

StarPlane, a GRID application controlled photonic switching network

Cees de Laat

SURFnet

BSIK

EU

University of Amsterdam

SURFnet



SARA
TI
TNO
NCF



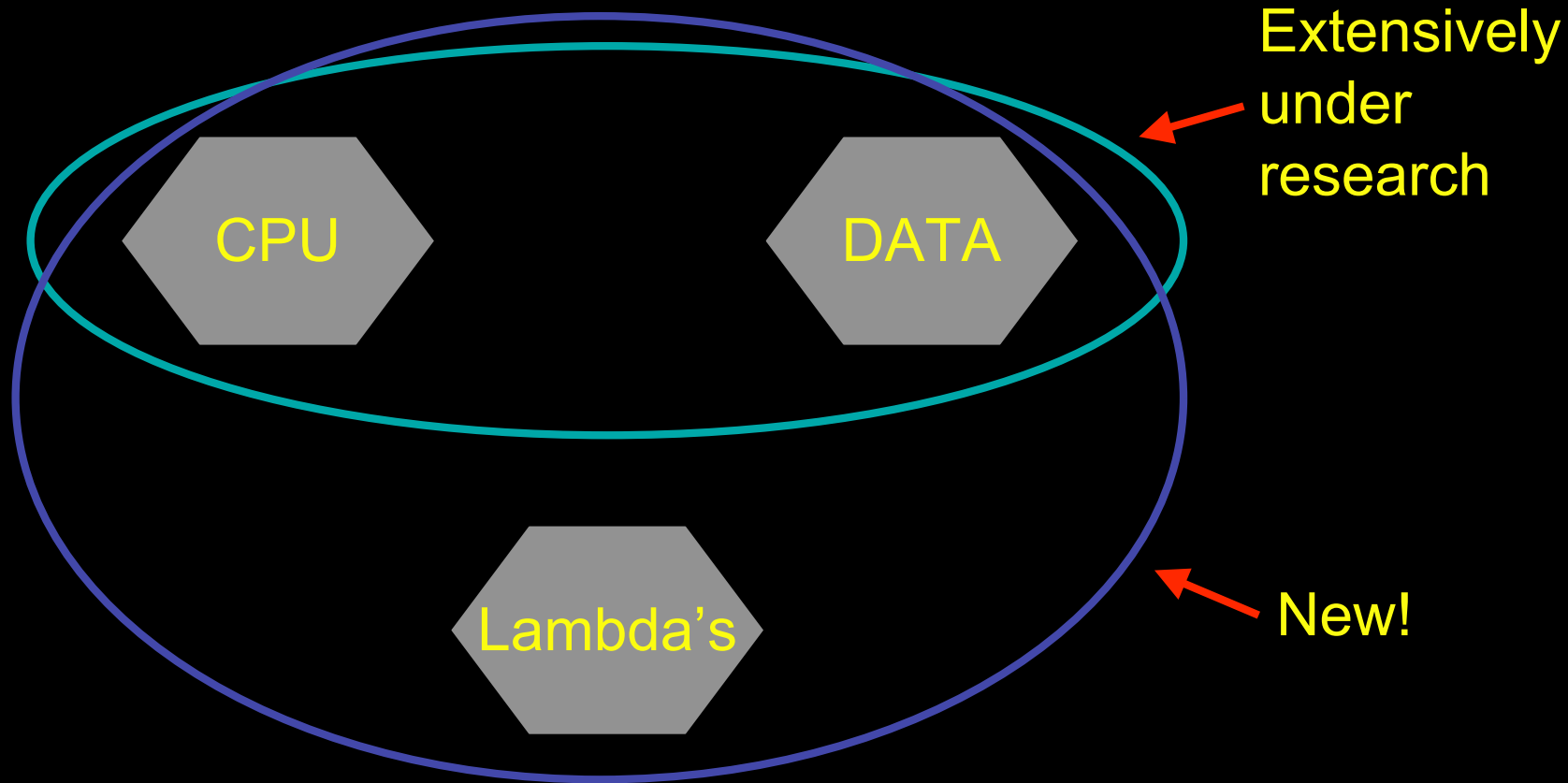
StarPlane Approach

(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.

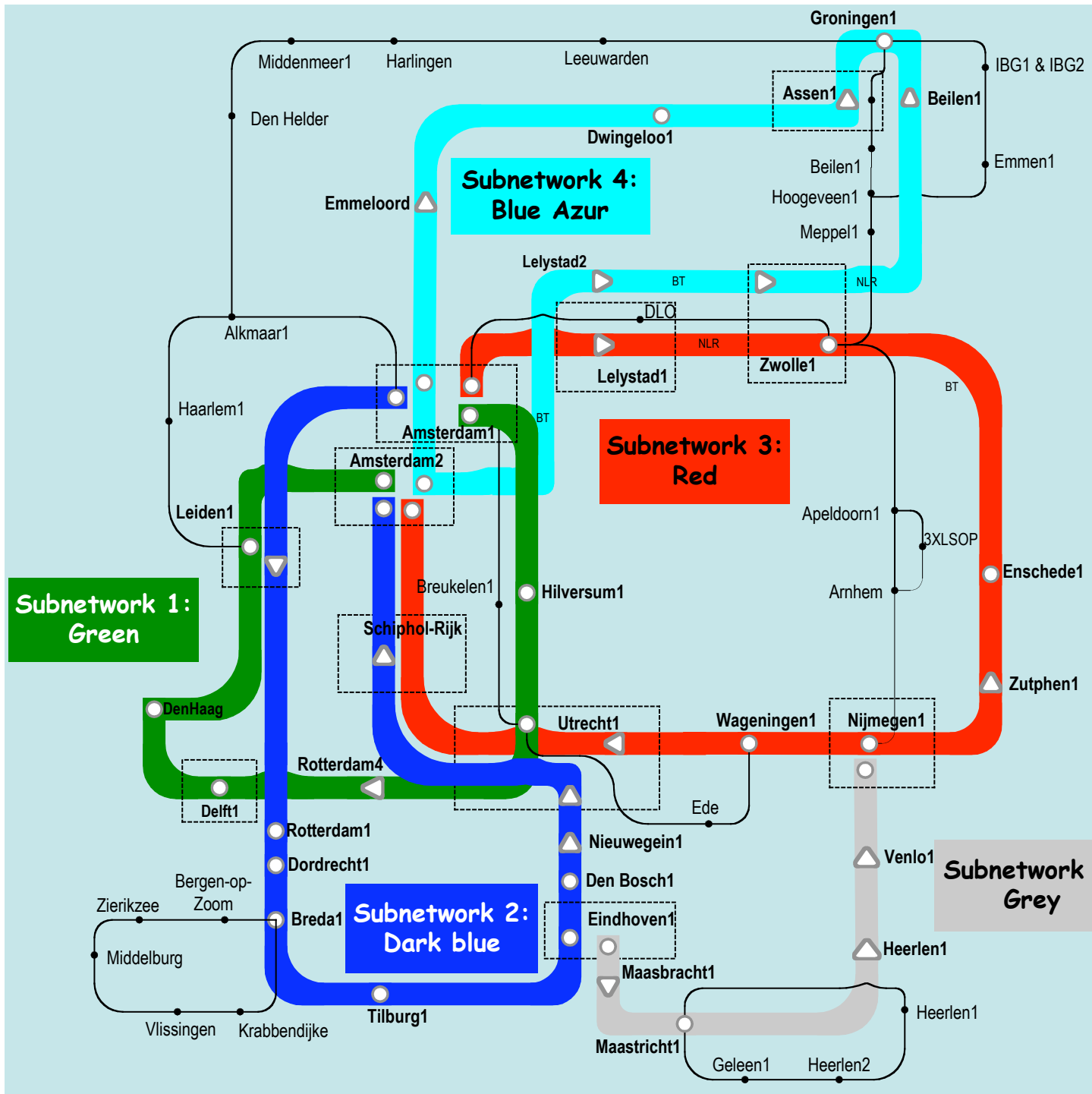
with a user base of ~750K users

> 6000 km
comparable
to railway
system

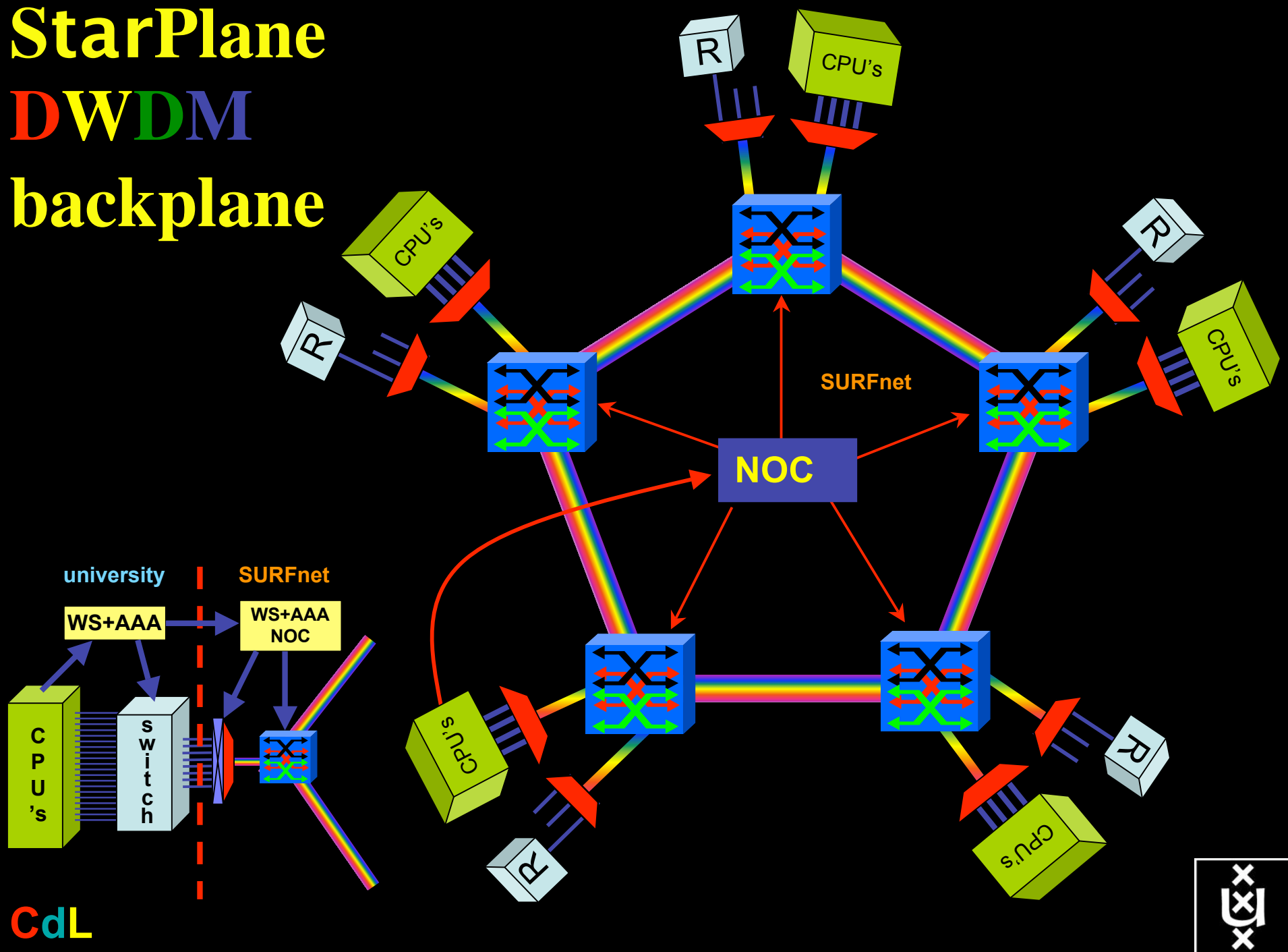


Common Photonic Layer (CPL) in SURFnet6

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

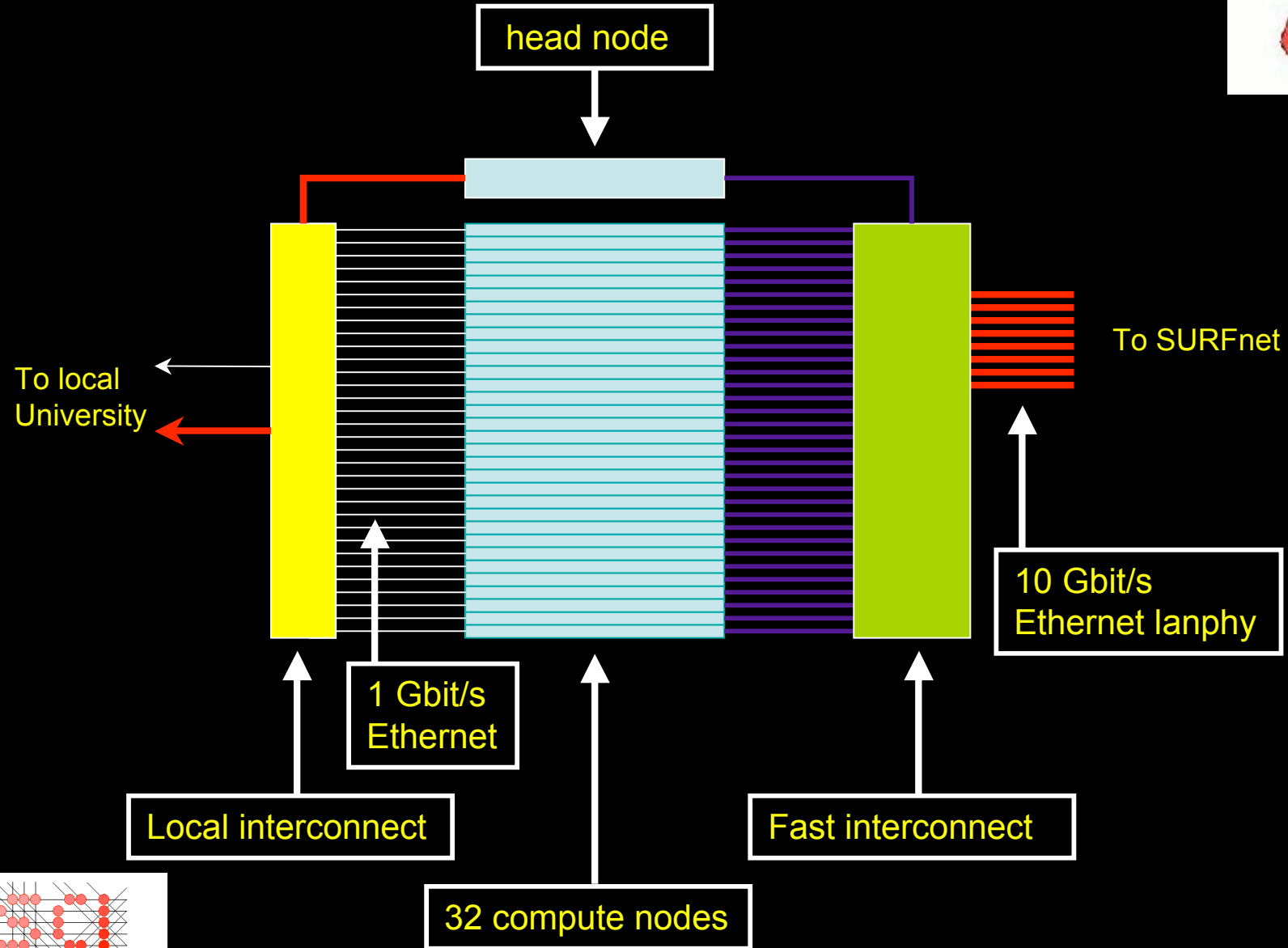


StarPlane DWDM backplane



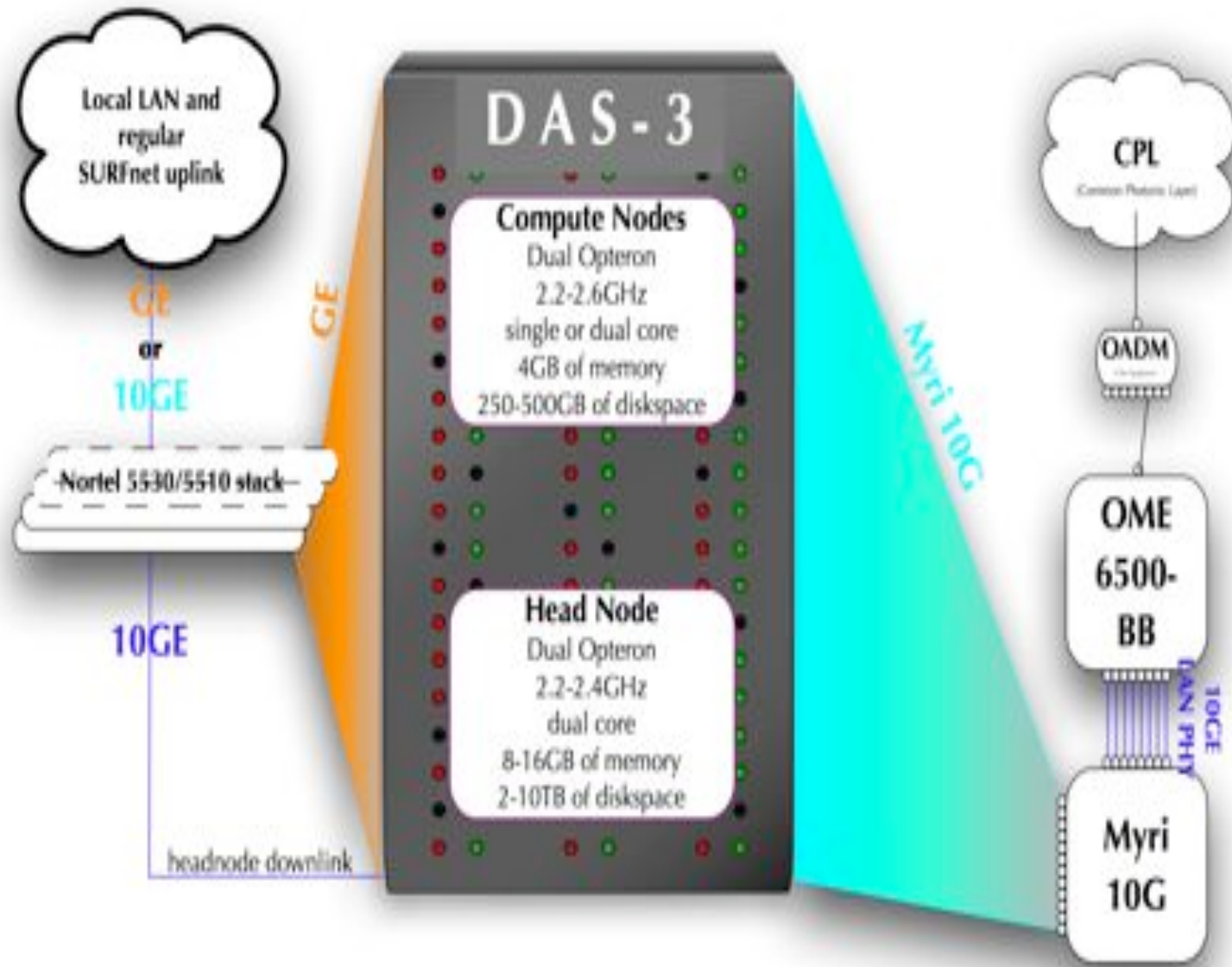
DAS-3 Cluster Tender

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

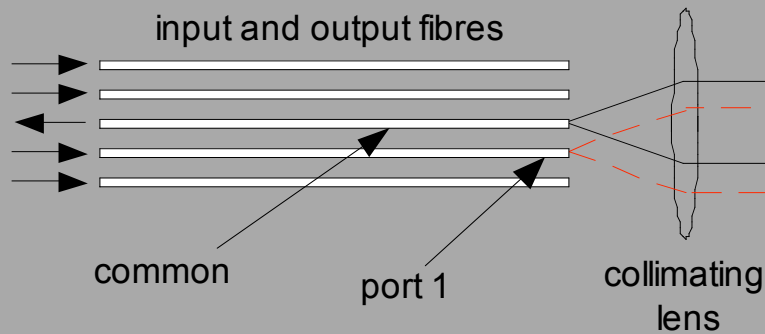


Heterogeneous clusters

	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)	



Module Operation

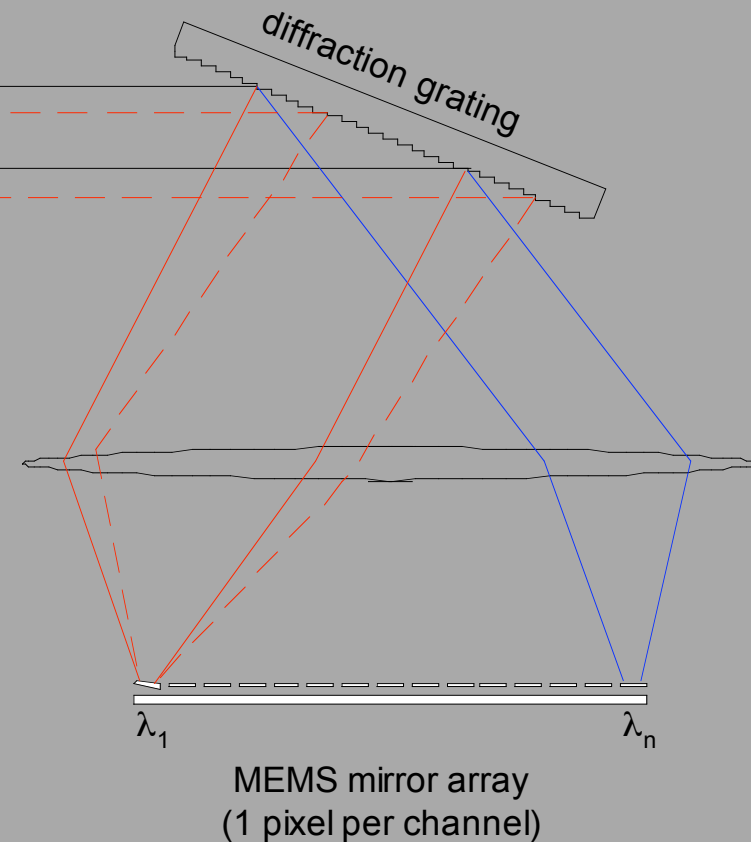


> this schematic shows

- 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

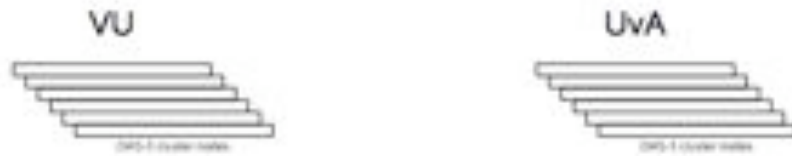


What makes StarPlane possible

- 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

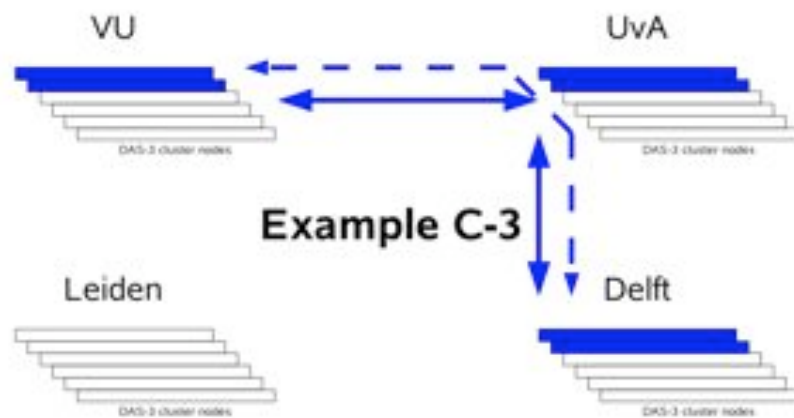
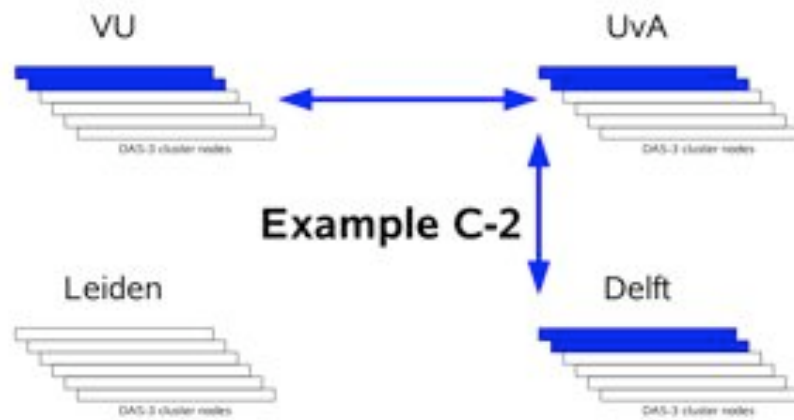
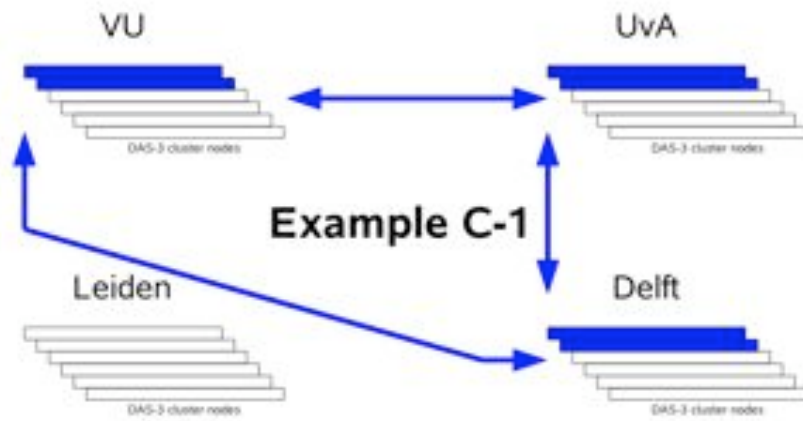


Example A



Example B

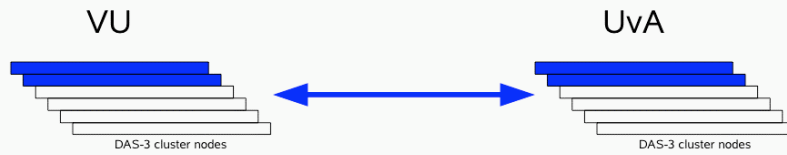




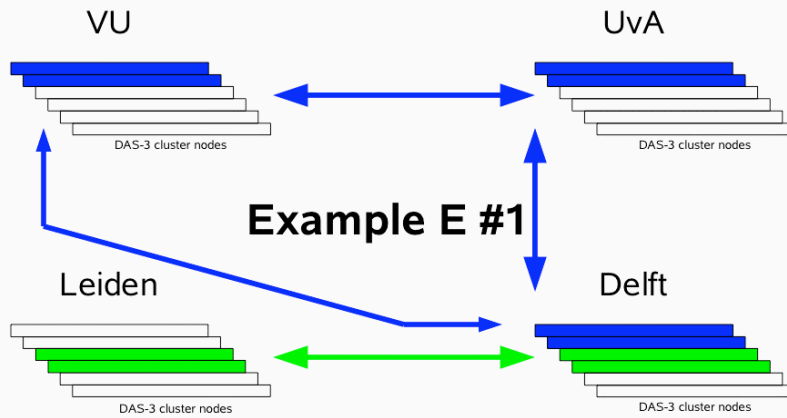
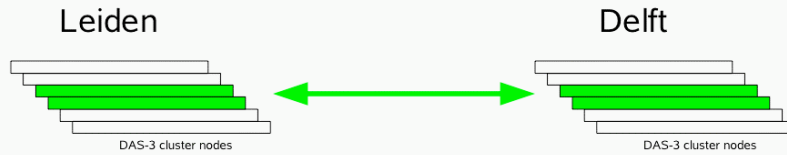
More traffic engineering



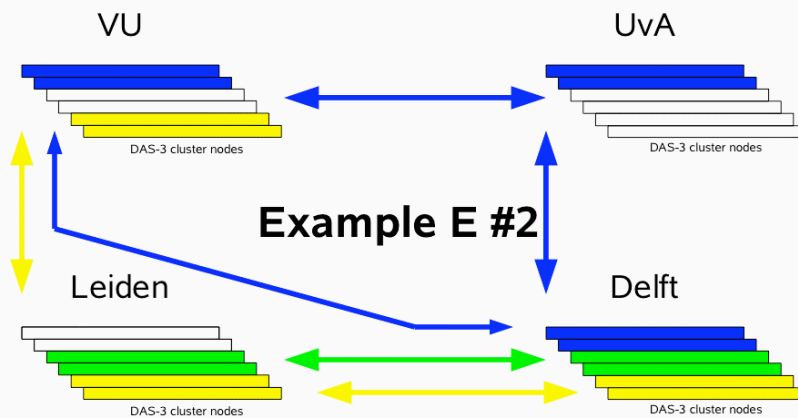
Even more traffic engineering



Example D



Example E #1



Example E #2



Risks

what have we today

what to avoid





No Change
Minimum Credit
Billing \$3
For questions, comments, or info
(800) 484-7825...
Office Hours: 9:00 AM -

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

surfnet



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)



StarPlane Overview



StarPlane is a NWO funded project of the University of Amsterdam (UvA) and the Vrije Universiteit (VU). The project uses the photonic network of the Dutch National Research and Education Network - SURFnet - that connects the five compute clusters at the four sites (UvA, VU, TU Delft and Universiteit Leiden). The vision is to allow part of the photonic network infrastructure of SURFnet to be manipulated by Grid applications to optimize the performance of specific In-Science applications.

Network Topology



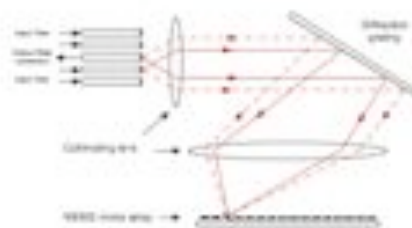
StarPlane is built on top of the photonic SURFnet; SURFnet's sixth generation network built with Nortel and Avici equipment. StarPlane utilizes a full 3-band within CPL, #8, starting with four and later eight 10Gbps wavelengths at each site. Two sets of Wavelength-Selectable Switches (WSS) placed in 3-degree branch setups allow for seamless 'optical switch-routing' of signals between all five DAS-3 clusters. DWDM filter equipment at all four sites breaks out 3-band #8 using an OMS6500-BB for signal conversion before entering the DAS-3/StarPlane WAN equipment. The proposed control plane for the optical/photonic side is a combination of Dynamic Resource Allocation Control (DRAC), CMPLS and Generic AAA.

DAS-3 / StarPlane Site Design



Four of the five DAS-3 clusters use Myrinet's Myri-10C interfaces, switches and 8-port 10GE LAN PHY bridging switch cards for their WAN interconnect. The DAS-3 cluster in Delft will use a stack of Nortel SEDs to bundle 20Gbps links per node into eight 10GE LAN PHY interfaces for its WAN interconnect. All eight 10GE LAN PHY connections from each cluster will be connected to a SURFnet's Nortel OMS6500-BB. The DAS-3 clusters are AMD Opteron based, per site differences in memory, CPU speed, single or dual core, mass-storage and LAN interconnect are to satisfy local computational, area of expertise and research needs.

Wavelength Selectable Switch



A Wavelength Selectable Switch (SURFnet uses a Nortel WSS) can select wavelengths and send each different wavelength to a specific output port. The StarPlane project novelly utilizes those switches to realize the redirection of certain wavelengths and establish point-to-point lightpaths between nodes at short time-scales. By manipulating the WSS and other optical devices, StarPlane is also able to change the logical topology of the network in real time (subsecond) and allocate the network resources on-demand to multiple applications at different clusters.

Visit poster session



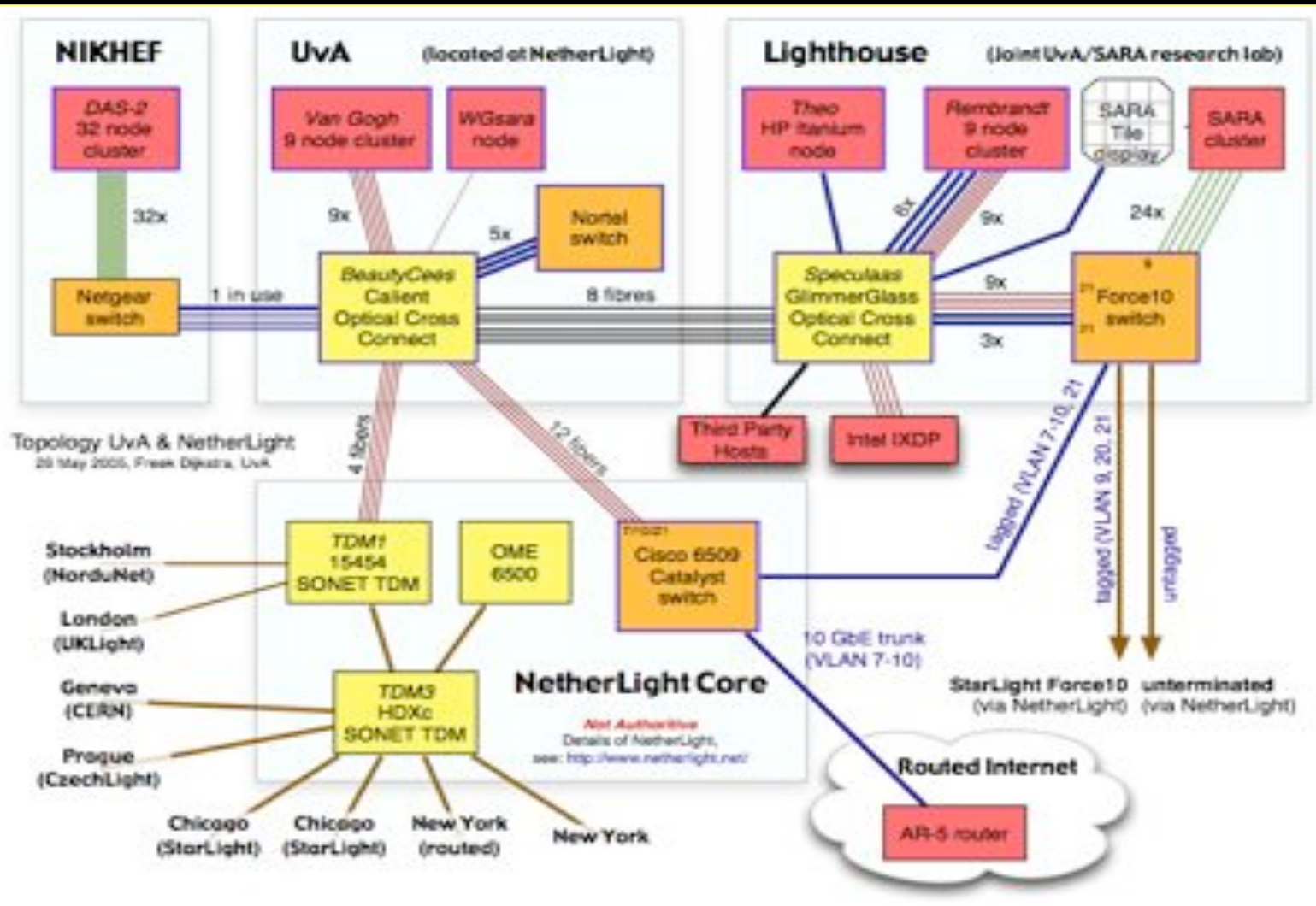
Key issue #1:
how to describe such networks?

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



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

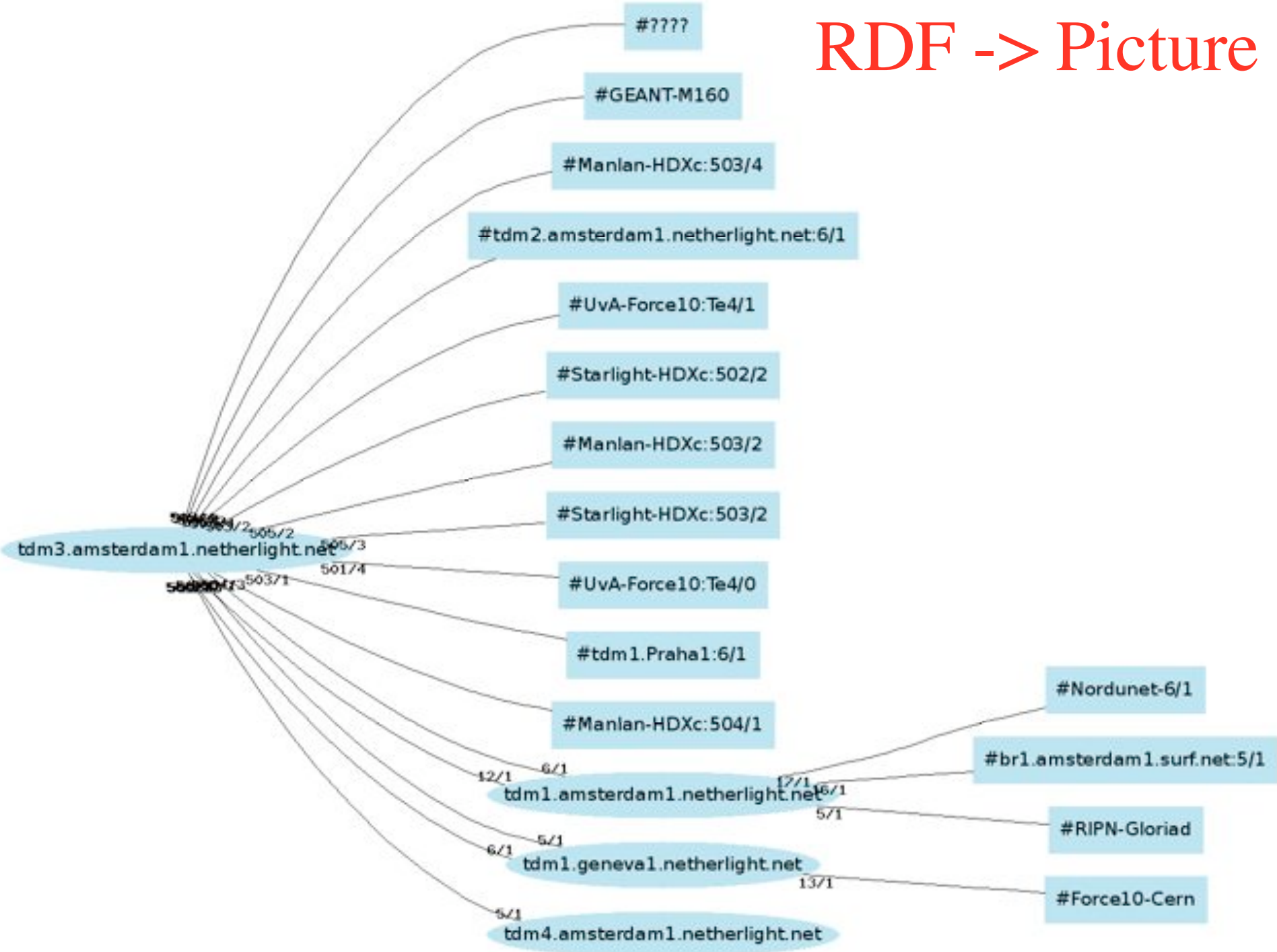
```
<ndl:Device rdf:about="#Vangogh3">  
  <ndl:name>Vangogh3</ndl:name>  
  <rb:isOfType>ComputingElement</rb:isOfType>  
  <ndl:locatedAt rdf:resource="#Lighthouse"/>  
  <ndl:hasInterface rdf:resource="#Vangogh3:eth2"/>  
</ndl:Device>
```

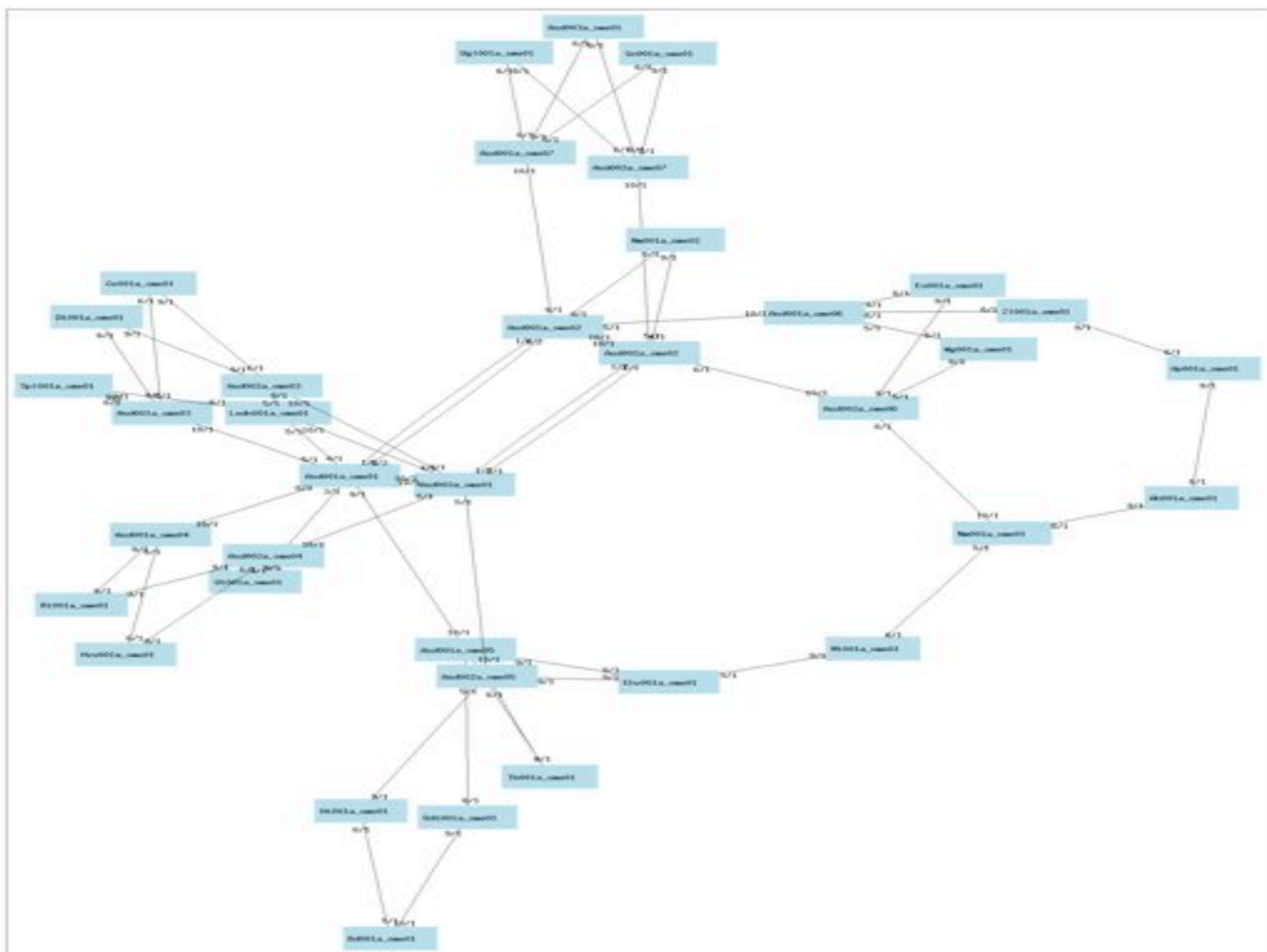
NetherLight in RDF

```
<?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:resource="#tdm3.amsterdam1.netherlight.net:503/2"/>
    <ndl:hasInterface rdf:resource="#tdm3.amsterdam1.netherlight.net:503/3"/>
    <!-- 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





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">
```


Resource Brokering: Your Ticket Into NetherLight

Application architecture:



Lambda networking allows the creation of application specific light paths.

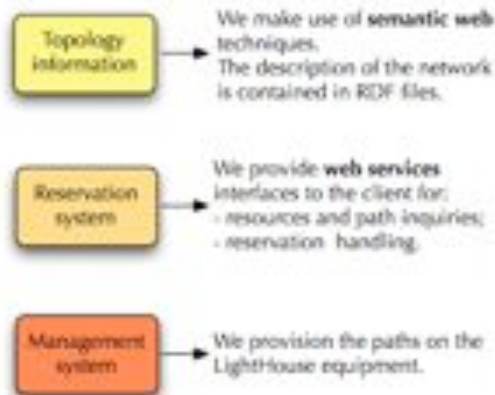
Lambda networking facilities empower users to request services and provision **end-to-end light paths** if and when they need it.

NetherLight, located in Amsterdam, The Netherlands, is one of such facilities.

The Amsterdam **LightHouse** is a joint research laboratory of the UvA and SARA.

Resources in the LightHouse can be used by collaborators to prove the concepts of hybrid networks.

Lightpath setup components:



Semantic web

The Network Description Language, an RDF Schema, describes networks in a standard, interoperable way.

Web Services

A WSDL file describes the interfaces to the service available to clients. Clients can interact with the service directly or via a portal.

Our SC|05 demonstration

We show the setup of dynamic connections between two computing nodes through the LightHouse/ NetherLight Optical Exchange.

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 !!!!

