Internet Innovation to support Science & Education.

Cees de Laat

EU COMMIT UVA

NWO
PID/EFRO
SURFnet
NLESC
TNO



... more data!



Internet developments





... more realtime!







... more data! Speed DATA Volume You Tube twitter > Scalable myspace Linked in Secure more users!

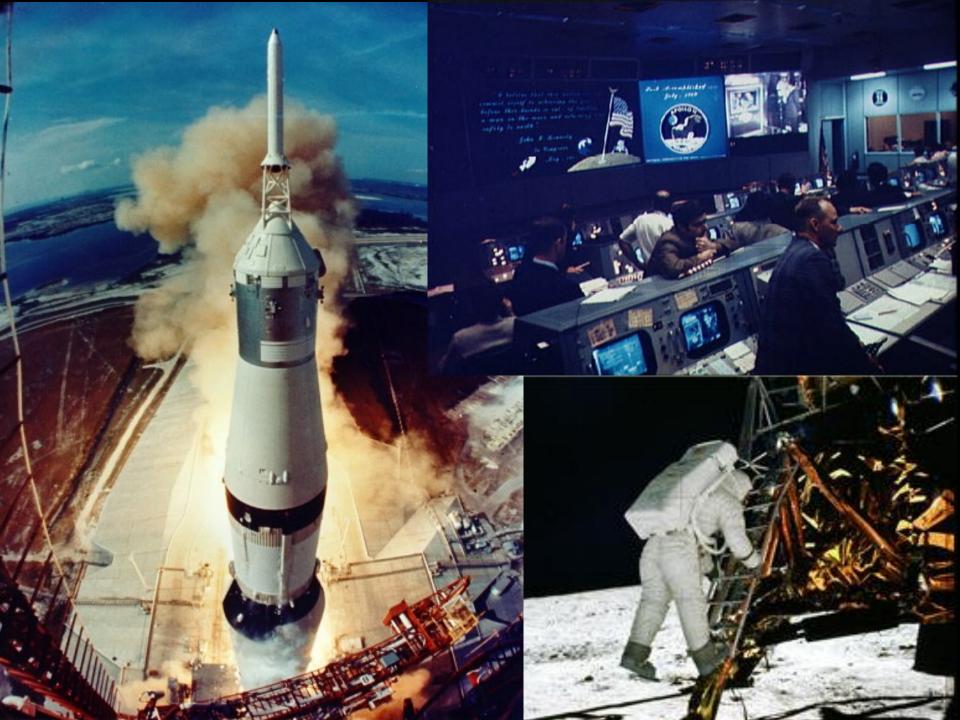
Internet developments



Real-timere realtime!



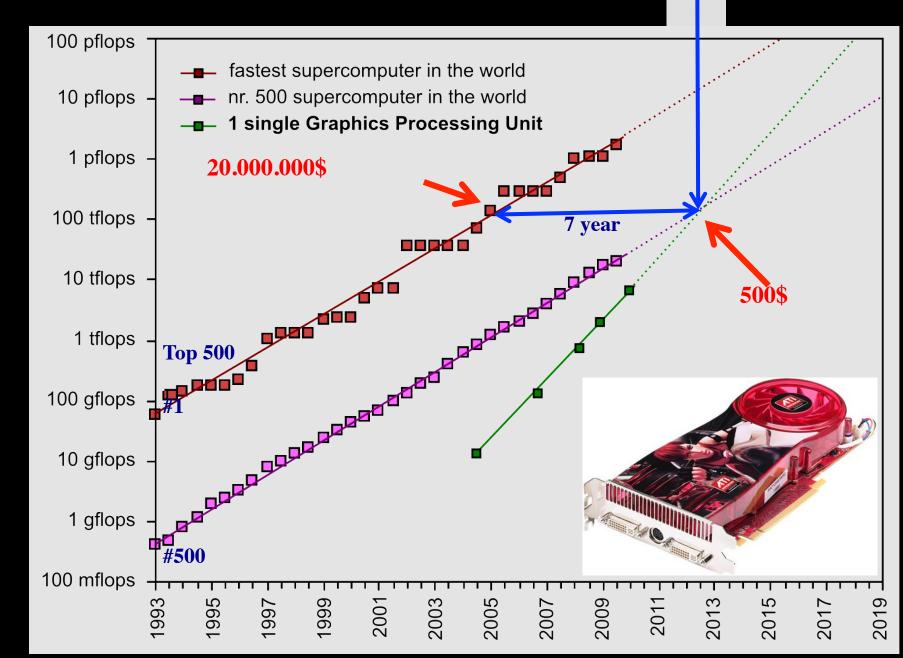




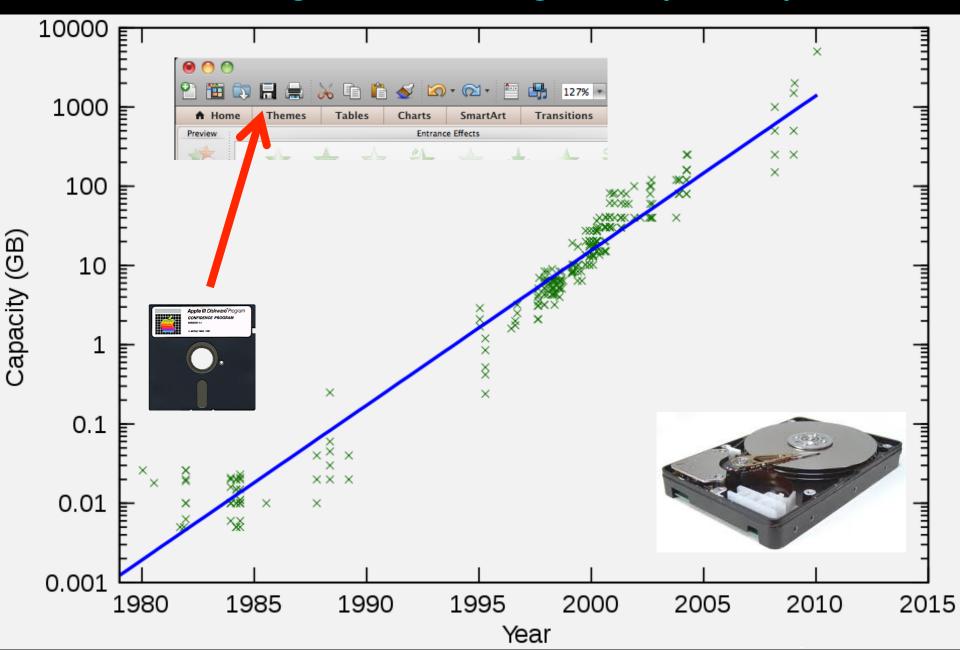


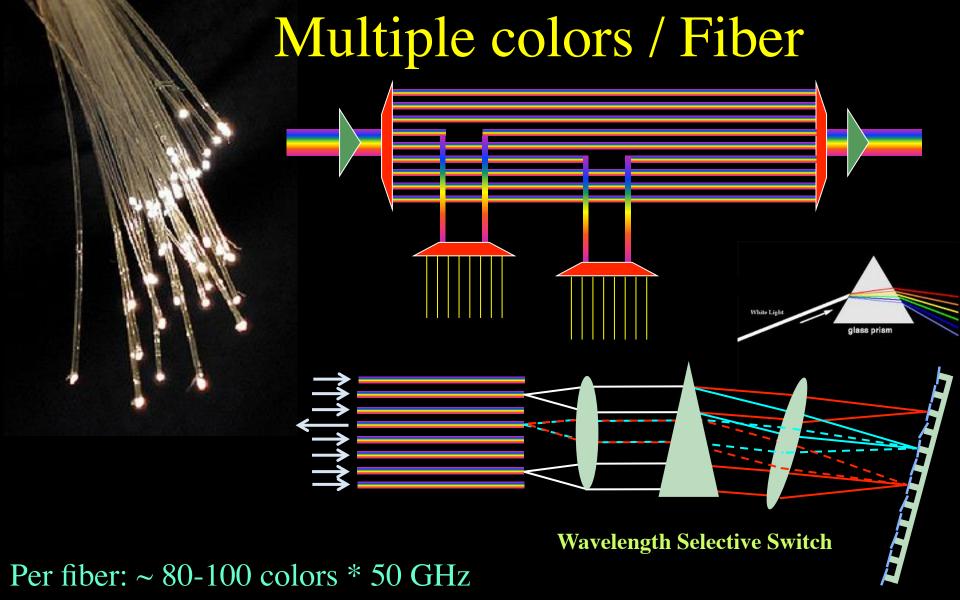


GPU cards are distruptive!



Data storage: doubling every 1.5 year!





Per color: 10 - 40 - 100 Gbit/s

BW * Distance $\sim 2*10^{17}$ 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



Next Generation Wireless LAN Technology 802.1 ac 1 Gbps throughput with

Published By admin under Network Devices Tags: 1gbps throughput, 1gbps wireless, 1gbps wireless tans, 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

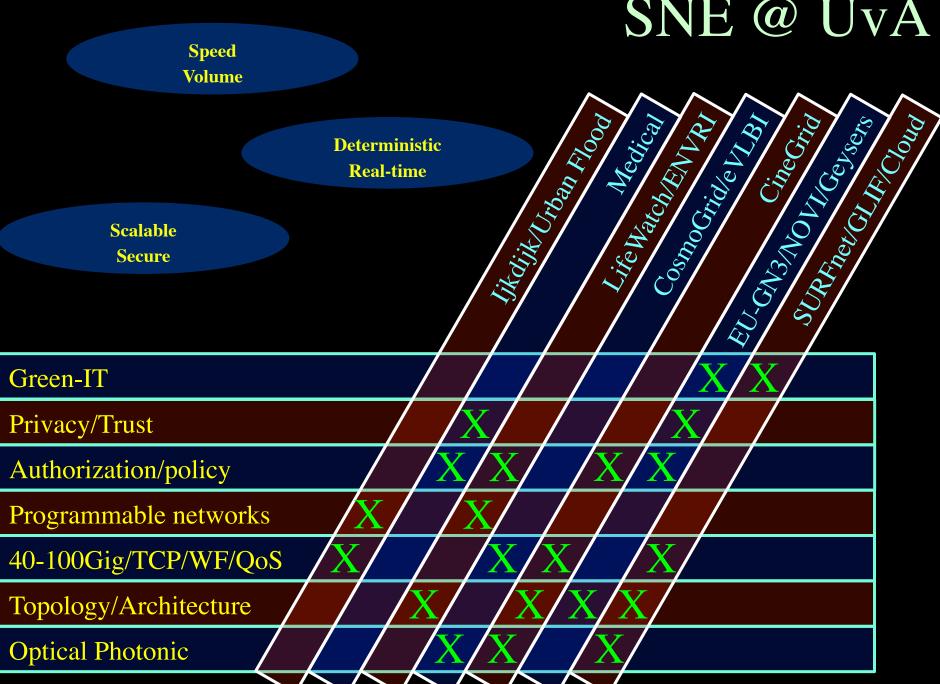




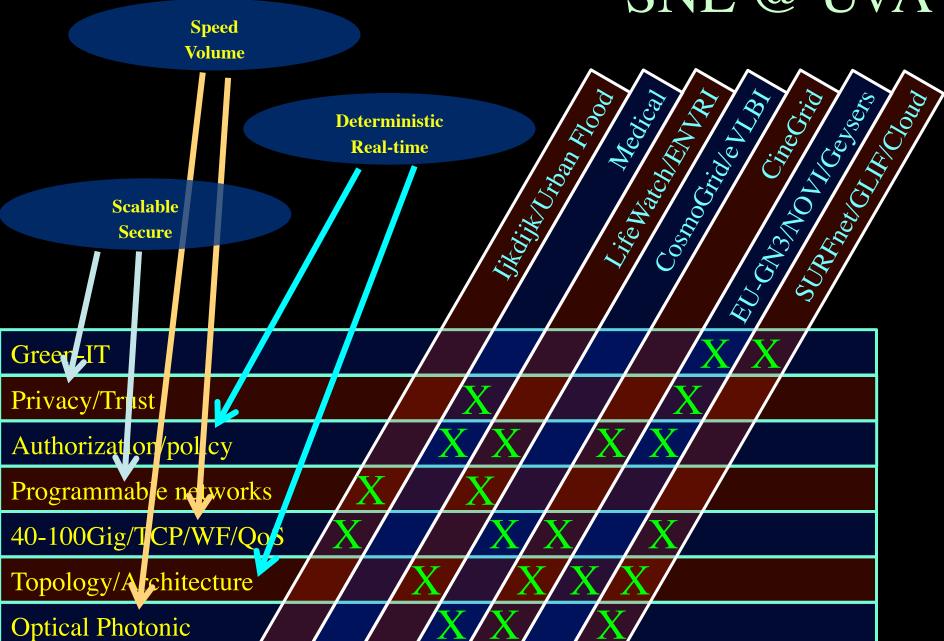
COPYRIGHT: MORTEN INGEMANN

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.

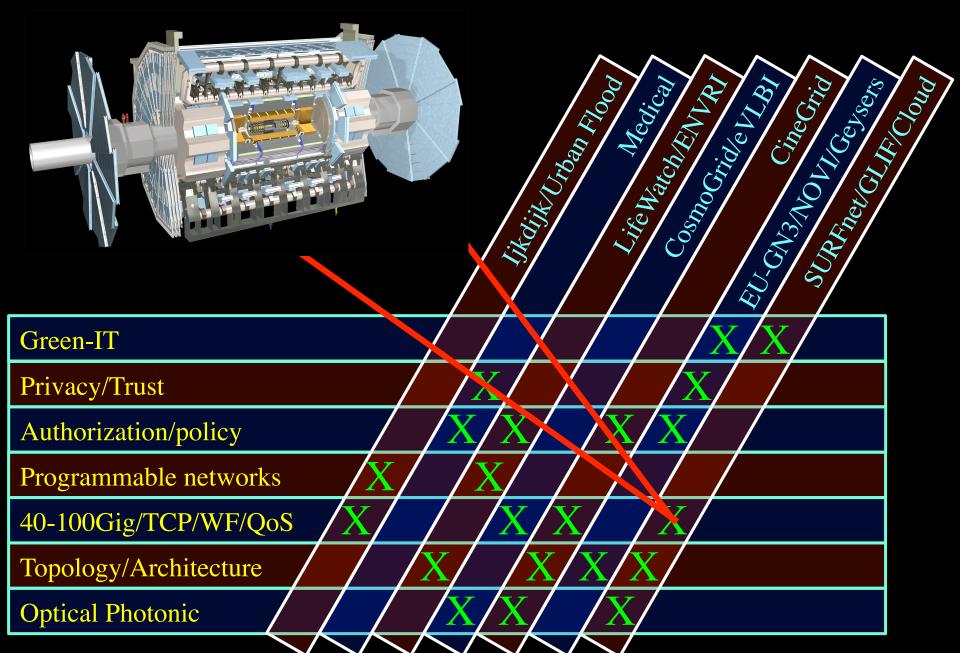
SNE @ UvA



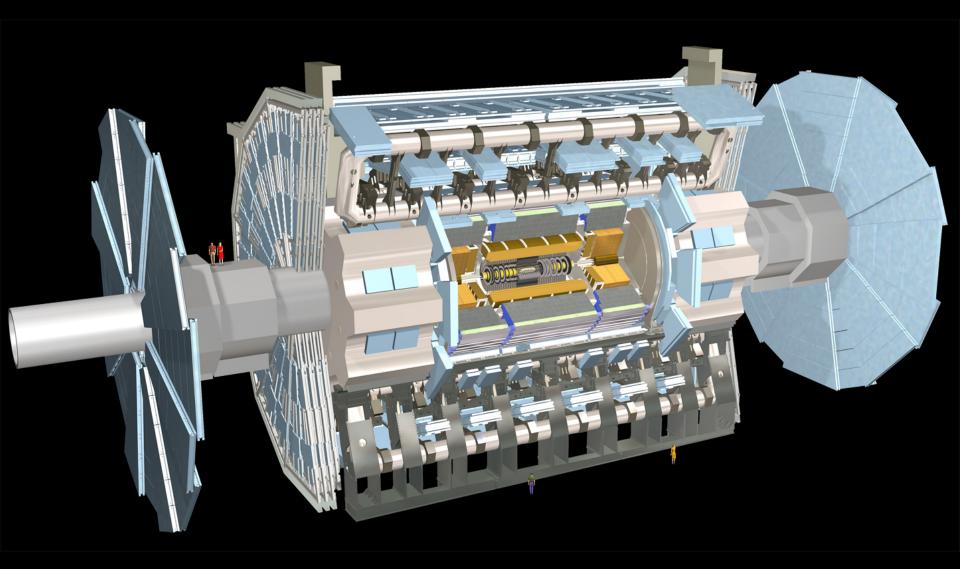
SNE @ UvA **Speed** Volume



SNE @ UvA



ATLAS detector @ CERN Geneve



ATLAS detector @ CERN Geneve





Physics data cache

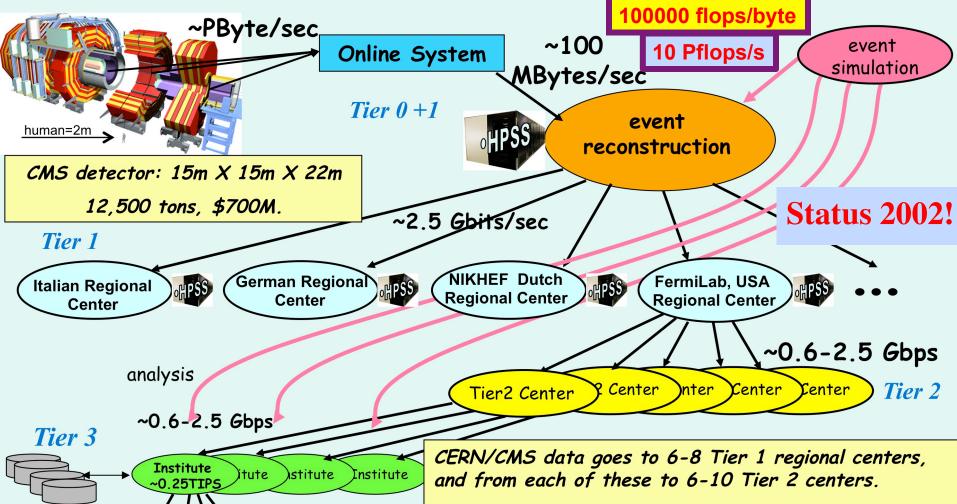
100 - 1000 Mbits/sec

Tier 4

LHC Data Grid Hierarchy

CMS as example, Atlas is similar





Courtesy Harvey Newman,
CalTech and CERN

Workstations

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

more channels.

Physicists work on analysis "channels" at 135 institutes.

Each institute has ~10 physicists working on one or

- 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 2011

 $\Sigma A = \Sigma B = \Sigma C \approx 1 \text{ Tb/s}$

However:

A -> all connects

B -> on several

C -> just a few (SP, LHC, LOFAR)

B

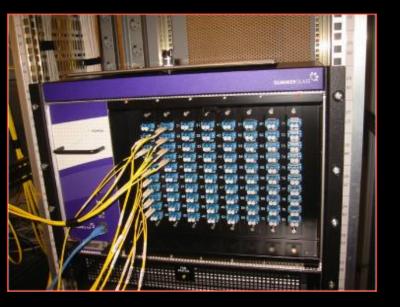
C

A

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 $A \rightarrow L3$, $B \rightarrow L2$, $C \rightarrow L1$ and L2
- Give each packet in the network the service it needs, but no more!

$L1 \approx 2-3 \text{ k}\text{/port}$



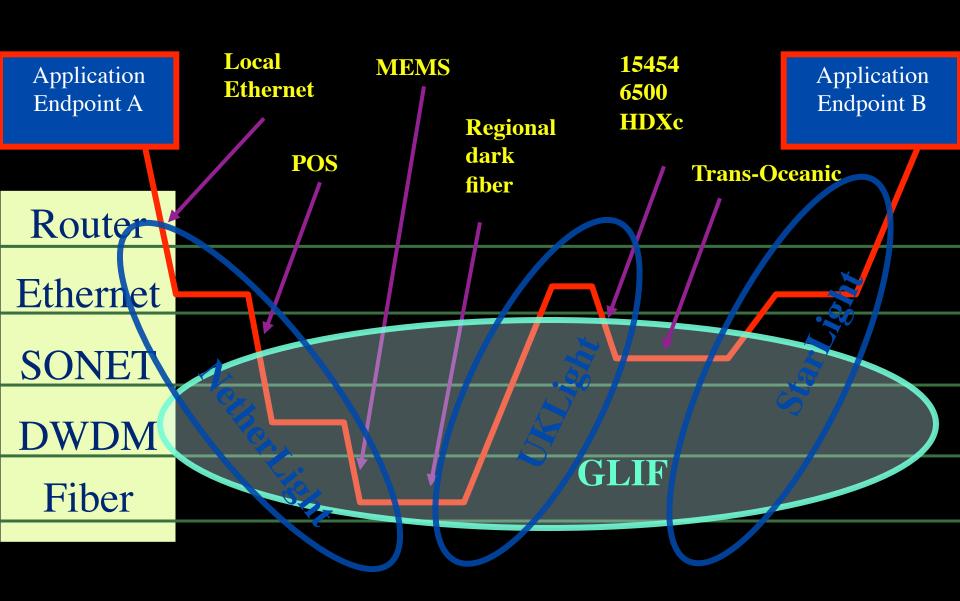
 $L2 \approx 5-8 \text{ k}/\text{port}$



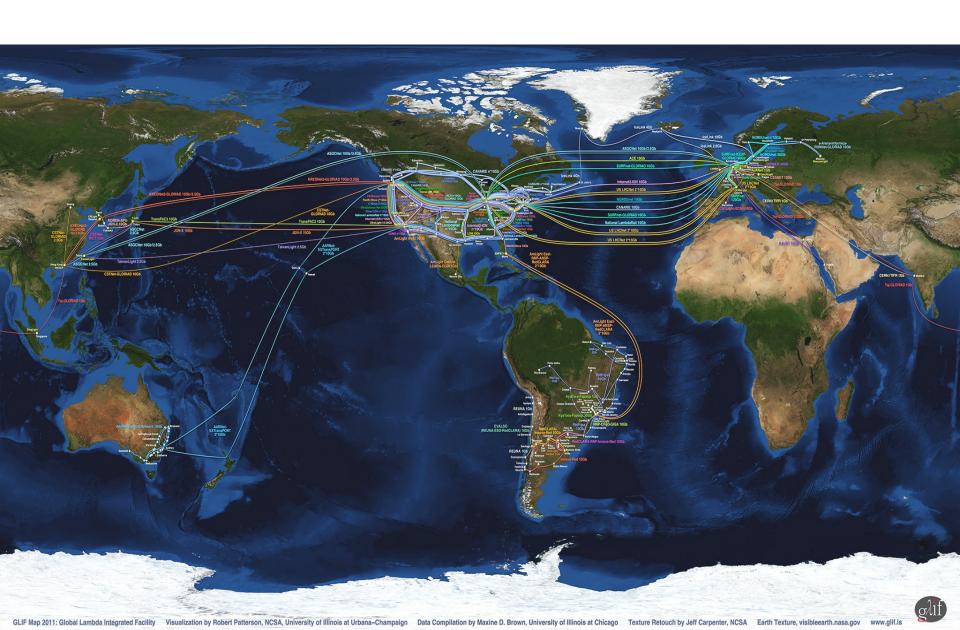
 $L3 \approx 75 + k\$/port$



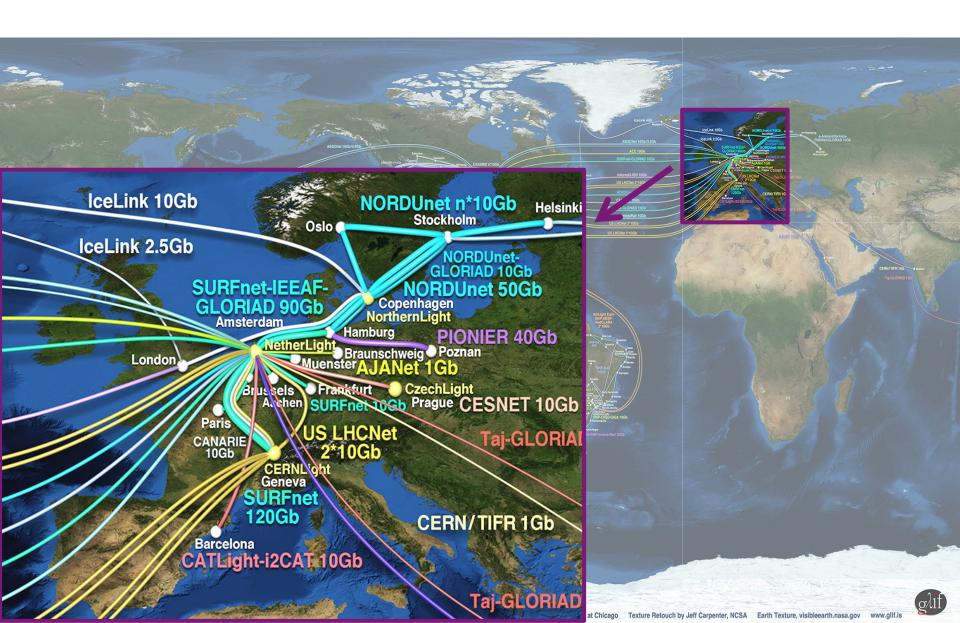
How low can you go?

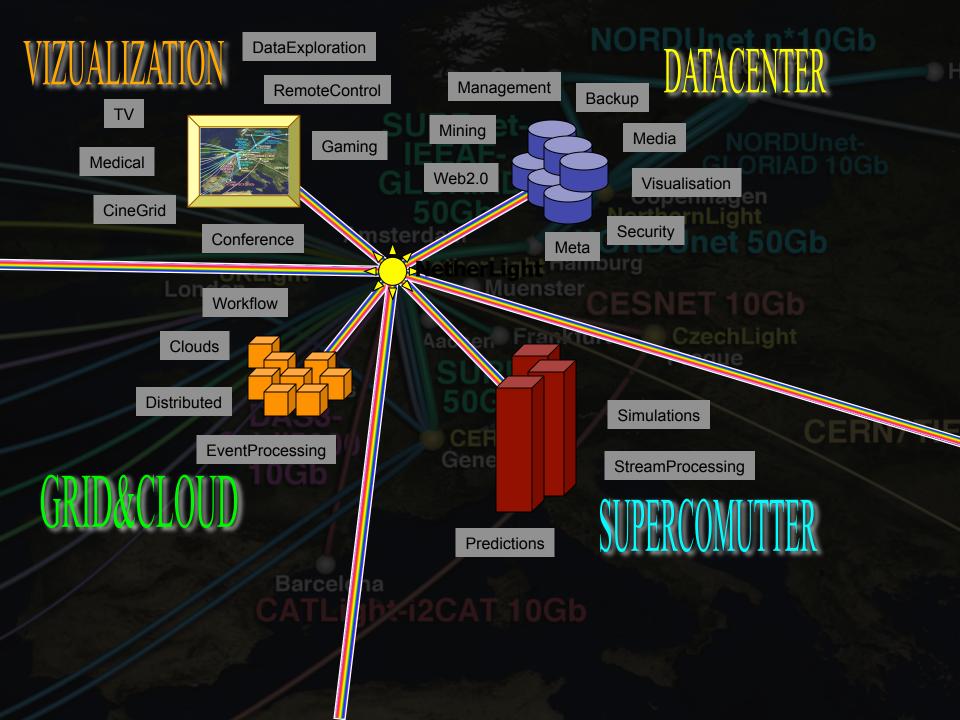


The GLIF – lightpaths around the world



The GLIF – lightpaths around the world







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



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

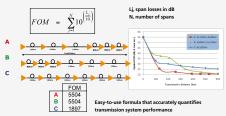
Alien wavelength challenges

- 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.

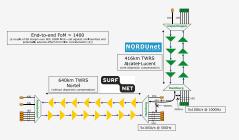
New method to present fiber link quality, FoM (Figure of Marit)

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.



Transmission system setup

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



Test results



Frror-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.



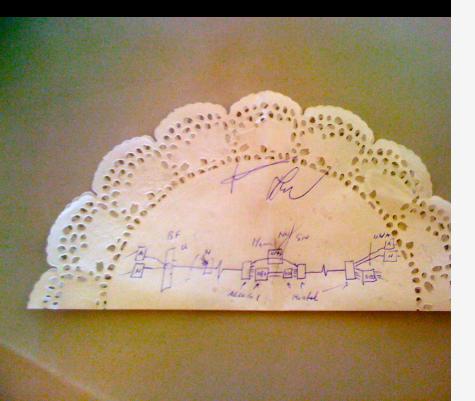






REFERENCES

Alien light From idea to realisation!



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



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

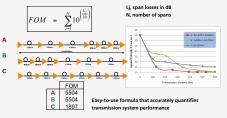
Alien wavelength challenges

- 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.

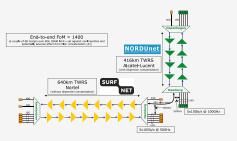
New method to present fiber link quality, FoM (Figure of Marit)

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.



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.









REFERENCES

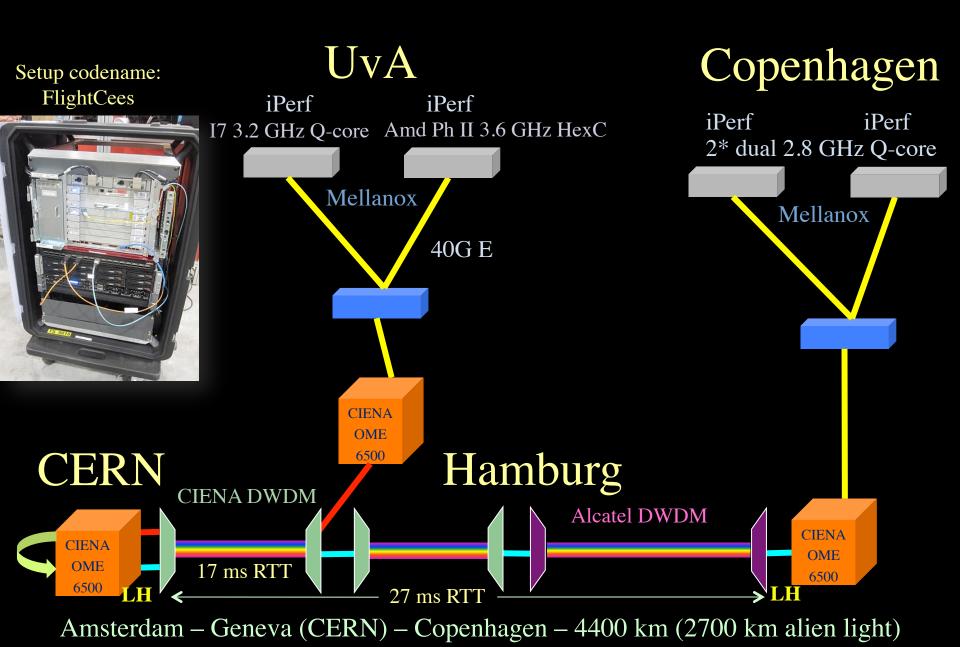
[1] OPERATIONAL SOLUTIONS FOR AN OPER DWIDM LATER 7.0. SERVICE FIRE, OF CA009 1 [2]. ATAIL OF FIGURE TRANSPORT SERVICES, SARRABERA SMIRIT, OF CO.

[3] OPERASIONAS OF ALL-OPTICAL CORE NETWORKS, "ADDRESS LOSS DAN CARL ENGINEER, ECC.2009 1 [4] NORTELISARIEST INTERNAL COMMUNICATION

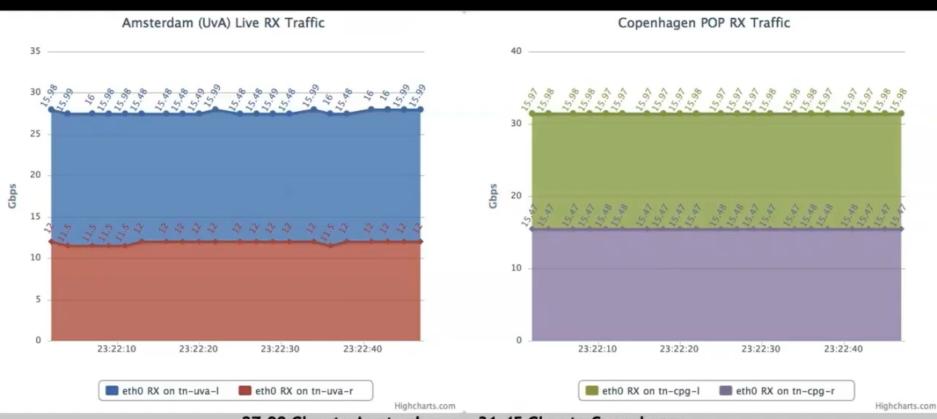
WE ARE GRATEFUL TO NOBDUNET FOR PROVIDING US WITH BANDWIDTH ON THEIR DWDM LINK FOR THIS EXPERIMENT AND ALSO FOR THEIR SUPPORT AND ASSISTANCE

DUBING THE EXPERIMENTS. WE ALSO ACKNOWLEDGE TELINIOUS AND NORTHER FOR THEIR THE FRANCH OWN KAND SMULTATION SHOW.

ClearStream @ TNC2011



http://tnc11.delaat.net



27.99 Gbps to Amsterdam <-> 31.45 Gbps to Copenhagen

Total Throughput 59.44 Gbps RTT 44.010 ms

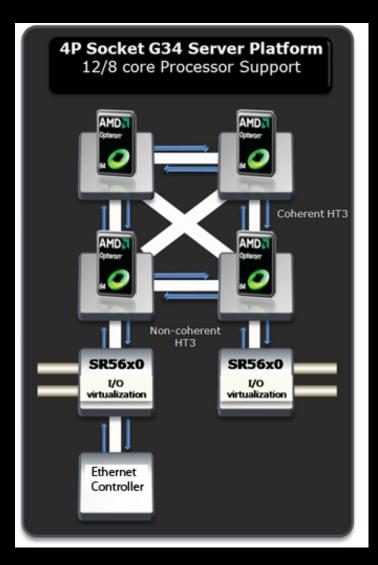
Results (rtt = 17 ms)

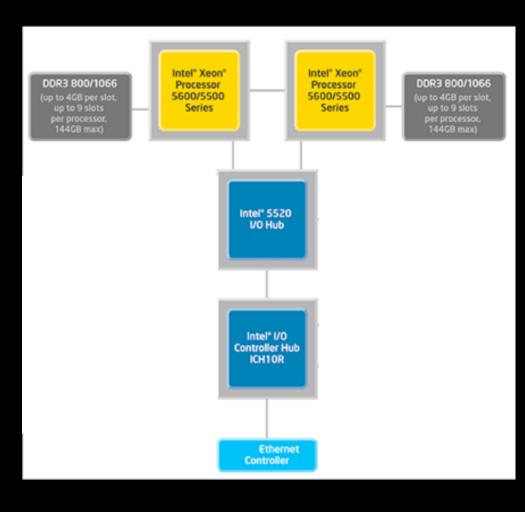
- Single flow iPerf 1 core -> 21 Gbps
- Single flow iPerf 1 core <> -> 15+15 Gbps
- Multi flow iPerf 2 cores -> 25 Gbps
- Multi flow iPerf 2 cores <> -> 23+23 Gbps
- DiViNe \Leftrightarrow -> 11 Gbps
- Multi flow iPerf + DiVine -> 35 Gbps
- Multi flow iPerf + DiVine <> -> 35 + 35 Gbps

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

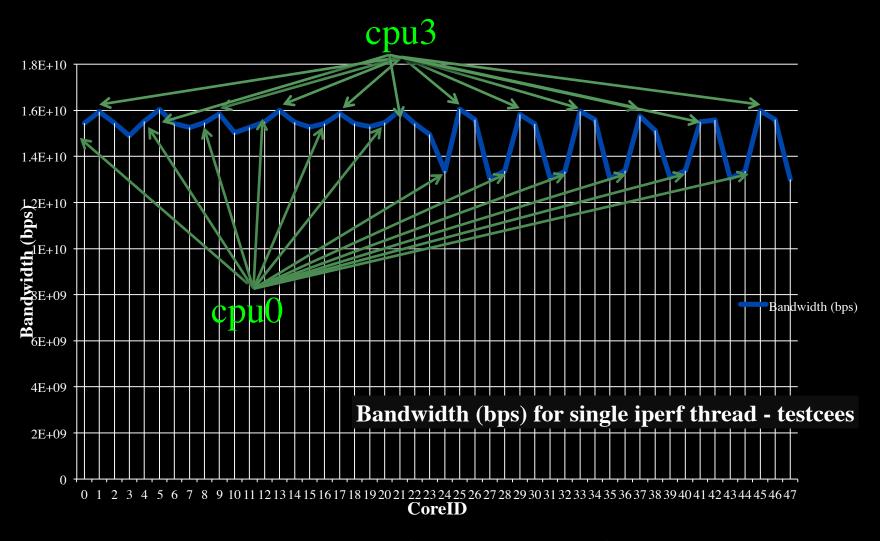
Server Architecture





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





IJKDIJK

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

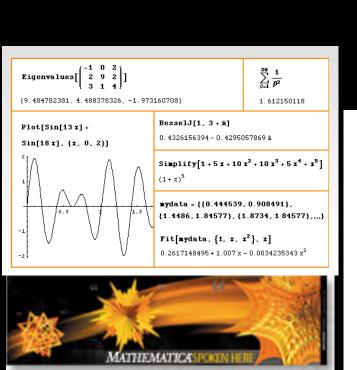
Sensor grid: instrument the dikes

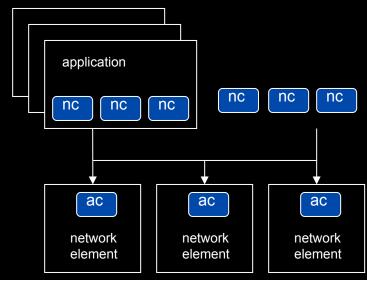
First controlled breach occurred on sept 27th '08:

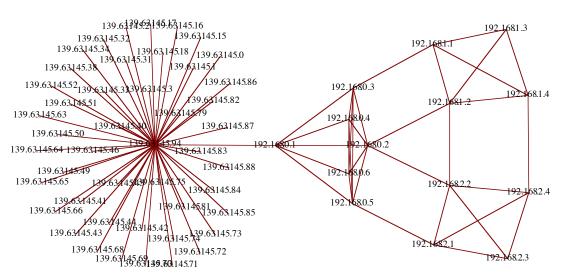


User Programmable Virtualized Networks.

- 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

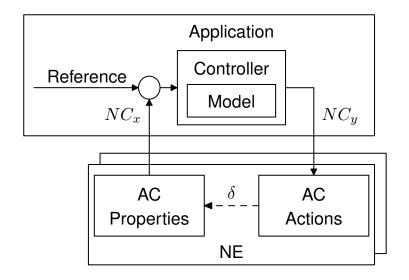


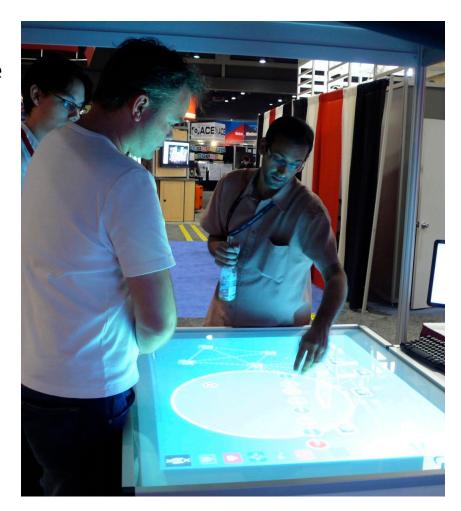




In the Intercloud virtual servers and networks become software

- 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

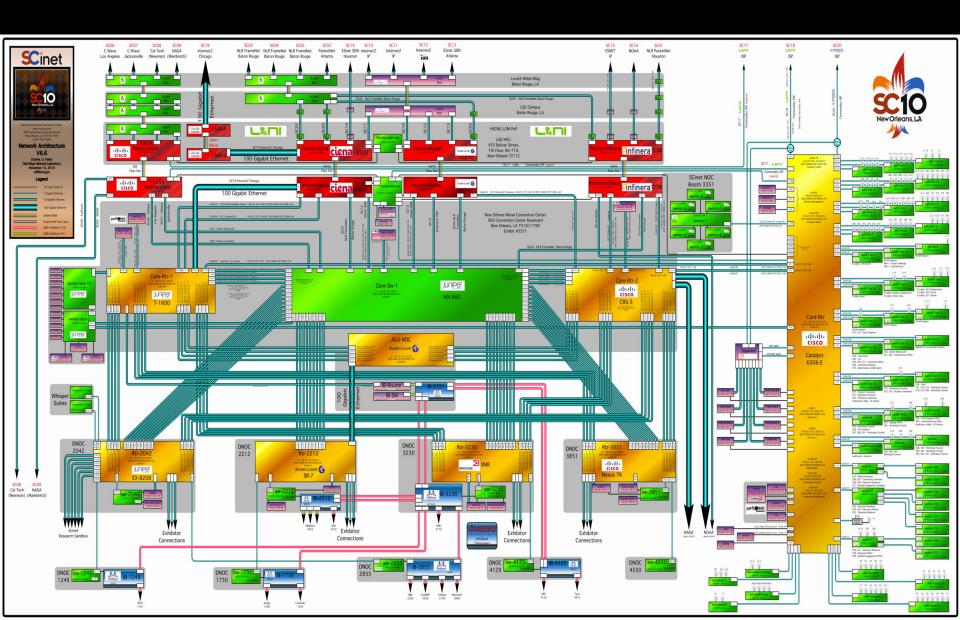
Cees de Laat 1

Robert Meijer1,2

- ¹ University of Amsterdam, Amsterdam The Netherlands
- ² TNO Information and Communication Technology, Groningen, The Netherlands

Mastering Complexity SNE @ UvA EU-GN3/NOVII/GUSGrs SURPENCEL AFT COMA Jikajik Urban Filood Life Watch LEW PRI CosmoGride VI BI Medical Green-IT Privacy/Trust Authorization/policy Programmable networks 40-100Gig/TCP/WF/QoS Topology/Architecture Optical Photonic

Complex eInfrastructure @ SC10





for

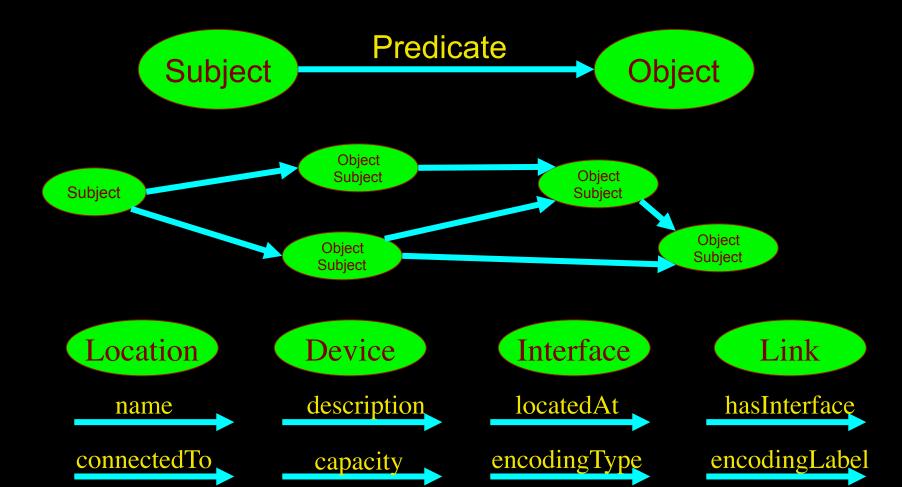
We investigate: complex networks!



LinkedIN for Infrastructure



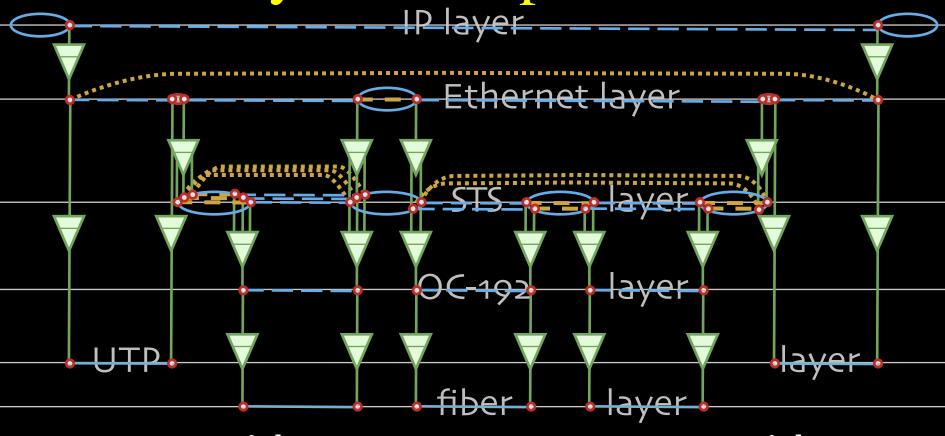
- From semantic Web / Resource Description Framework.
- The RDF uses XML as an interchange syntax.
- Data is described by triplets (Friend of a Friend):



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:resourd<!-- all the interfaces of TDM3.amsterdam1.netherlight.net -->
    <ndl:hasInterface rdf:resource
    <ndl:hasInterface rdf:resourd<ndl:Interface rdf:about="#tdm3.amsterdam1.netherlight.net:501/1">
    <ndl:hasInterface rdf:resource
                                            <ndl:name>tdm3.amsterdam1.netherlight.net:POS501/1</ndl:name>
    <ndl:hasInterface rdf:resource
                                            <ndl:connectedTo rdf:resource="#tdm4.amsterdam1.netherlight.net:5/1"/>
    <ndl:hasInterface rdf:resourd </ndl:Interface>
    <ndl:hasInterface rdf:resourd<ndl:Interface rdf:about="#tdm3.amsterdam1.netherlight.net:501/2">
    <ndl:hasInterface rdf:resource
                                            <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



End host

Université du Quebec SONET switch with Ethernet intf.



Ethernet & SONET switch



SONET switch

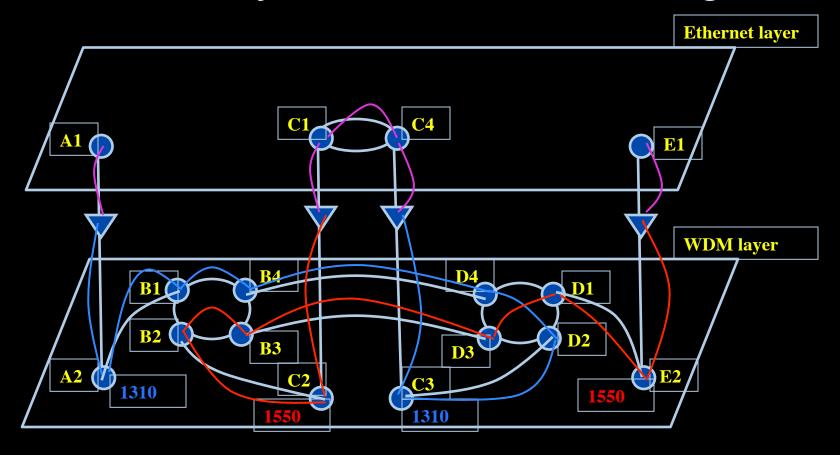


SONET switch with Ethernet intf.

NetherLight Amsterdam End host



Multi-layer Network PathFinding

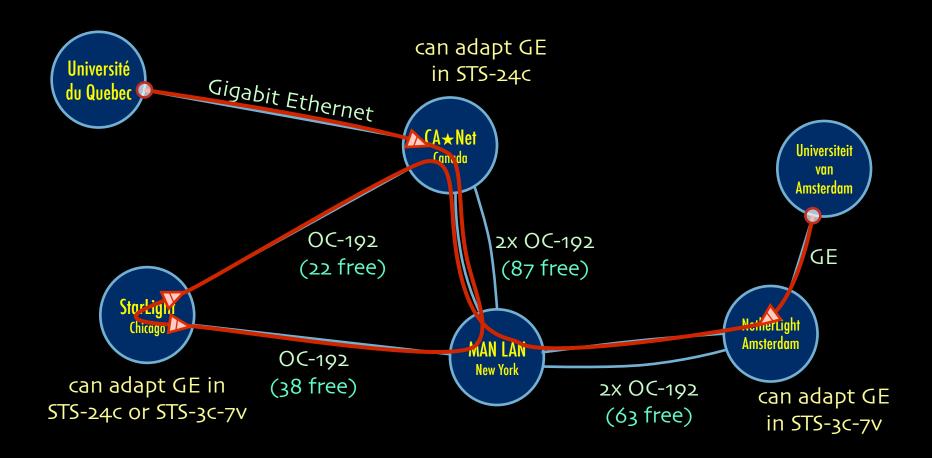


Path between interfaces A1 and E1:

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

Scaling: Combinatorial problem

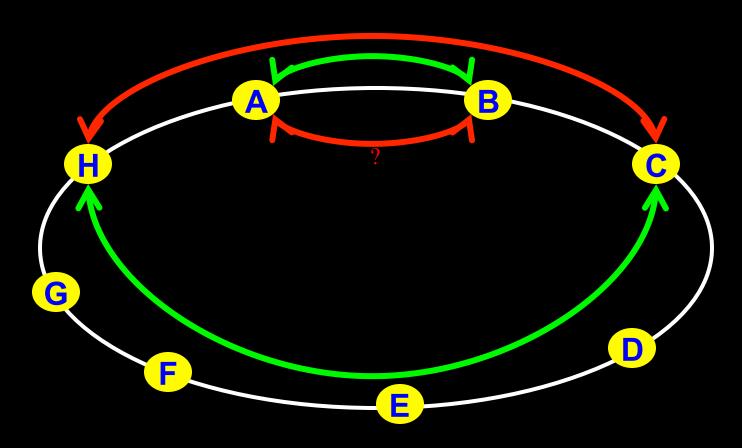
A weird example



The Problem

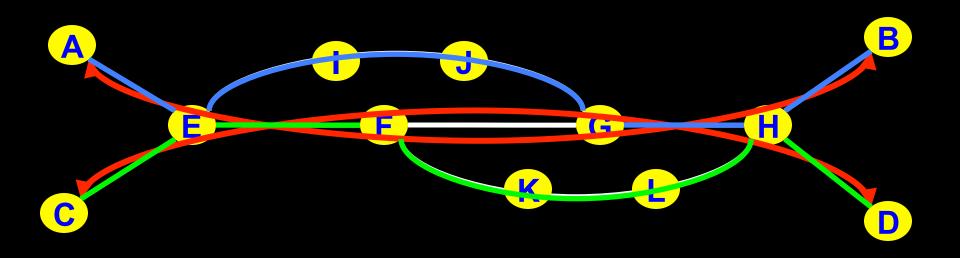
I want HC and AB Success depends on the order

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



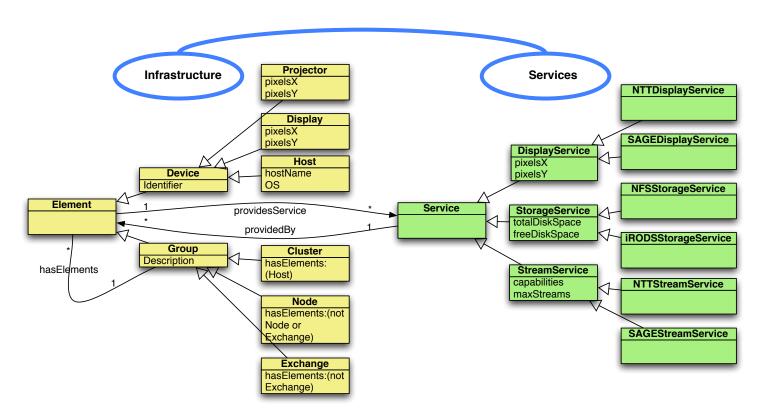
Another one ©

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.

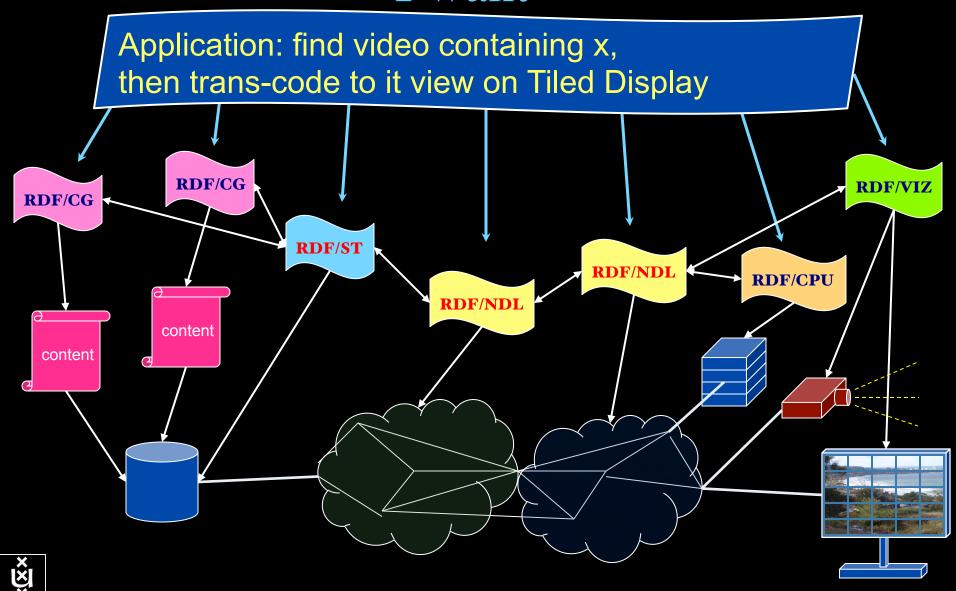


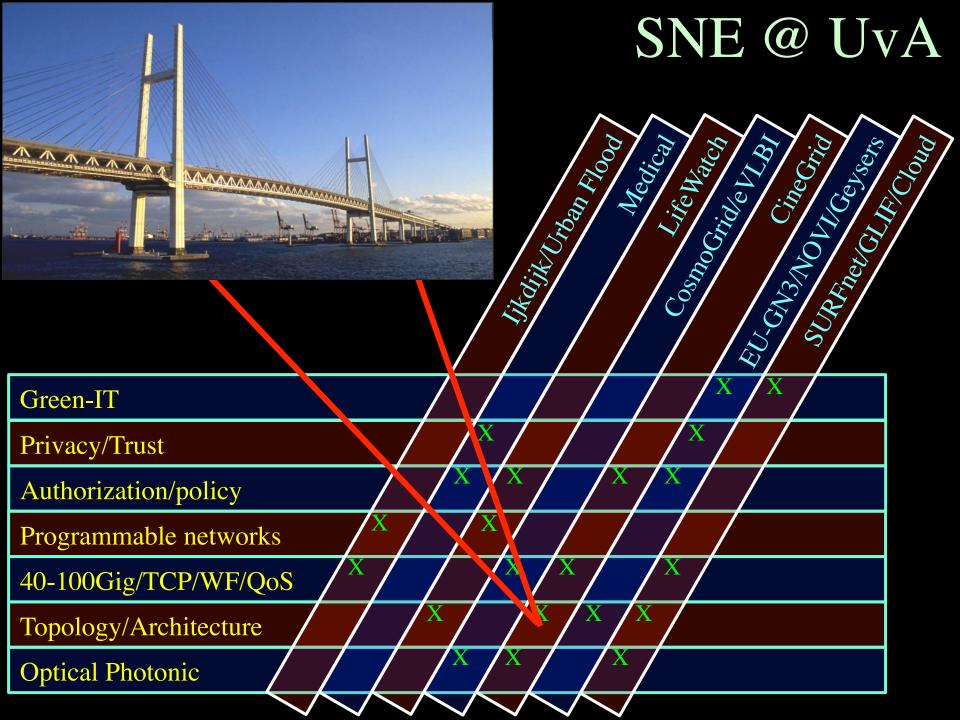
J. van der Ham, F. Dijkstra, P. Grosso, R. van der Pol, A. Toonk, C. de Laat *A distributed topology information system for optical networks based on the semantic web*,

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)

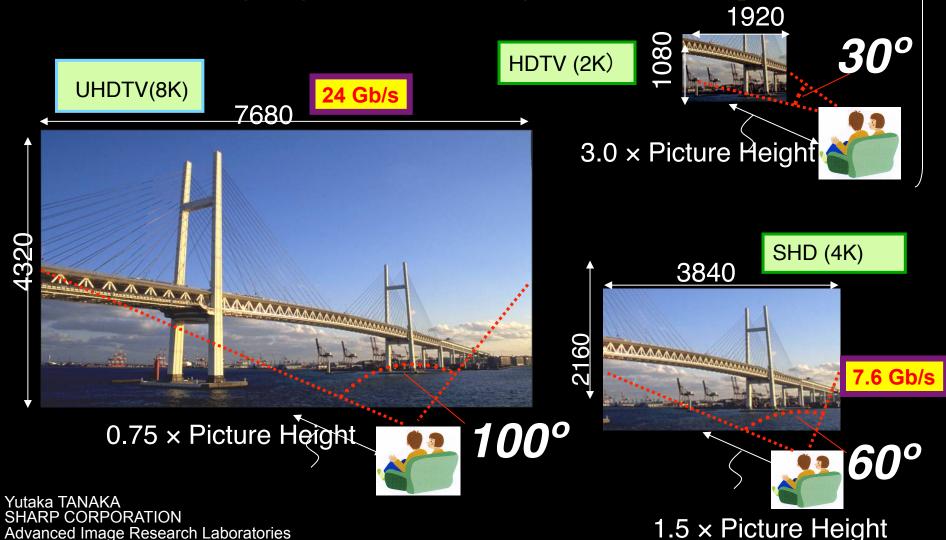
RDF describing Infrastructure "I want"





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









US and International OptIPortal Sites

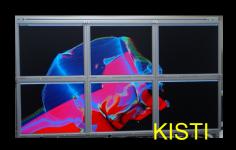






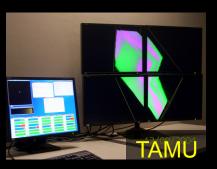


















The "Dead Cat" demo

1 Mflops/byte



SC2004, Pittsburgh, Nov. 6 to 12, 2004 iGrid2005, San Diego, sept. 2005

Many thanks to:
AMC
SARA
GigaPort
UVA/AIR
Silicon Graphics,
Inc.
Zoölogisch Museum



M. Scarpa, R.G. Belleman, P.M.A. Sloot and C.T.A.M. de Laat, "Highly Interactive Distributed Visualization", iGrid2005 special issue, Future Generation Computer Systems, volume 22 issue 8, pp. 896-900 (2006).







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

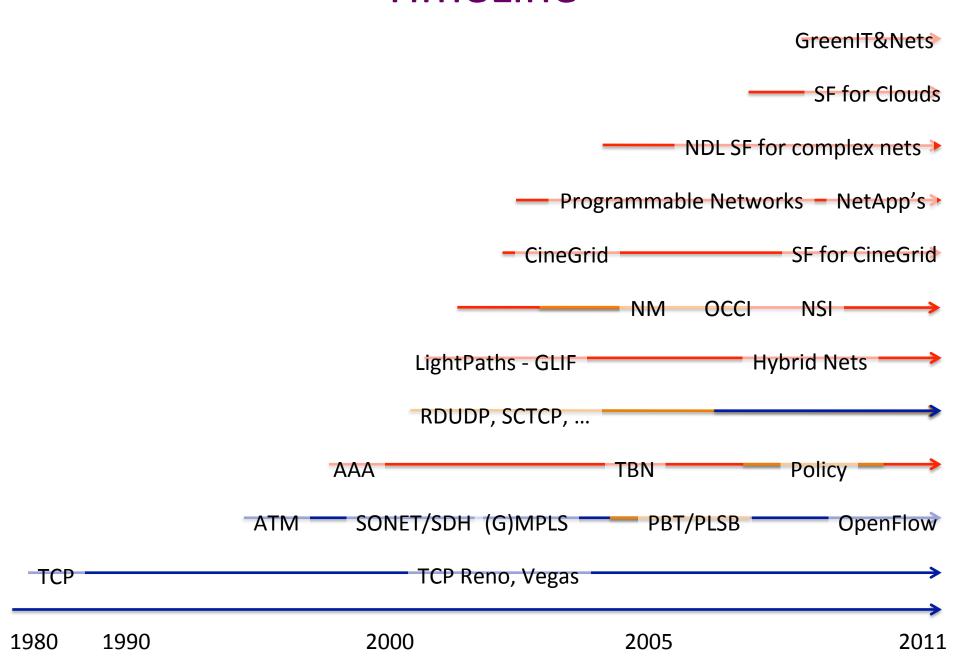
- 1. Energy Efficient Communication
- 2. Separation of Identity and Address
- 3. Location Awareness
- 4. Explicit Support for Client-Server Traffic and Distributed Services
- 5. Person-to-Person Communication
- 6. Security
- 7. Control, Management, and Data Plane separation
- 8. Isolation
- 9. Symmetric/Asymmetric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

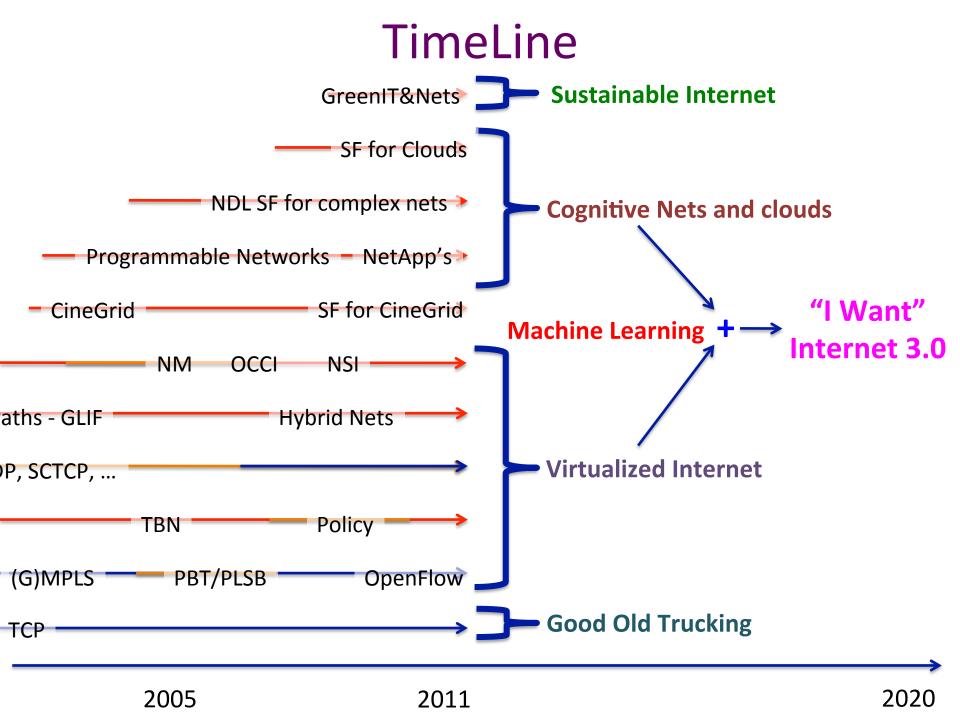
Nice to have:

- 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

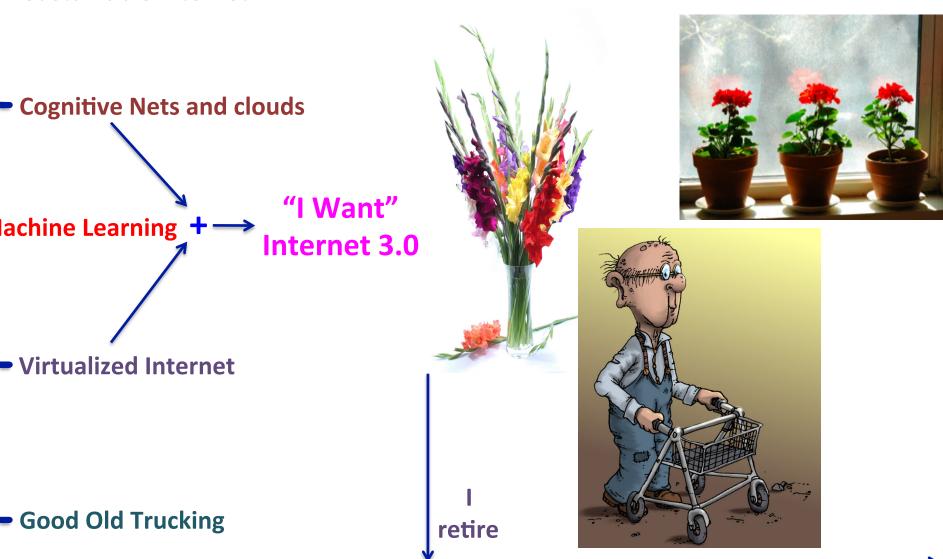
TimeLine



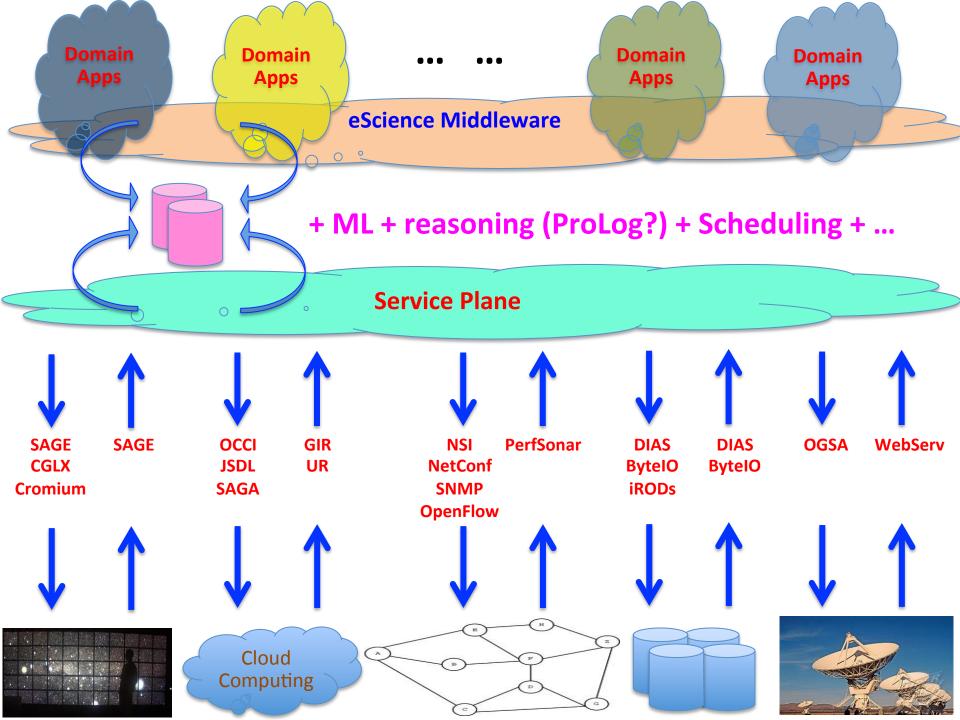


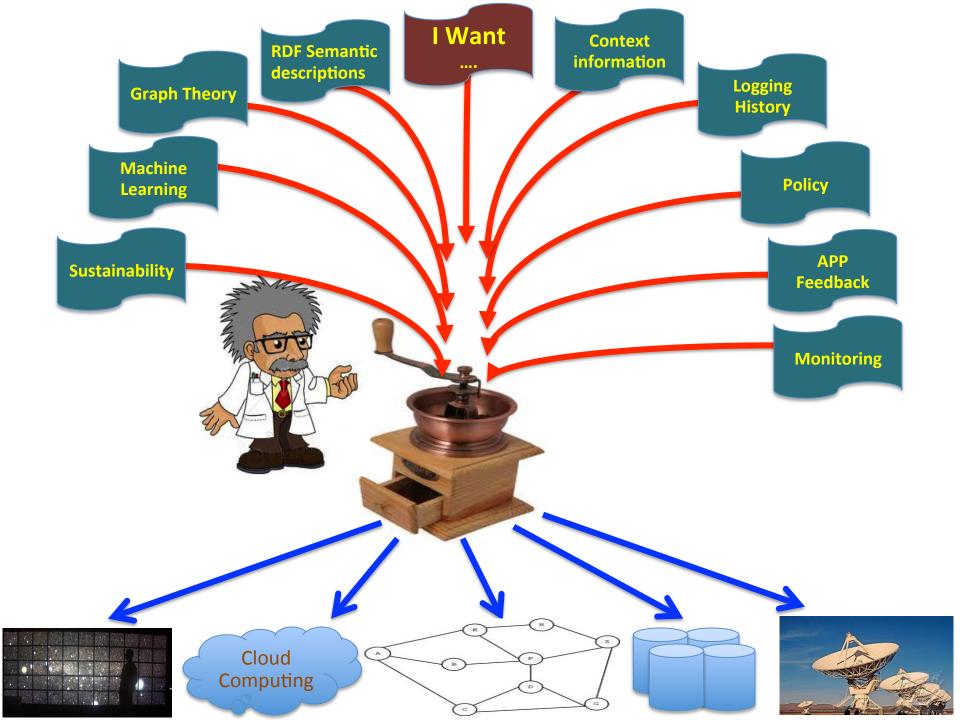
TimeLine

Sustainable Internet



2020 2040





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...

ECO-Scheduling



Questions?

CookReport feb 2009 and feb-mar 2010

november '08
interview with
Kees Neggers (SURFnet),
Cees de Laat (UvA)

and furthermore on november '09

Wim Liebrandt (SURF), Bob Hertzberger (UvA) and Hans Dijkman (UvA)

BSIK projects
GigaPort &
VL-e / e-Science





The COOK Report

On Internet Protocol



I did not talk about:

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

http://ext.delaat.net/

Slides thanks to:

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

