Lambda-Grid projects in 2007 and beyond

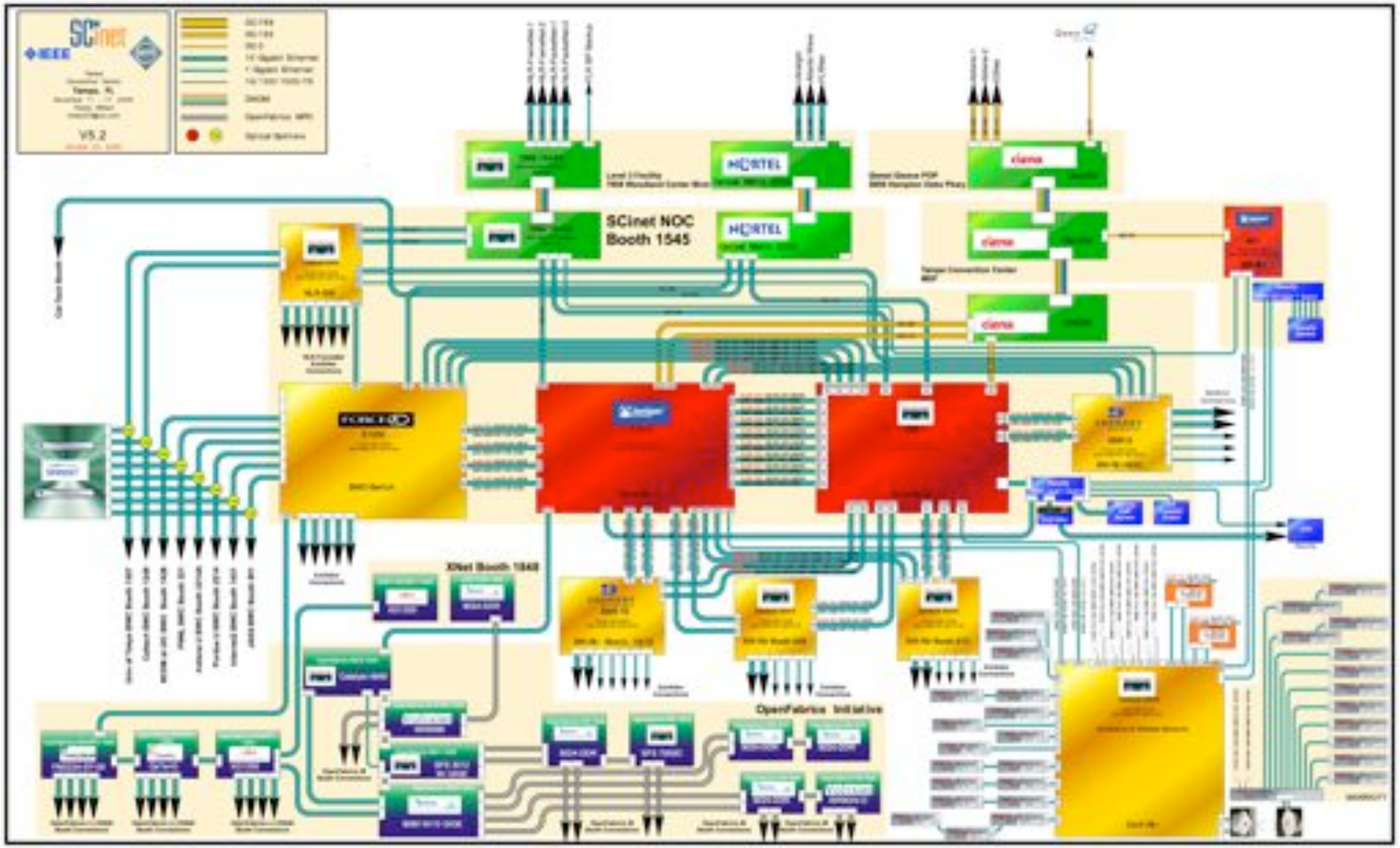
Cees de Laat

University of Amsterdam

What is missing in e-Infrastructure from the e-Science viewpoint?

- Useful ubiquitous access to photonic networks
  - first mile problems
- Grid programming models which go beyond treating the communication as Virtual Private Networks
- Scaleable optical/photonic network resources preventing cost explosions

Architecture SC06



What is the next hot topic in engineering e-Infrastructures?

- Middleware is the key to unlock the tremendous capacity in dark fiber networks
  - RDF, policy, addressing & routing
  - make these networks functions in WFM systems
- Utilize the capacity
  - few Tbit/sec/fiber => few 100 times 10 Gbit/s
- reduce cost and complexity of grooming and switching
- power per bit, power per multiplication, etc.
  - 250 W/10Gbit -> few times 25 kW/fiber/side for >L0
  - costs ~ 1 kEuro per kW per year

GigaPort-NG (ends december 2008)

- Plans for 2007 from UvA:
  - Hybrid networking structure
    - Network modeling
    - Extensions to NDL
      - multi layer
    - NDL-based multi-domain path provisioning
    - OGF Standardization
      - nml-wg
      - glif
  - SN7 brainstorming

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.

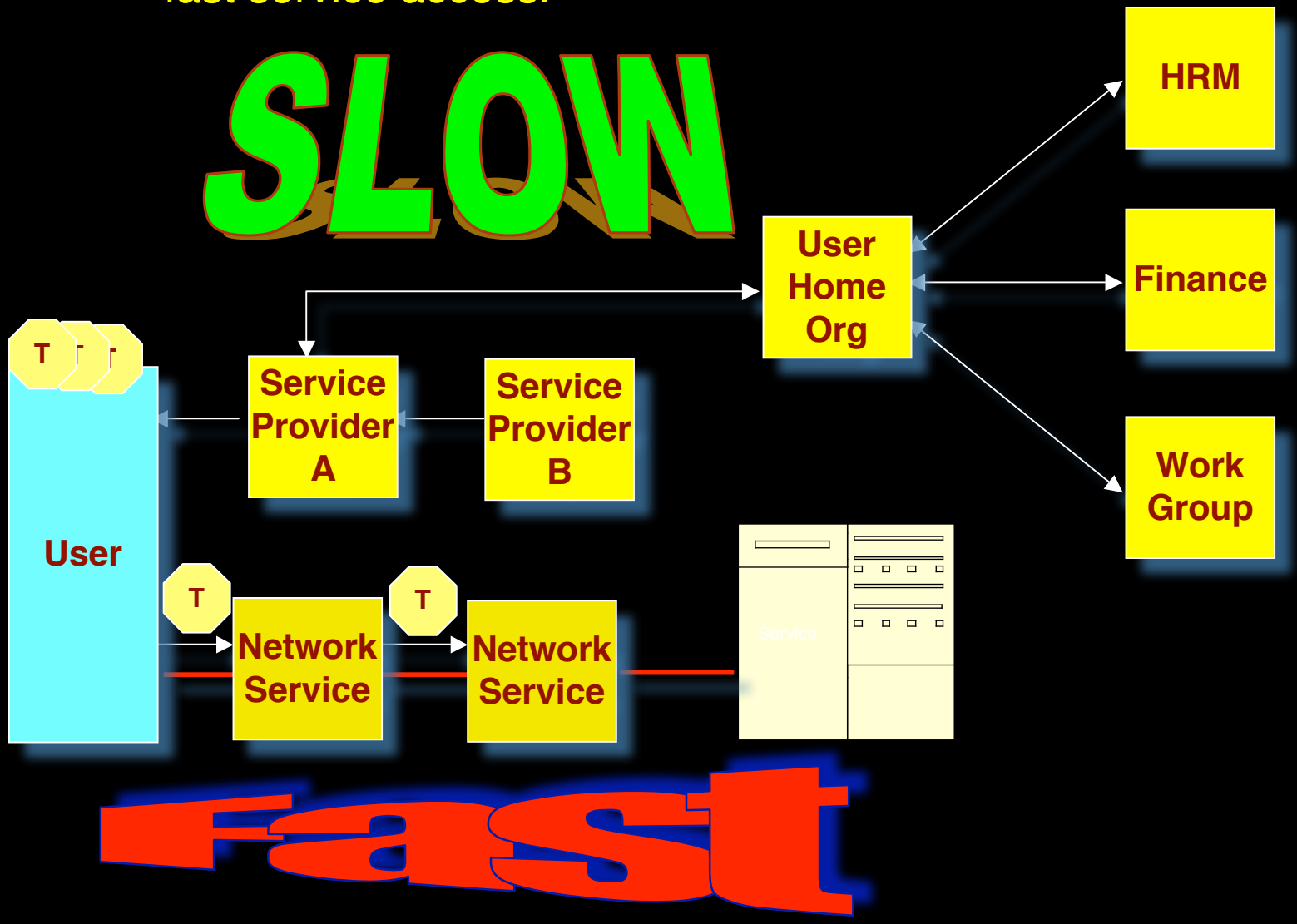
Google map and NDL…

…the GLIF connections described by NDL.

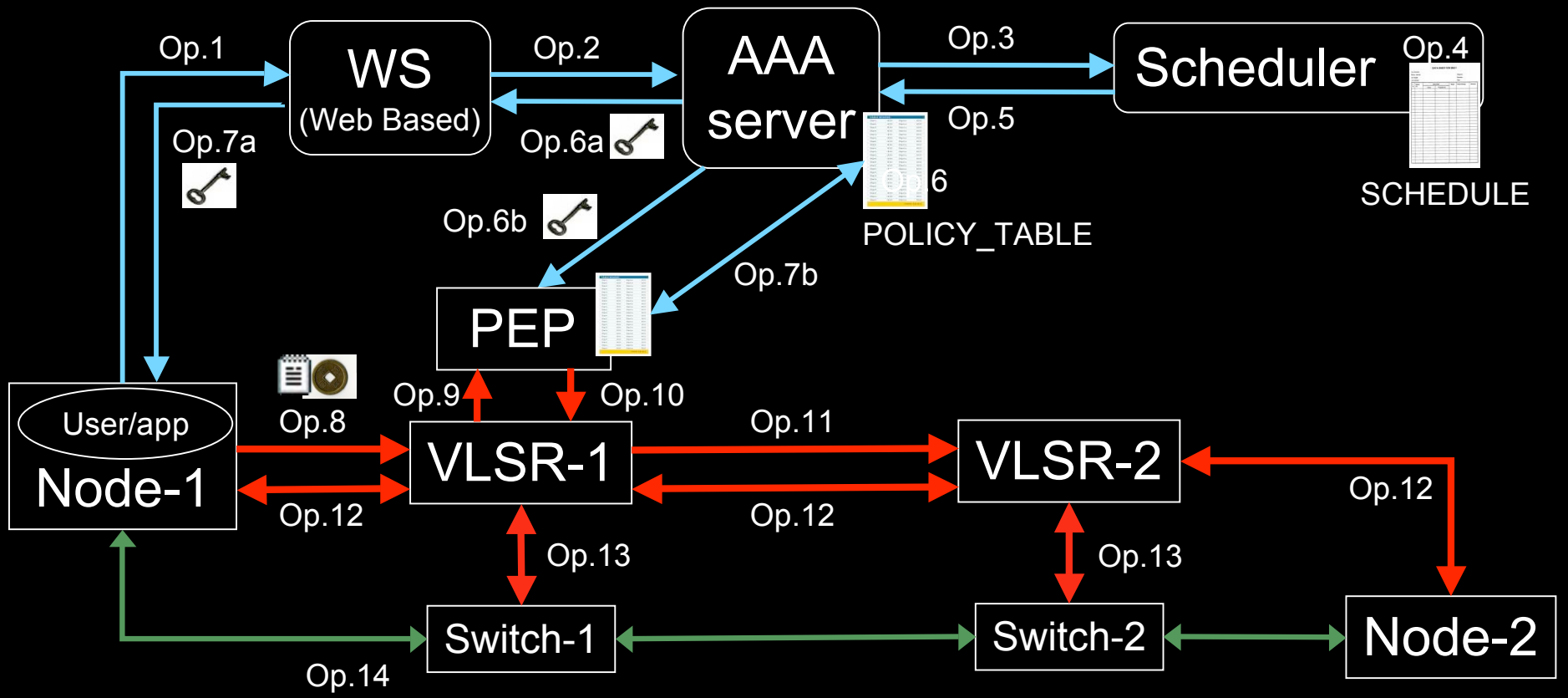
GigaPort-NG

- AAA
  - Token based GMPLS
  - AAA toolkit

TBN: split (time consuming) service authorization process from service access using secure tokens in order to allow fast service access.



Workflow for TBN in GMPLS with DRAGON



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.

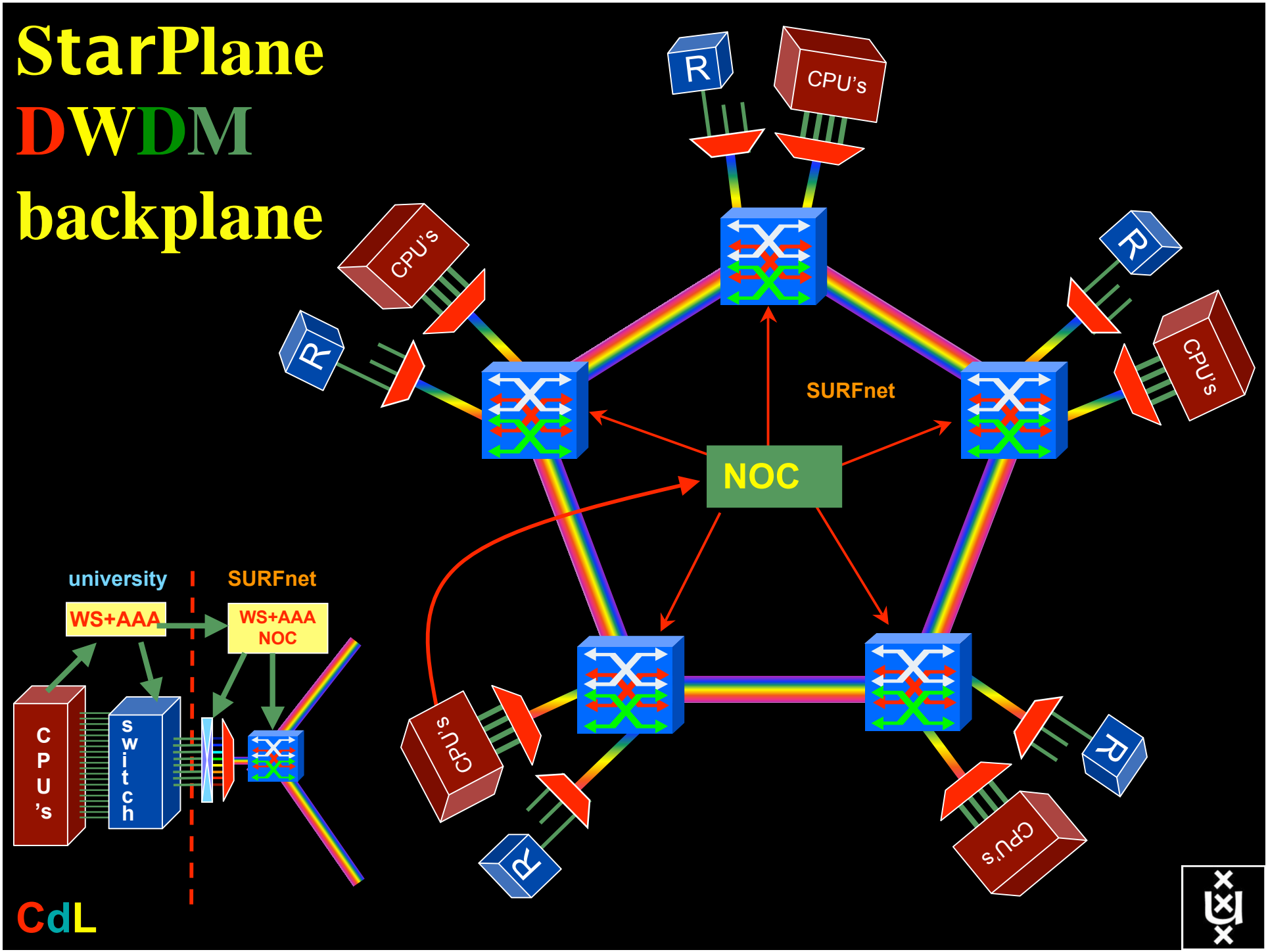
6a. AAA server issues an encrypted key to the WS.
6b. AAA server passes the same key to the PEP.
7a. WS passes the key to the user.
7b. AAA server interacts with PEP to update the local POLICY_TABLE on the PEP.

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

GigaPort-NG

- AAA
  - Token based GMPLS
  - AAA toolkit
- Hybrid networking applications
  - StarPlane
    - StarPlane network engineering
    - StarPlane control plane architecture

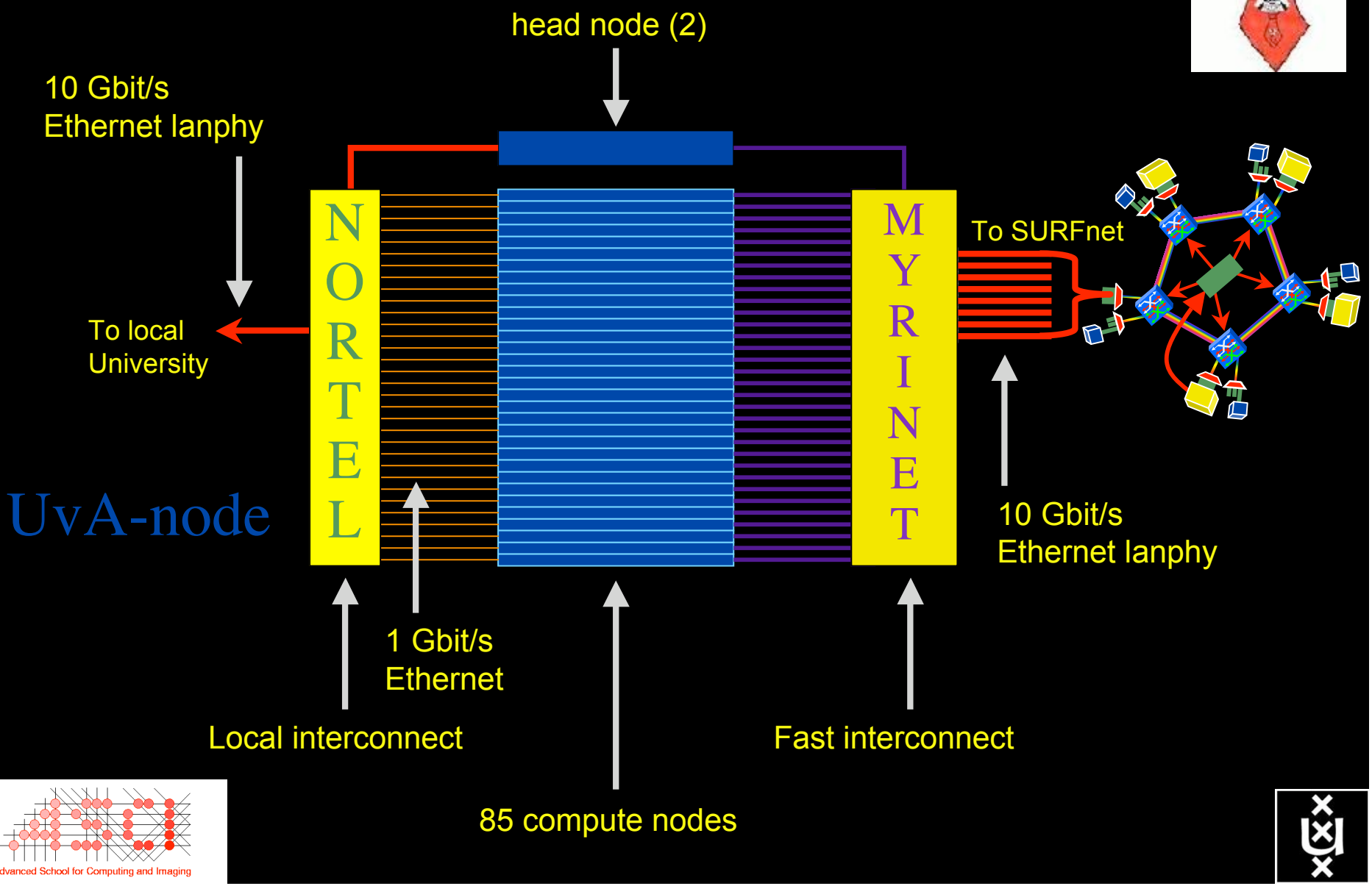
StarPlane
DWDM
backplane



CdL

DAS-3 Cluster Tender

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



Advanced School for Computing and Imaging

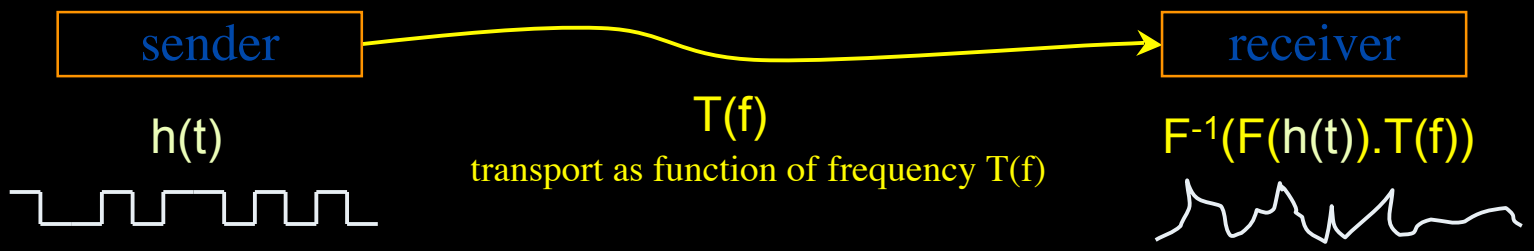
Heterogeneous clusters

(# of unused ports)

	LU	TUD	UvA-VLE	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	46.4 GHz
* memory	16GB	16GB	8GB	16GB	8GB	64GB
* Myri 10G	1		1	1	1	40 Gb/s
* 10GE	1	1	1	1	1	50 Gb/s
Compute	32	68	40 (+1)	46	85	271
* storage	400GB	250GB	250GB	2x250GB	250GB	84 TB
* CPU	2x2.6GHz	2x2.4GHz	2x2.2GHz DC	2x2.4GHz	2x2.4GHz DC	1.9 THz
* memory	4GB	4GB	4GB	4GB	4GB	1048 GB
* Myri 10G	1		1	1	1	2030 Gb/s
Myrinet						
* 10G ports	33 (7)		41	47	86 (2)	2070 Gb/s
* 10GE ports	8		8	8	8	320 Gb/s
Nortel						
* 1GE ports	32 (16)	136 (8)	40 (8)	46 (2)	85 (11)	339 Gb/s
* 10GE ports	1 (1)	9 (3)	2	2	1 (1)	

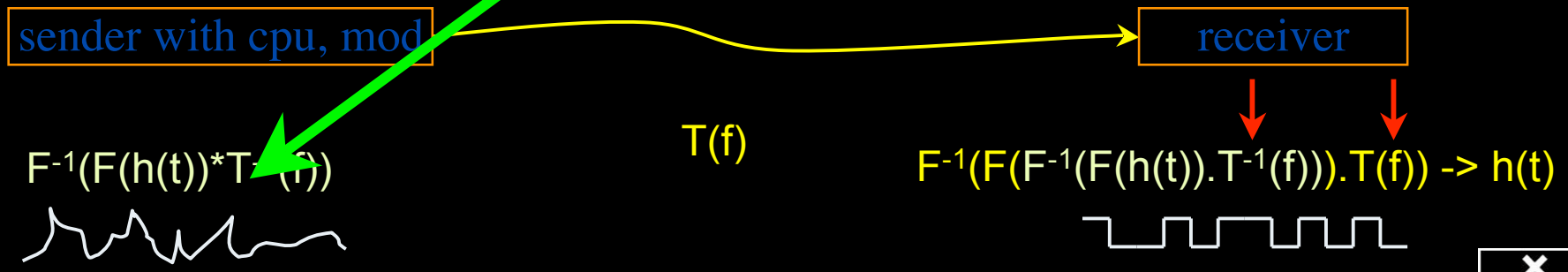
Dispersion compensating modem: eDCO from NORTEL

(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)

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

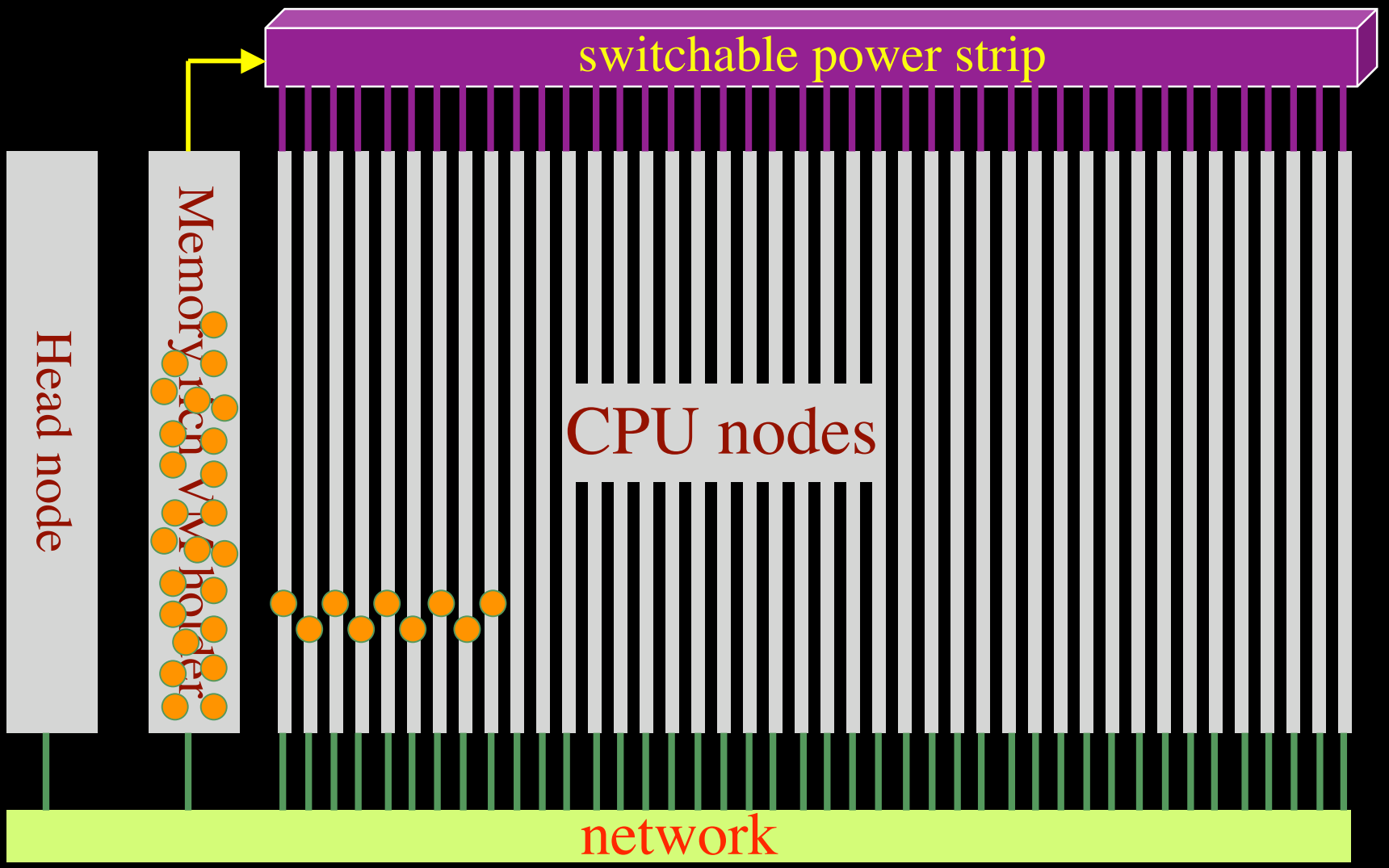
Power is a big issue

- 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!

Power outages are a big problem!

- on average about one outage per year
  - once the generator not starting/taking over
    - -> batteries
  - this summer weekend explosion of cable
    - -> generator fine!
- battery power for 5 minutes, generator to take over
- priorities for emergency power/cooling
- asked to shutdown compute nodes if temp rises

VM opportunity



GigaPort-NG

- AAA
  - Token based GMPLS
  - AAA toolkit

- Hybrid networking applications
  - StarPlane
    - StarPlane network engineering
    - StarPlane control plane architecture
  - Smallest University for proof of concepts
    - 'Optical' University of the Netherlands
    - MALAN - dark fiber
  - CineGrid
    - CineGrid: photonic distribution of digital cinema

OptIPuter related activities at SARA

Frame from Elephant's Dream. 1080p 720p and PAL formats superimposed for reference. Scaled up to 8000x3600 for OptIPortal use. For future: Investigate rendering at native resolution.

Bram Stolk – Paul Wielinga
OptIPuter AHM – Calit2 – January 21-23, 2007

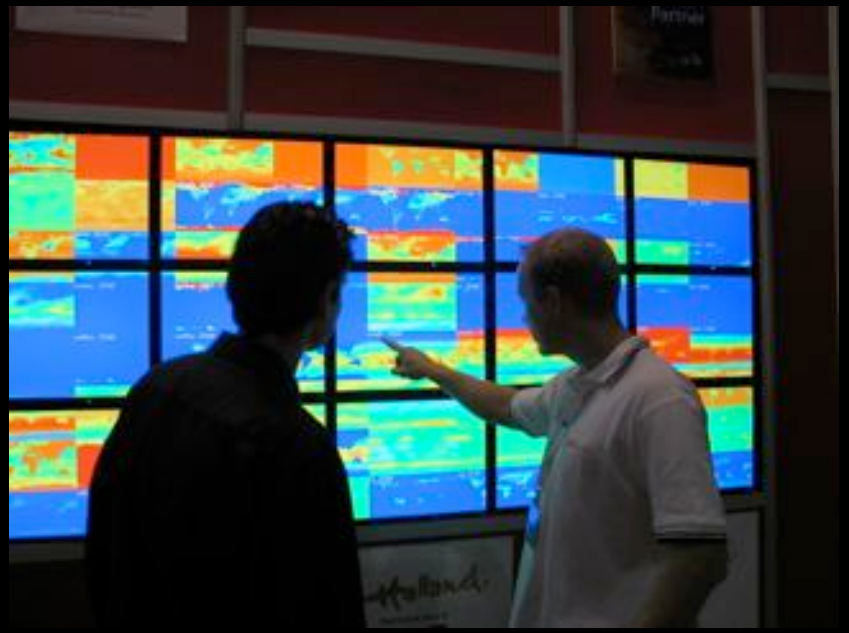
sara

OptIPuter related activities at SARA

- SAGE-to-SAGE collaboration at SC06 (hi-def vid)
- Deployment of eBio lab for UvA: sage equiped 5x4 panel
- Visualization of 50 TByte climate data on OptIPortal
- Phosphorus: EU project on application controlled lambdas
- Kick-starting 4K cinema: CineGrid Amsterdam
  * Testbed creation by [UvA, SARA] in LightHouse
  * First basic Demonstrator :4K content creation:
    scene from Elephant's Dream.

> *Elephant's Dream is the world's first Open Source movie. It's also the first that was released at 1080p on both HD-DVD and BluRay. Rendered at 4K and displayed on SARA's Tiled Panel Display. Elephant's Dream was rendered using the Blender Open Source software.*

SC06: SAGE demo, Climate data demo



sara

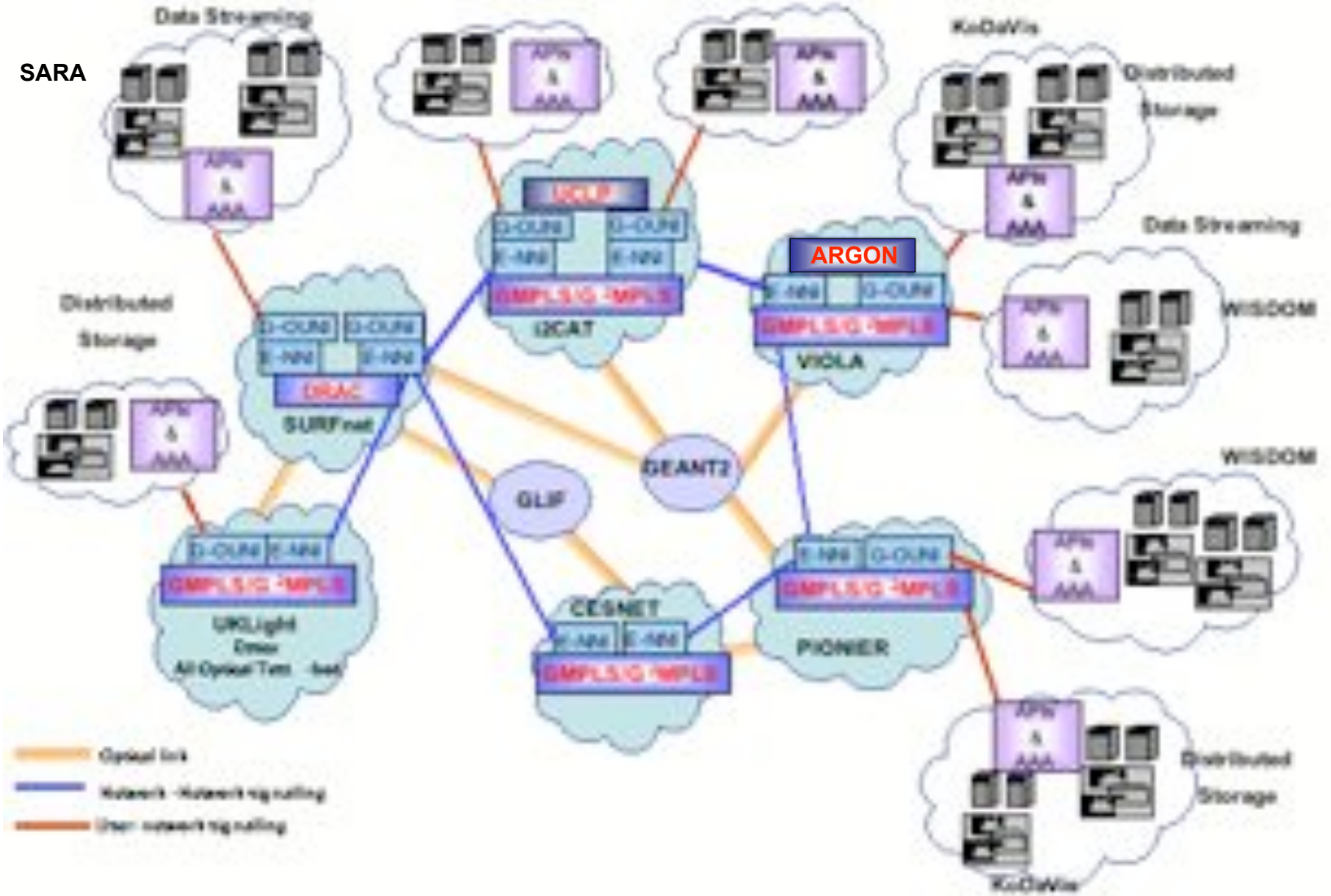
Lambda User Controlled Infrastructure For European Research

PHOSPHORUS

Cees de Laat (University of Amsterdam)
on behalf of the Phosphorus collaboration.

Phosphorus

European Multi-Domain Test-Bed Including Phosphorus Planned Developments



The phosphorus Project Key Features/Objectives (1)

- Demonstrate on demand service delivery across multi-domain/multi-vendor research network test-beds on a European and Worldwide scale. The test-bed will include:
  - EU NRENs: SURFnet, CESNET, PIONIER as well national test-beds (VIOLA, OptiCAT, UKLight)
  - GN2, GLIF and Cross Border Dark Fibre connectivity infrastructure
  - GMPLS, UCLP, DRAC and ARGON control and management planes
  - Multi-vendor equipment environment (ADVA, HITACHI, NORTEL, Vendor's equipment in the participating NREN infrastructure)

The phosphorus Project Key Features/Objectives (2)

- Develop integration between application middleware and transport networks, based on three planes:
  - Service plane:
    - Middleware extensions and APIs to expose network and Grid resources and make reservations of those resources (NDL)
    - Policy mechanisms (AAA) for networks participating in a global hybrid network infrastructure, allowing both network resource owners and applications to have a stake in the decision to allocate specific network resources
  - Network Resource Provisioning plane:
    - Adaptation of existing Network Resource Provisioning Systems (NRPS) to support the framework of the project
    - Implementation of interfaces between different NRPS to allow multi-domain interoperability with phosphorus's resource reservation system
  - Control plane:
    - Enhancements of the GMPLS Control Plane (G^2MPLS) to provide optical network resources as first-class Grid resource
    - Interworking of GMPLS-controlled network domains with NRPS-based domains, i.e. interoperability between G^2MPLS and UCLP, DRAC and ARGON

Questions ?

OptIPuter