StarPlane, a GRID application controlled photonic switching network

Cees de Laat

SURFnet
BSIK
EU

University of Amsterdam

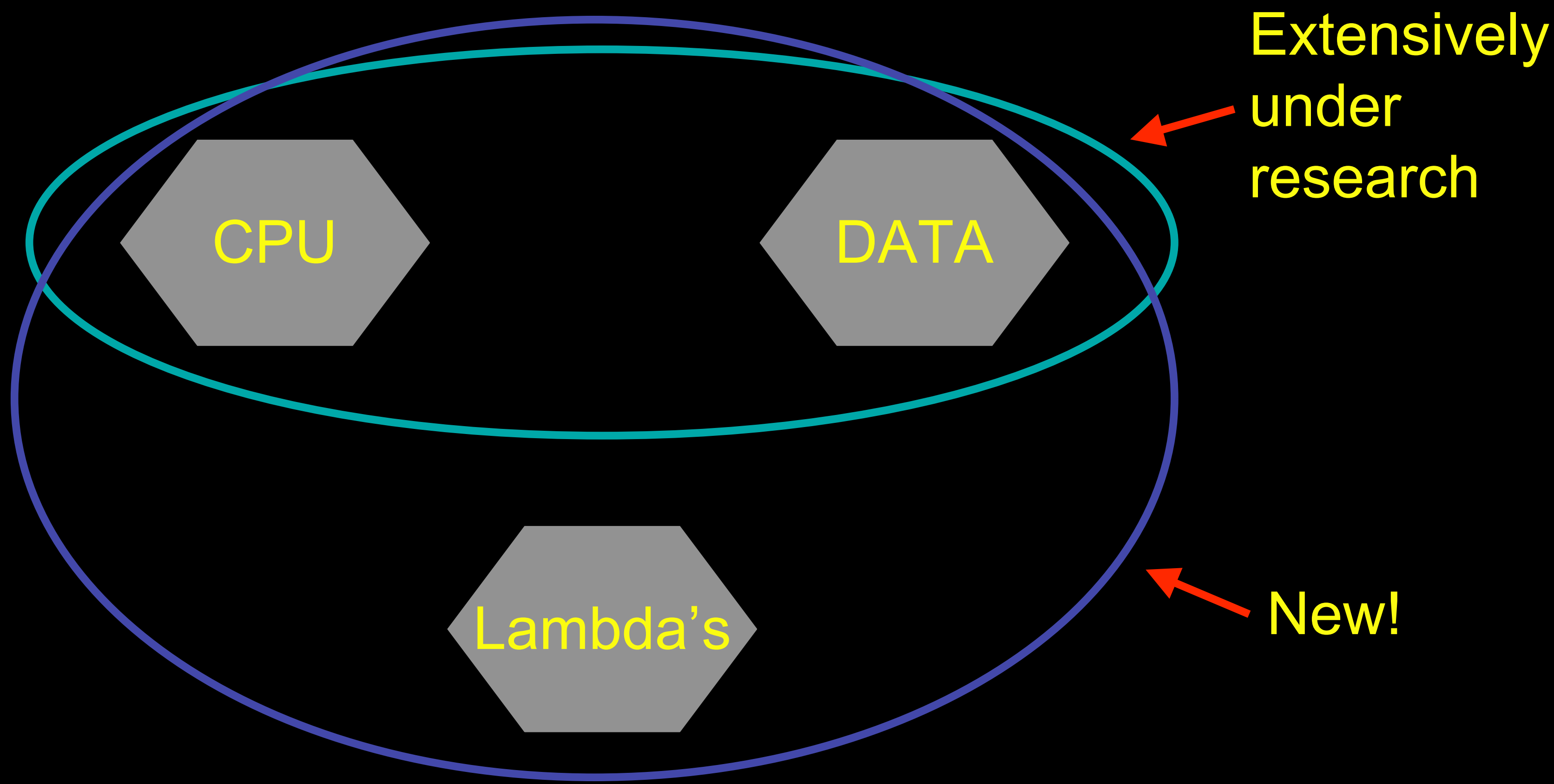
SARA
TI
TNO
NCF

StarPlane Approach

(PG)

- StarPlane is a NWO funded project with major contributions from SURFnet and NORTEL.
- The vision is to allow part of the photonic network infrastructure of SURFnet6 to be manipulated by Grid applications to optimize the performance of specific e-Science applications.
- StarPlane will use the physical infrastructure provided by SURFnet6 and the distributed supercomputer DAS-3.
- The novelty: to give flexibility directly to the applications by allowing them to choose the logical topology in real time, ultimately with subsecond lambda switching times.

GRID-Colocation problem space



StarPlane applications

- Large 'stand-alone' file transfers
  - User-driven file transfers
  - Nightly backups
  - Transfer of medical data files (MRI)

- Large file (speedier) Stage-in/Stage-out
  - MEG modeling
  - Analysis of video data

- Application with static bandwidth requirements
  - Distributed game-tree search
  - Remote data access for analysis of video data
  - Remote visualization

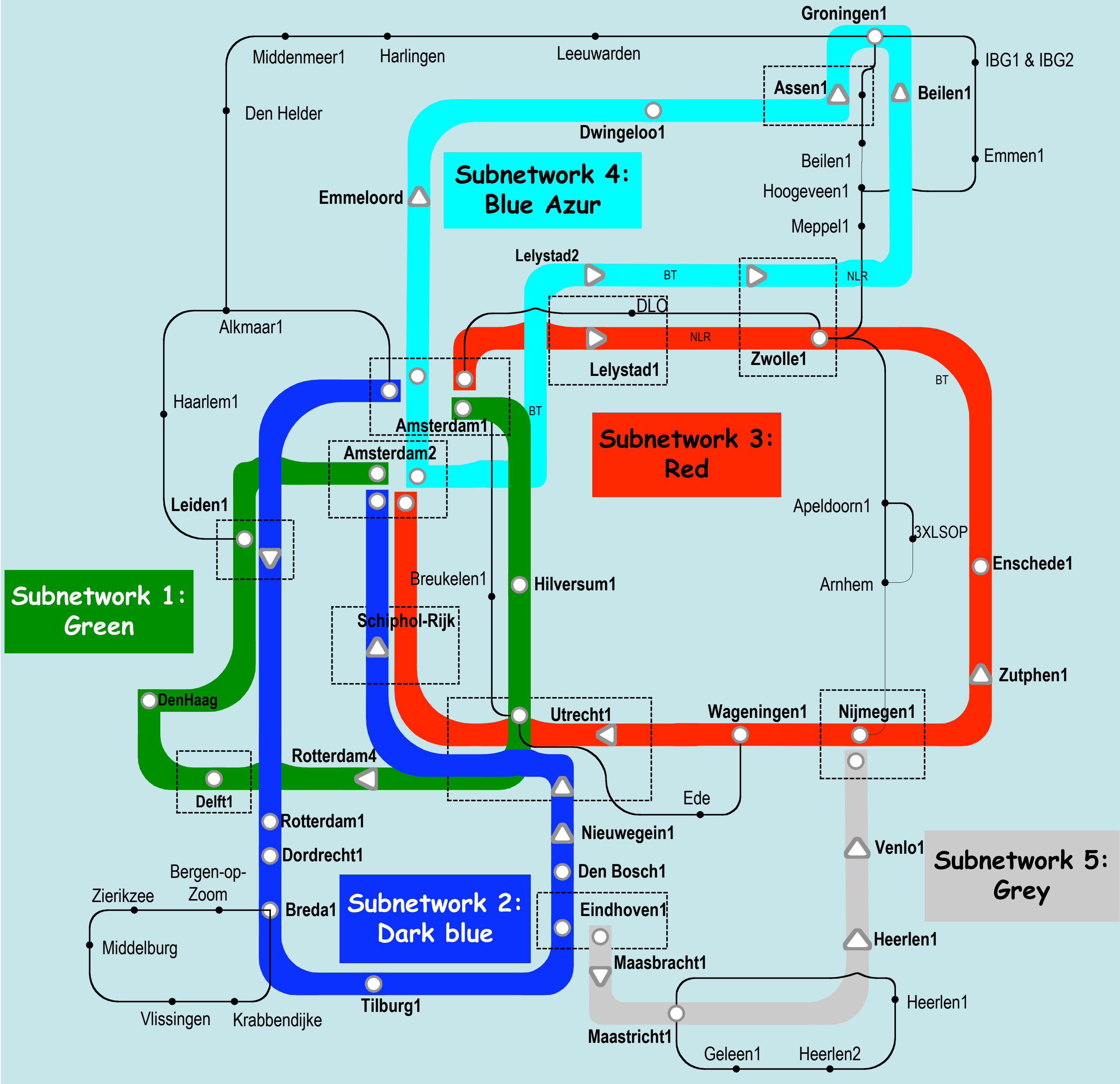
- Applications with dynamic bandwidth requirements
  - Remote data access for MEG modeling
  - SCARI

In The Netherlands SURFnet connects between 180:

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

with a user base of ~750K users

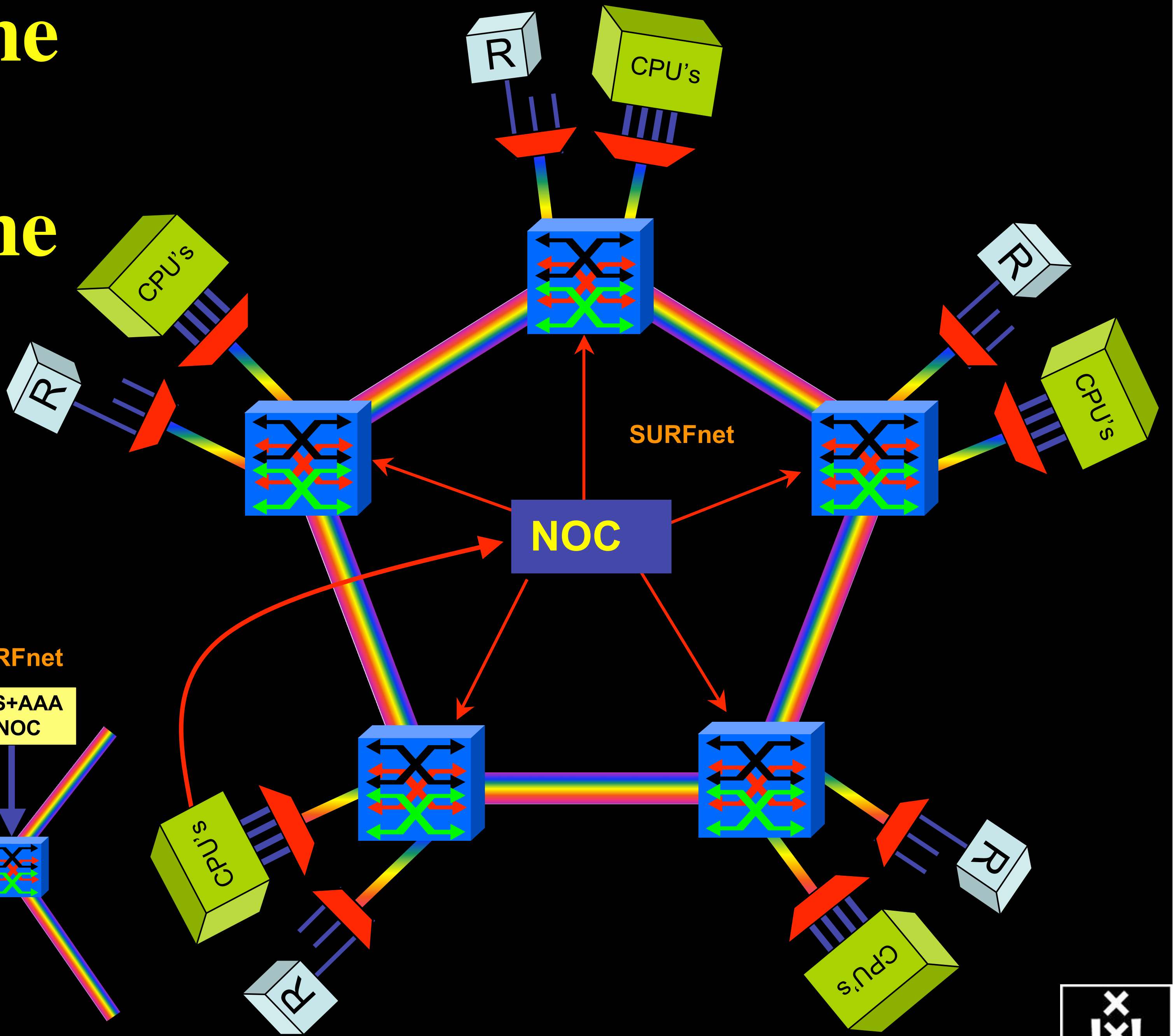
> 6000 km comparable to railway system

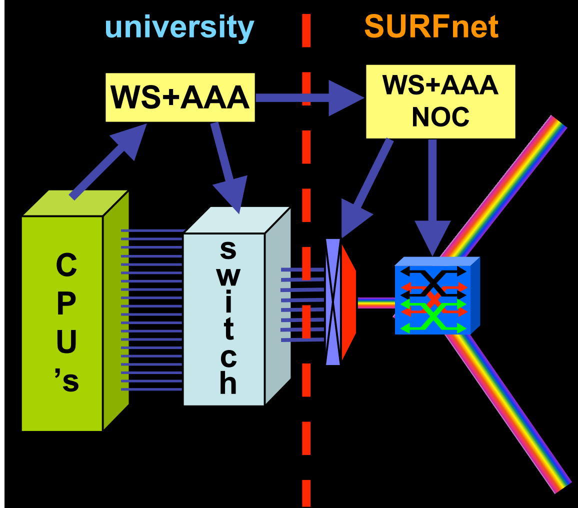


Common Photonic Layer (CPL) in SURFnet6

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

StarPlane DWDM backplane

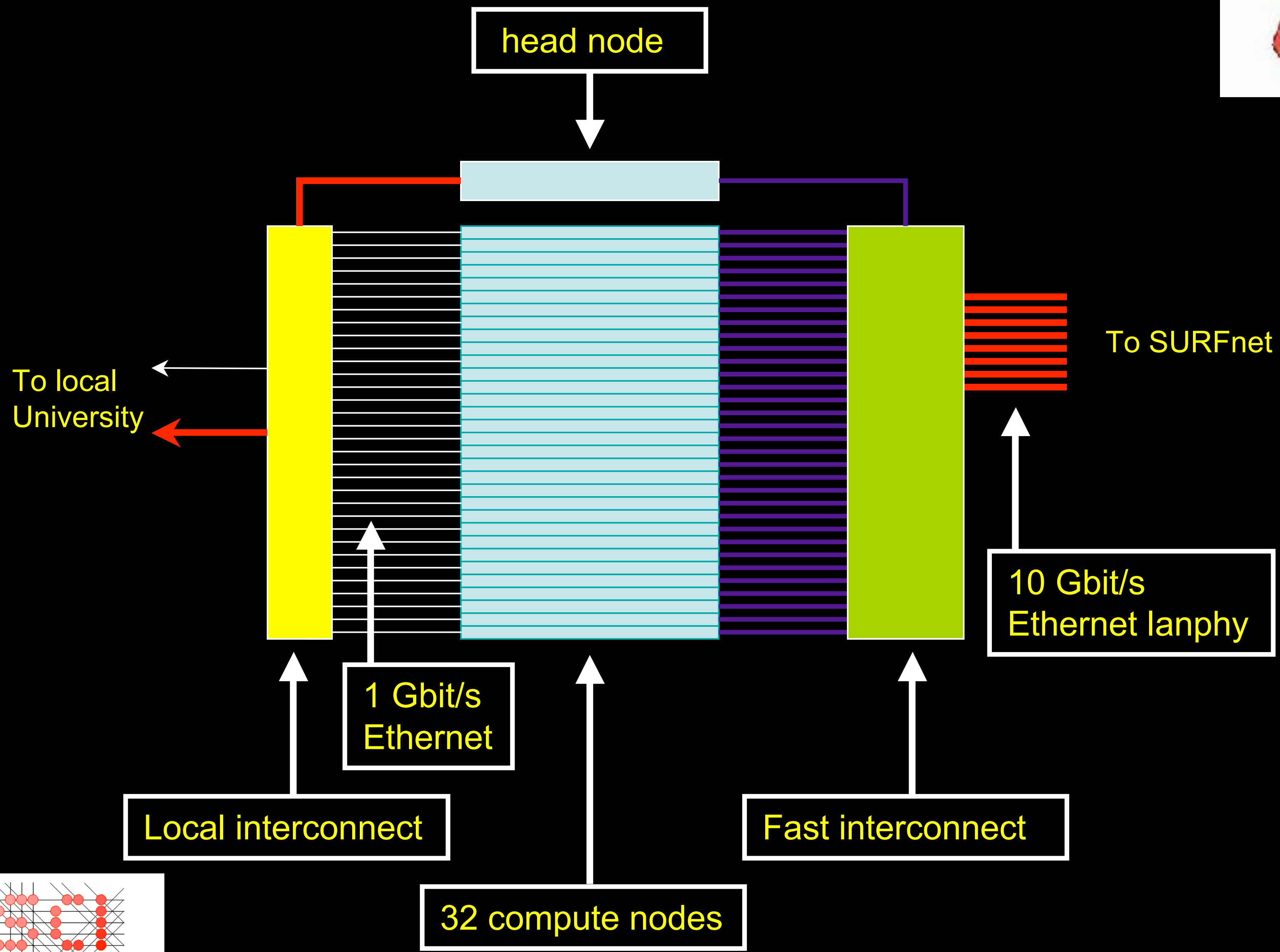




CdL

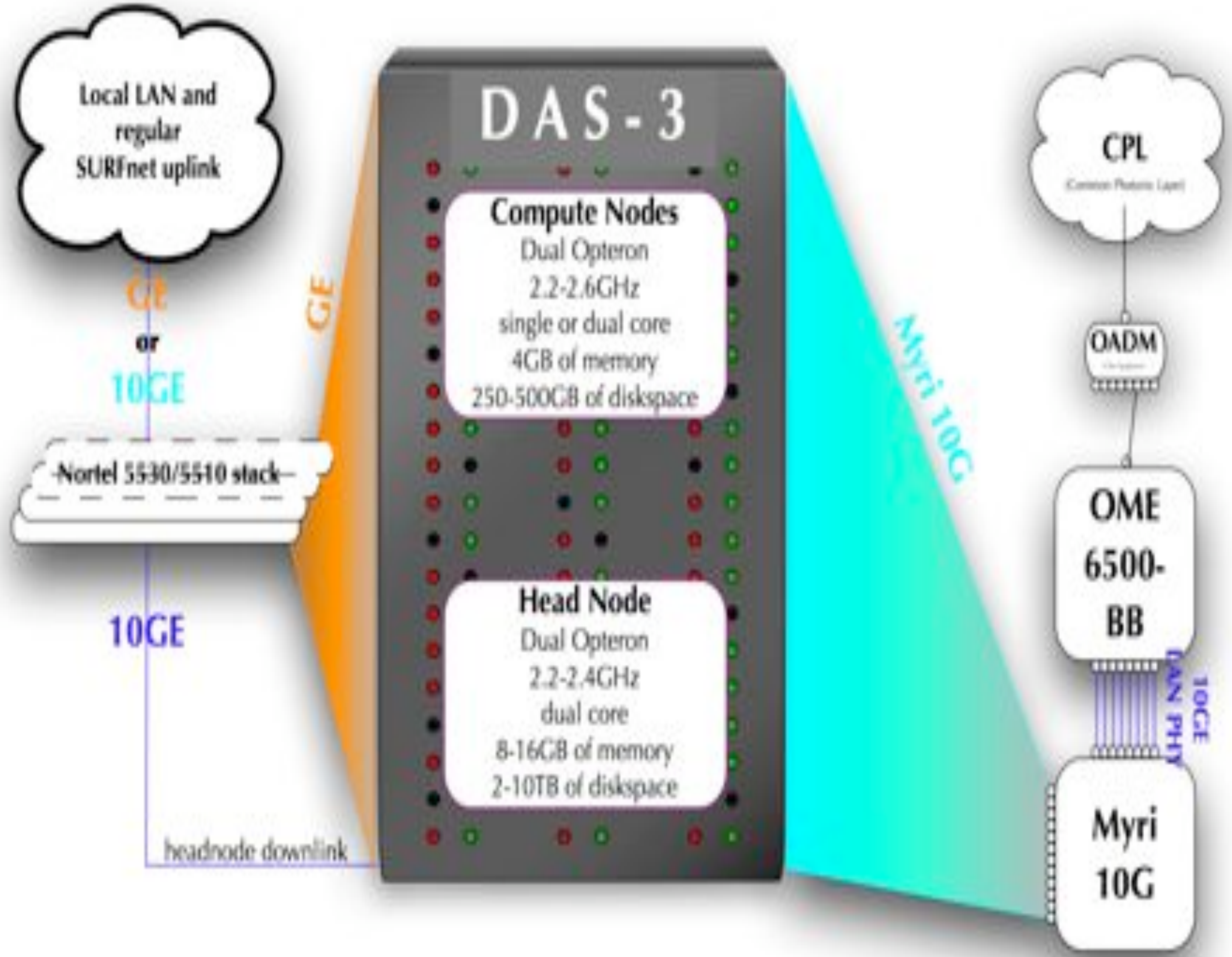
DAS-3 Cluster Tender

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



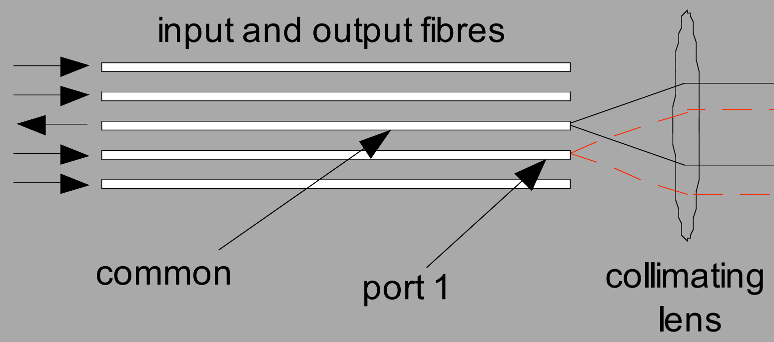
Heterogeneous clusters

	LU	TUD	UvA	UvA-MN	VU	TOTALS
Head						
* storage	10TB	5TB	2TB	2TB	10TB	29TB
* CPU	2x2.4GHz DC	2x2.4GHz DC	2x2.2GHz DC	2x2.2GHz DC	2x2.4GHz DC	
* memory	16GB	16GB	8GB	16GB	8GB	64GB
* Myri 10G	1		1	1	1	
* 10GE	1	1	1	1	1	
Compute	32	68	40 (1)	46	85	271
* storage	400GB	250GB	250GB	2x250GB	250GB	84TB
* CPU	2x2.6GHz	2x2.4GHz	2x2.2GHz DC	2x2.4GHz	2x2.4GHz DC	1.9THz
* memory	4GB	4GB	4GB	4GB	4GB	1048GB
* Myri 10G	1		1	1	1	
Myrinet						
* 10G ports	33 (7)		41	47	86 (2)	
* 10GE ports	8		8	8	8	320
Nortel						
* 1GE ports	32 (16)	136 (8)	40 (8)	46 (2)	85 (11)	339
* 10GE ports	1 (1)	9 (3)	2	2	1 (1)	



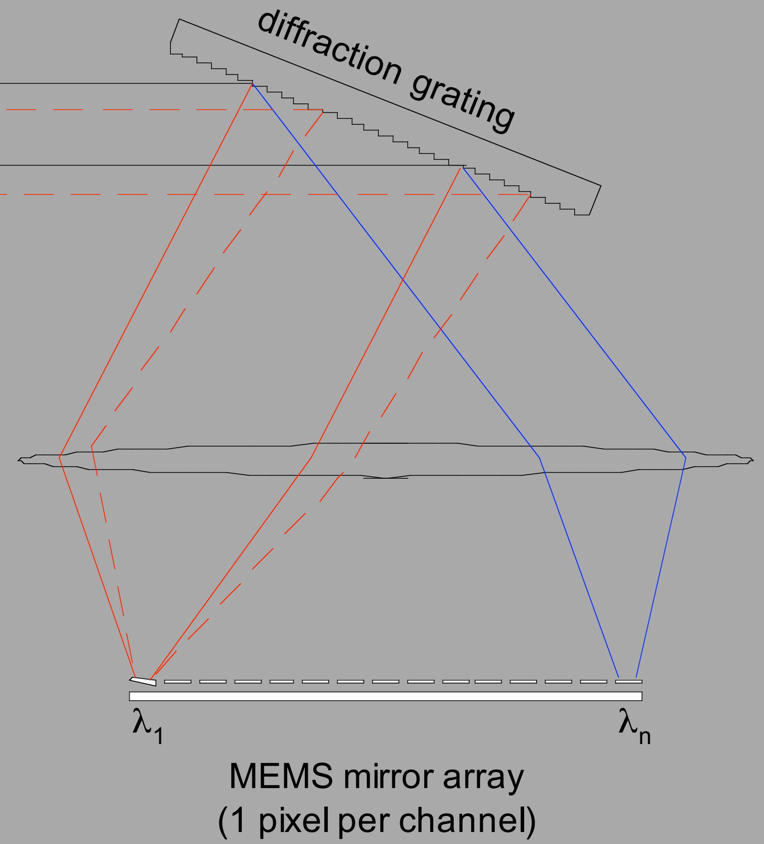
Local LAN and regular SURFnet uplink
GE
or
10GE
Nortel 5530/5510 stack
10GE
headnode downlink
GE
DAS-3
Compute Nodes
Dual Opteron
2.2-2.6GHz
single or dual core
4GB of memory
250-500GB of diskspace
Head Node
Dual Opteron
2.2-2.4GHz
dual core
8-16GB of memory
2-10TB of diskspace
Myri 10G
CPL
OADM
OME 6500-BB
10GE
LAN PHY
Myri 10G

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

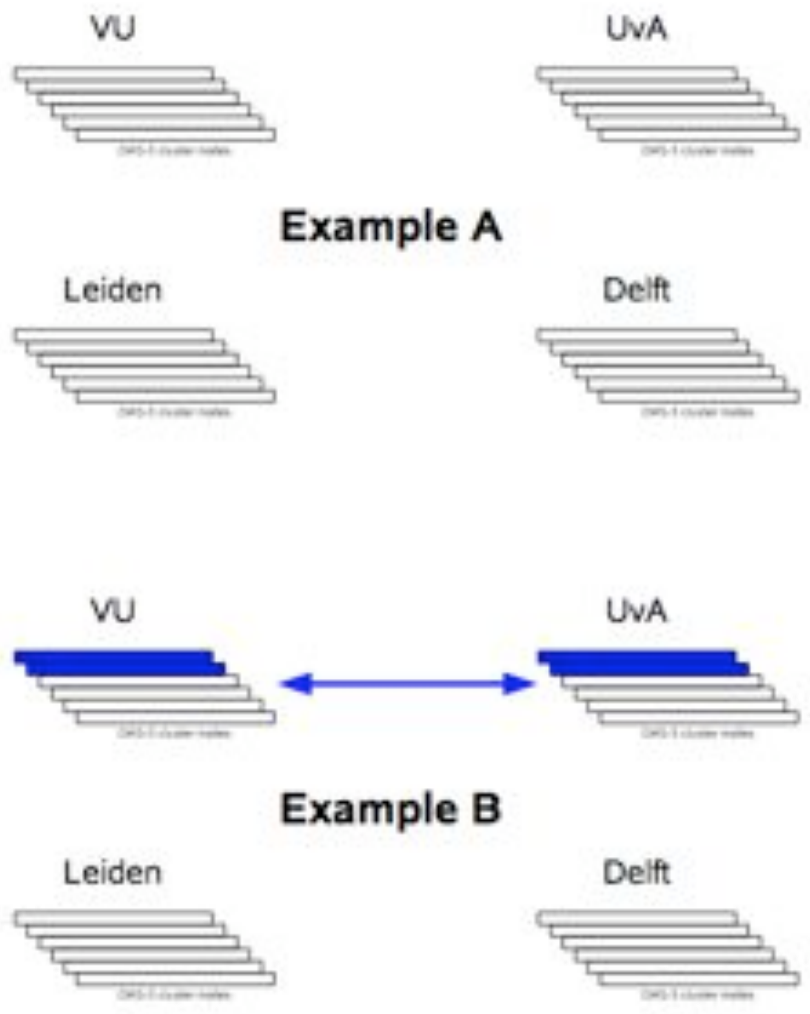


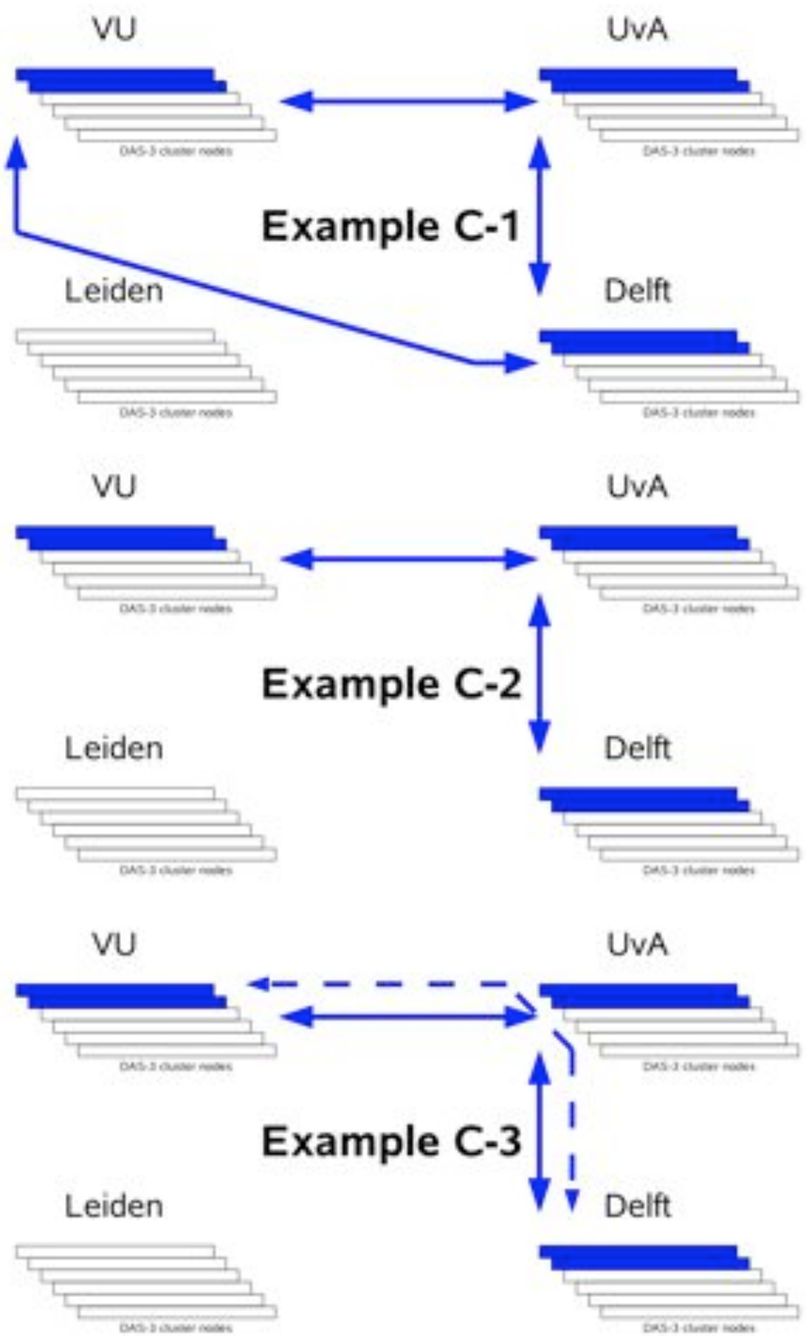
ref Eric Bernier, NORTEL

What makes StarPlane possible

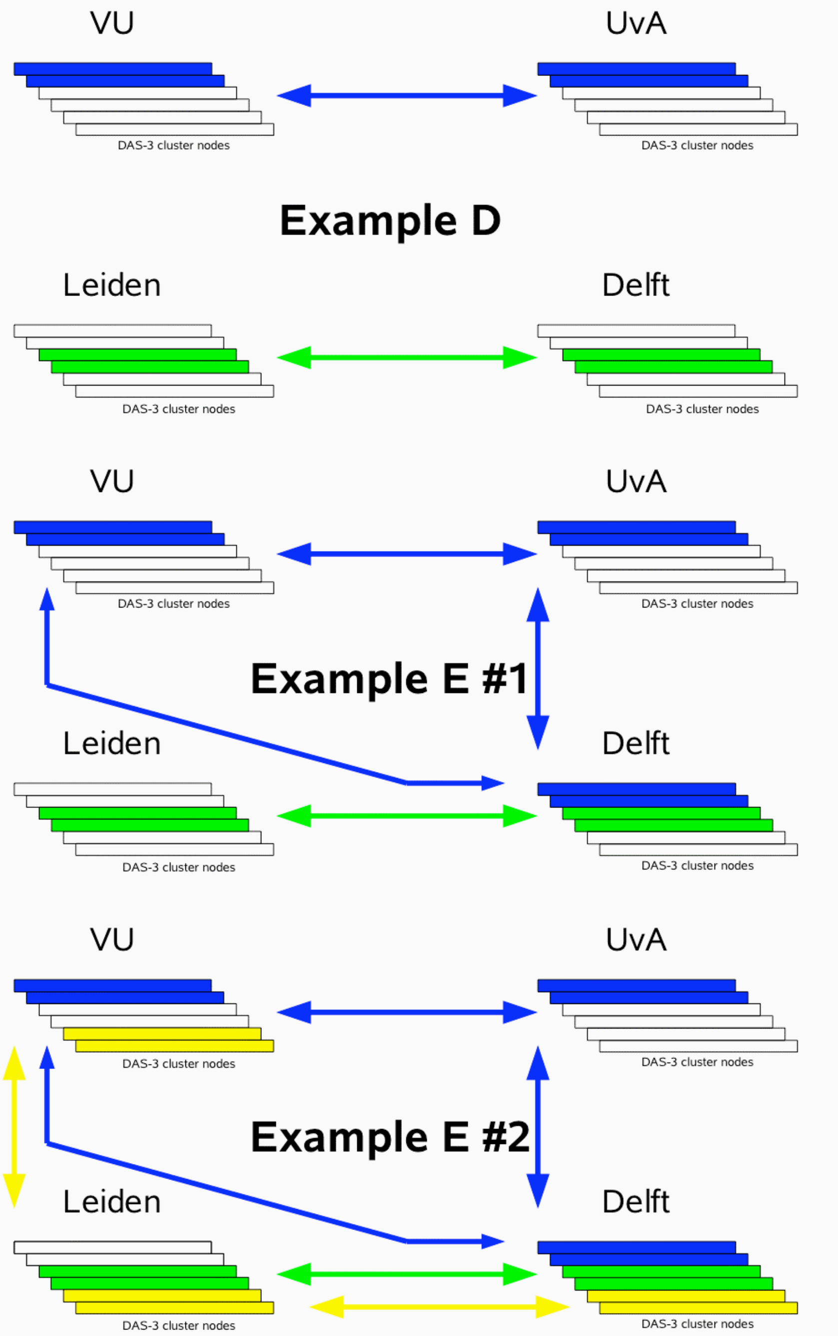
- Wavelength Selective Switches
- Sandbox by confining StarPlane to a band
- Optimization of the controls to turn on/off a Lambda
- electronic Dynamically Compensating Optics (eDCO)
- traffic engineering

Traffic engineering





More traffic engineering



Even more traffic engineering

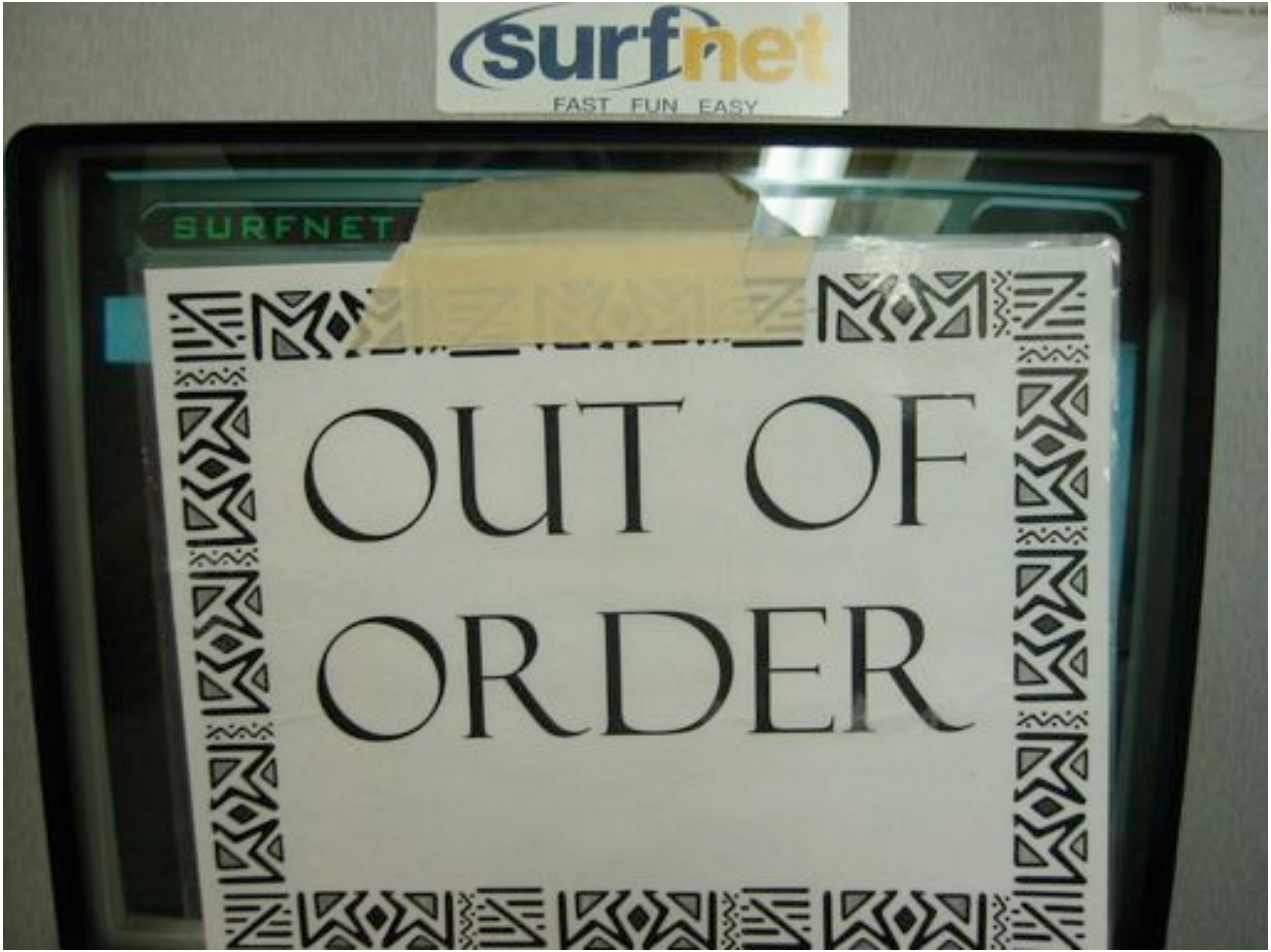
Risks

what have we today

what to avoid

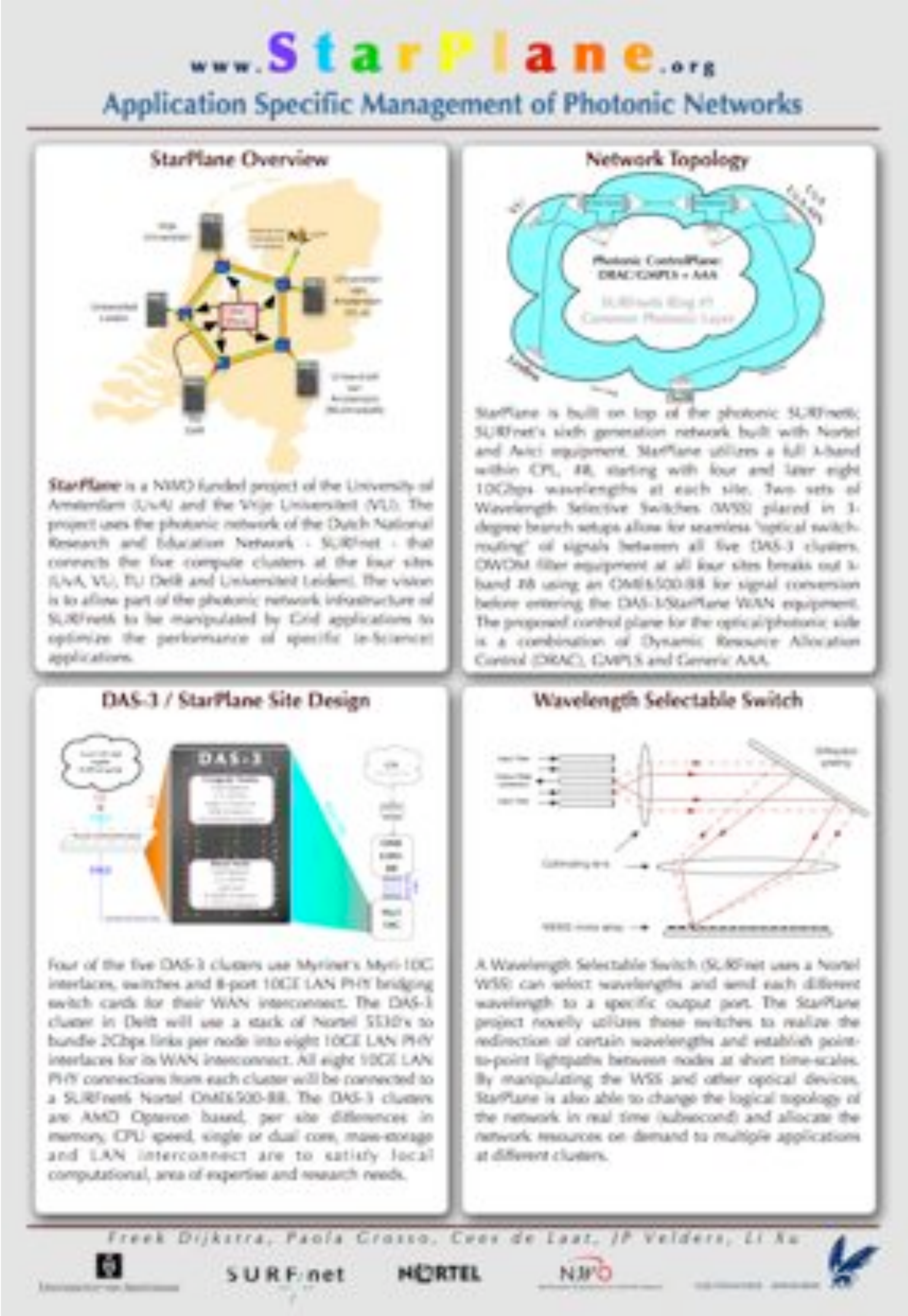

surfnet
FAST FUN EASY
SURFNET PREMIERE
HELP
Three Easy Steps :
Click the START button
Insert money...
$0.25 per minute...
Example :
$1 = 4 minutes
$5 = 20 minutes
No change is provided!
Surf the web!


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


surfnet
FAST FUN EASY
SURFNET
OUT OF
ORDER

What do we need

- vlan's
- trunking
- spanning tree modified?
- mac in mac?
- source routing modified
- Policy interfaces
- AAA interaction (EduRoam, Shibboleth)



Visit poster session

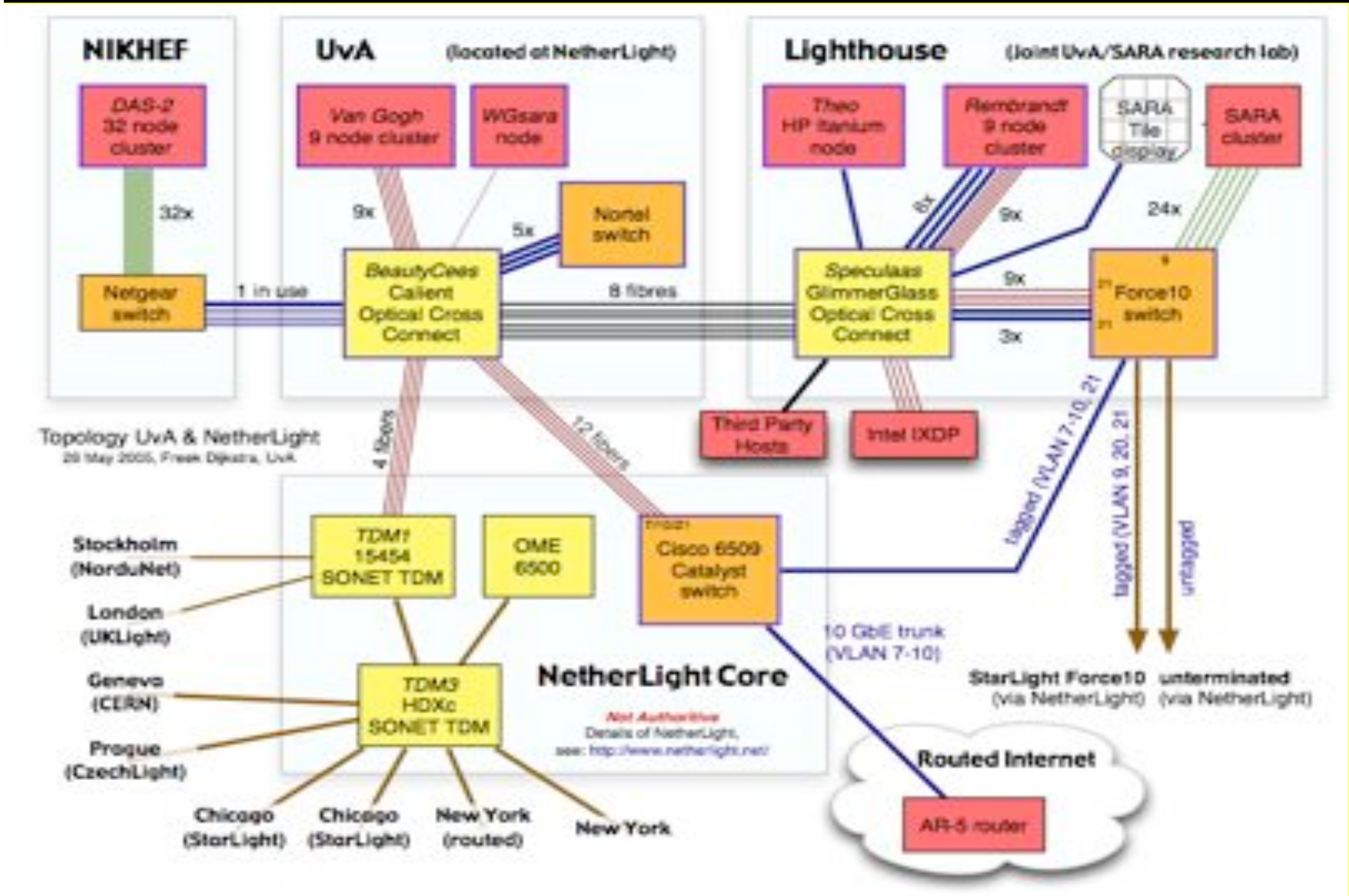
Key issue #1:
how to describe such networks?

UvA/SARA LightHouse

A joint network research lab of the University of Amsterdam and SARA.

Connects end resources to NetherLight.

Proof of concept e.g. tier 0/1, webservices, GSP



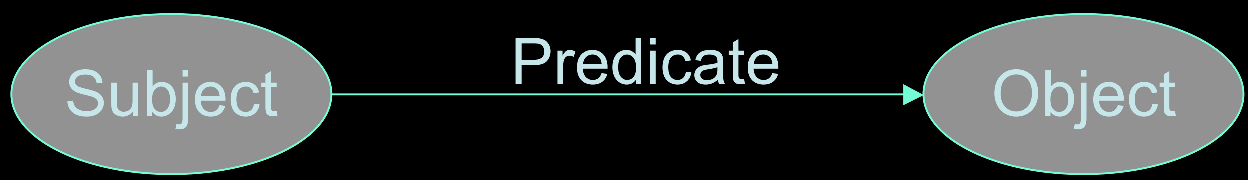
sara
Computing & Networking Services

Semantic web

“a universal medium for the exchange of data where data can be shared and processed by automated tools as well as by people”

The Resource Description Framework (RDF) uses XML as an interchange syntax.

Data is described by triplets:



NDL - Network Description Language

A way to describe network resources using RDF.

Parser can use the data to:

- generate network maps
- provide information to schedulers

```
<ndl:Device rdf:about="#Vangogh3">
    <ndl:name>Vangogh3</ndl:name>
    <rb:isOfType>ComputingElement</rb:isOfType>
   <ndl:locatedAt rdf:resource="#Lighthouse"/>
   <ndl:hasInterface rdf:resource="#Vangogh3:eth2"/>
</ndl:Device>
```

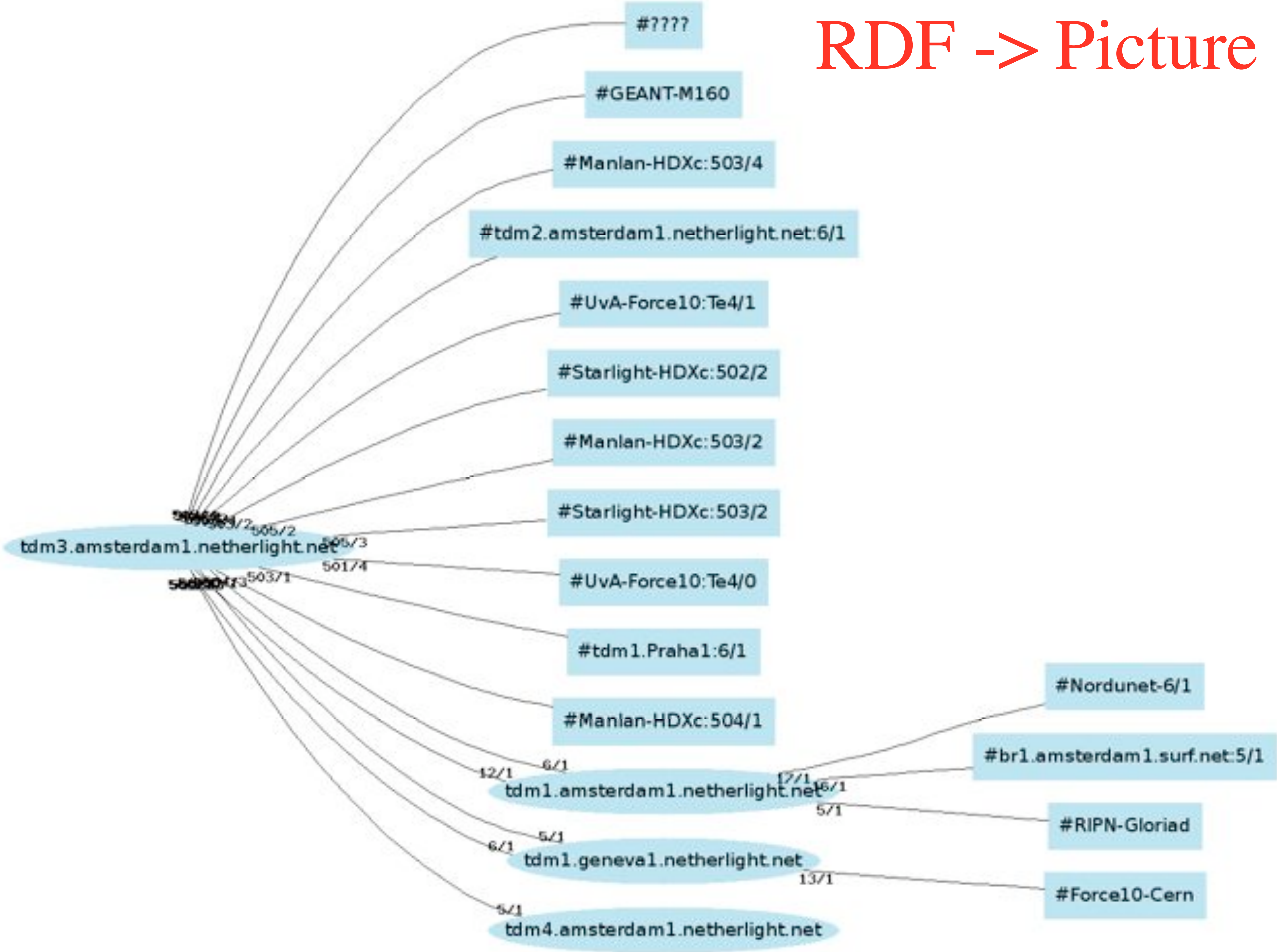
NetherLight in RDF

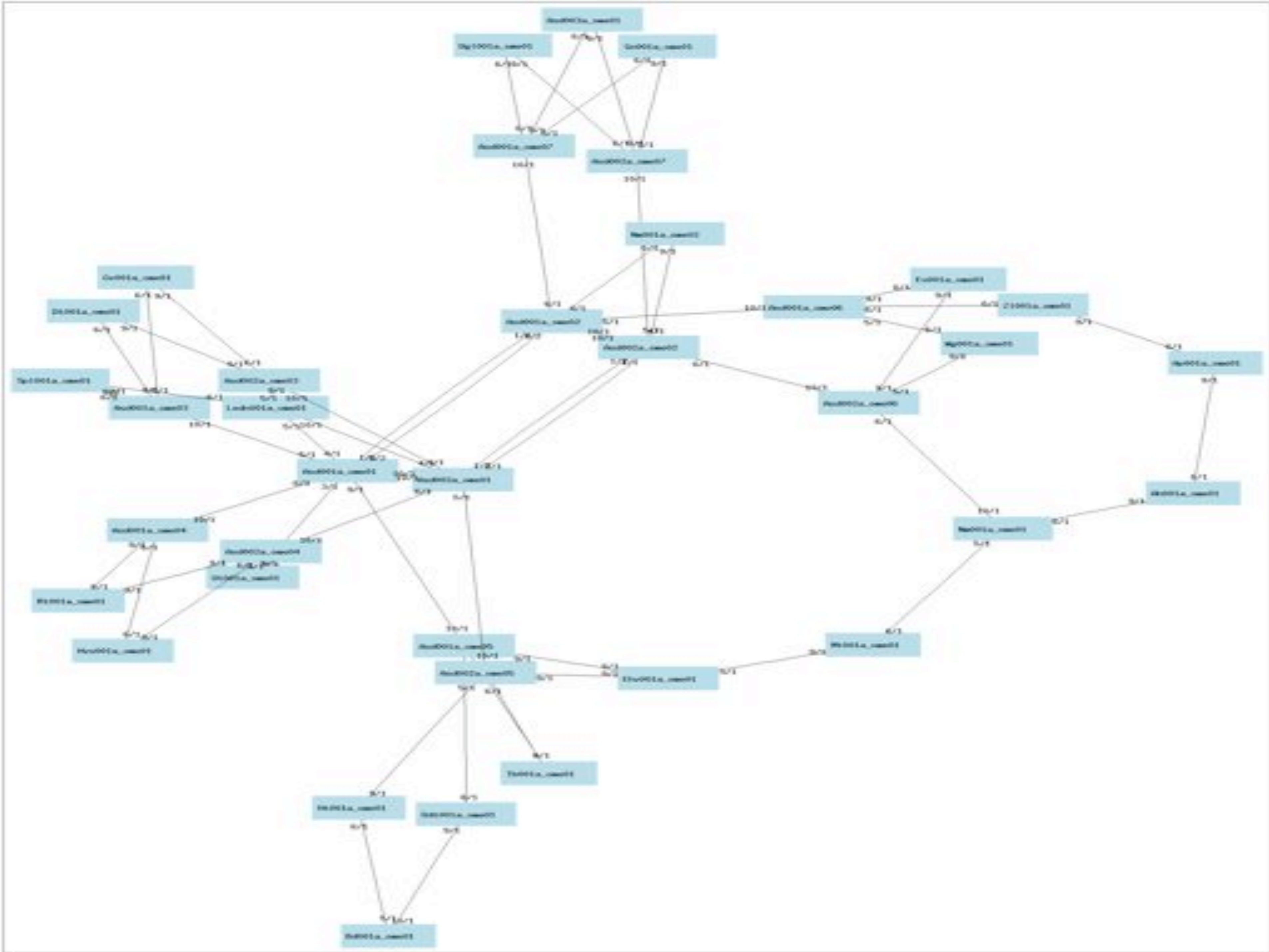
```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
      xmlns:ndl="http://www.science.uva.nl/research/air/ndl#">
<!-- Description of Netherlight -->
<ndl:Location rdf:about="#Netherlight">
      <ndl:name>Netherlight Optical Exchange</ndl:name>
</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
```

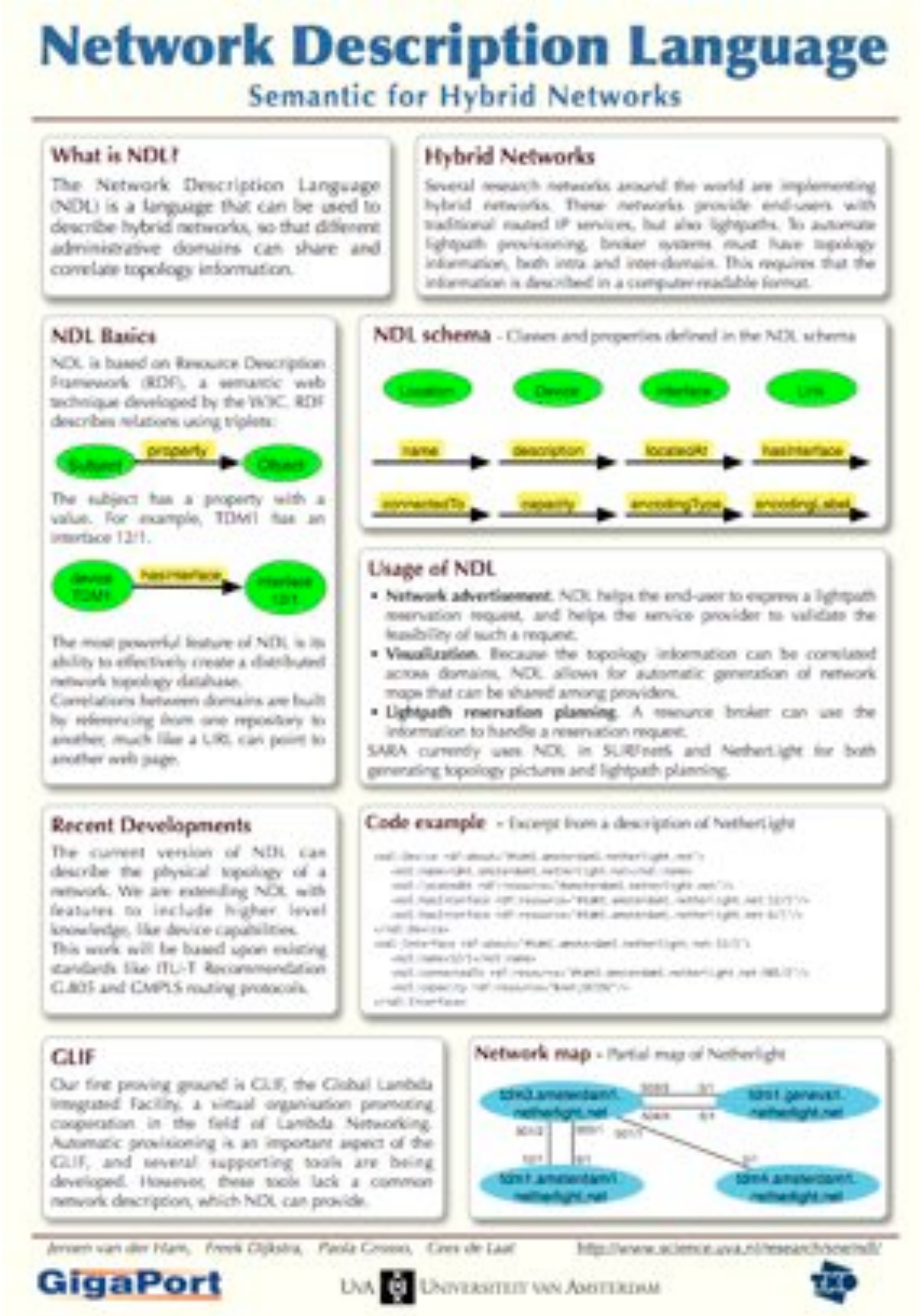
```
<!-- all the interfaces of TDM3.amsterdam1.netherlight.net -->

<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>
```

RDF -> Picture


#????
#GEANT-M160
#Manlan-HDXc:503/4
#tdm2.amsterdam1.netherlight.net:6/1
#UvA-Force10:Te4/1
#Starlight-HDXc:502/2
#Manlan-HDXc:503/2
#Starlight-HDXc:503/2
tdm3.amsterdam1.netherlight.net
505/2
505/3
501/4
503/1
#UvA-Force10:Te4/0
#tdm1.Praha1:6/1
#Manlan-HDXc:504/1
#Nordunet-6/1
#br1.amsterdam1.surf.net:5/1
12/1
6/1
17/1
16/1
tdm1.amsterdam1.netherlight.net
5/1
#RIPN-Gloriad
5/1
6/1
tdm1.geneva1.netherlight.net
13/1
#Force10-Cern
5/1
tdm4.amsterdam1.netherlight.net





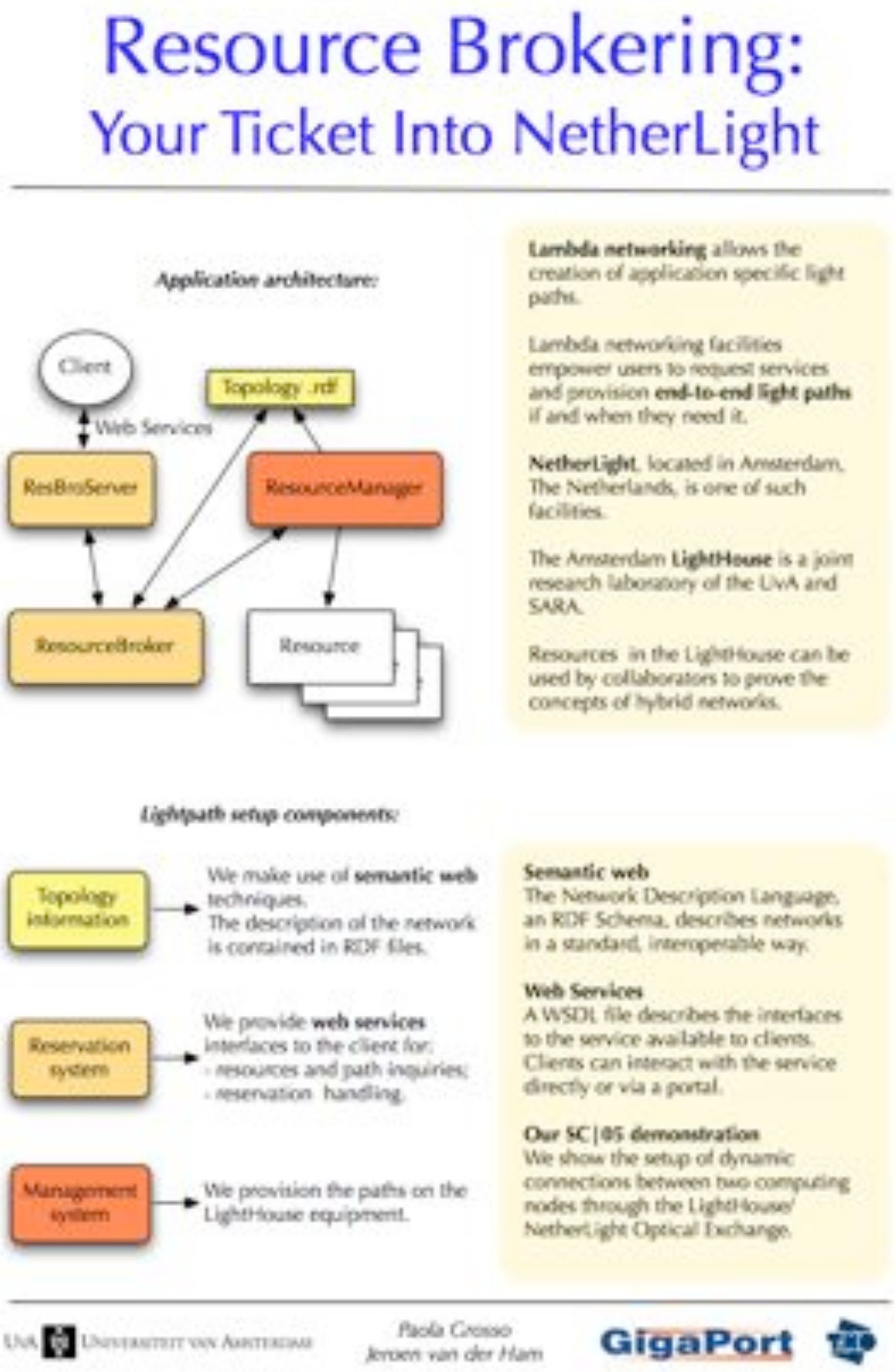
Visit poster session

Key issue #2:
How to book resources on such networks?

Web services

Web services interfaces provide the API for the reservation framework:

<wsdl:operation name="getResourceInformation">
<wsdl:operation name="getResourceList">
<wsdl:operation name="getTypeList">
<wsdl:operation name="getResourcesOfType">
<wsdl:operation name="reservePath">
<wsdl:operation name="getPossiblePaths">
<wsdl:operation name="isPathAvailable">
<wsdl:operation name="confirmPathReservation">
<wsdl:operation name="cancelPathReservation">



Poster @ SC2005
ask
Paola Grosso

Questions ?

(lesson learned)

Credits:

- Leon Gommans, Paola Grosso, Marten Hoekstra, Arie Taal, Freek Dijkstra, Bert Andree, Jeroen van der Ham, Hans Blom, Yuri Demchenko, Fred Wan, Karst Koymans, Martijn Steenbakkers Jaap van Ginkel, Li Xu
- SURFnet / GigaPort, Kees Neggers, Erik-Jan Bos, et al!
- NORTEL: Franco Travostino, Kim Roberts, Rod Wilson
- SARA: Anwar Osseryan, Paul Wielinga, Pieter de Boer, Ronald van der Pol, teams
- Joe Mambretti, Bill stArnaud, GLIF community
- Tom & Maxine & Larry, Laurin, OptIPuter, OnVector team !!!!

sara Computing & Networking Services

Advanced School for Computing and Imaging

SURFnet