GigaPort-RON dec 2008

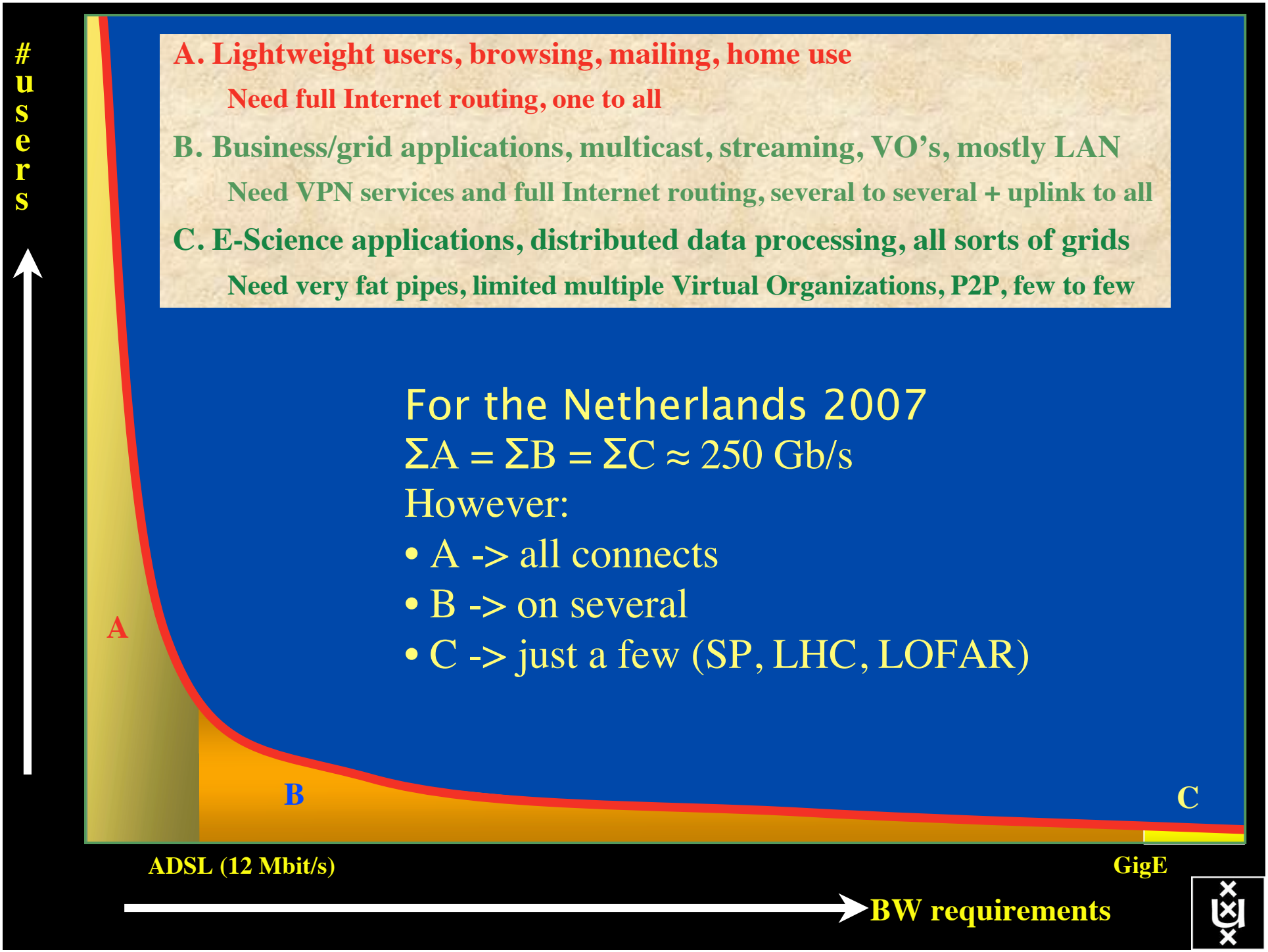
From Routed to Hybrid Networking

Cees de Laat

University of Amsterdam

GP - Plans 2004-2008

1. Hybrid networking structure
   - Network Architecture
   - Optical Internet Exchange Architecture
   - Network Modeling <NDL, Pathfinding>
   - Fault Isolation
2. Network transport protocols
   - UDP - TCP
   - Protocol testbed
   - LinkLocal Addressing
3. Optical networking applications
   - StarPlane
   - eVLBI
   - Smallest University for proof of concepts
   - CineGrid
   - CosmoGrid
4. Authorization, Authentication and Accounting in Networking and Grids
   - AAA & schedule server
   - WS security
   - Multi domain token based implementations
   - Cross domain LightPath setup
5. Testbed LightHouse, SC0X, iGrid, GLIF, OGF, Terena, ...


#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

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

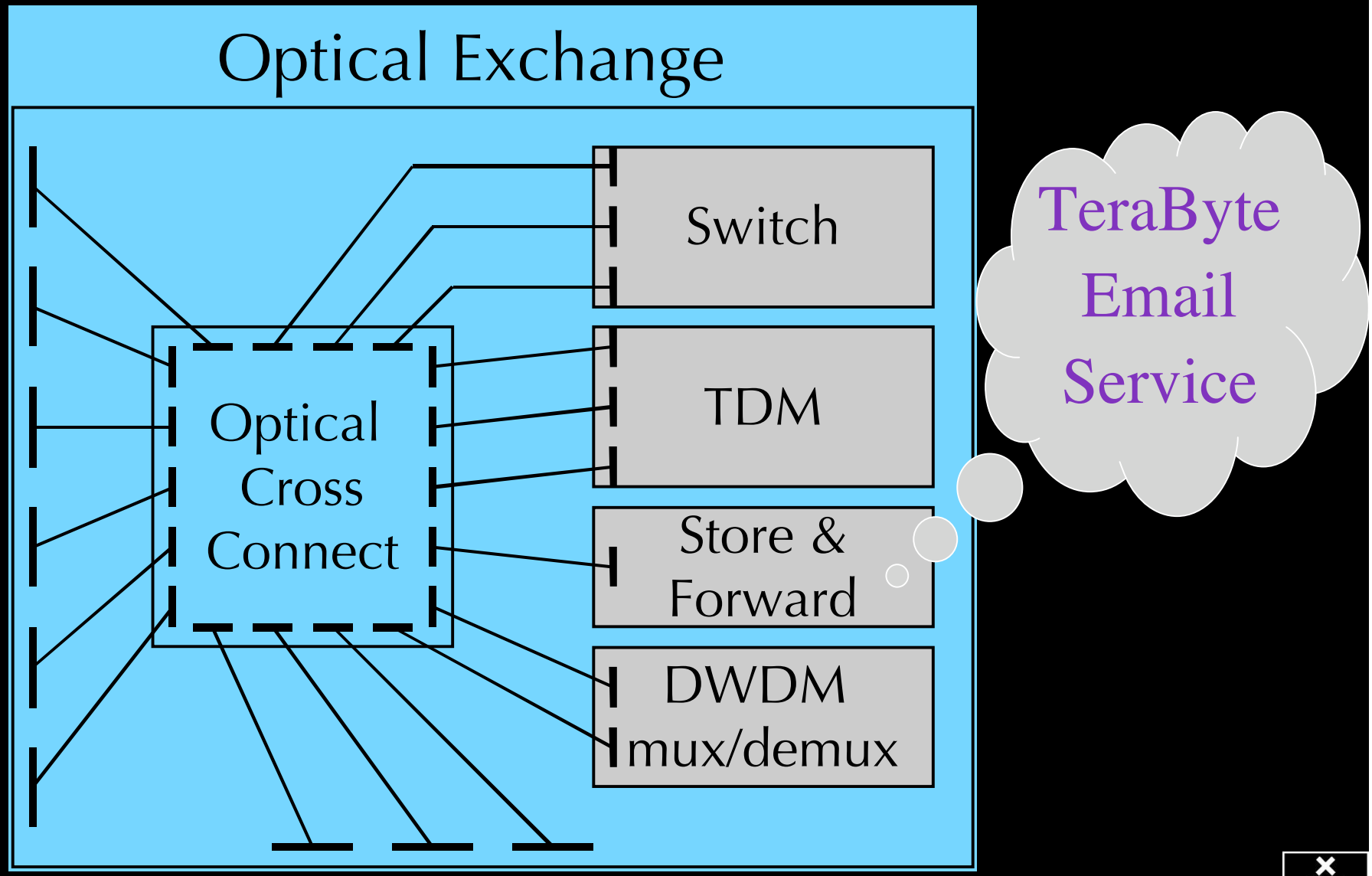
L2 ≈ 5-8 k$/port

L3 ≈ 75+ k$/port

Hybrid Network Paradigm

- Capability to handle datatransport on different OSI layers
- Most NREN's now offer end-to-end Lightpath services to their users
- Last 2 years tremendous progress in control plane implementations.
- Commercial Internet world has already >20.000 WSS's installed (ECOC2008)
- Our differentiating factor: put user in charge!

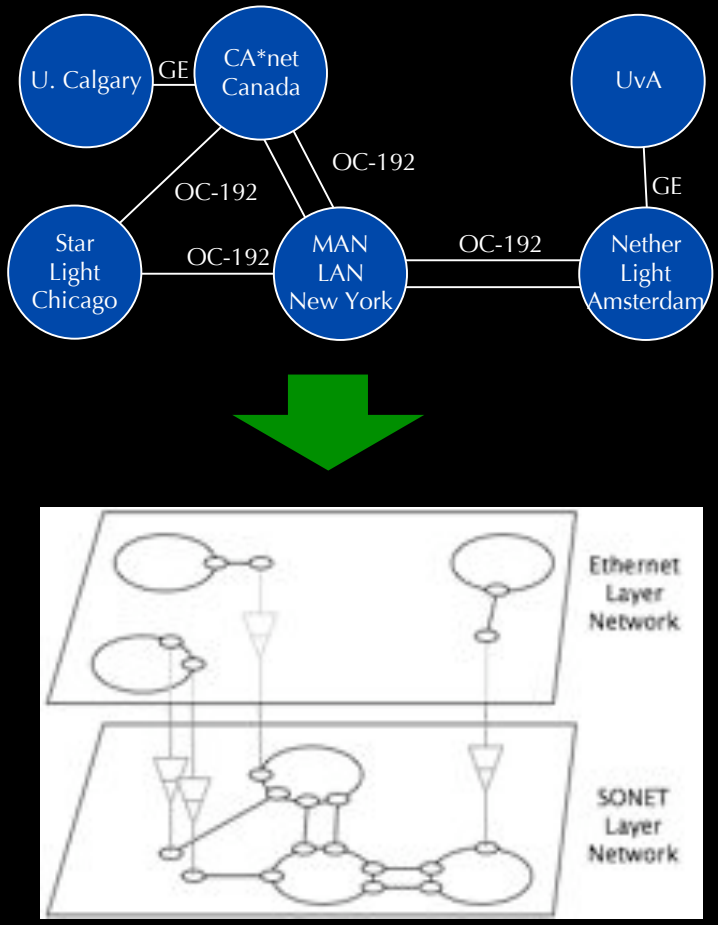
Optical Exchange as Black Box



Ref: gridnets paper by Freek Dijkstra, Cees de Laat

NDL Multilayer Extension

- ITU-T G.805 describes functional elements (e.g. adaptation, termination functions, link connections, etc.) to describe **network connections**.
- We extended these function elements (e.g. with potential adaptation functions) to describes **networks**.
- We created a model to map actual network elements to functional elements.
- Defined a simple algebra to define correctness of a connection
- We created a NDL extension to describe these functional elements.



Simplified model to map network elements to functional elements

OGF NML-WG

Open Grid Forum - Network Markup Language workgroup

Chairs:

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

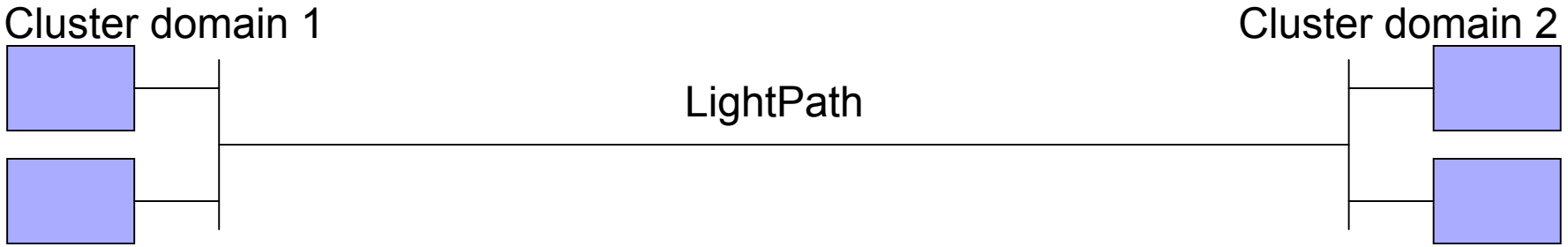
Purpose:

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

IP configuration in Optical Networks

- Problem: After a LightPath has been created, time is spent to manually configure IP addresses. Can this be done automatically?
- DHCP will not work out-of-the-box, since it is not clear which domain should run it.
- Possible solution: self-assigned IP addresses (RFC3927 for IPv4 or RFC1971 for IPv6)
- How to discover the target IP address or service?

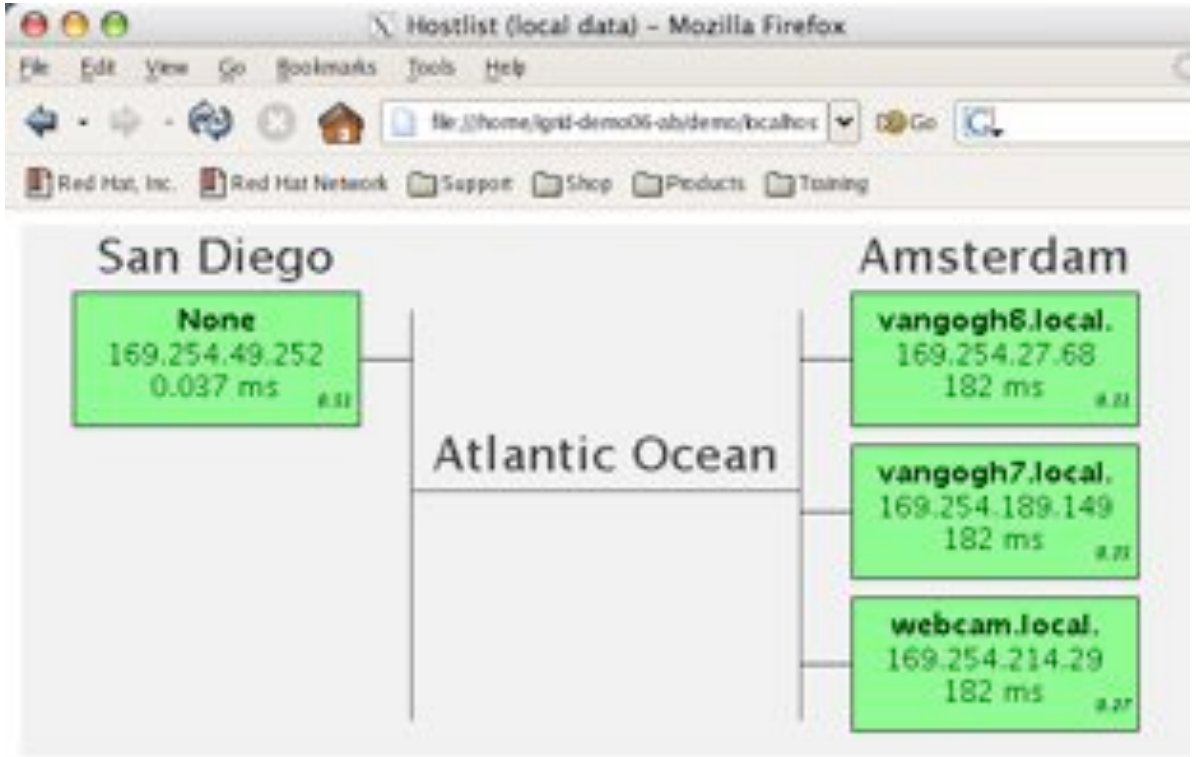


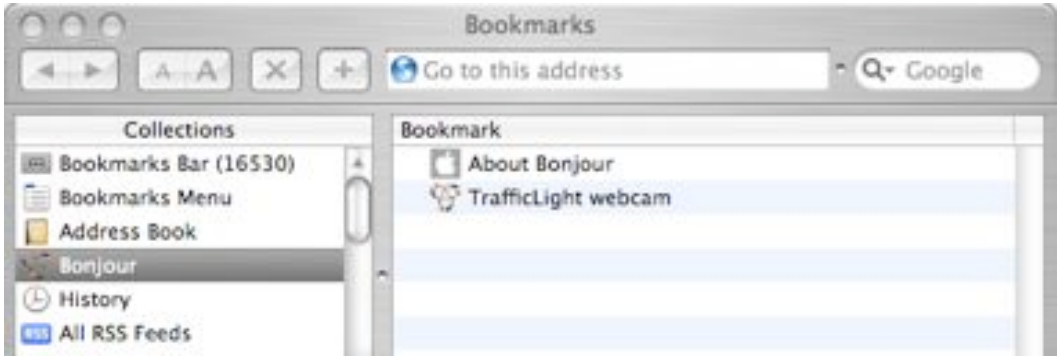
Technologies and Implementations

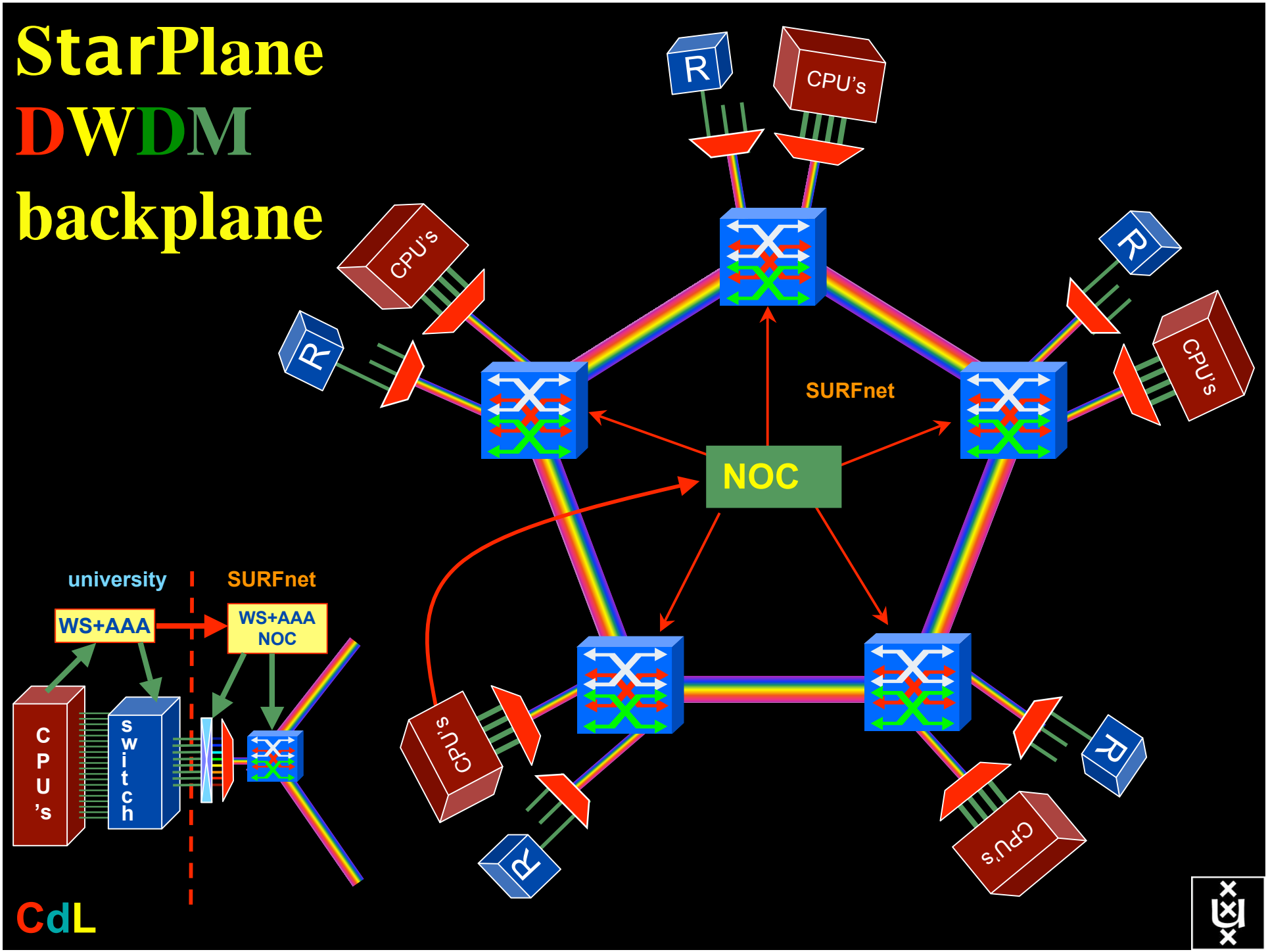
- Use Zero Configuration protocols
  - **Automatic configuration of IP addresses**
    - RFC3927 for IPv4 or RFC1971 for IPv6
  - **Name lookup of hosts**
    - Multicast DNS (mDNS) or Link-Local Multicast Name Resolution (LLMNR)
  - **Discovery of services**
    - DNS Service Discovery (DNS-SD), or Simple Service Discovery Protocol (SSDP, in UPnP), or Service Location Protocol (SLP) (or even UDDI, SDP, Salutation, or Jini)
- Three software suites, used multiple implementations:
  - RFC3927: ZCIP and autoip for Linux, native in OS X and Windows
  - mDNS: mDNSResponder, tmdns, and Porchdog mDNS
  - hooking gethostby*() to use mDNS: tmdns and libnss_mdns

Demonstration

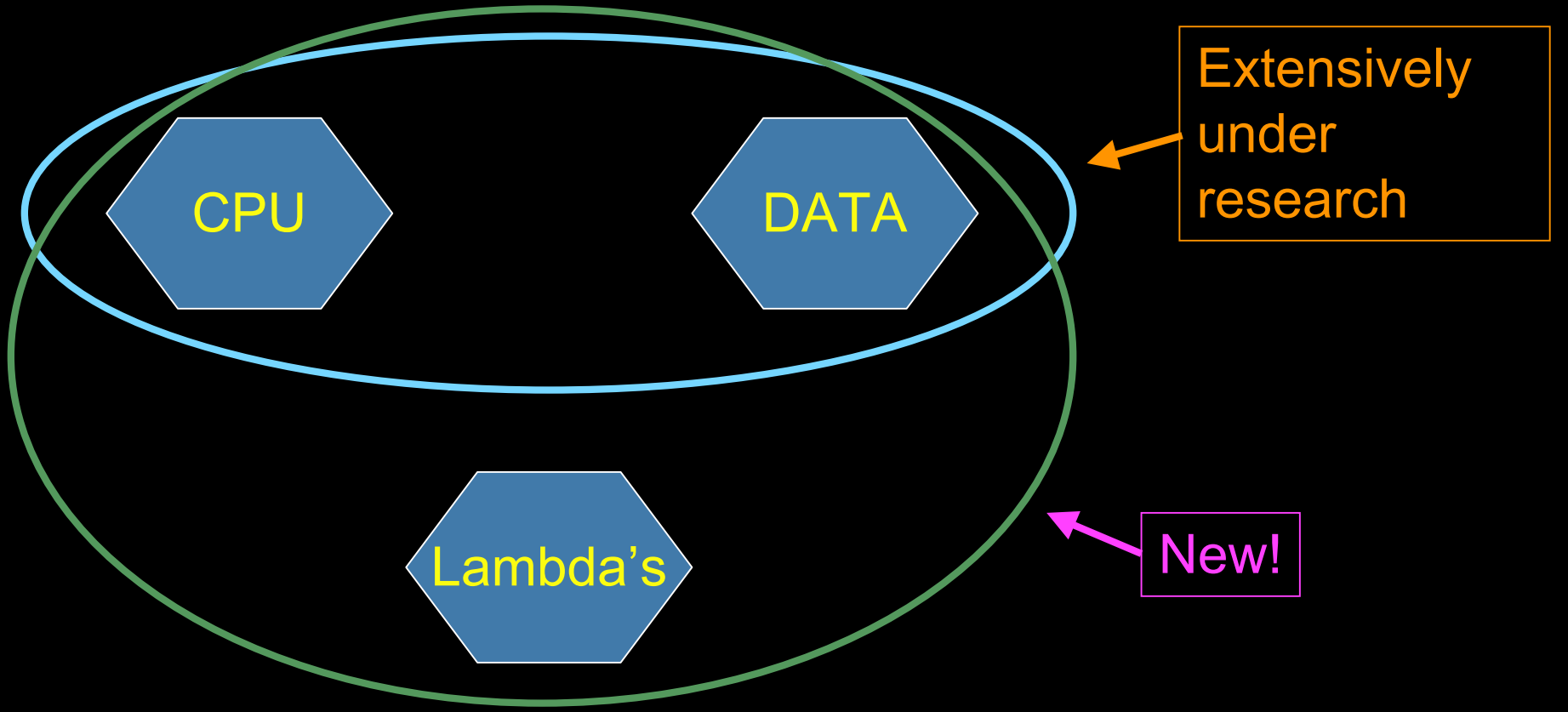
- Used broadcast ping to discover hosts
- Used multicast DNS and gethostbyaddr() hook to discover hostnames
- Tested IP collisions
- Also demonstrated service discovery through DNS



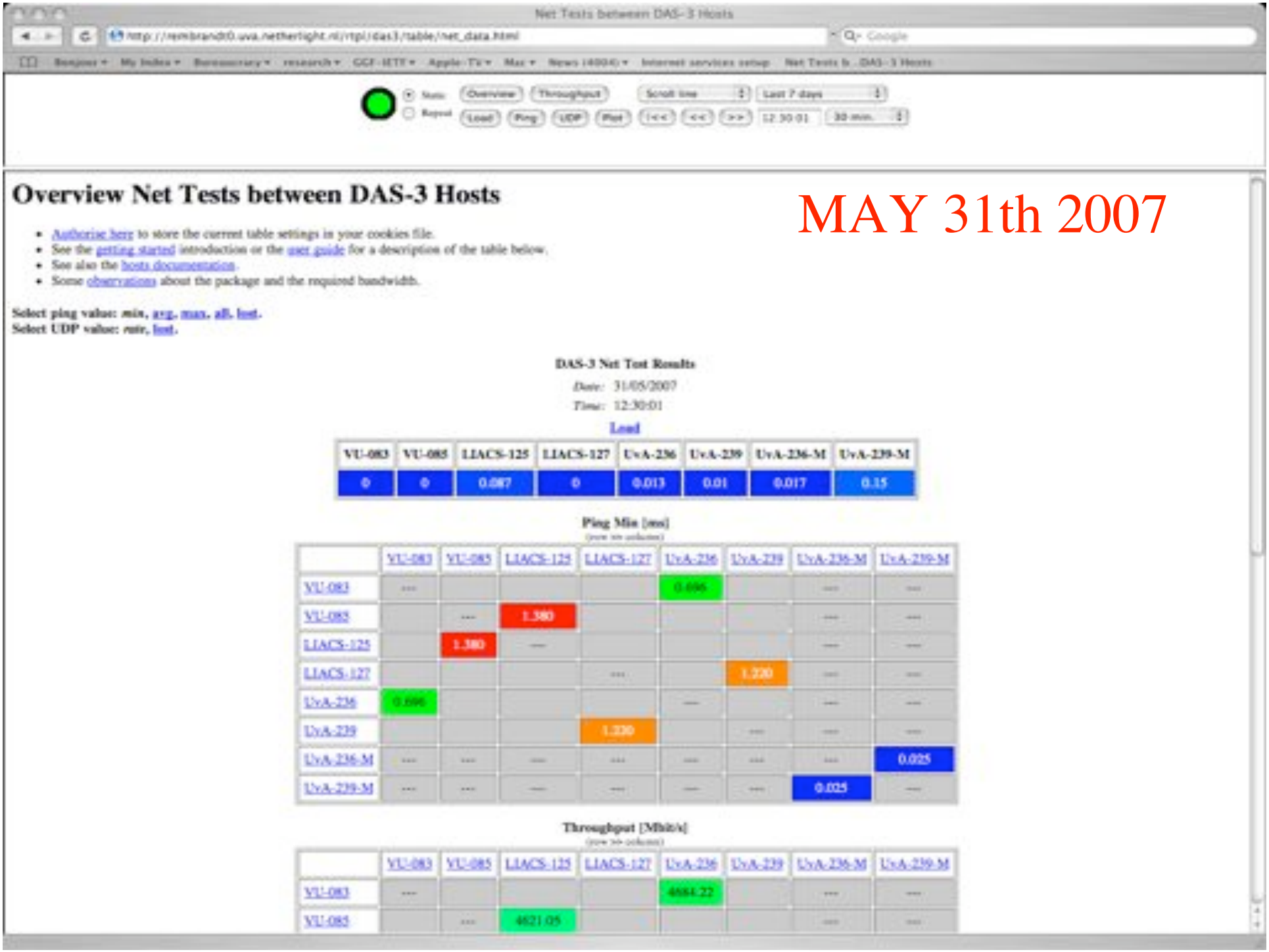


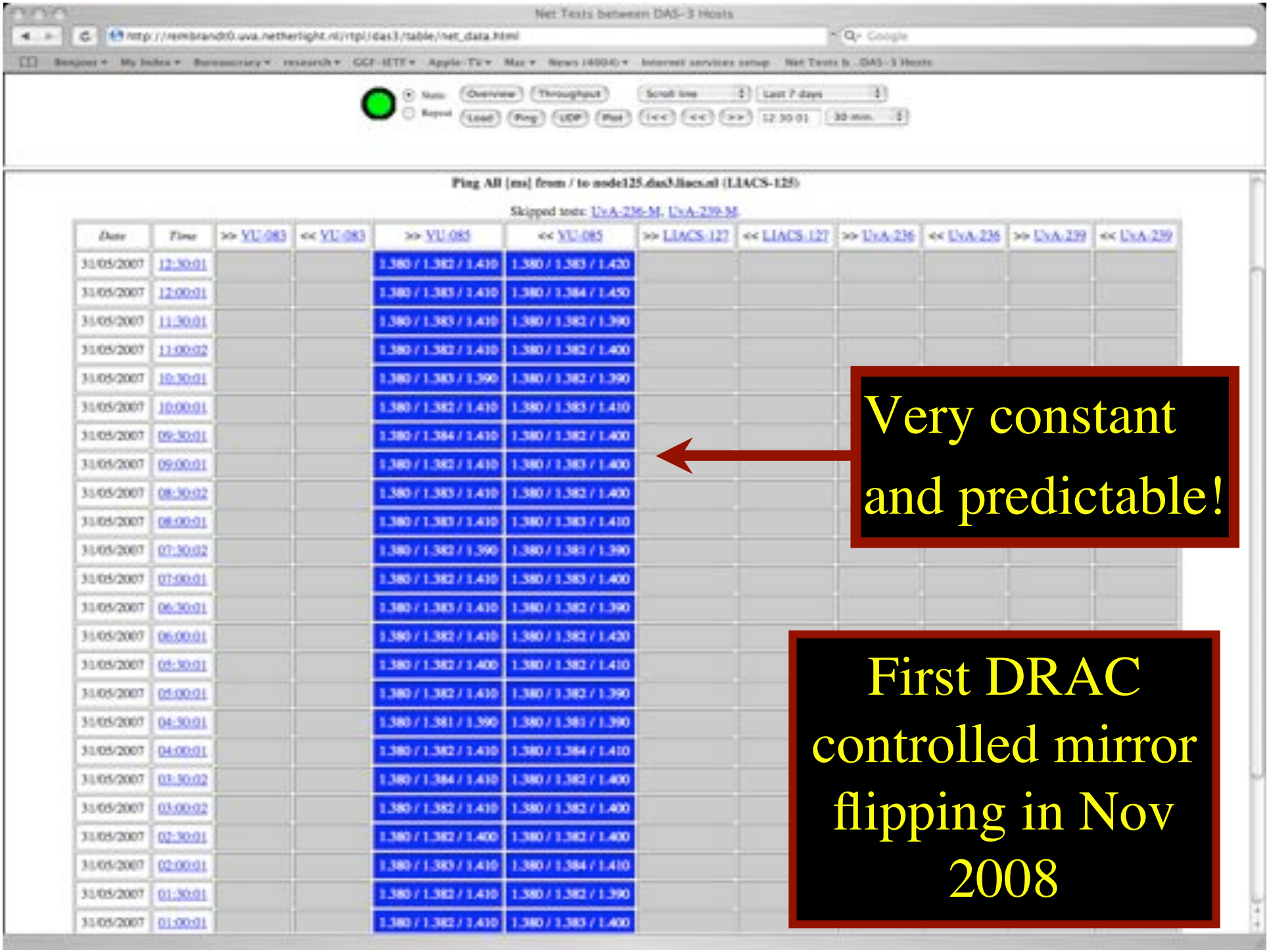

StarPlane
DWDM
backplane
R
CPU's
SURFnet
NOC
university
WS+AAA
WS+AAA
NOC
CPU's
switch
CdL

GRID Co-scheduling problem space

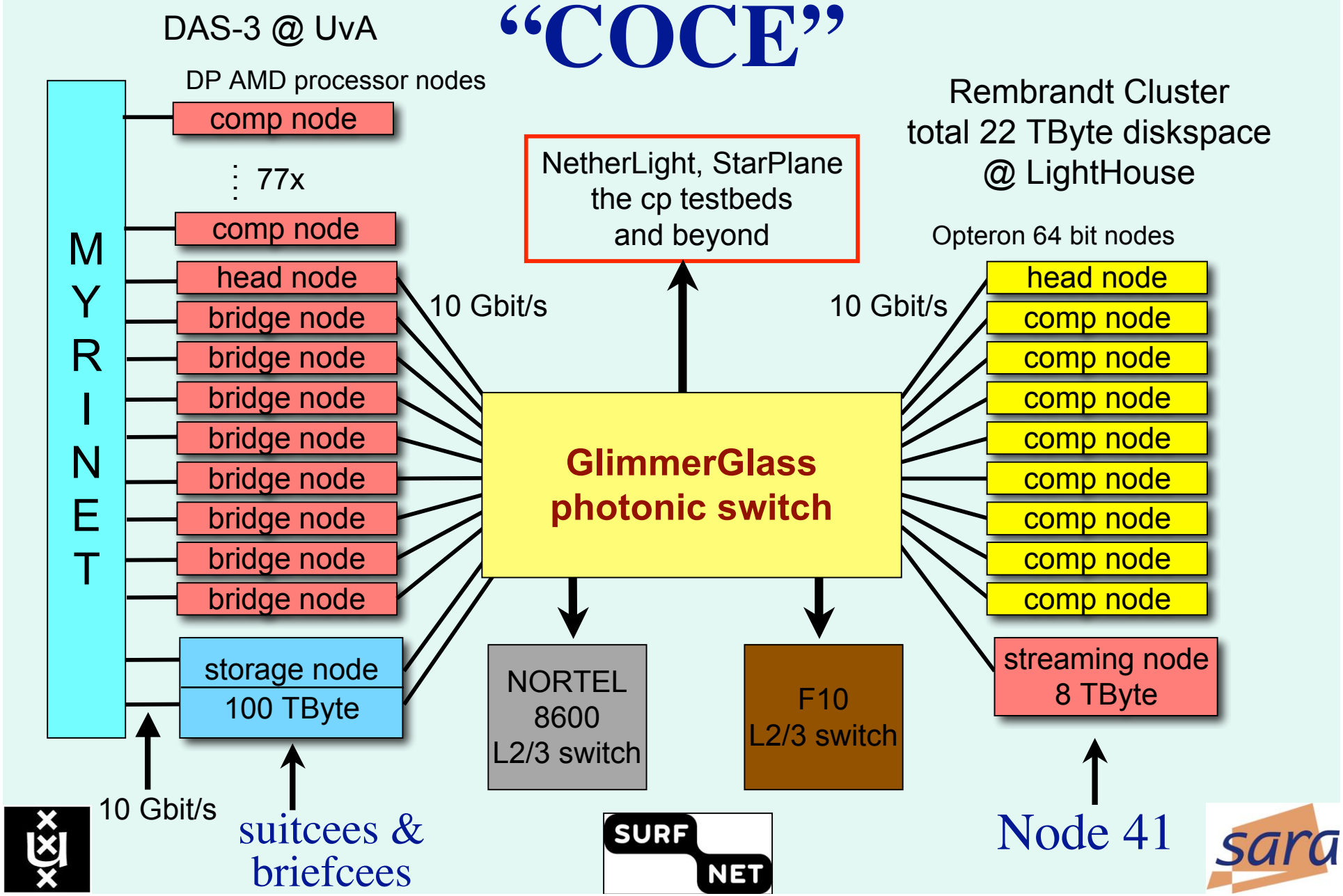


The StarPlane vision is to give flexibility directly to the applications by allowing them to choose the logical topology in real time, ultimately with sub-second lambda switching times on part of the SURFnet6 infrastructure.

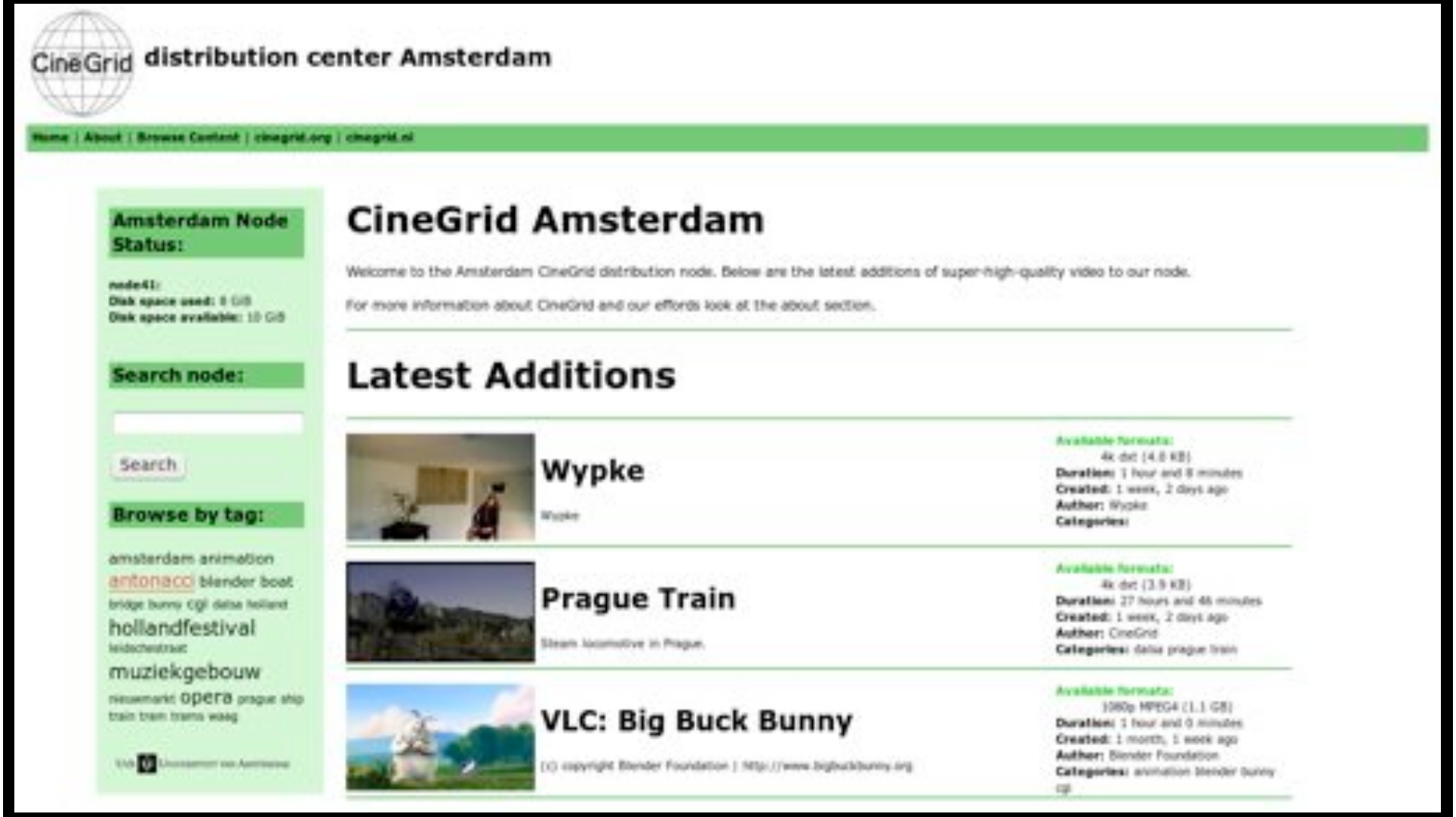

Net Tests between DAS-3 Hosts
Overview
Throughput
Load
Ping
UDP
Plot
Overview Net Tests between DAS-3 Hosts
MAY 31th 2007
DAS-3 Net Test Results
Date: 31/05/2007
Time: 12:30:01
Load
VU-083 VU-085 LIACS-125 LIACS-127 UvA-236 UvA-239 UvA-236-M UvA-239-M
0 0 0.087 0 0.013 0.01 0.017 0.15
Ping Min [ms]
VU-083 VU-085 LIACS-125 LIACS-127 UvA-236 UvA-239 UvA-236-M UvA-239-M
0.696
1.380
1.380
1.220
0.696
1.220
0.025
0.025
Throughput [Mbit/s]
VU-083 VU-085 LIACS-125 LIACS-127 UvA-236 UvA-239 UvA-236-M UvA-239-M
4684.22
4621.05


Net Tests between DAS-3 Hosts
Overview
Throughput
Load
Ping
UDP
Plot
Ping AB [ms] from / to node125.das3.liacs.nl (LIACS-125)
Skipped tests: UvA-236-M, UvA-239-M
Very constant and predictable!
First DRAC controlled mirror flipping in Nov 2008

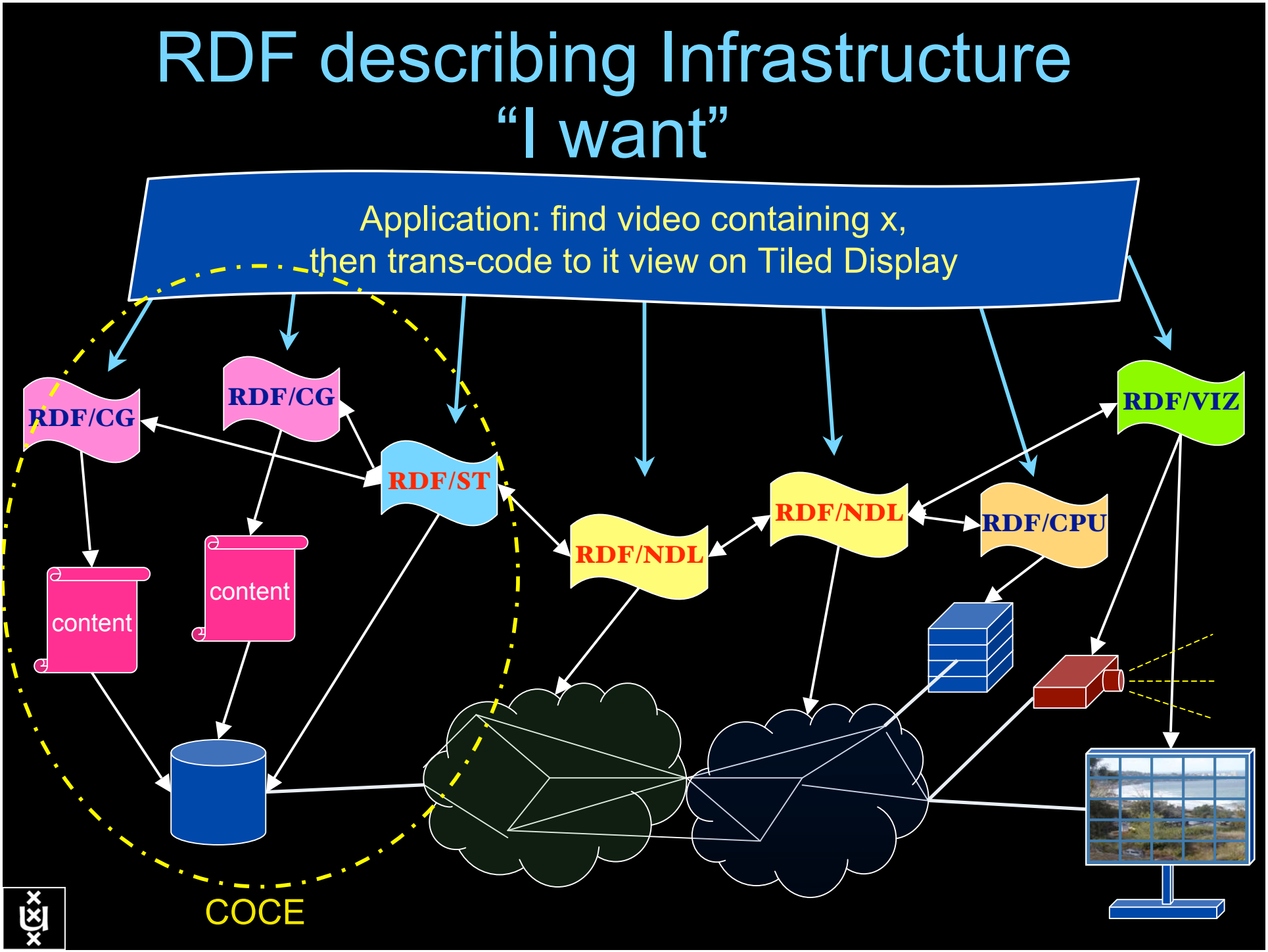
Amsterdam CineGrid S/F node "COCE"



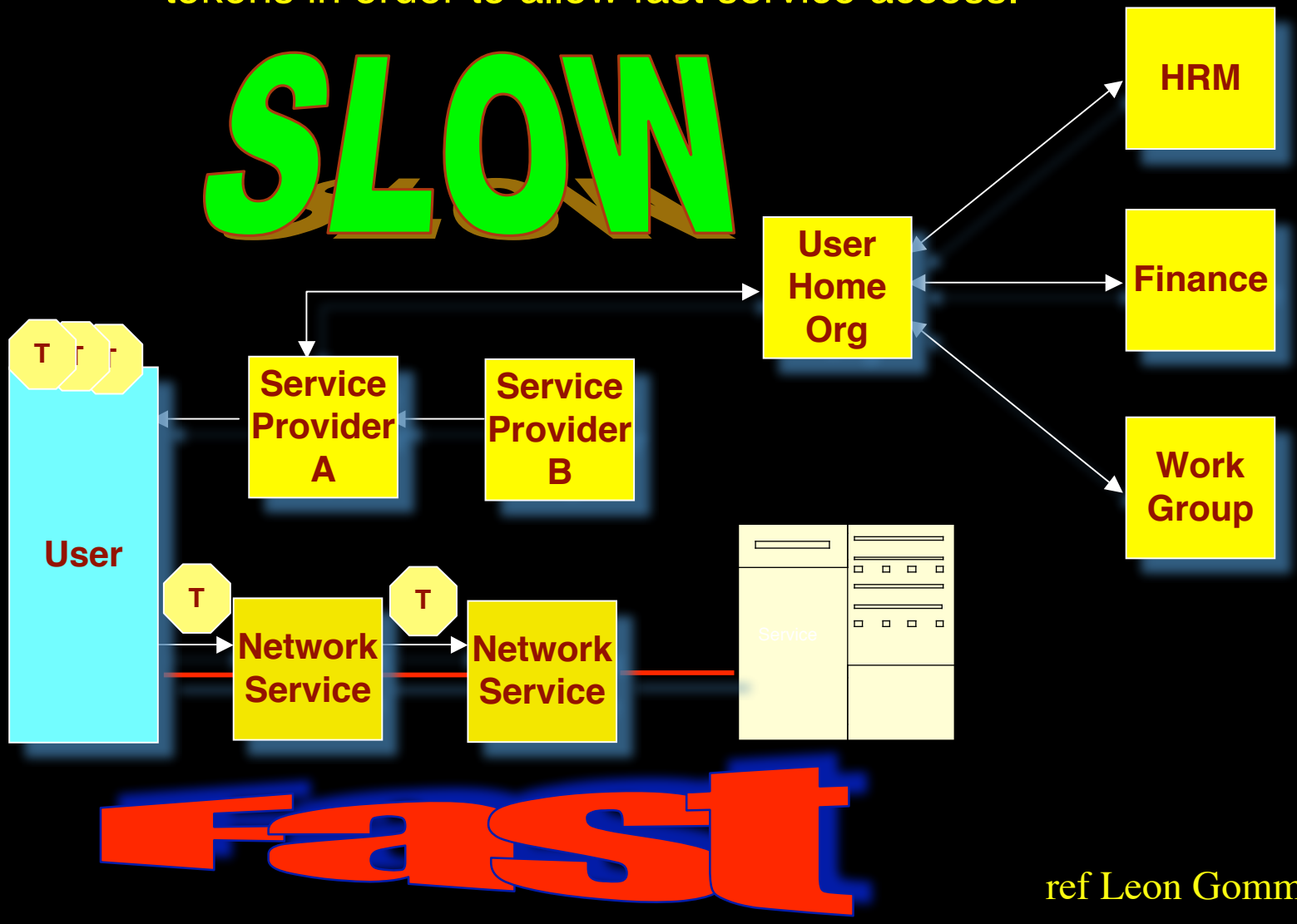
CineGrid portal



RDF describing Infrastructure
"I want"



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

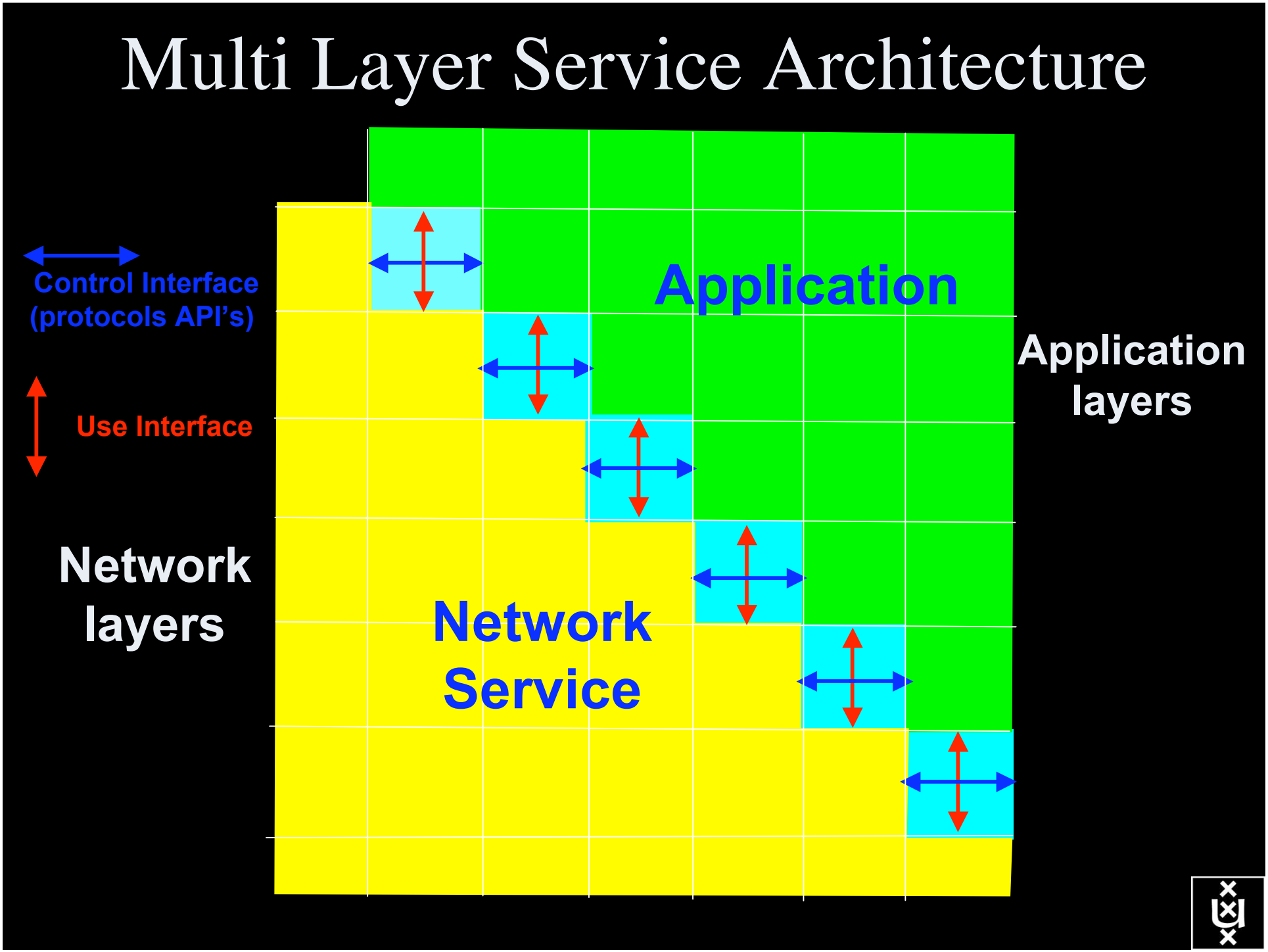


ref Leon Gommans

The HighLights

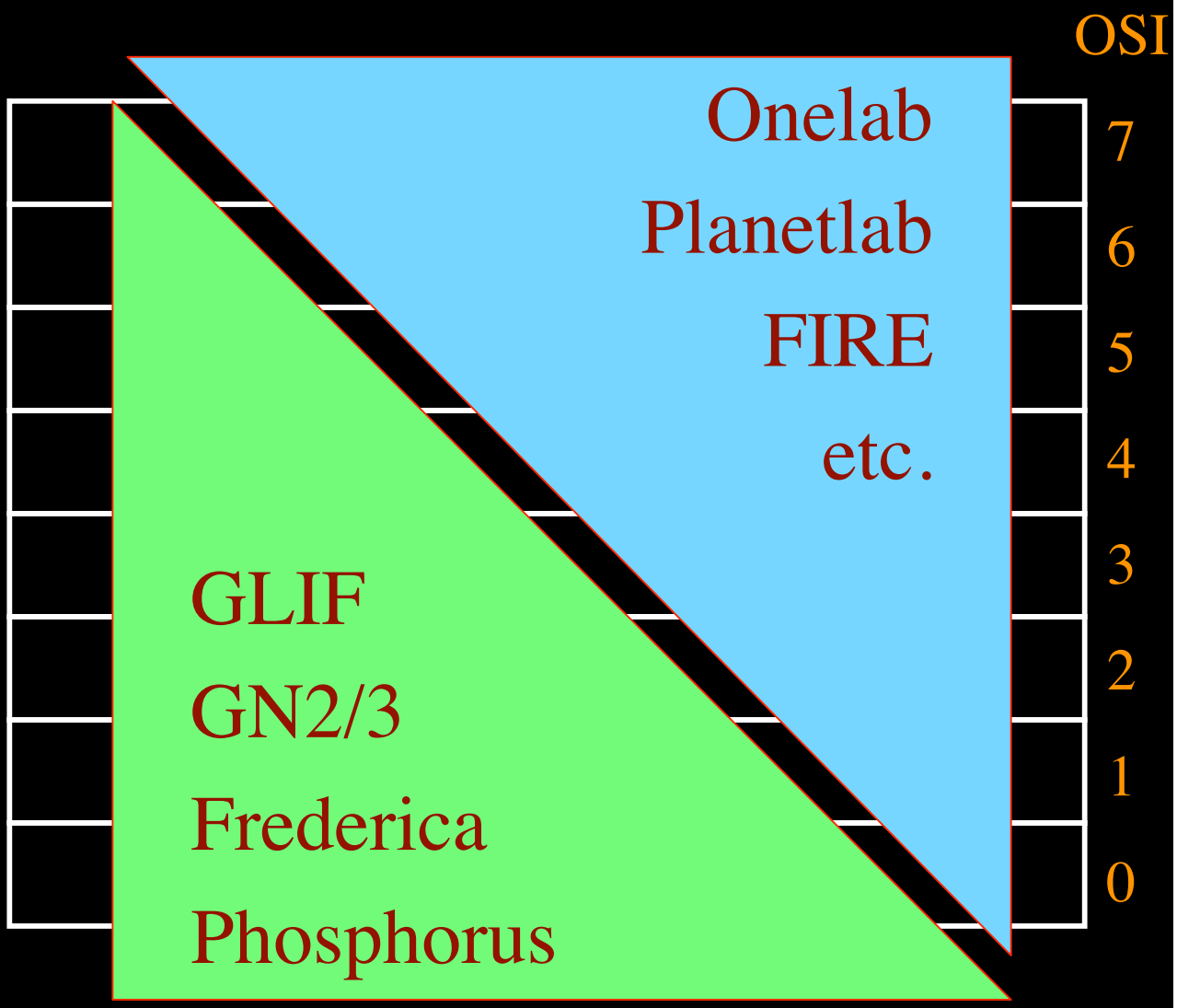
- StarPlane first DRAC WSS flip nov 2008
- NDL Multilayer pathfinding is being adopted
- Multi domain simulation NDL
- NDL & PROLOG
- Token based networking for inter domain GMPLS
- TBN solves problems for PhosPhorus-I2 interworking
- DRAC - IDC - Harmony LightPath setup
- SCARIe AuthoBAHN StarPlane demo
- HPDMnet High Quality video switching
- CineGrid Streaming, Storage and Forwarding
- Dark fiber SARA and SNE master extended to Oslo
- Programmable network demonstration with touch-table
- CineGrid portal streaming with PBT for QoS

Multi Layer Service Architecture



My view

- needs repeatable experiment
- needs QoS & lightpaths
- needs capacity and capability
- needs infrastructure descriptions

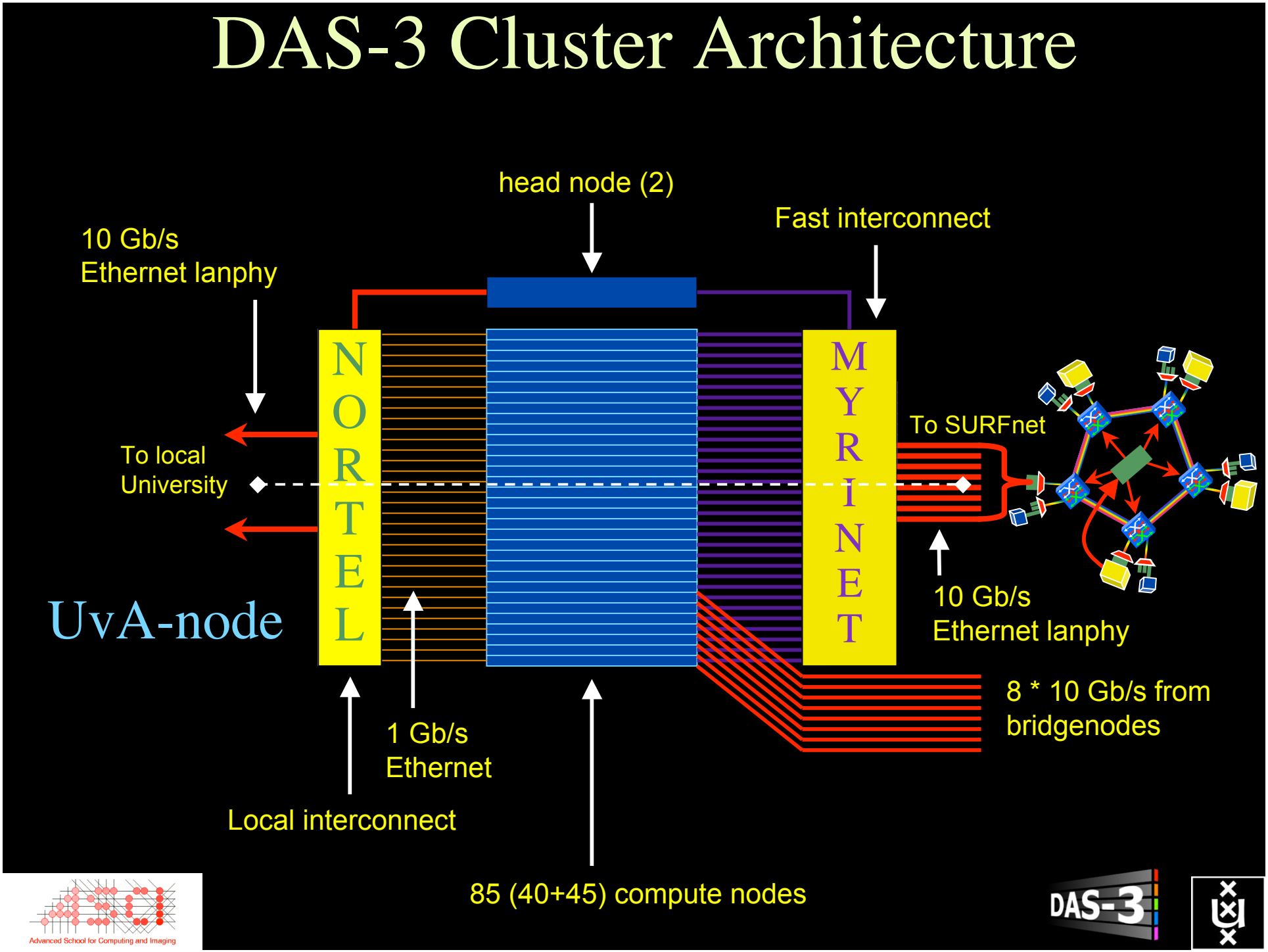


Sensor grid: instrumenting the dikes

First controlled breach occurred on sept 27th ‘08:

- 30000 sensors (microphones) to cover Dutch dikes
- focus on problem area when breach is to occur

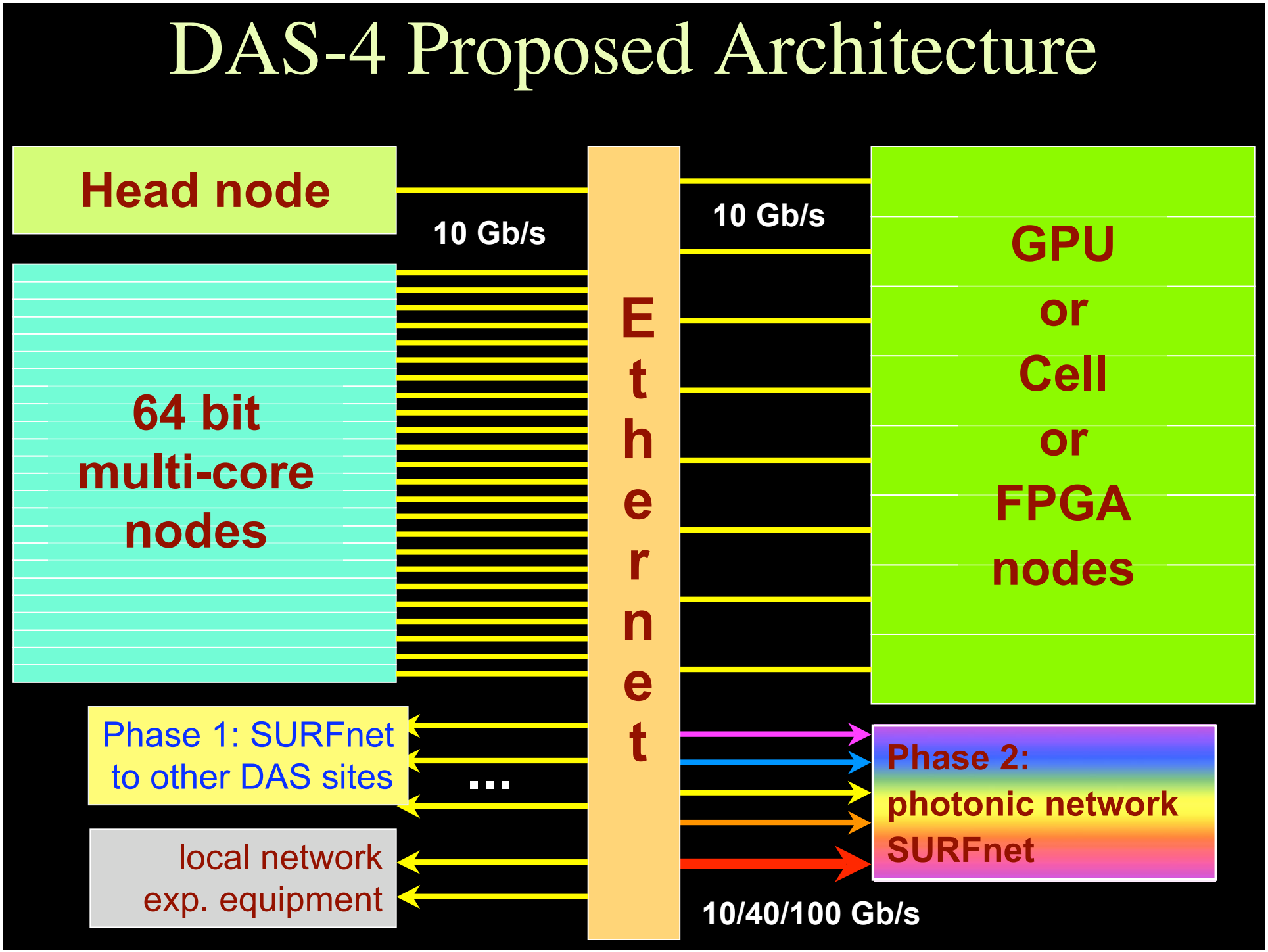
DAS-3 Cluster Architecture



Advanced School for Computing and Imaging

DAS-3

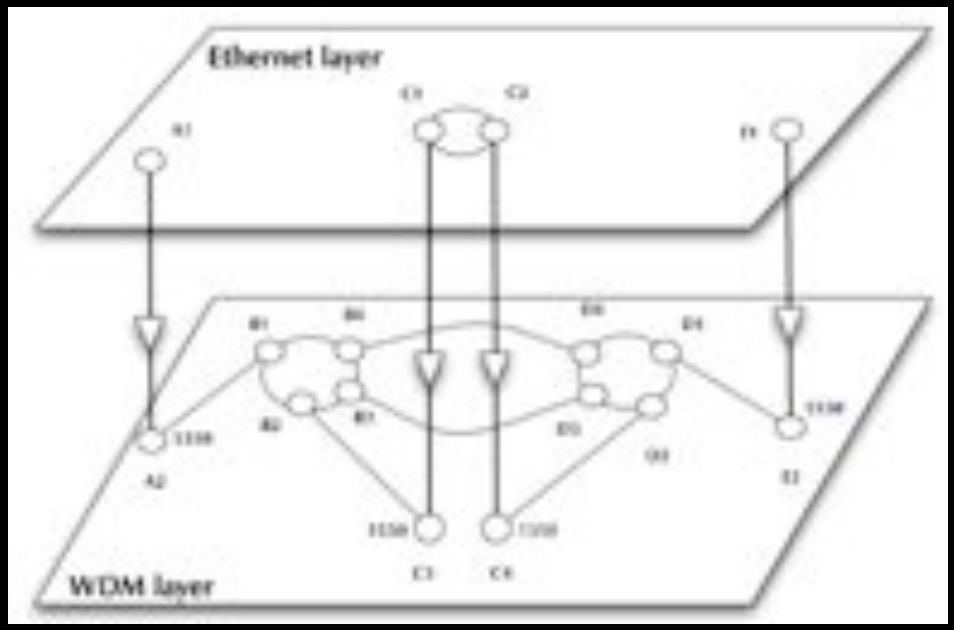
DAS-4 Proposed Architecture

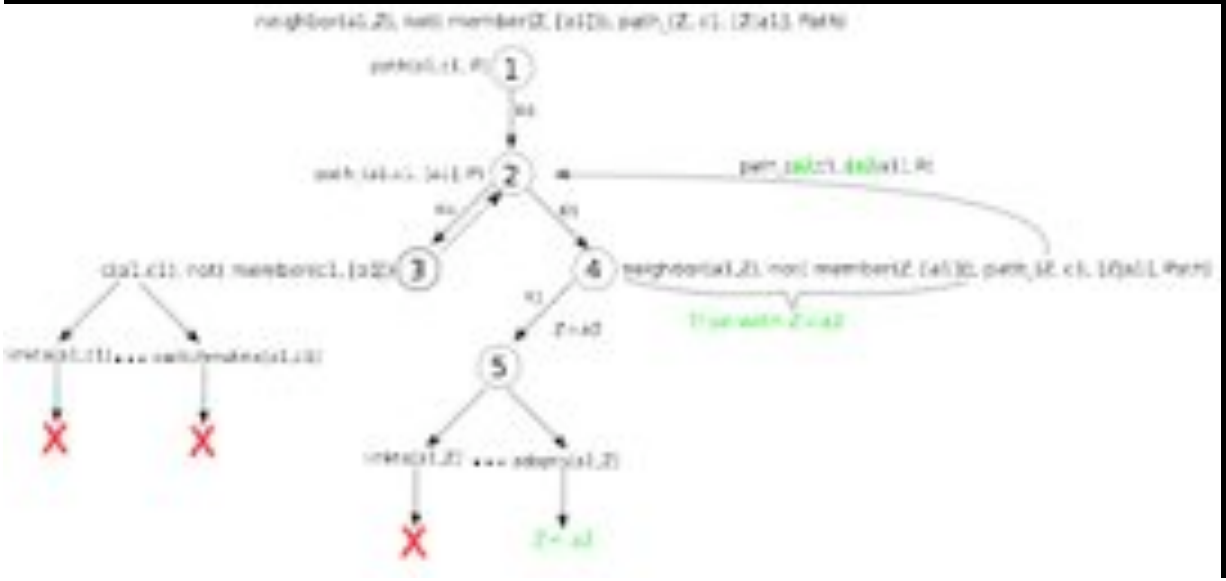


NDL + PROLOG

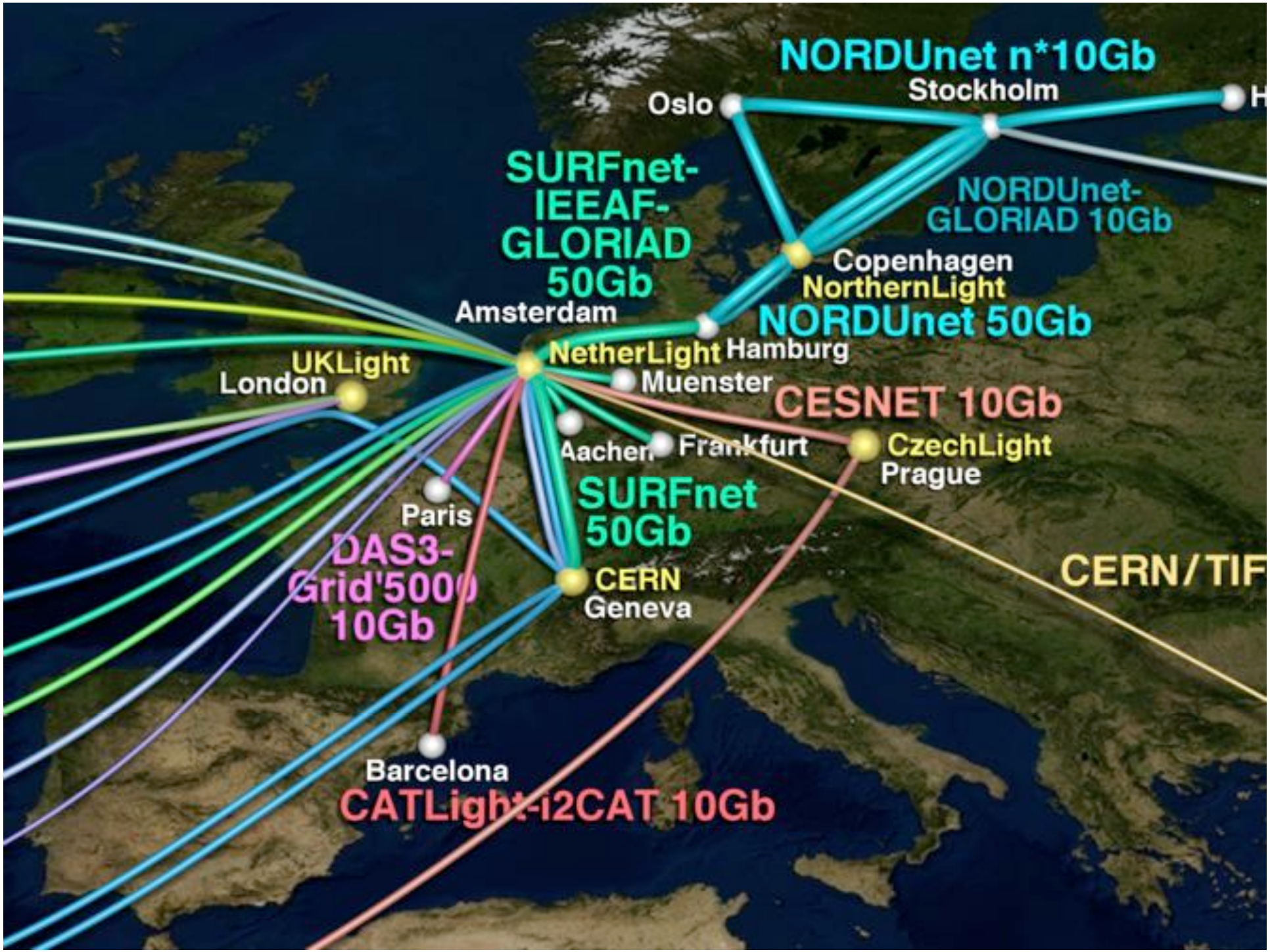
Research Questions:

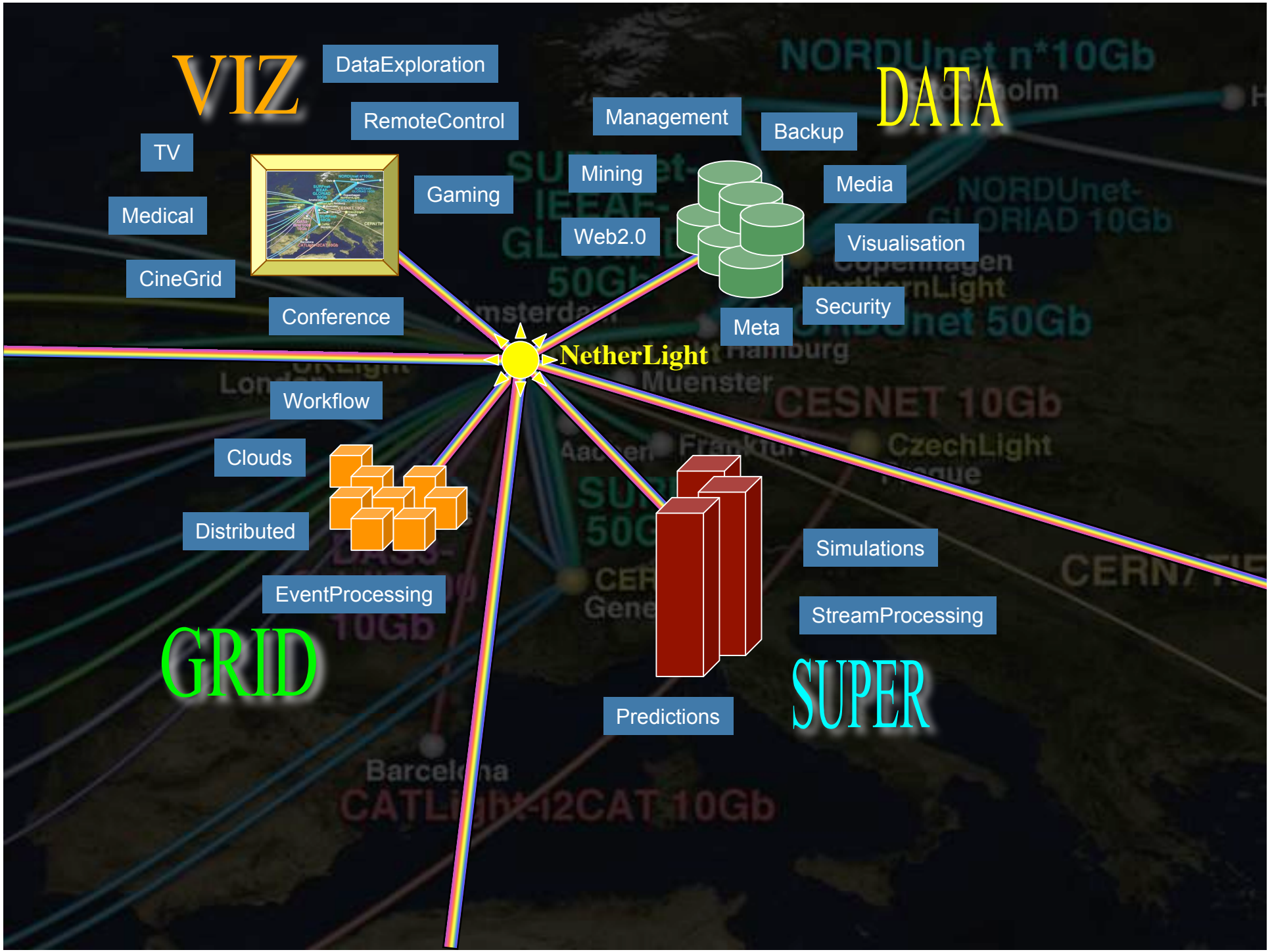
- order of requests
- complex requests
- Usable leftovers





- Reason about graphs
- Find sub-graphs that comply with rules


NORDUnet n*10Gb
Stockholm
Oslo
SURFnet-
IEEAF-
GLORIAD
50Gb
NORDUnet-
GLORIAD 10Gb
Copenhagen
NorthernLight
NORDUnet 50Gb
Amsterdam
NetherLight
Hamburg
UKLight
London
Muenster
CESNET 10Gb
Aachen
Frankfurt
CzechLight
Prague
SURFnet
50Gb
Paris
DAS3-
Grid'5000
10Gb
CERN
Geneva
CERN/TIF
Barcelona
CATLight-i2CAT 10Gb


VIZ
DataExploration
RemoteControl
TV
Medical
Gaming
CineGrid
Conference
DATA
Management
Backup
Mining
Media
Web2.0
Visualisation
Security
Meta
NetherLight
Workflow
Clouds
Distributed
EventProcessing
GRID
Simulations
StreamProcessing
SUPER
Predictions

The Dutch Booth #2409 at SC 2008 nov 17 – 20 2008 live !

This is an archived page, see you next year in Portland, Oregon.

The Dutch Booth #2603 at SC 2008, nov 15 - 21, Austin (Texas) (made by C.T. de Laat)

This page is best viewed with FireFox. Click on photo for a film recording.

NWO Netherlands Organisation for Scientific Research | UNIVERSITEIT VAN AMSTERDAM | SURFNET | sara | NCF

TNO | BiG Grid | NORTEL | NIKHEF | vl-e

Interviews

The TouchTable Interview by Rudolf Strijkers	The Tiled Panel Interview by Paul Wielinga	The Phosphorus Interview by Fred Wan
The HPDMnet Introduction Interview by Joe mambretti	The HPDMnet demo Interview by Herve Guy and Joe Mambretti	

SC08

TouchTable Demonstration @ SC08

Interactive programmable networks

Themes for next years

- Network modeling and simulation
- Cross domain Alien Light switching
- Green-Light
- Network and infrastructure descriptions & WEB2.0
- Reasoning about services
- Cloud Data - Computing
- Web Services based Authorization
- Network Services Interface (N-S and E-W)
- Fault tolerance, Fault isolation, monitoring
- eScience integrated services
- Data and Media specific services

RON evaluation

- The good
  - Lightweight bureaucracy
  - Adapt research to new insights
  - freedom for excellent ideas
- The bad
  - often difficult access to testbed, need “my own”
  - review process@SURFnet not clear but journals and community count for us
  - dissemination of results to production undefined (kzkr)
- The ugly
  - enormous delays in some network capabilities (example: StarPlane, DRAC)
  - DRAC pretty closed environment

Questions ?