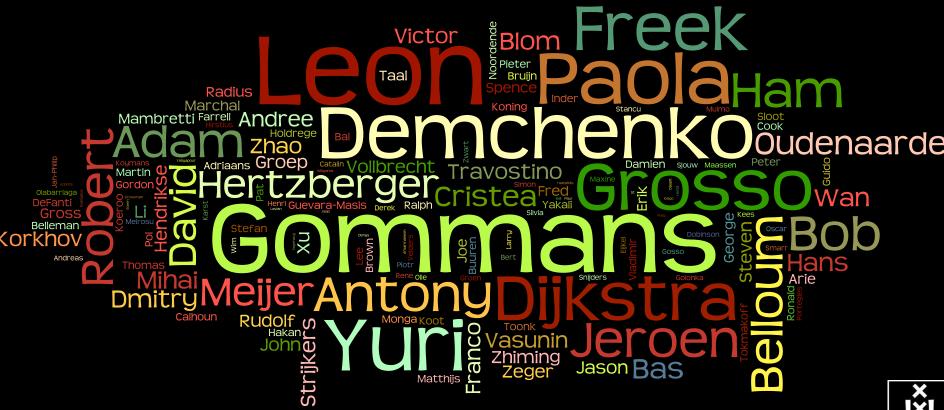
Supporting e-Science

Cees de Laat



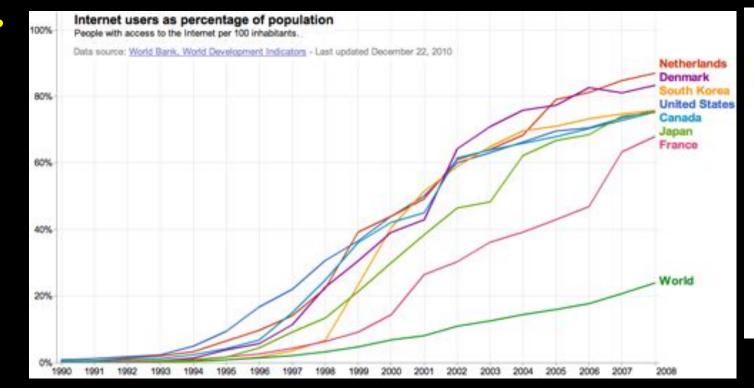


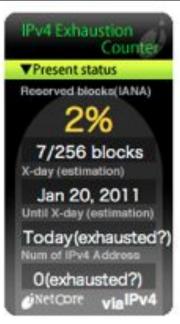
Internet From a network experiment that never ended (Vint Cerf)

- 1974: for the first time the word **internet** (*RFC 675 Specification of Internet Transmission Control Program*) [note -> Open process!]
- 1981: the **TCP/IP** standard was ready to be adopted (*RFC* 791,792,793)

To a network for society

• 1989: WWW was born





June 8th @ UvA

• Jan 2011 → IANA IPv4 address space depleted! →

Ipv6day.nl

Internet is a Billion - Business!

Google	197	guardian.co.uk Monday 3 January 2011
Amazon	83	News Sport Comment Culture Business Money Life & style
Facebook	50	News Technology Facebook Facebook's value swells to \$50bn after
BAIDU	37	Goldman Sachs investment Deal underlines Facebook's power and fuels rumours that Mark
eBay	36	Zuckerberg is preparing a stock market flotation
Yahoo	22	e.g.: Exxon Mobil 368
PriceLine	21	Apple Inc. 333
SalesForce	18	

F5 Networks

CheckPoint

NetFlix

Expedia



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

Internet developments



Real-timere realtime!



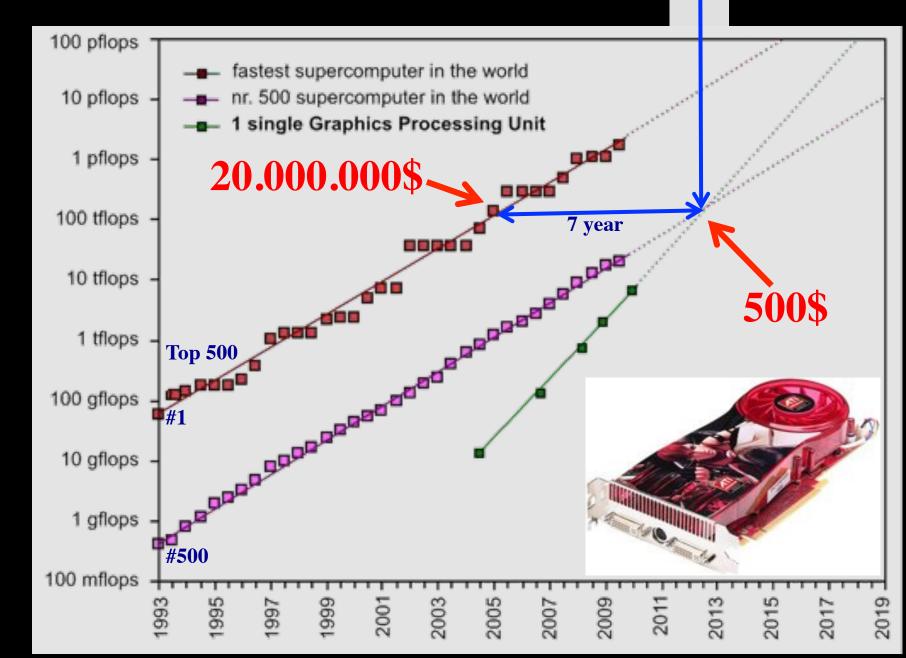




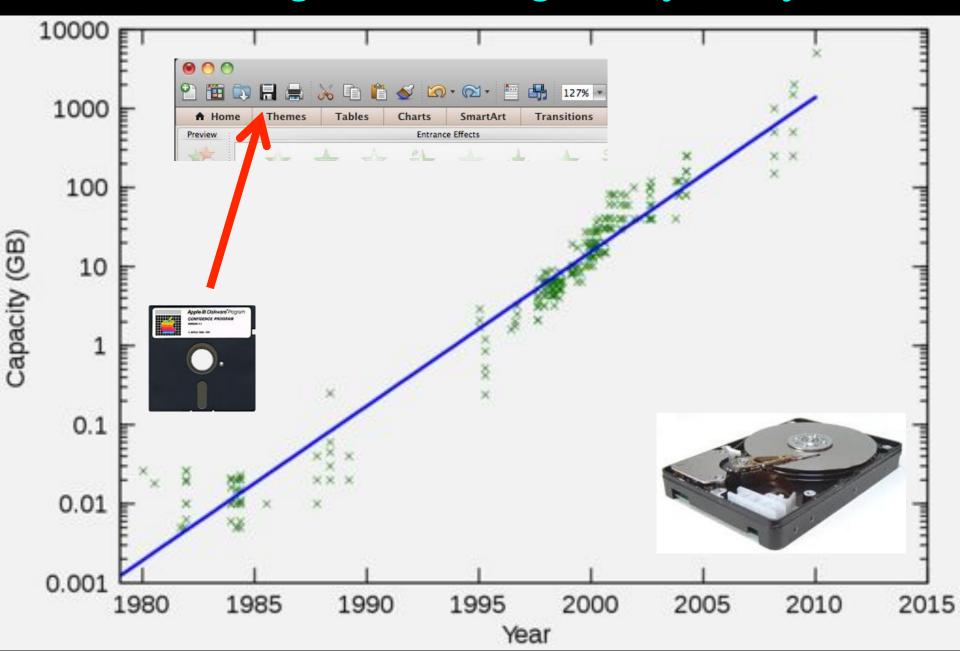




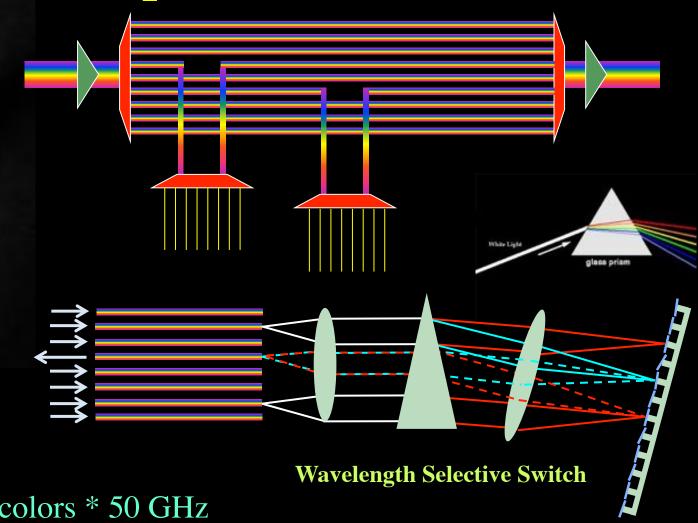
GPU cards are distruptive!



Data storage: doubling every 1.5 year!



Multiple colors / Fiber



Per fiber: ~ 80-100 colors * 50 GHz

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 s

HOME

CONTACT US

PRIVACY POLICY

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



Next Generation Wireless LAN Technology 802.11ac 1 Gbps throughput with

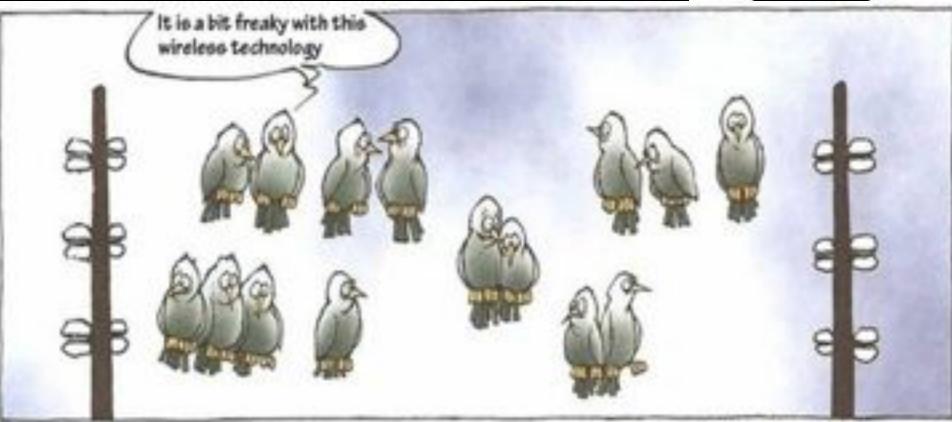
Published By admin under Network Devices Tags: 1gbps throughput, 1gbps wireless, 1gbps wireless tans, generation, new generation, technologies, technology, throughput, wireless, wireless land

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

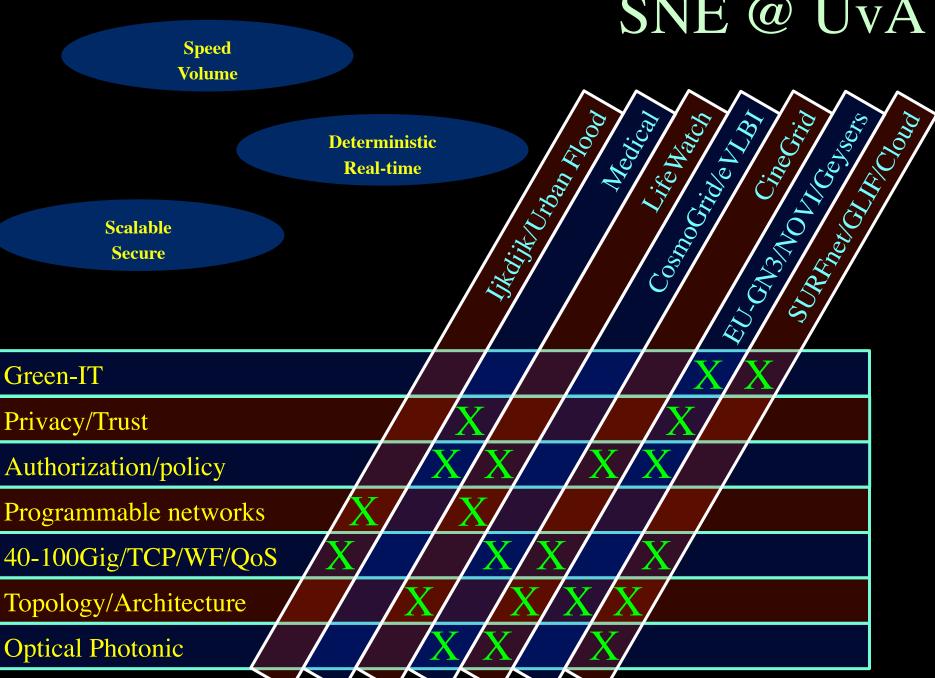




COPYRIGHT: MORTEN MICEMANN

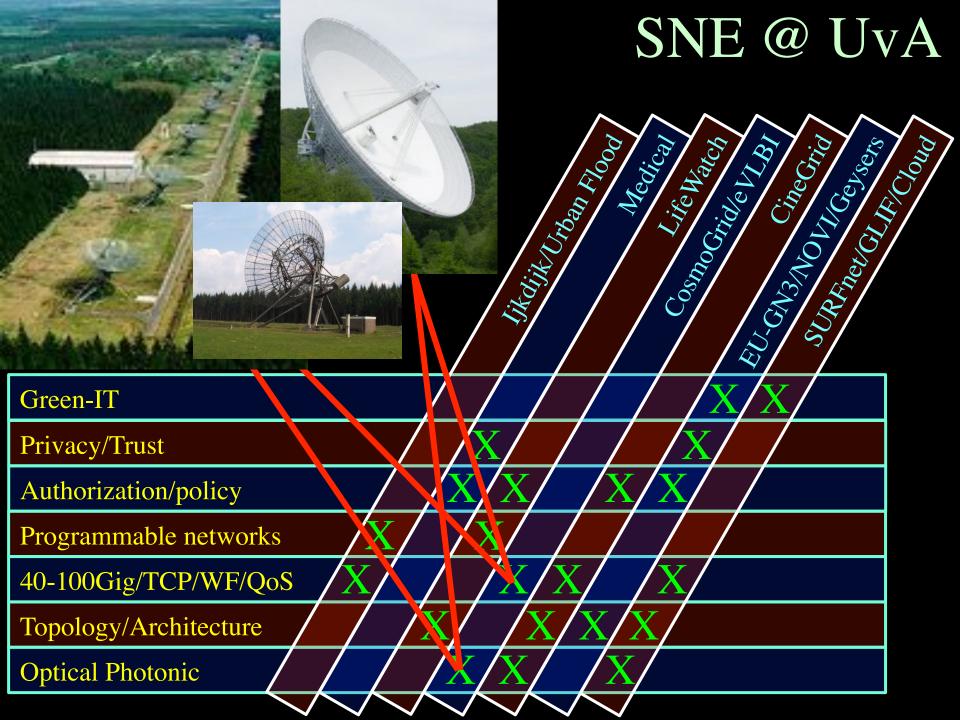
protocol LAN due to the easy comparison and convenience in the digital home. While consumer PC products has just started to migrate to a much higher bandwidth of 802.11n wireless LAN now working on next-generation standard definition is already in progress.

SNE @ UvA



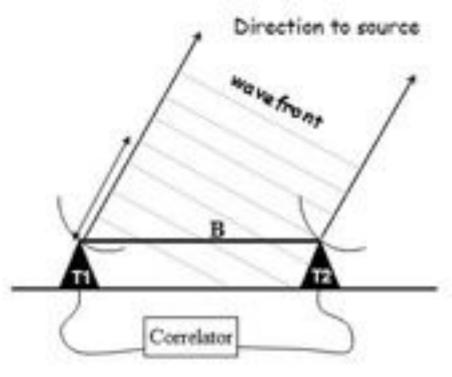
SNE @ UvA **Speed** Volume EU-GN3/NOVII/GUSGrs SURP WOUGLIF! Cloud Jikdijik Urban Filood Cosmo Gride VI BI Life Watch Medical **Deterministic Real-time Scalable** Secure Green-IT Privacy/Trust Authorizat or pol Programmab e networks 40-100Gig/TCP/WF/Qo Topology/Architecture

Optical Photonic

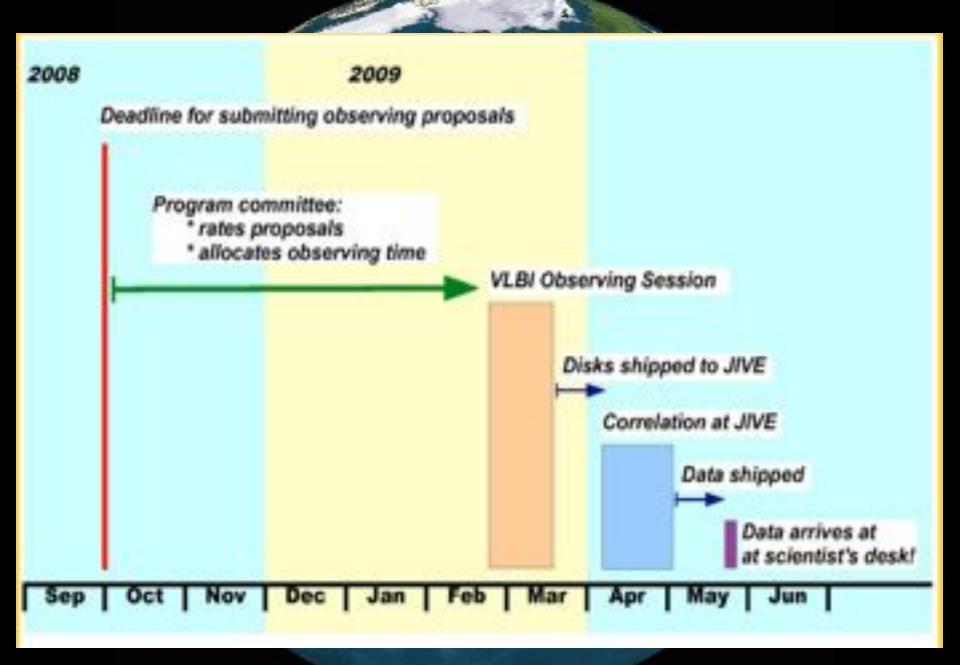


e -Very Large Base Interferometer

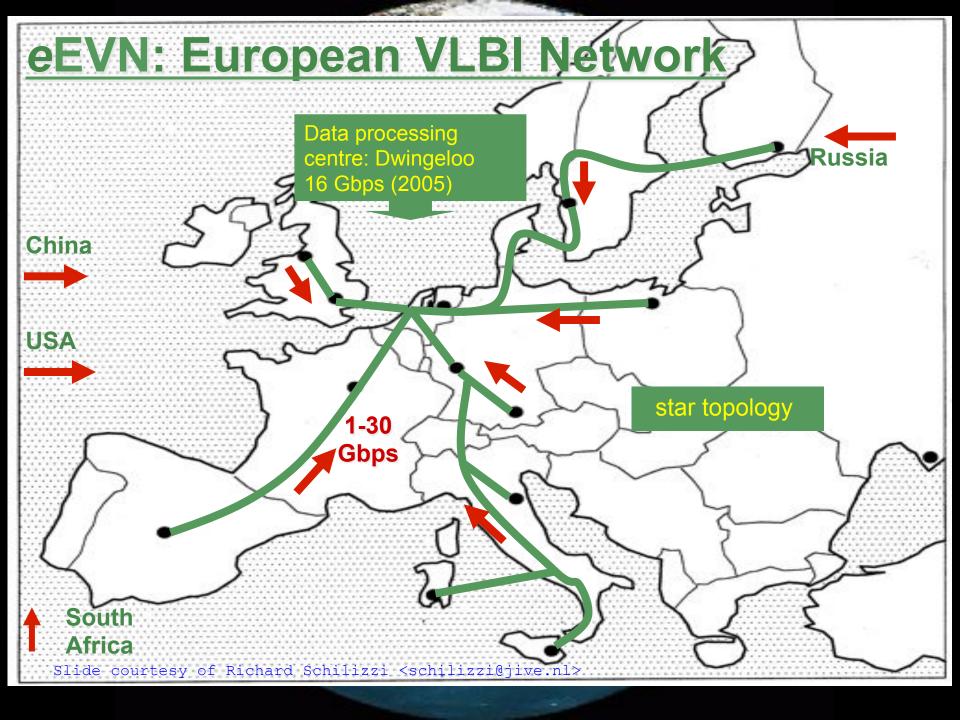


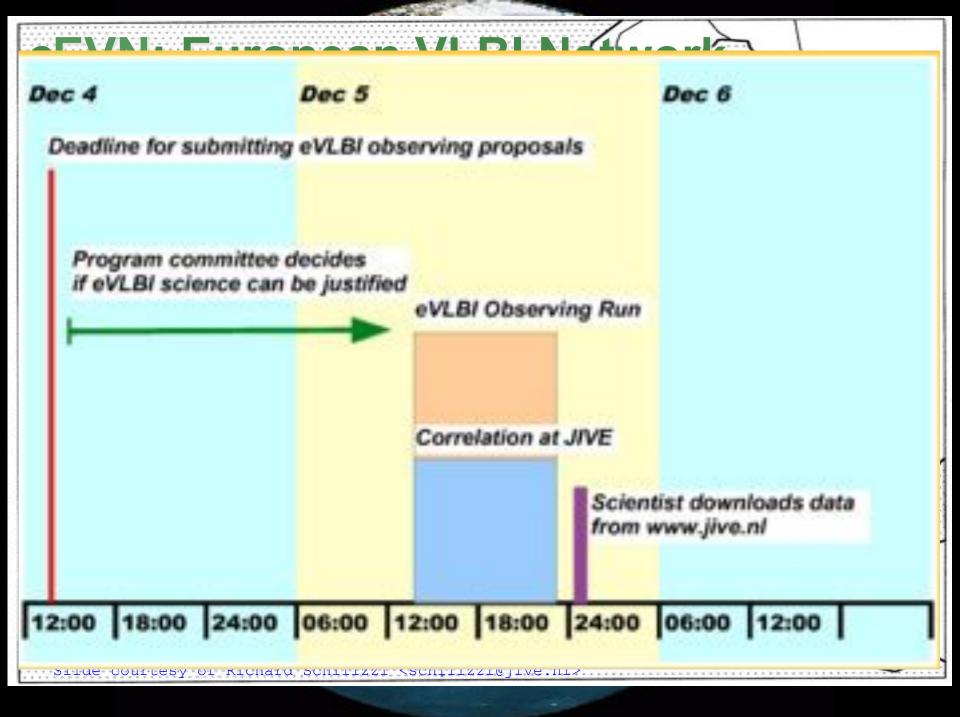






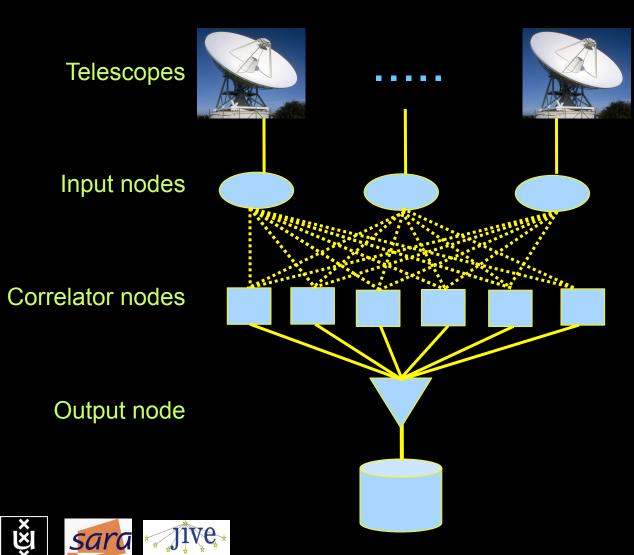






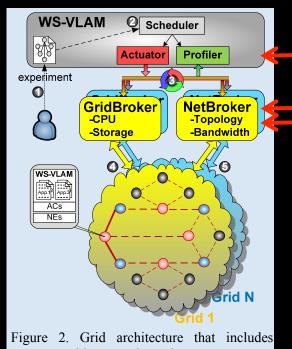
The SCARIe project

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:



programmable network services.





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

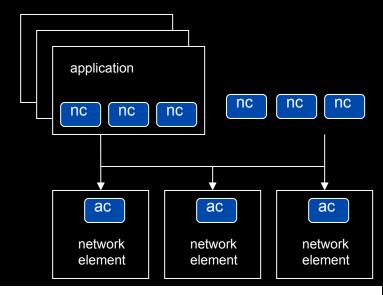
- What constitutes a Tb/s network?
- think back to teraflop computing!
 - MPI turns a room full of pc's in a teraflop machine
- massive parallel channels in hosts, NIC's
- TeraApps programming model supported by

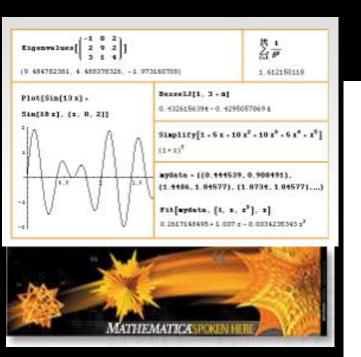
```
– TFlops -> MPI / Globus / Cloud
```

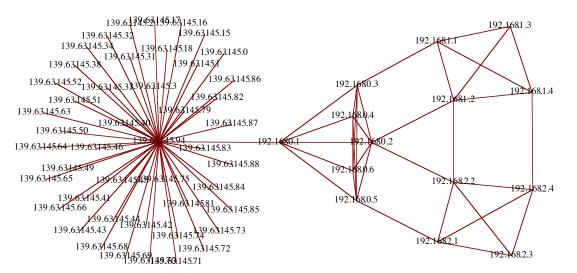
- TBytes -> DAIS / MONETdb ...
- TPixelsSAGE
- TSensors -> LOFAR, LHC, LOOKING, CineGrid, ...
- Tbit/s -> ?
- ? -> Programmable Networks

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



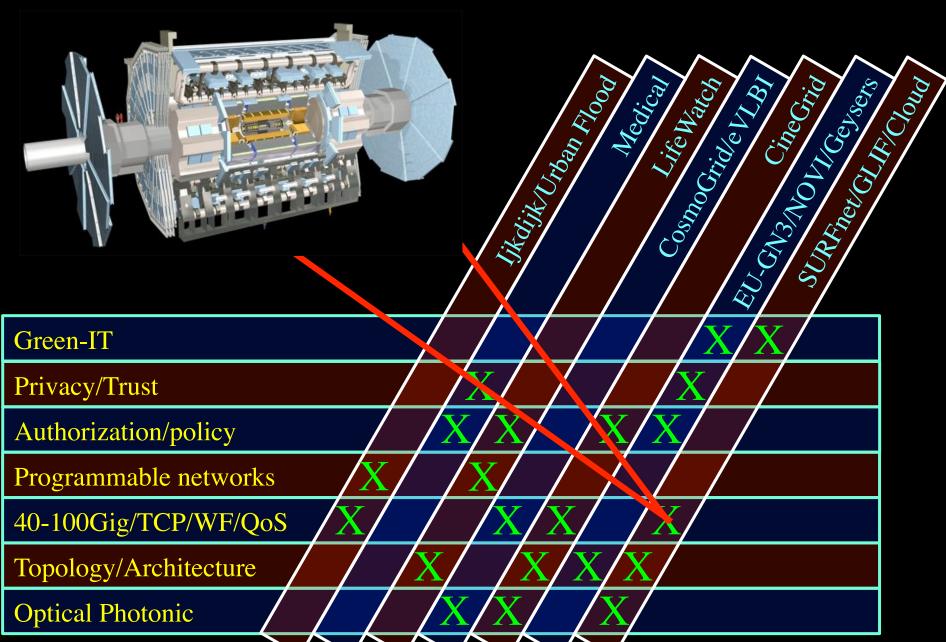




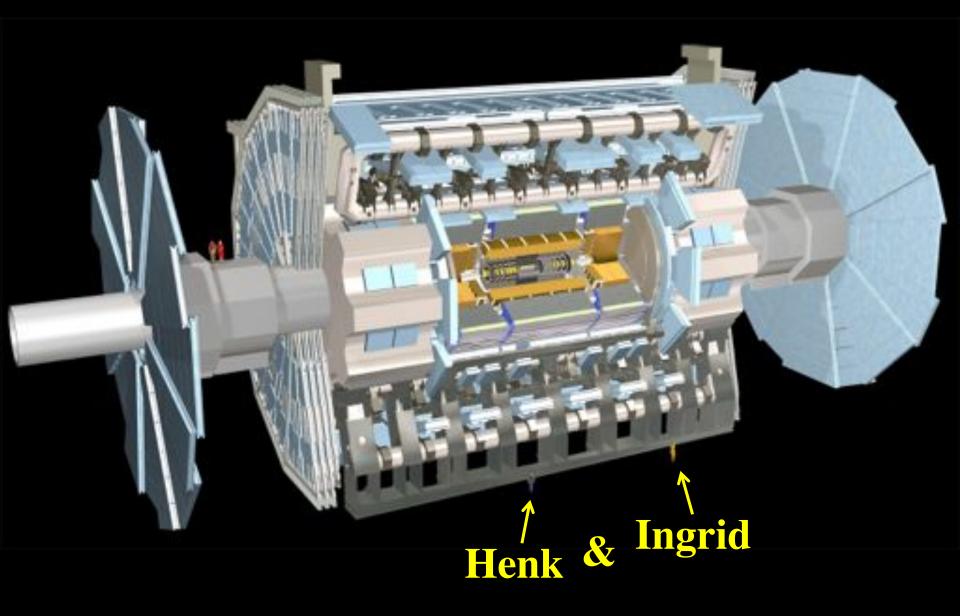
TouchTable Demonstration @ SC08



SNE @ UvA



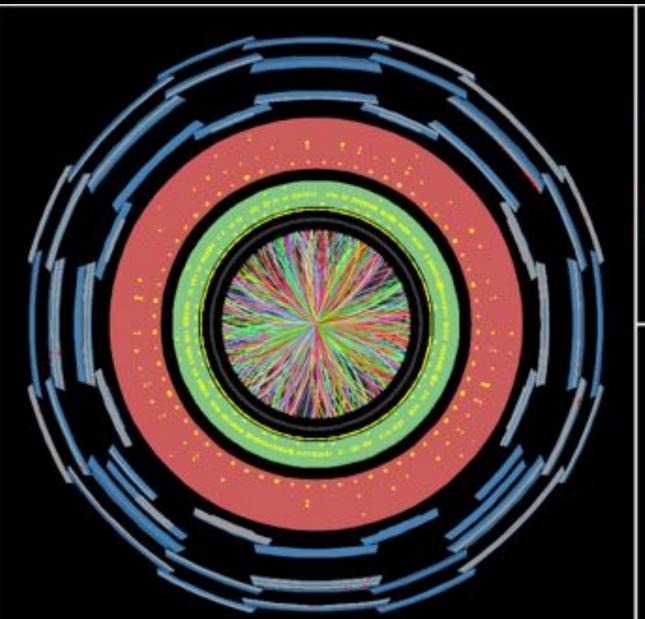
ATLAS detector @ CERN Geneve

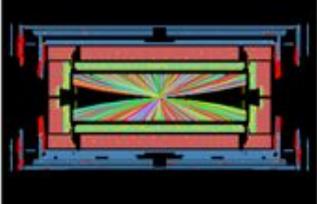


ATLAS detector @ CERN Geneve



One Heavy Ion Collision in Atlas!







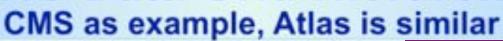
Run Number: 170482, Event Number: 3936308

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

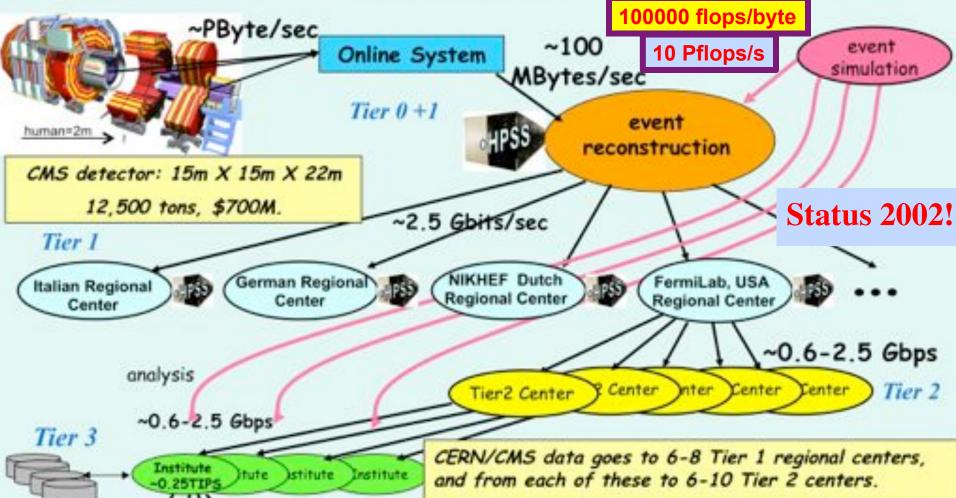
Snapshot of a heavy ion collision directly from the ATLAS experiment



LHC Data Grid Hierarchy







Tier 4 Courtesy Harvey Newman, Workstations CalTech and CERN

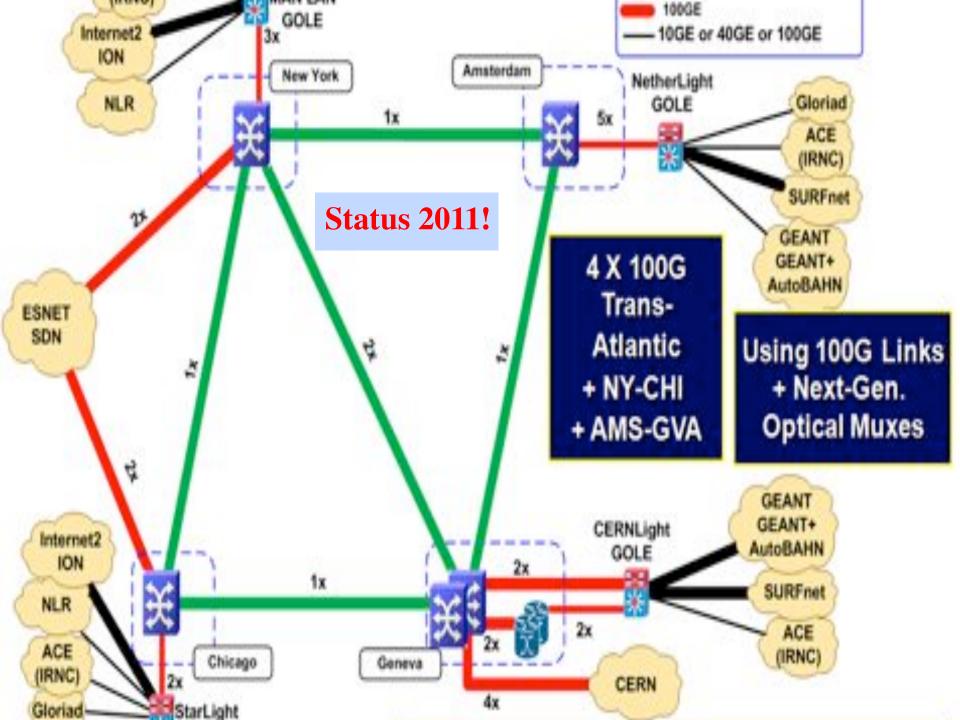
100 - 1000 Mbits/sec

Physics data cache

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 20year experiment in which DOE is a major player.

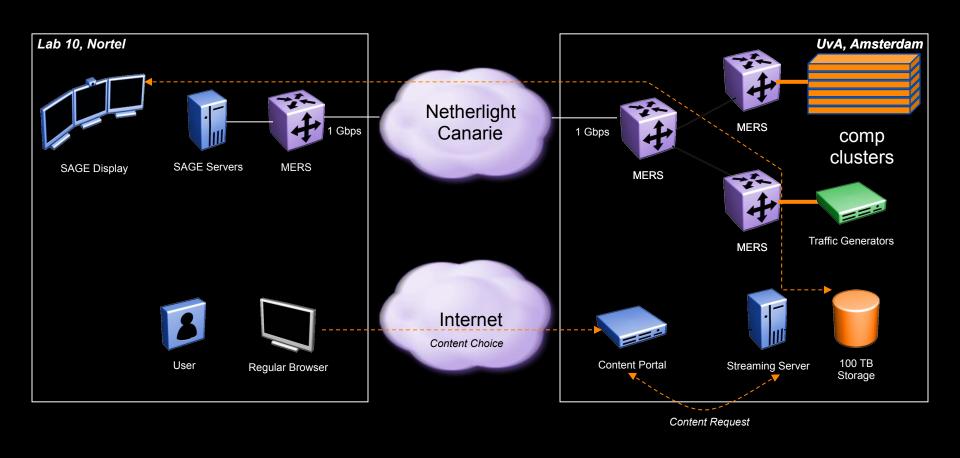


Big and small flows don't go well together on the same wire!



Diagram for SAGE video streaming to ATS





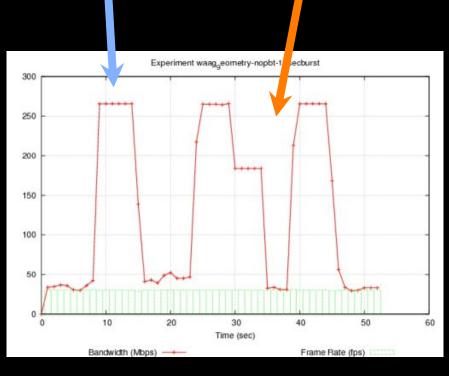
-Nortel- CIENA Confidential

Experimental Data

Sage without background traffic

Sage with background traffic





Experiment waag_geometry-pbt-10secburst

300
250
200
150
100
50
Time (sec)

Frame Rate (fps)

10 Second Traffic bursts with No PBT

10 Second Traffic bursts with PBT

PBT is <u>SIMPLE</u> and <u>EFFECTIVE</u> technology to build a shared Media-Ready Network



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

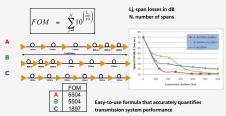
Alien wavelength challenges

- 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.

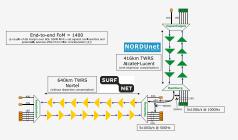
New method to present fiber link quality, FoM (Figure of Merit)

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.



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.





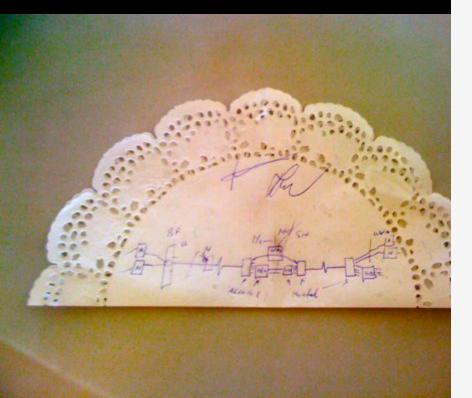




REFERENCES

[1] "OPEXAINMAS OF ALL-OPTICAL CORN INFORMATION CORN INFO

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

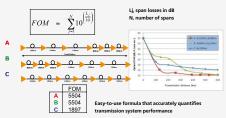
Alien wavelength challenges

- 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.

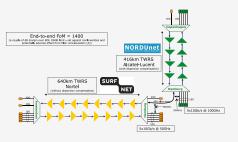
New method to present fiber link quality, FoM (Figure of Marit)

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.





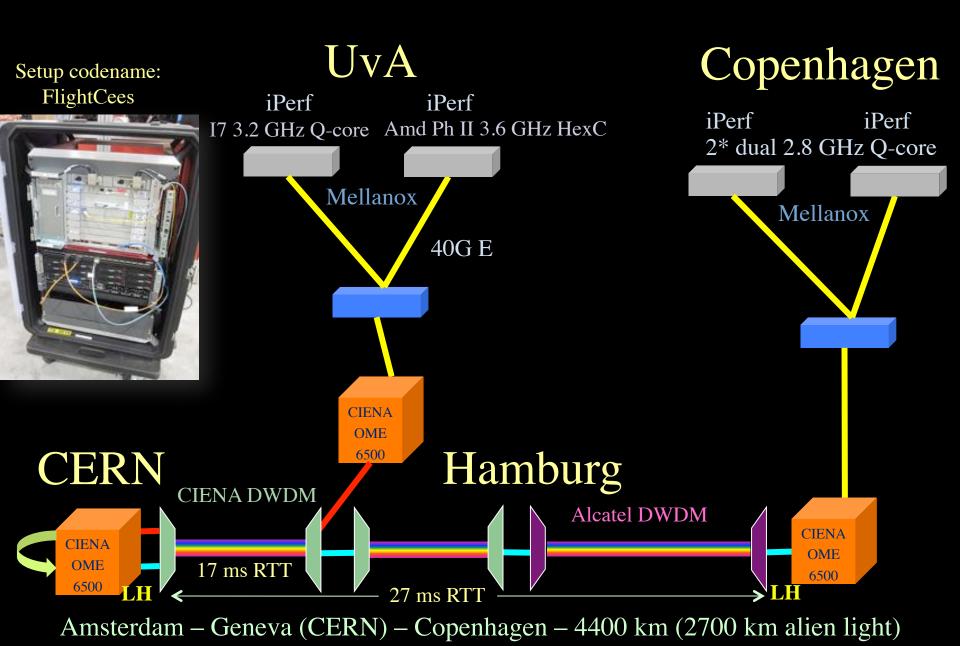




REFERENCES

[3] "OPEX SAVINGES OF ALL-OPTICAL ORGEN ETWORKS," CEREINE LET AL, OF CAUSE 11, [2] "A LET OFF (IT ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES, "BARBARKA E. SHEN LET ALL HEARDS OFF ESWINES," BARBARKA E. SHEN LET ALL HEARDS OFF E

ClearStream @ TNC2011



Demo setup codename: FlightCees



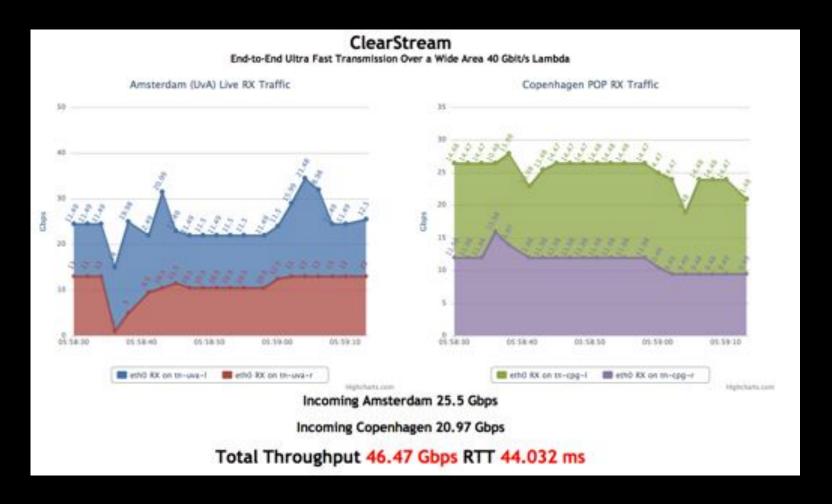
Ciena ActiveFlex(OME) 6500

Broadcom 40GE 18 port L2 Ethernet Switch

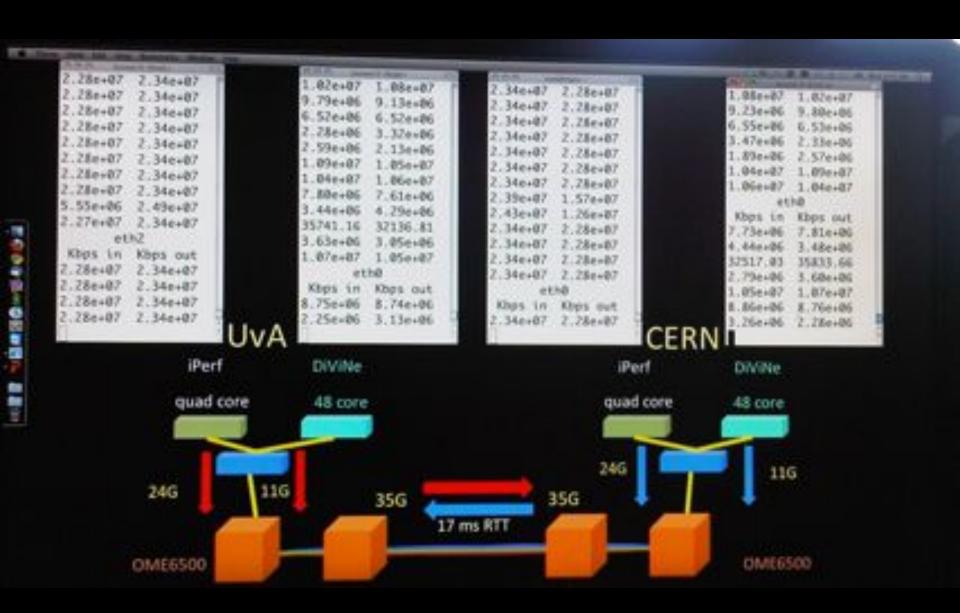
Supermicro Intel Server

Dell R815 Server

Visit CIENA Booth surf to http://tnc11.delaat.net



From GLIF October 2010 @ CERN



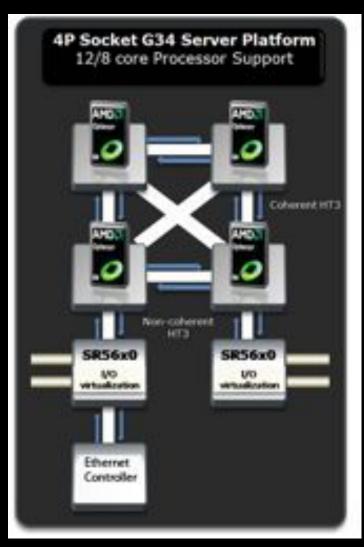
Results (rtt = 17 ms)

- 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

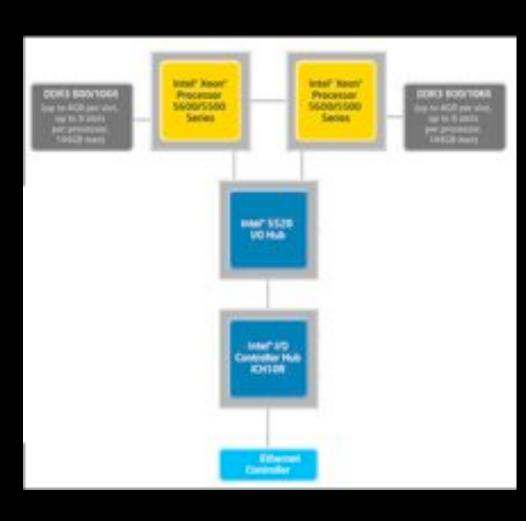
Performance Explained

- 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

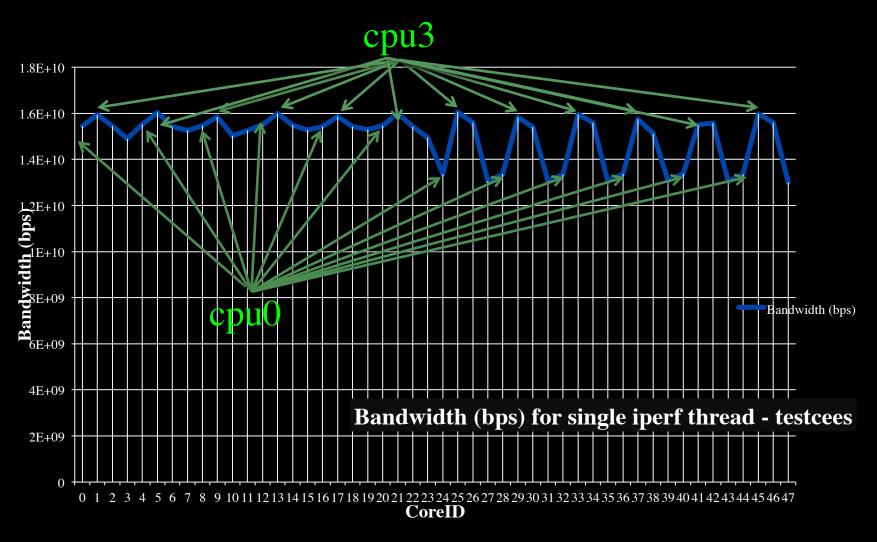


DELL R815 4 x AMD Opteron 6100



Supermicro X8DTT-HIBQF 2 x Intel Xeon

CPU Topology benchmark



We used numactl to bind iperf to cores







for

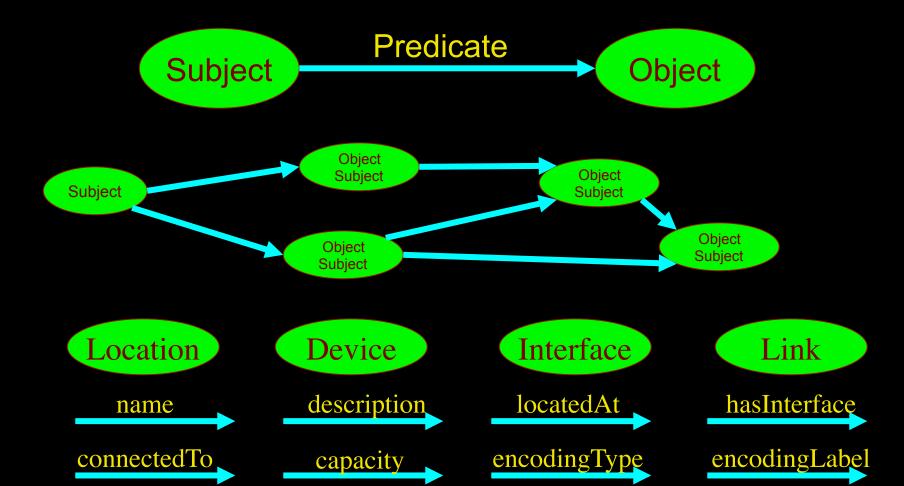
We investigate: complex networks!



LinkedIN for Infrastructure



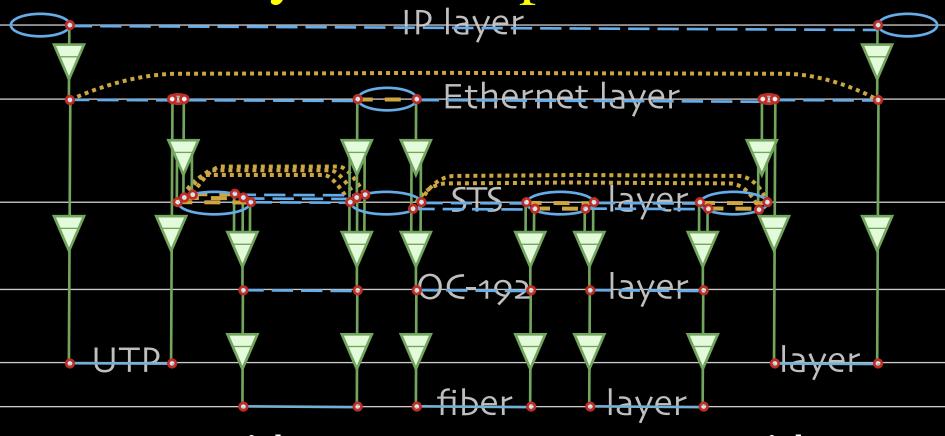
- From semantic Web / Resource Description Framework.
- The RDF uses XML as an interchange syntax.
- Data is described by triplets (Friend of a Friend):



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

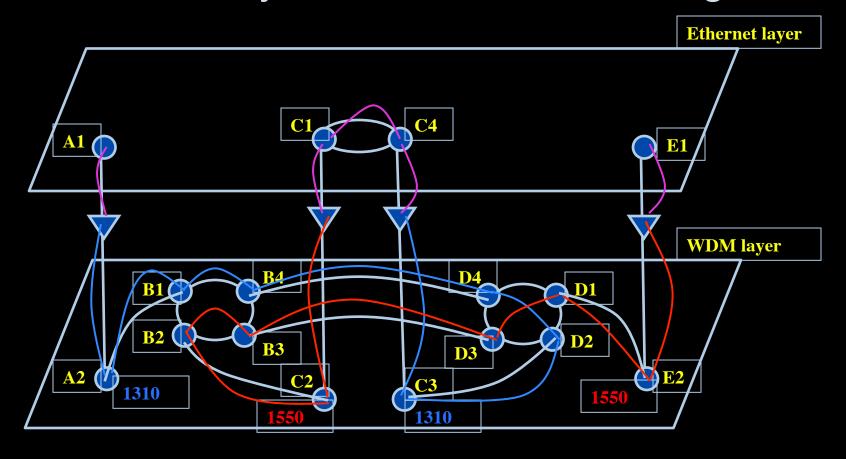


SONET switch with Ethernet intf.

NetherLight Amsterdam End host

Universiteit van Amsterdam

Multi-layer Network PathFinding

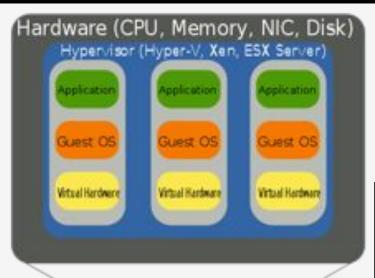


Path between interfaces A1 and E1:

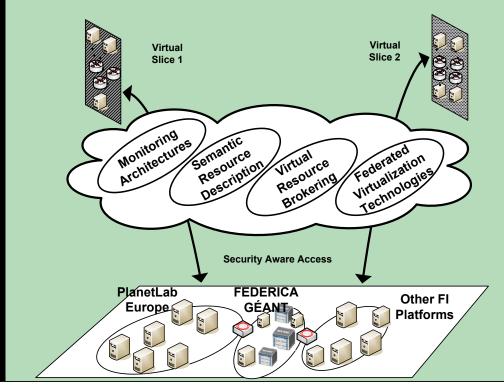
A1-A2-B1-B4-D4-D2-C3-C4-C1-C2-B2-B3-D3-D1-E2-E1

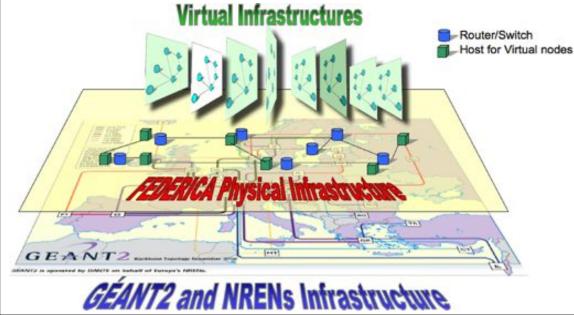
Scaling: Combinatorial problem

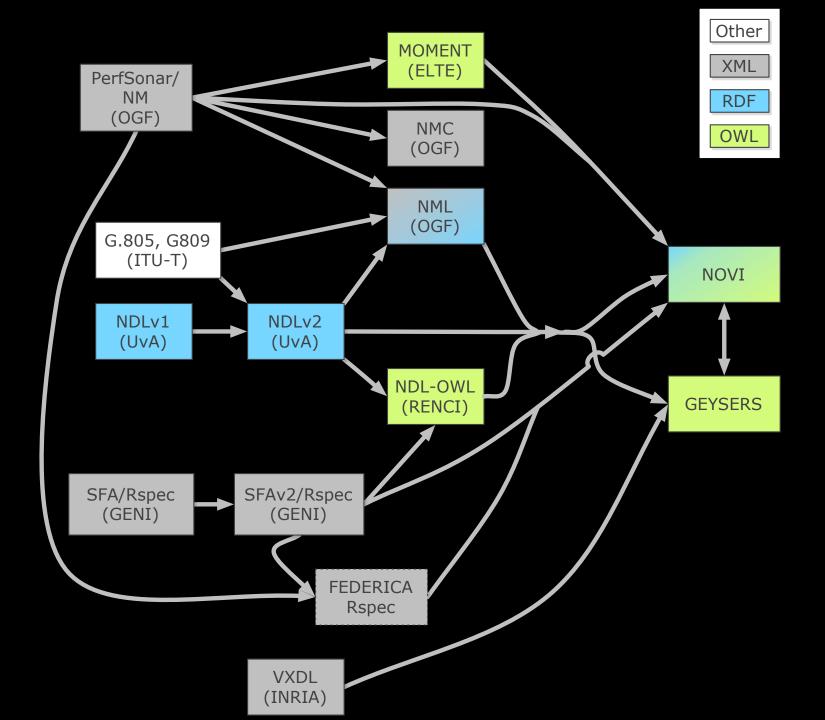
Virtualisatie van infrastructuur & QoS



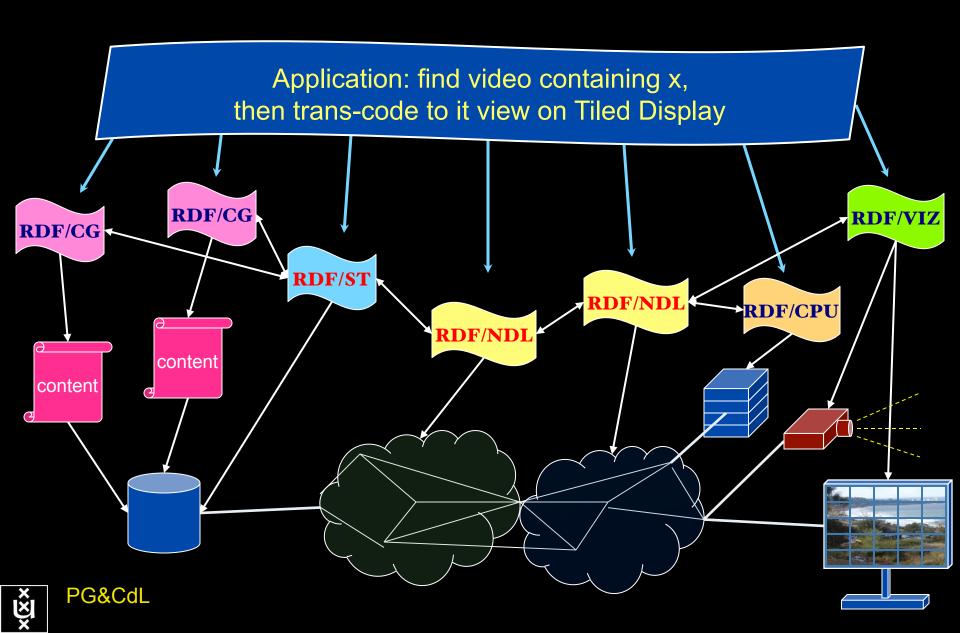








RDF describing Infrastructure



Applications and Networks become aware of each other!

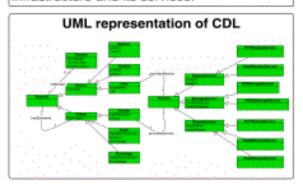
CineGrid Description Language

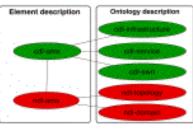
CineGrid is an initiative to facilitate the exchange, storage and display of high-quality digital media.

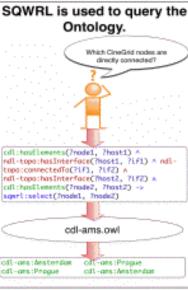
The CineGrid Description Language (CDL) describes CineGrid resources. Streaming, display and storage components are organized in a hierarchical way.

CDL has bindings to the NDL ontology that enables descriptions of network components and their interconnections.

With CDL we can reason on the CineGrid infrastructure and its services.







CDL links to NDL using the owl:SameAs property. CDL defines the services, NDL the network interfaces and links. The combination of the two ontologies identifies the host pairs that support matching services via existing network connections.



Balgh Honings viralghöberinnen ava. mir, Pamia Greener vip. germanbern mir

CineGrid portal

100 Tbyte

Cache & Store & Forward



distribution center Amsterdam

Hume | About | Browse Content | cinegrid.org | cinegrid.nl

Amsterdam Node Status:

nede41:

Disk space used: 6 G/S Disk space available: 10 G/S

Search node:

Search.

Browse by tag:

amsterdam animation antonacci blender boat tridge burns Cgl data totand hollandfestival inducentus:

muziekgebouw

rieuwmarkt OPEF8 prague ship train trams wasg



CineGrid Amsterdam

Welcome to the Amsterdam CineGrid distribution node. Below are the latest additions of super-high-quality video to our node.

for more information about CineCrid and our effords look at the about section.

Latest Additions



Wypke

Walke

Available formatic: 4c do: (4.0 KB)

Duration: I hour and 8 minutes. Created: 1 week, 2 days age

Author: Wypks Categories



Prague Train

Steam locamotive in Progue.



VLC: Big Buck Bunny

(ii) aspyright Brender Foundation (Ntp://www.brgbuckburns.org

Available formats:

4k de (3.9 kg)

Durationi 27 hours and 46 monutes

Created: 1 week, 2 days ago

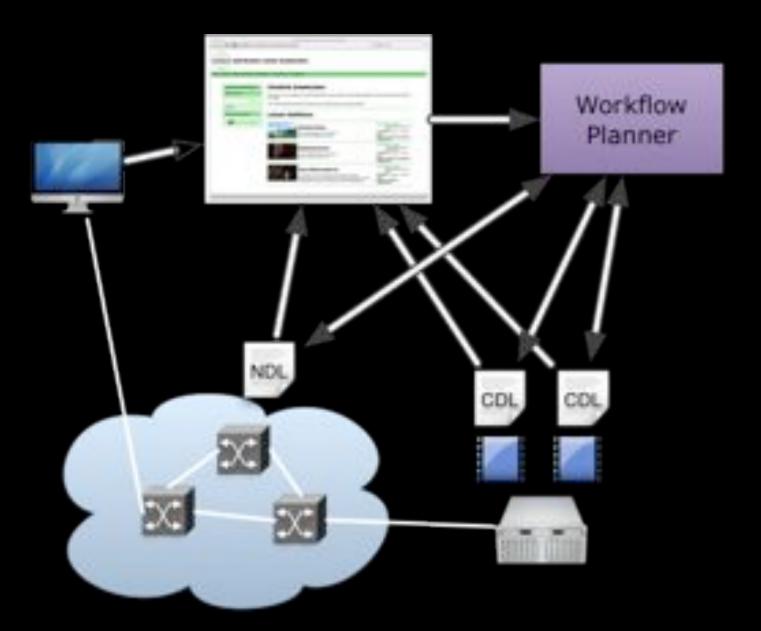
Author: CireCrid

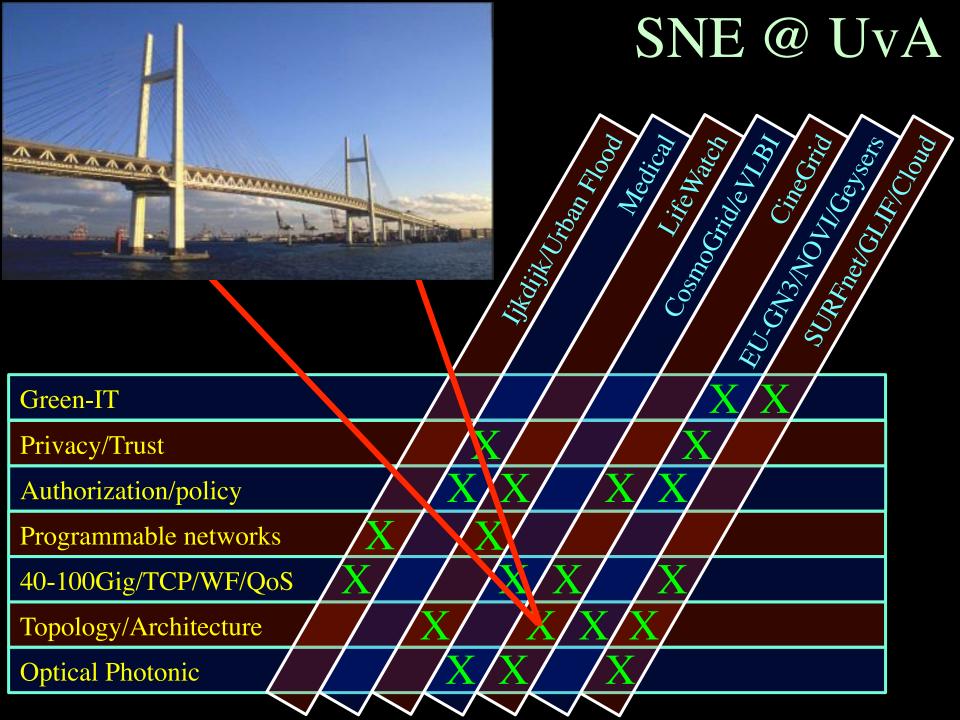
Categories: datus prague train

Available formats:

1060p HPEG4 (1.1 GB)
Durations 1 hour and 0 minutes
Created: 1 month, 1 seek ago
Author: Blender Foundation
Categories: animation blender burny

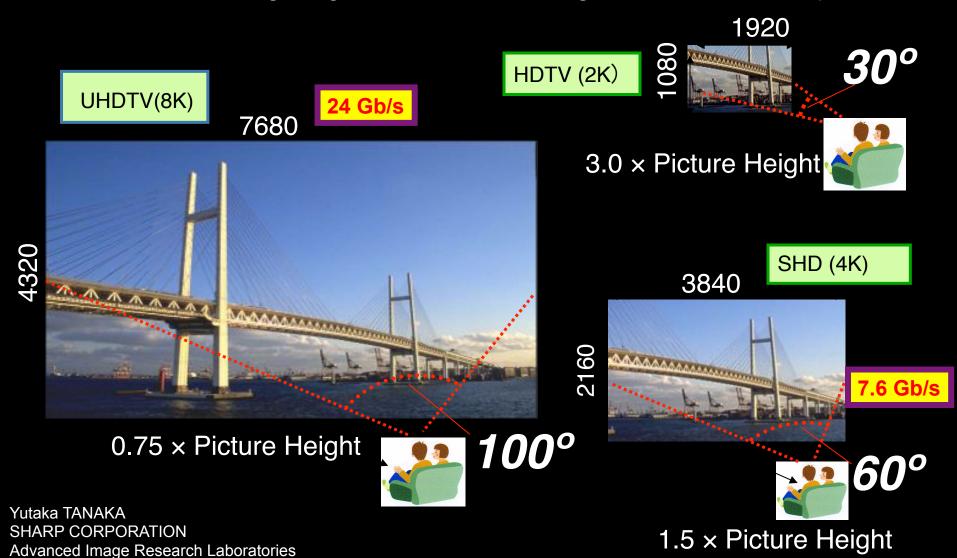
CineGrid Workflow Planner





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









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

Hybrid Networking <-> Computing

Routers



← → Supercomputers

Ethernet switches $\leftarrow \rightarrow$ Grid & Cloud



Photonic transport GPU's



What matters:

Energy consumption/multiplication

Energy consumption/bit transported

ECO-Scheduling











SARA LA

Pieken-in-de-Delta

SURFINE

SUVA

NWO

Q&A

- http://tnc11.delaat.net
- http://ipv6day.nl/ (win an iPad2!)
- http://cees.delaat.net/smartgreen/index.html
- http://cees.delaat.net/netbuf/index.html
- http://cees.delaat.net/ssdtcp/index.html
- Etc.