Lambda-Grid developments

RDF/NDL, AAA and StarPlane

Cees de Laat

SURFnet

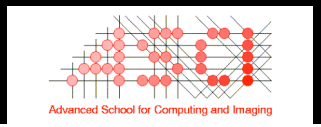
BSIK

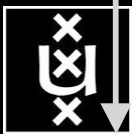
EU

University of Amsterdam

SARA

TNO
NCF

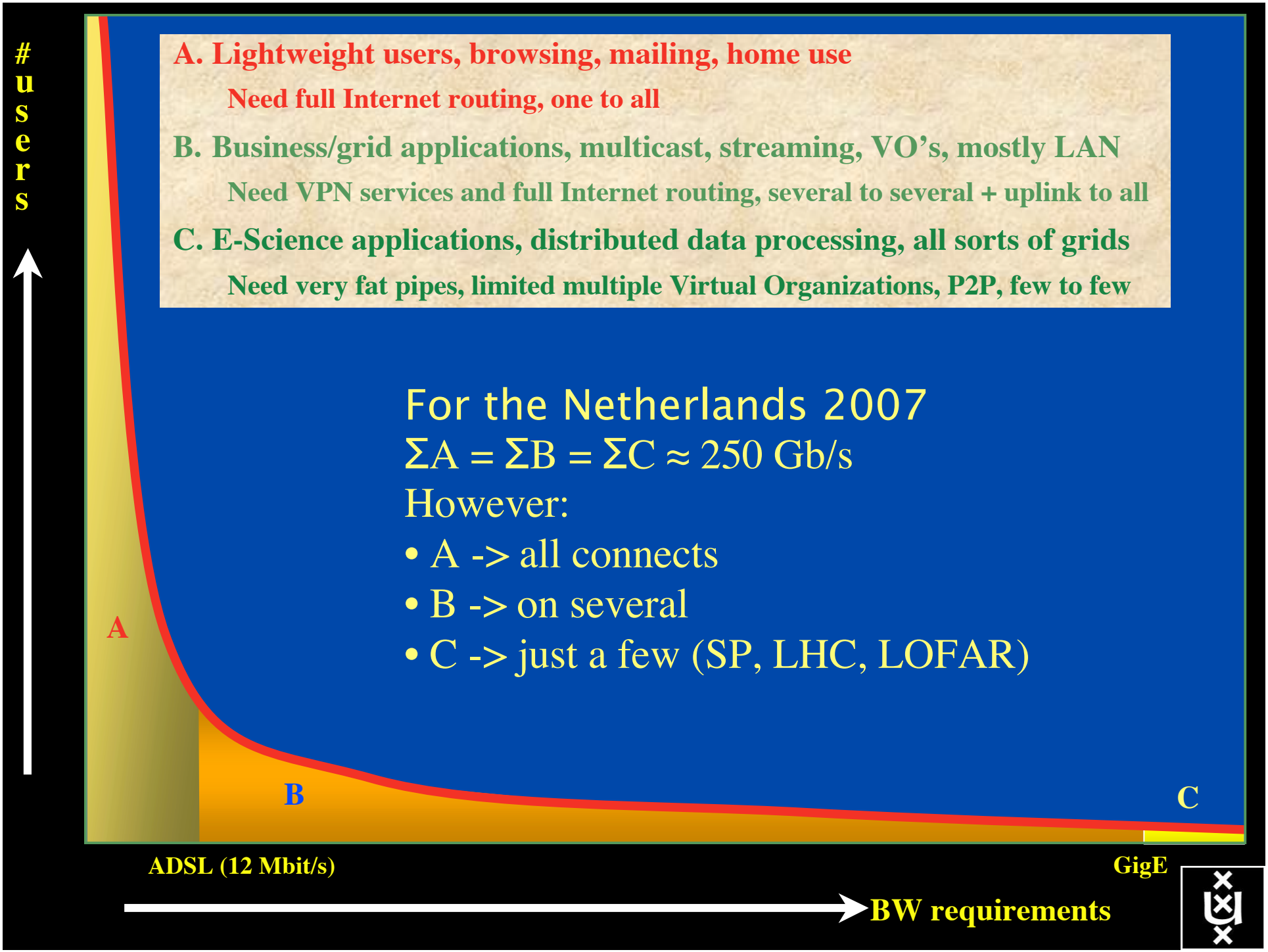




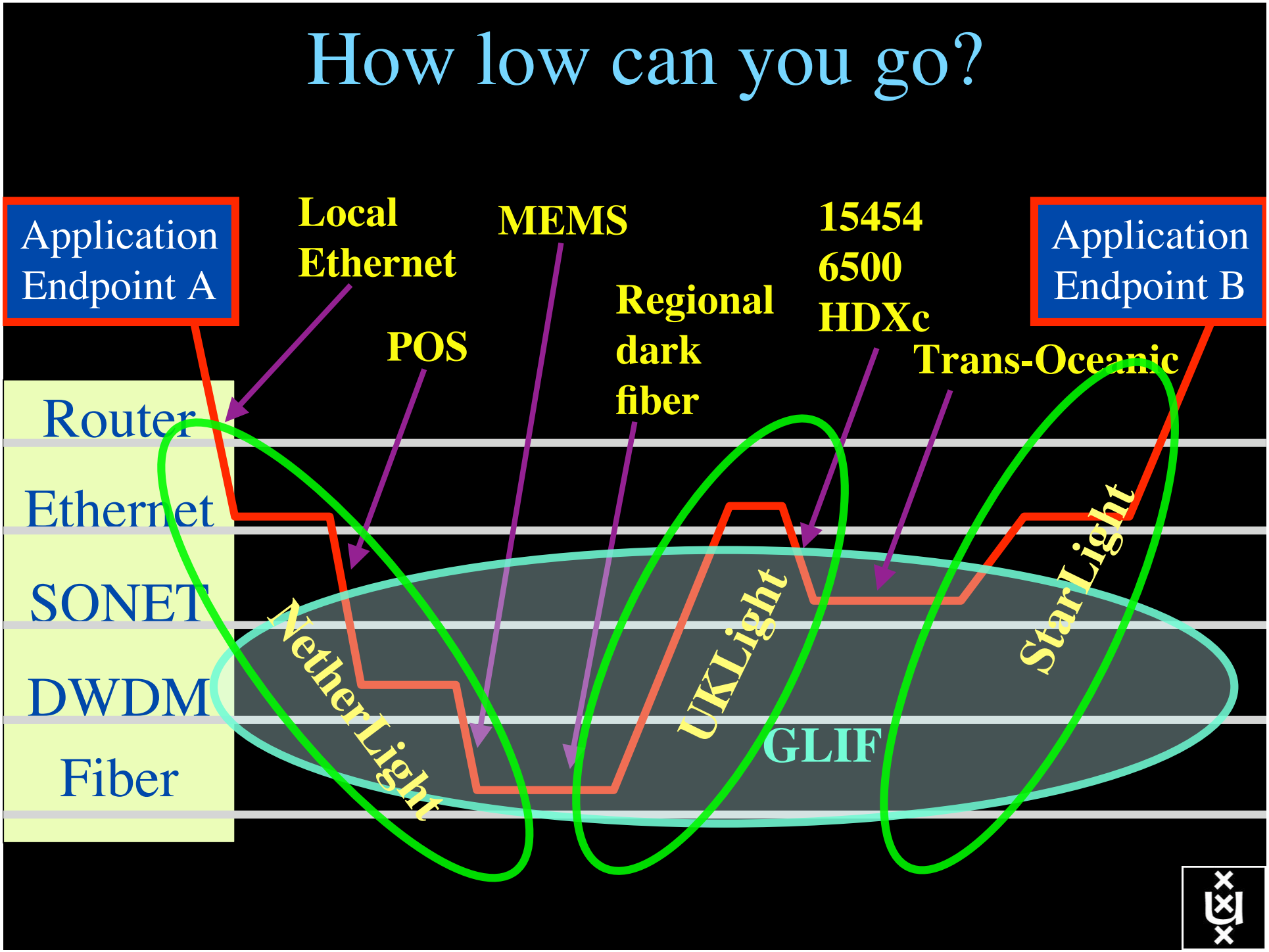
Contents

This page is intentionally left blank

- Ref: www.this-page-intentionally-left-blank.org


#users
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 2007
ΣA = ΣB = ΣC ≈ 250 Gb/s
However:
• A -> all connects
• B -> on several
• C -> just a few (SP, LHC, LOFAR)
A
B
C
ADSL (12 Mbit/s)
GigE
BW requirements

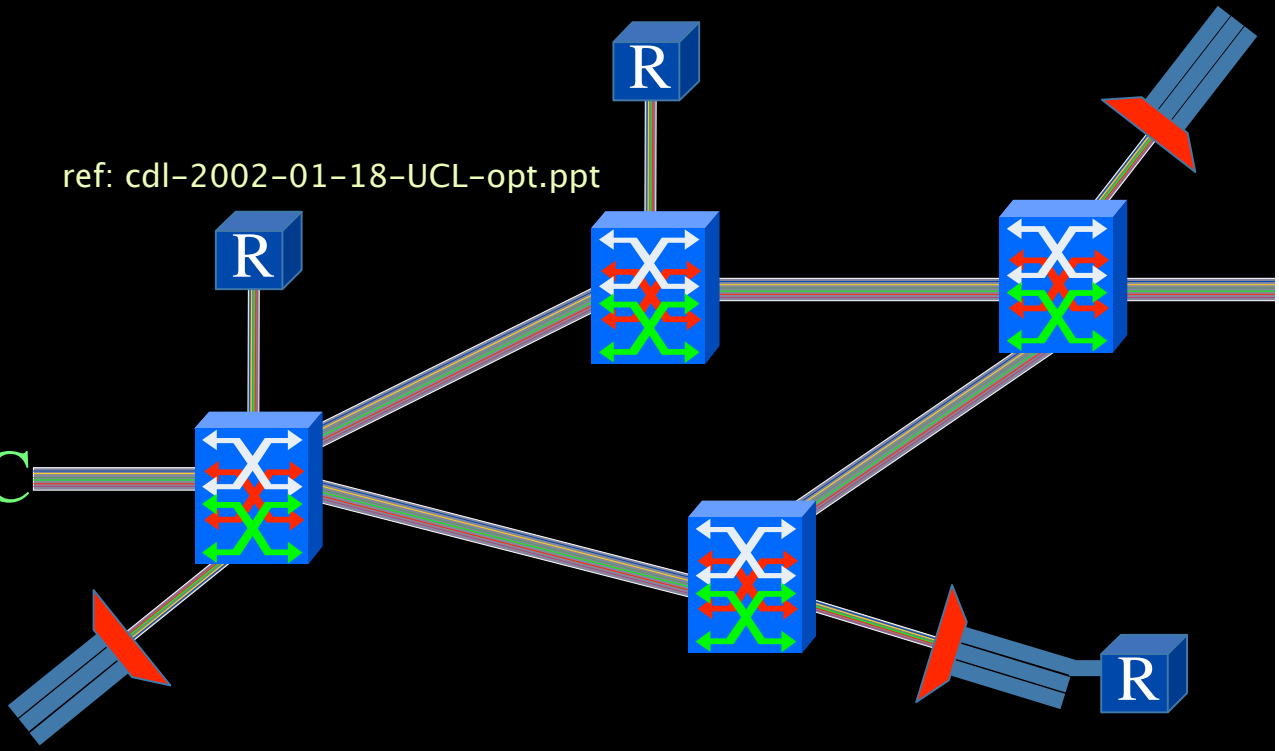
How low can you go?



History

SURFnet6 Architecture discussions 2001-2002

- photonic backbone
- L1 - L3 services
- NORTEL
- Static provisioning
- Summer 2004 K&C
- NWO-GLANCE
- StarPlane



- The StarPlane vision is to give flexibility directly to the applications by allowing them to choose the logical topology in real time, ultimately with sub-second lambda switching times on part of the SURFnet6 infrastructure.

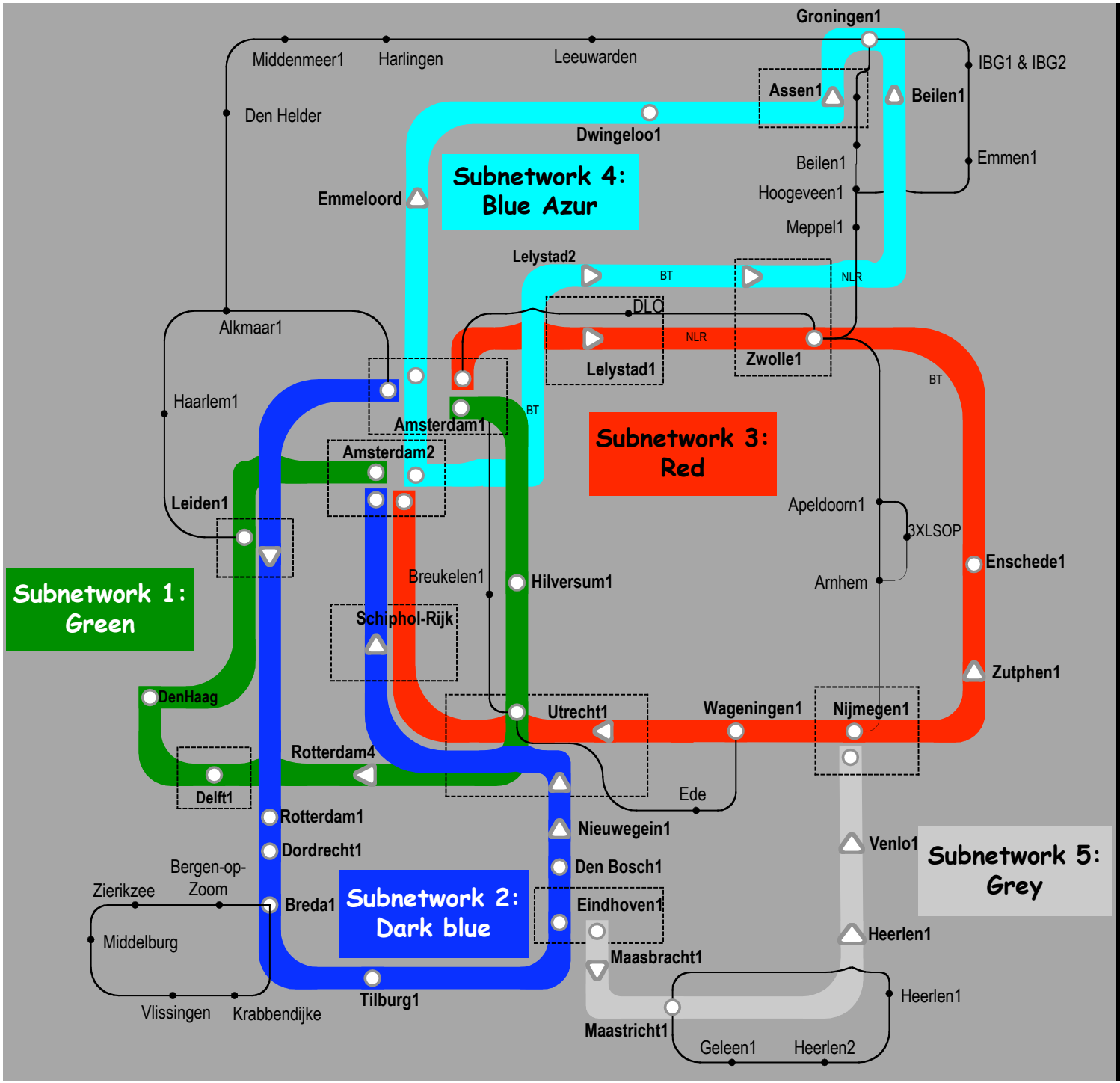
In The Netherlands SURFnet connects between 180:

- universities;
- academic hospitals;
- most polytechnics;
- research centers.

with an indirect ~750K user base

~ 6000 km

scale comparable to railway system

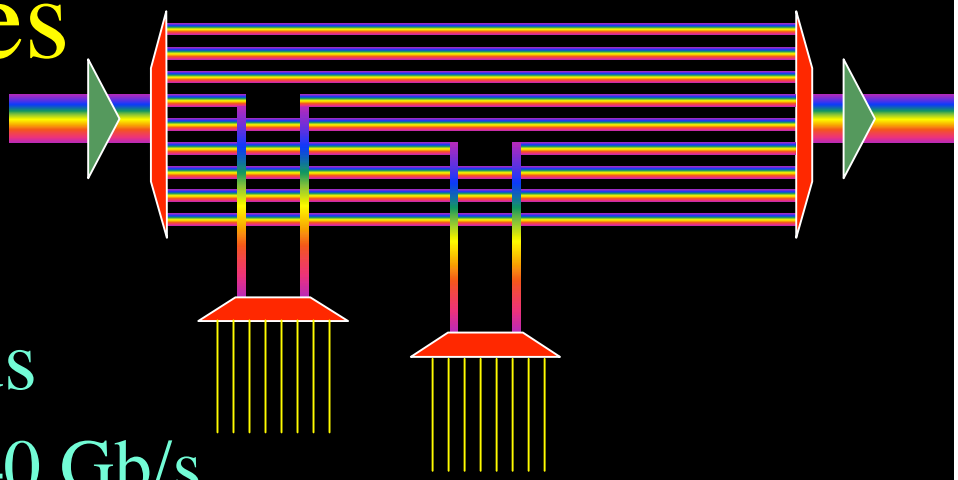


Common Photonic Layer (CPL) in SURFnet6

supports up to
72 Lambda's of
10 G each
40 G soon.

SURFnet

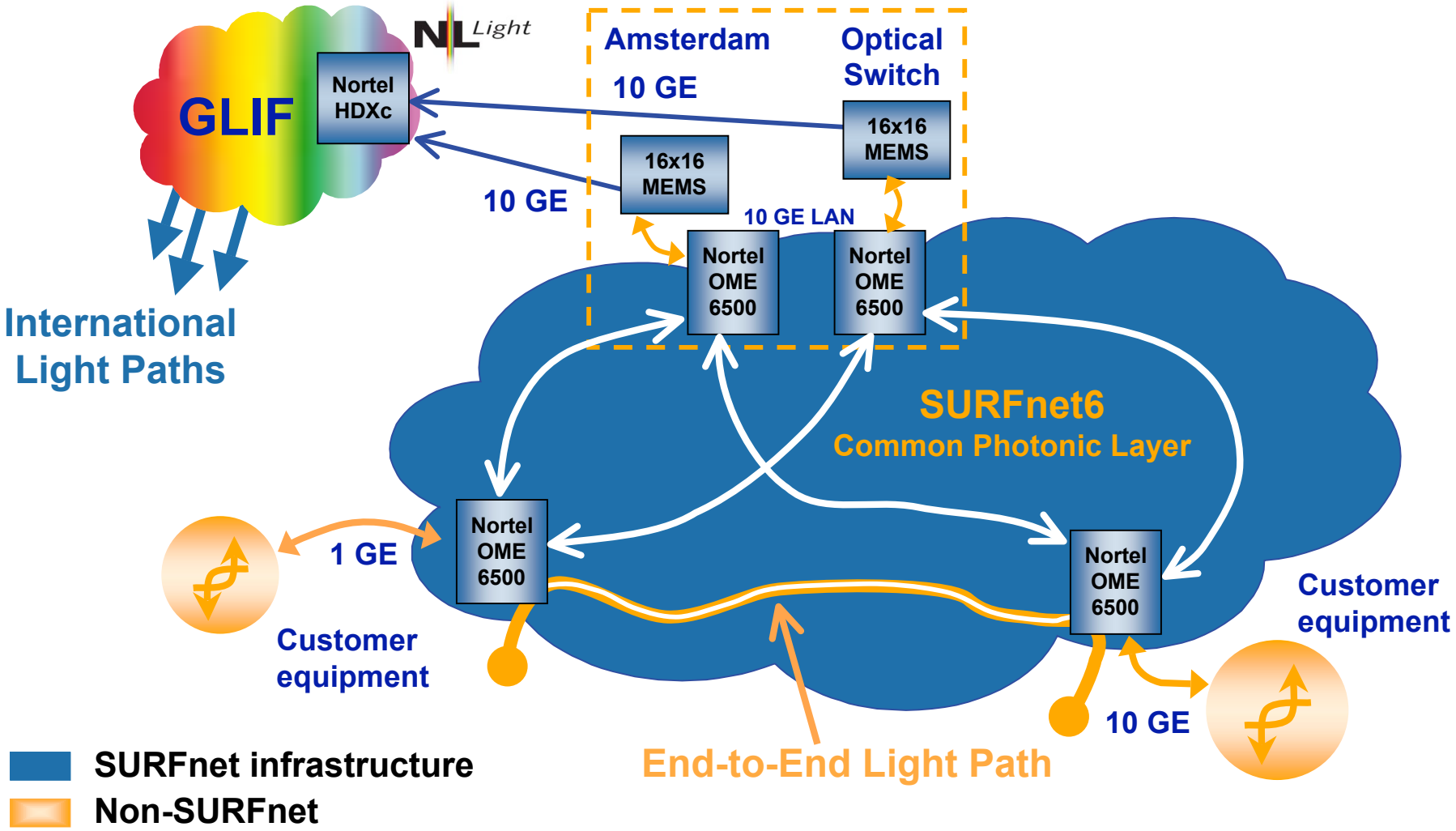
SURFnet 6 principles



- Based on dark fiber
- 4 DWDM rings of 9 bands
  - Each capable of 10, later 40 Gb/s
  - each 4 (100 GHz spacing) or 8 (50 GHz spacing) colors
- Universities each have 1 band to connect their Routers +LightPaths
- Connect with 1 or 10 Gb/s Ethernet LanPhy
- Routing in Amsterdam in 2 core POP's!
- International connectivity in Amsterdam
- Lambda service between ring POP's and to NetherLight

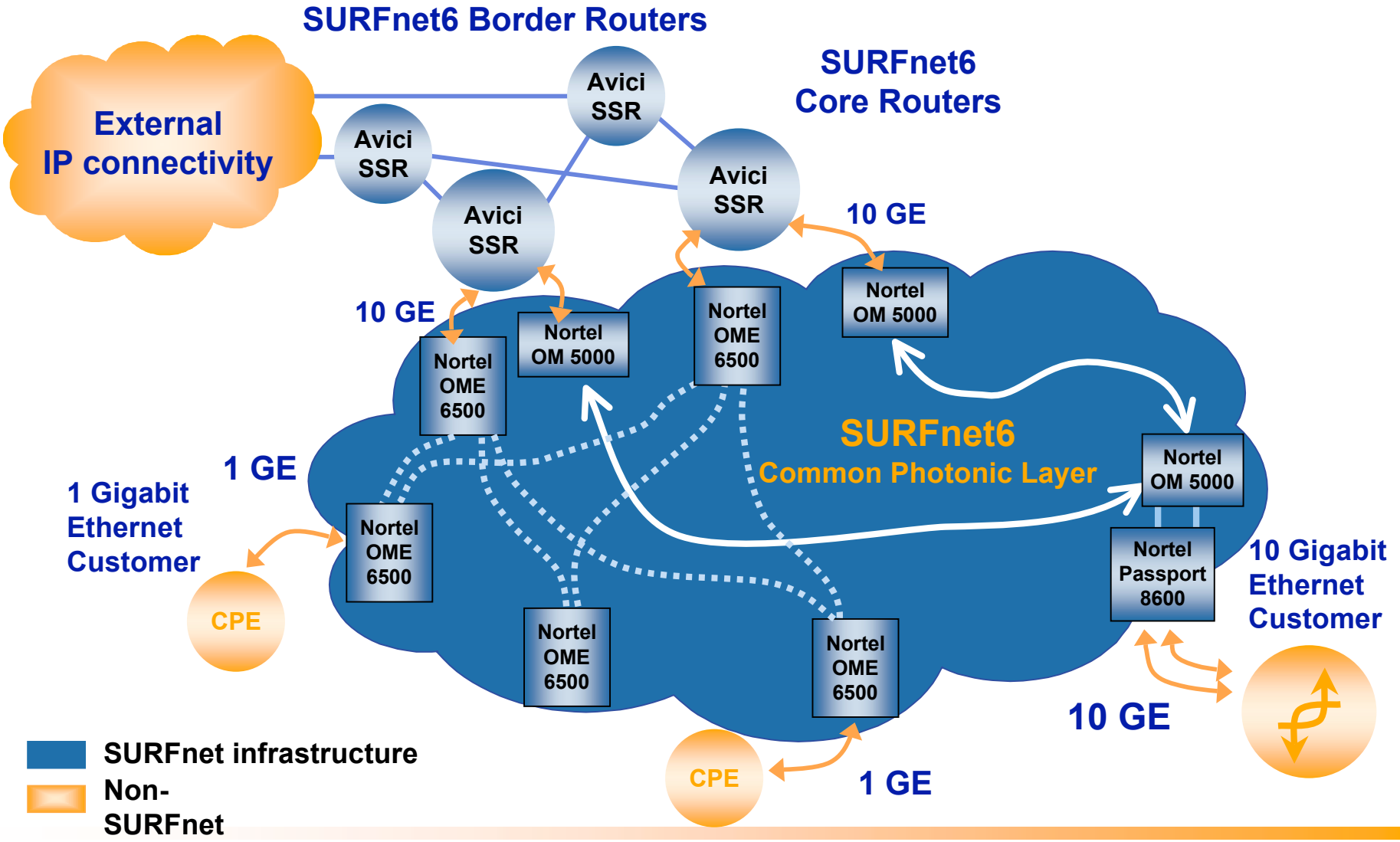
GigaPort

SURFnet6: Light Path Provisioning implementation



SURFnet

SURFnet6: IP network implementation

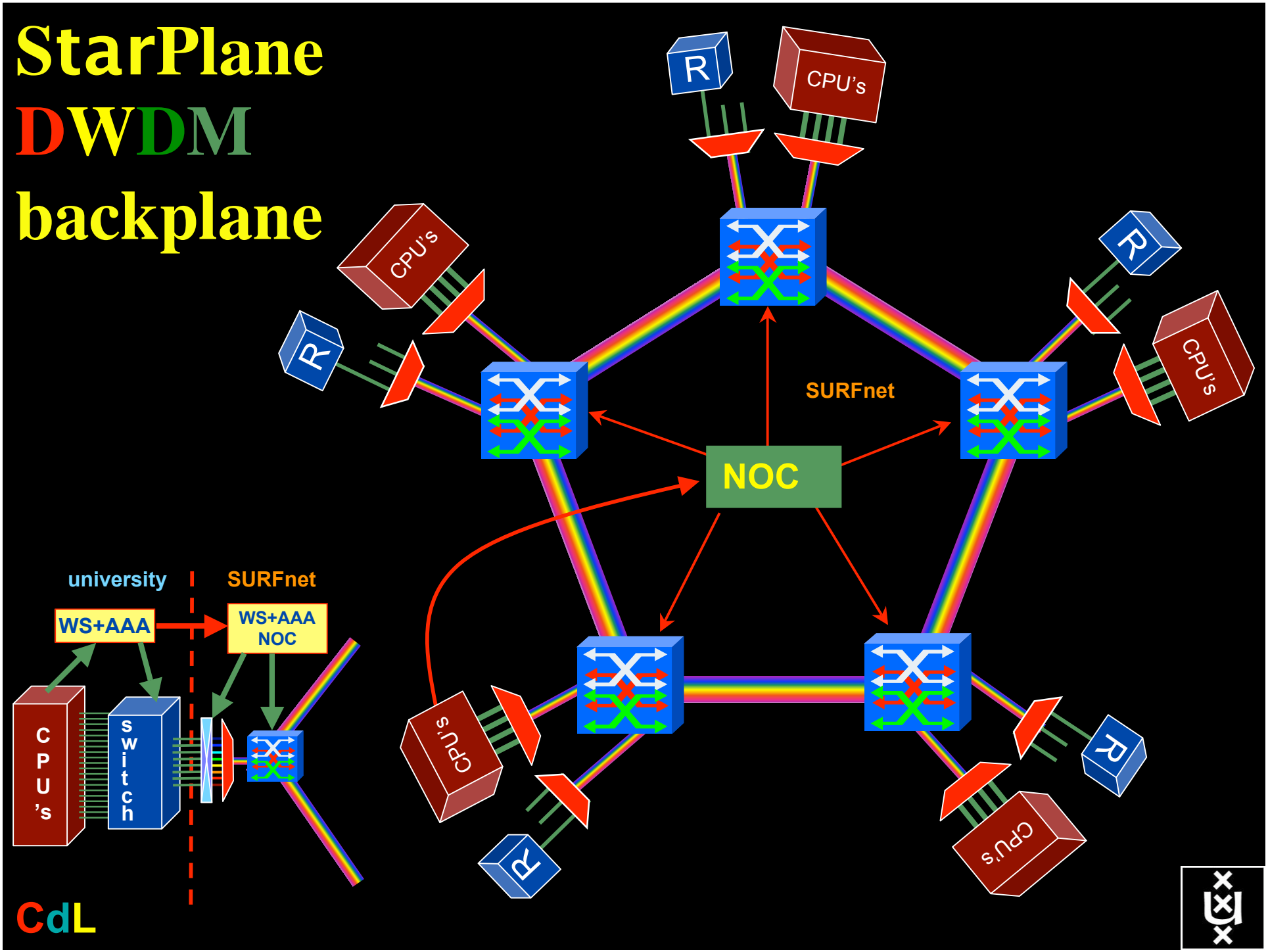


TeleCity2, Amsterdam

GigaPort

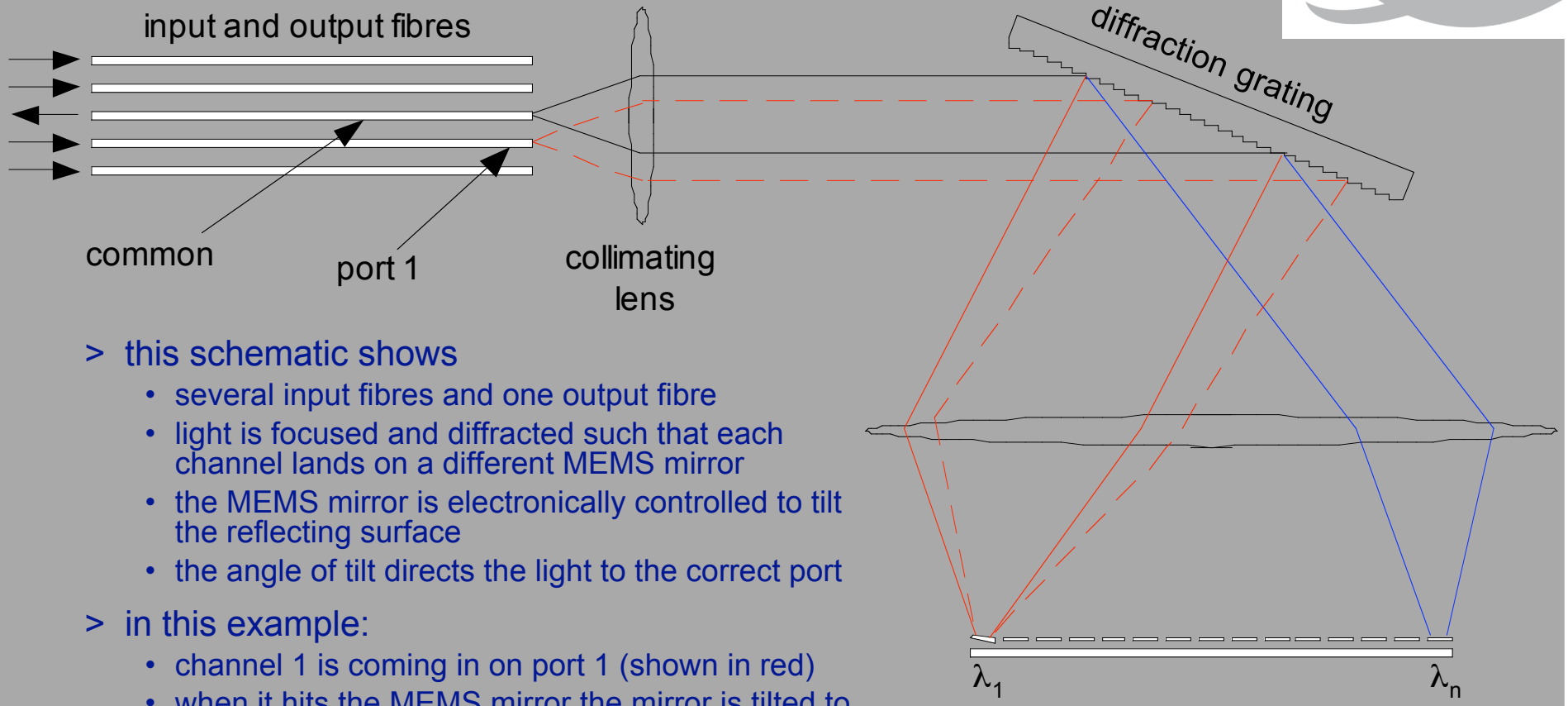
SURFnet

StarPlane
DWDM
backplane



CdL

Module Operation



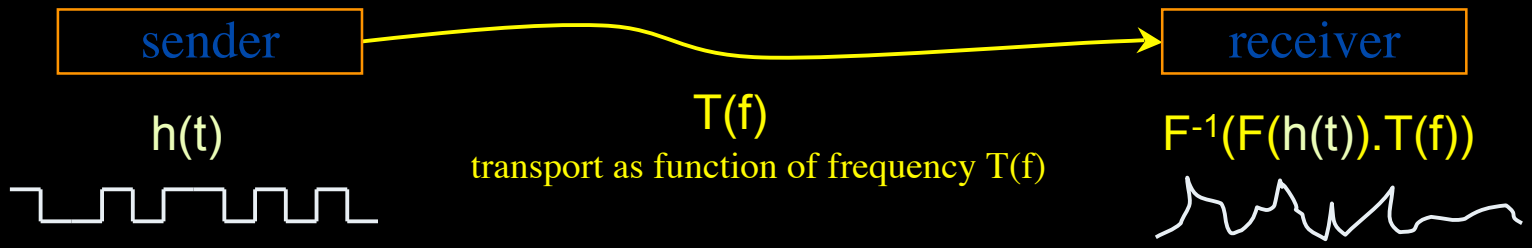
> this schematic shows
  - several input fibres and one output fibre
  - light is focused and diffracted such that each channel lands on a different MEMS mirror
  - the MEMS mirror is electronically controlled to tilt the reflecting surface
  - the angle of tilt directs the light to the correct port

> in this example:
  - channel 1 is coming in on port 1 (shown in red)
  - when it hits the MEMS mirror the mirror is tilted to direct this channel from port 1 to the common
  - only port 1 satisfies this angle, therefore all other ports are blocked

ref Eric Bernier, NORTEL

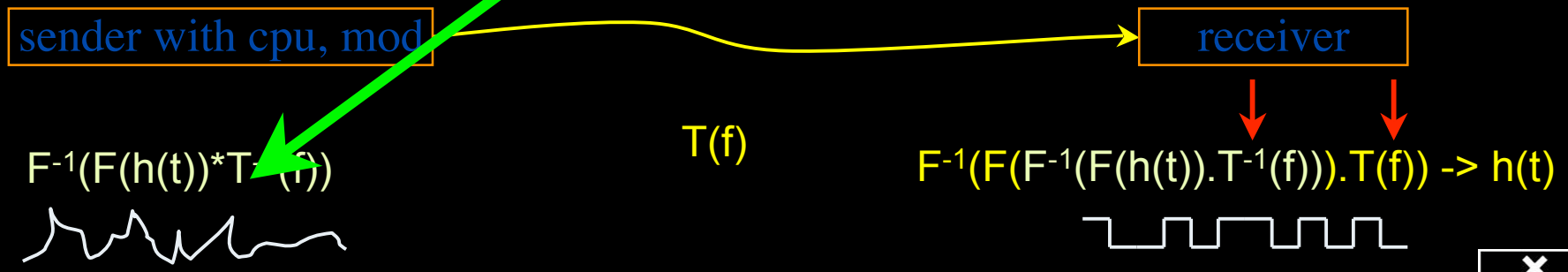
Dispersion compensating modem: eDCO from NORTEL

(Try to Google eDCO :-)



Solution in 5 easy steps for dummy's :

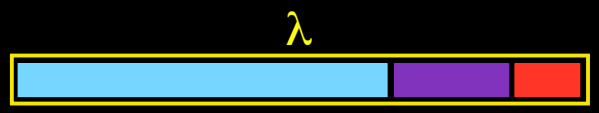
1. try to figure out T(f) by trial and error
2. invert T(f) -> $T^{-1}(f)$
3. computationally multiply $T^{-1}(f)$ with Fourier transform of bit pattern to send
4. inverse Fourier transform the result from frequency to time space
5. modulate laser with resulting $h'(t) = F^{-1}(F(h(t)).T^{-1}(f))$



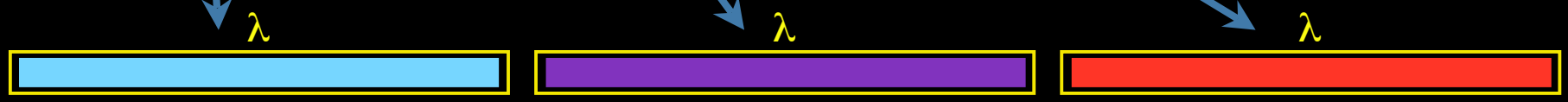
(ps. due to power ~ square E the signal to send **looks** like uncompensated received but is not)

QOS in a non destructive way!

- Destructive QOS:
  - have a link or λ
  - set part of it aside for a lucky few under higher priority
  - rest gets less service

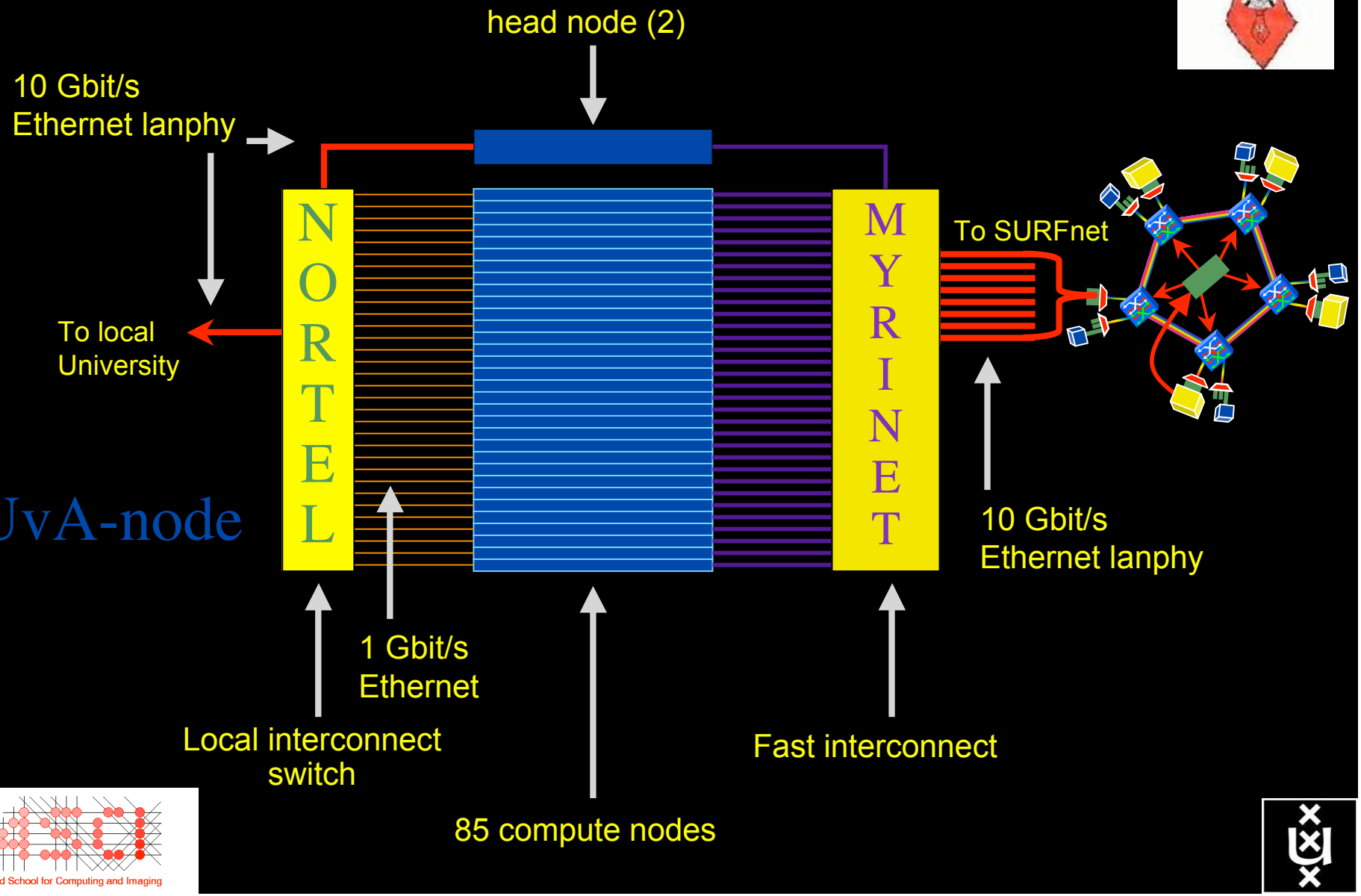


- Constructive QOS:
  - have a λ
  - add other λ's as needed on separate colors
  - move the lucky ones over there
  - rest gets also a bit happier!



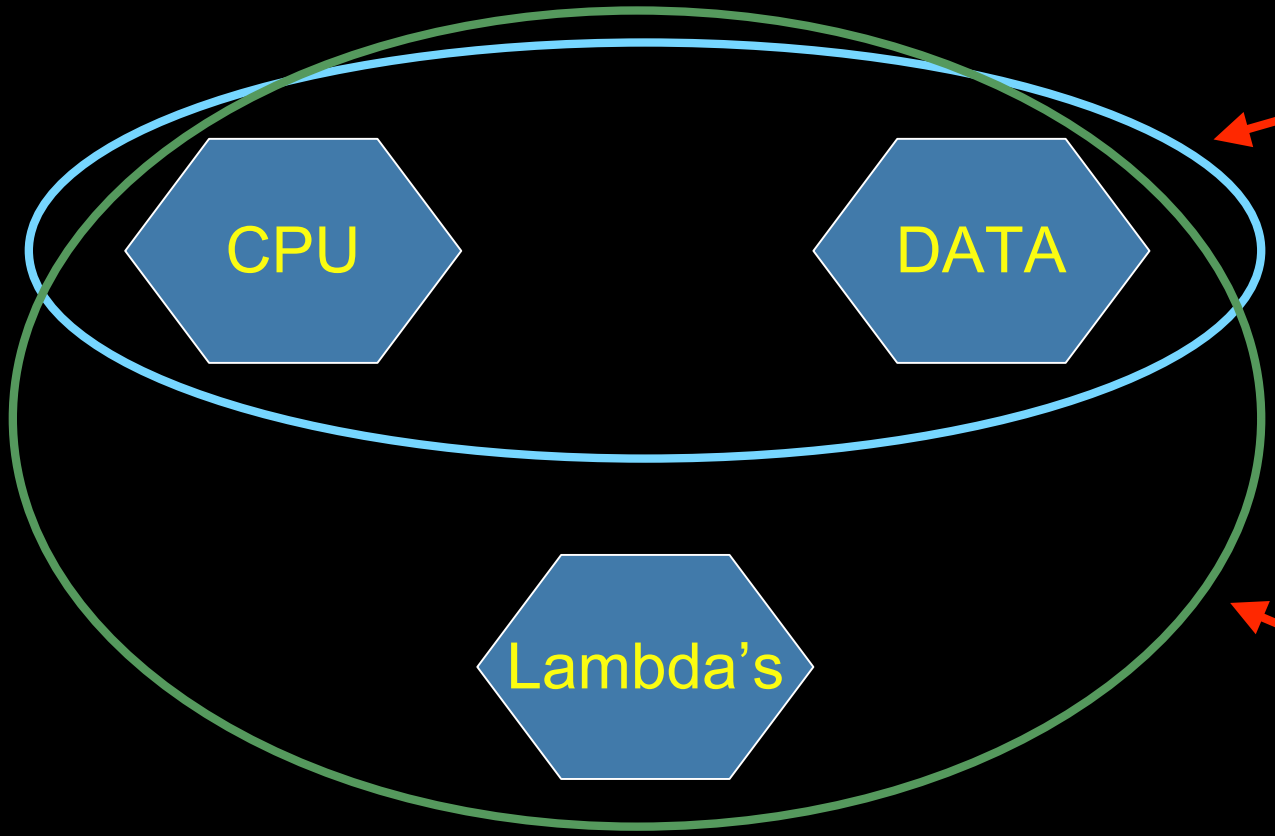
DAS-3 Cluster

http://www.clustervision.com/pr_das3_uk.html



Advanced School for Computing and Imaging

GRID-Colocation problem space



Extensively under research

New!

What makes StarPlane fly?

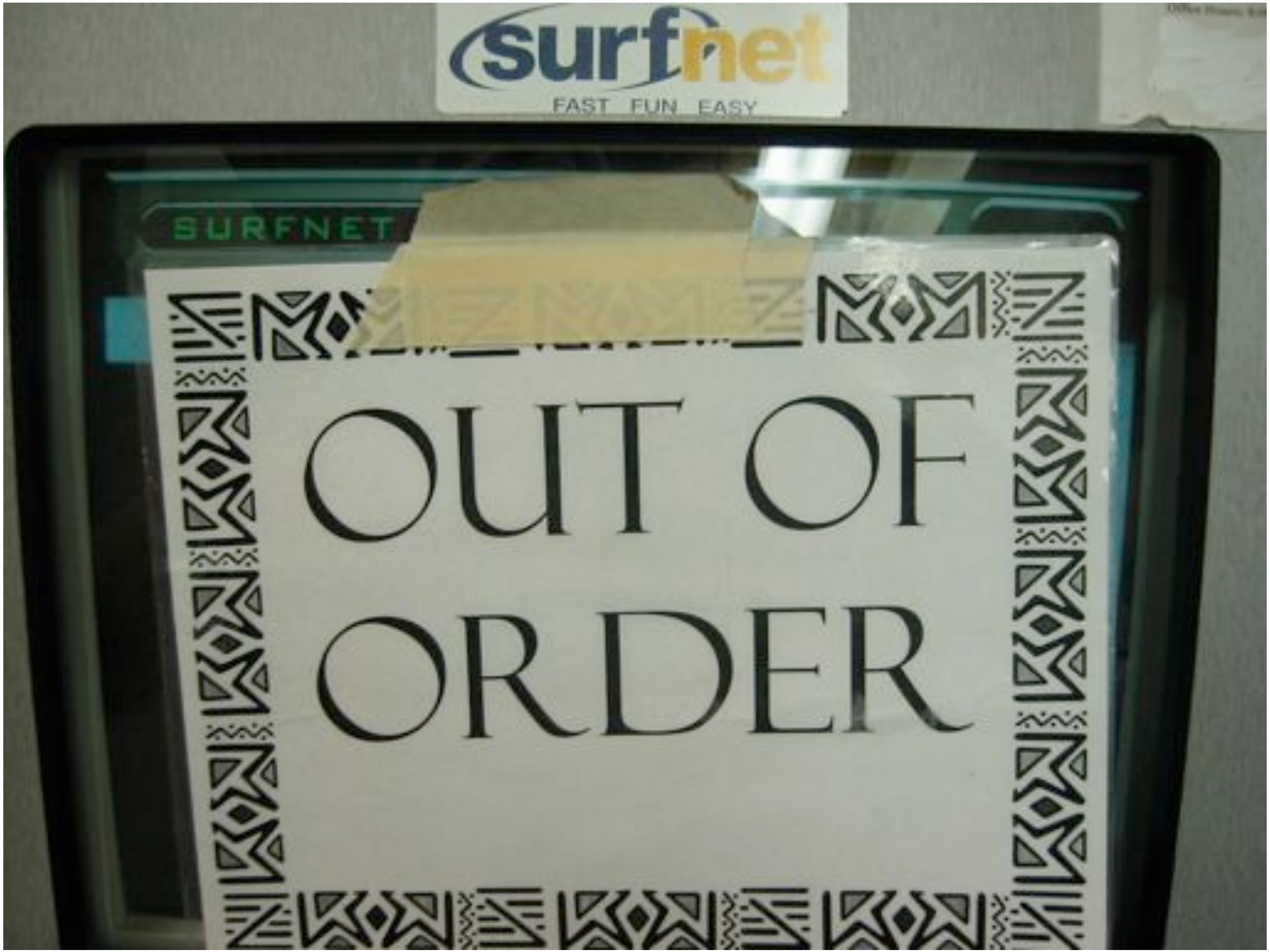
- Wavelength Selective Switches
  - for the “low cost” photonics
- Sandbox by confining StarPlane to one band
  - for experimenting on a production network
- Optimization of the controls to turn on/off a Lambda
  - direct access to part of the controls at the NOC
- electronic Dynamically Compensating Optics (eDCO)
  - to compensate for changing lengths of the path
- traffic engineering
  - to create the OPN topologies needed by the applications
- Open Source GMPLS
  - to facillitate policy enabled cross domain signalling

Power is a big issue

- UvA cluster uses (max) 30 kWh
- 1 kWh ~ 0.1 €
- per year -> 26 k€/y
- add cooling 50% -> 39 k€/y
- Emergency power system -> 50 k€/y
- per rack 10 kWh is now normal
- **YOU BURN ABOUT HALF THE CLUSTER OVER ITS LIFETIME!**


surfnet
FAST. FUN. EASY.
SURFNET PREMIERE
HELP
net
Three Easy Steps :
Click the START button
Insert money...
$0.25 per minute...
Example :
$1 = 4 minutes
$5 = 20 minutes
No change is provided!
Surf the web!


surfnet
FAST FUN EASY
SURFNET PREMIERE
HELP
019
2
ROYAL MAIL
FRANKING
SERVICE
2nd
Paid
Check your
email
here!
Click the Start Button to begin


surfnet
FAST FUN EASY
SURFNET
OUT OF
ORDER

Simple service access

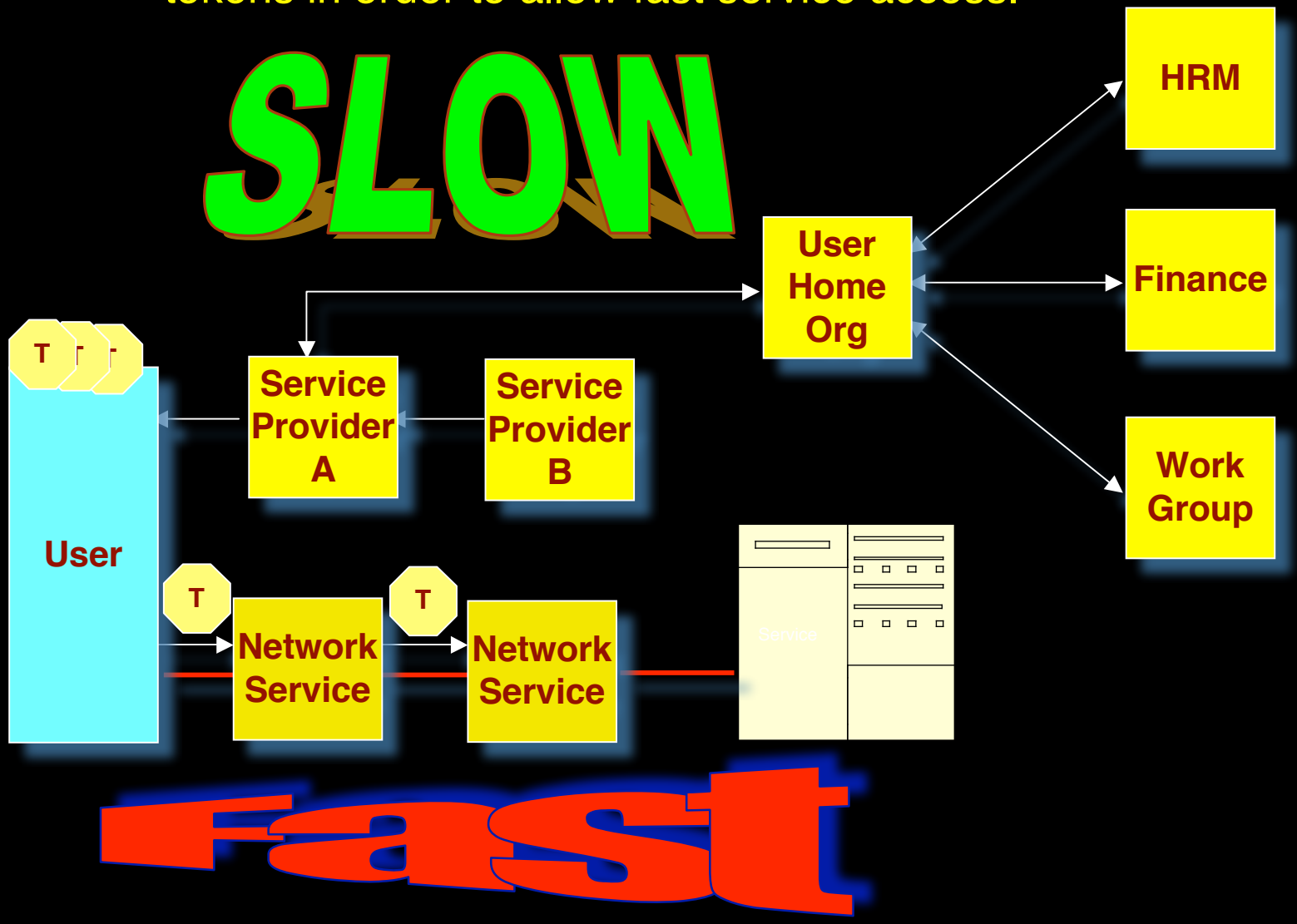
Pitlochry, Scotland - Summer 2005

Universiteit van Amsterdam

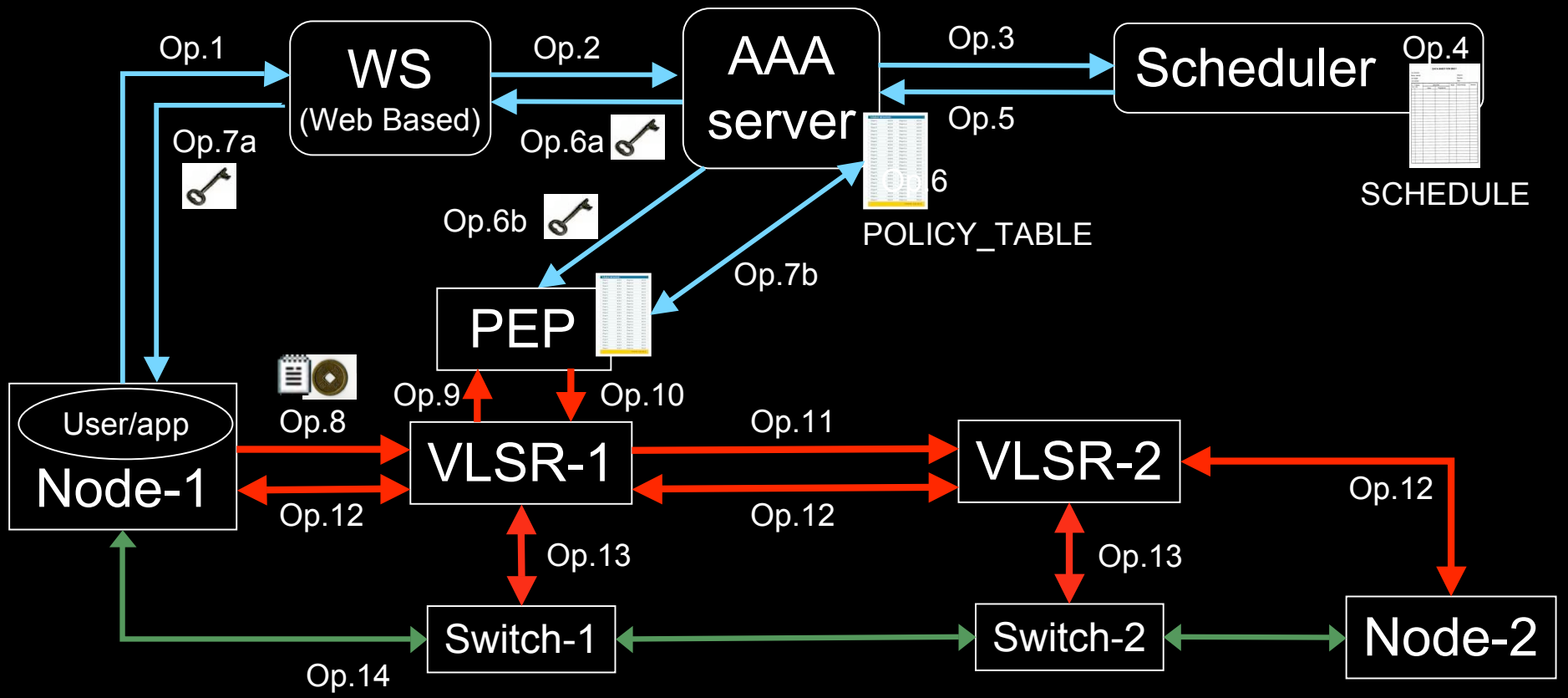
GigaPort

NextGRID

Use AAA concept to split (time consuming) service authorization process from service access using secure tokens in order to allow fast service access.



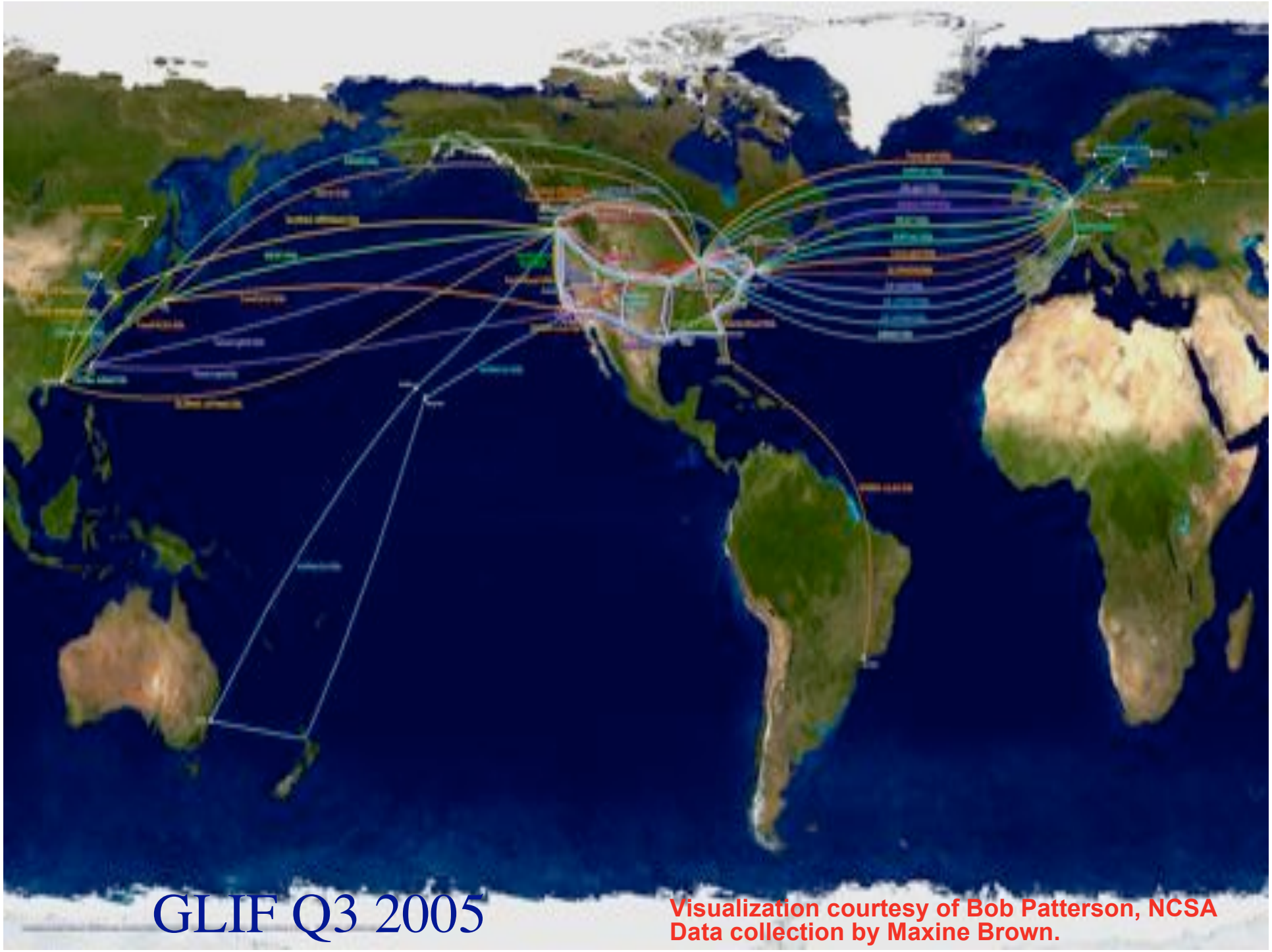
DRAGON GMPLS & TBN Demo, SC06 Tampa



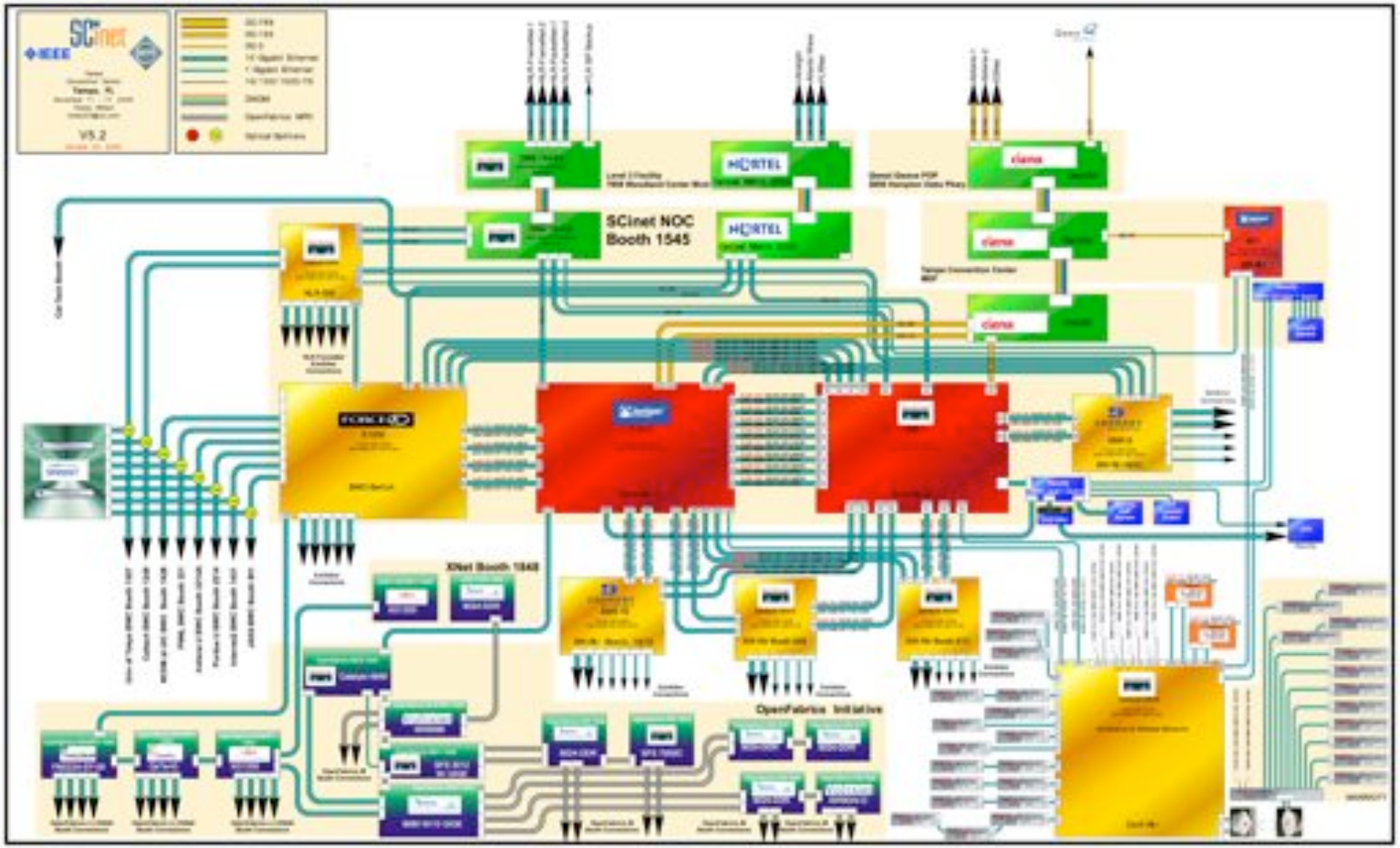
1. User (on Node1) requests a path via web to the WS.
2. WS sends the XML requests to the AAA server.
3. AAA server calculates a hashed index number and submits a request to the Scheduler.
4. Scheduler checks the SCHEDULE and add new entry.
5. Scheduler confirms the reservation to the AAA.
6. AAA server updates the POLICY_TABLE.

6a. AAA server issues an encrypted key to the WS.
6b. AAA server passes the same key to the PEP.
7a. WS passes the key to the user.
7b. AAA server interacts with PEP to update the local POLICY_TABLE on the PEP.

8. User constructs the RSVP message with extra Token data by using the key and sends to VLSR-1.
9. VLSR-1 queries PEP whether the Token in the RSVP message is valid.
10. PEP checks in the local POLICY_TABLE and return YES.
11. When VLSR-1 receives YES from PEP, it forwards the RSVP message.
12. All nodes process RSVP message(forwarding/response)
13. The Ethernet switches are configured
14. LSP is set up and traffic can flow

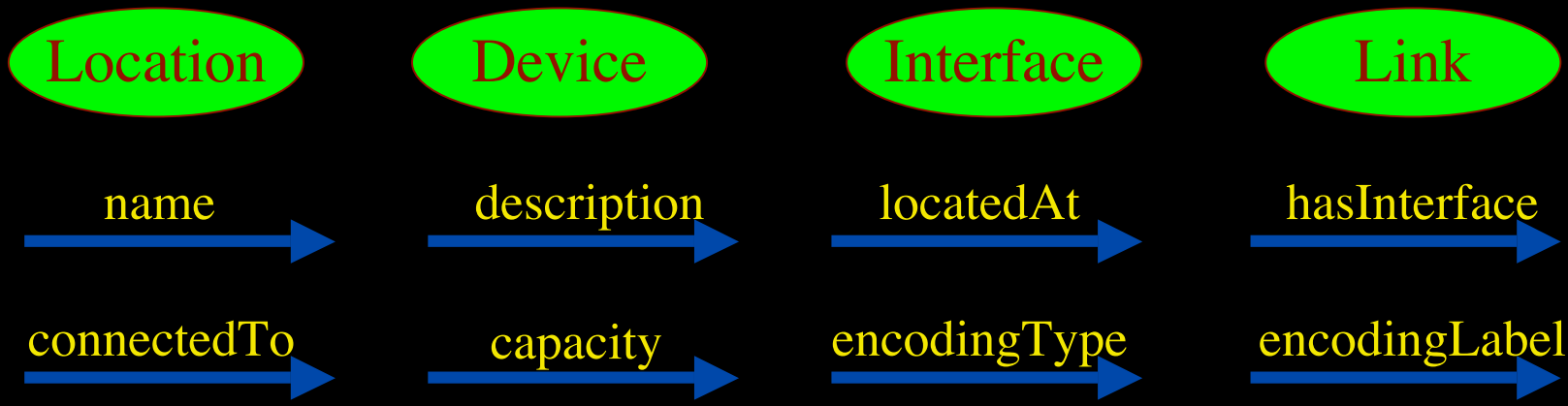

GLIF Q3 2005
Visualization courtesy of Bob Patterson, NCSA
Data collection by Maxine Brown.

Architecture SC06

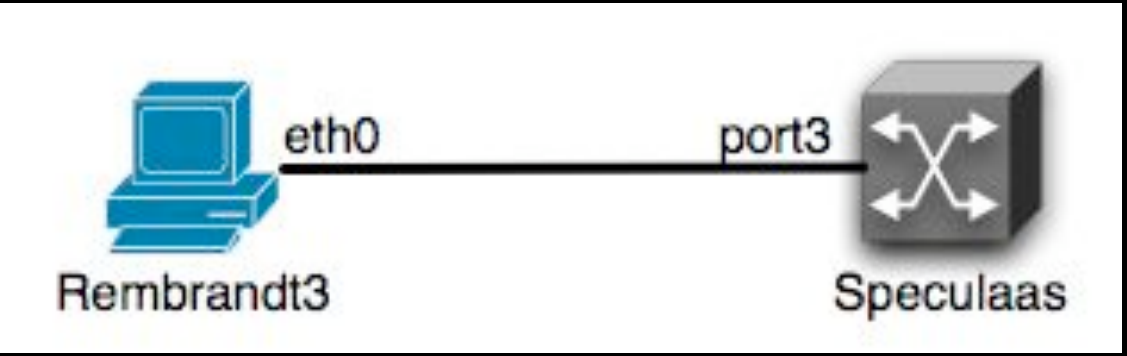


Network Description Language

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



NDL Example



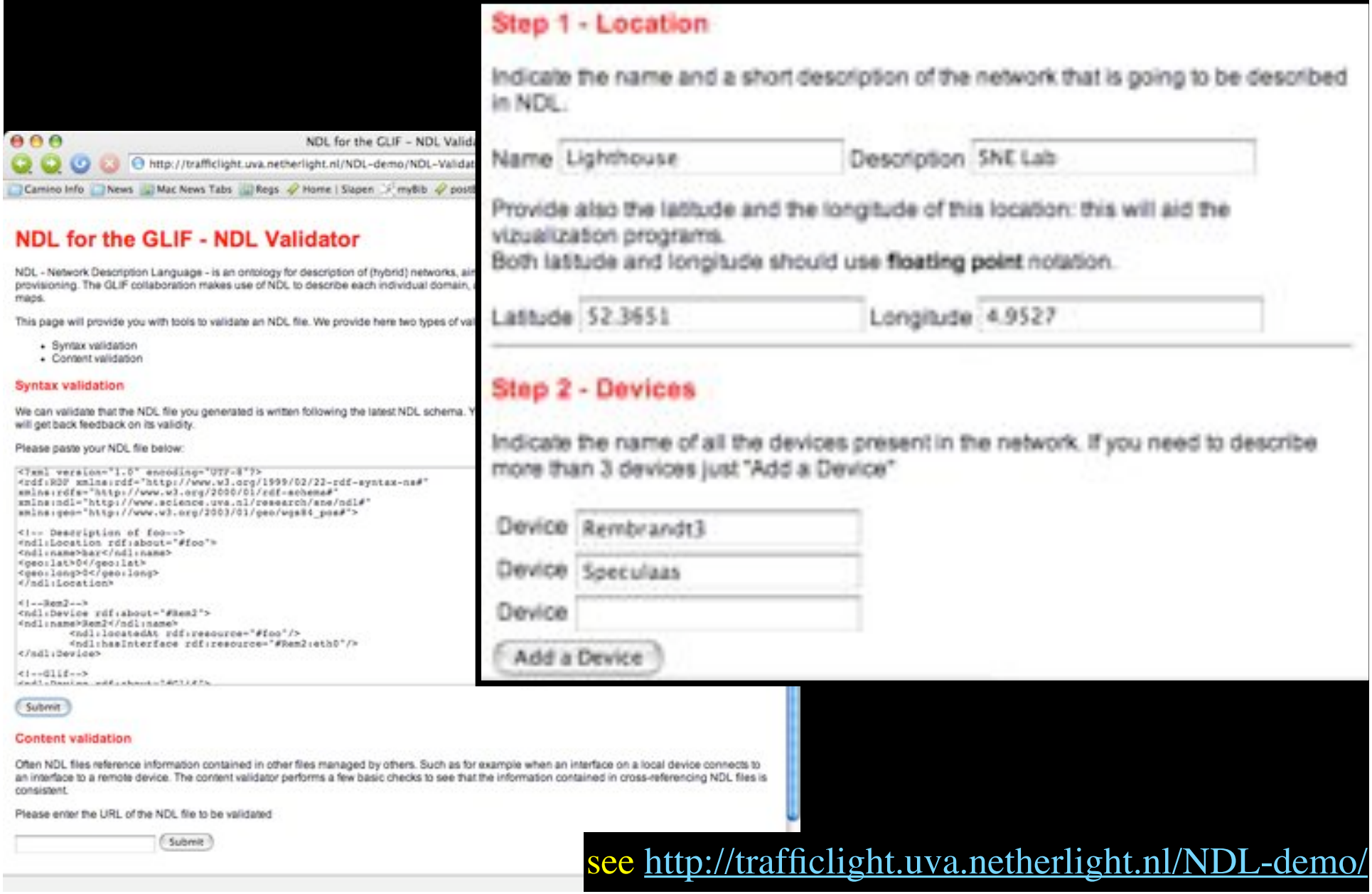

```
<ndl:Device rdf:about="#Rembrandt3">
    <ndl:name>Rembrandt3</ndl:name>
    <ndl:locatedAt rdf:resource="#Lighthouse"/>
    <ndl:hasInterface rdf:resource="#Rembrandt3:eth0"/>
</ndl:Device>
<ndl:Interface rdf:about="#Rembrandt3:eth0">
    <ndl:name>Rembrandt3:eth0</ndl:name>
    <ndl:connectedTo rdf:resource="#Speculaas:port3"/>
</ndl:Interface>
```

NetherLight in RDF

```
<?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:resour
      <ndl:hasInterface rdf:resour
      <ndl:hasInterface rdf:resour
      <ndl:hasInterface rdf:resour
      <ndl:hasInterface rdf:resour
      <ndl:hasInterface rdf:resour
      <ndl:hasInterface rdf:resour
      <ndl:hasInterface rdf:resour
```

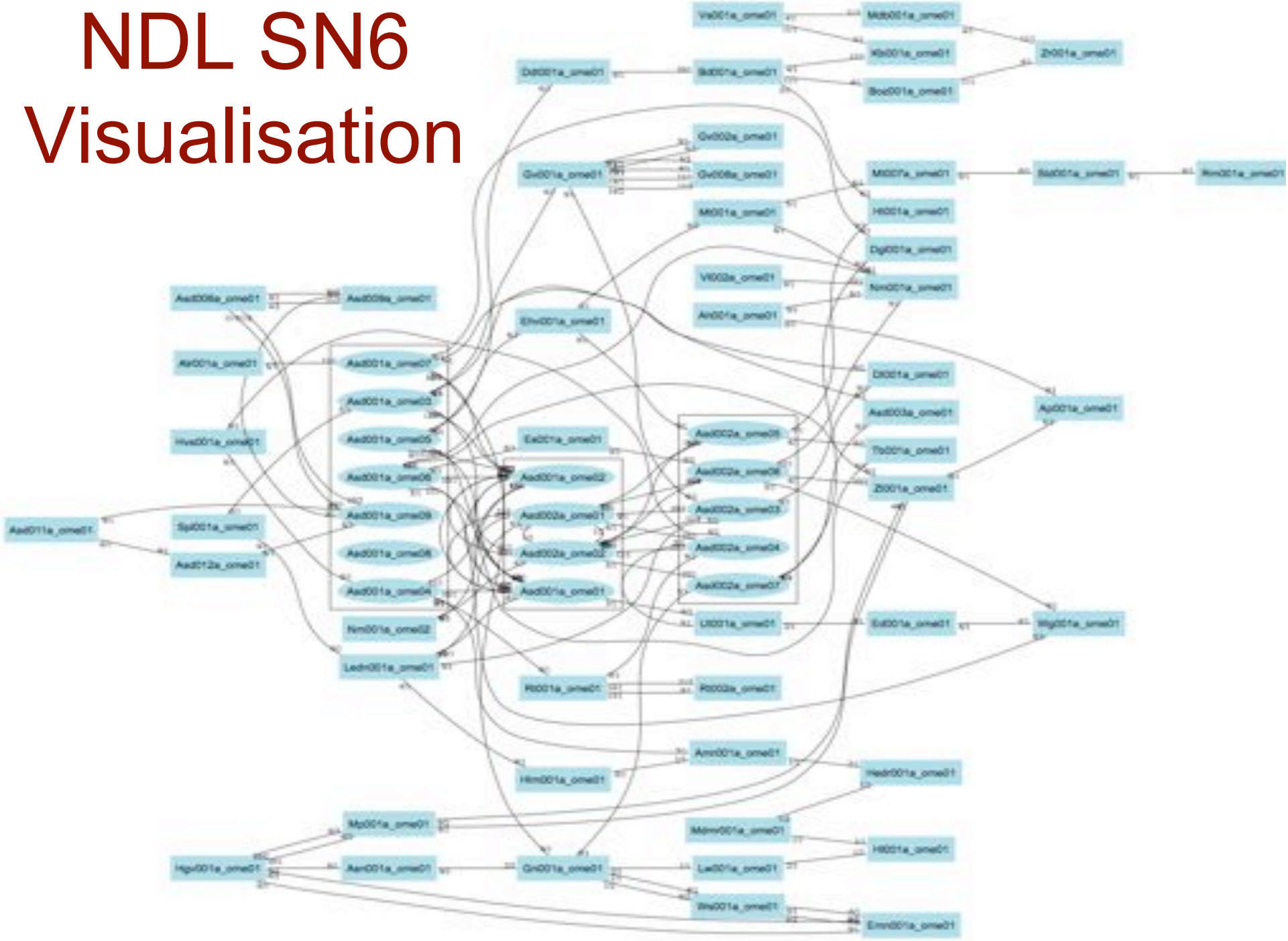
```
<!-- all the interfaces of TDM3.amsterdam1.netherlight.net -->

<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>
```

NDL Generator and Validator



see http://trafficlight.uva.netherlight.nl/NDL-demo/

NDL SN6 Visualisation



Current status: NDL

NDL - Network Description Language - an RDF based model for hybrid network descriptions.

It leverages all the semantic web tools, to provide:

- parsing of the RDF files
- graphs and visualization of connections and lightpaths
- lightpath provisioning support at inter and intra domain level.

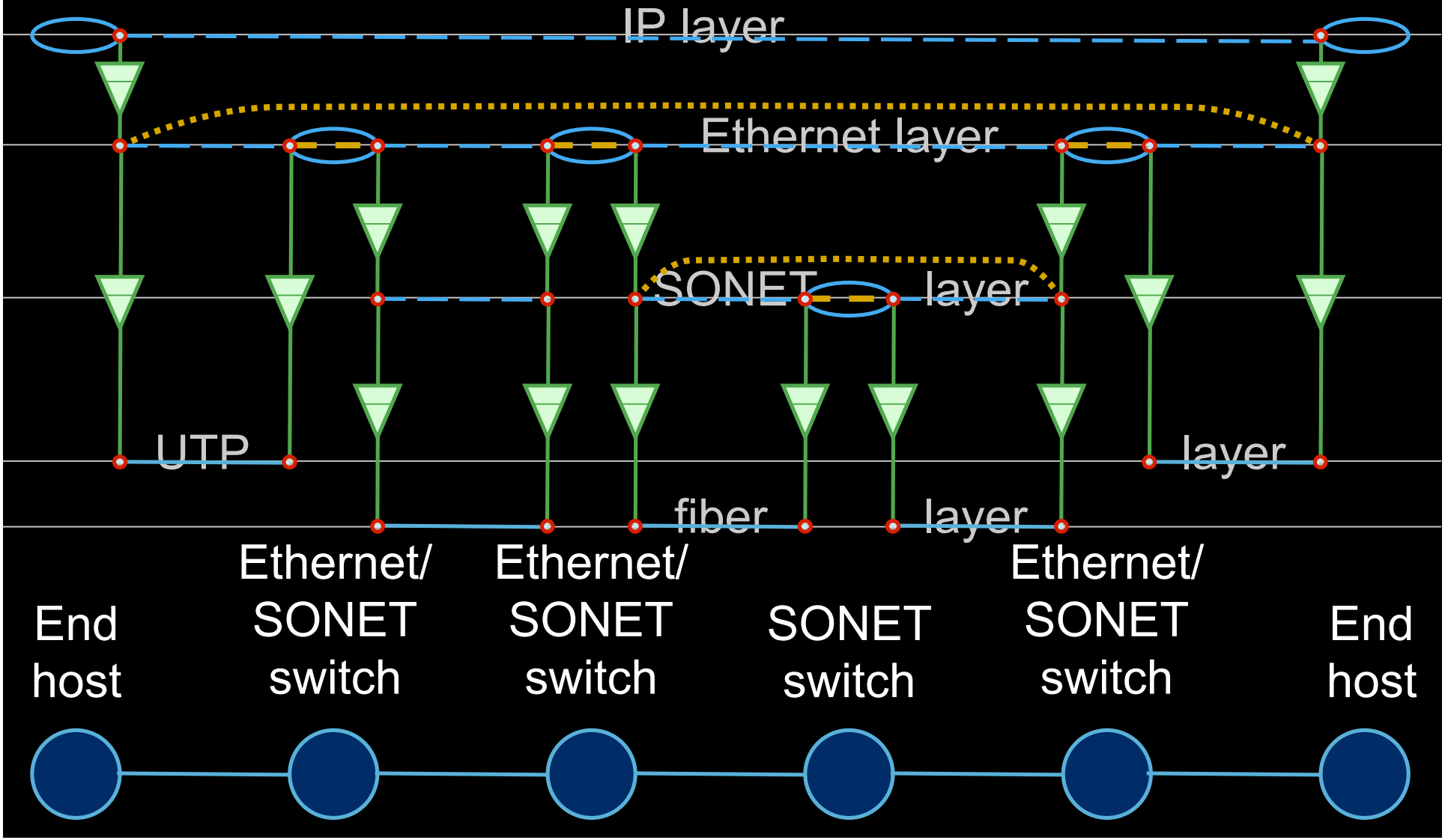
Latest developments were presented at the GLIF meeting in Sep. '06.

Google map and NDL…

…the GLIF connections described by NDL.

Multi-layer extensions to NDL

Layer schema based on G.805



OGF NML-WG

Open Grid Forum - Network Markup Language workgroup

Chairs:

- Paola Grosso – Universiteit van Amsterdam
- Martin Swany – University of Delaware

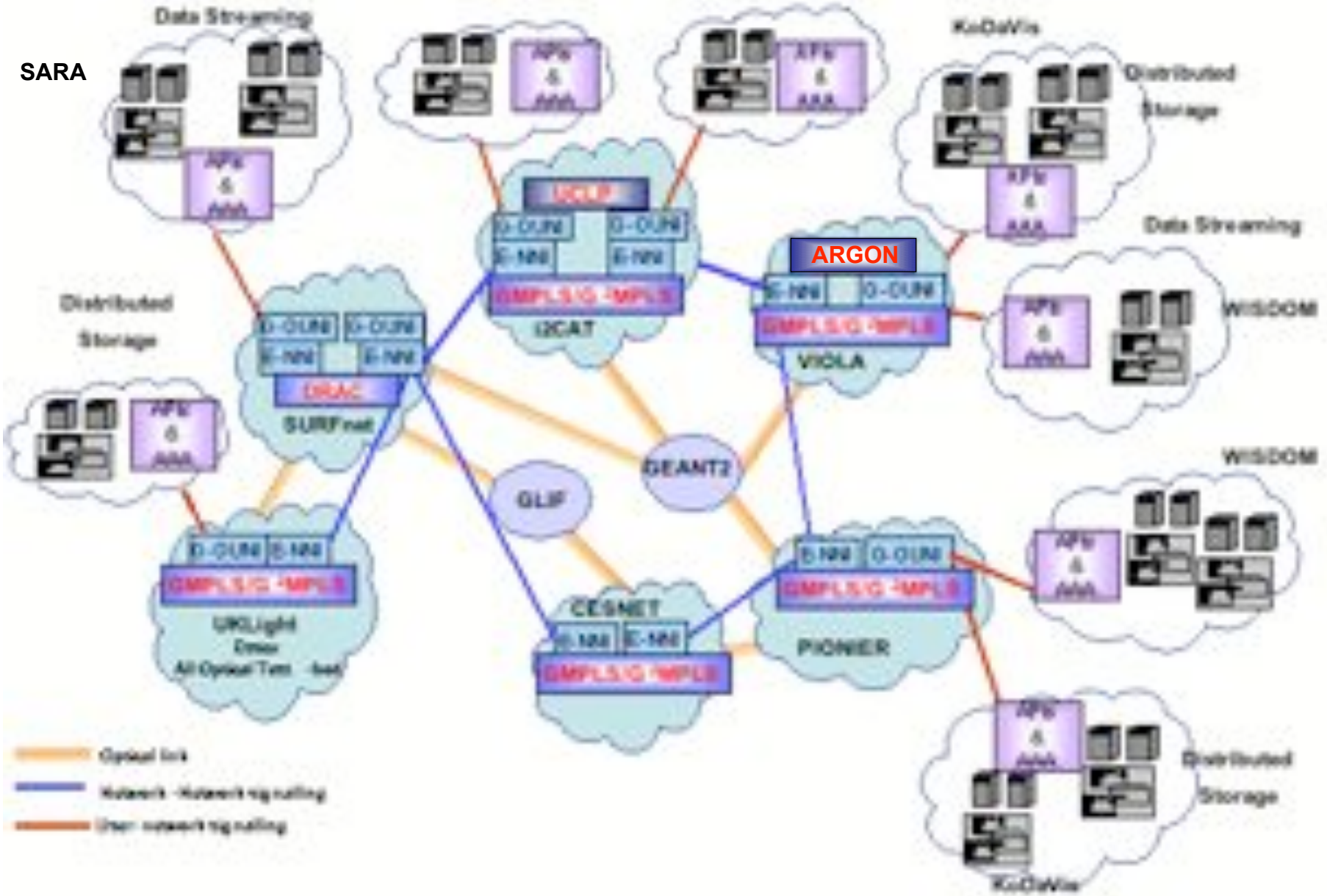
Purpose:

To describe network topologies, so that the outcome is a standardized network description ontology and schema, facilitating interoperability between different projects.

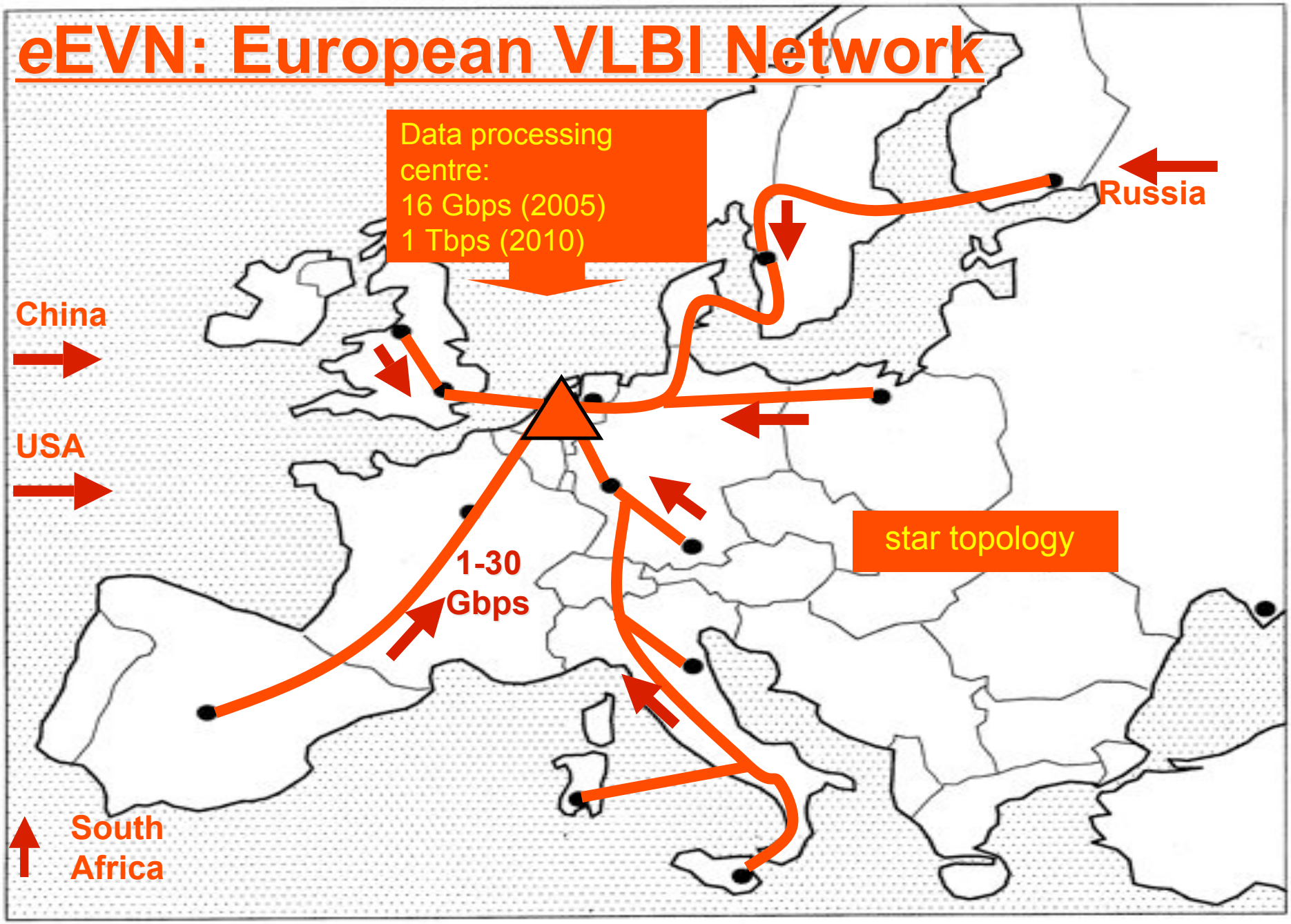
https://forge.gridforum.org/sf/projects/nml-wg

Phosphorus

European Multi-Domain Test-Bed Including Phosphorus Planned Developments

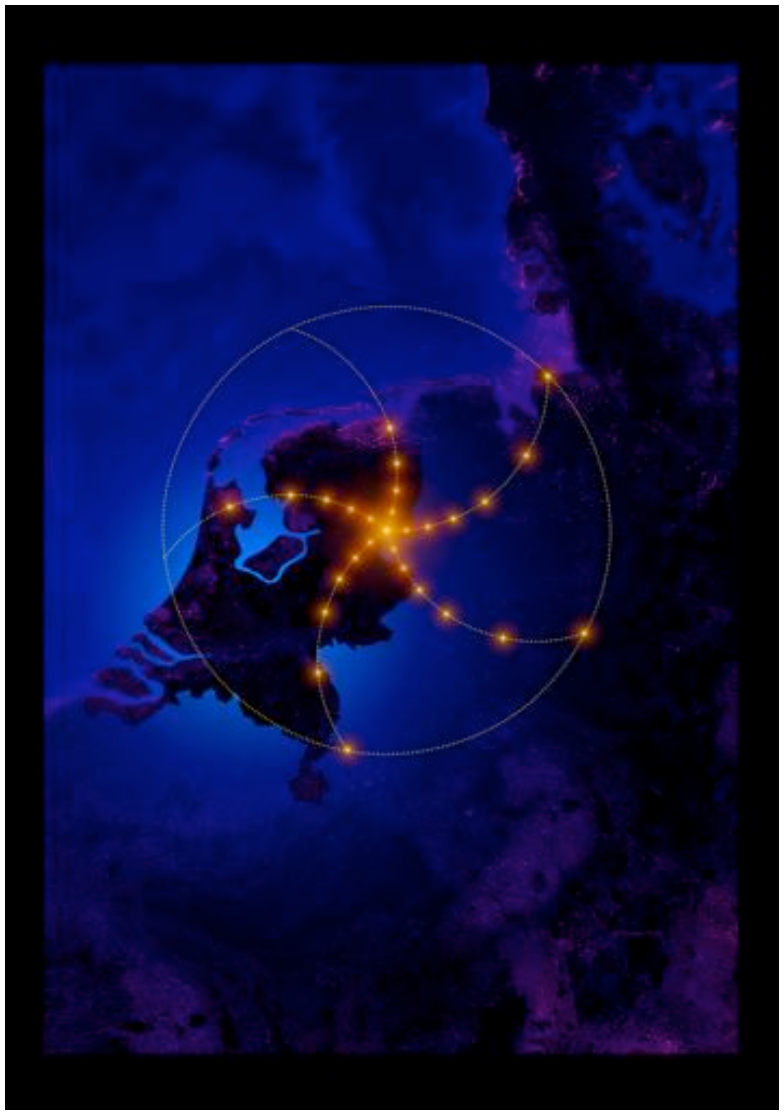


eEVN: European VLBI Network



This slide courtesy of Richard Schilizzi <schilizzi@jive.nl>

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

Slide courtesy of Marco de Vos (LOFAR)

The “Dead Cat” demo

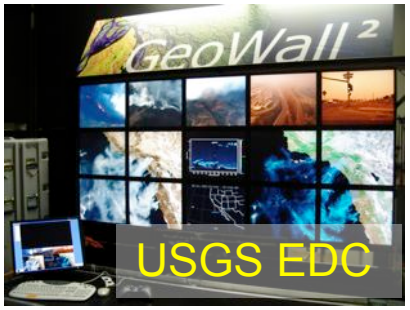
SC2004 & iGrid2005

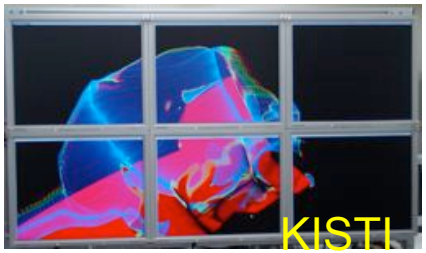
SC2004,
Pittsburgh,
Nov. 6 to 12, 2004
iGrid2005,
San Diego,
sept. 2005

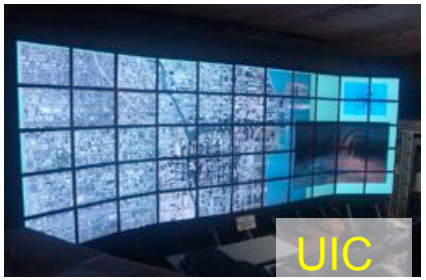
Produced by:
Michael Scarpa
Robert Belleman
Peter Sloot

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

US and International OptIPortal Sites

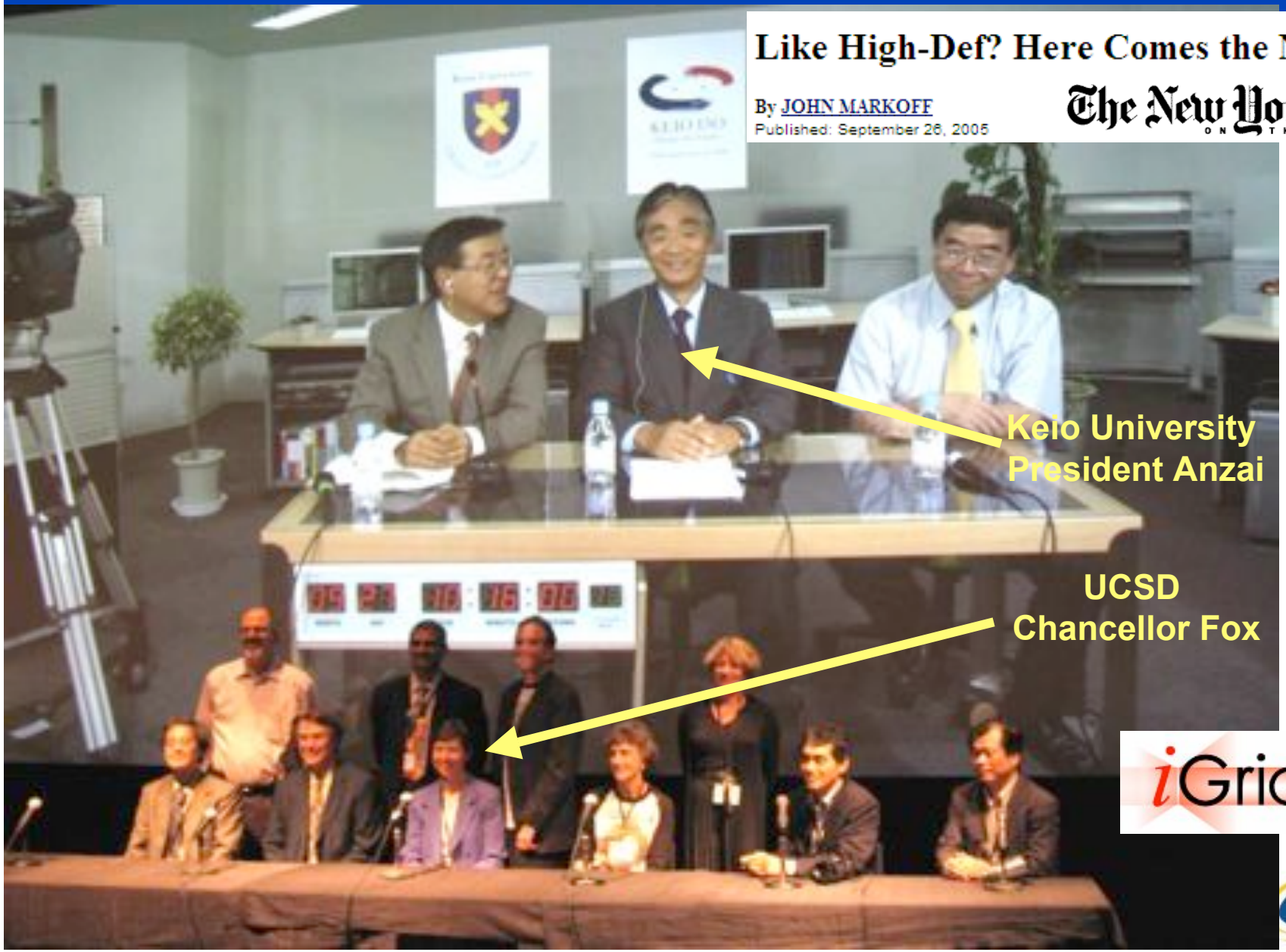






UCSD
UCIrvine

Keio/Calit2 Collaboration: Trans-Pacific 4K Teleconference



Like High-Def? Here Comes the Next Level

By JOHN MARKOFF
Published: September 26, 2005

The New York Times ON THE WEB

Keio University
President Anzai

UCSD
Chancellor Fox

Used
1Gbps
Dedicated

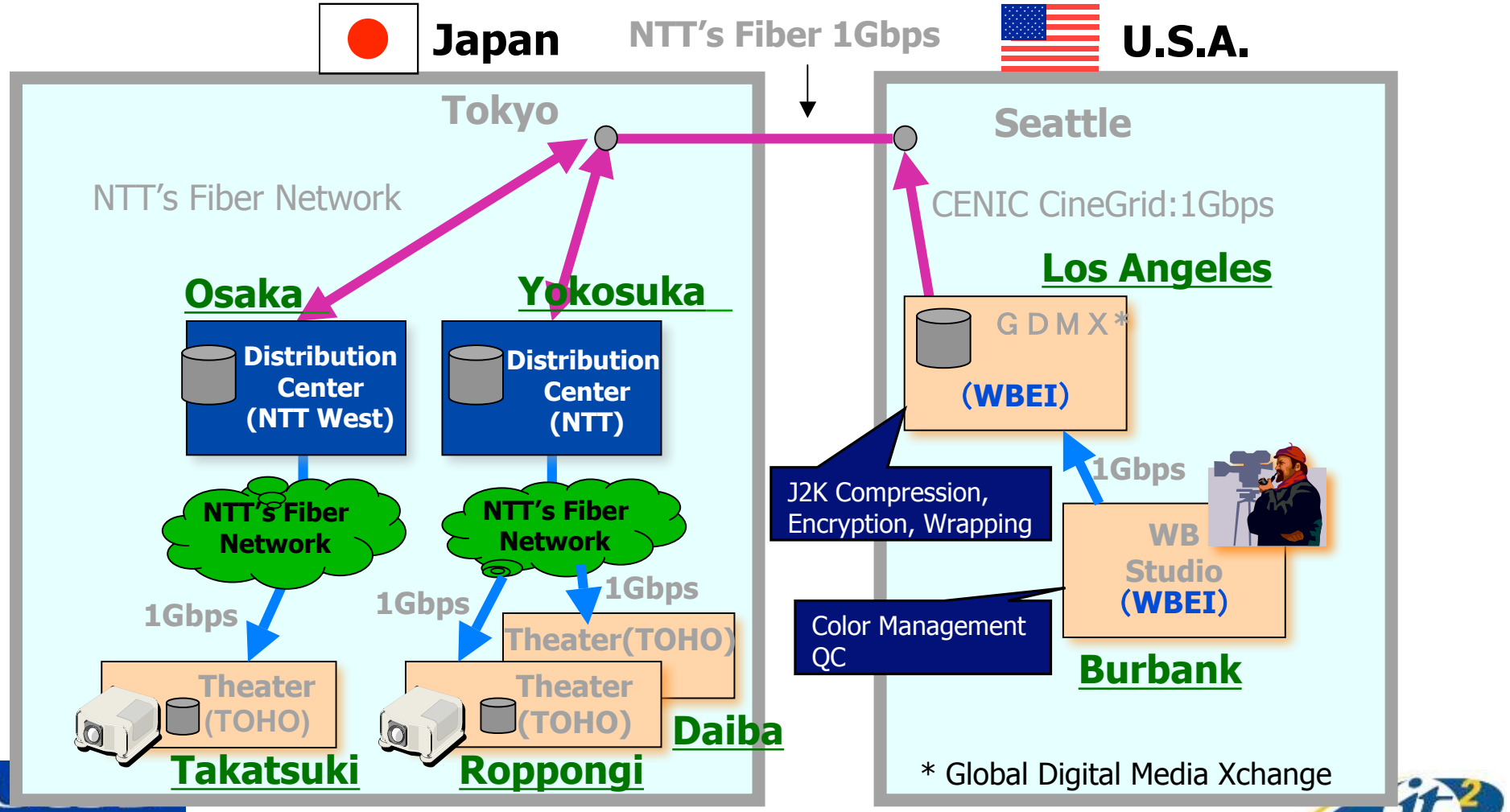
Sony
NTT
SGI

iGrid2005

it²

Network for “4K Pure Cinema” Trial

DCP is directly transferred from GDMX in LA to distribution centers in Japan via fiber network. Within Japan, DCP is distributed from the distribution centers to TOHO theaters. Key is distributed from Osaka center, based on the contract between WB Japan and TOHO cinemas.



UCIrvine

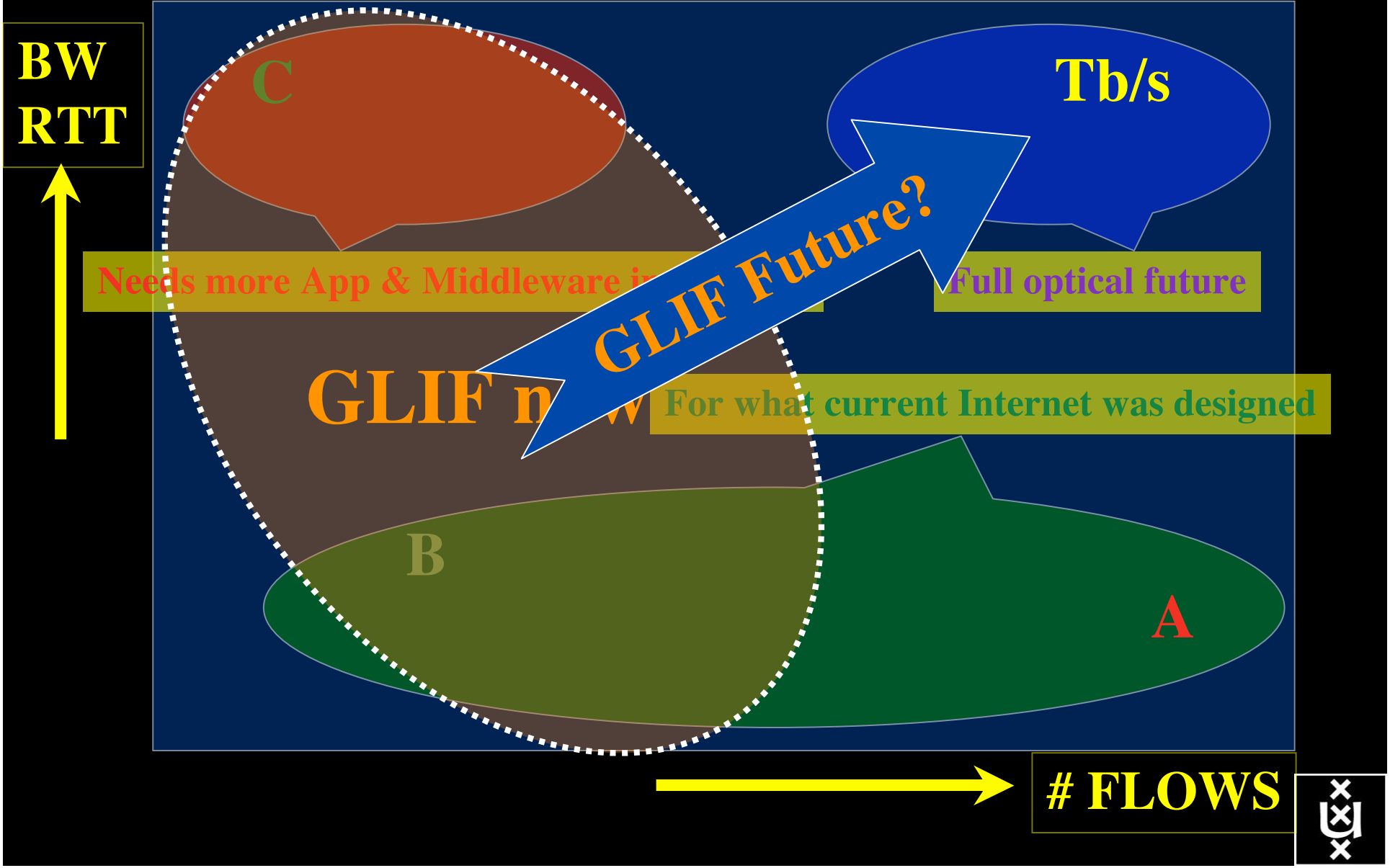
CineGrid@SARA

TeraThinking

- What constitutes a Tb/s network?
- CALIT2 has 8000 Gigabit drops ?->? Terabit Lan?
- look at 80 core Intel processor
  - cut it in two, left and right communicate 8 TB/s
- think back to teraflop computing!
  - MPI makes it a teraflop machine
- massive parallel channels in hosts, NIC's
- TeraApps programming model supported by
  - TFlops -> MPI / Globus
  - TBytes -> OGSA/DAIS
  - TPixels -> SAGE
  - TSensors -> LOFAR, LHC, LOOKING, CineGrid, ...
  - Tbit/s -> ?

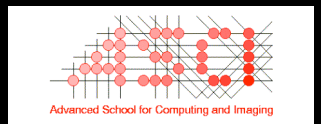
ref Larry Smarr & CdL

Transport of flows



Questions ?


sara
Computing & Networking Services


Advanced School for Computing and Imaging


SURFnet