Internet Innovation to support Science & Education.

Cees de Laat

EU

COMMIT

UvA

NWO

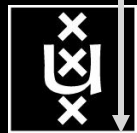
PID/EFRO

SURFnet

NLESC

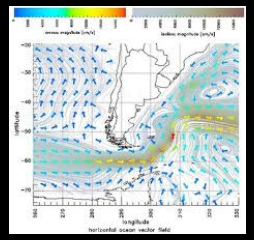
TNO

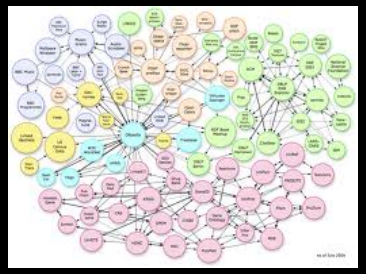
NWO/ncf



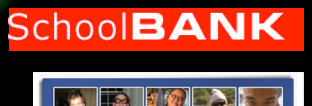
Internet developments

... more data!





... more realtime!



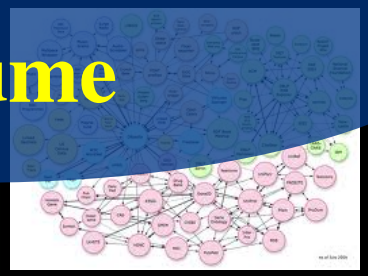
... more users!

Internet developments

... more data!

Speed

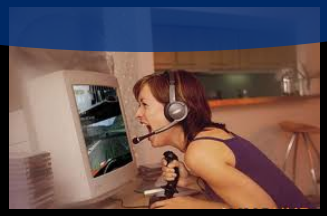
Volume



Deterministic

Real-time

more realtime!



Scalable

Secure

... more users!

Internet is a Billion - Business!

Google	197
Amazon	83
Facebook	50
BAIDU	37
eBay	36
Yahoo	22
PriceLine	21
SalesForce	18
F5 Networks	11
CheckPoint	9
NetFlix	9
Expedia	7


guardian.co.uk

Monday 3 January 2011

News | Sport | Comment | Culture | Business | Money | Life & style

News › Technology › Facebook

Facebook's value swells to $50bn after Goldman Sachs investment

Deal underlines Facebook's power and fuels rumours that Mark Zuckerberg is preparing a stock market flotation

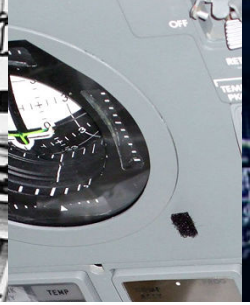
e.g.: Exxon Mobil 368

Apple Inc. 565 (2001: 333)

1 miljard in 100$ biljetten

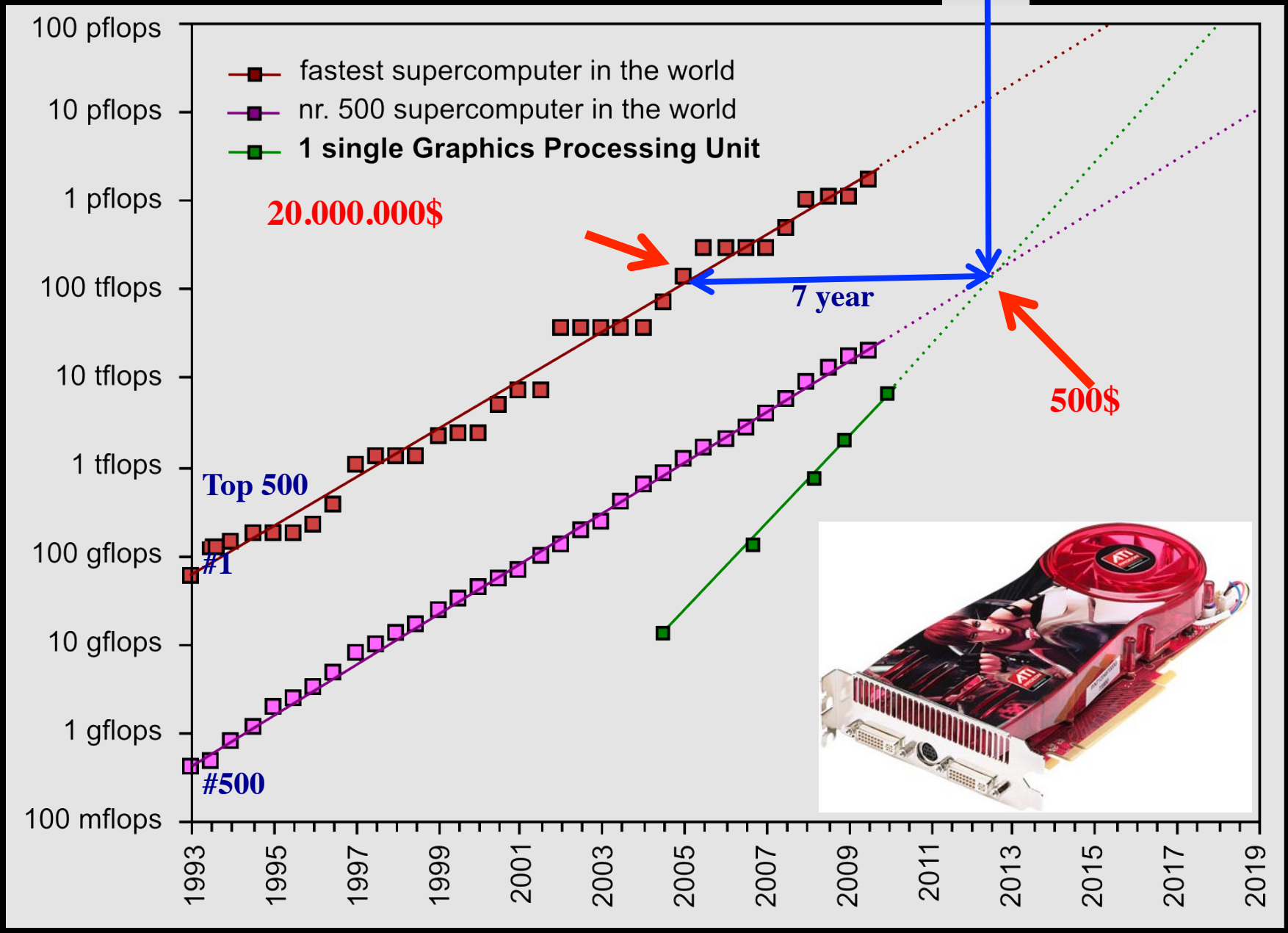

TEMP
GIMBAL LOCK
PROG
RESTART
TRACKER
VERB
NOUN
+ 7 8 9
− 4 5 6
0 1 2


UPLINK ACTY
TEMP
AUTO
GIMBAL LOCK
PROG
RESTART
NO ATT
TRACKER
STBY
KEY REL
OPR ERR
COMP ACTY
PROG
VERB
NOUN
VERB
NOUN
+ 7 8 9 CLR
− 4 5 6 STBY
1 2 3 KEY REL
ENTR
RSET

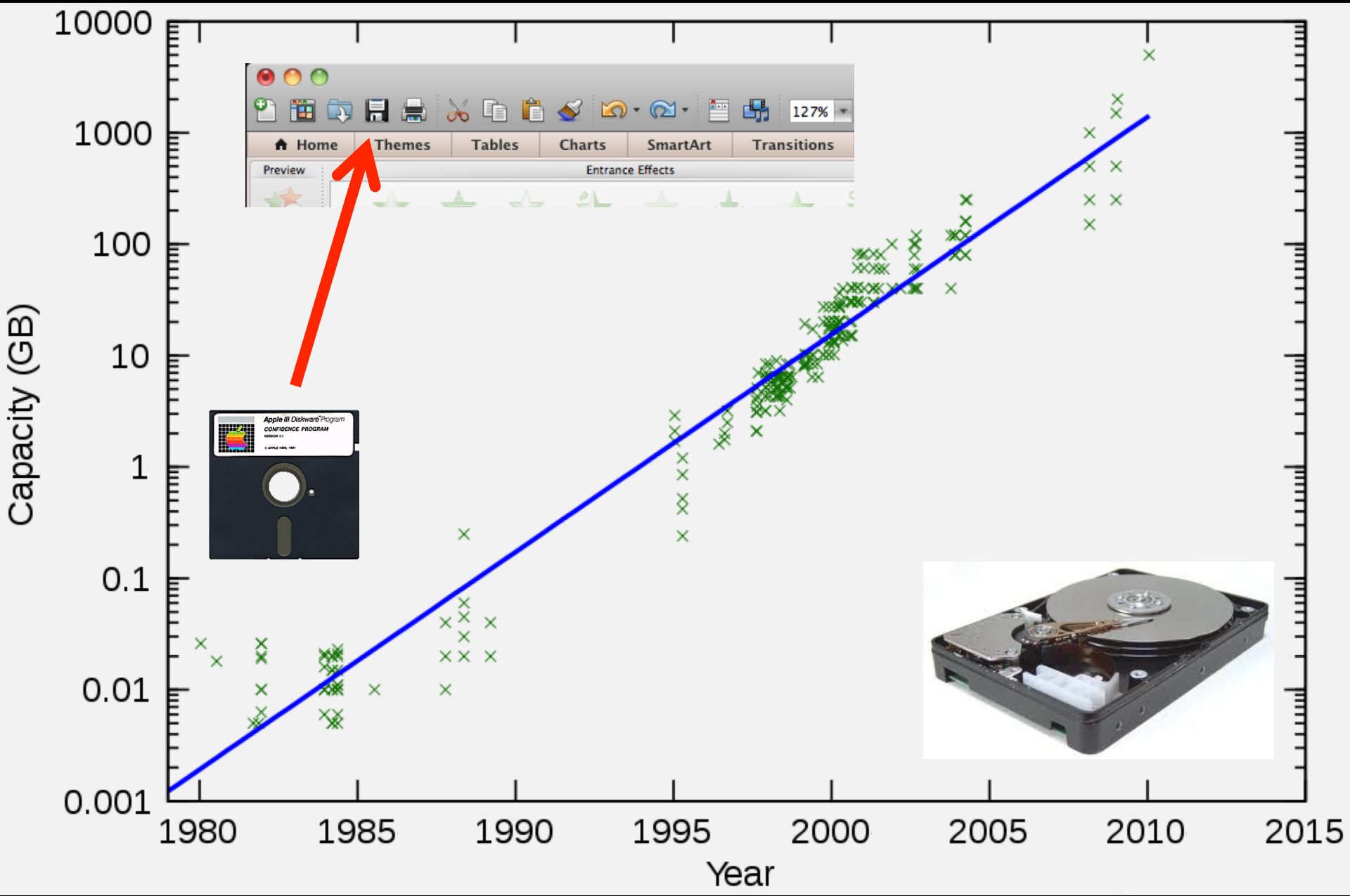



AT&T 3G
9:42 AM
SMS
Text
Tuesday
9
Calendar
Photos
Camera
YouTube
Stocks
280
Maps
73°
Weather
Clock
Calculator
Notes
Settings
iTunes
App Store
Phone
Mail
Safari
iPod

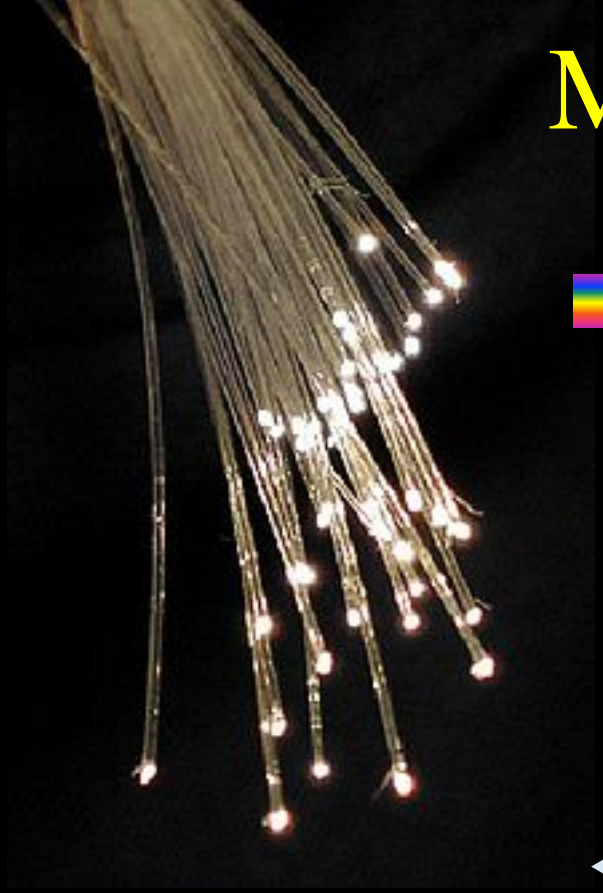
GPU cards are distruptive!

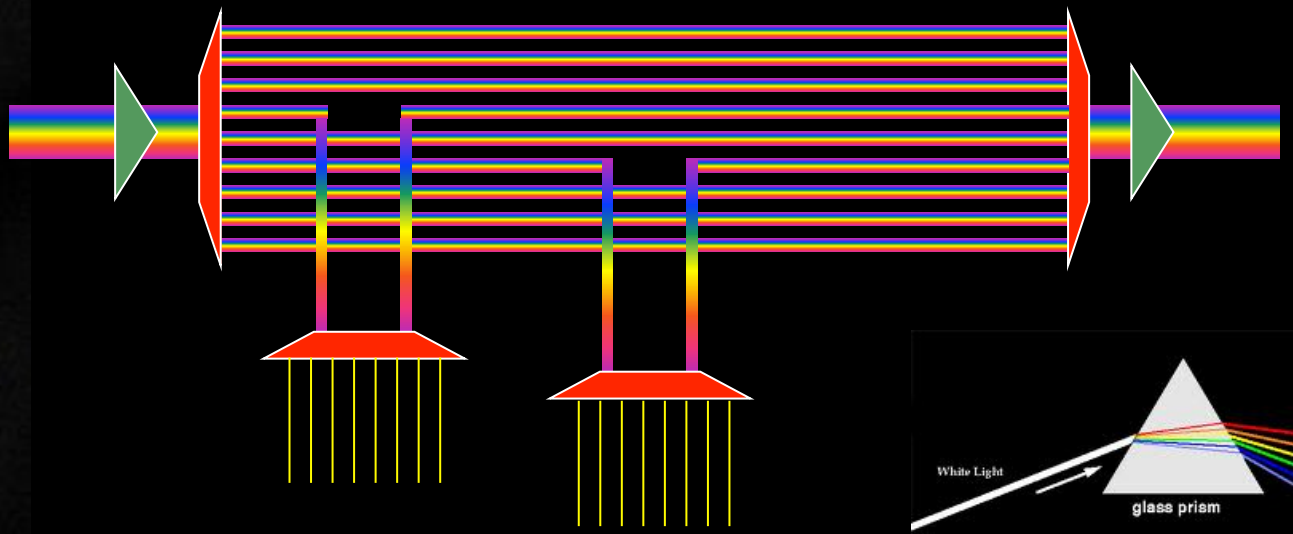


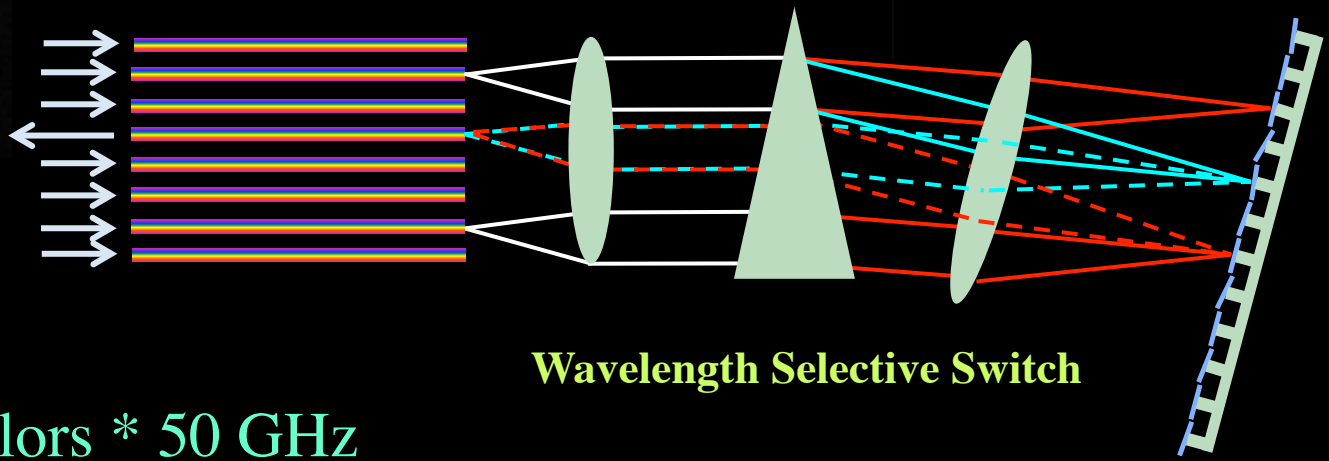
Data storage: doubling every 1.5 year!



Multiple colors / Fiber







Wavelength Selective Switch

Per fiber: ~ 80-100 colors * 50 GHz

Per color: 10 – 40 – 100 Gbit/s

BW * Distance ~ $2*10^{17}$ bm/s

New: Hollow Fiber!

➔ less RTT!

Wireless Networks

Digital technology reviews

Tech XO provied latest Digital Technology reviews like digital camara,digital lens reviews,digital

HOME | CONTACT US | PRIVACY POLICY

You Are Here : Digital Technology Reviews » Network Devices » Next Generation Throughput With

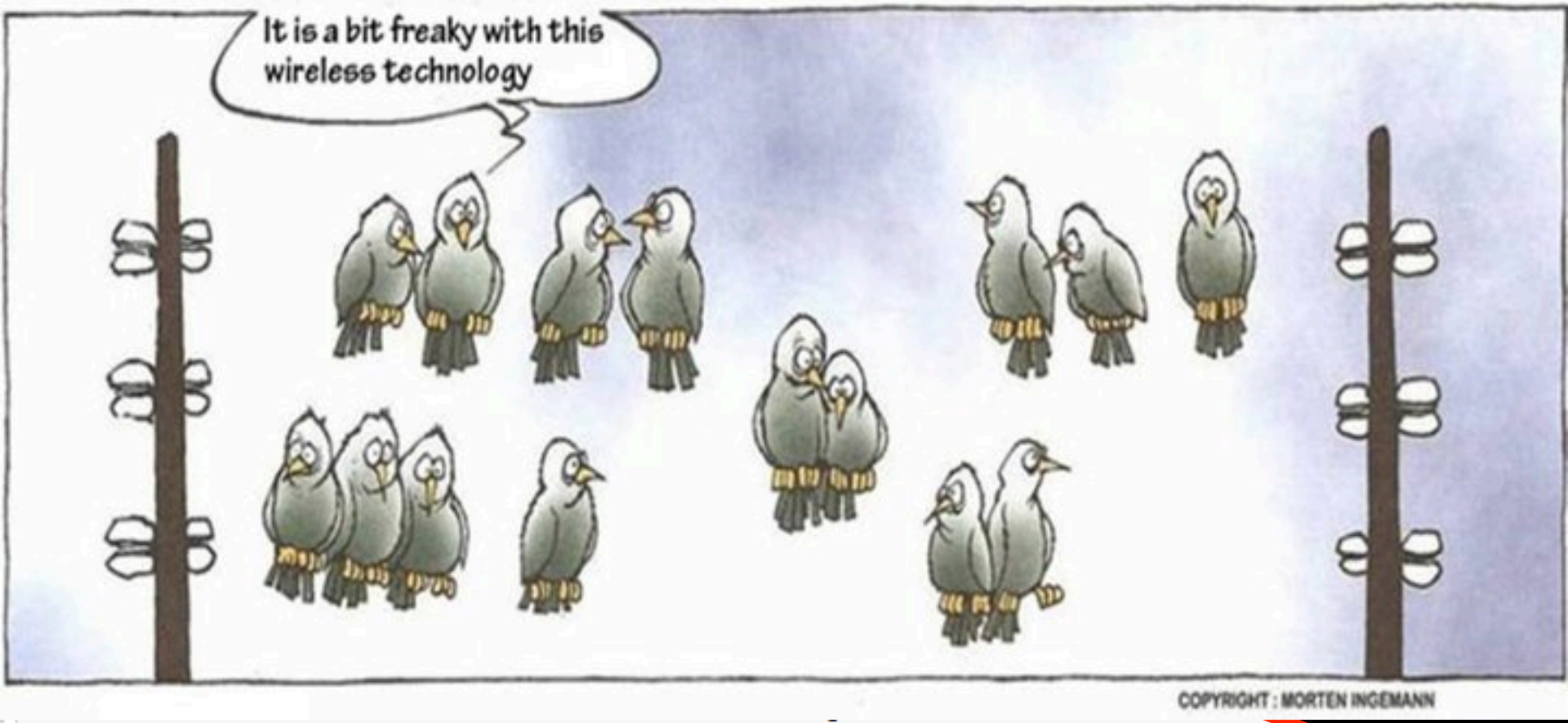
SEP 06

Next Generation Wireless LAN Technology 802.11ac 1 Gbps throughput with

Published By admin under Network Devices Tags: 1gbps throughput, 1gbps wireless, 1gbps wireless lans, generation, new generation, technologies, technology, throughput, wireless, wireless lan

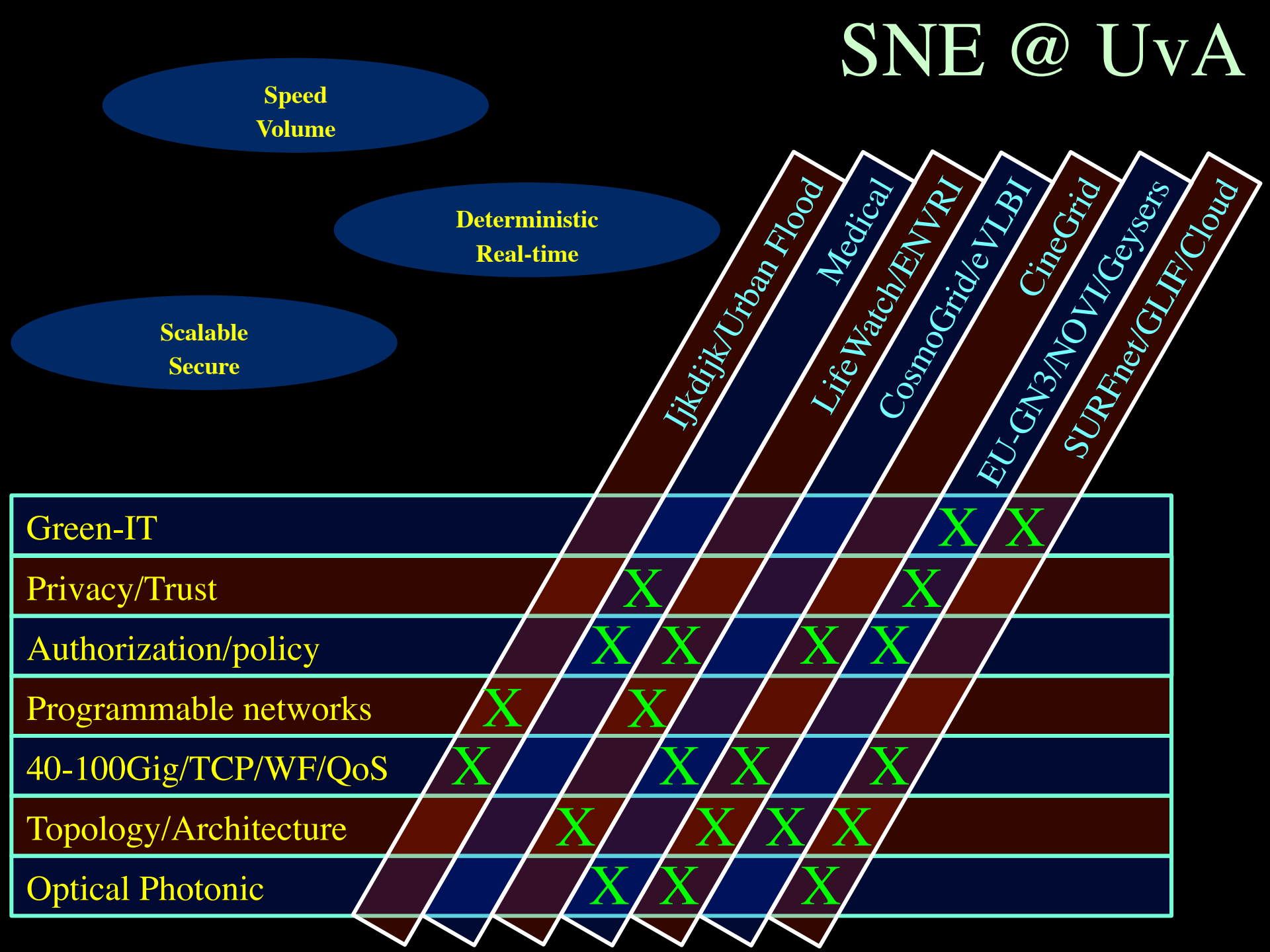
WiFi is one of the most preferred communication protocol LAN due to the easy comparison and convenience in the digital home. While consumer PC products has just started to migrate to a much higher bandwidth of 802.11n wireless LAN now working on next-generation standard definition is already in progress.

Wireless Networks

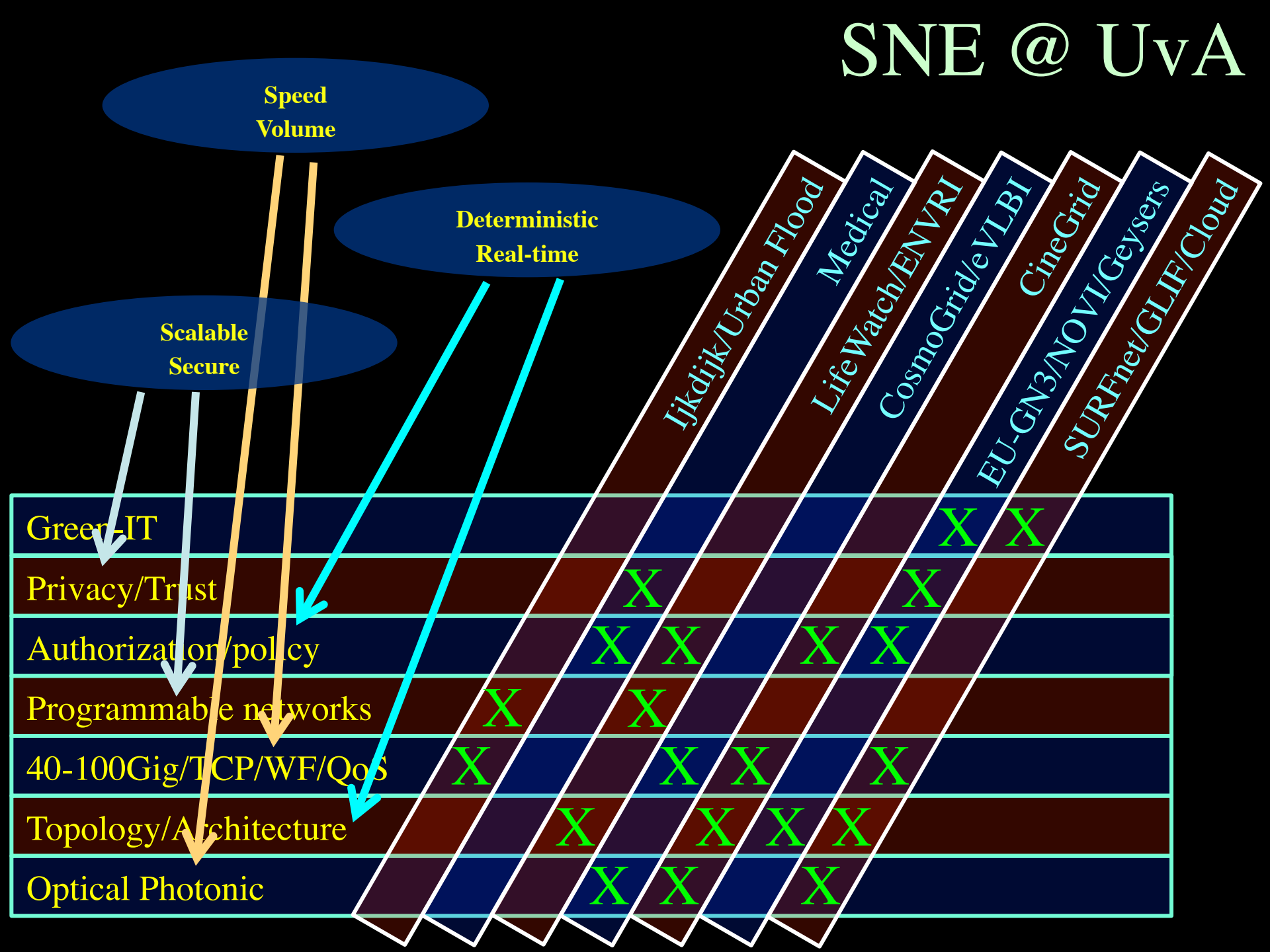


protocol LAN due to the easy comparison and convenience in the digital home. While consumer PC products has just started to migrate to a much higher bandwidth of 802.11n wireless LAN now working on next-generation standard definition is already in progress.

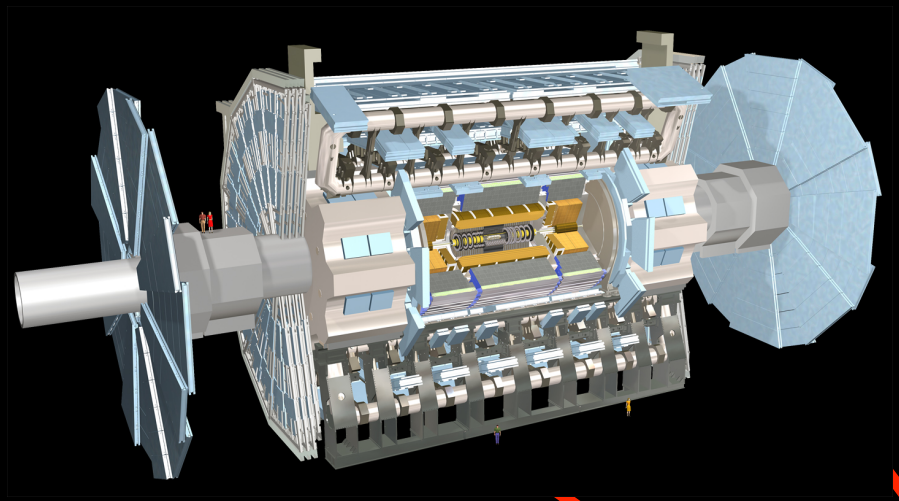
SNE @ UvA


Speed
Volume
Deterministic
Real-time
Scalable
Secure
Ijkdijk/Urban Flood
Medical
LifeWatch/ENVRI
CosmoGrid/eVLBI
CineGrid
EU-GN3/NOVI/Geysers
SURFnet/GLIF/Cloud
Green-IT
Privacy/Trust
Authorization/policy
Programmable networks
40-100Gig/TCP/WF/QoS
Topology/Architecture
Optical Photonic

SNE @ UvA

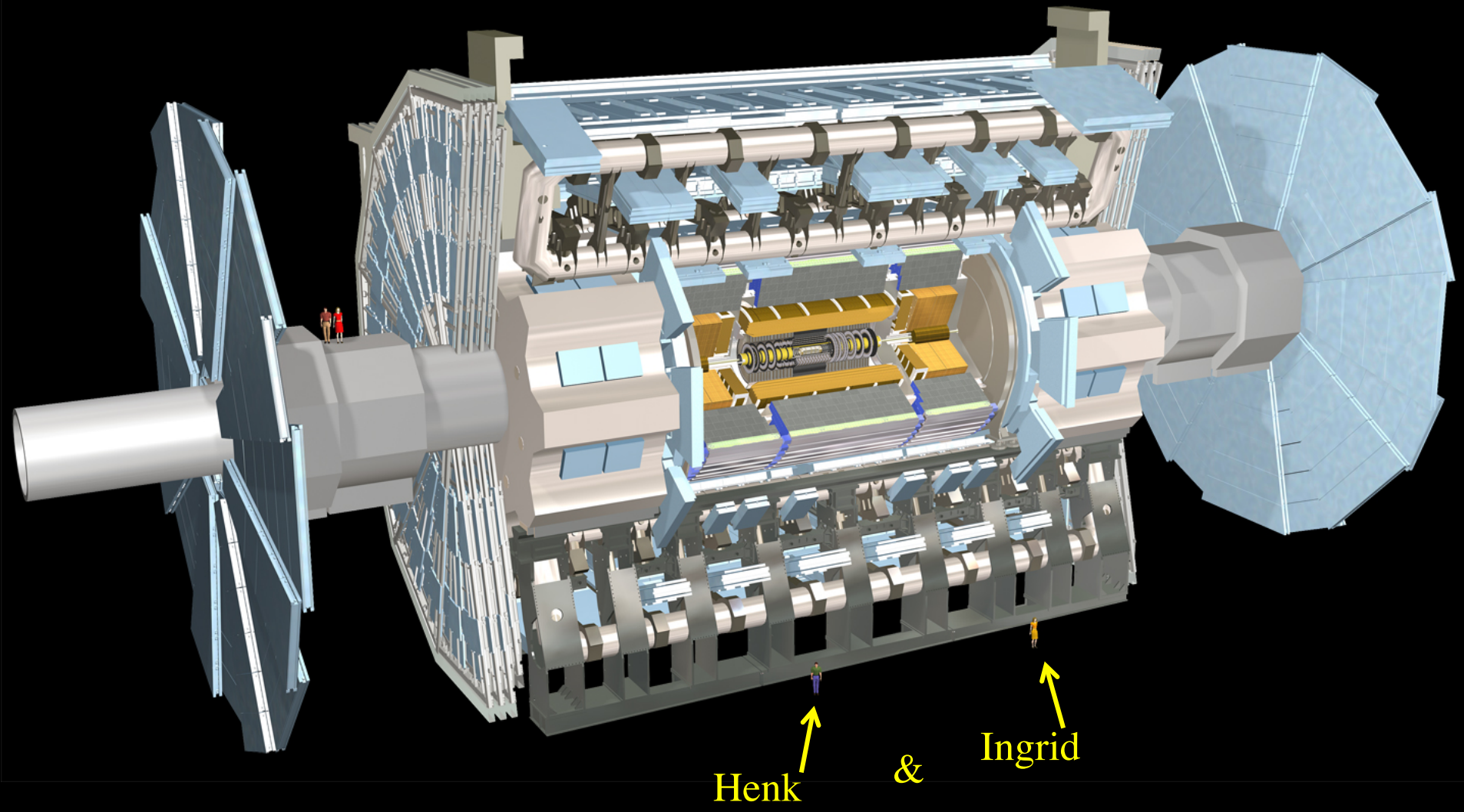

Speed
Volume
Deterministic
Real-time
Scalable
Secure
Ijkdijk/Urban Flood
Medical
LifeWatch/ENVRI
CosmoGrid/eVLBI
CineGrid
EU-GN3/NOVI/Geysers
SURFnet/GLIF/Cloud
Green-IT
Privacy/Trust
Authorization/policy
Programmable networks
40-100Gig/TCP/WF/QoS
Topology/Architecture
Optical Photonic

SNE @ UvA



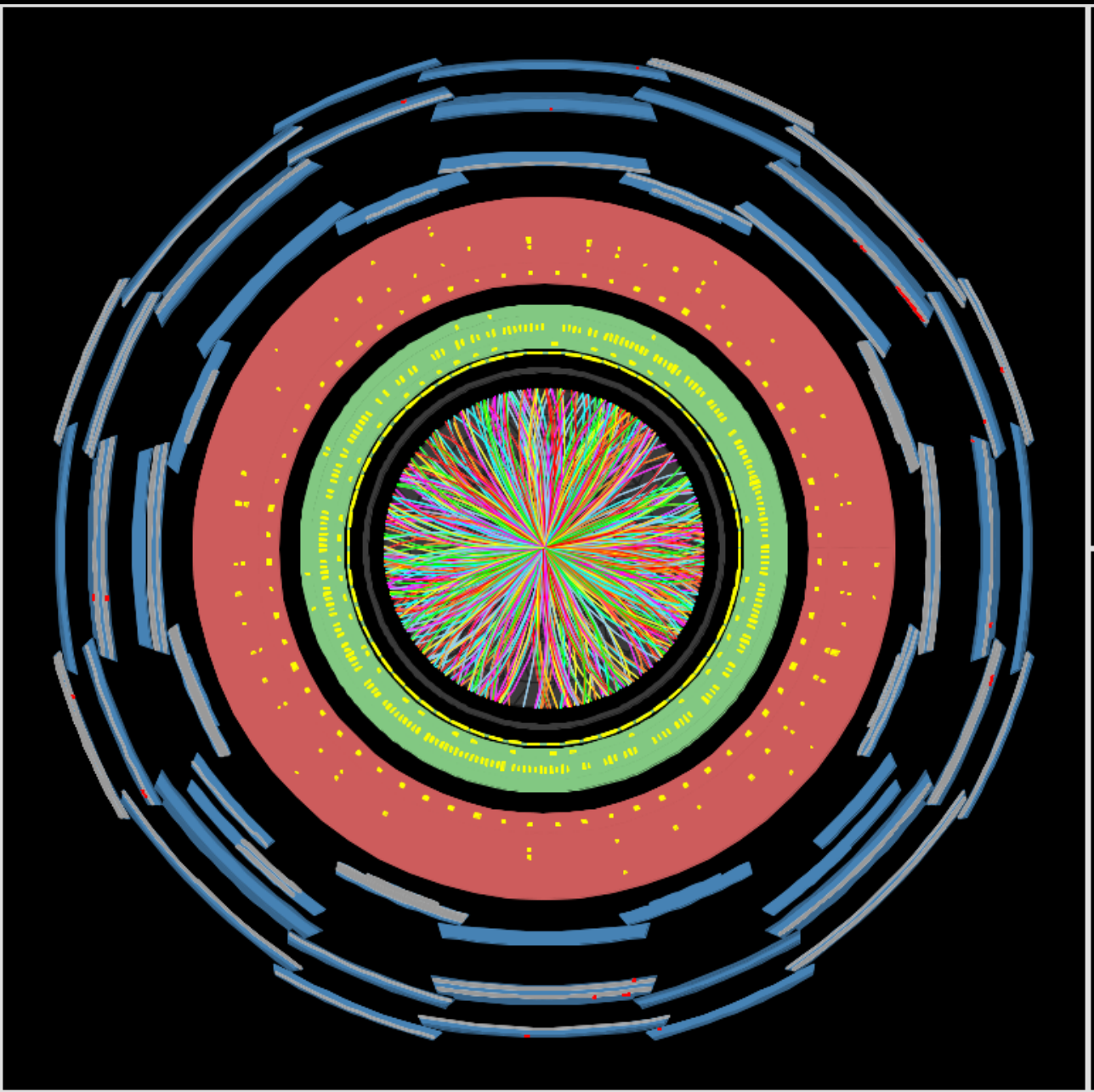
	Ijkdijk/Urban Flood	Medical	LifeWatch/ENVRI	CosmoGrid/eVLBI	CineGrid	EU-GN3/NOVI/Geysers	SURFnet/GLIF/Cloud
Green-IT						X	X
Privacy/Trust		X			X		
Authorization/policy		X	X		X	X	
Programmable networks	X		X				
40-100Gig/TCP/WF/QoS	X		X	X		X	
Topology/Architecture		X		X	X	X	
Optical Photonic		X	X		X		

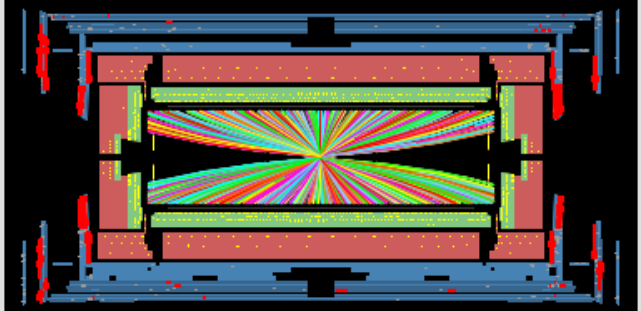
ATLAS detector @ CERN Geneve



ATLAS detector @ CERN Geneve

One Heavy Ion Collision in Atlas!





ATLAS EXPERIMENT

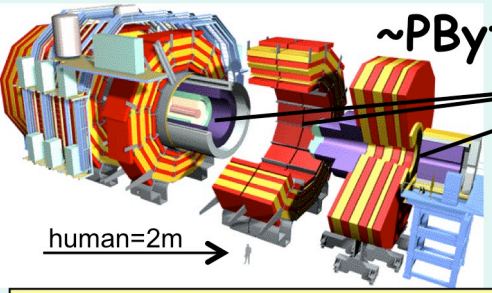
Run Number: 170482, Event Number: 3936308

Date: 2010-12-06 17:21:31 CET

Snapshot of a heavy ion collision directly from the ATLAS experiment

LHC Data Grid Hierarchy

CMS as example, Atlas is similar



~PByte/sec

Online System

~100 MBytes/sec

100000 flops/byte

10 Pflops/s

event simulation

Tier 0 +1

HPSS

event reconstruction

CMS detector: 15m X 15m X 22m

12,500 tons, $700M.

~2.5 Gbits/sec

Status 2002!

Tier 1

Italian Regional Center

German Regional Center

NIKHEF Dutch Regional Center

FermiLab, USA Regional Center

...

~0.6-2.5 Gbps

analysis

Tier2 Center

Tier 2

~0.6-2.5 Gbps

Tier 3

Institute ~0.25TIPS

Physics data cache

100 - 1000 Mbits/sec

Tier 4

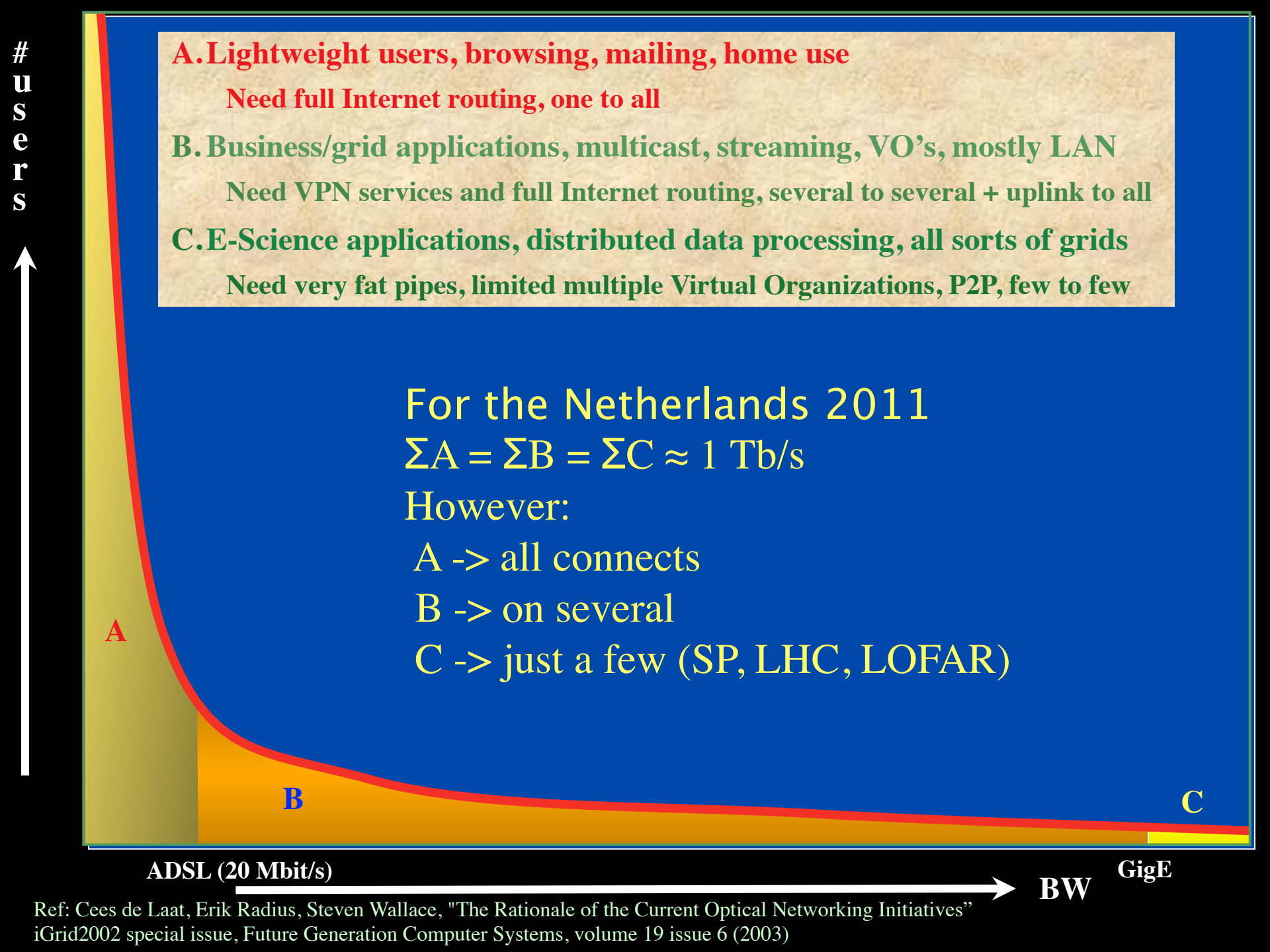
Workstations

Courtesy Harvey Newman, CalTech and CERN

CERN/CMS data goes to 6-8 Tier 1 regional centers, and from each of these to 6-10 Tier 2 centers.

Physicists work on analysis "channels" at 135 institutes. Each institute has ~10 physicists working on one or more channels.

2000 physicists in 31 countries are involved in this 20-year experiment in which DOE is a major player.

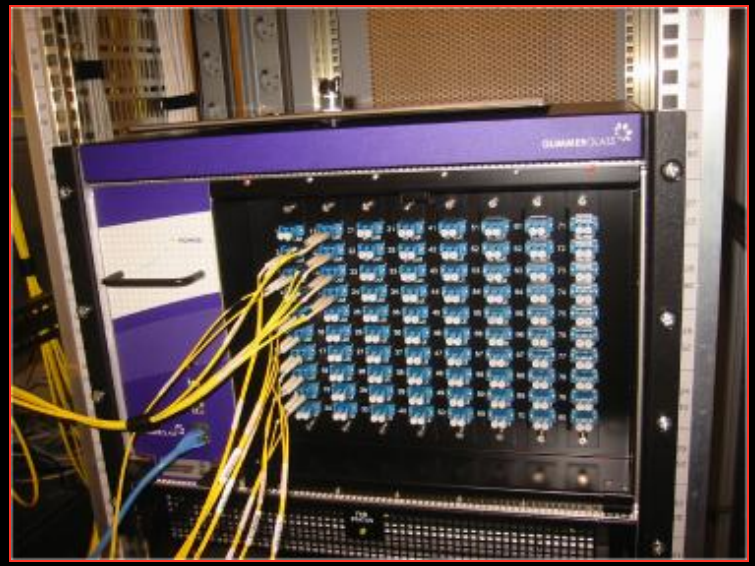


Ref: Cees de Laat, Erik Radius, Steven Wallace, "The Rationale of the Current Optical Networking Initiatives" iGrid2002 special issue, Future Generation Computer Systems, volume 19 issue 6 (2003)

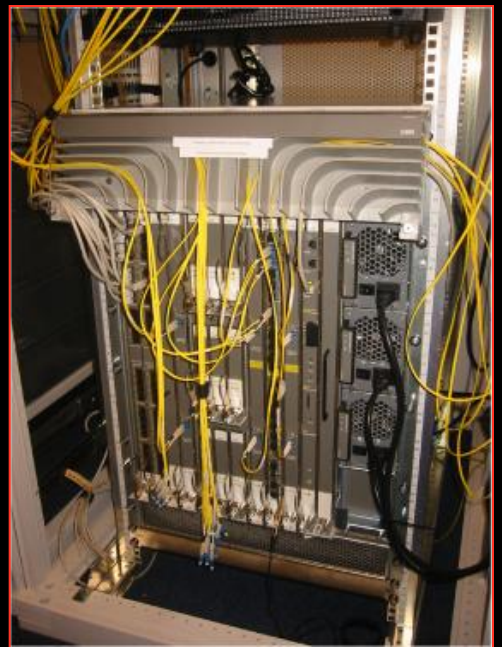
Towards Hybrid Networking!

- 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

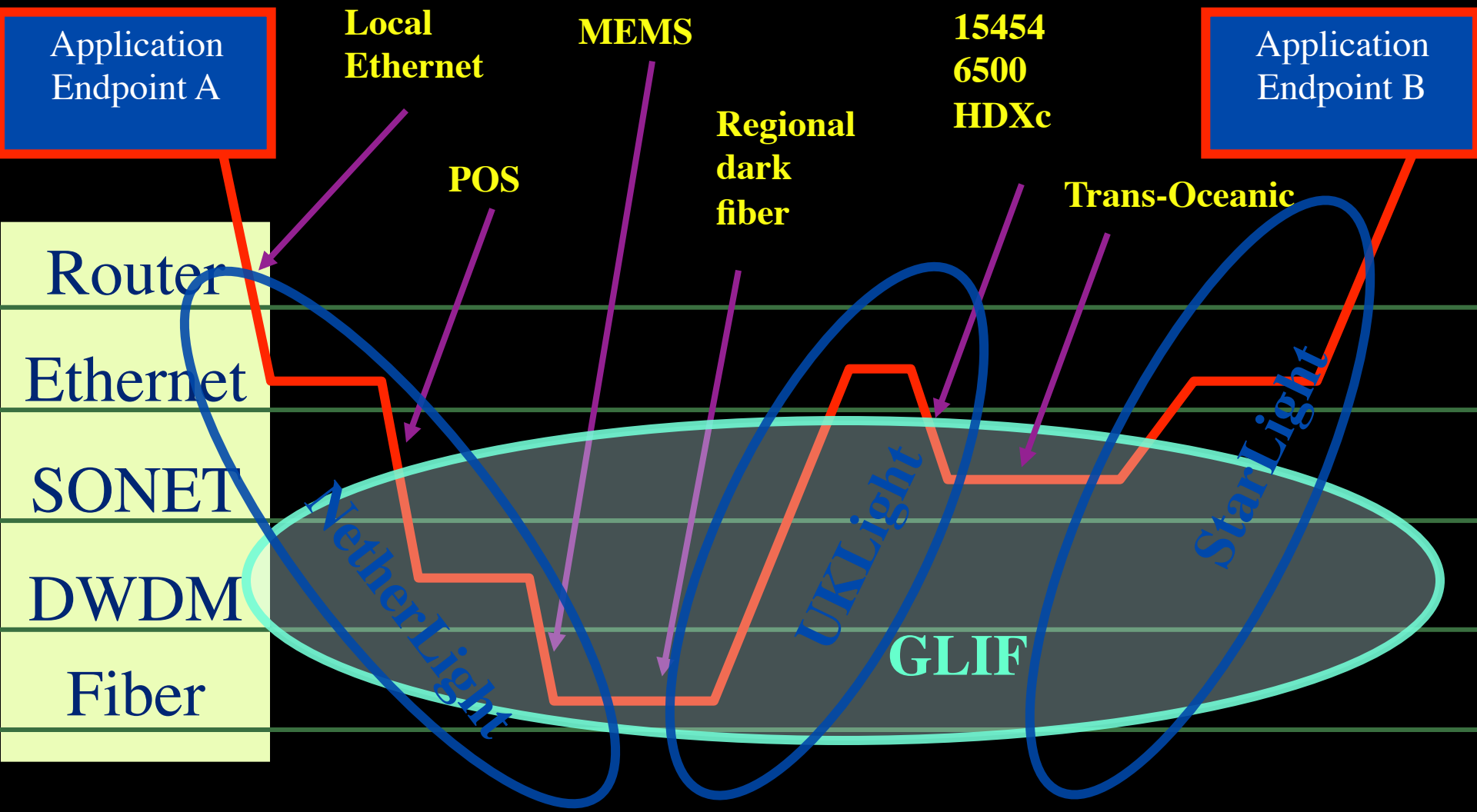


L2 ≈ 5-8 k$/port

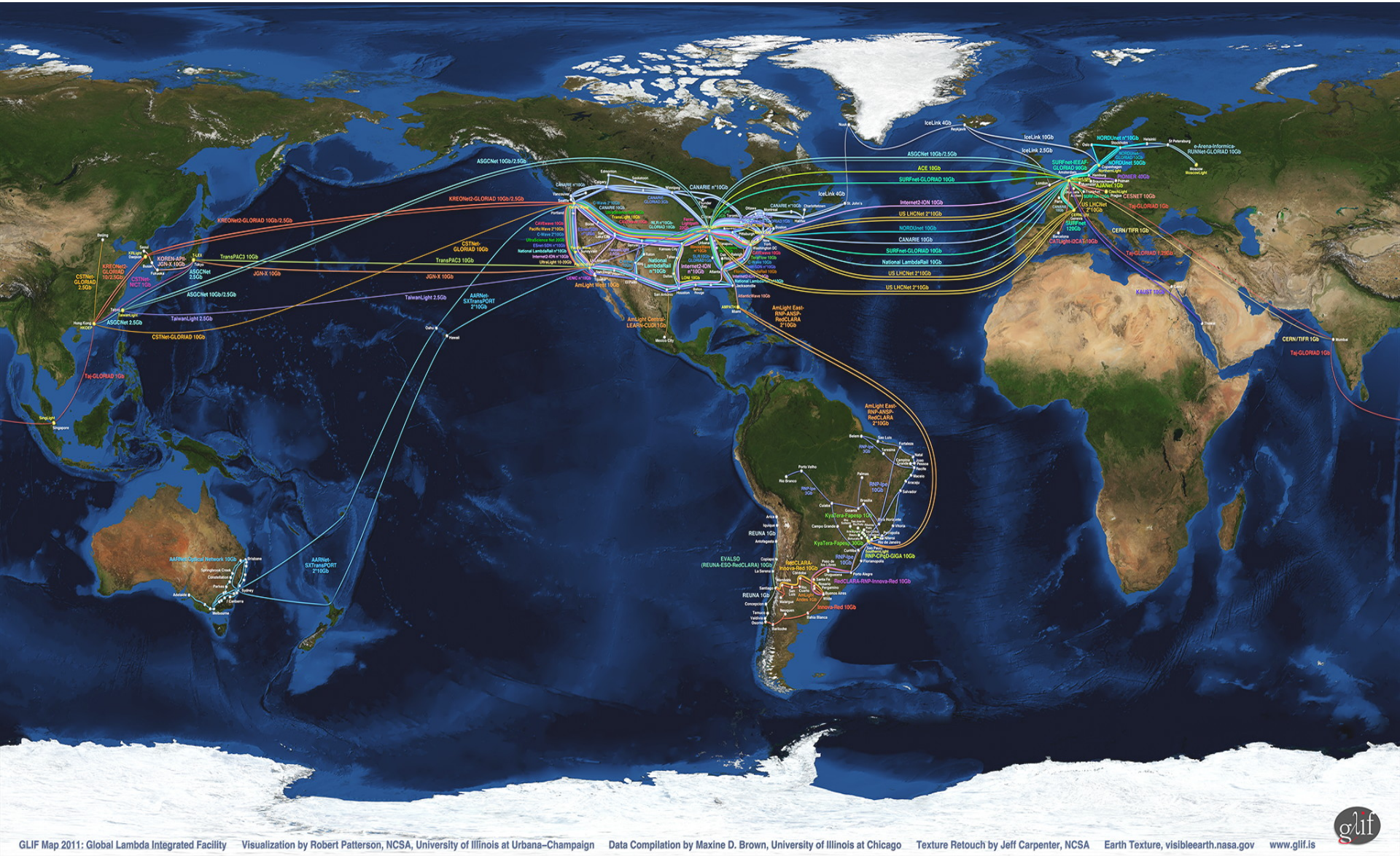


L3 ≈ 75+ k$/port

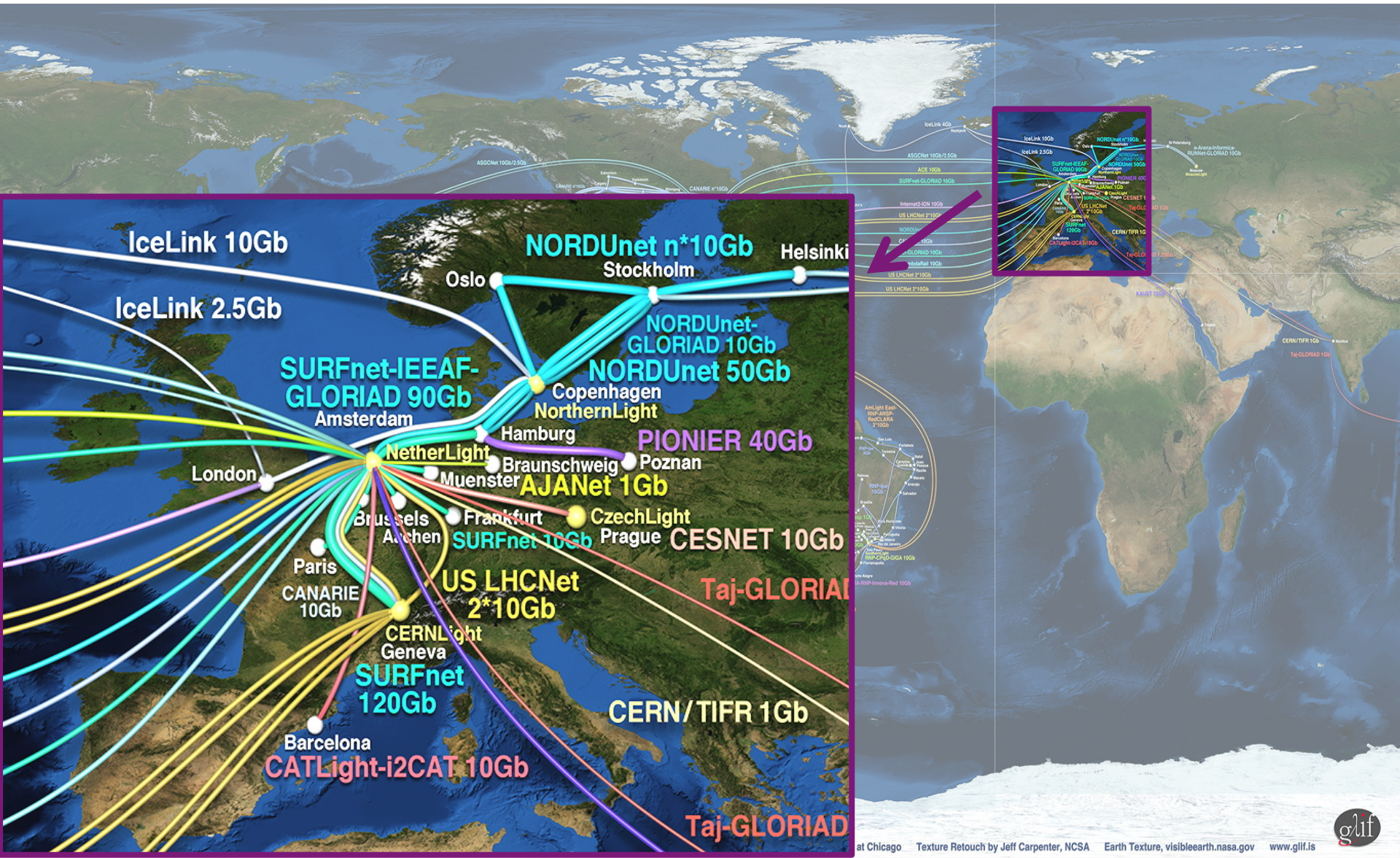
How low can you go?

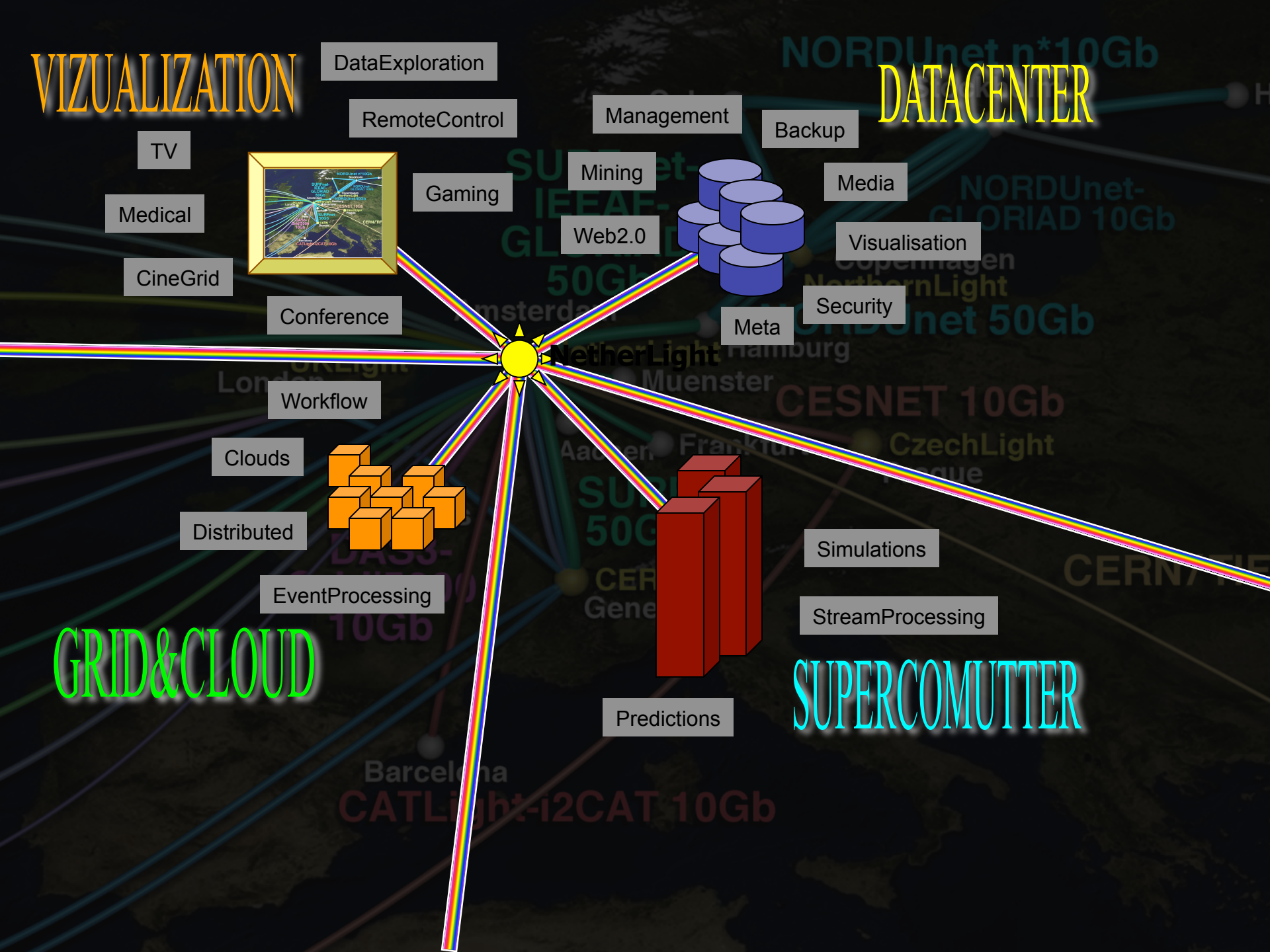


The GLIF – lightpaths around the world



The GLIF – lightpaths around the world




VIZUALIZATION
DataExploration
RemoteControl
TV
Medical
CineGrid
Gaming
Conference
DATACENTER
Management
Backup
Mining
Media
Web2.0
Visualisation
Security
Meta
NetherLight
Workflow
Clouds
Distributed
EventProcessing
GRID&CLOUD
Simulations
StreamProcessing
Predictions
SUPERCOMUTTER

In The Netherlands SURFnet connects between 180:

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

with an indirect ~750K user base

~ 8860 km

scale comparable to railway system

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.

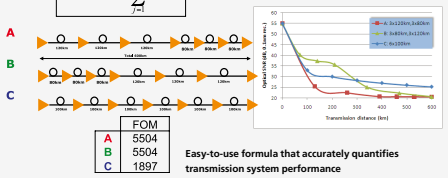
In this demonstration we will investigate the performance of a 40Gb/s PM-QPSK alien wavelength installed on a 10Gb/s DWDM infrastructure.

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.

$$FOM = \sum_{j=1}^{N} 10^{\left(\frac{L_j}{10}\right)}$$

Lj, span losses in dB
N, number of spans

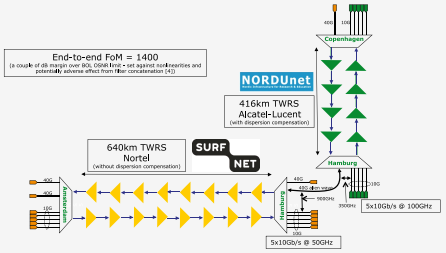


	FOM
A	5504
B	5504
C	1897

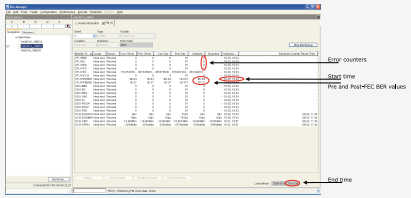
Easy-to-use formula that accurately quantifies transmission system performance

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

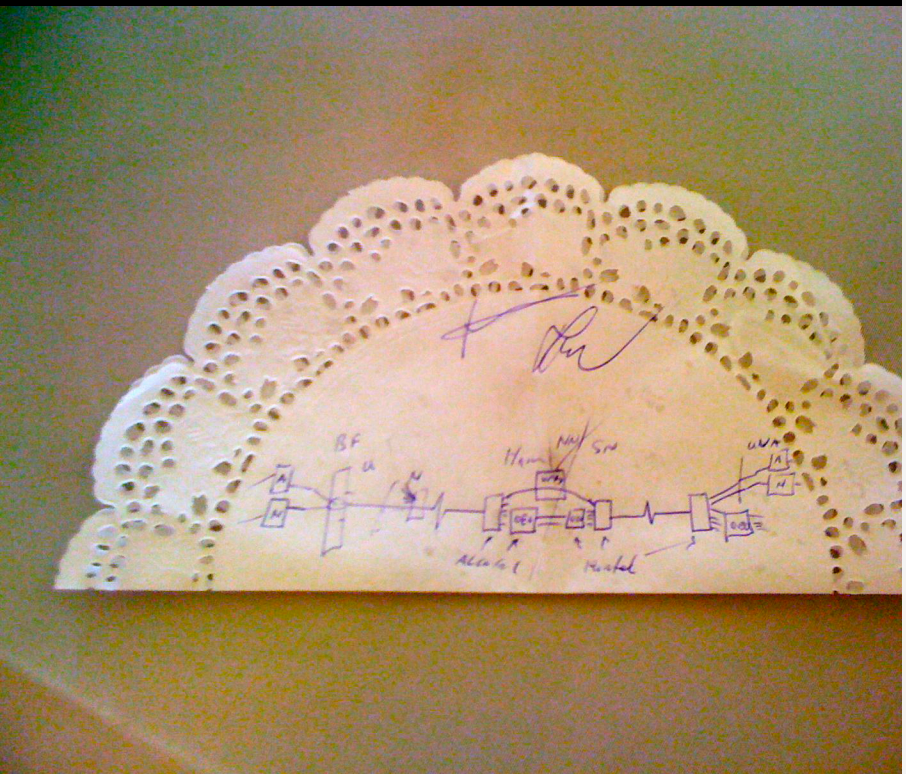
- We have investigated experimentally the all-optical transmission of a 40Gb/s PM-QPSK alien wavelength via a concatenated native and third party DWDM system that both were carrying live 10Gb/s wavelengths.
- The end-to-end transmission system consisted of 1056 km of TWRS (TrueWave Reduced Slope) transmission 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 telindus together with ISTY SURF NET

REFERENCES [1] "OPERATIONAL SOLUTIONS FOR AN OPEN DWDM LAYER", O. GERSTEL ET AL, OFC'2009 | [2] "AT&T OPTICAL TRANSPORT SERVICES", BARBARA E. SMITH, OFC'09 [3] "OPEX SAVINGS OF ALL-OPTICAL CORE NETWORKS", ANDREW LORD AND CARL ENGINEER, ECOC2009 | [4] NORTEL/SURFNET INTERNAL COMMUNICATION

ACKNOWLEDGEMENTS WE ARE GRATEFUL TO NORDUNET FOR PROVIDING US WITH BANDWIDTH ON THEIR DWDM LINK FOR THIS EXPERIMENT AND ALSO FOR THEIR SUPPORT AND ASSISTANCE DURING THE EXPERIMENTS. WE ALSO ACKNOWLEDGE TELINDUS AND NORTEL FOR THEIR INTEGRATION WORK AND SIMULATION SUPPORT

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.

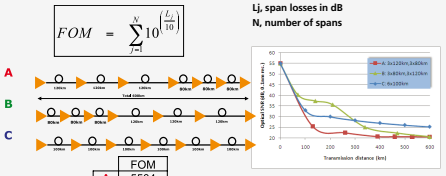
In this demonstration we will investigate the performance of a 40Gb/s PM-QPSK alien wavelength installed on a 10Gb/s DWDM infrastructure.

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.

$$FOM = \sum_{j=1}^{N} 10^{\left(\frac{L_j}{10}\right)}$$

Lj, span losses in dB
N, number of spans

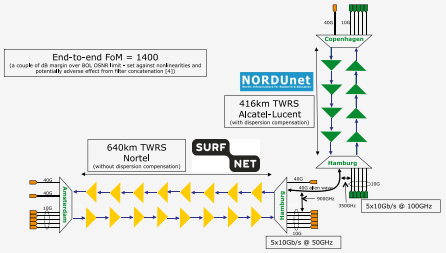


	FOM
A	5504
B	5504
C	1897

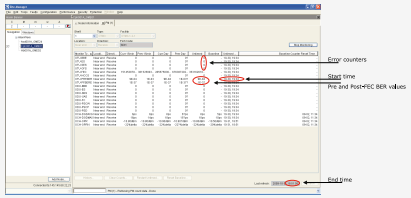
Easy-to-use formula that accurately quantifies transmission system performance

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 telindus together with ISTY SURF NET

REFERENCES [1] "OPERATIONAL SOLUTIONS FOR AN OPEN DWDM LAYER", O. GERSTEL ET AL, OFC'2009 | [2] "AT&T OPTICAL TRANSPORT SERVICES", BARBARA E. SMITH, OFC'09
[3] "OPEX SAVINGS OF ALL-OPTICAL CORE NETWORKS", ANDREW LORD AND CARL ENGINEER, ECOC2009 | [4] NORTEL/SURFNET INTERNAL COMMUNICATION

ACKNOWLEDGEMENTS WE ARE GRATEFUL TO NORDUNET FOR PROVIDING US WITH BANDWIDTH ON THEIR DWDM LINK FOR THIS EXPERIMENT AND ALSO FOR THEIR SUPPORT AND ASSISTANCE DURING THE EXPERIMENTS. WE ALSO ACKNOWLEDGE TELINDUS AND NORTEL FOR THEIR INTEGRATION WORK AND SIMULATION SUPPORT

ClearStream @ TNC2011

Setup codename:
FlightCees

UvA

iPerf
I7 3.2 GHz Q-core

iPerf
Amd Ph II 3.6 GHz HexC

Mellanox

40G E

Copenhagen

iPerf
2* dual 2.8 GHz Q-core

iPerf

Mellanox

CIENA
OME
6500

CERN

Hamburg

CIENA DWDM

Alcatel DWDM

CIENA
OME
6500

CIENA
OME
6500

17 ms RTT

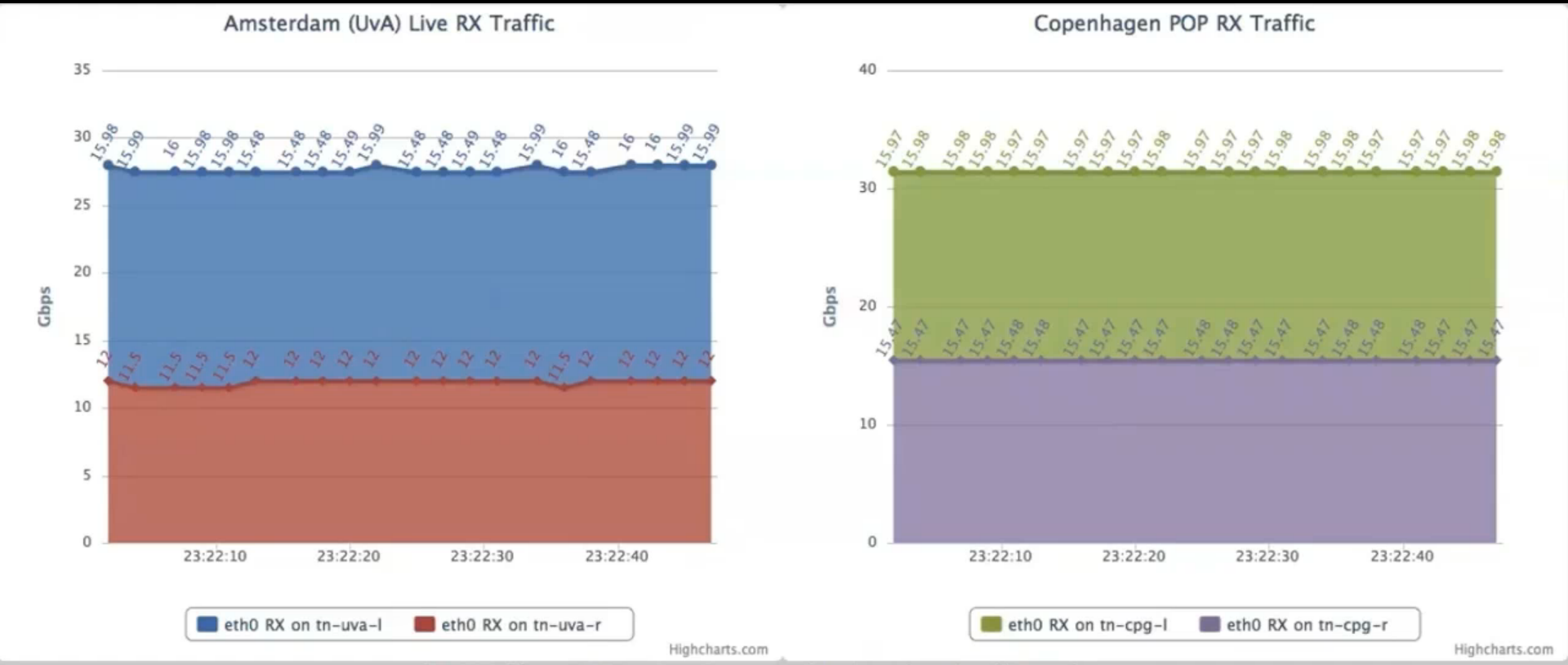
27 ms RTT

LH

LH

Amsterdam – Geneva (CERN) – Copenhagen – 4400 km (2700 km alien light)

http://tnc11.delaat.net



27.99 Gbps to Amsterdam <-> 31.45 Gbps to Copenhagen

Total Throughput 59.44 Gbps RTT 44.010 ms

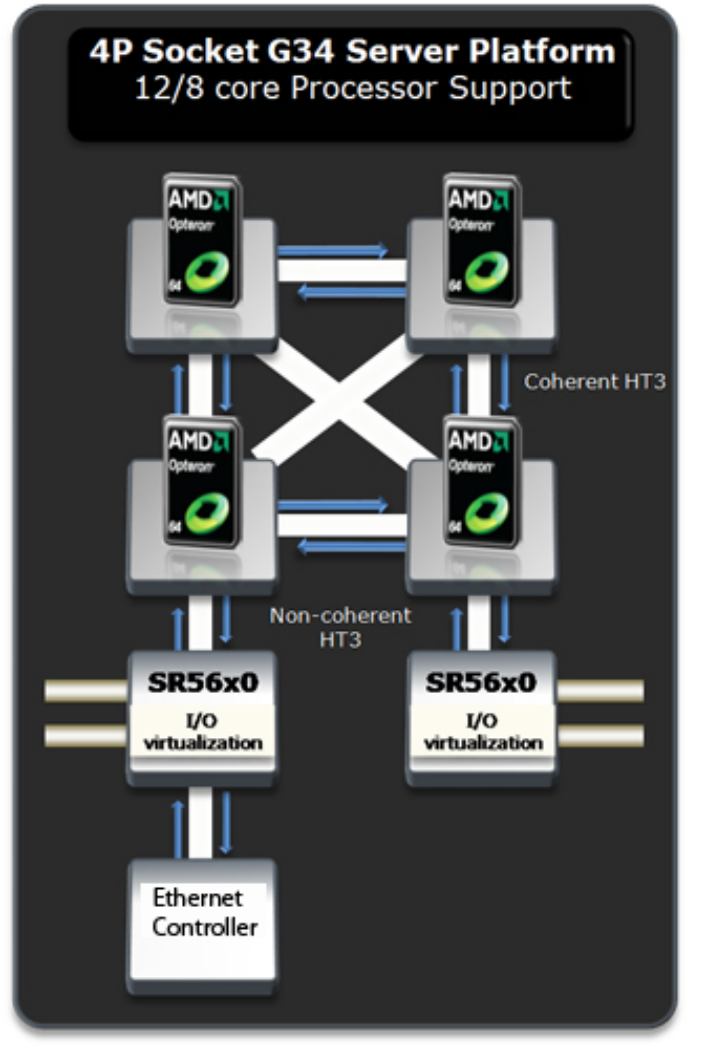
Results Ams-CERN oct '10 (rtt = 17 ms)

- Single flow iPerf 1 core -> 21 Gbps
- Single flow iPerf 1 core <> -> 15+15 Gbps
- Multi flow iPerf 2 cores -> 25 Gbps
- Multi flow iPerf 2 cores <> -> 23+23 Gbps
- DiViNe <> -> 11 Gbps
- Multi flow iPerf + DiVine -> 35 Gbps
- Multi flow iPerf + DiVine <> -> 35 + 35 Gbps

Performance Explained

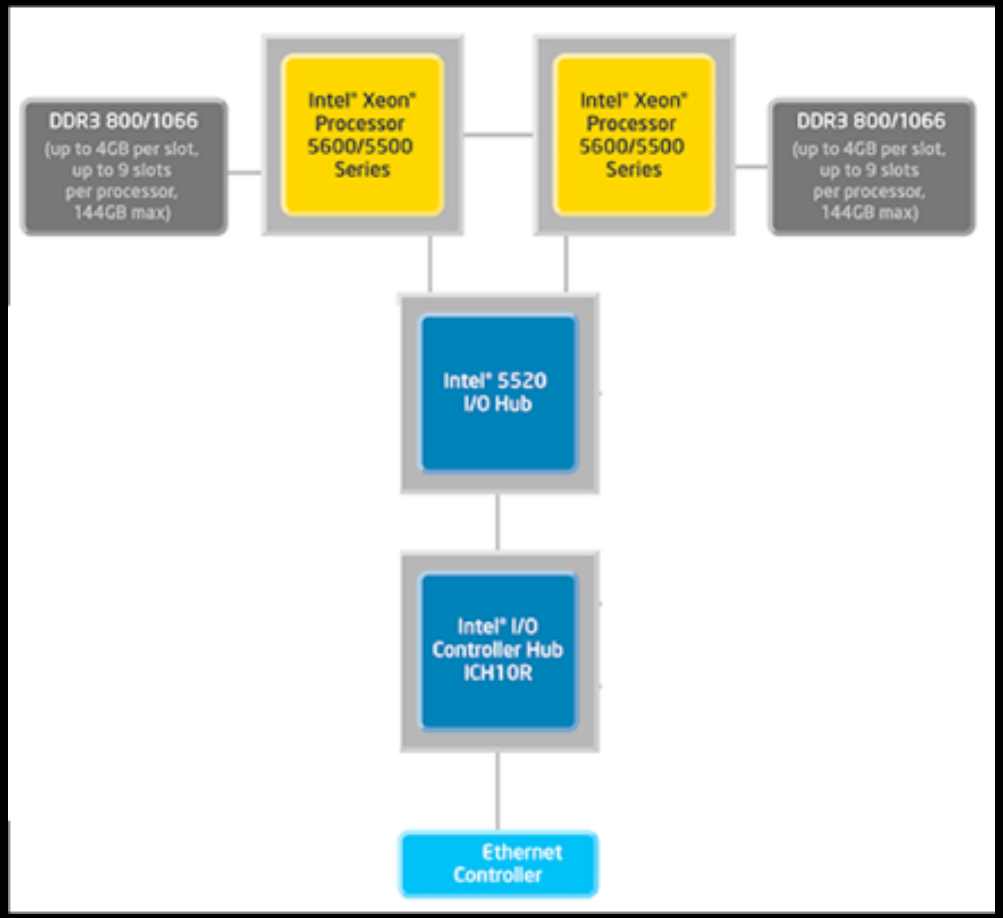
- Mellanox 40GE card is PCI-E 2.0 8x (5GT/s)
- 40Gbit/s raw throughput but ….
- PCI-E is a network-like protocol
  - 8/10 bit encoding -> 25% overhead -> 32Gbit/s maximum data throughput
  - Routing information
- Extra overhead from IP/Ethernet framing
- Server architecture matters!
  - 4P system performed worse in multithreaded iperf

Server Architecture



DELL R815

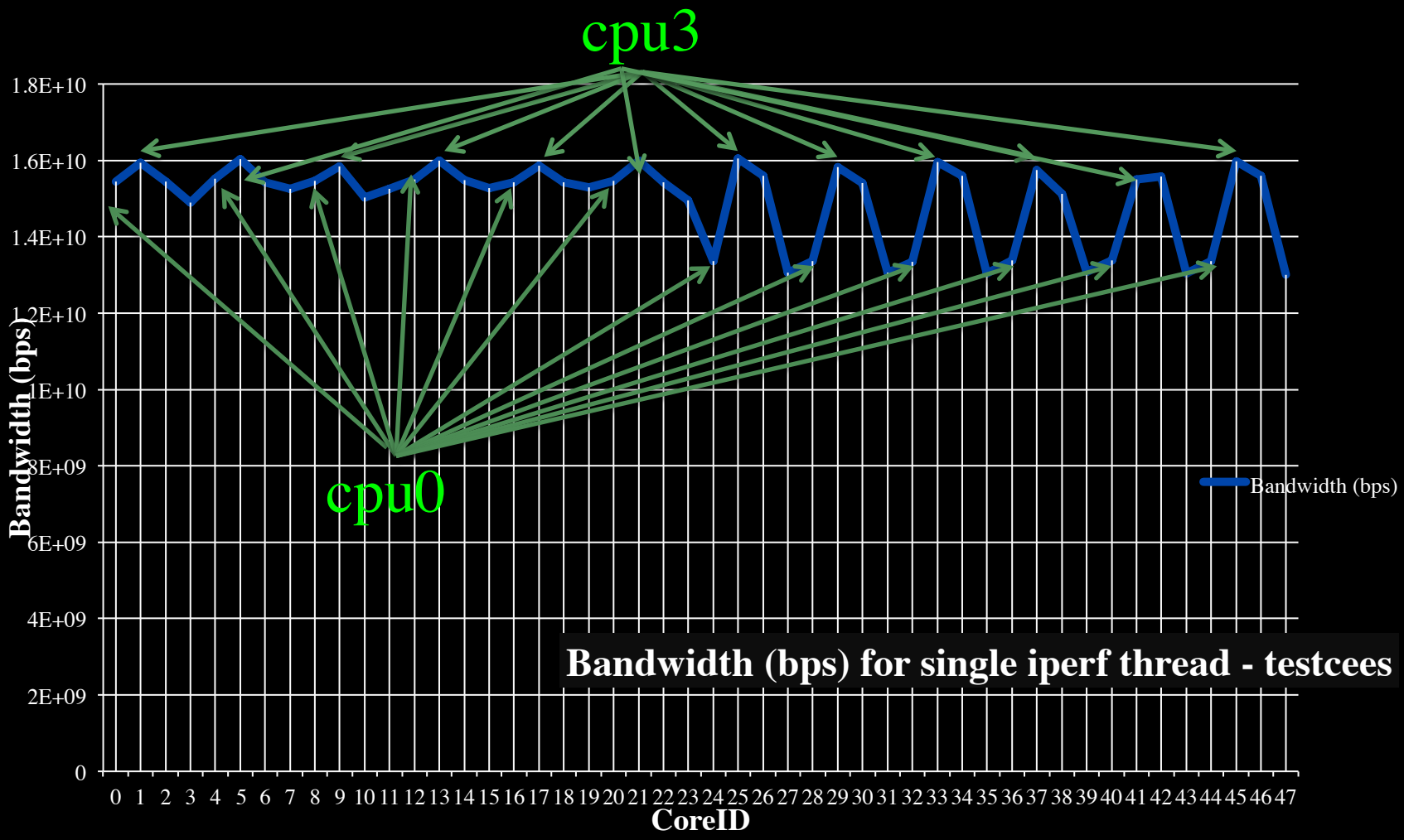
4 x AMD Opteron 6100



Supermicro X8DTT-HIBQF

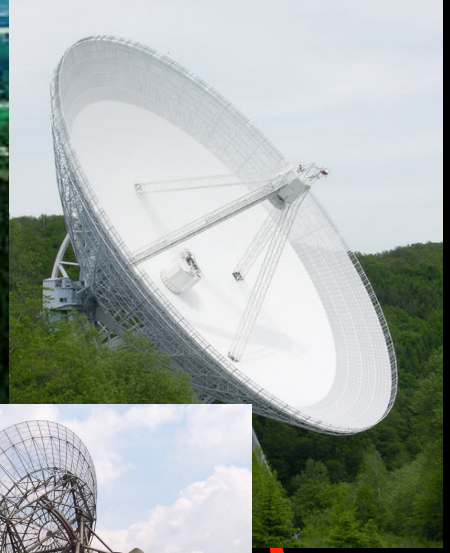
2 x Intel Xeon

CPU Topology benchmark



We used numactl to bind iperf to cores

SNE @ UvA



	Ijkdijk/Urban Flood	Medical	LifeWatch	CosmoGrid/eVLBI	CineGrid	EU-GN3/NOVI/Geysers	SURFnet/GLIF/Cloud
Green-IT						X	X
Privacy/Trust			X		X		
Authorization/policy		X	X		X	X	
Programmable networks	X		X				
40-100Gig/TCP/WF/QoS	X		X	X		X	
Topology/Architecture		X		X	X	X	
Optical Photonic		X	X		X		

LOFAR as a Sensor Network

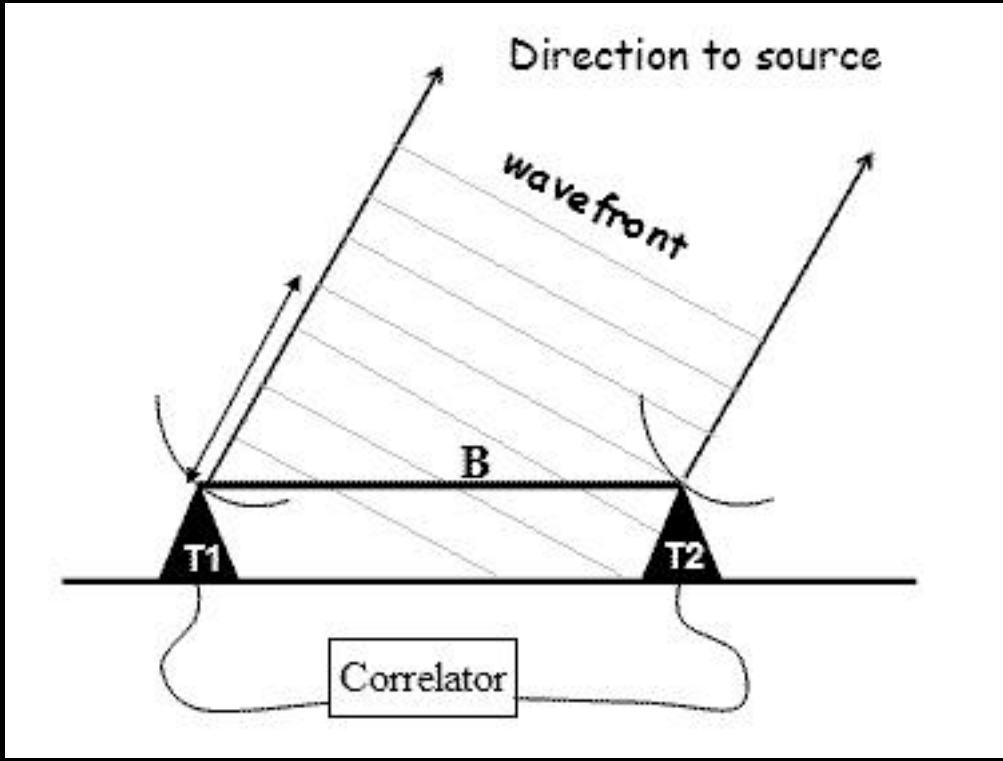
20 flops/byte

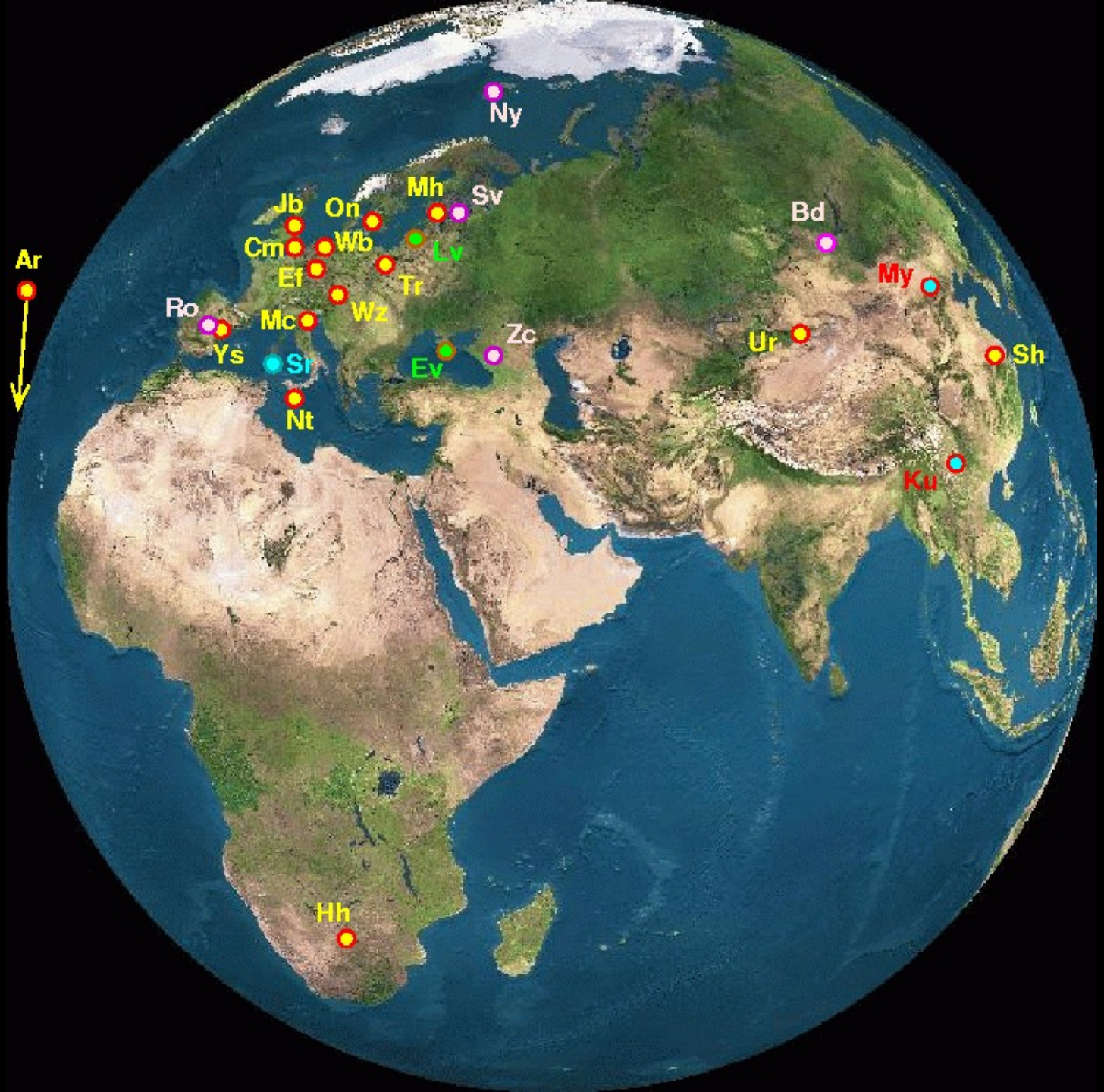
- LOFAR is a large distributed research infrastructure:

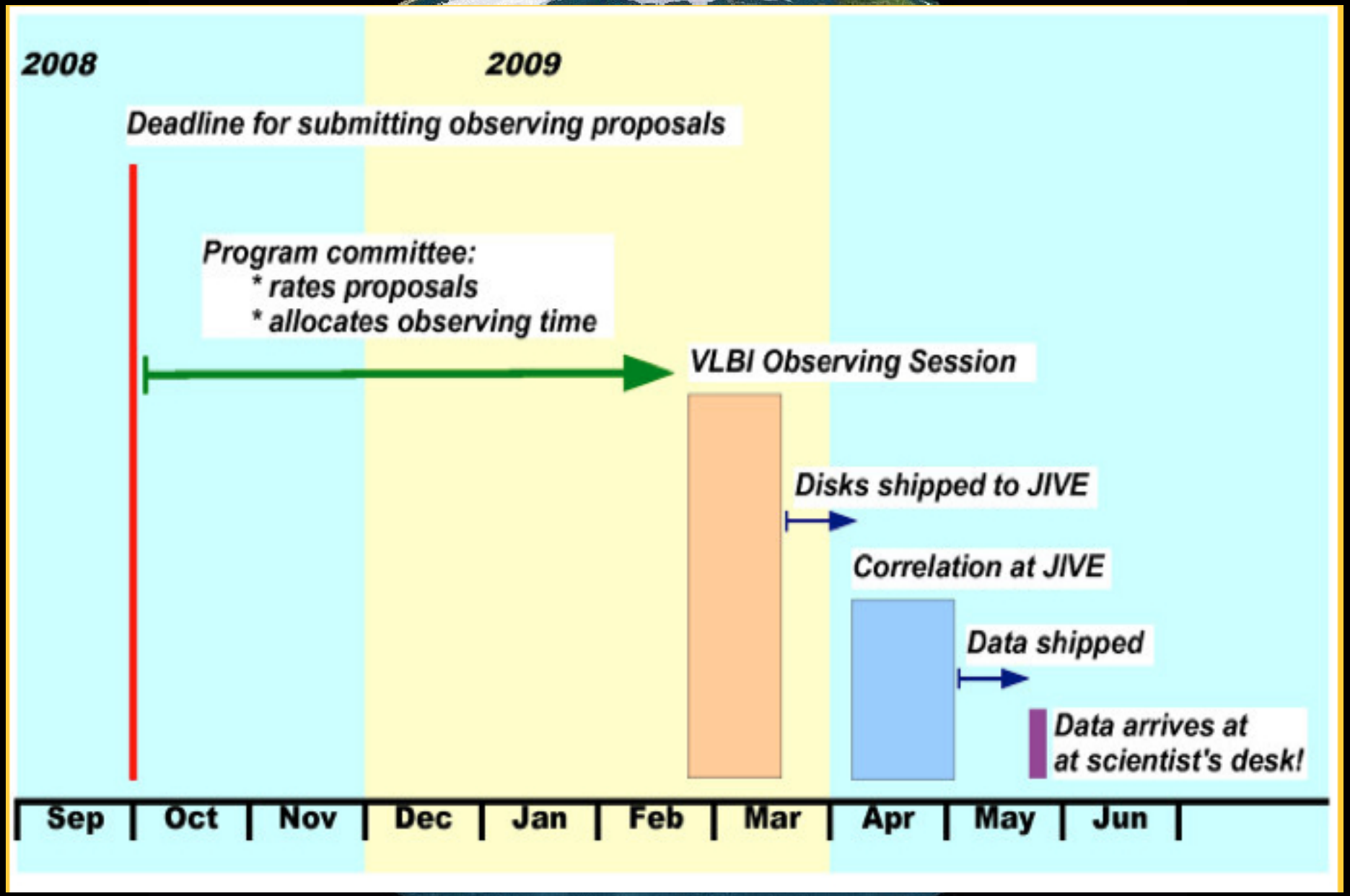
2 Tflops/s

  - Astronomy:
    - >100 phased array stations
    - Combined in aperture synthesis array
    - 13,000 small “LF” antennas
    - 13,000 small “HF” tiles
  - Geophysics:
    - 18 vibration sensors per station
    - Infrasound detector per station
  - >20 Tbit/s generated digitally
  - >40 Tflop/s supercomputer
  - innovative software systems
    - new calibration approaches
    - full distributed control
    - VO and Grid integration
    - datamining and visualisation

e -Very Large Base Interferometer



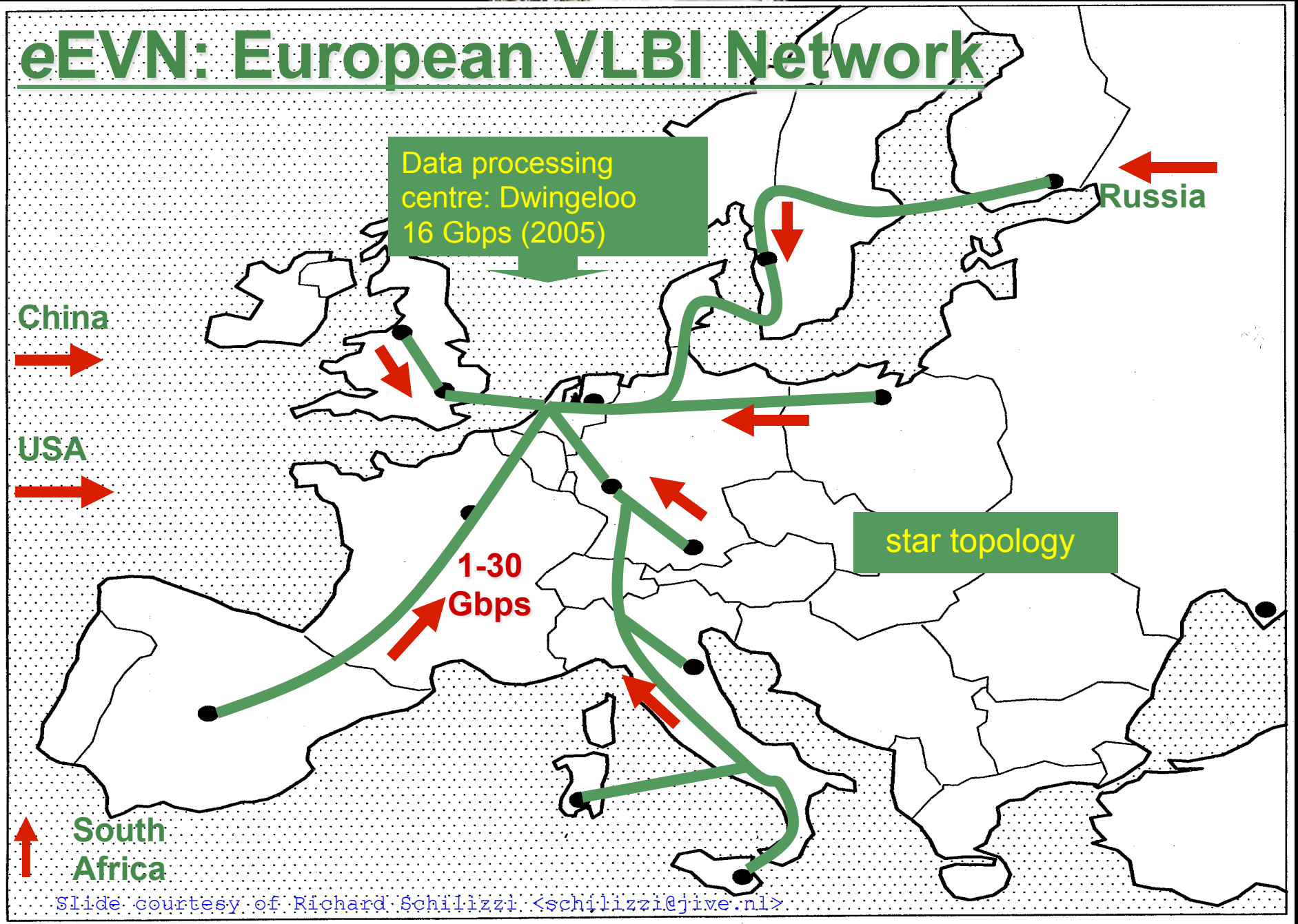

Ny
Mh
Sv
Jb
On
Cm
Wb
Lv
Ef
Tr
Wz
Ro
Mc
Ys
Sr
Nt
Ev
Zc
Ar
Bd
My
Ur
Sh
Ku
Hh


2008
2009
Deadline for submitting observing proposals
Program committee:
* rates proposals
* allocates observing time
VLBI Observing Session
Disks shipped to JIVE
Correlation at JIVE
Data shipped
Data arrives at
at scientist's desk!
Sep
Oct
Nov
Dec
Jan
Feb
Mar
Apr
May
Jun

2008

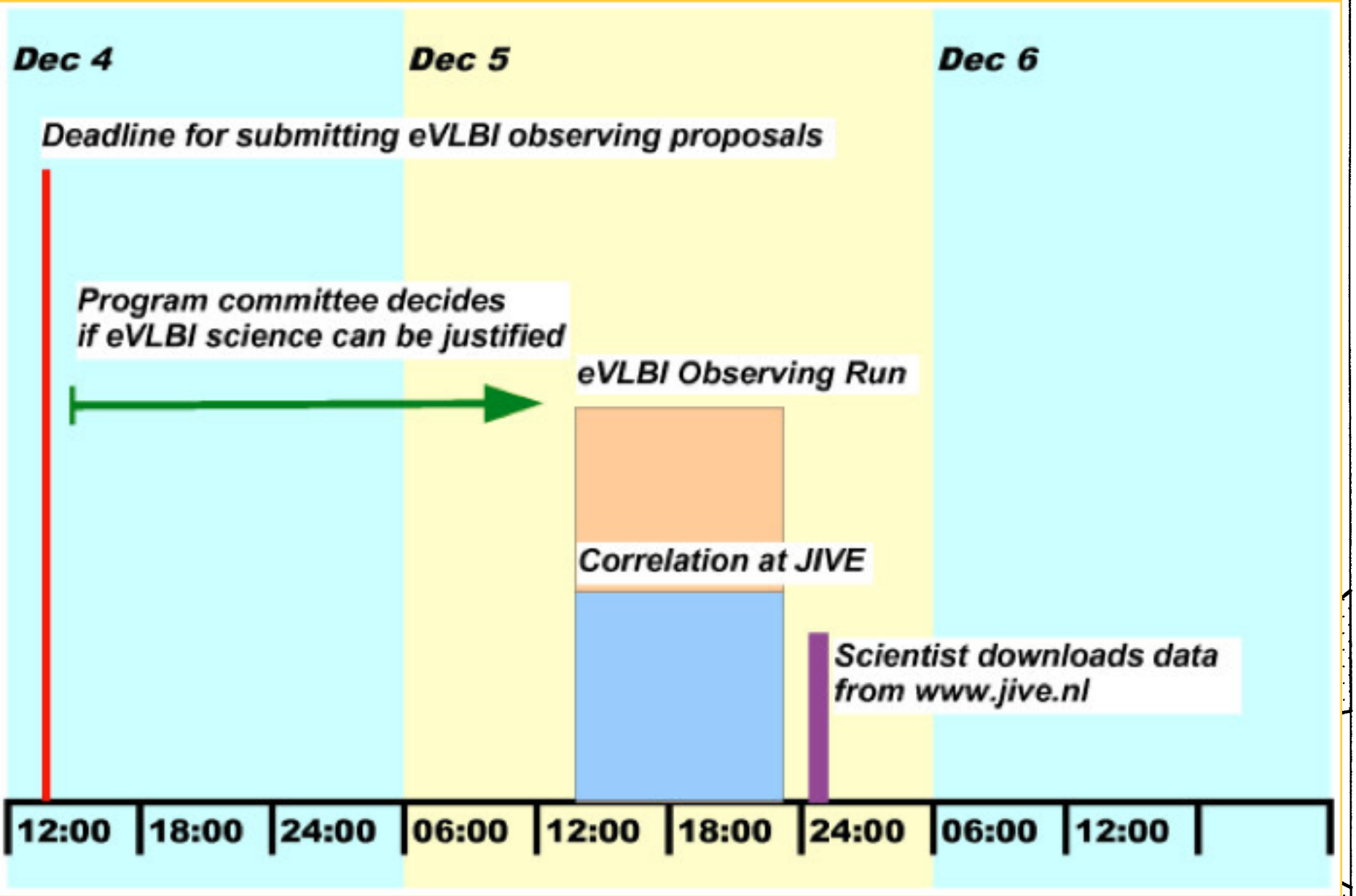
2009

eEVN: European VLBI Network



Slide courtesy of Richard Schilizzi <schilizzi@jive.nl>

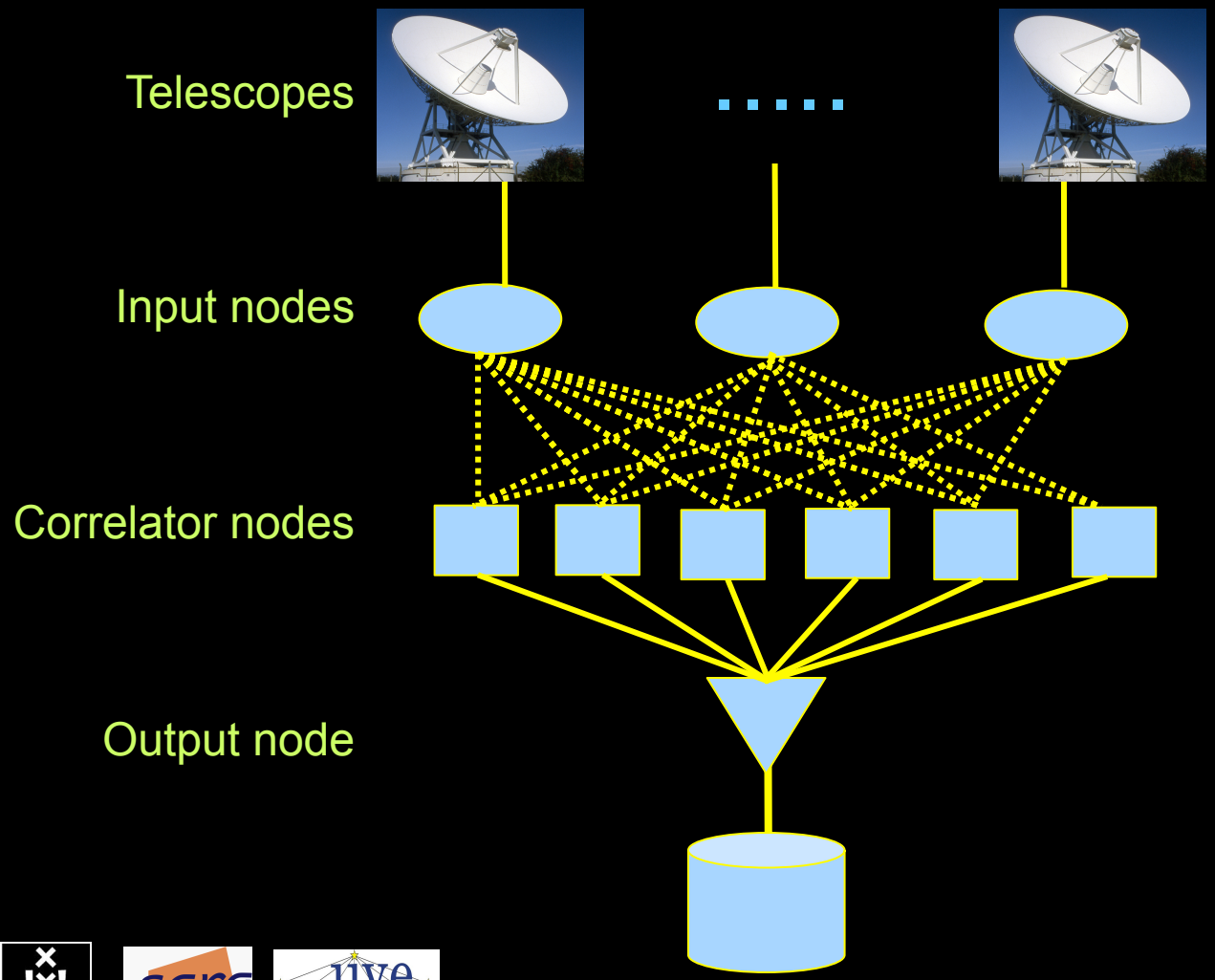
eEVN: European VLBI Network



Slide courtesy of Richard Schilizzi <schilizzi@jive.nl>

The SCARIe project

SCARIe: a research project to create a Software Correlator for e-VLBI.
VLBI Correlation: signal processing technique to get high precision image from spatially distributed radio-telescope.



16 Gbit/s - 2 Tflop ➔ THIS IS A DATA FLOW PROBLEM !!!

Research:

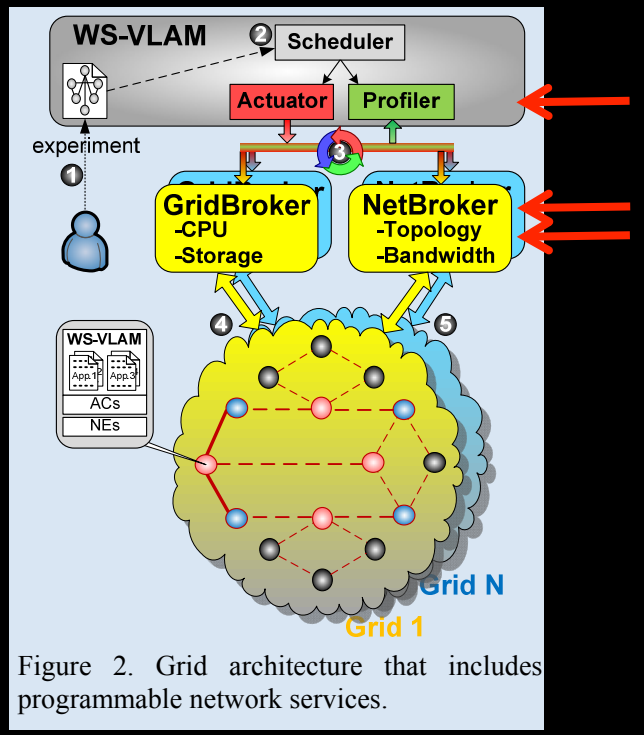
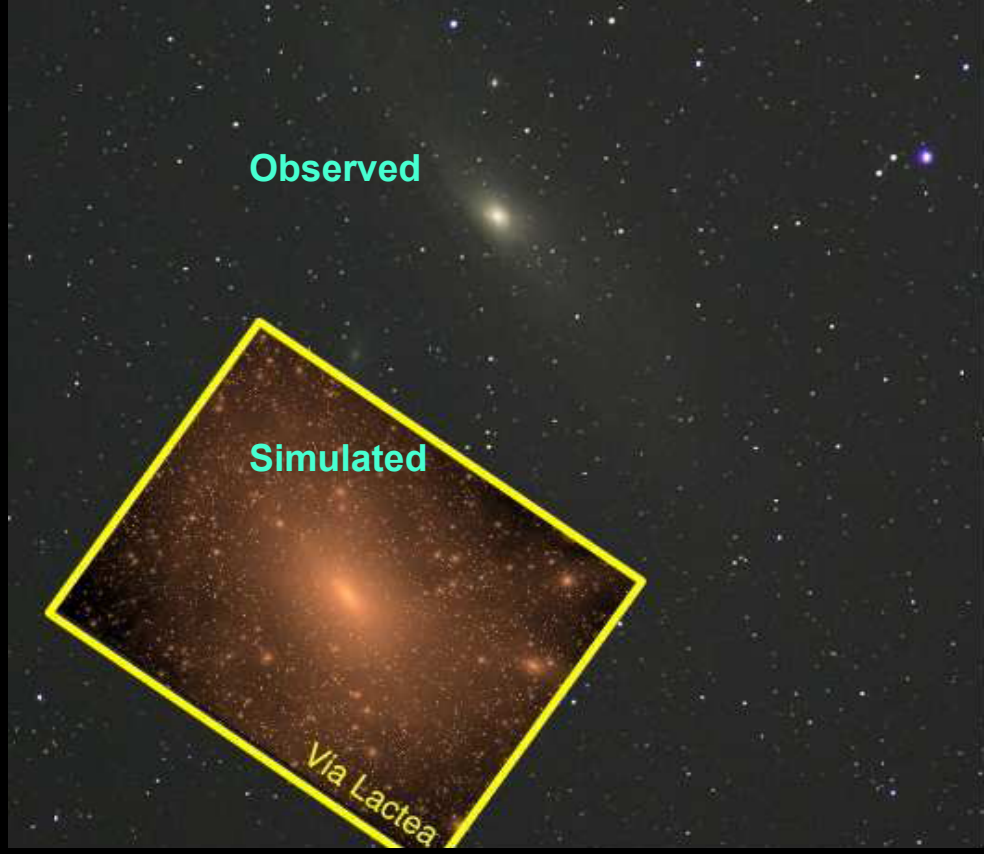


Figure 2. Grid architecture that includes programmable network services.

CosmoGrid

- Motivation:
  previous simulations found >100 times more substructure than is observed!



- Simulate large structure formation in the Universe
  - Dark Energy (cosmological constant)
  - Dark Matter (particles)
- Method: Cosmological *N*-body code
- Computation: Intercontinental SuperComputer Grid

The hardware setup

10 Mflops/byte

1 Eflops/s

- 2 supercomputers :
  - 1 in Amsterdam (60 Tflops Power6 @ SARA)
  - 1 in Tokyo (30 Tflops Cray XD0-4 @ CFCA)
- Both computers are connected via an intercontinental optical 10 Gbit/s network

SNE @ UvA

	Ijkdijk/Urban Flood	Medical	LifeWatch/ENVRI	CosmoGrid/eVLBI	CineGrid	EU-GN3/NOVI/Geysers	SURFnet/GLIF/Cloud
Green-IT						X	X
Privacy/Trust		X			X		
Authorization/policy		X	X		X	X	
Programmable networks	X		X				
40-100Gig/TCP/WF/QoS	X		X	X		X	
Topology/Architecture		X		X	X	X	
Optical Photonic		X	X		X		

IJKDIJK

Sensors: 15000km* 800 bps/m ->12 Gbit/s to cover all Dutch dikes

Sensor grid: instrument the dikes

First controlled breach occurred on sept 27th ‘08:

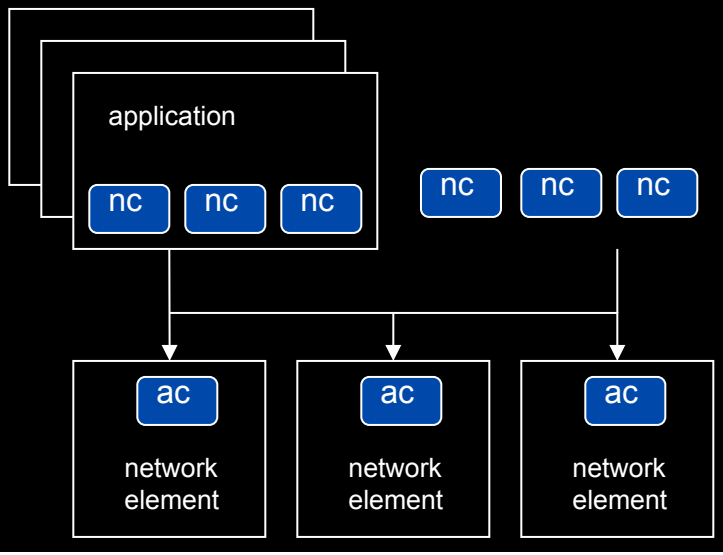
Tera-Thinking

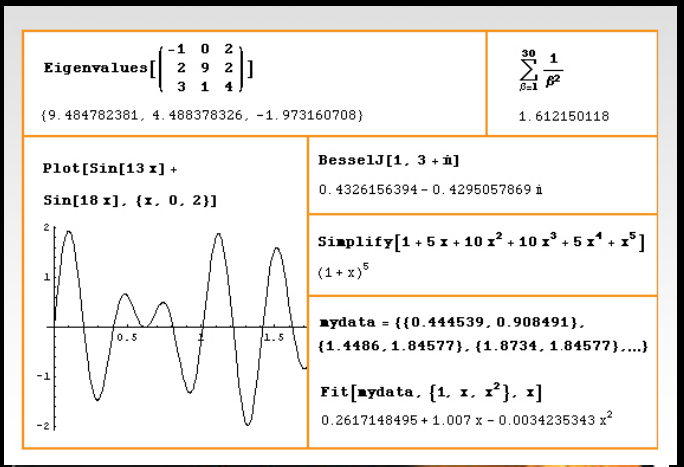
- What constitutes a Tb/s network?
- think back to teraflop computing!
  - MPI turns a room full of pc's in a teraflop machine
- massive parallel channels in hosts, NIC's
- TeraApps programming model supported by
  - TFlops -> MPI / Globus / Cloud
  - TBytes -> DAIS / MONETdb …
  - TPixels -> SAGE
  - TSensors -> LOFAR, LHC, LOOKING, CineGrid, ...
  - Tbit/s -> ?

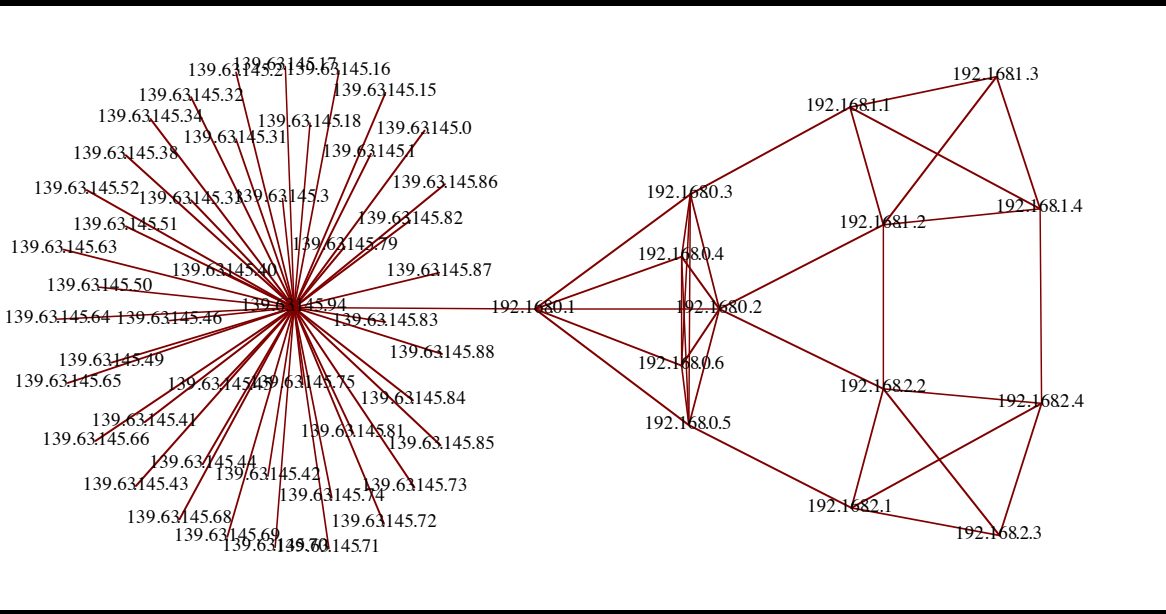
  - ? -> Programmable Networks

User Programmable Virtualized Networks.

- The network is virtualized as a collection of resources
- UPVNs enable network resources to be programmed as part of the application
- Mathematica interacts with virtualized networks using UPVNs and optimize network + computation



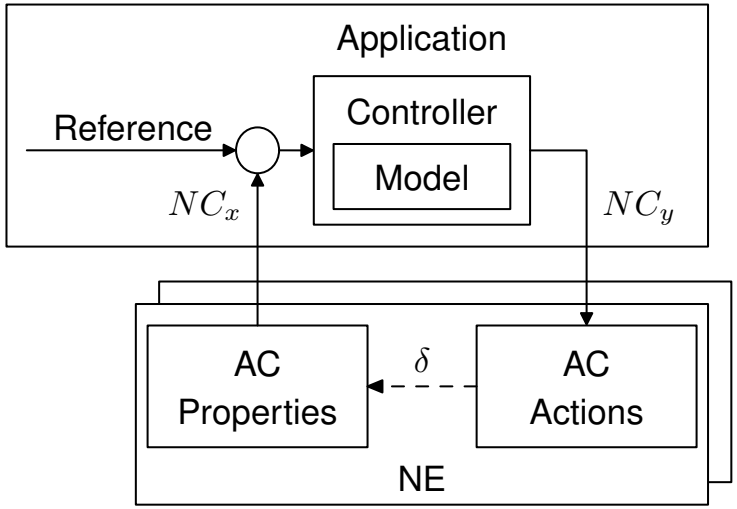




ref: Robert J. Meijer, Rudolf J. Strijkers, Leon Gommans, Cees de Laat, User Programmable Virtualiized Networks, accepted for publication to the IEEE e-Science 2006 conference Amsterdam.

In the Intercloud virtual servers and networks become software

- Virtual Internets adapt to the environment, grow to demand, iterate to specific designs
- Network support for application specific interconnections are merely opitimizations: Openflow, active networks, cisco distributed switch
- But how to control the control loop?

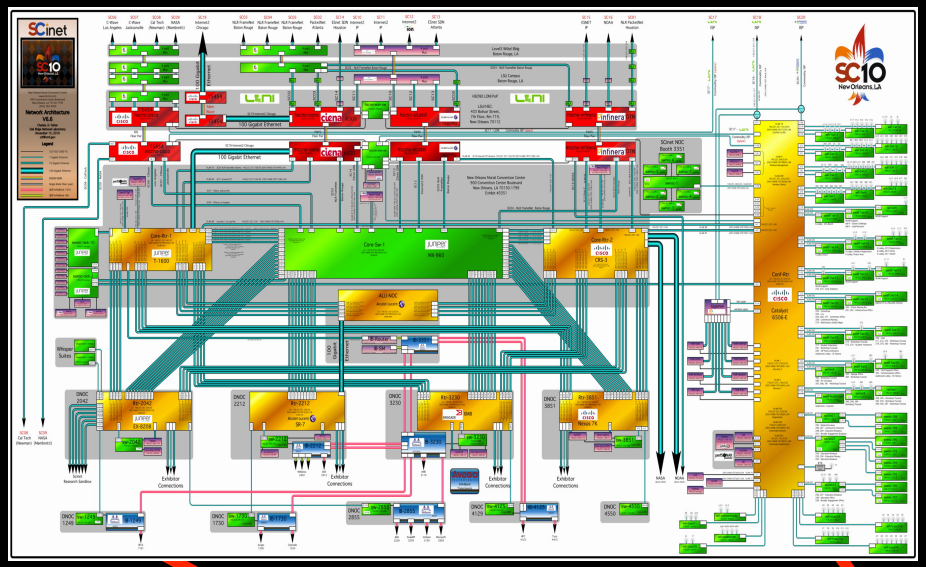


Interactive Networks

Rudolf Strijkers [1,2]
Marc X. Makkes [1,2]
Mihai Christea [1]
Laurence Muller [1]
Robert Belleman [1]
Cees de Laat [1]
Robert Meijer [1,2]
[1] University of Amsterdam, Amsterdam The Netherlands
[2] TNO Information and Communication Technology, Groningen, The Netherlands

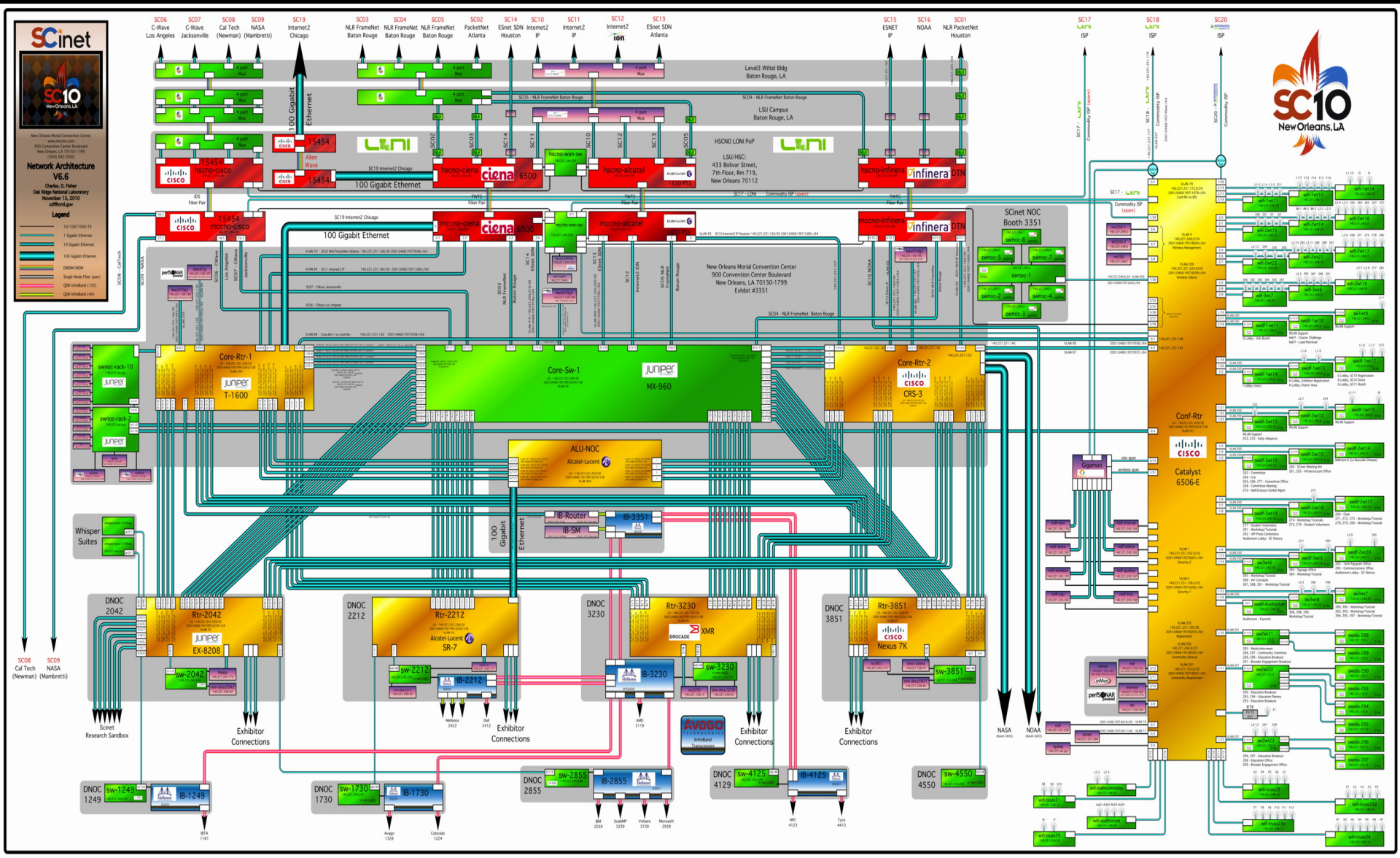
Mastering Complexity

SNE @ UvA



	Ijkdijk/Urban Flood	Medical	LifeWatch/ENVRI	CosmoGrid/eVLBI	CineGrid	EU-GN3/NOVI/Geysers	SURFnet/GLIF/Cloud
Green-IT						X	X
Privacy/Trust		X			X		
Authorization/policy		X	X		X	X	
Programmable networks	X		X				
40-100Gig/TCP/WF/QoS	X		X	X		X	
Topology/Architecture		X		X	X	X	
Optical Photonic		X	X		X		

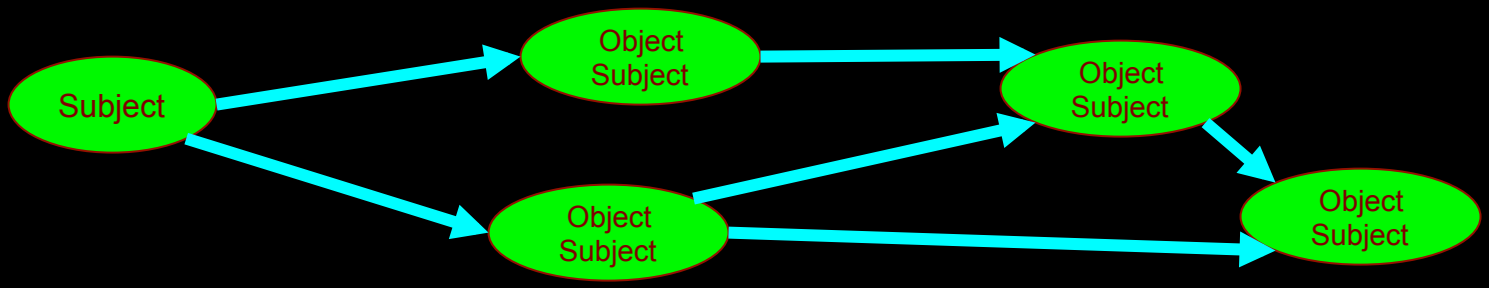
Complex eInfrastructure @ SC10

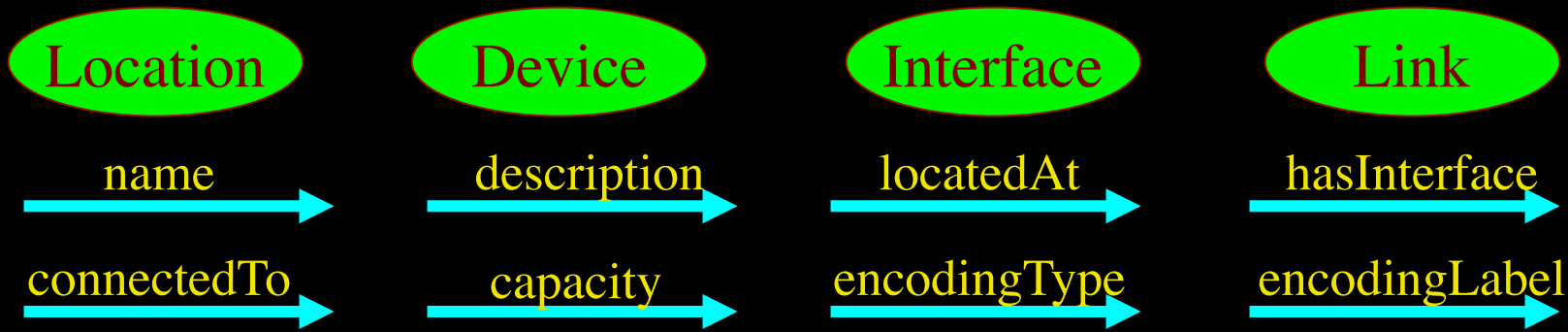


We investigate: for complex networks!

LinkedIN for Infrastructure

- From semantic Web / Resource Description Framework.
- The RDF uses XML as an interchange syntax.
- Data is described by triplets (Friend of a Friend):



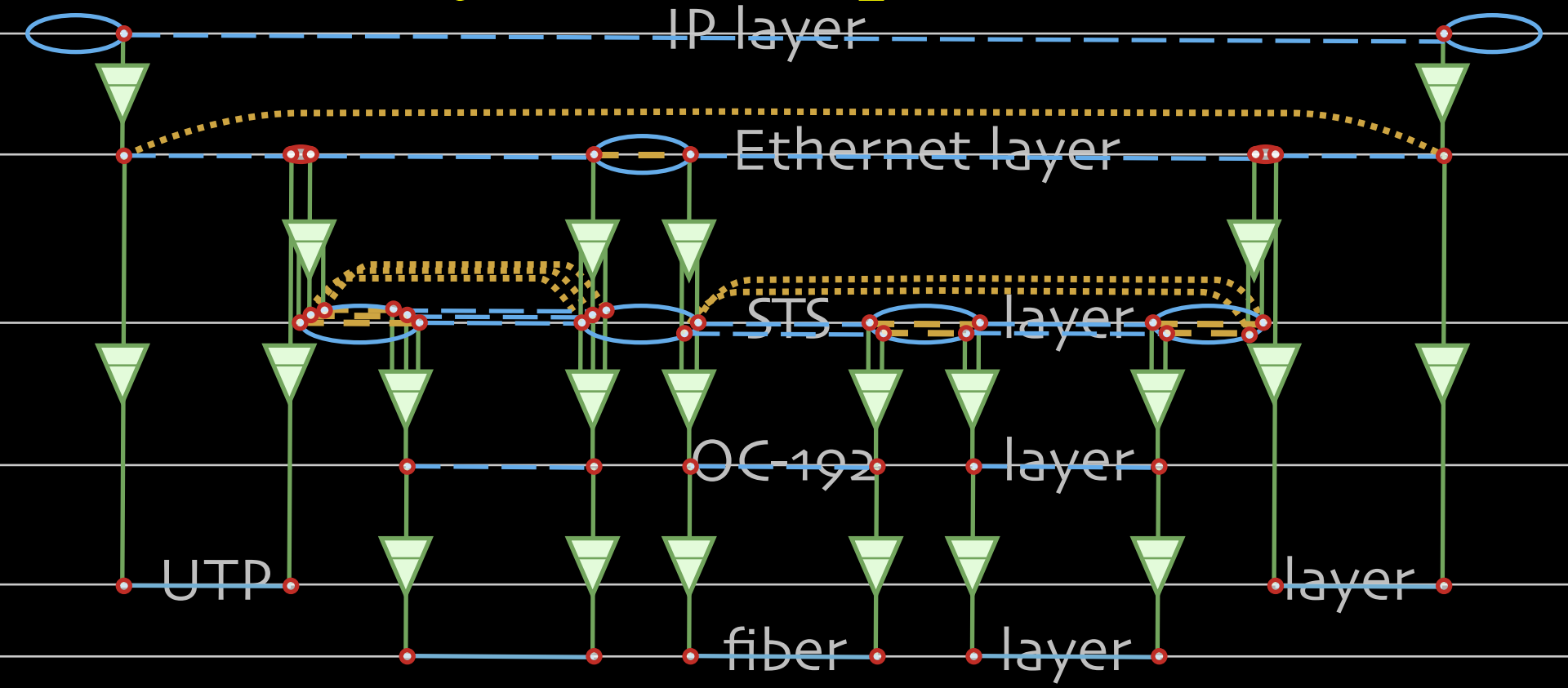


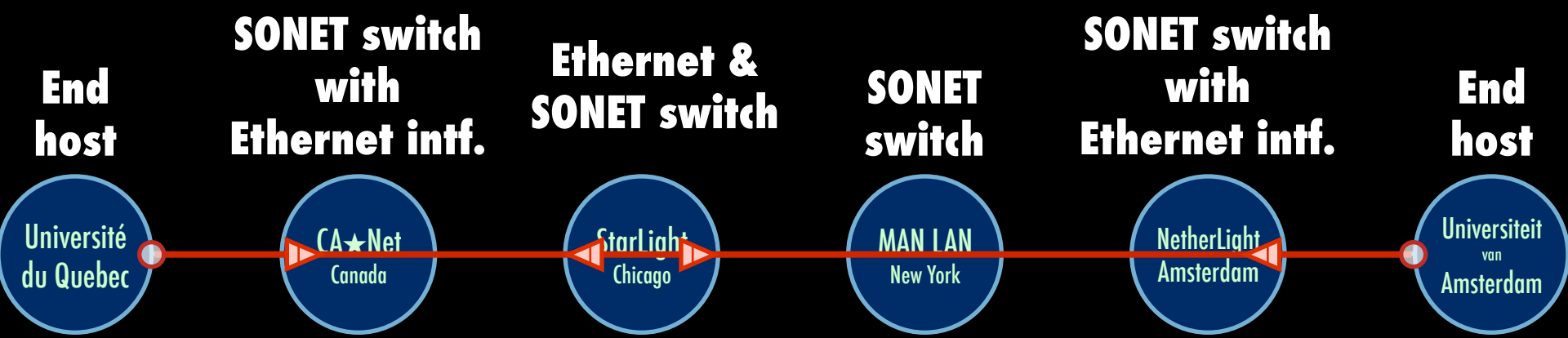
NetherLight in RDF

```
<?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:resourc
      <ndl:hasInterface rdf:resourc
      <ndl:hasInterface rdf:resourc
      <ndl:hasInterface rdf:resourc
      <ndl:hasInterface rdf:resourc
      <ndl:hasInterface rdf:resourc
      <ndl:hasInterface rdf:resourc
```

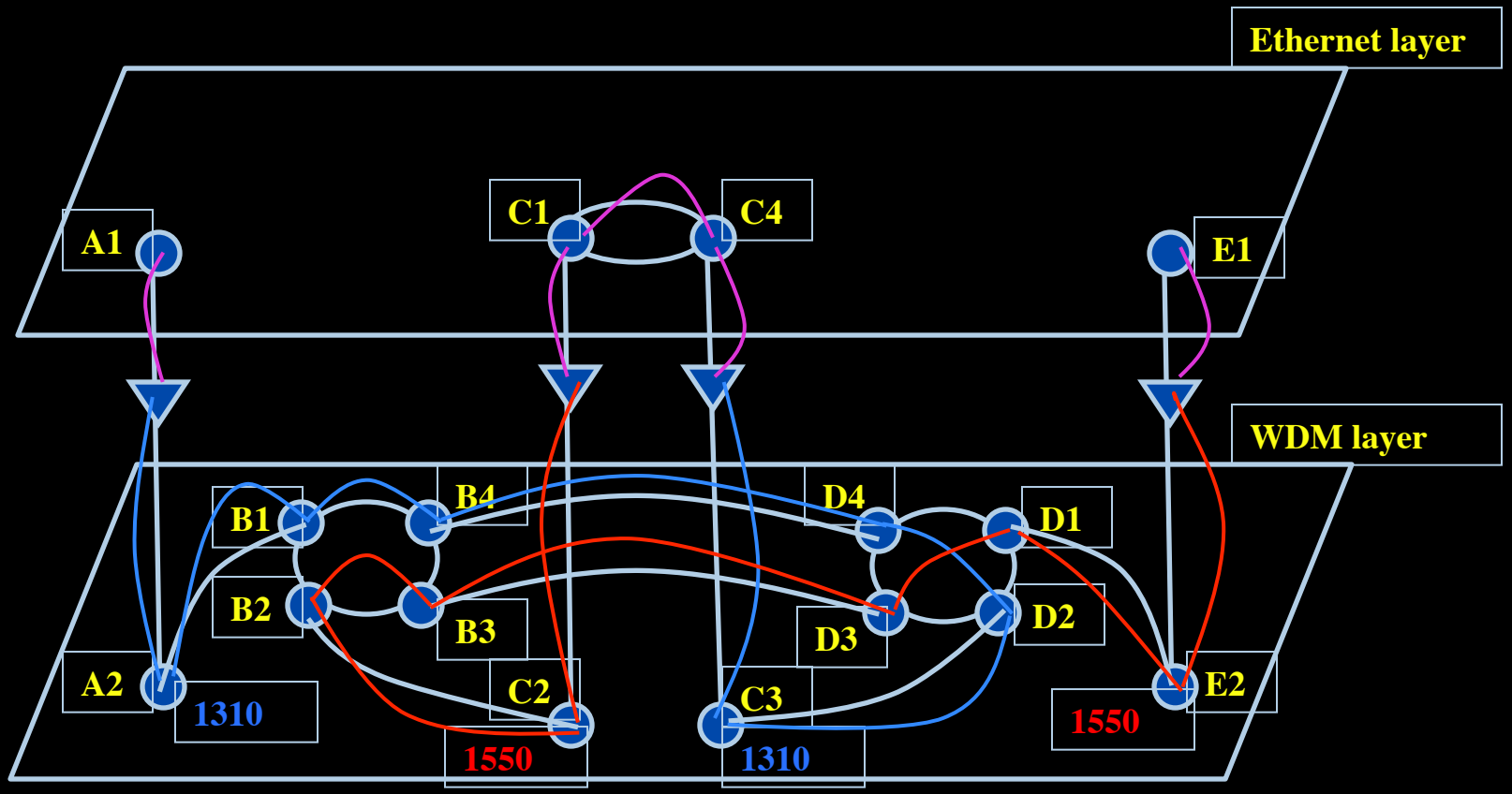
```
<!-- all the interfaces of TDM3.amsterdam1.netherlight.net -->

<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>
```

Multi-layer descriptions in NDL





Multi-layer Network PathFinding

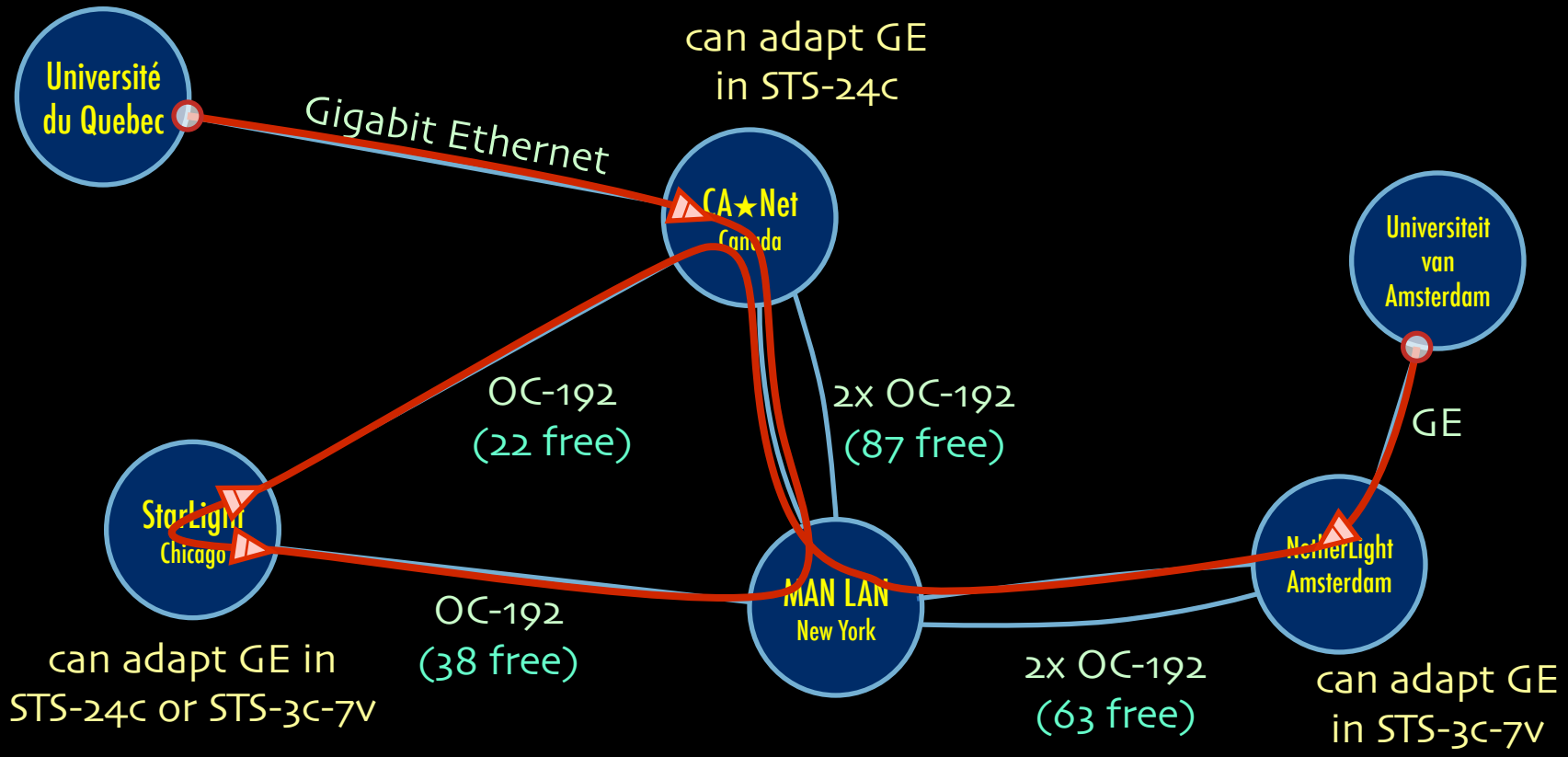


Path between interfaces A1 and E1:

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

Scaling: Combinatorial problem

A weird example



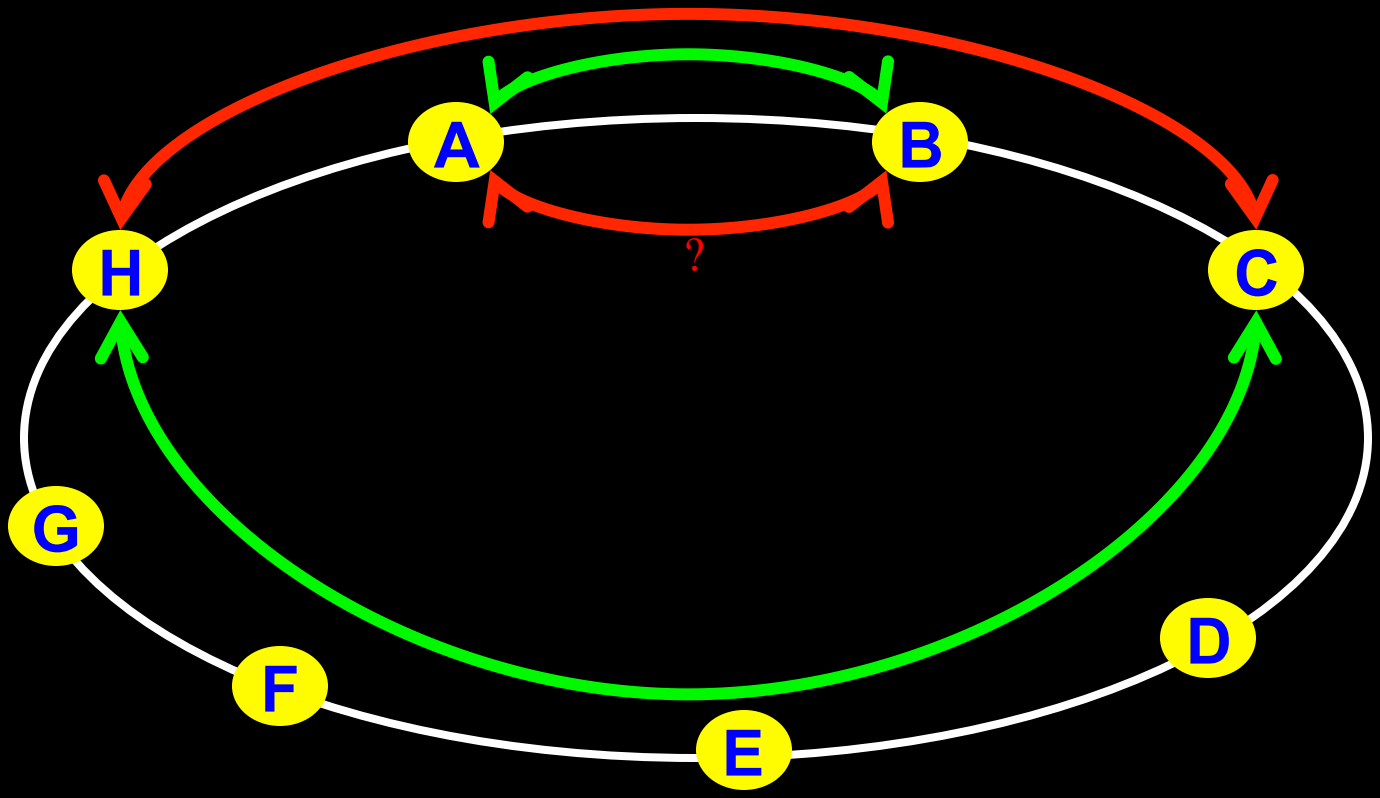
Thanks to Freek Dijkstra & team

The Problem

I want HC and AB

Success depends on the order

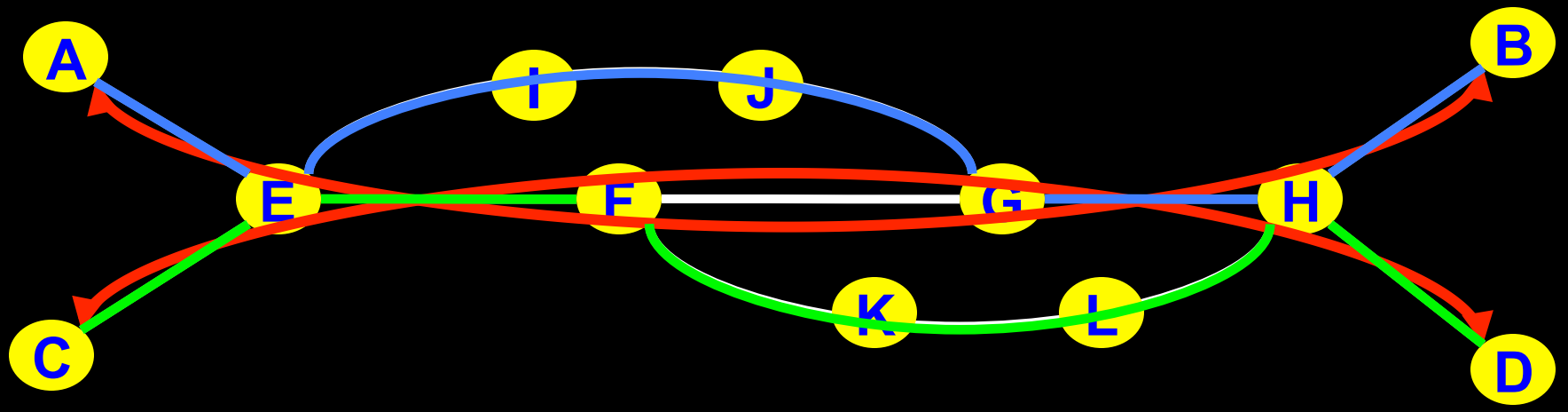
Wouldn't it be nice if I could request [HC, AB, ...]



Another one ☺

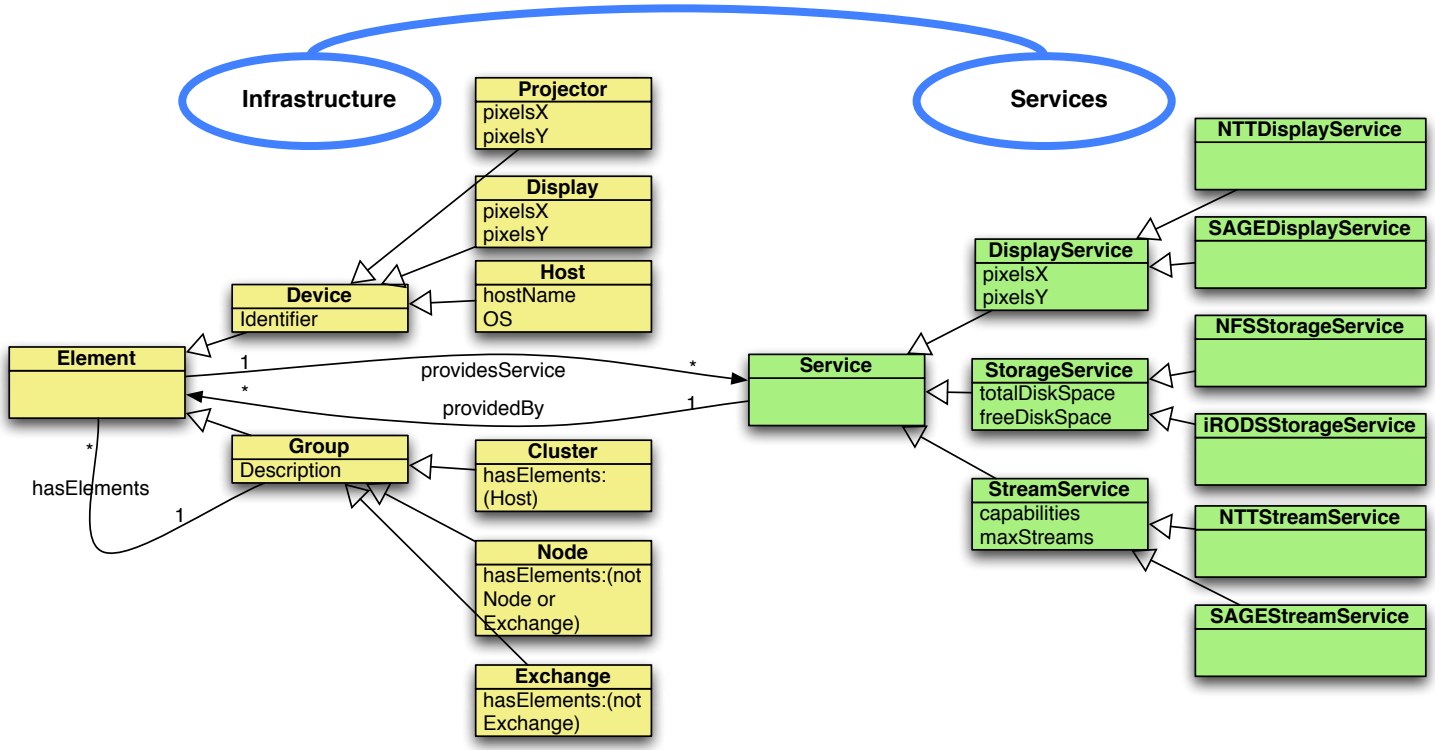
I want AB and CD

Success does not even depend on the order!!!



Information Modeling

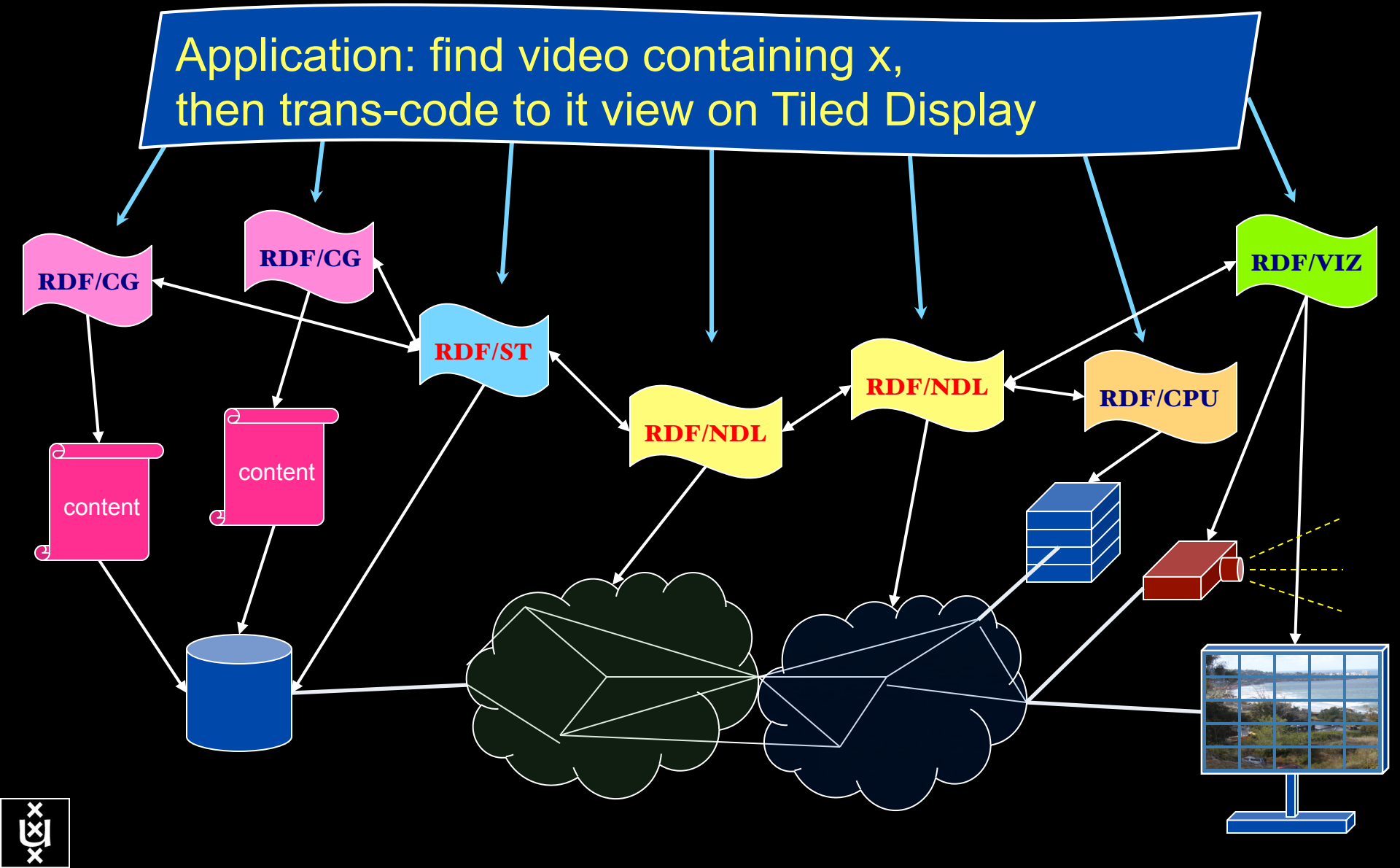
Define a common information model for ***infrastructures*** and ***services***. Base it on Semantic Web.



J. van der Ham, F. Dijkstra, P. Grosso, R. van der Pol, A. Toonk, C. de Laat
A distributed topology information system for optical networks based on the semantic web,
In: Elsevier Journal on Optical Switching and Networking, Volume 5, Issues 2-3, June 2008, Pages 85-93

R.Koning, P.Grosso and C.de Laat
Using ontologies for resource description in the CineGrid Exchange
In: Future Generation Computer Systems (2010)

RDF describing Infrastructure "I want"



SNE @ UvA

	Ijkdijk/Urban Flood	Medical	LifeWatch/ENVRI	CosmoGrid/eVLBI	EU-GN3/NOVI/Geysers	CineGrid	SURFnet/GLIF/Cloud
Green-IT						X	X
Privacy/Trust		X			X		
Authorization/policy		X	X		X	X	
Programmable networks	X		X				
40-100Gig/TCP/WF/QoS	X		X	X		X	
Topology/Architecture		X		X	X	X	
Optical Photonic			X	X		X	

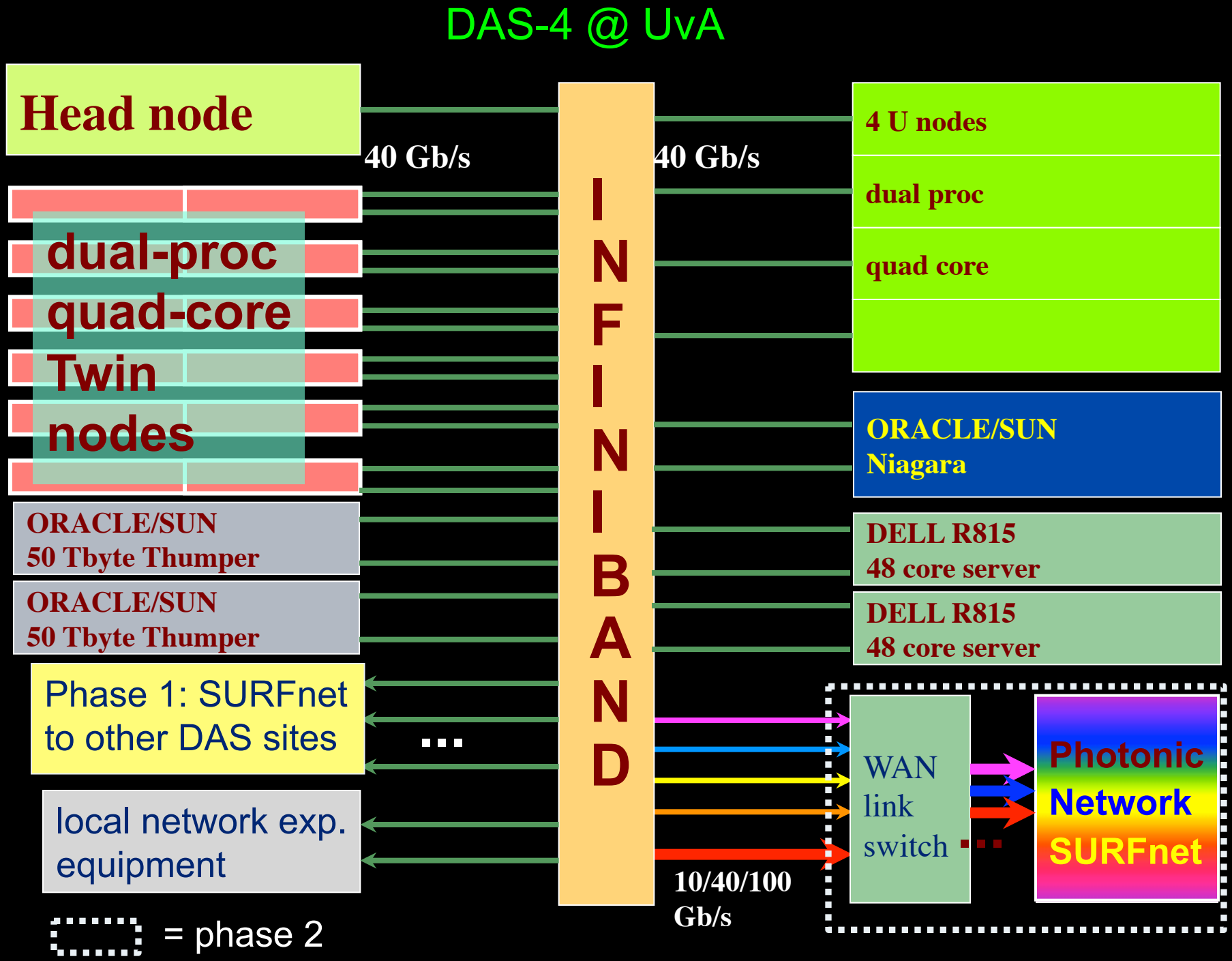
Partners in GreenClouds

- Free University of Amsterdam
  - Henri Bal
- (really free) University of Amsterdam
  - Paola Grosso, Cees de Laat
- SARA
  - Axel Berg
- In context of:
  - ASCI
  - DAS4

GreenClouds @ VU & UvA

- The GreenClouds project studies how to reduce the energy footprint of modern High Performance Computing systems (like Clouds) that are distributed, elastically scalable, and contain a variety of hardware (accelerators and hybrid networks). The project takes a system-level approach and studies the problem of how to map high-performance applications onto such distributed systems, taking both performance and energy consumption into account.

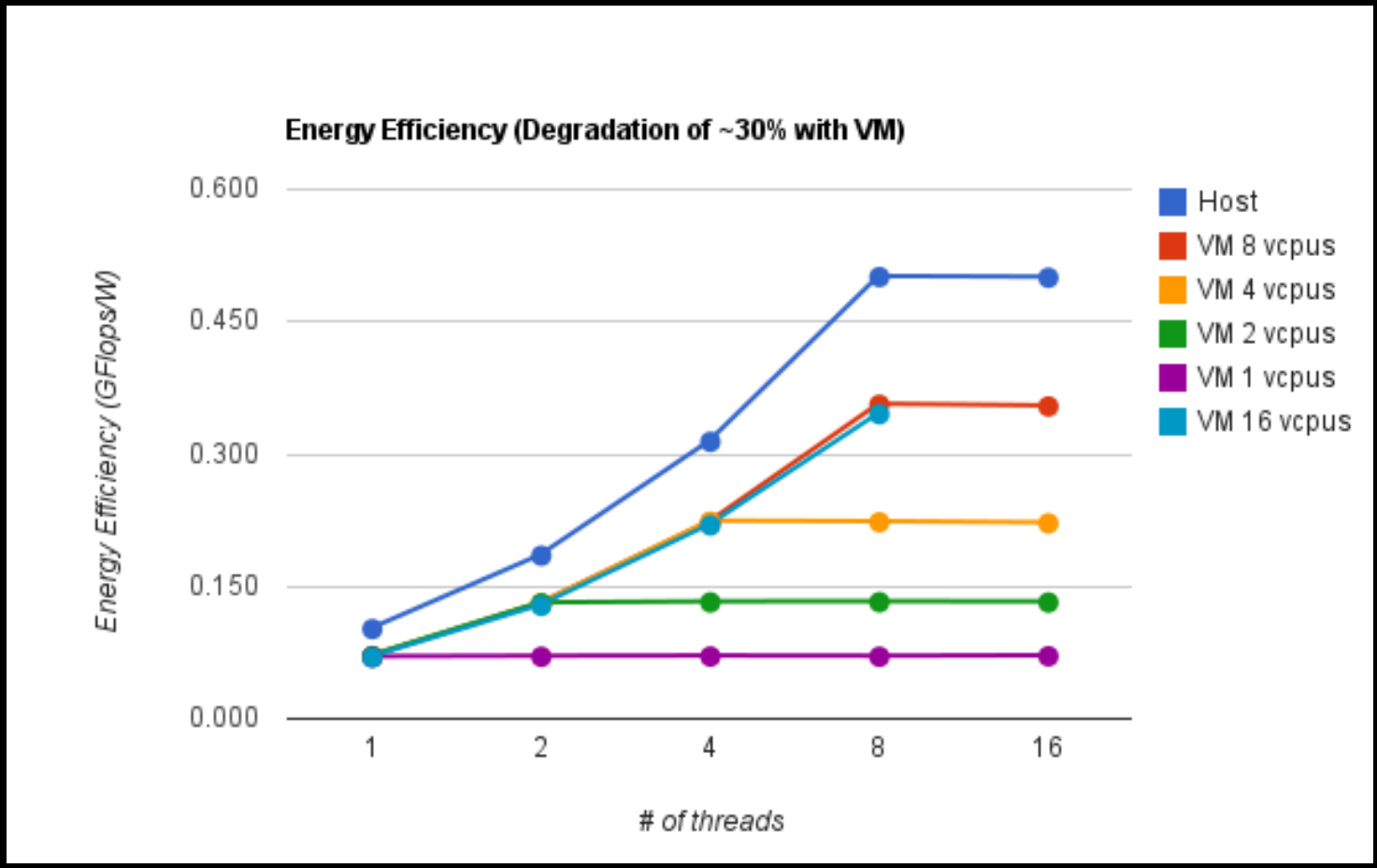
- We will explore three ideas to reduce energy:
  1. Exploit the diversity of computing architectures (e.g. GPUs, multicores) to run computations on those architectures that perform them in the most energy-efficient way;
  2. Dynamically adapt the number of resources to the application needs accounting for computational and energy efficiency;
  3. Use optical and photonic networks to transport data and computations in a more energy-efficient way.


DAS-4 @ UvA
Head node
40 Gb/s
dual-proc
quad-core
Twin
nodes
ORACLE/SUN
50 Tbyte Thumper
ORACLE/SUN
50 Tbyte Thumper
Phase 1: SURFnet
to other DAS sites
...
local network exp.
equipment
= phase 2
INFINIBAND
40 Gb/s
4 U nodes
dual proc
quad core
ORACLE/SUN
Niagara
DELL R815
48 core server
DELL R815
48 core server
10/40/100
Gb/s
WAN
link
switch
Photonic
Network
SURFnet

GreenClouds @ VU & UvA

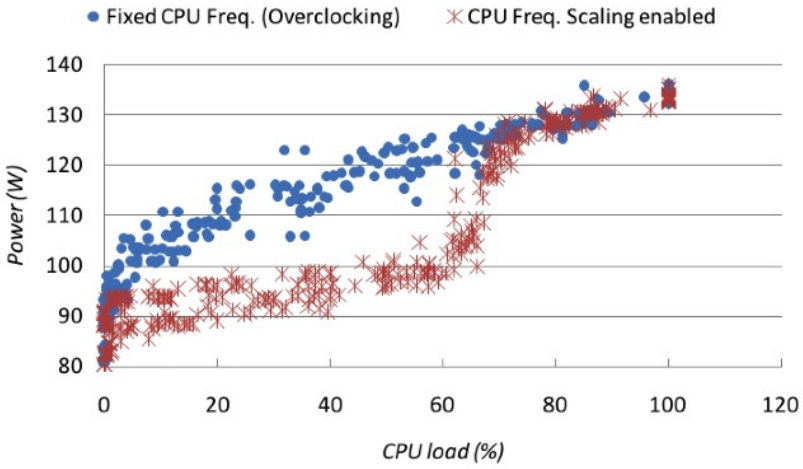
- GreenClouds Knowledge Base System (GKBS) based on semantic web technology (NDL – alike)
  - detailed information on the energy characteristics of various applications (previous execution runs
  - Information on different parts of the distributed system, including the network.
- Determine classes of applications that can reduce their energy consumption using accelerators
- study energy reductions through dynamic adaptation of computing and networking resources.

The project will make extensive use of the DAS-4 infrastructure, which is a wide-area testbed for computer scientists, to be equipped with many types of accelerators, a photonic network, and energy sensors.

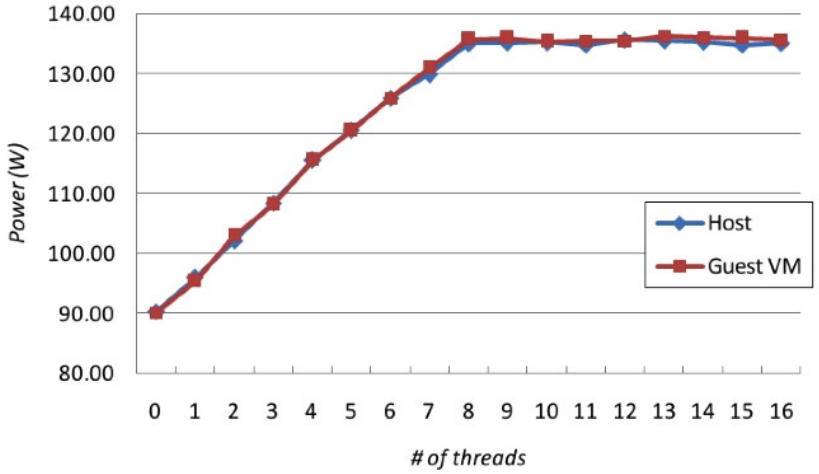


Each benchmark is run with the same amount of memory.
The degradation in energy efficiency of VMs is around 30% compared with the host.

CPU



Gradual increase of CPU load on all available cores

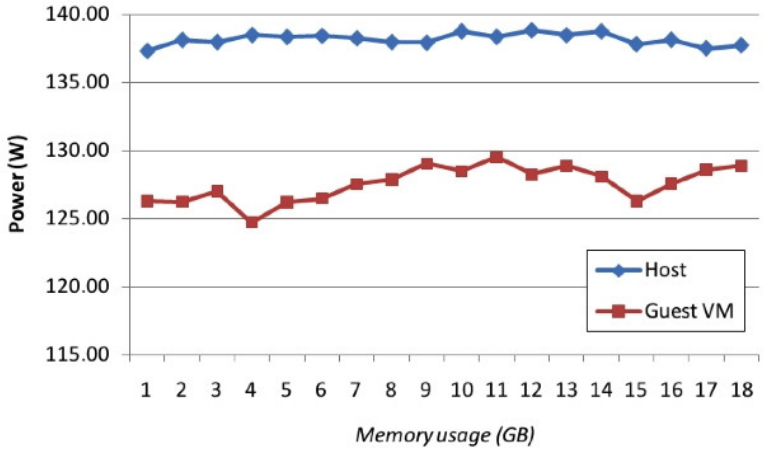


Gradual increase of number of cores, where each core is at its maximum usage

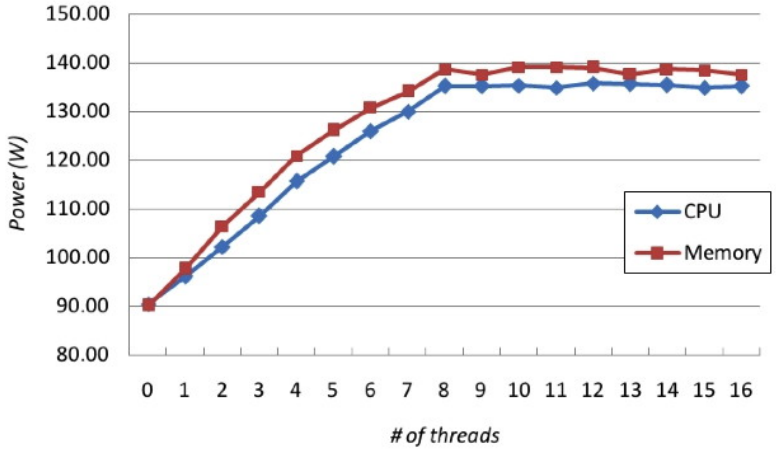
Observations

- Power usage is linear to the CPU load.
- No significant differences in power usage of a VM and its host.

Memory



Varying memory usage



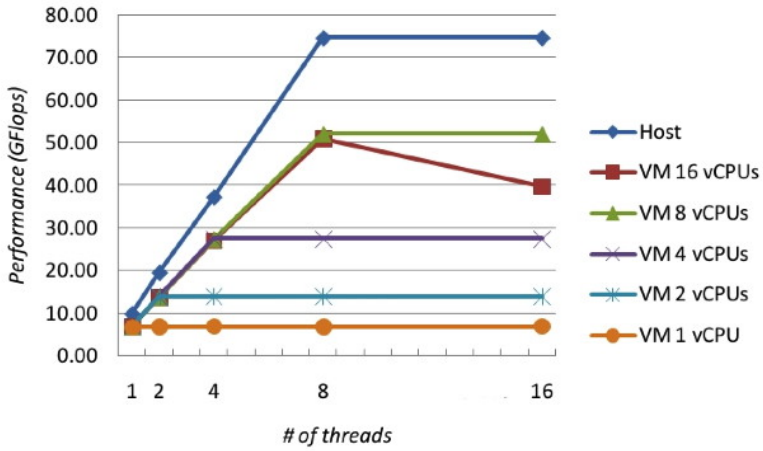
Memory and CPU stress tests

Observations

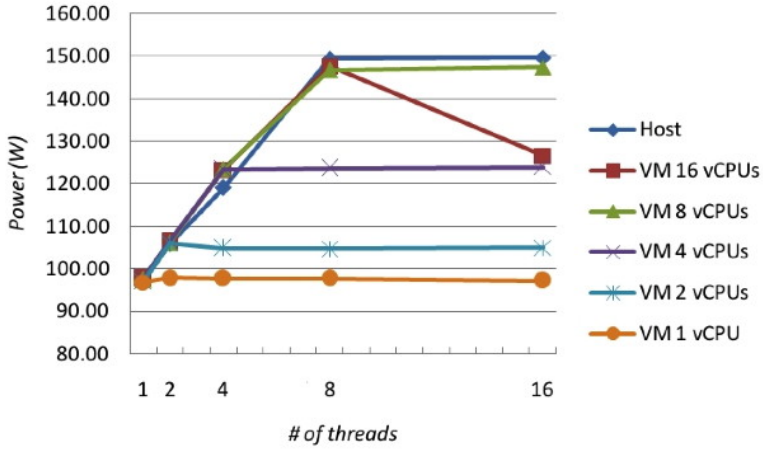
- Nearly constant power usage of memory
- Variation is less than 10% of total power usage

Overall benchmarks

Floating-point operation (Linpack) test



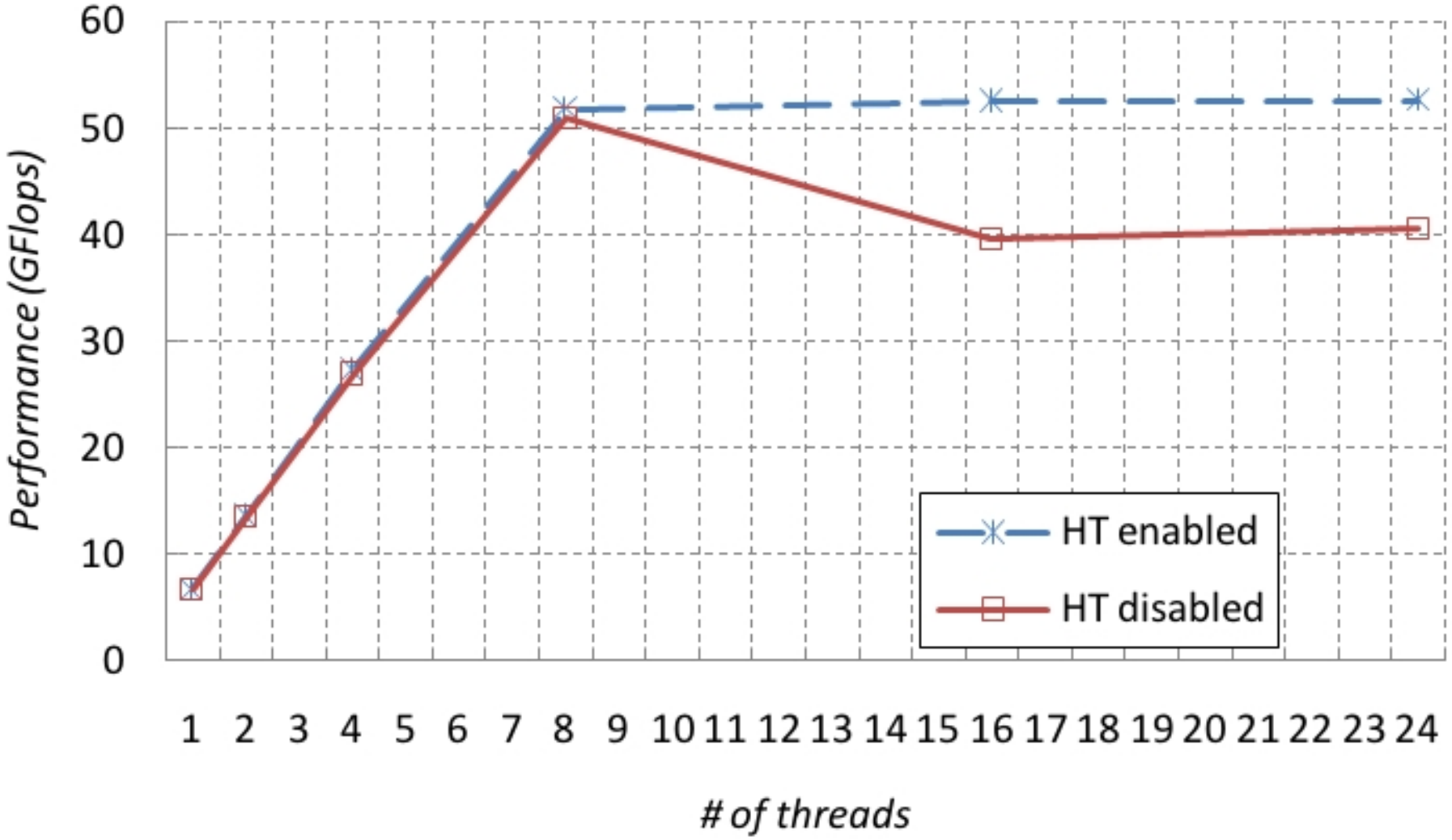
Performance



Power usage

Observations

- Performance $\propto$ CPU load (# of threads).
- Power usage is nearly linear to CPU load.
- Abnormal result for over-committed VM (i.e. with 16 vCPUs).

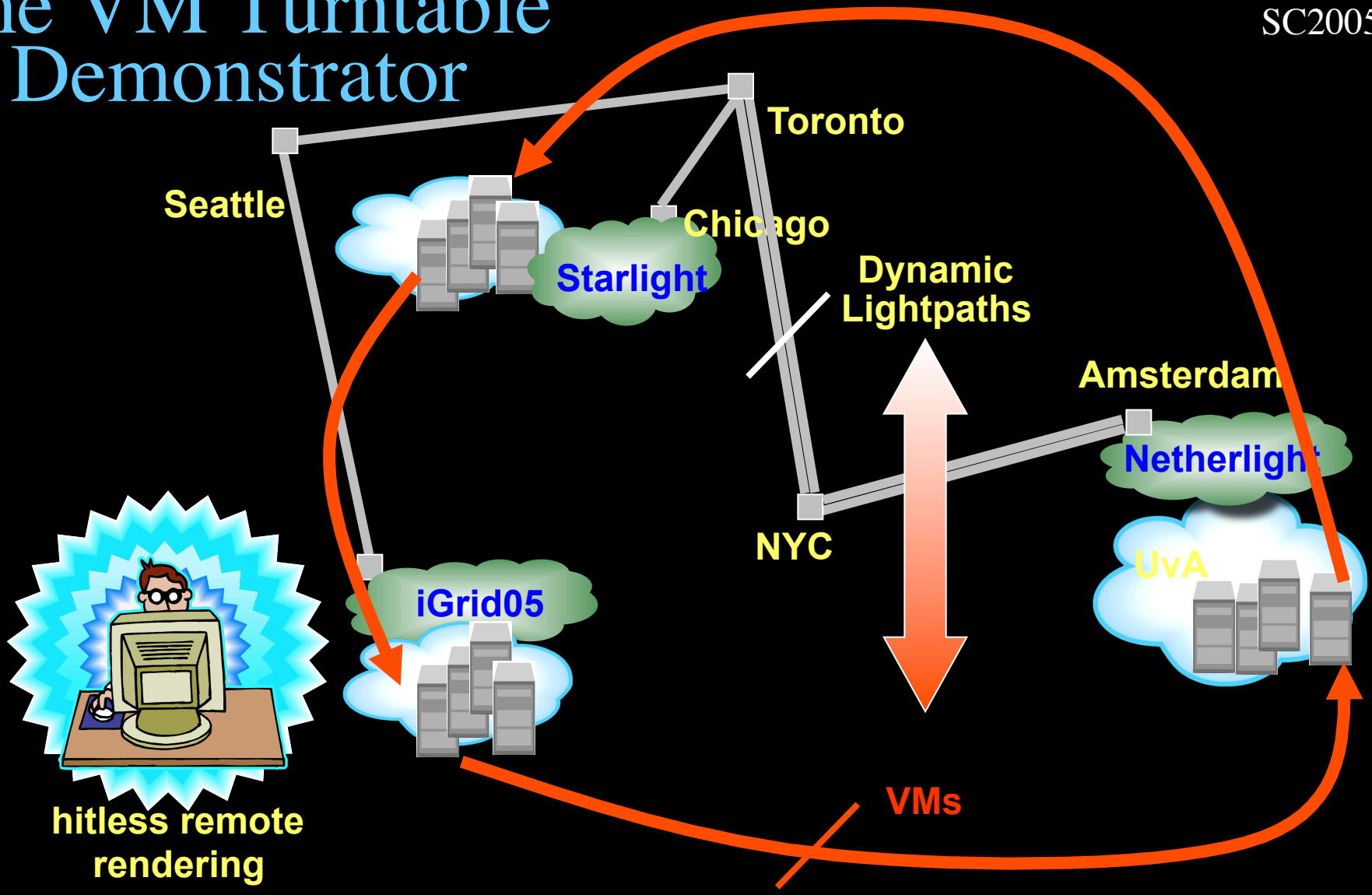

60
50
40
30
20
10
0
Performance (GFlops)
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
# of threads
HT enabled
HT disabled

Semantic web approach in GreenClouds

- Distributed info system describing current and historical load on infrastructure including parameters of jobs running
- Describe contextual parameters (energy sources, etc.)
- Dynamically optimize and migrate if context changes

The VM Turntable Demonstrator

iGrid2005
SC2005



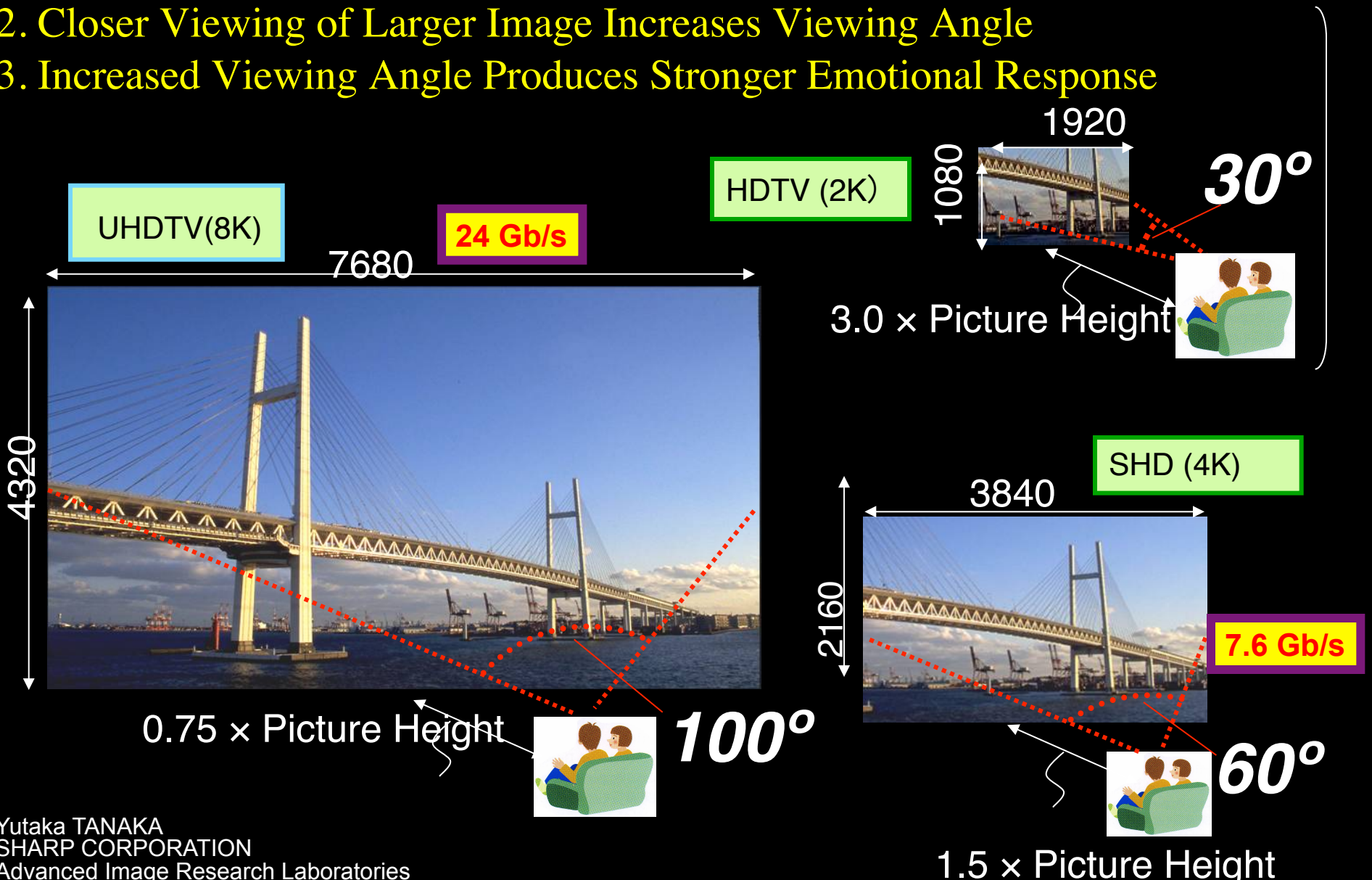
The VMs that are live-migrated run an iterative search-refine-search workflow against data stored in different databases at the various locations. A user in San Diego gets hitless rendering of search progress as VMs spin around

SNE @ UvA

	Ijkdijk/Urban Flood	Medical	LifeWatch/ENVRI	CosmoGrid/eVLBI	CineGrid	EU-GN3/NOVI/Geysers	SURFnet/GLIF/Cloud
Green-IT						X	X
Privacy/Trust		X			X		
Authorization/policy		X	X		X	X	
Programmable networks	X		X				
40-100Gig/TCP/WF/QoS	X		X	X		X	
Topology/Architecture		X		X	X	X	
Optical Photonic			X	X		X	

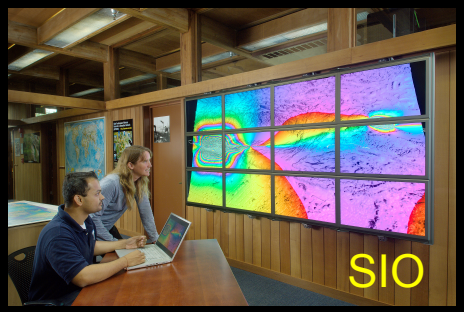
Why is more resolution is better?

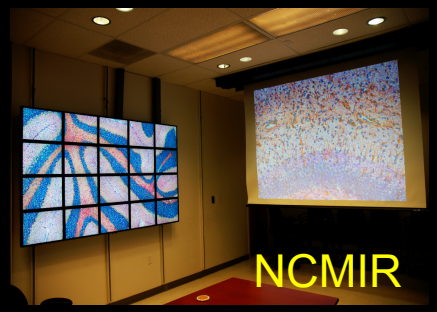
1. More Resolution Allows Closer Viewing of Larger Image
2. Closer Viewing of Larger Image Increases Viewing Angle
3. Increased Viewing Angle Produces Stronger Emotional Response

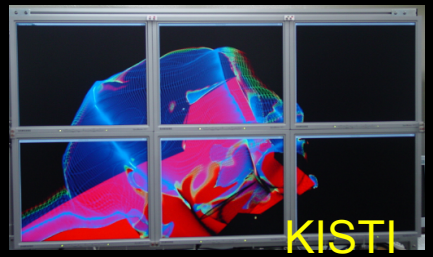


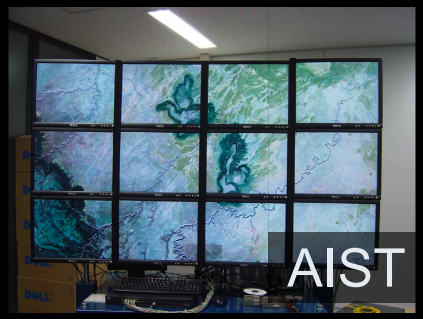
Yutaka TANAKA
SHARP CORPORATION
Advanced Image Research Laboratories

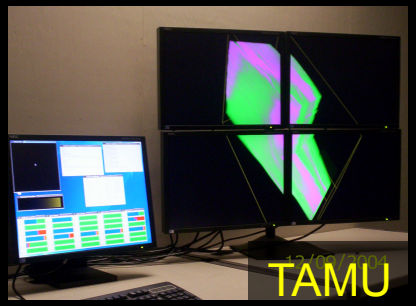
US and International OptIPortal Sites











Real time, multiple 10 Gb/s

Why?

I want to:

"Show Big Bug Bunny in 4K on my Tiled Display using green Infrastructure"

- Big Bugs Bunny can be on multiple servers on the Internet.
- Movie may need processing / recoding to get to 4K for Tiled Display.
- Needs deterministic Green infrastructure for Quality of Experience.
- Consumer / Scientist does not want to know the underlying details.
  ➔ His refrigerator also just works.

The Ten Problems with the Internet

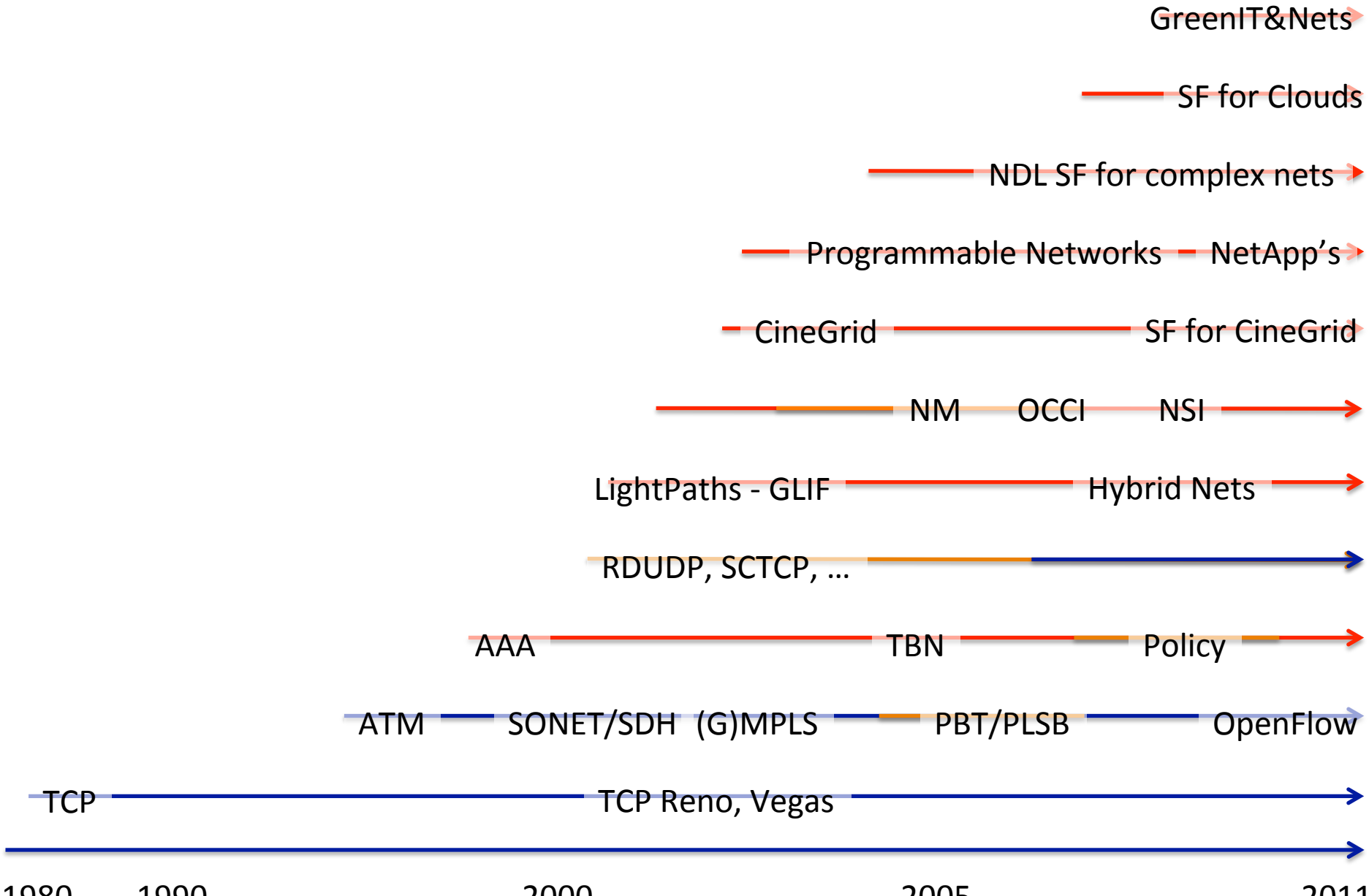
1. **Energy Efficient Communication**
2. Separation of Identity and Address
3. Location Awareness
4. **Explicit Support for Client-Server Traffic and Distributed Services**
5. Person-to-Person Communication
6. Security
7. **Control, Management, and Data Plane separation**
8. **Isolation**
9. Symmetric/Asymmetric Protocols
10. **Quality of Service**

Nice to have:

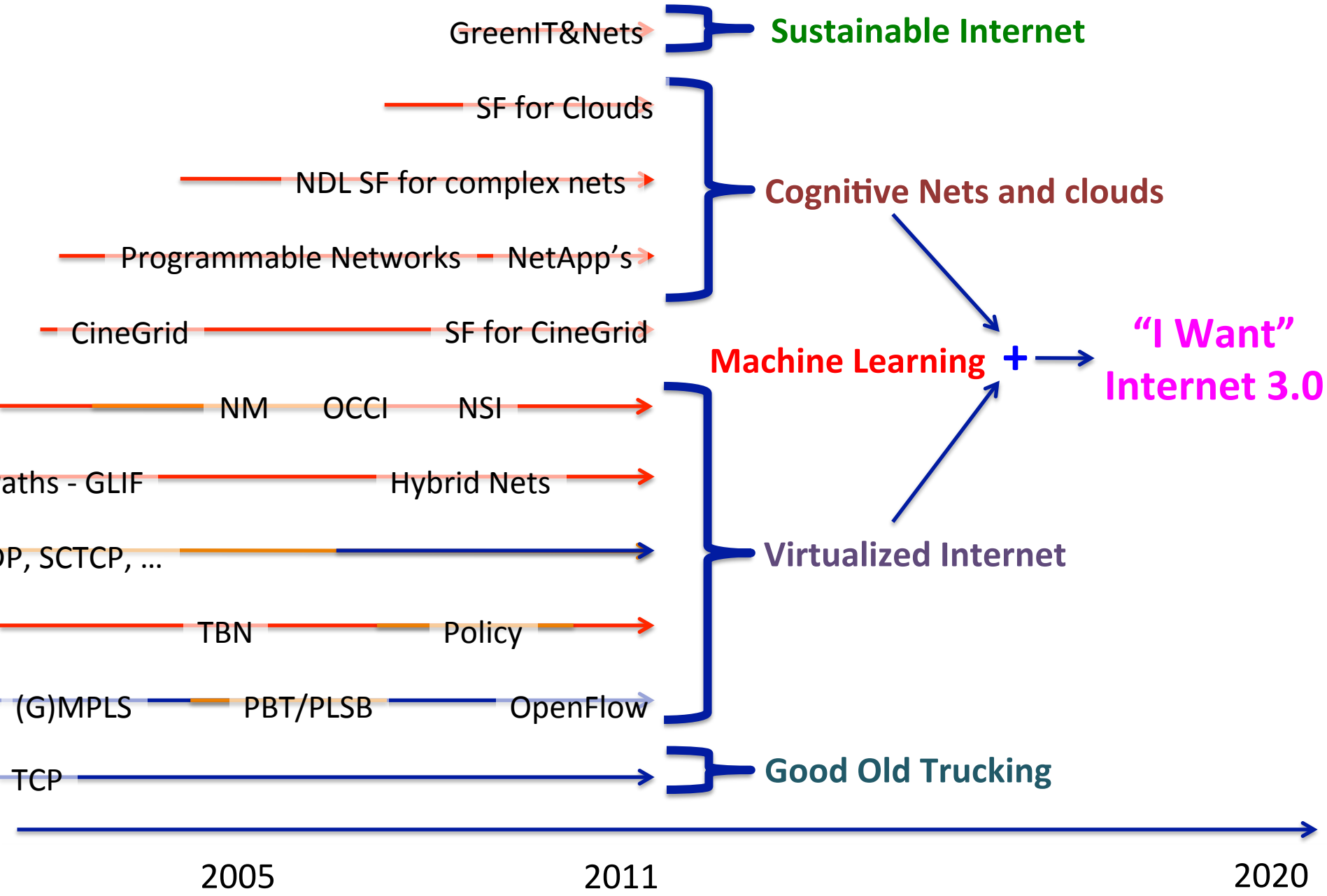
- Global Routing with Local Control of Naming and Addressing
- **Real Time Services**
- **Cross-Layer Communication**
- Manycast
- Receiver Control
- Support for Data Aggregation and Transformation
- **Support for Streaming Data**
- **Virtualization**

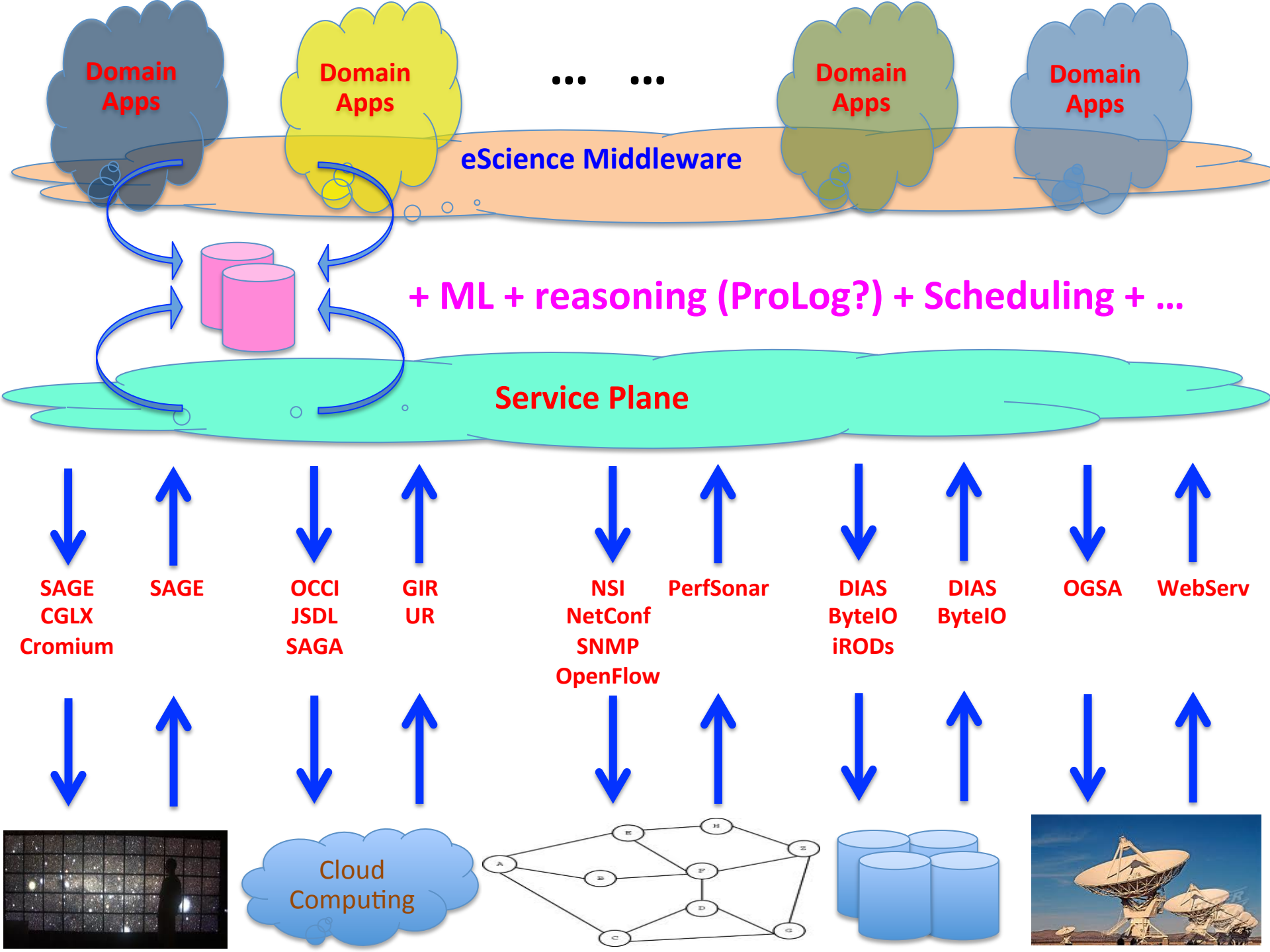
ref: Raj Jain, "Internet 3.0: Ten Problems with Current Internet Architecture and Solutions for the Next Generation", Military Communications Conference, 2006. MILCOM 2006. IEEE

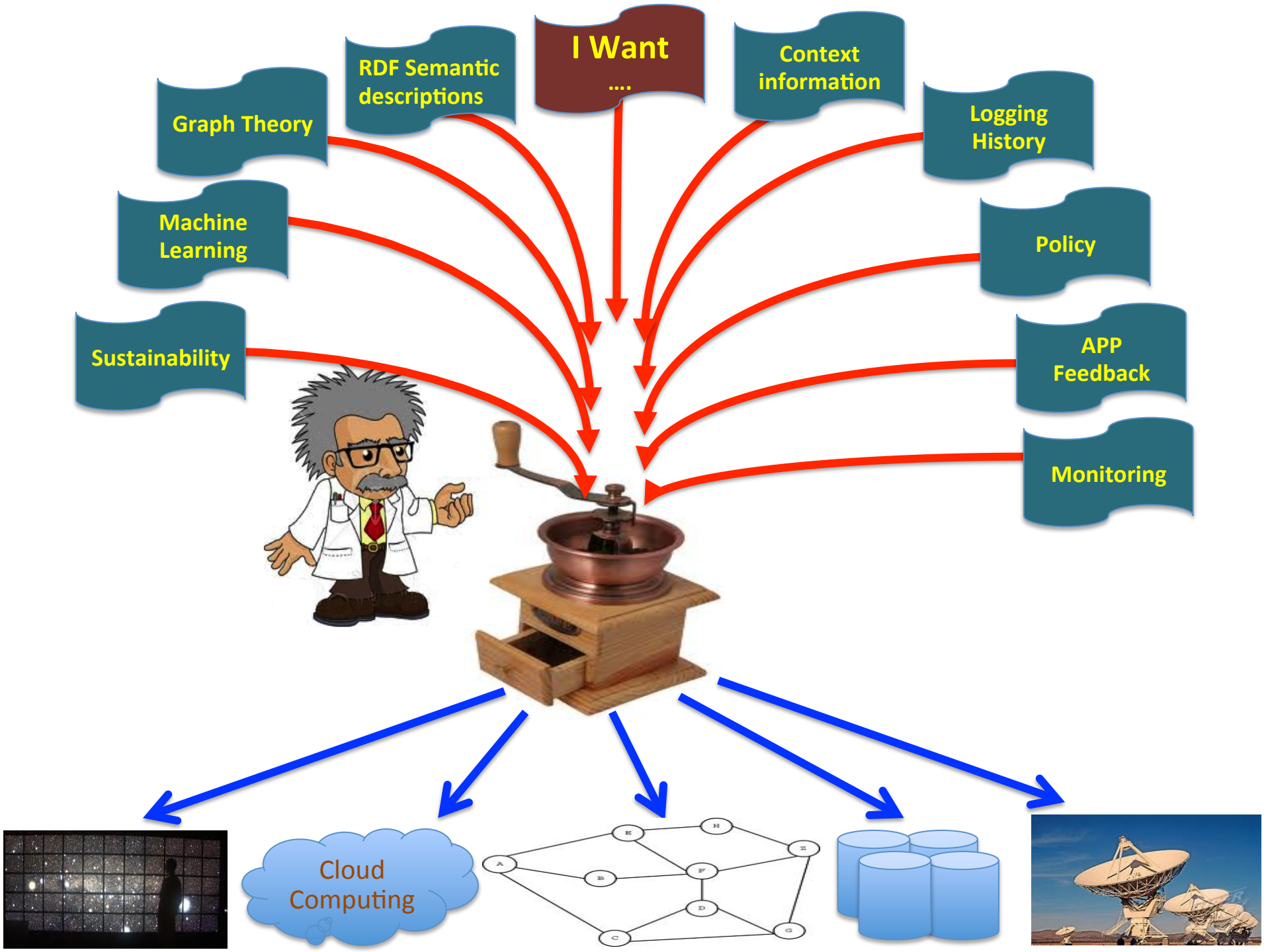
TimeLine


GreenIT&Nets
SF for Clouds
NDL SF for complex nets
Programmable Networks
NetApp's
CineGrid
SF for CineGrid
NM
OCCI
NSI
LightPaths - GLIF
Hybrid Nets
RDUDP, SCTCP, …
AAA
TBN
Policy
ATM
SONET/SDH
(G)MPLS
PBT/PLSB
OpenFlow
TCP
TCP Reno, Vegas
1980
1990
2000
2005
2011

TimeLine


GreenIT&Nets
Sustainable Internet
SF for Clouds
NDL SF for complex nets
Programmable Networks
NetApp's
Cognitive Nets and clouds
CineGrid
SF for CineGrid
Machine Learning
+
"I Want"
Internet 3.0
NM
OCCI
NSI
aths - GLIF
Hybrid Nets
P, SCTCP, …
Virtualized Internet
TBN
Policy
(G)MPLS
PBT/PLSB
OpenFlow
TCP
Good Old Trucking
2005
2011
2020


Domain Apps
Domain Apps
... ...
Domain Apps
Domain Apps
eScience Middleware
+ ML + reasoning (ProLog?) + Scheduling + …
Service Plane
SAGE
CGLX
Cromium
SAGE
OCCI
JSDL
SAGA
GIR
UR
NSI
NetConf
SNMP
OpenFlow
PerfSonar
DIAS
ByteIO
iRODs
DIAS
ByteIO
OGSA
WebServ
Cloud Computing


I Want
....
Graph Theory
RDF Semantic descriptions
Context information
Logging History
Machine Learning
Policy
Sustainability
APP Feedback
Monitoring
Cloud Computing

Challenges

- Data – Data – Data
  - Archiving, publication, searchable, transport, self-describing, DB innovations needed, multi disciplinary use
- Virtualisation
  - Another layer of indeterminism
- Greening the Infrastructure
  - e.g. Department Of Less Energy: http://www.ecrinitiative.org/pdfs/ECR_3_0_1.pdf
- Disruptive developments
  - BufferBloath, Revisiting TCP, influence of SSD's & GPU's
  - Multi layer Glif Open Exchange model
  - Invariants in LightPaths (been there done that ☺)
    - X25, ATM, SONET/SDH, Lambda's, MPLS-TE, VLAN's, PBT, OpenFlow, ….
  - Authorization & Trust & Security and Privacy

The Way Forward!

- Nowadays scientific computing and data is dwarfed by commercial & cloud, there is also no scientific water, scientific power.
  - Understand how to work with elastic clouds
  - Trust & Policy & Firewalling on VM/Cloud level
- Technology cycles are 3 – 5 year
  - Do not try to unify but prepare for diversity
  - Hybrid computing & networking
  - Compete on implementation & agree on interfaces and protocols
- Limitation on natural resources and disruptive events
  - Energy becomes big issue
  - Follow the sun
  - Avoid single points of failure (aka Amazon, Blackberry, …)
  - Better very loosly coupled than totally unified integrated…

Hybrid Networking <-> Computing

Routers ⇐ ⇒ Supercomputers

Ethernet switches ⇐ ⇒ Grid & Cloud

Photonic transport ⇐ ⇒ GPU's

What matters:

- Energy consumption/multiplication
- Energy consumption/bit transported

ECO-Scheduling

Q & A

I did not talk about:

- CineGrid, digital Cinema on CI
- Knowlegde complexity
- Security & privacy
- AAA
- …

http://ext.delaat.net/

Slides thanks to:

- Paola Grosso
- Sponsors see slide 1. ☺
- SNE Team & friends, see below