# OnVector 2010

# The Power of Change!

Cees de Laat

EU

SURFnet

BSIK

NWO

University of Amsterdam

TNO

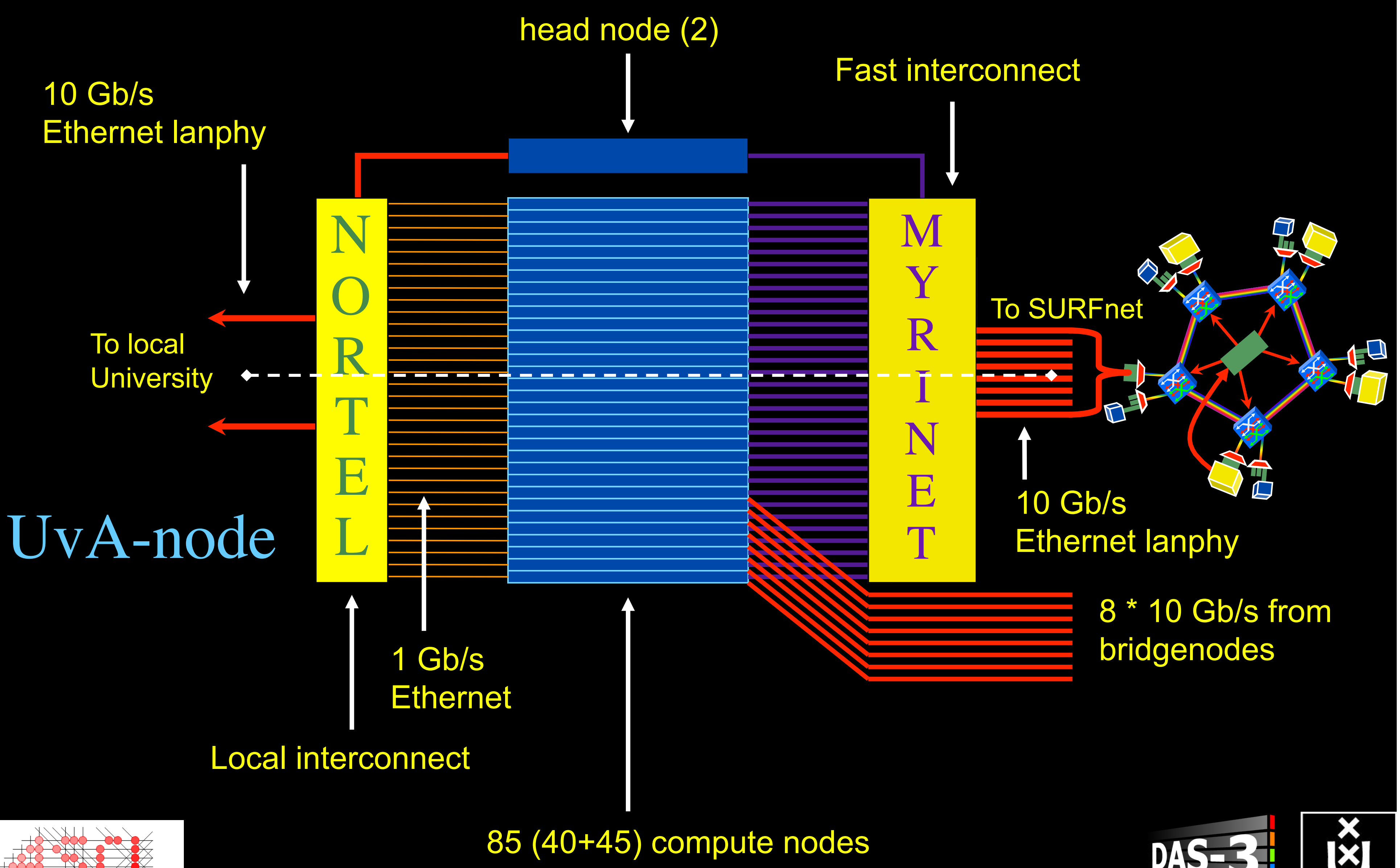
NCF

SURF NET

# Themes for next years

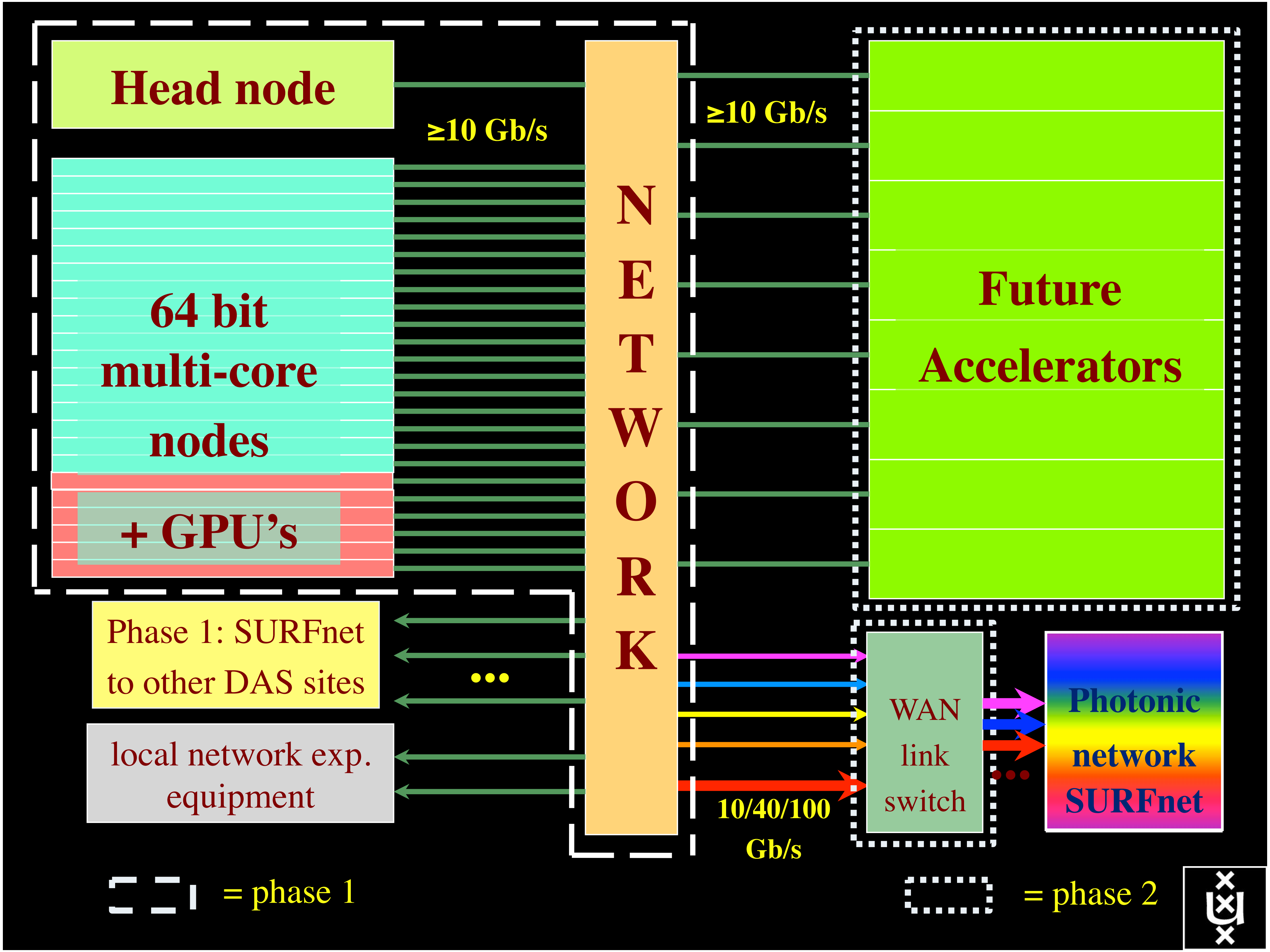
- 40 and 100 Gbit/s
- Network modeling and simulation
- Cross domain Alien Light switching
- GreenLight - GreenSonar
- Network and infrastructure descriptions & WEB2.0
- Reasoning about services
- Cloud Data - Computing
- Web Services based Authorization
- Network Services Interface (N-S and E-W)
- Fault tolerance, Fault isolation, Monitoring
- eScience integrated services
- Data and Media specific services
- → Smart e-Infrastructure

# DAS-3 Cluster Architecture

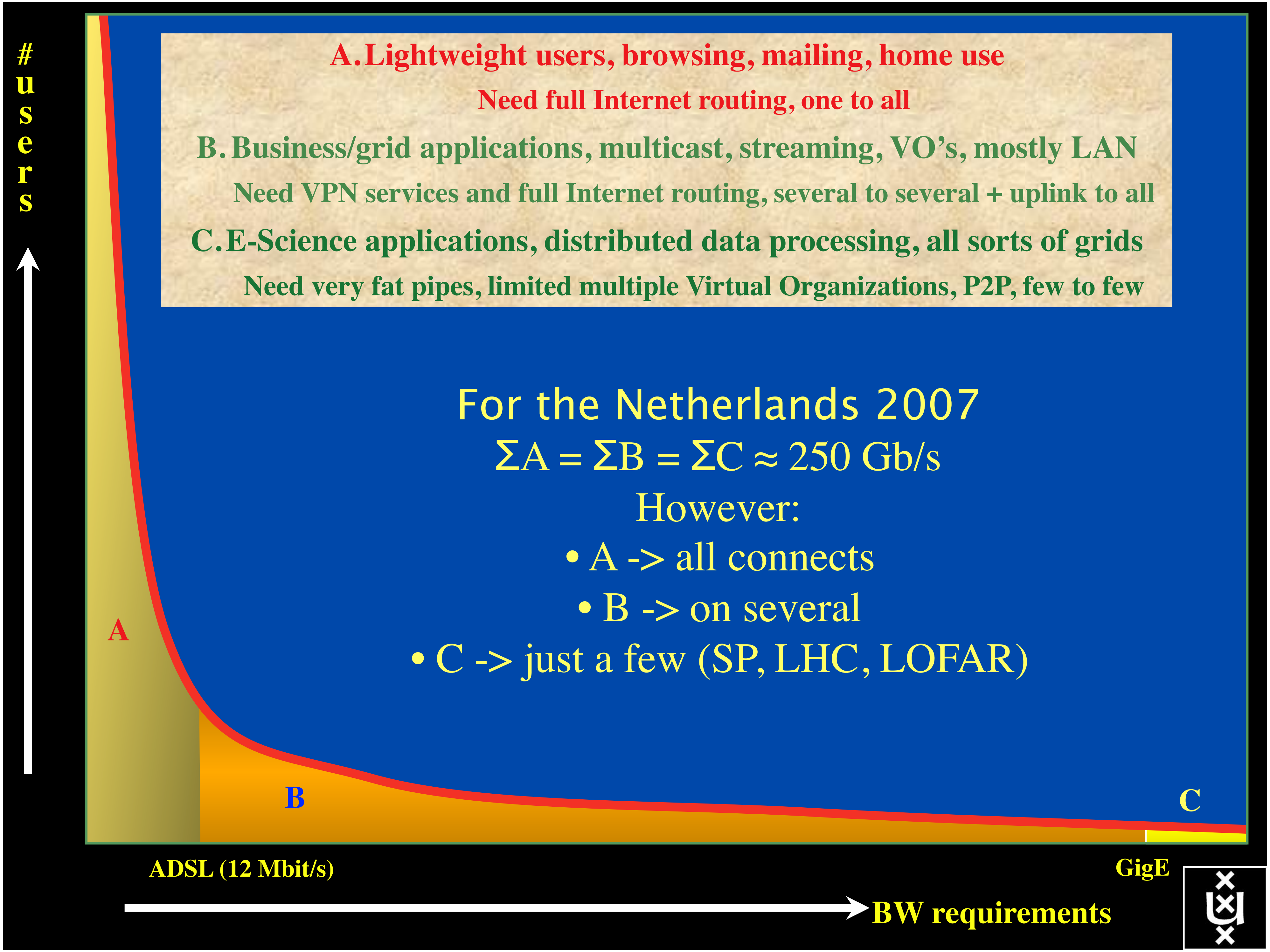


# Power is a big issue

- UvA cluster uses (max) 30 kWh
- 1 kWh ~ 0.1 €
- per year -> 26 k€/y
- add cooling 50% -> 39 k€/y
- Emergency power system -> 60 k€/y
- over 4 year = 240 kEuro for a 500 kEuro set.
- per rack 15 kWh is now normal
- YOU BURN HALF THE CLUSTER OVER ITS LIFETIME!



Head node
≥10 Gb/s
64 bit
multi-core
nodes
+ GPU's
N E T W O R K
≥10 Gb/s
Future
Accelerators
Phase 1: SURFnet
to other DAS sites
...
local network exp.
equipment
WAN
link
switch
10/40/100
Gb/s
Photonic
network
SURFnet
= phase 1
= phase 2

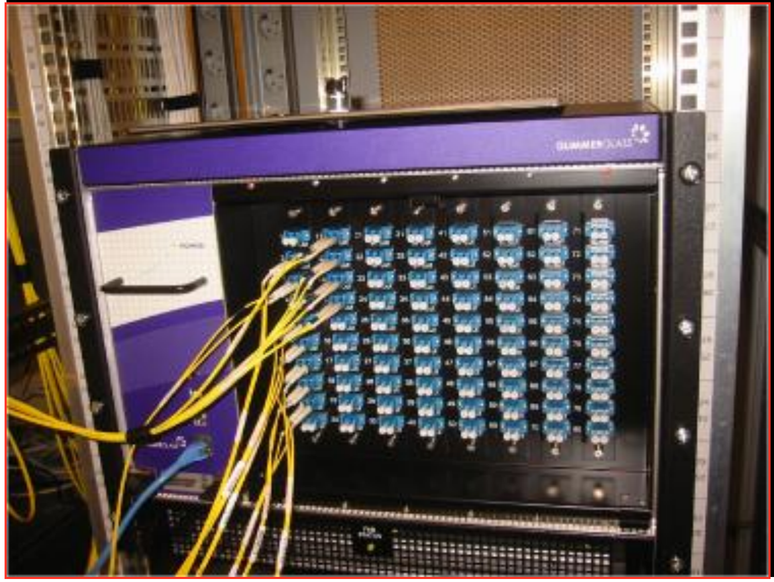


A. Lightweight users, browsing, mailing, home use
Need full Internet routing, one to all
B. Business/grid applications, multicast, streaming, VO’s, mostly LAN
Need VPN services and full Internet routing, several to several + uplink to all
C. E-Science applications, distributed data processing, all sorts 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)
# users
A
B
C
ADSL (12 Mbit/s)
GigE
BW requirements

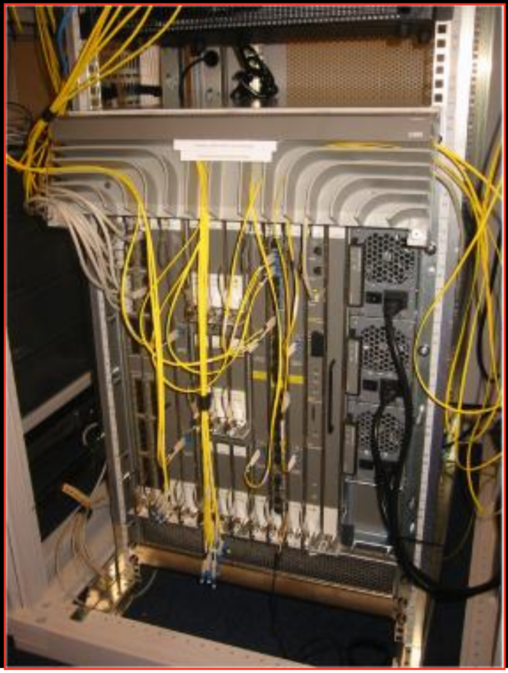
# Towards Hybrid Networking!

- 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 computing

Routers ← → Supercomputers

Ethernet switches ← → Grid & Cloud

Photonic transport ← → GPU's

What matters:

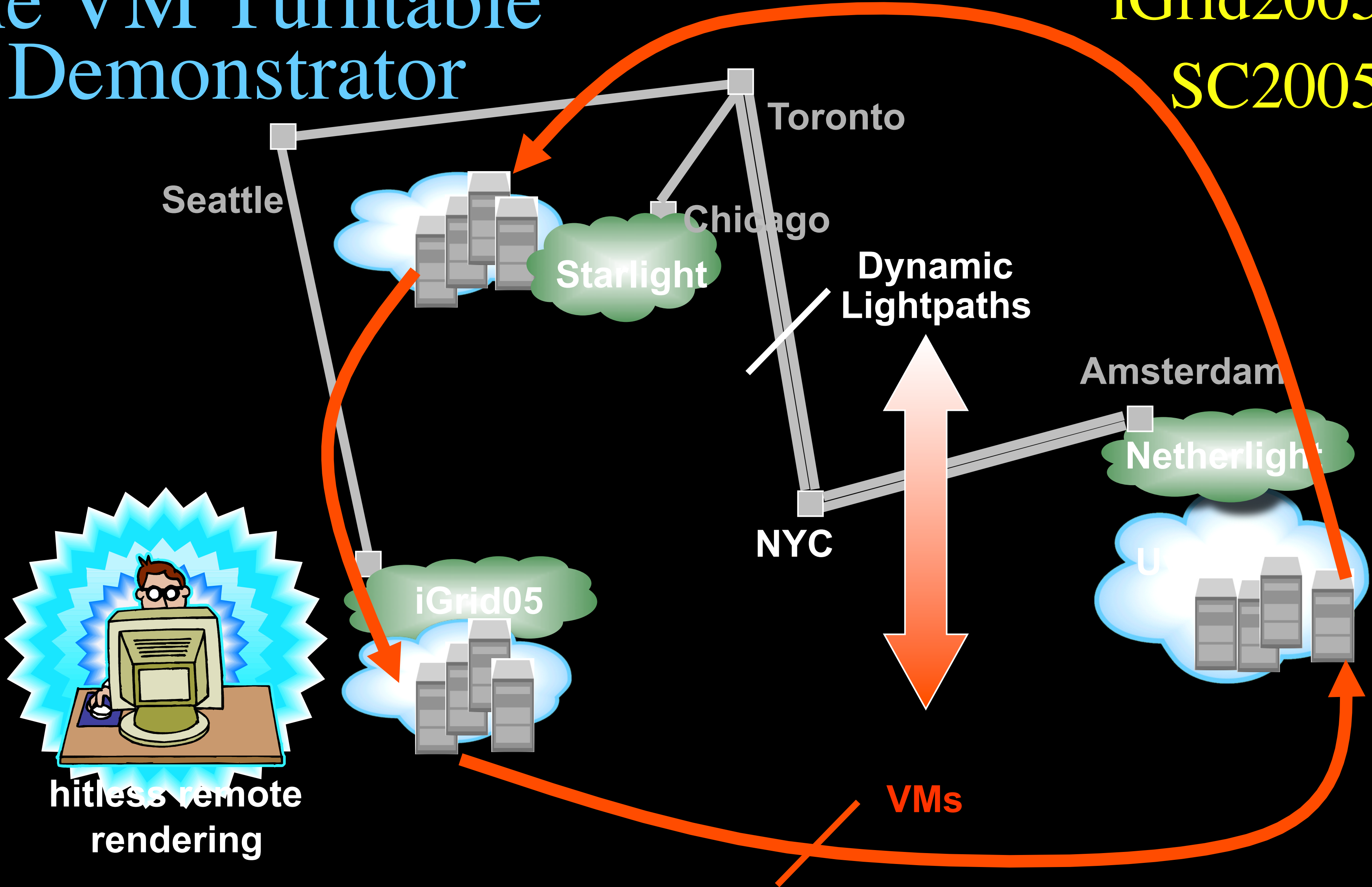
Energy consumption/multiplication

Energy consumption/bit transported

# The VM Turntable Demonstrator

iGrid2005

SC2005

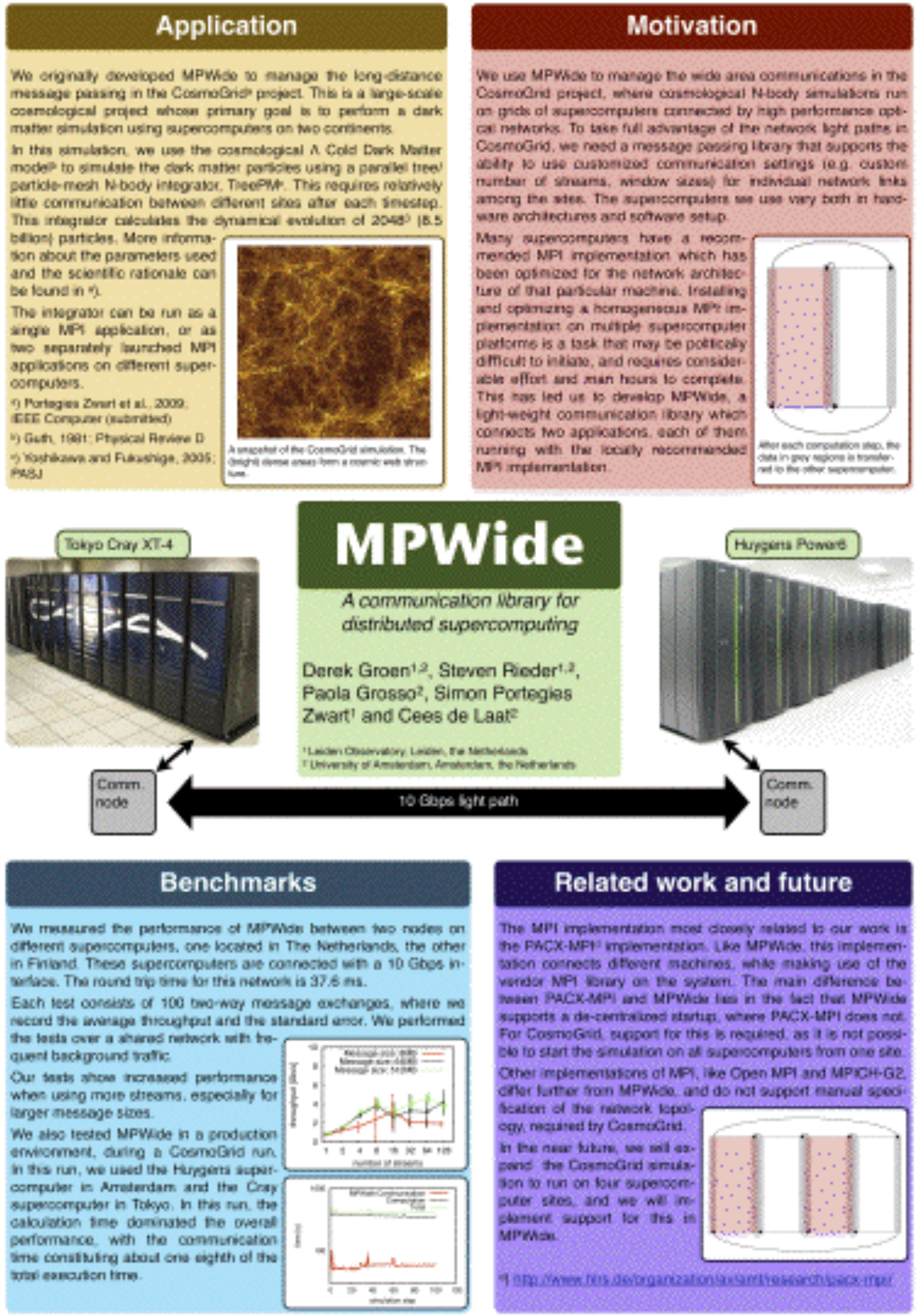



**The VMs that are live-migrated run an iterative search-refine-search workflow against data stored in different databases at the various locations. A user in San Diego gets hitless rendering of search progress as VMs spin around**

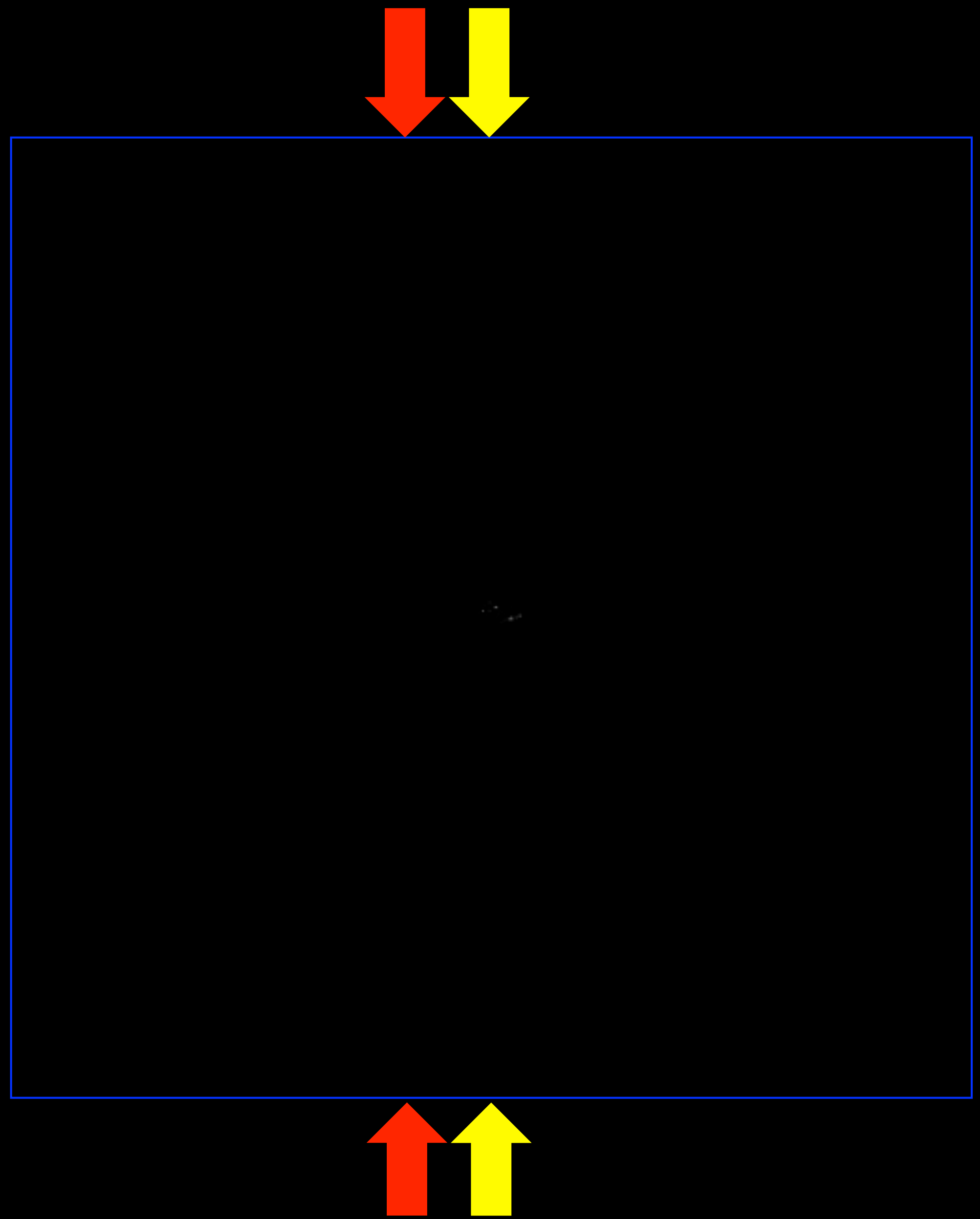
# CosmoGrid

# Supercomputing Grid across Continents and Oceans

## And yes, it works!

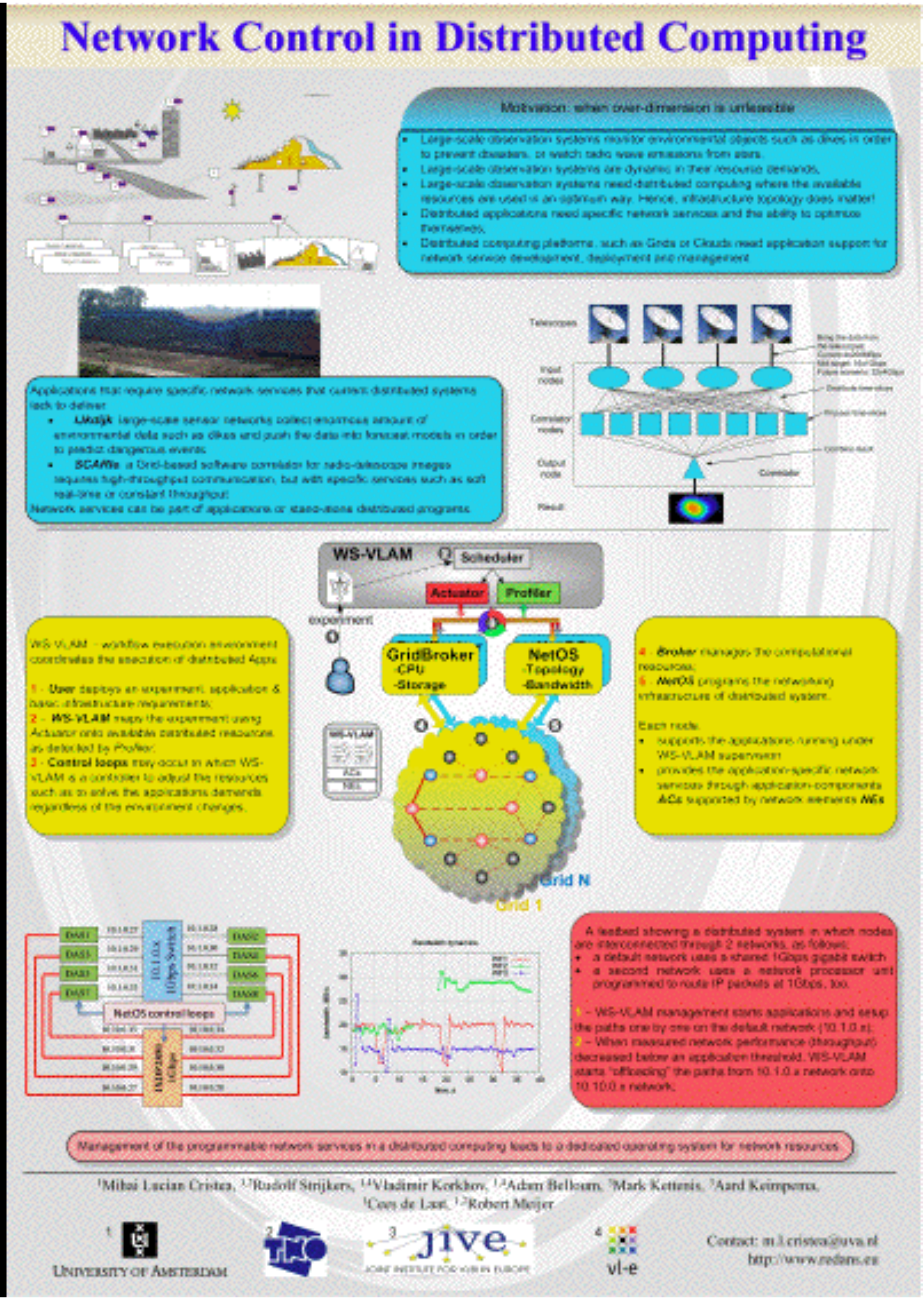


# Auto-balancing Supers

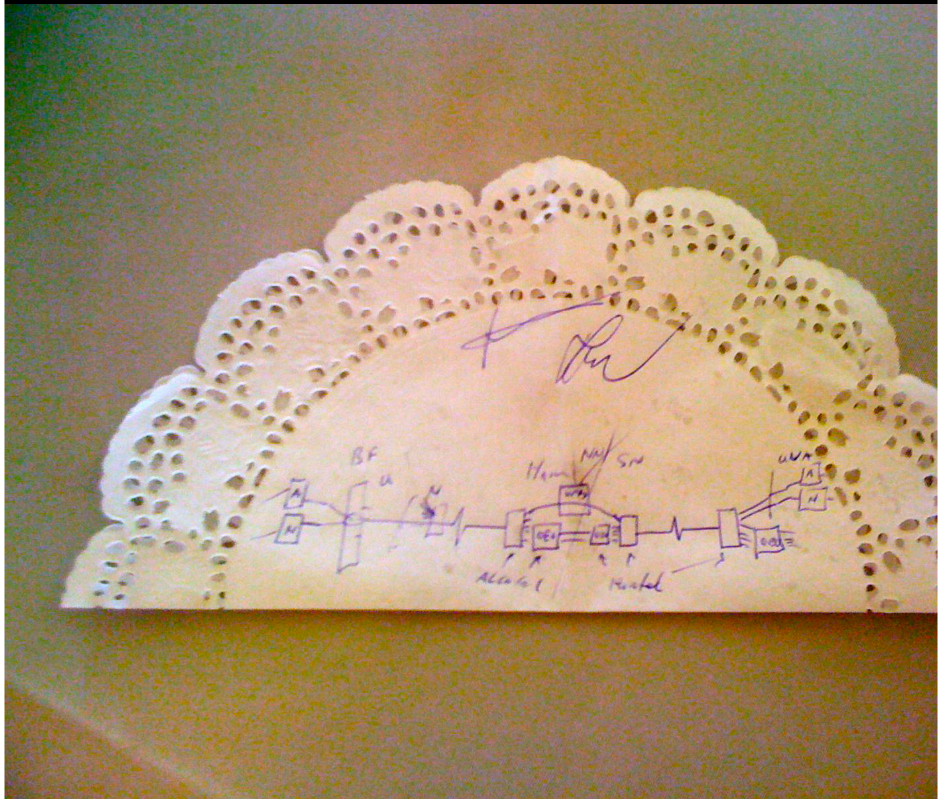


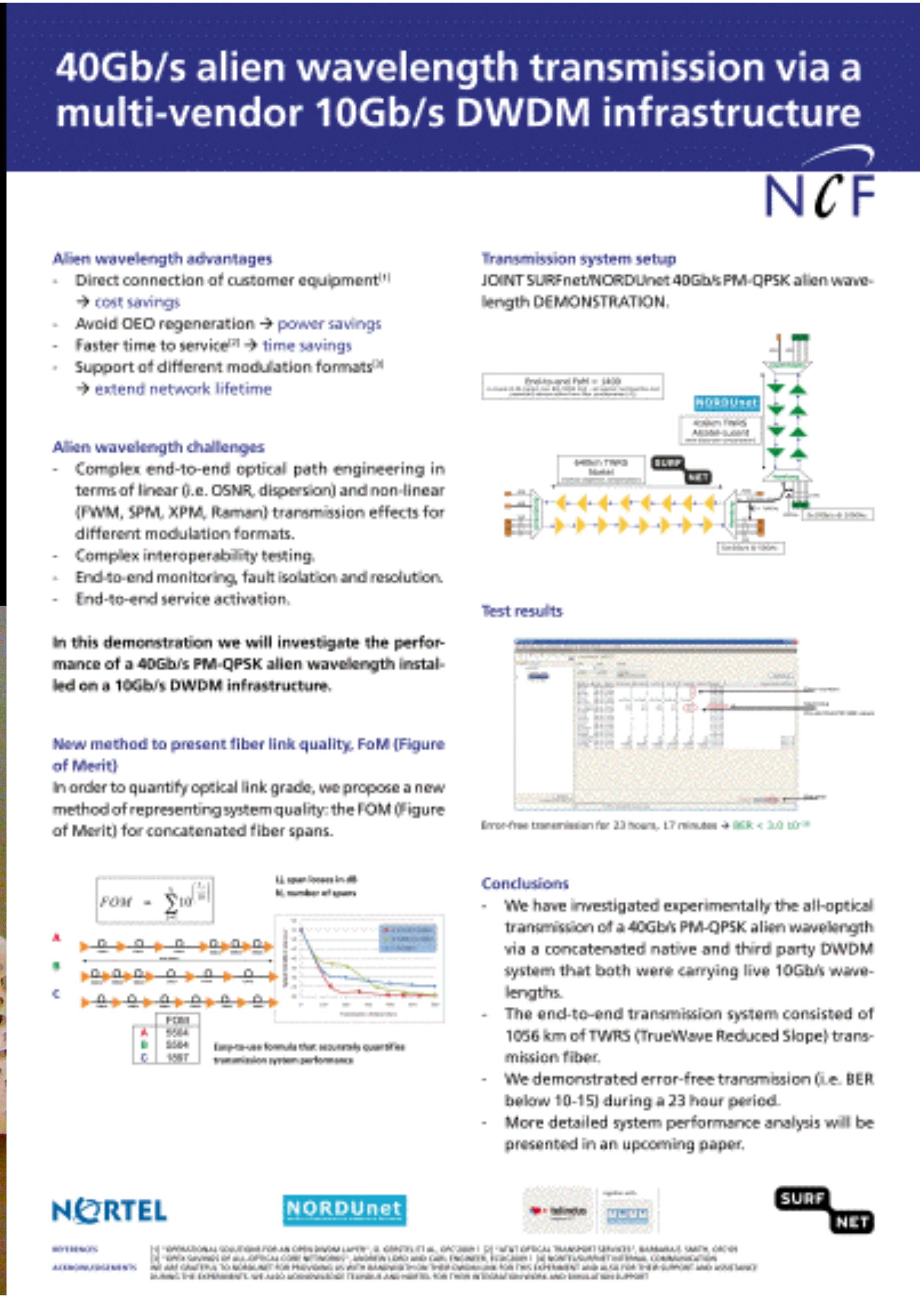
# Interactive programmable networks

# SCARIe Programmable networks to distribute work



# Alien light
# From idea to realisation!





## 40Gb/s alien wavelength transmission via a multi-vendor 10Gb/s DWDM infrastructure

NCF

### Alien wavelength advantages

- Direct connection of customer equipment[1]
  → cost savings
- Avoid OEO regeneration → power savings
- Faster time to service[2] → time savings
- Support of different modulation formats[3]
  → extend network lifetime

### Alien wavelength challenges

- Complex end-to-end optical path engineering in terms of linear (i.e. OSNR, dispersion) and non-linear (FWM, SPM, XPM, Raman) transmission effects for different modulation formats.
- Complex interoperability testing.
- End-to-end monitoring, fault isolation and resolution.
- End-to-end service activation.

**In this demonstration we will investigate the performance of a 40Gb/s PM-QPSK alien wavelength installed on a 10Gb/s DWDM infrastructure.**

### New method to present fiber link quality, FoM (Figure of Merit)

In order to quantify optical link grade, we propose a new method of representing system quality: the FOM (Figure of Merit) for concatenated fiber spans.

Easy-to-use formula that accurately quantifies transmission system performance

### Transmission system setup

JOINT SURFnet/NORDUnet 40Gb/s PM-QPSK alien wavelength DEMONSTRATION.

### Test results

Error-free transmission for 23 hours, 17 minutes → BER < 3.0 10⁻¹⁶

### Conclusions

- We have investigated experimentally the all-optical transmission of a 40Gb/s PM-QPSK alien wavelength via a concatenated native and third party DWDM system that both were carrying live 10Gb/s wavelengths.
- The end-to-end transmission system consisted of 1056 km of TWRS (TrueWave Reduced Slope) transmission fiber.
- We demonstrated error-free transmission (i.e. BER below 10-15) during a 23 hour period.
- More detailed system performance analysis will be presented in an upcoming paper.

NORTEL NORDUnet SURF NET

# Path finding in multi-domain multi-layer networks

A new approach based on declarative programming

P. Grosso, A. Taal, L. Xu, J. v/d Ham, C. de Laat

# Multi layer multi domain networks

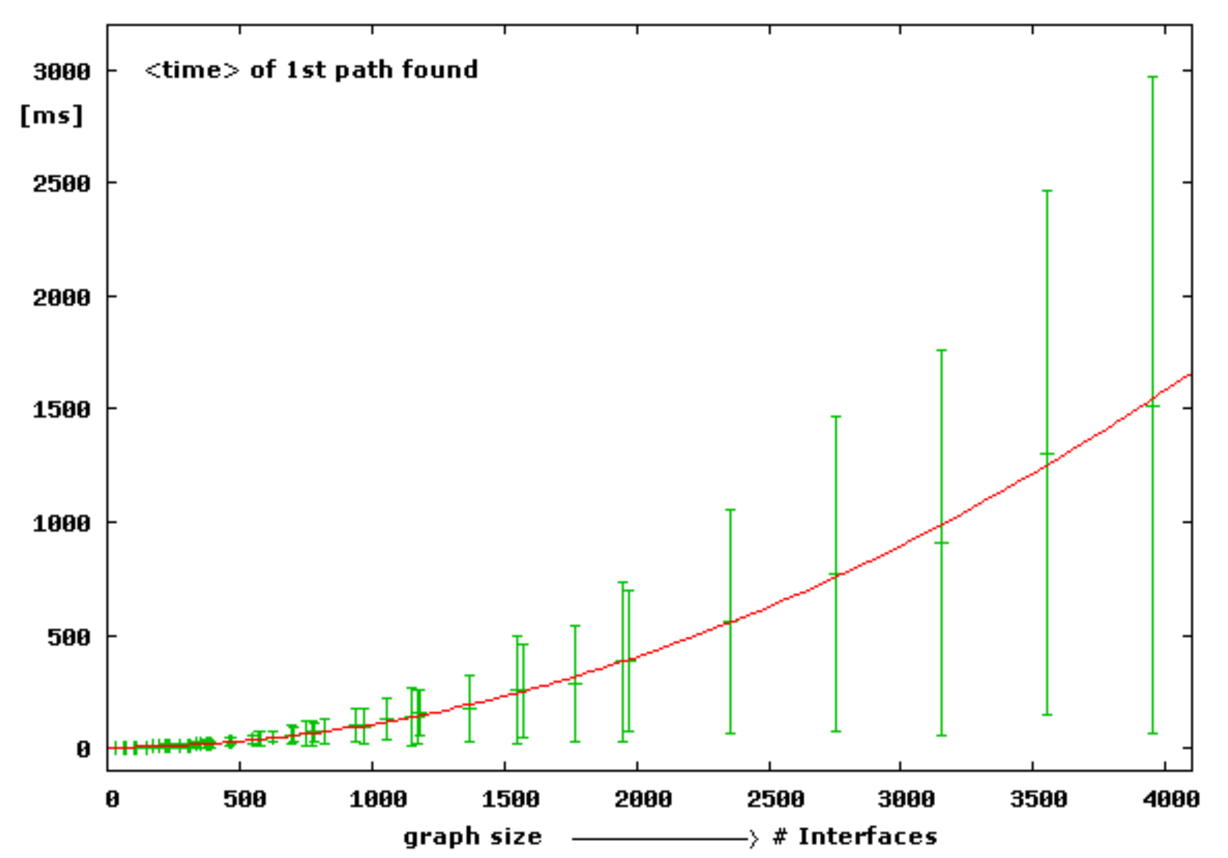
The networks for e-Science where applications use dedicated optical circuits.

*Is declarative programming more suitable to find paths in multi-domain multi-layer networks? Especially in presence of constraints and complex requests?*

Our approach:

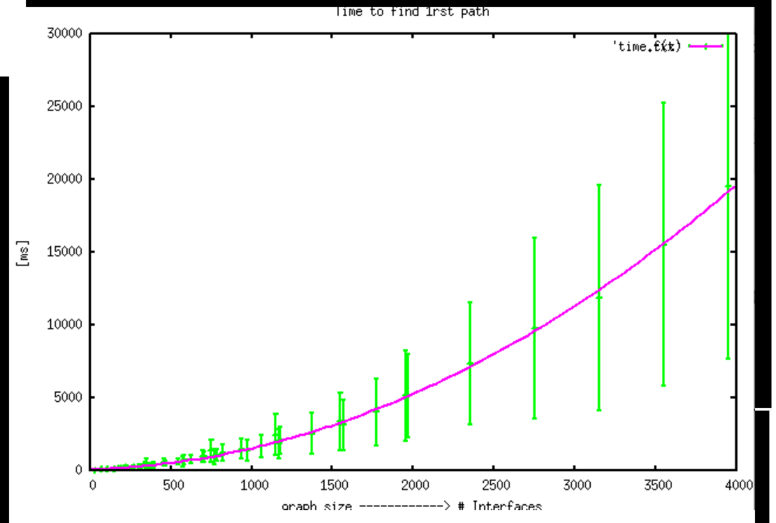
1. We generate BA network graphs with a varying number of domains and nodes. Barabasi-Albert scale free graphs are a good representation of these networks.
2. We represent the graphs in NDL – Network Description Language, the RDF schemas.
3. We load the RDF files in Prolog and Python programs
4. We perform a modified DFS –Depth First Search- algorithm to find paths.

# Single layer networks: results




- *Number of interfaces,*
- *given N nodes per domain D*
- *4*(D-2) + D*4*(N-2) for D > 2*

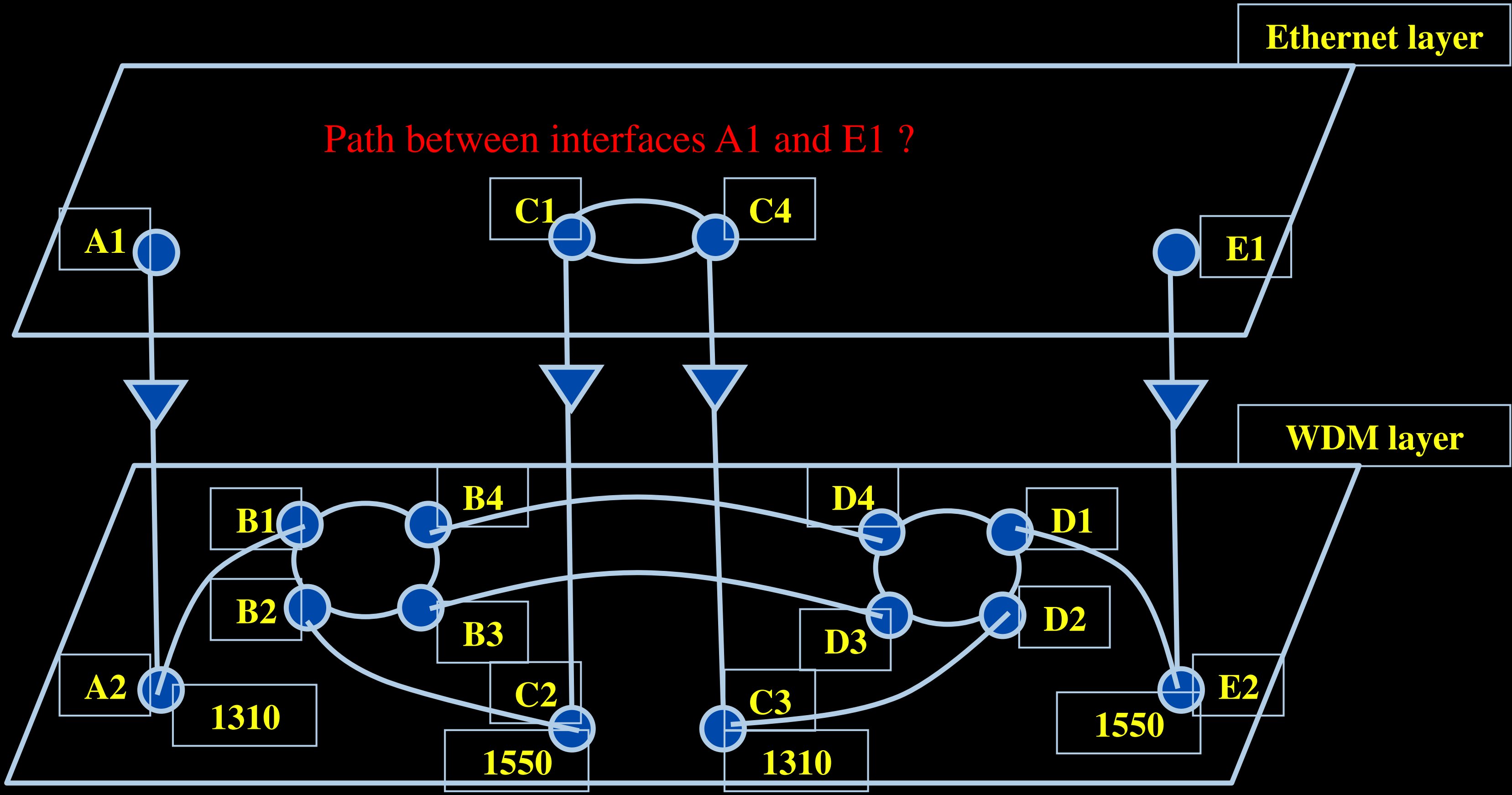
## *Pynt-based DFS*




## *Prolog DFS*

- Prolog time to find first path shorter than Python time.
- We observe a quadratic dependence.
- Length of paths found comparable.

# Multi-layer network




**Prolog rule:**

```
linkedto( Intf1, Intf2, CurrWav ):-                      %-- is there a link between  Intf1 and Intf2 for wavelength CurrWav ?
    rdf_db:rdf( Intf1, ndl:'layer', Layer ),             %-- get layer of interface Intf1 → Layer
    Layer == 'wdm#LambdaNetworkElement',                 %-- are we at the WDM-layer ?
    rdf_db:rdf( Intf1, ndl:'linkedTo', Intf2 ),          %-- is Intf1 linked to Intf2 in the RDF file?
    rdf_db:rdf( Intf2, wdm:'wavelength', W2 ),           %-- get wavelength of Intf2 → W2
    compatible_wavelengths( CurrWav, W2 ).               %-- is CurrWav compatible with W2 ?
```

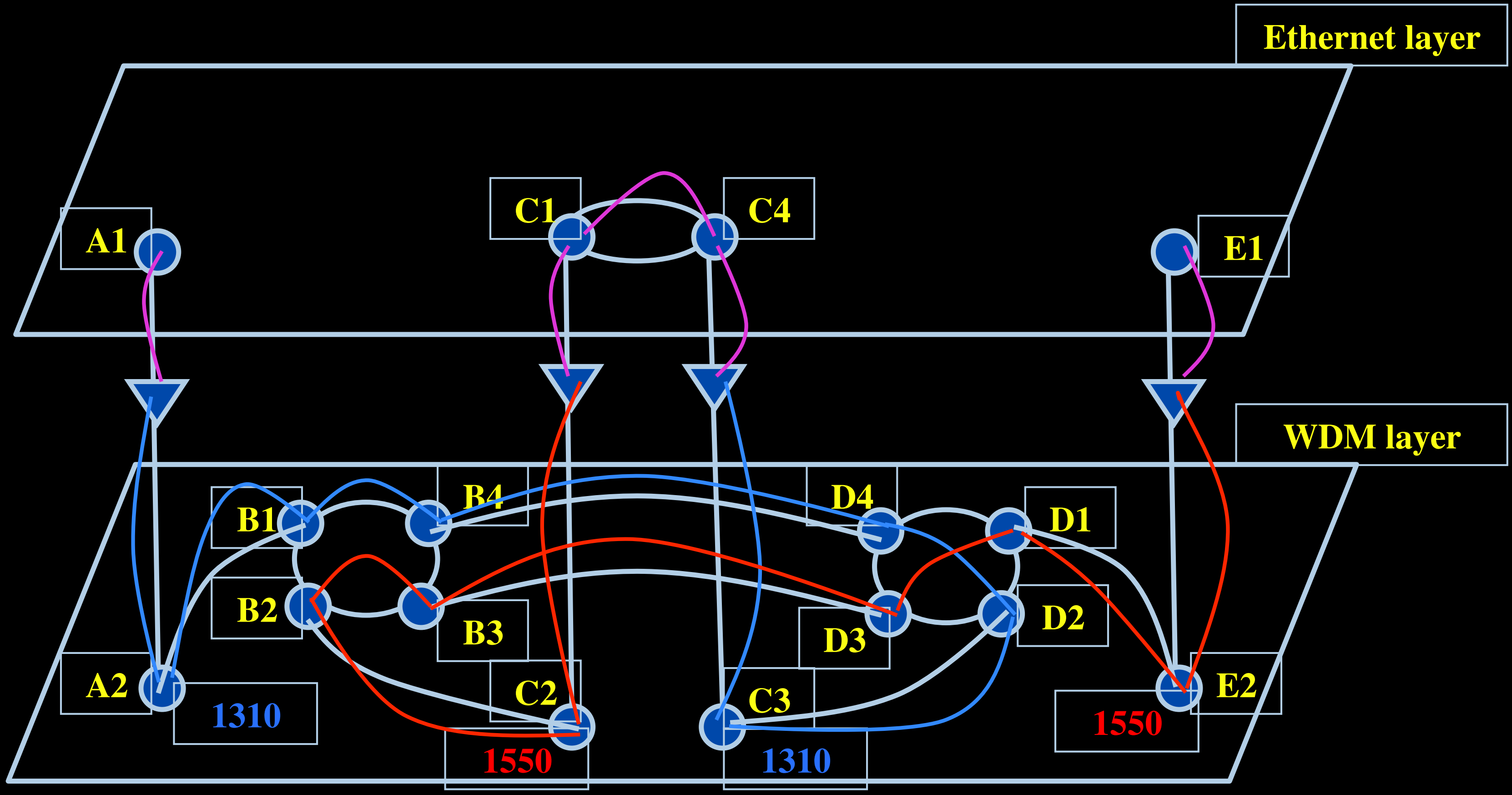
**linkedto( B4, D4, CurrWav ) is true for any value of CurrWav**

**linkedto( D2, C3, CurrWav) is true if CurrWav == 1310**

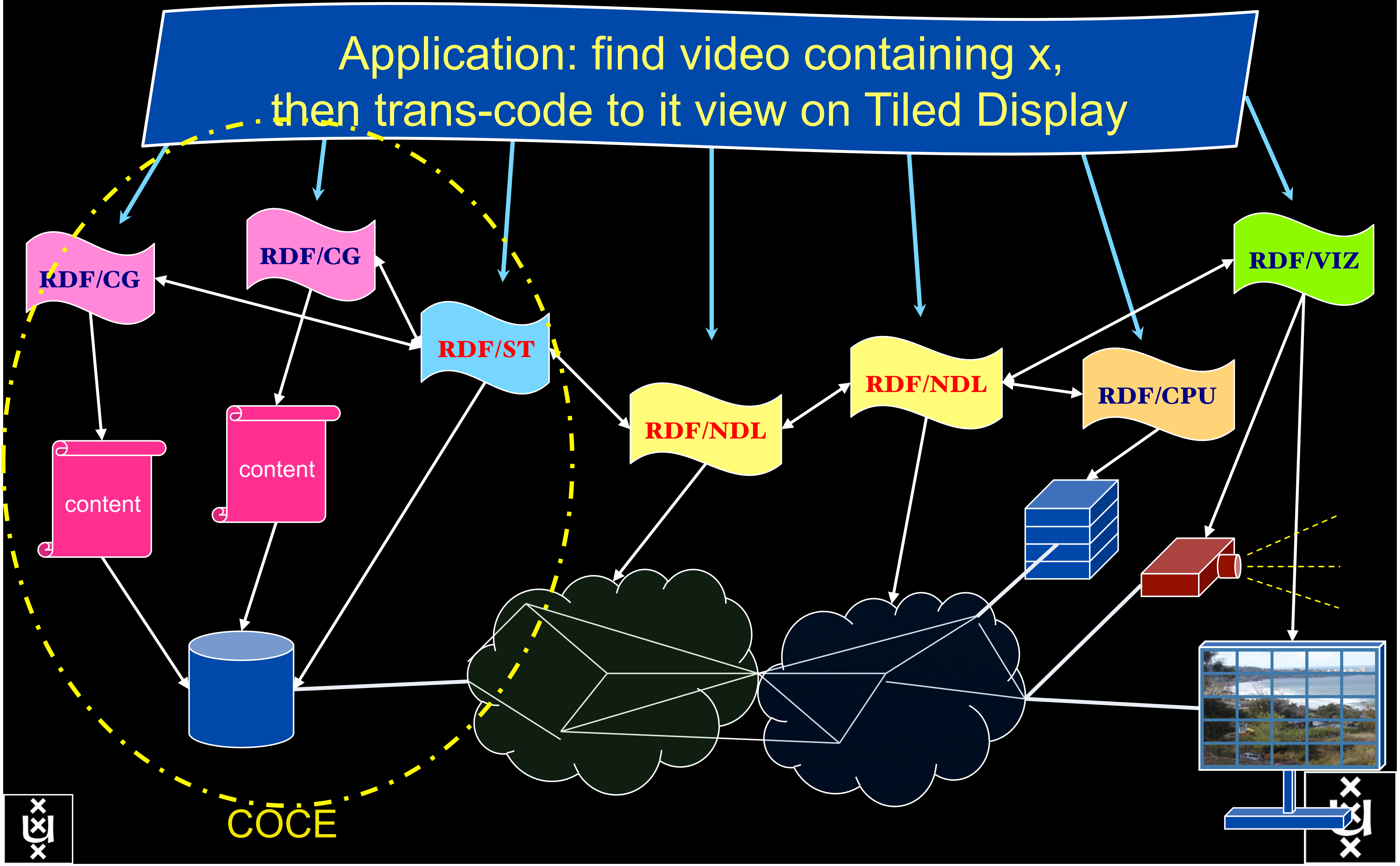
# Multi-layer

Path between interfaces A1 and E1:

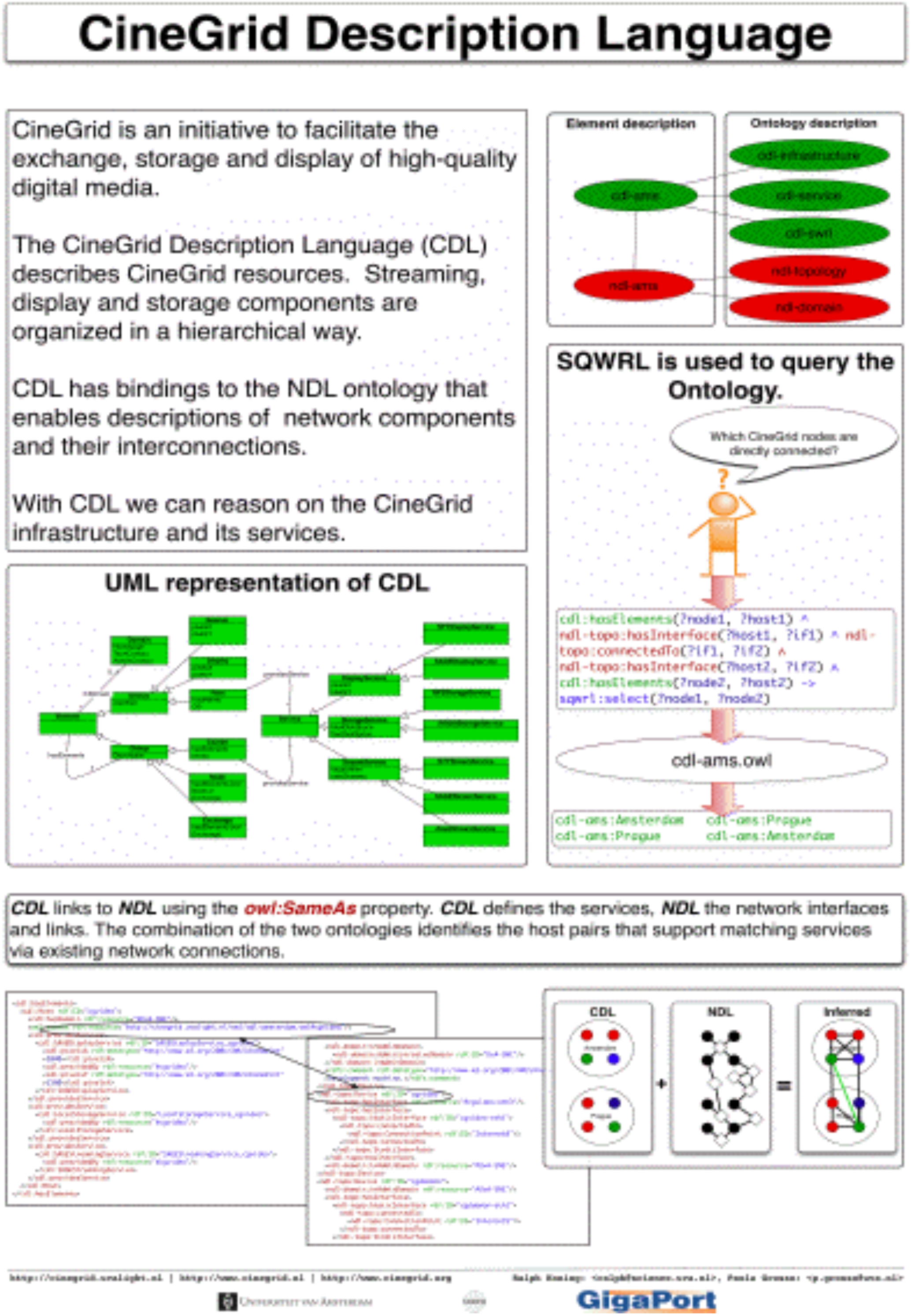
A1-A2-B1-B4-D4-D2-C3-C4-C1-C2-B2-B3-D3-D1-E2-E1



# RDF describing Infrastructure
# "I want"



Applications and Networks become aware of each other!



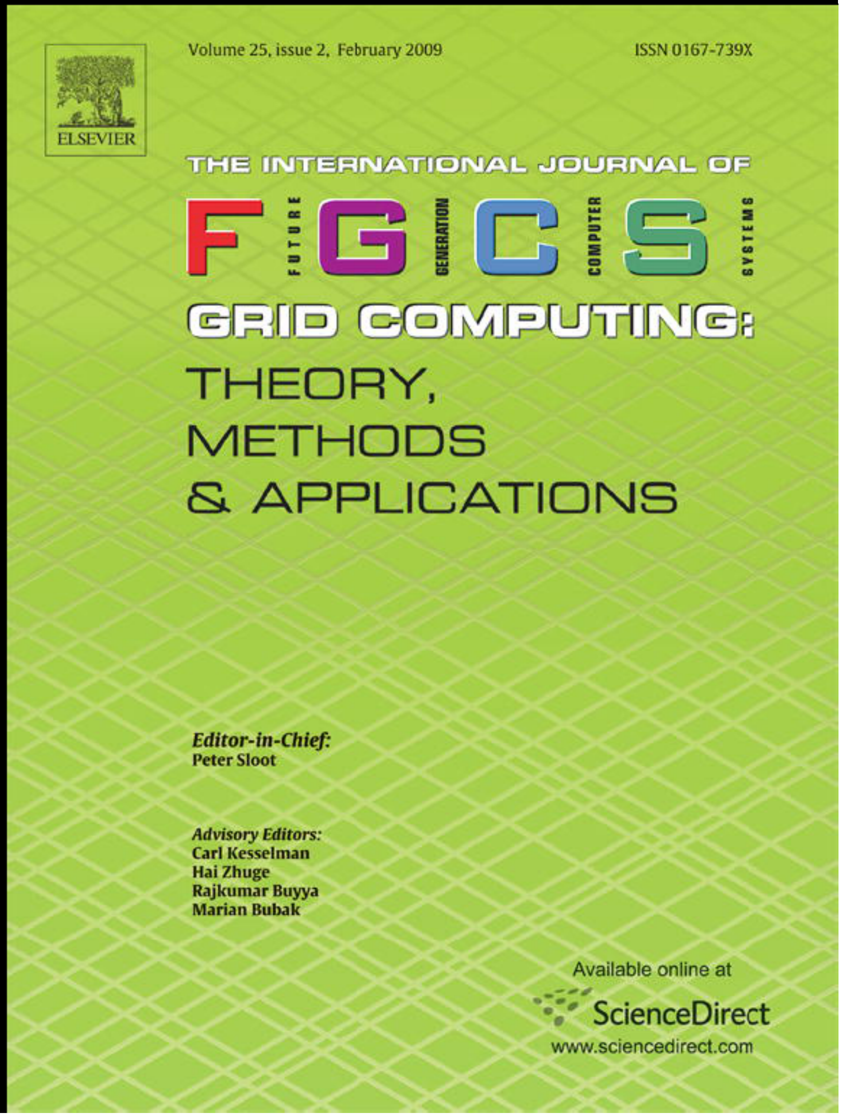
# Last Thoughts

- Energy consumption is the main issue
- Cloud Computing as solution
- We did Hybrid networking
  - now hybrid computing, what else?
- Network photonics developments
- GreenSonar (aka PerfSonar)
- Smart energy conscious infrastructure

# Need for Scientific Publications!
# Call for papers!

- Guest Editors:
  Naohisa Ohta & Paul Hearty
  & Cees de Laat
- Special section on CineGrid!
- 6-8 papers in a section
- Submission via:
  http://ees.elsevier.com/fgcs/

- CineGrid section submission site is up
- Info: delaat@uva.nl
- Submission deadline March 1st 2010



***The Power of Change?***

***OR***

***The Change of Power!***

***sc09.delaat.net***

***Questions ?***

p.s. On teleportation: look at prof. Eric Verlinde's work on emergent phenomena relating bit density, entropy and gravity!

**http://staff.science.uva.nl/%7Eerikv/**

SURF NET