality: the FOM (Figure of Merit) for concatenated fiber spans.

\[
FOM = \sum_{i=1}^{N} \frac{L_i}{10}
\]

A

B

C

Easy-to-use formula that accurately quantifies transmission system performance

Transmission system setup
JOINT SURFnet/NORDUnet 40Gb/s PM-QPSK alien wavelength DEMONSTRATION.

Test results
Error-free transmission for 23 hours, 17 minutes \(\rightarrow\) BER \(\times 3.0 \times 10^{-14}\)

Conclusions
- We have investigated experimentally the all-optical transmission of a 40Gb/s PM-QPSK alien wavelength via a concatenated native and third party DWDM system that both were carrying live 10Gb/s wavelengths.
- The end-to-end transmission system consisted of 1056 km of TWRS (TrueWave Reduced Slope) transmission fiber.
- We demonstrated error-free transmission (i.e. BER below 10^{-15}) during a 23 hour period.
- More detailed system performance analysis will be presented in an upcoming paper.

REFERENCES

ACKNOWLEDGMENTS
We are grateful to NORDUnet for providing us with bandwidth on their fiber for use in this experiment and also for their support and assistance during the experiments. We also acknowledge Teleco and Nortel for their infrastructure work and simulation support.
Alien light From idea to realisation!

40Gb/s alien wavelength transmission via a multi-vendor 10Gb/s DWDM infrastructure

Alien wavelength advantages
- Direct connection of customer equipment \(^{(1)}\) ➔ cost savings
- Avoid OEO regeneration ➔ power savings
- Faster time to service \(^{(2)}\) ➔ time savings
- Support of different modulation formats \(^{(3)}\) ➔ extend network lifetime

Transmission system setup
JOINT SURFnet/NORDUnet 40Gb/s PM-QPSK alien wavelength DEMONSTRATION.

Test results
Error-free transmission for 23 hours, 17 minutes ➔ BER × 10\(^{-18}\)

Conclusions
- We have investigated experimentally the all-optical transmission of a 40Gb/s PM-QPSK alien wavelength via a concatenated native and third party DWDM system that both were carrying live 10Gb/s wavelengths.
- The end-to-end transmission system consisted of 1056 km of TWRS (TrueWave Reduced Slope) transmission fiber.
- We demonstrated error-free transmission (i.e. BER below 10\(^{-15}\)) during a 23 hour period.
- More detailed system performance analysis will be presented in an upcoming paper.

References
[1] "OPERATIONAL SOLUTIONS FOR AN ENTERPRISE LAYER", C. GERSTE et al. in C'2009
[2] "NEXT OPTICAL TRANSPORT NETWORKS: MANAGING DUAL OPTICAL OPTIONS" 
[4] "NORDUNET OPTICAL BACKHAUL COMMUNICATION"
We are grateful to NORDUNET for providing us with bandwidth on their network for this experiment and also for their support and assistance during the experiments. We also acknowledge TELEFÓNICA and NORTTEL for their infrastructure work and simulation support.
ClearStream @ TNC2011

Setup codename: FlightCees

UvA
- iPerf
- I7 3.2 GHz Q-core
- Amd Ph II 3.6 GHz HexC
- Mellanox
- 40G E

Copenhagen
- iPerf
- 2* dual 2.8 GHz Q-core
- Mellanox

CERN
- CIENA OME 6500
- CIENA DWDM
- LH
- 17 ms RTT

Hamburg
- CIENA OME 6500
- CIENA DWDM
- LH
- 27 ms RTT

Amsterdam – Geneva (CERN) – Copenhagen – 4400 km (2700 km alien light)
Demo setup codename: FlightCees

- Ciena ActiveFlex(OME) 6500
- Broadcom 40GE 18 port L2 Ethernet Switch
- Supermicro Intel Server
- Dell R815 Server
Visit CIENA Booth
surf to http://tnc11.delaat.net

ClearStream
End-to-End Ultra Fast Transmission Over a Wide Area 40 Gbit/s Lambda

Amsterdam (UvA) Live RX Traffic

Copenhagen POP RX Traffic

Incoming Amsterdam 25.5 Gbps
Incoming Copenhagen 20.97 Gbps
Total Throughput 46.47 Gbps RTT 44.032 ms
Results (rtt = 17 ms)

- **Single flow iPerf 1 core** -> 21 Gbps
- **Single flow iPerf 1 core** <-> -> 15+15 Gbps
- **Multi flow iPerf 2 cores** -> 25 Gbps
- **Multi flow iPerf 2 cores** <-> -> 23+23 Gbps
- **DiViNe** <-> -> 11 Gbps
- **Multi flow iPerf + DiVine** -> 35 Gbps
- **Multi flow iPerf + DiVine** <-> -> 35 + 35 Gbps
Performance Explained

• Mellanox 40GE card is PCI-E 2.0 8x (5GT/s)
• 40Gbit/s raw throughput but …. 
• PCI-E is a network-like protocol
  – 8/10 bit encoding -> 25% overhead -> 32Gbit/s maximum data throughput
  – Routing information
• Extra overhead from IP/Ethernet framing
• Server architecture matters!
  – 4P system performed worse in multithreaded iperf
Server Architecture

DELL R815
4 x AMD Opteron 6100

Supermicro X8DTT-HIBQF
2 x Intel Xeon
CPU Topology benchmark

We used numactl to bind iperf to cores
We investigate: for complex networks!
LinkedIN for Infrastructure

- From semantic Web / Resource Description Framework.
- The RDF uses XML as an interchange syntax.
- Data is described by triplets (Friend of a Friend):

```
Object   Predicate   Subject

Subject Object   Subject

Object Subject   Object Subject

Object Subject   Object Subject

Object Subject   Object Subject

Location

Device

Interface

Link

name

description

locatedAt

hasInterface

connectedTo

capacity

encodingType

encodingLabel
```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:ndl="http://www.science.uva.nl/research/air/ndl#">
    <!-- Description of Netherlight -->
    <ndl:Location rdf:about="#Netherlight">
        <ndl:name>Netherlight Optical Exchange</ndl:name>
    </ndl:Location>
    <!-- TDM3.amsterdam1.netherlight.net -->
    <ndl:Device rdf:about="#tdm3.amsterdam1.netherlight.net">
        <ndl:name>tdm3.amsterdam1.netherlight.net</ndl:name>
        <ndl:locatedAt rdf:resource="#amsterdam1.netherlight.net"/>
        <ndl:hasInterface rdf:resource="#tdm3.amsterdam1.netherlight.net:501/1"/>
        <ndl:hasInterface rdf:resource="#tdm3.amsterdam1.netherlight.net:501/3"/>
        <ndl:hasInterface rdf:resource="#tdm3.amsterdam1.netherlight.net:501/4"/>
        <ndl:hasInterface rdf:resource="#tdm3.amsterdam1.netherlight.net:503/1"/>
        <ndl:hasInterface rdf:resource="#tdm3.amsterdam1.netherlight.net:503/2"/>
        <ndl:hasInterface rdf:resource="#tdm3.amsterdam1.netherlight.net:503/3"/>
        <ndl:hasInterface rdf:resource="#tdm3.amsterdam1.netherlight.net:504/1"/>
        <ndl:hasInterface rdf:resource="#tdm3.amsterdam1.netherlight.net:504/2"/>
        <ndl:hasInterface rdf:resource="#tdm3.amsterdam1.netherlight.net:504/3"/>
        <ndl:hasInterface rdf:resource="#tdm3.amsterdam1.netherlight.net:504/4"/>
        <ndl:hasInterface rdf:resource="#tdm3.amsterdam1.netherlight.net:501/2"/>
    </ndl:Device>
    <!-- all the interfaces of TDM3.amsterdam1.netherlight.net -->
    <ndl:Interface rdf:about="#tdm3.amsterdam1.netherlight.net:501/1">
        <ndl:name>tdm3.amsterdam1.netherlight.net:POS501/1</ndl:name>
        <ndl:connectedTo rdf:resource="#tdm4.amsterdam1.netherlight.net:5/1"/>
    </ndl:Interface>
    <ndl:Interface rdf:about="#tdm3.amsterdam1.netherlight.net:501/2">
        <ndl:name>tdm3.amsterdam1.netherlight.net:POS501/2</ndl:name>
        <ndl:connectedTo rdf:resource="#tdm1.amsterdam1.netherlight.net:12/1"/>
    </ndl:Interface>
</rdf:RDF>
Multi-layer descriptions in NDL

IP layer

Ethernet layer

STS layer

OC-192 layer

SONET switch with Ethernet intf.

End host

Université du Quebec

CA Net Canada

StarLight Chicago

Ethernet & SONET switch

SONET switch

SONET switch with Ethernet intf.

End host

Universiteit van Amsterdam

NetherLight Amsterdam

MAN LAN New York

New York

Universiteit van Amsterdam

Université du Quebec

CA Net Canada

StarLight Chicago

Ethernet & SONET switch

SONET switch

SONET switch with Ethernet intf.

End host

Universiteit van Amsterdam

NetherLight Amsterdam

MAN LAN New York

New York

Université du Quebec

CA Net Canada

StarLight Chicago

Ethernet & SONET switch

SONET switch

SONET switch with Ethernet intf.

End host
Path between interfaces A1 and E1:
A1-A2-B1-B4-D4-D2-C3-C4-C1-C2-B2-B3-D3-D1-E2-E1

Scaling: Combinatorial problem
A weird example

- Université du Quebec
- CA★Net Canada
- MAN LAN New York
- Universiteit van Amsterdam
- StarLight Chicago
- NetherLight Amsterdam

Connecting lines:
- Gigabit Ethernet
- OC-192 (22 free)
- 2x OC-192 (87 free)
- 2x OC-192 (63 free)

Adaptation options:
- can adapt GE in STS-24c
- can adapt GE in STS-3c-7v

Thanks to Freek Dijkstra & team!
The Problem

I want HC and AB
Success depends on the order
Wouldn’t it be nice if I could request [HC, AB, ...]
Another one 😊

I want AB and CD
Success does not even depend on the order!!!
Virtualisatie van infrastructuur & QoS
RDF describing Infrastructure

Application: find video containing x, then trans-code to it view on Tiled Display

RDF/CG

RDF/CG

RDF/ST

RDF/NDL

RDF/NDL

RDF/VIZ

RDF/CPU

content

database

PG&CdL
Applications and Networks become aware of each other!
CineGrid portal

Cache & Store & Forward

100 Tbyte

CineGrid Amsterdam

Welcome to the Amsterdam CineGrid distribution node. Below are the latest additions of super-high-quality video to our node.

For more information about CineGrid and our efforts look at the about section.

Latest Additions

Wypke

Available format: 4k dci (4.6 KB)
Duration: 1 hour and 8 minutes
Created: 1-week, 2 days ago
Author: Wypke
Categories:

Prague Train

Steam locomotive in Prague.

Available format: 4k dci (3.9 KB)
Duration: 27 hours and 46 minutes
Created: 1-week, 2 days ago
Author: CineGrid
Categories: dclis, prague train

VLC: Big Buck Bunny

Available format: 1080p MPEG4 (1.1 GB)
Duration: 1 hour and 9 minutes
Created: 1 month, 1 week ago
Author: Blender Foundation
Categories: animation, blender bunny, cg
CineGrid Workflow Planner
### Green-IT
- IJkdijik/Urban Flood
- Medical LifeWatch
- EU-GN3/NOVI/Geyser
- SURFnet/GLIF/Cloud

### Privacy/Trust
- IJkdijik/Urban Flood
- Medical LifeWatch
- EU-GN3/NOVI/Geyser
- SURFnet/GLIF/Cloud

### Authorization/policy
- IJkdijik/Urban Flood
- Medical LifeWatch
- EU-GN3/NOVI/Geyser
- SURFnet/GLIF/Cloud

### Programmable networks
- IJkdijik/Urban Flood
- Medical LifeWatch
- EU-GN3/NOVI/Geyser
- SURFnet/GLIF/Cloud

### 40-100Gig/TCP/WF/QoS
- IJkdijik/Urban Flood
- Medical LifeWatch
- EU-GN3/NOVI/Geyser
- SURFnet/GLIF/Cloud

### Topology/Architecture
- IJkdijik/Urban Flood
- Medical LifeWatch
- EU-GN3/NOVI/Geyser
- SURFnet/GLIF/Cloud

### Optical Photonic
- IJkdijik/Urban Flood
- Medical LifeWatch
- EU-GN3/NOVI/Geyser
- SURFnet/GLIF/Cloud
Why is more resolution better?

1. More Resolution Allows Closer Viewing of Larger Image
2. Closer Viewing of Larger Image Increases Viewing Angle
3. Increased Viewing Angle Produces Stronger Emotional Response

- **UHDTV (8K)**
  - Resolution: 7680
  - Bandwidth: 24 Gb/s
  - Viewing Angle: 100°

- **HDTV (2K)**
  - Resolution: 1920
  - Bandwidth: 7.6 Gb/s
  - Viewing Angle: 60°

- **SHD (4K)**
  - Resolution: 3840
  - Bandwidth: 3.0 Gb/s
  - Viewing Angle: 30°
Red End
Robin Noorda & Bethany de Forest
Hey at still.

We're almost done. Sshh...
US and International OptIPortal Sites

USGS EDC

NCSA & TRECC

SIO

NCMIR

AIST

RINCON & Nortel

SARA

KISTI

UCI

TAMU

CALIT2

Real time, multiple 10 Gb/s
The “Dead Cat” demo

SC2004, Pittsburgh, Nov. 6 to 12, 2004
iGrid2005, San Diego, Sept. 2005

Many thanks to:
AMC
SARA
GigaPort
UvA/AIR
Silicon Graphics, Inc.
Zoölogisch Museum

Use AAA concept to split (time consuming) service authorization process from service access using secure tokens in order to allow fast service access.
Token Based networking
Service Provider Domain Group

National / Federal Banks

Service Provider Domain Group

MasterCard (210 countries)

How is trust built at global scale? Can these concepts be reused in the Internet as a selectable Infrastructure Service?

Service Providers

Issuing Banks

Merchant Banks

Service Transaction Trust Layer

Card Holders
1.6 Billion

22.4 Billion transactions / year

Merchants
29 Million

_regulations_

_membership rules_

_contracts_

_transaction_
Challenges

• Data – Data – Data
  – Archiving, publication, searchable, transport, self-describing, DB innovations needed, multi disciplinary use

• Virtualisation
  – Another layer of indeterminism

• Greening the Infrastructure
  – e.g. Department Of Less Energy: [http://www.ecrinitiative.org/pdfs/ECR_3_0_1.pdf](http://www.ecrinitiative.org/pdfs/ECR_3_0_1.pdf)

• Disruptive developments
  – BufferBloath, Revisiting TCP, influence of SSD’s & GPU’s
  – Multi layer Glif Open Exchange model
  – Invariants in LightPaths (been there done that 😊)
    • X25, ATM, SONET/SDH, Lambda’s, MPLS-TE, VLAN’s, PBT, OpenFlow, ….
  – Authorization & Trust & Security and Privacy
Hybrid Networking <-> Computing

Routers ↔ Supercomputers

Ethernet switches ↔ Grid & Cloud

Photonic transport ↔ GPU’s

What matters:

Energy consumption/multiplication
Energy consumption/bit transported
ECO-Scheduling
Questions?

CookReport
feb 2009 and feb-mar 2010

november ’08
interview with
Kees Neggers (SURFnet),
Cees de Laat (UvA)

and furthermore
on november ’09

Wim Liebrandt (SURF),
Bob Hertzberger (UvA) and
Hans Dijkman (UvA)

BSIK projects
GigaPort &
VL-e / e-Science

ext.delaat.net