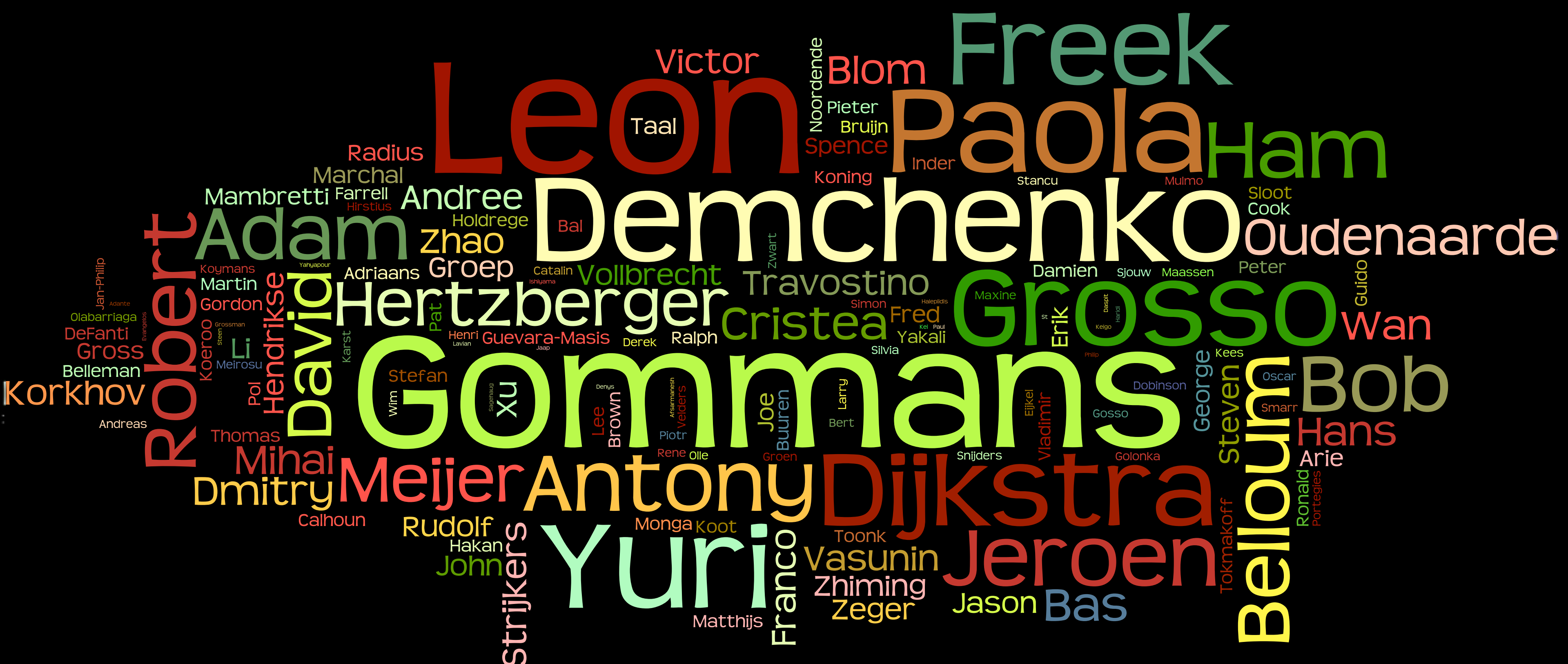
System and Network Engineering Research for Big Data Sciences

Cees de Laat



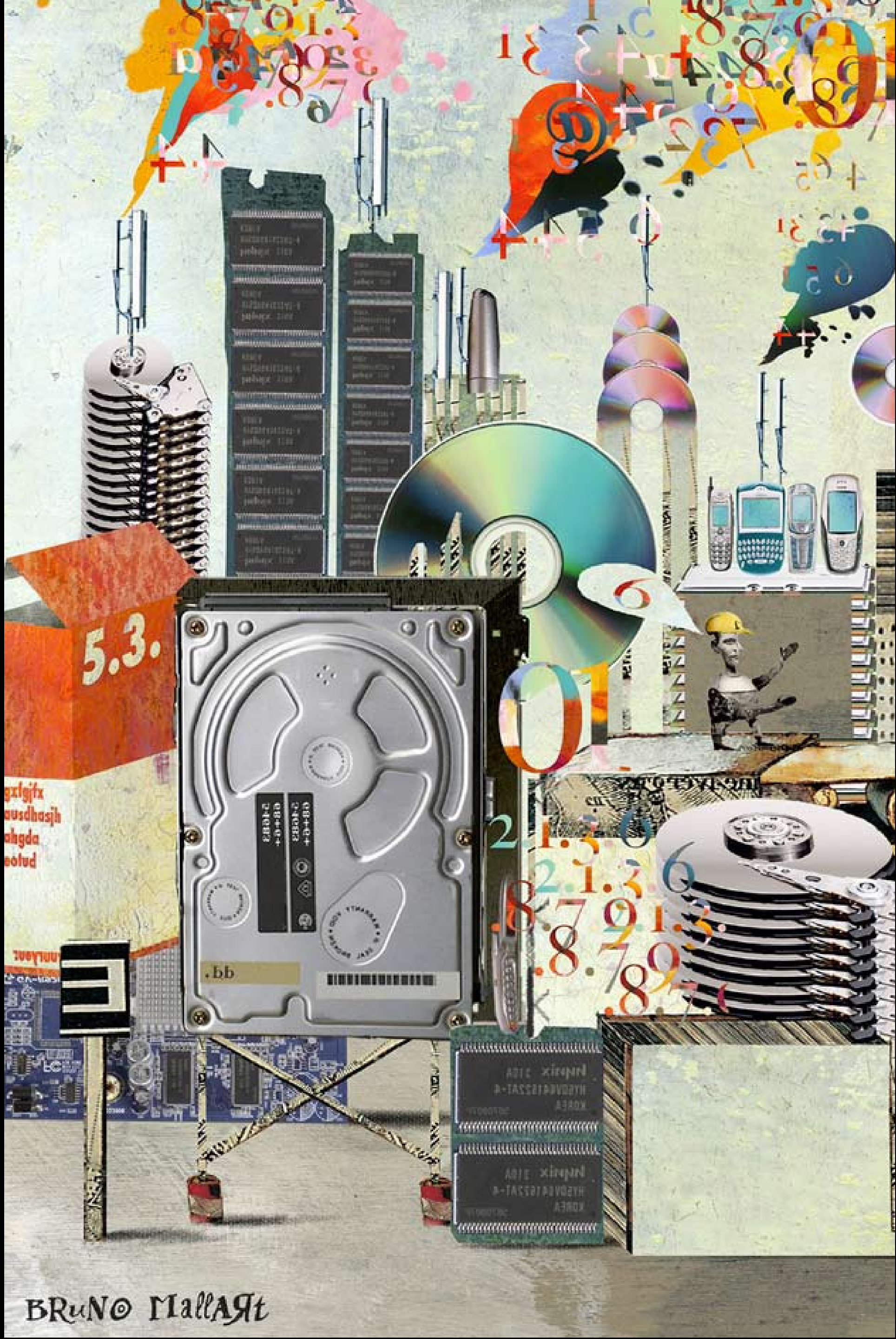
From King's Dutch Academy of Sciences

The Dutch Research Agenda

"Information technology (IT) now permeates all aspects of public, commercial, social, and personal life. bank cards, satnav, and weather radar... IT has become completely indispensable."

"But to guarantee the reliability and quality of constantly bigger and more complicated IT, we will need to find answers to some fundamental questions!"

http://www.knaw.nl/Content/Internet_KNAW/publicaties/pdf/20111029.pdf



Internet

From a network experiment that never ended (Vint Cerf)

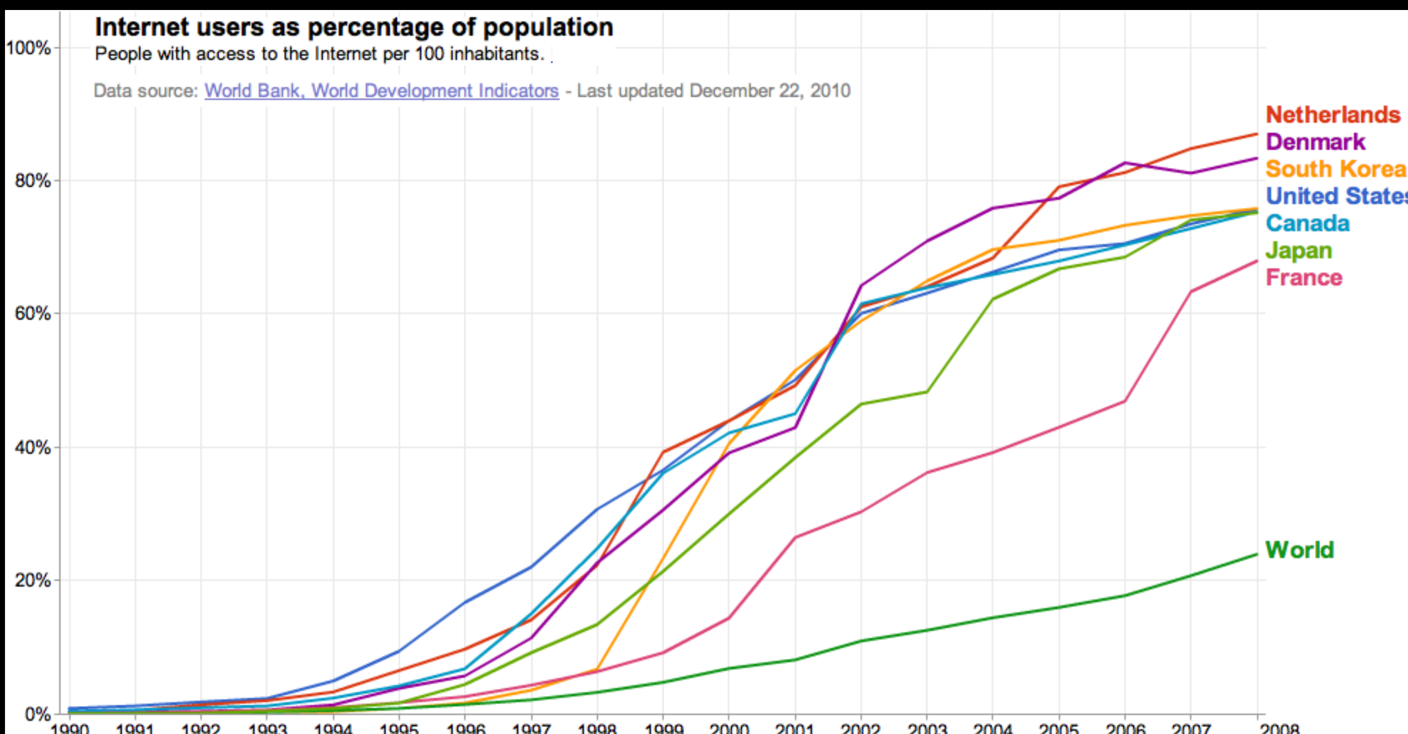
1974: for the first time the word **internet** (*RFC 675 - Specification of Internet Transmission Control Program) [note -> Open process!]*

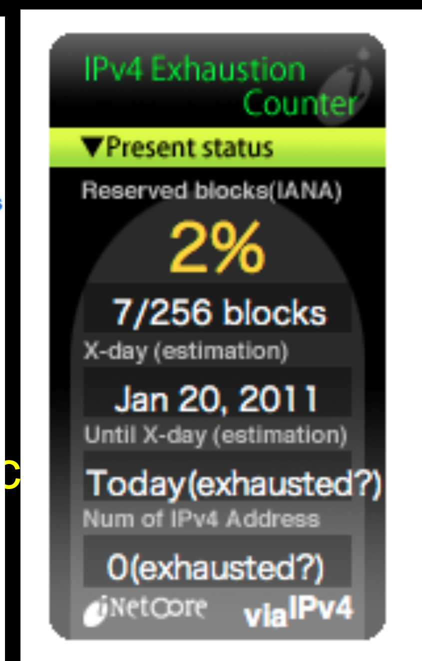
1981: the **TCP/IP** standard was ready to be adopted (*RFC 791,792,793*)

To a network for society

1989: WWW was born

2010

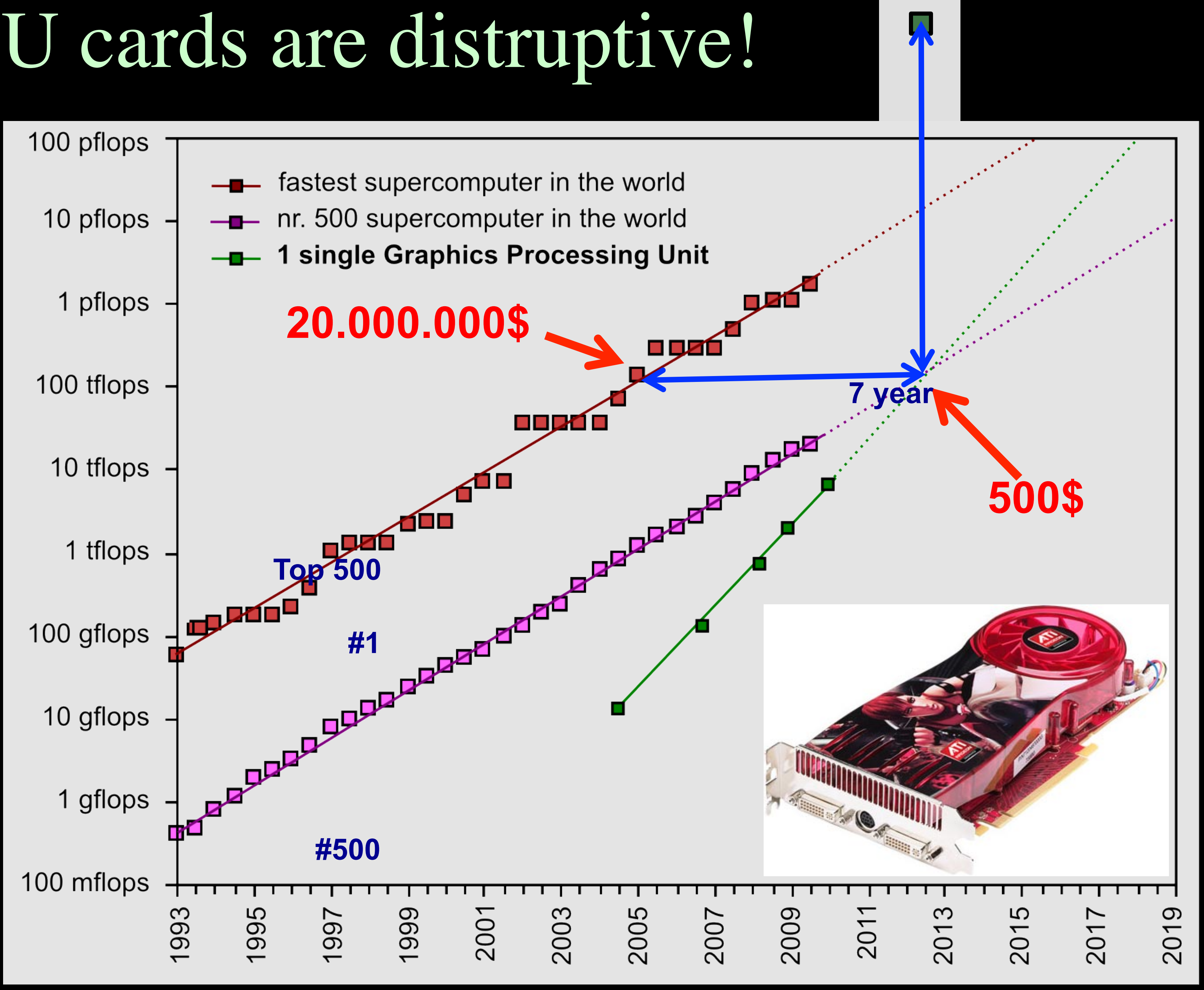




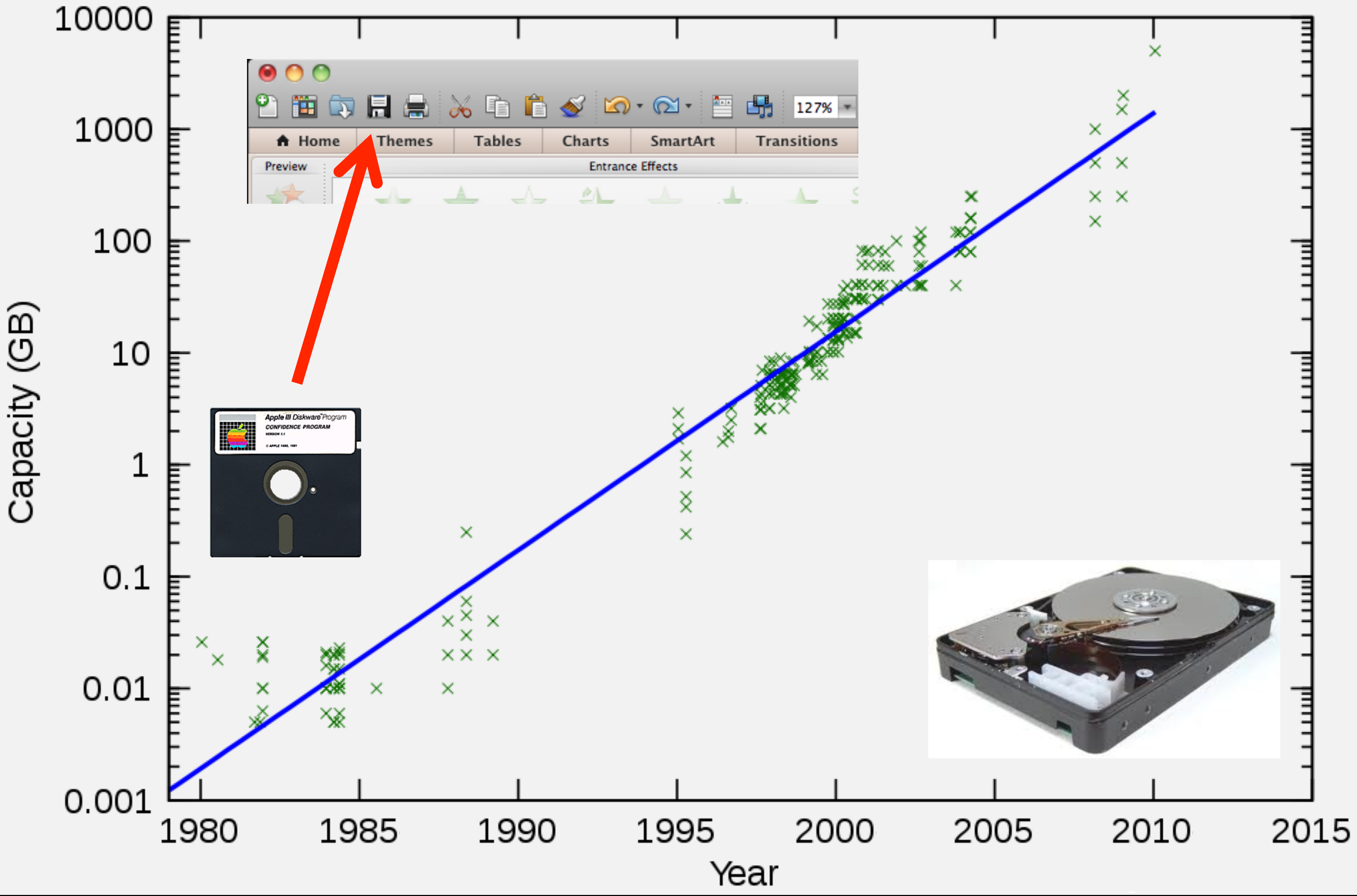
Ipv6day.nl


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

GPU cards are distruptive!



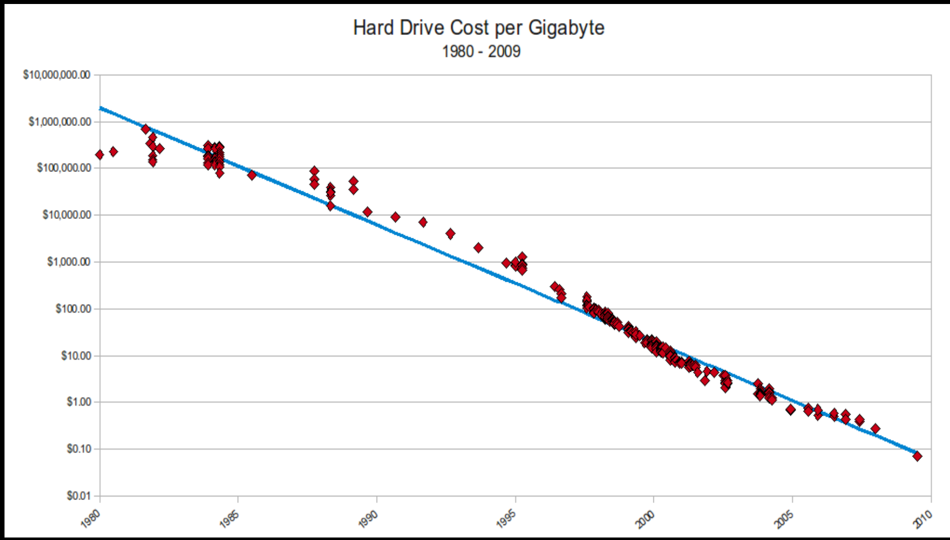
Data storage: doubling every 1.5 year!



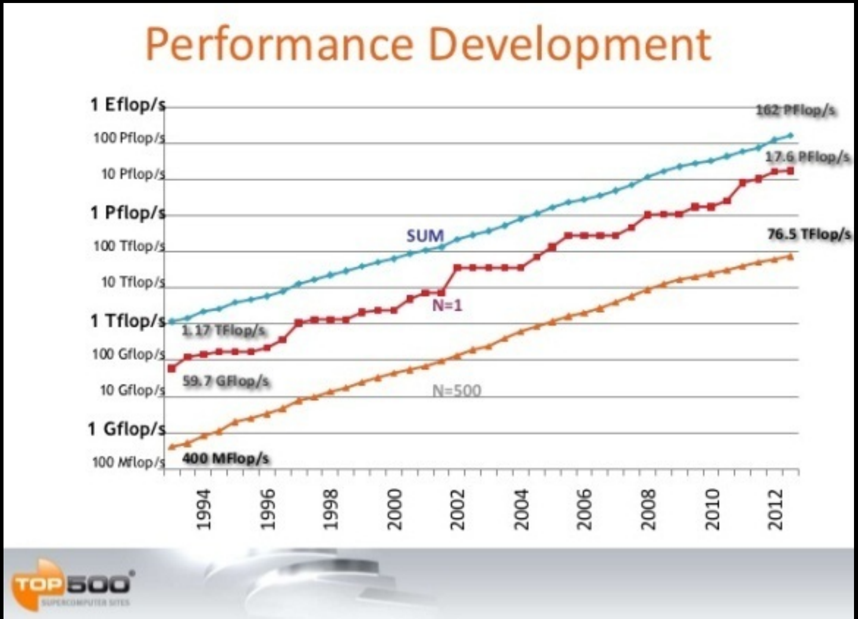
Reliable and Safe!

This omnipresence of IT makes us not only strong but also vulnerable.

- A virus, a hacker, or a system failure can instantly send digital shockwaves around the world.



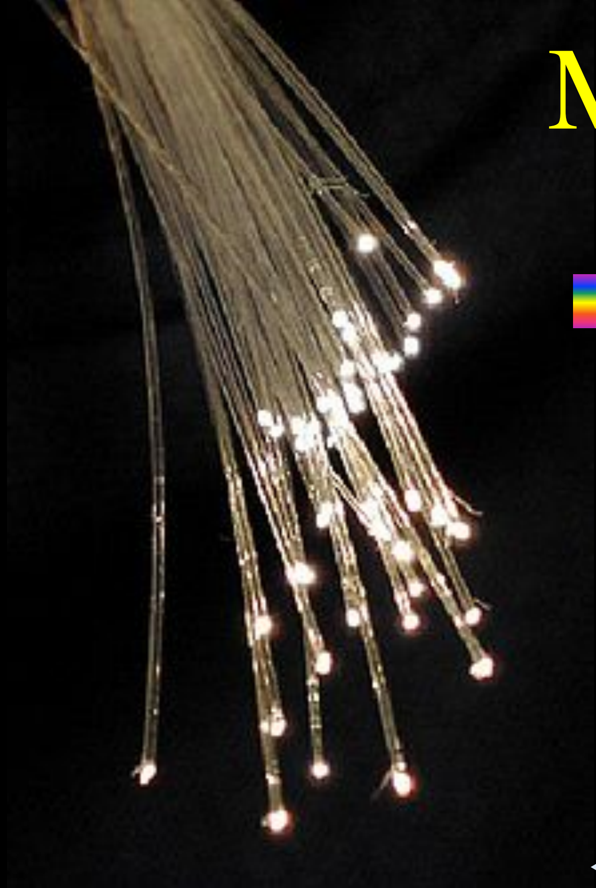
The hardware and software that allow all our systems to operate is becoming bigger and more complex all the time, and the capacity of networks and data storage is increasing by leaps and bounds.

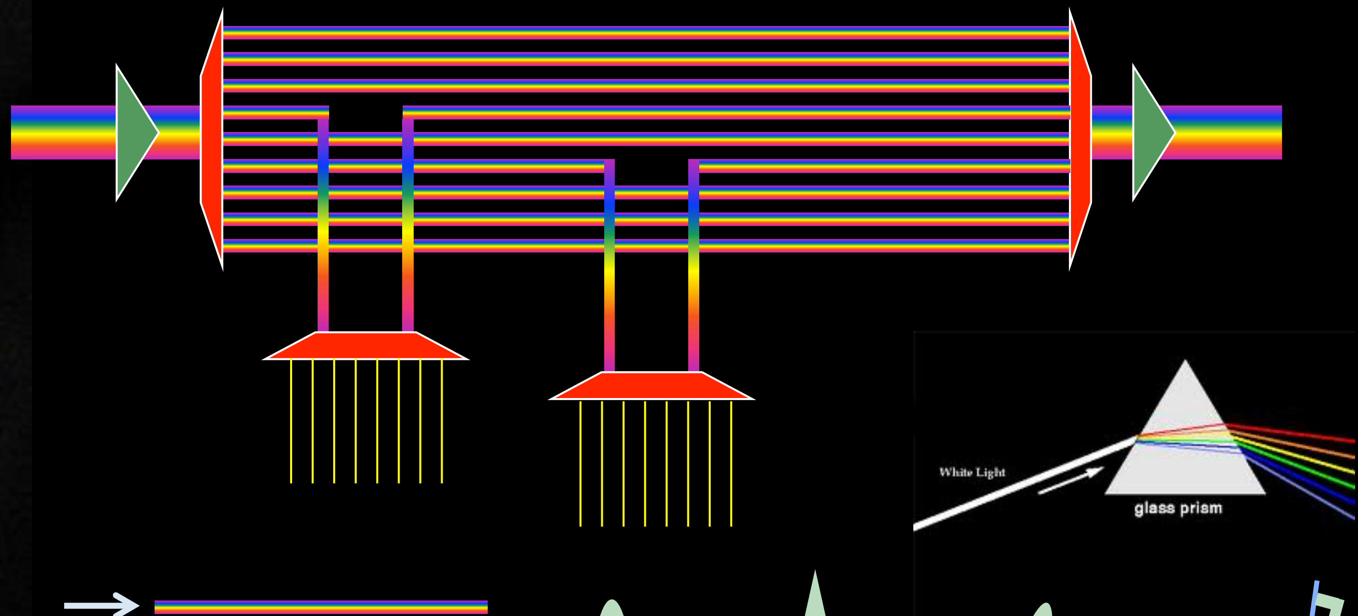


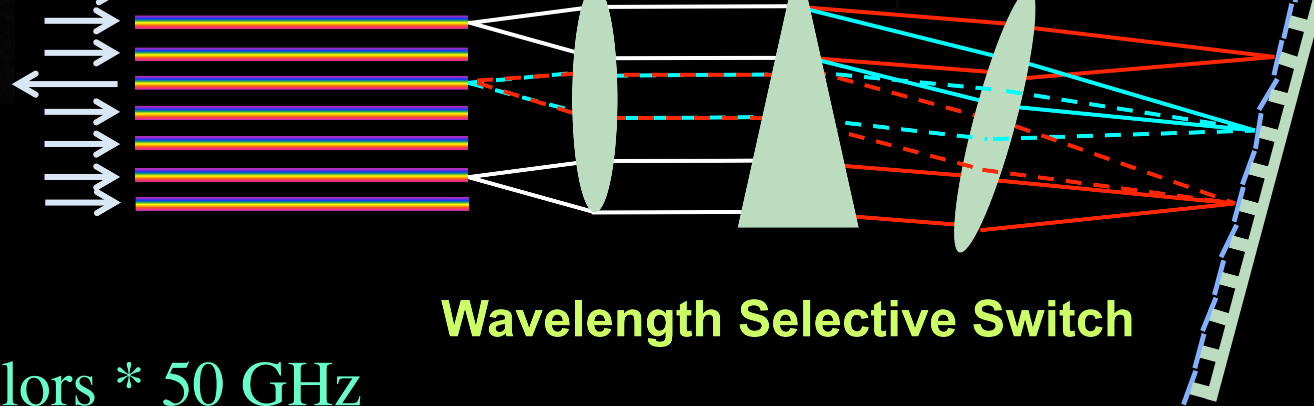
We will soon reach the limits of what is currently feasible and controllable.

http://www.knaw.nl/Content/Internet_KNAW/publicaties/pdf/20111029.pdf

Multiple colors / Fiber







Wavelength Selective Switch

Per fiber: ~ 80-100 colors * 50 GHz

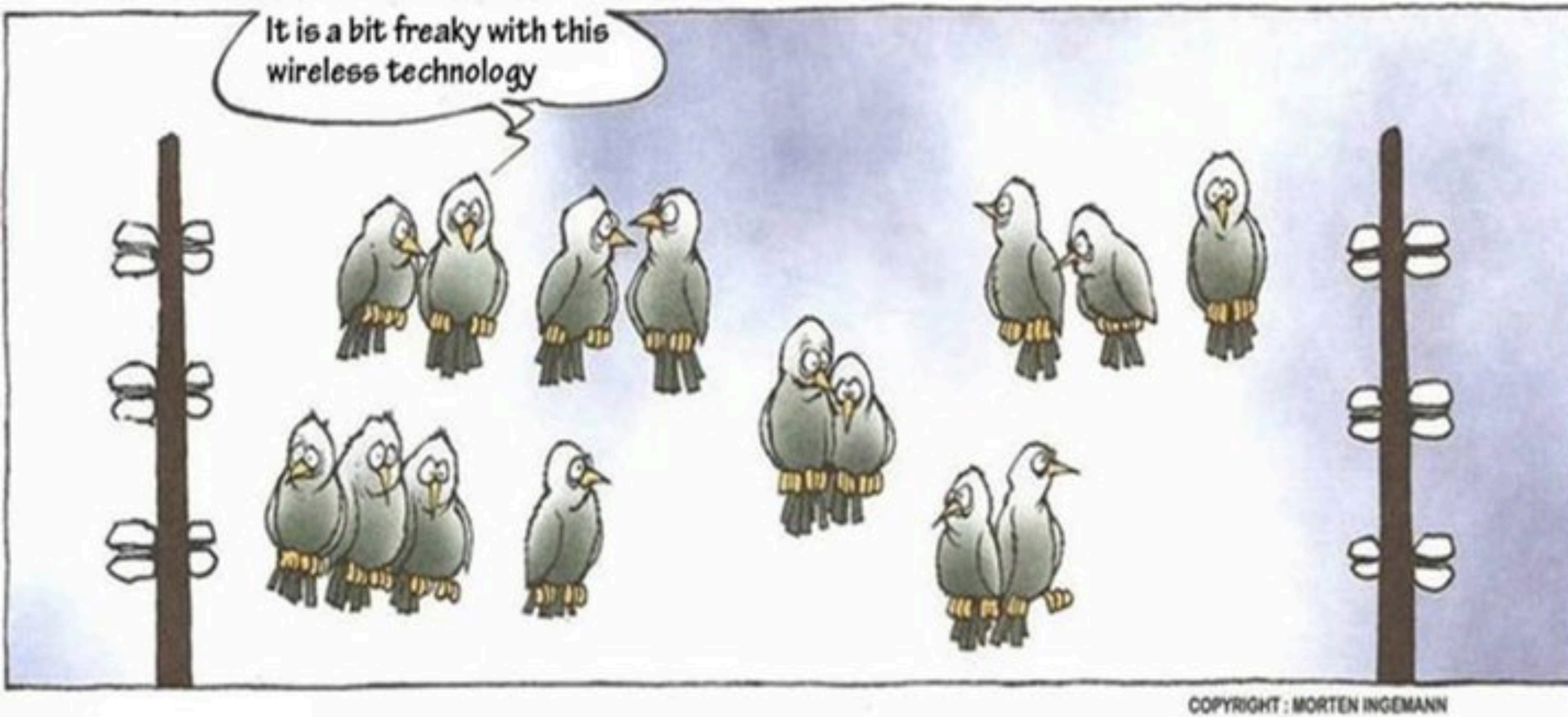
Per color: 10 – 40 – 100 Gbit/s

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

New: Hollow Fiber!

➔ less RTT!

Wireless Networks



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.

Reduction of Complexity by Integration

By combining services such as telephony, television, data, and computing capacity within a single network, we can cut down on complexity, energy consumption and maintenance.

- How can we describe and analyze complex information systems effectively?
- How can we specify and measure the quality and reliability of a system?
- How can we combine various different systems?
- How can we design systems in which separate processors can co-operate efficiently via mutual network connections within a much larger whole?
- Can we design information systems that can diagnose their own malfunctions and perhaps even repair them?
- How can we specify, predict, and measure system performance as effectively as possible?

SNE addresses a.o. the **highlighted** questions!

http://www.knaw.nl/Content/Internet_KNAW/publicaties/pdf/20111029.pdf

Mission

Can we create smart data processing infrastructures that are tailored to diverse application needs?

Mission

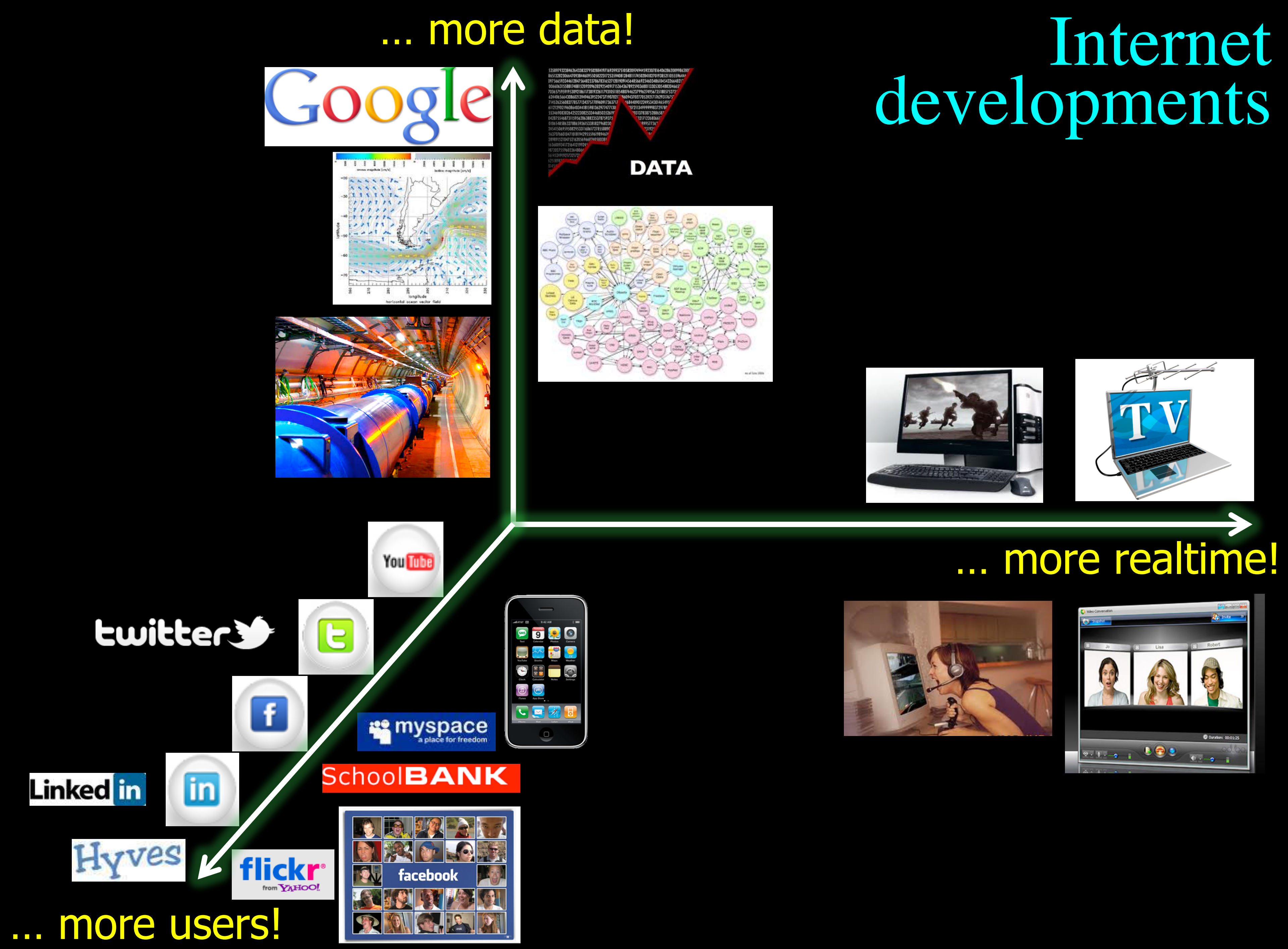
Can we create smart and safe data processing infrastructures that can be tailored to diverse application needs?

- *Capacity*
- *Capability*
- *Security*
- *Sustainability*
- *Resilience*

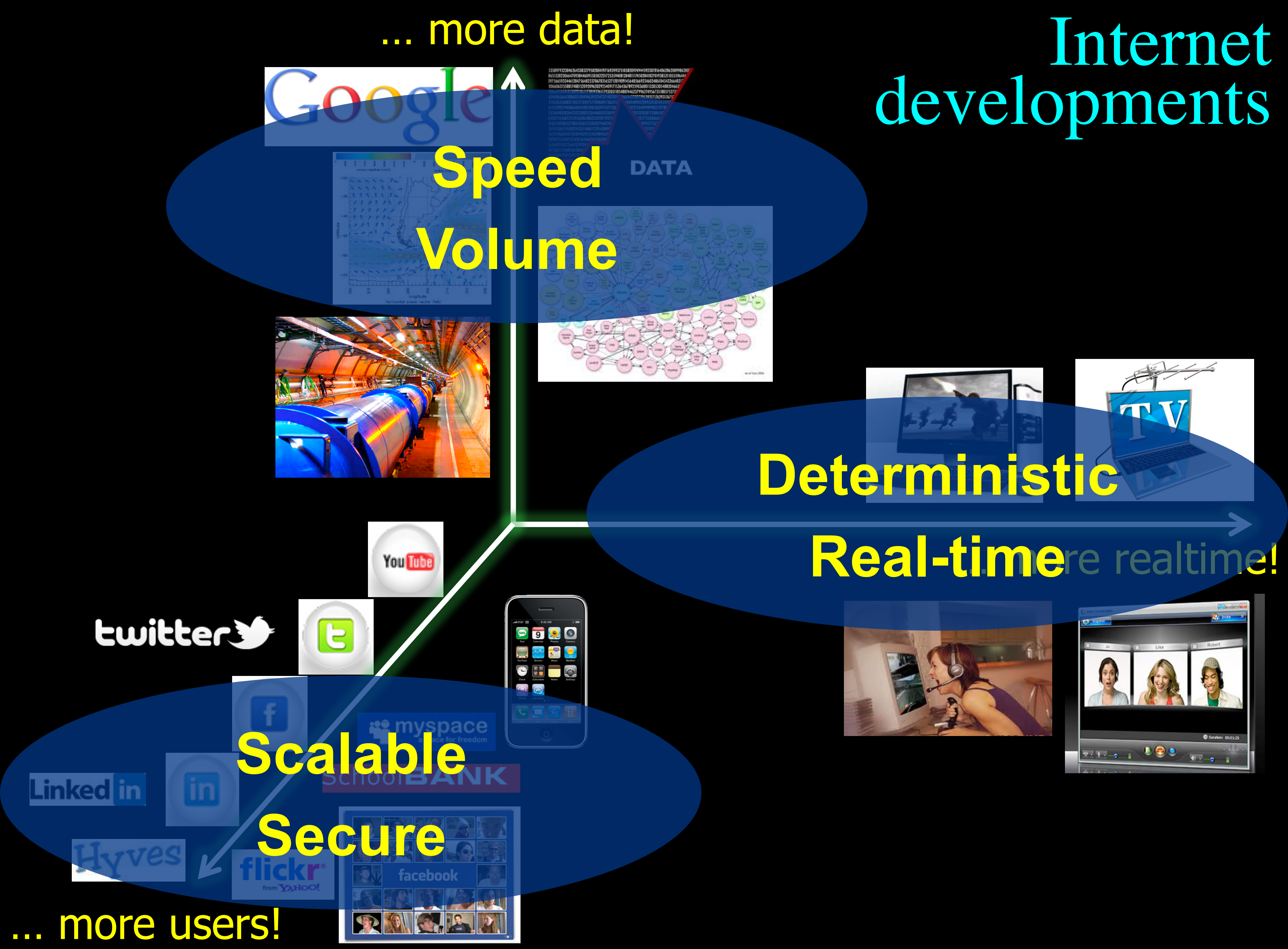
Mission

Can we create smart and safe data processing infrastructures that can be tailored to diverse application needs?

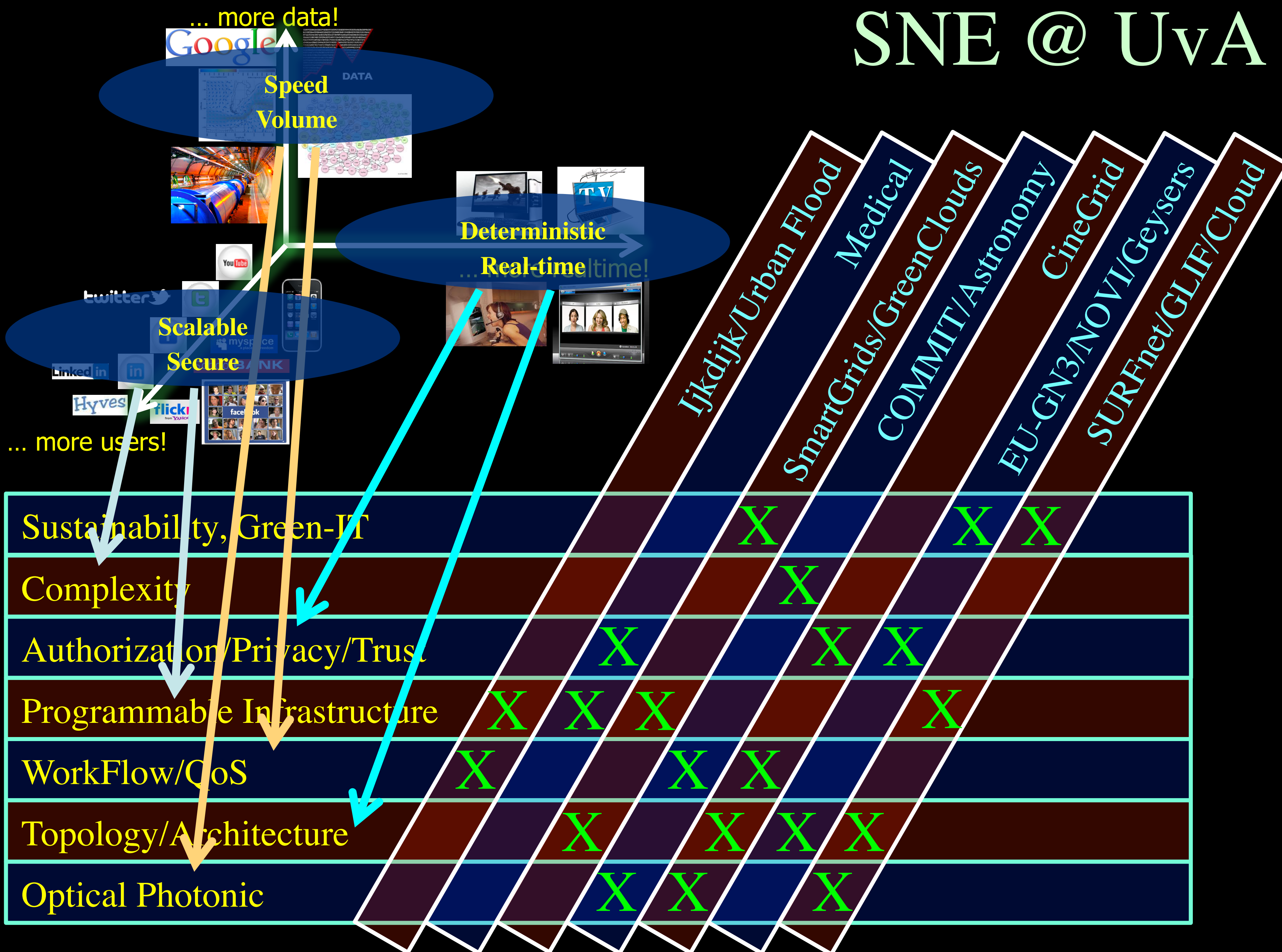
- *Capacity*
  - *Bandwidth on demand, QoS, architectures, photonics, performance*
- *Capability*
  - *Programmability, virtualization, complexity, semantics, workflows*
- *Security*
  - *Authorization, Anonymity, integrity of data in distributed data processing*
- *Sustainability*
  - *Greening infrastructure, awareness*
- *Resilience*
  - *Systems under attack, failures, disasters*


Internet developments
... more data!
Google
DATA
... more realtime!
TV
... more users!
You Tube
twitter
myspace
a place for freedom
SchoolBANK
Linked in
Hyves
flickr
from YAHOO!
facebook

Internet developments

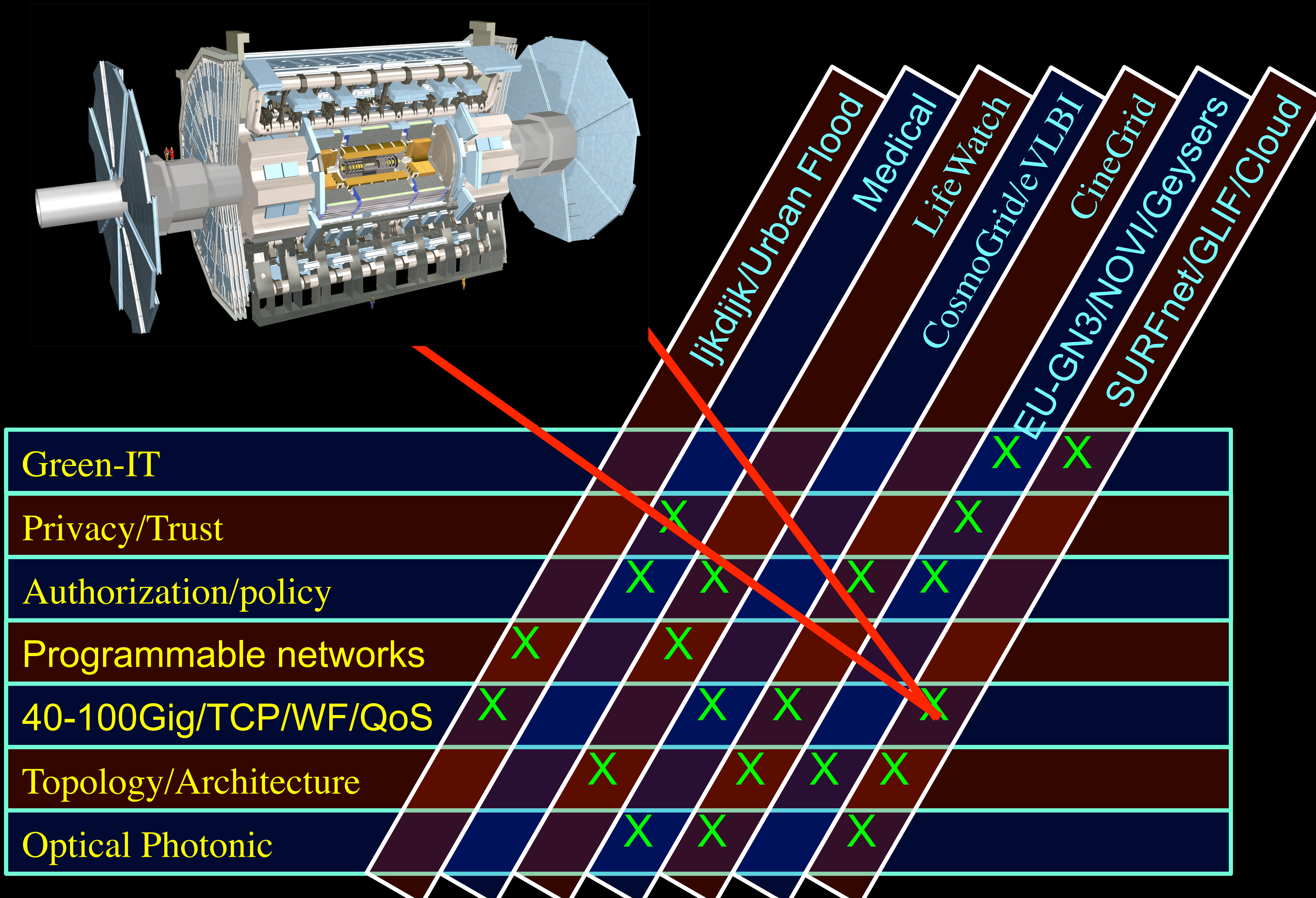


SNE @ UvA



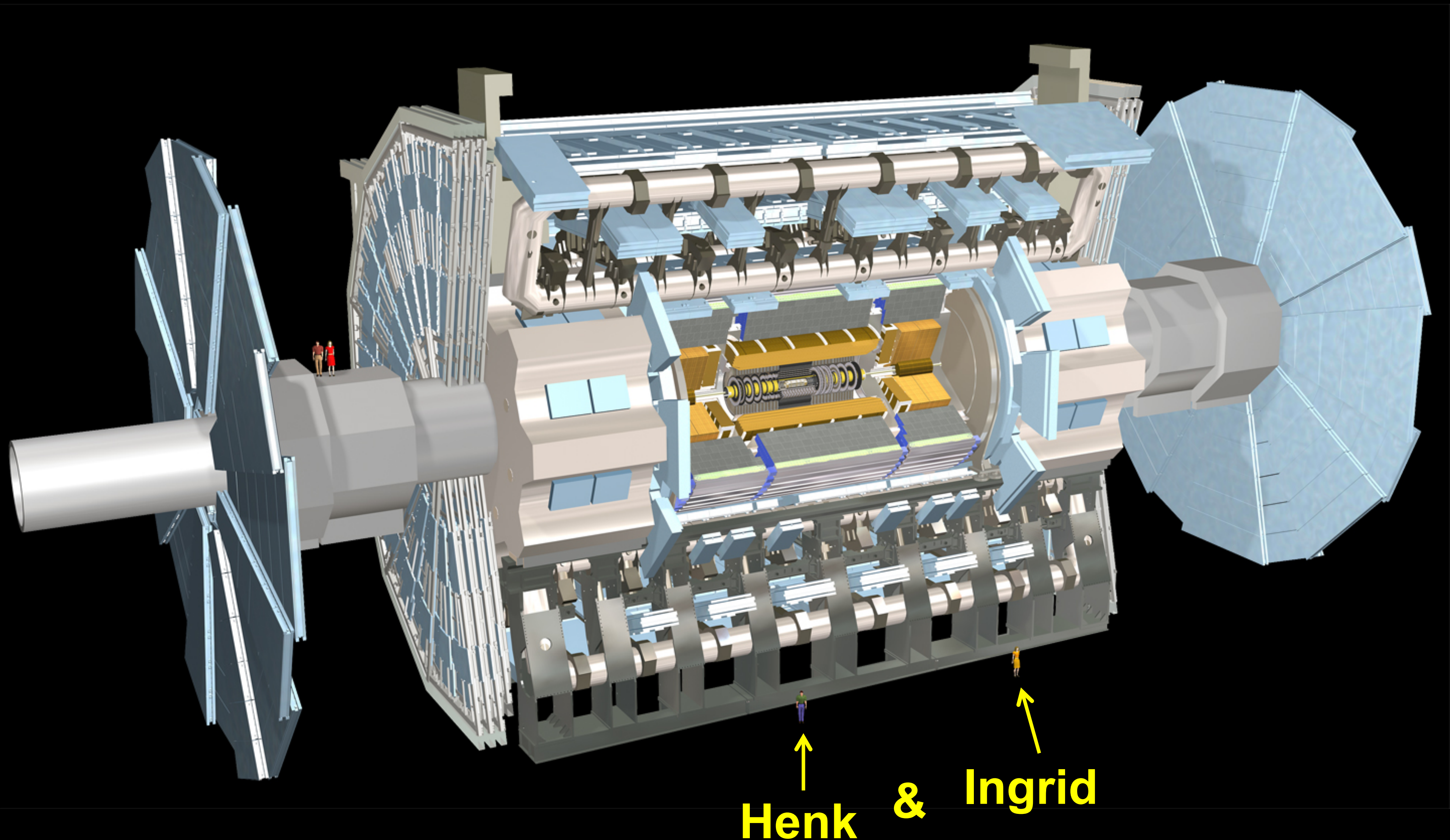
... more data!
Speed
Volume
DATA
Deterministic
Real-time
... more realtime!
Scalable
Secure
... more users!
Ijkdijk/Urban Flood
Medical
SmartGrids/GreenClouds
COMMIT/Astronomy
CineGrid
EU-GN3/NOVI/Geysers
SURFnet/GLIF/Cloud
Sustainability, Green-IT
Complexity
Authorization/Privacy/Trust
Programmable Infrastructure
WorkFlow/QoS
Topology/Architecture
Optical Photonic

SNE @ UvA



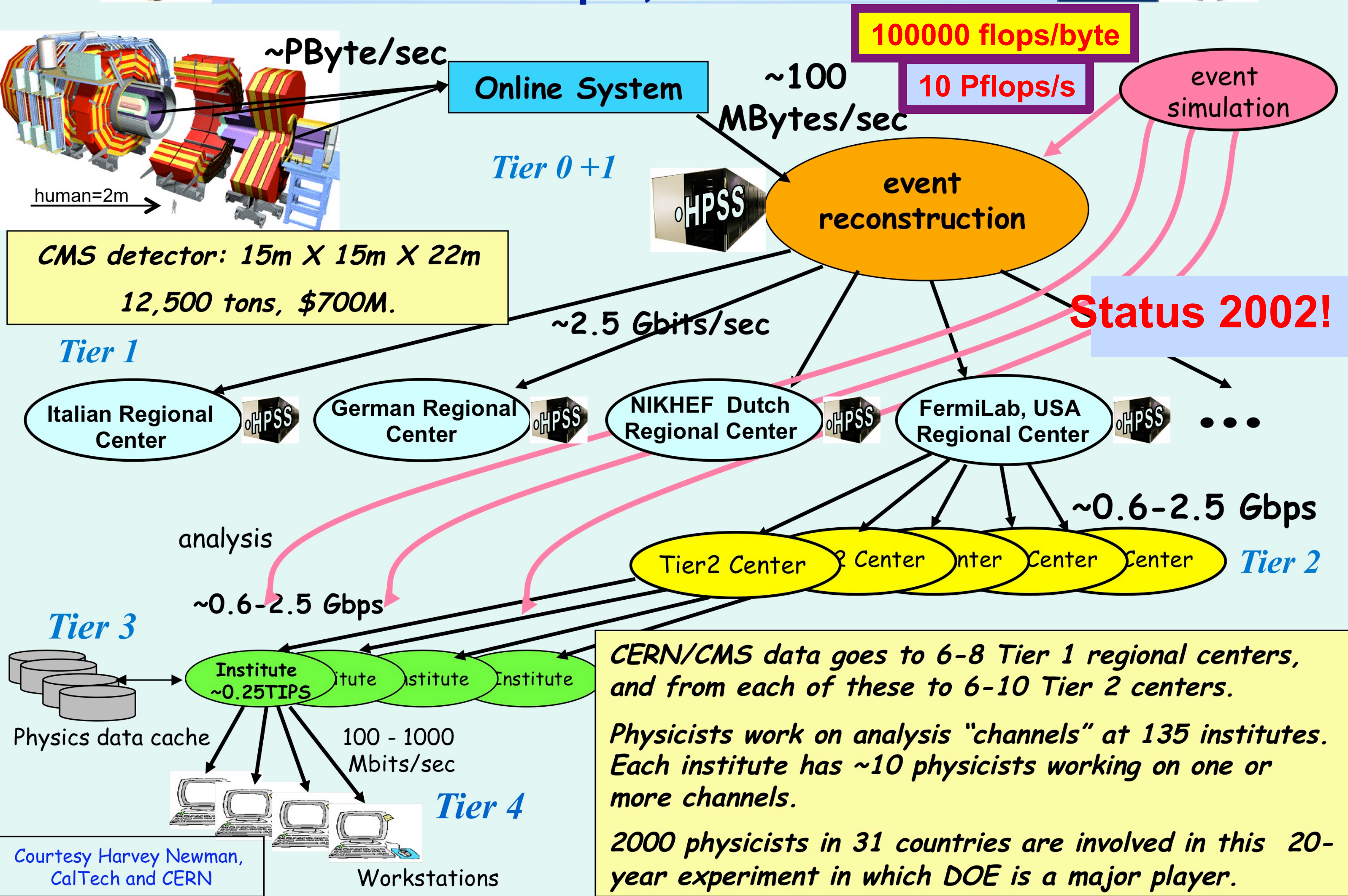
Ijkdijk/Urban Flood
Medical
LifeWatch
CosmoGrid/eVLBI
CineGrid
EU-GN3/NOVI/Geysers
SURFnet/GLIF/Cloud
Green-IT
Privacy/Trust
Authorization/policy
Programmable networks
40-100Gig/TCP/WF/QoS
Topology/Architecture
Optical Photonic

ATLAS detector @ CERN Geneve



ATLAS detector @ CERN Geneve

LHC Data Grid Hierarchy
CMS as example, Atlas is similar

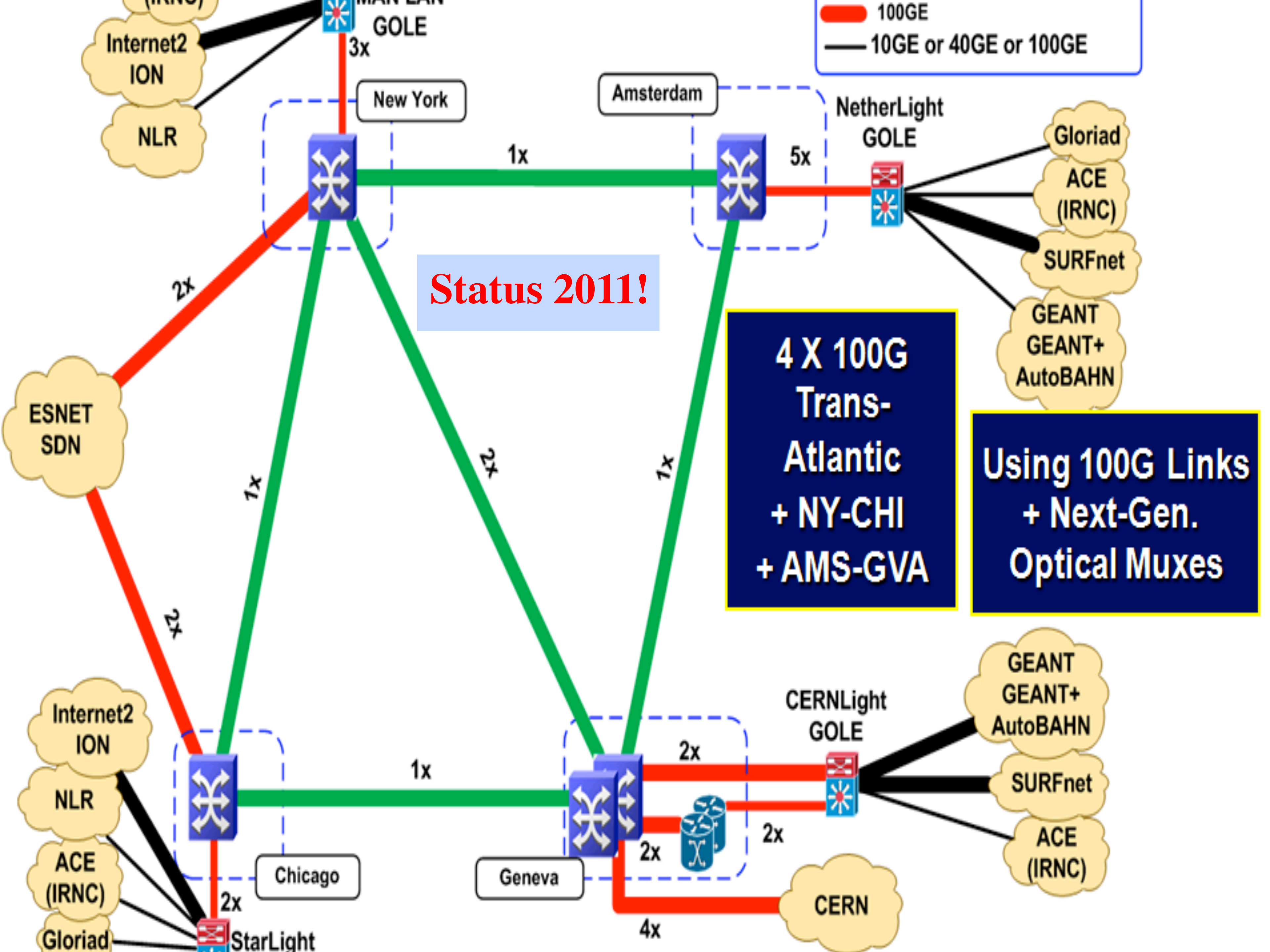


CERN/CMS data goes to 6-8 Tier 1 regional centers, and from each of these to 6-10 Tier 2 centers.

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.

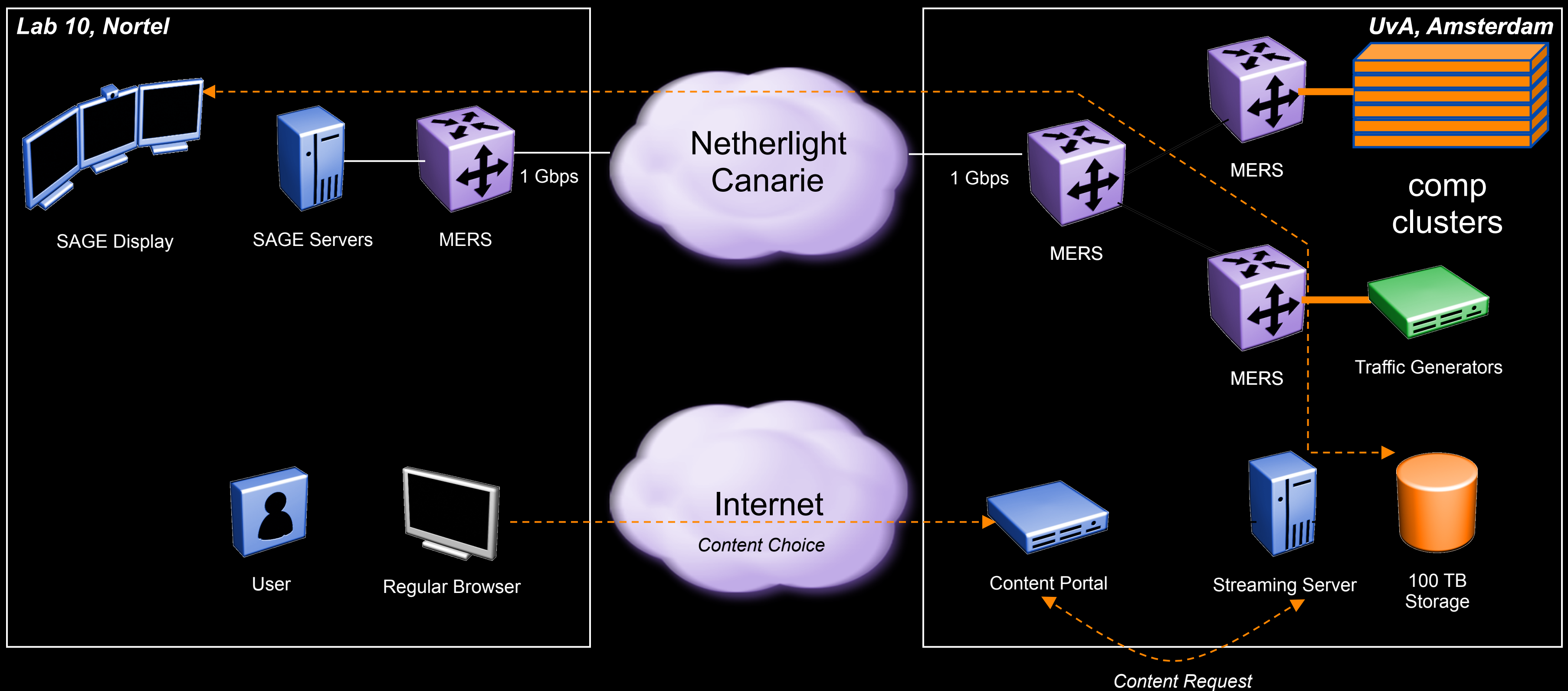
Courtesy Harvey Newman, CalTech and CERN



100GE
10GE or 40GE or 100GE
Internet2 ION
NLR
GOLE
3x
New York
Amsterdam
NetherLight GOLE
1x
5x
Gloriad
ACE (IRNC)
SURFnet
GEANT GEANT+ AutoBAHN
Status 2011!
2x
ESNET SDN
1x
2x
1x
4 X 100G Trans-Atlantic + NY-CHI + AMS-GVA
Using 100G Links + Next-Gen. Optical Muxes
2x
Internet2 ION
NLR
ACE (IRNC)
Gloriad
1x
CERNLight GOLE
GEANT GEANT+ AutoBAHN
SURFnet
ACE (IRNC)
2x
2x
2x
Chicago
Geneva
2x
StarLight
CERN
4x

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

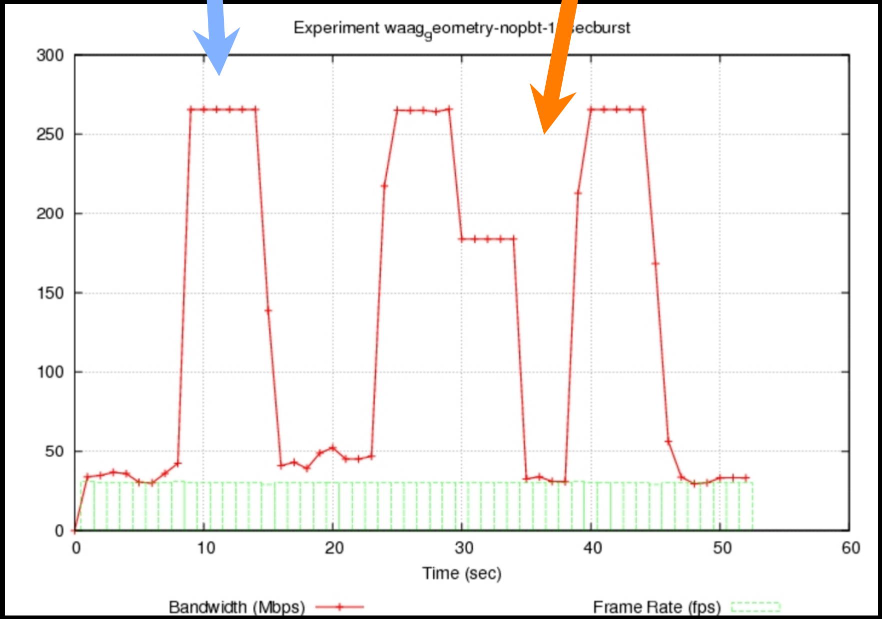
Diagram for SAGE video streaming to ATS



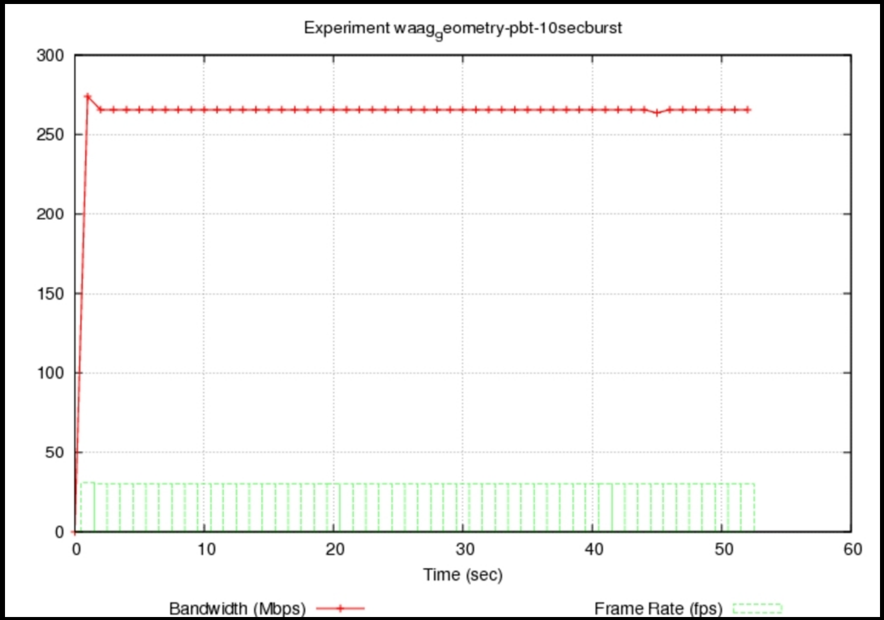
Experimental Data

Sage without background traffic

Sage with background traffic



10 Second Traffic bursts with No PBT



10 Second Traffic bursts with PBT

PBT is *SIMPLE* and *EFFECTIVE* 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

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.

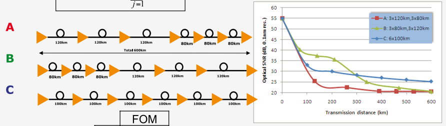
In this demonstration we will investigate the performance of a 40Gb/s PM-QPSK alien wavelength installed on a 10Gb/s DWDM infrastructure.

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

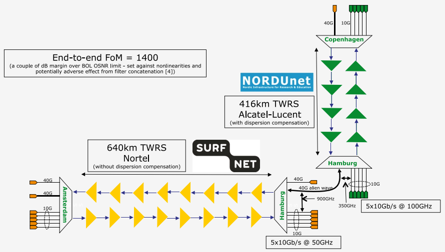


	FOM
A	5504
B	5504
C	1897

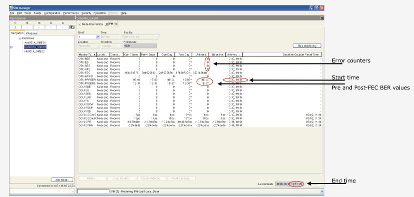
Easy-to-use formula that accurately quantifies transmission system performance

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.

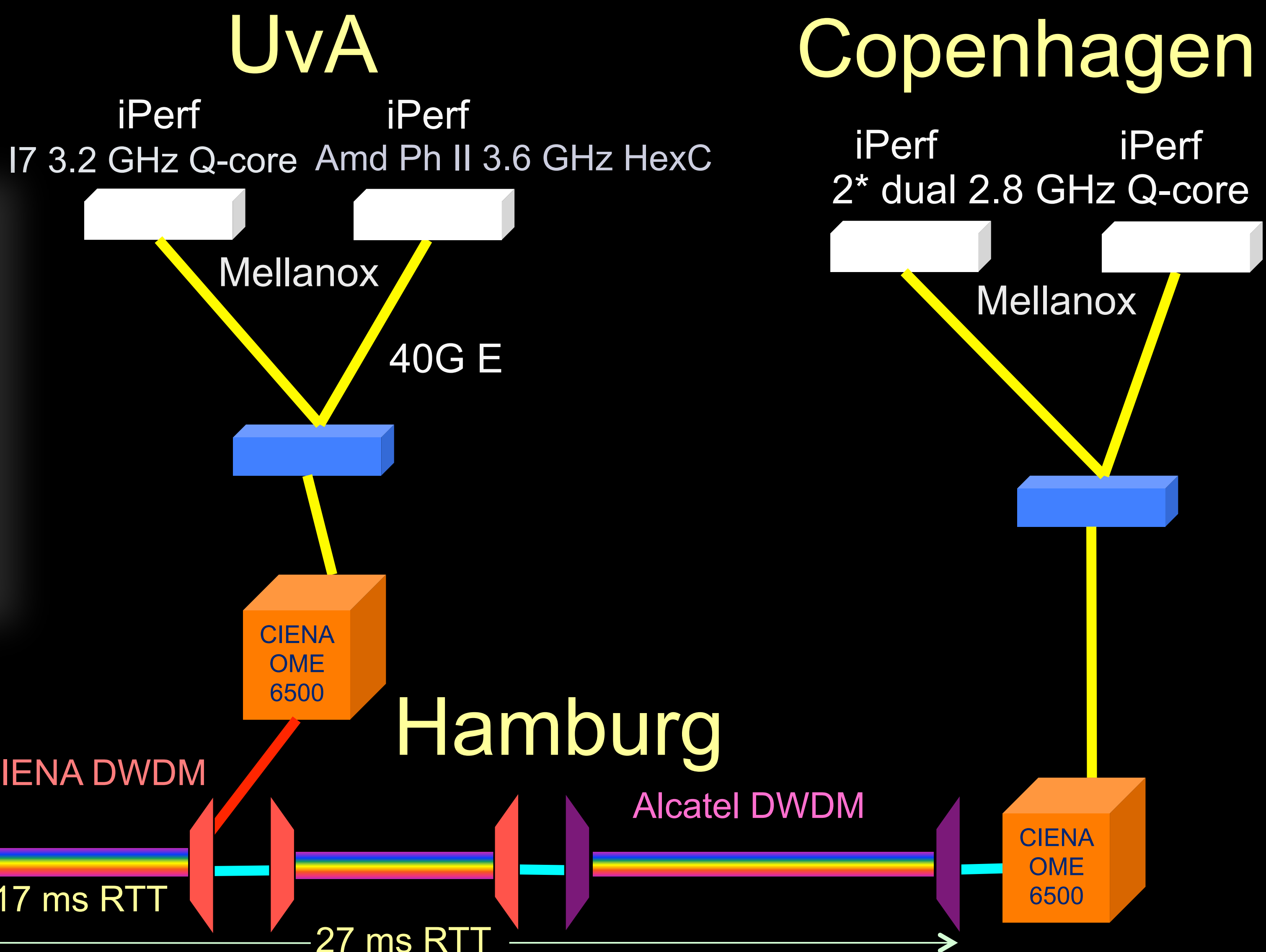
NORTEL | NORDUnet Nordic Infrastructure for Research & Education | telindus belgacom ICT together with ISIT THE STORAGE COMPANY | SURF NET

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

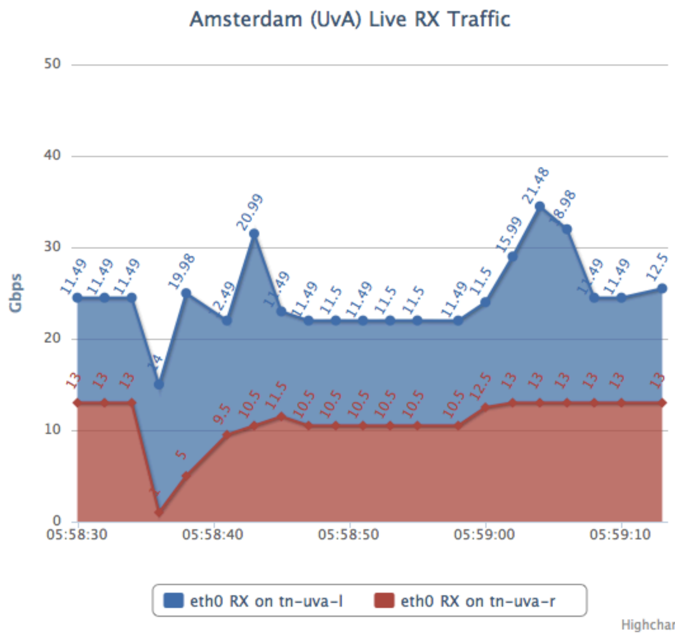


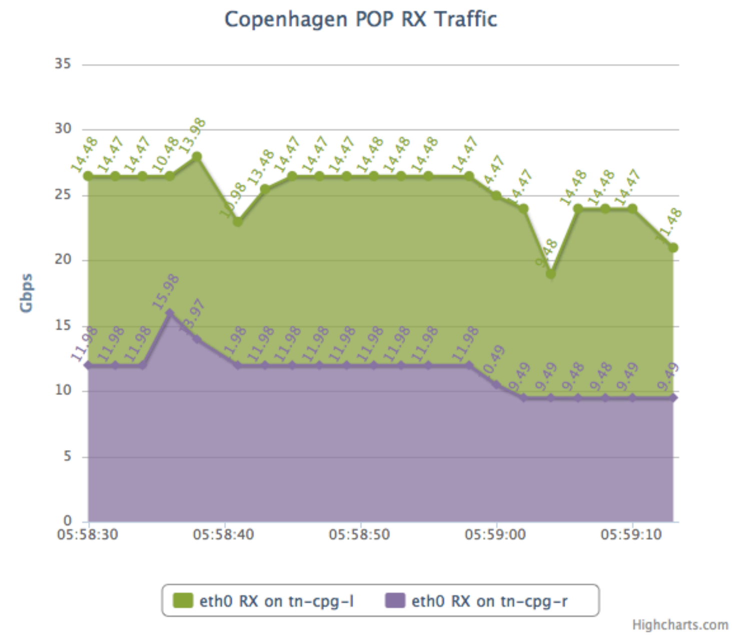
Amsterdam – Geneva (CERN) – Copenhagen – 4400 km (2700 km alien light)

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

ClearStream

End-to-End Ultra Fast Transmission Over a Wide Area 40 Gbit/s Lambda



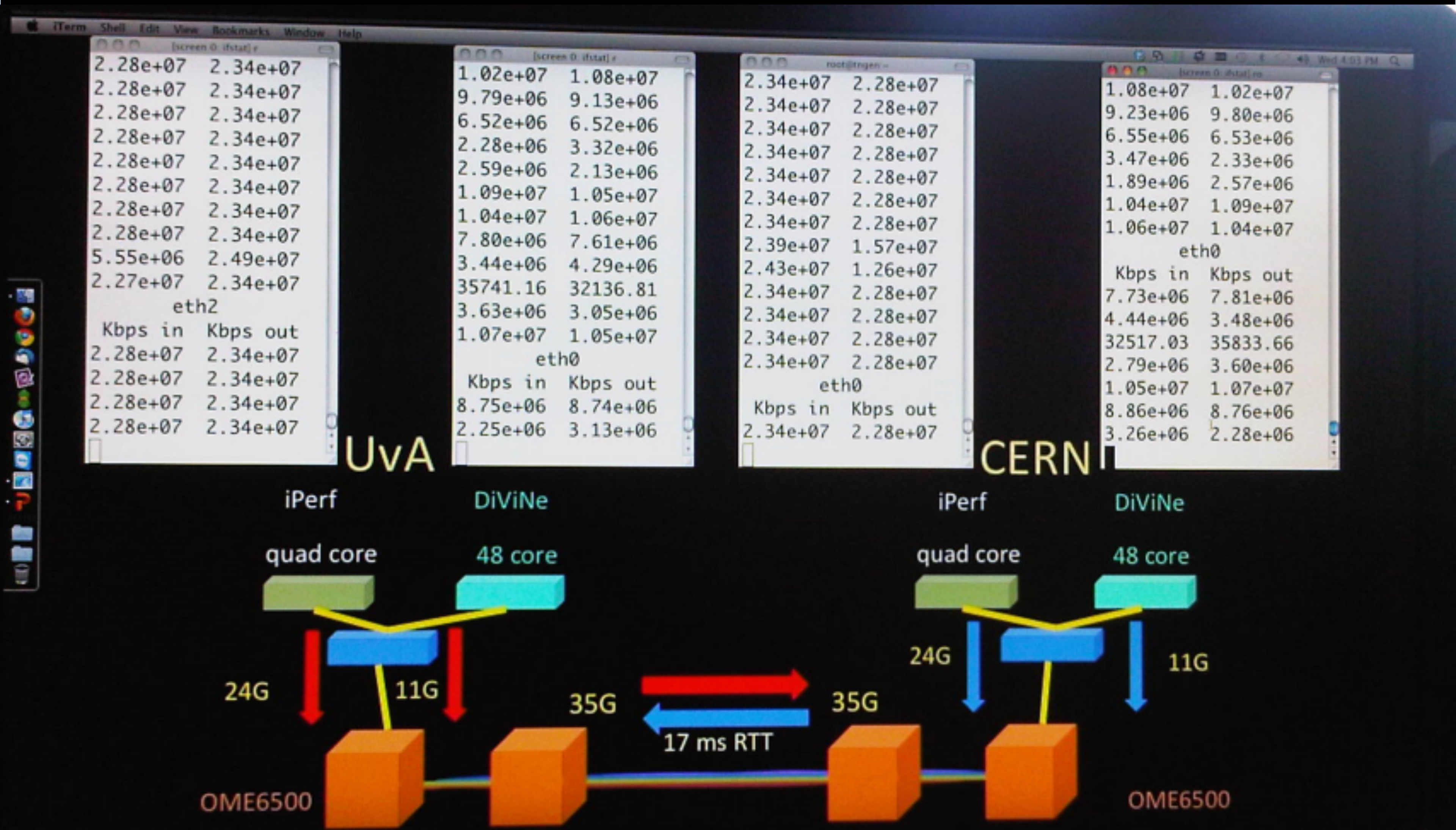


Incoming Amsterdam 25.5 Gbps

Incoming Copenhagen 20.97 Gbps

Total Throughput 46.47 Gbps RTT 44.032 ms

From GLIF October 2010 @ CERN



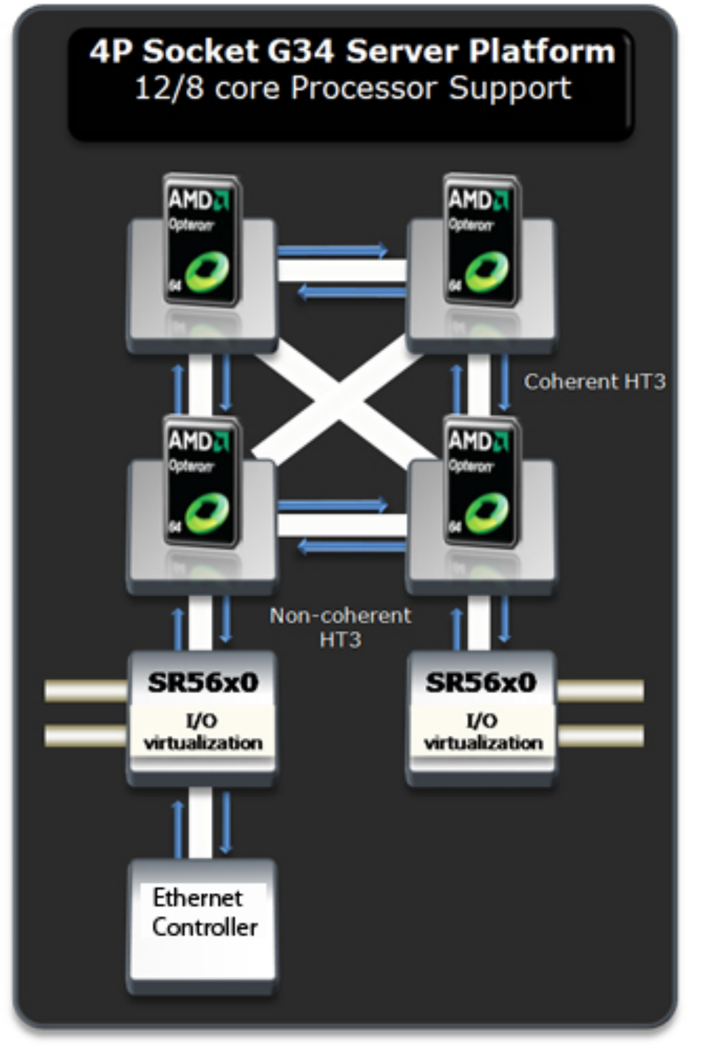
Results (rtt = 17 ms)

- 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 <> -> 11 Gbps
- Multi flow iPerf + DiVine -> 35 Gbps
- Multi flow iPerf + DiVine <> -> 35 + 35 Gbps

Performance Explained

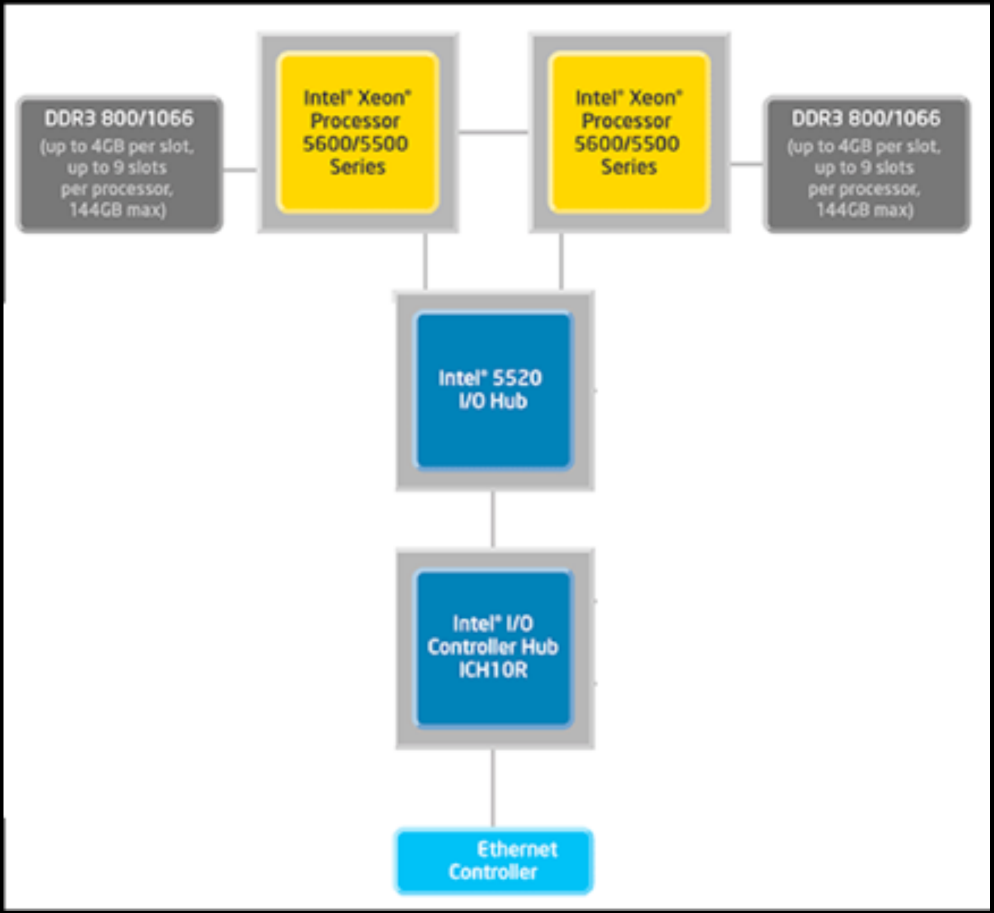
- Mellanox 40GE card is PCI-E 2.0 8x (5GT/s)
- 40Gbit/s raw throughput but ….
- PCI-E is a network-like protocol
  - 8/10 bit encoding -> 25% overhead -> 32Gbit/s maximum data throughput
  - Routing information
- Extra overhead from IP/Ethernet framing
- Server architecture matters!
  - 4P system performed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

Server Architecture



DELL R815

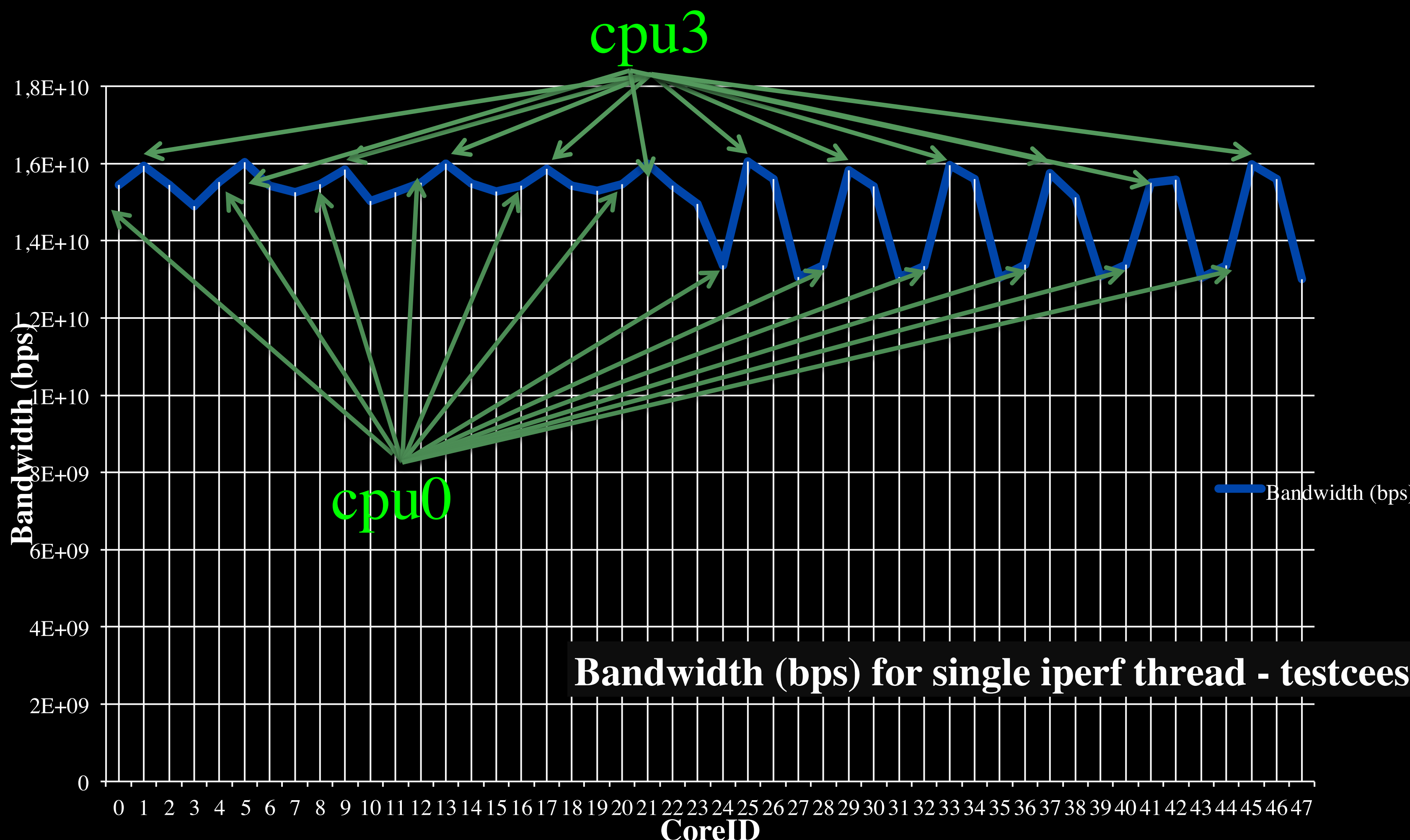
4 x AMD Opteron 6100



Supermicro X8DTT-HIBQF

2 x Intel Xeon

CPU Topology benchmark



We used numactl to bind iperf to cores

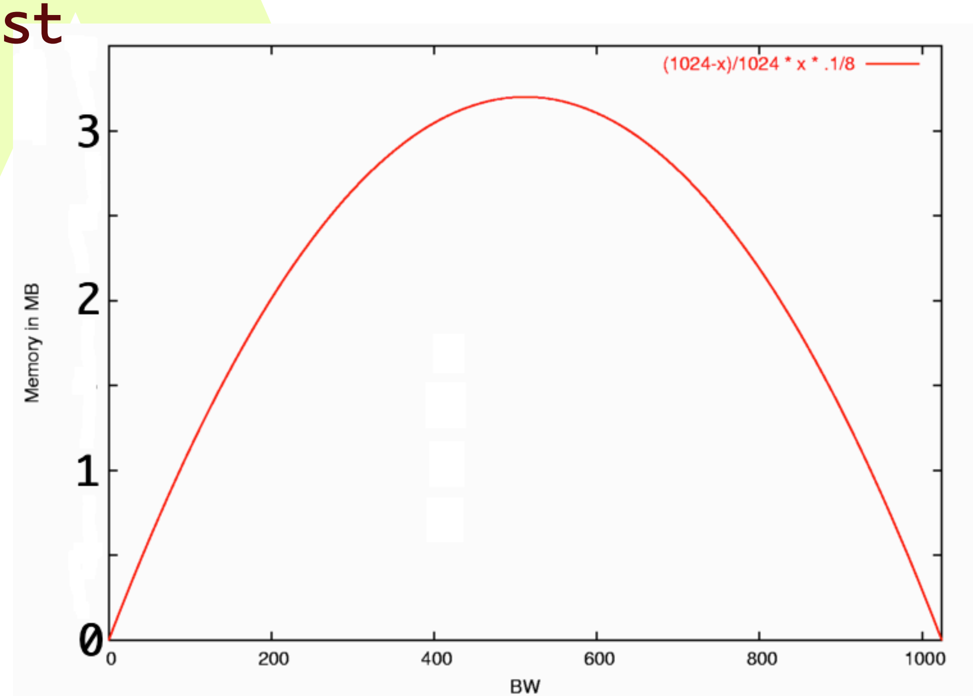
Layer - 2 requirements from 3/4

TCP is bursty due to sliding window protocol and slow start algorithm.

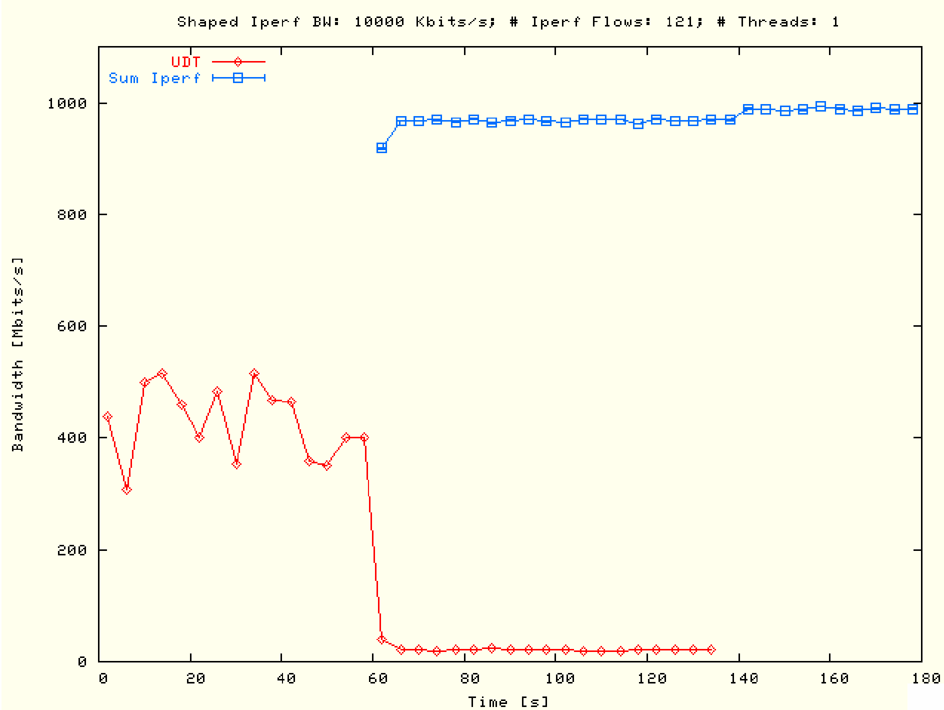
```
Window = BandWidth * RTT    &    BW == slow

                        fast - slow
Memory-at-bottleneck = ----------- * slow * RTT
                           fast
```

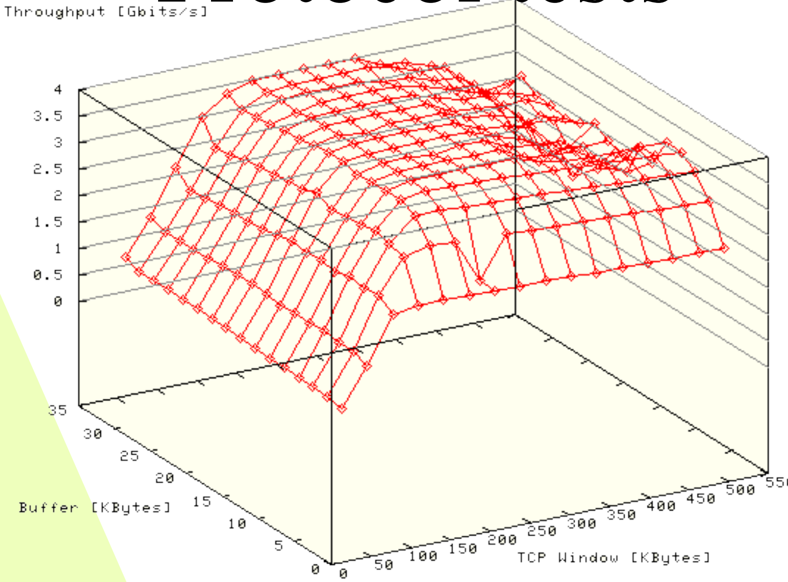
So pick from menu:

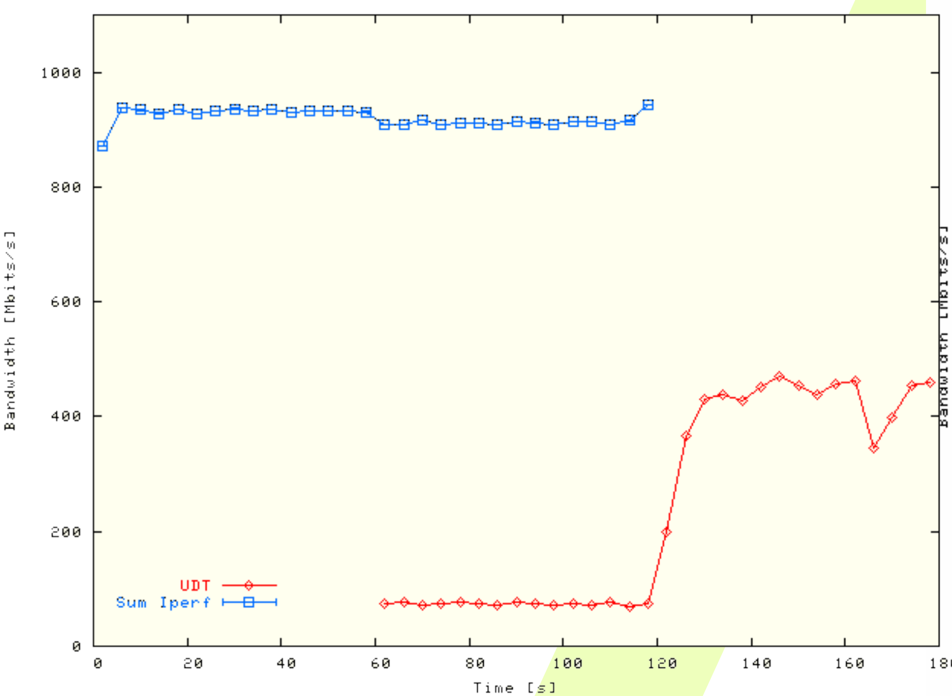
- *Flow control*
- *Traffic Shaping*
- *RED (Random Early Discard)*
- *Self clocking in TCP*
- *Deep memory*

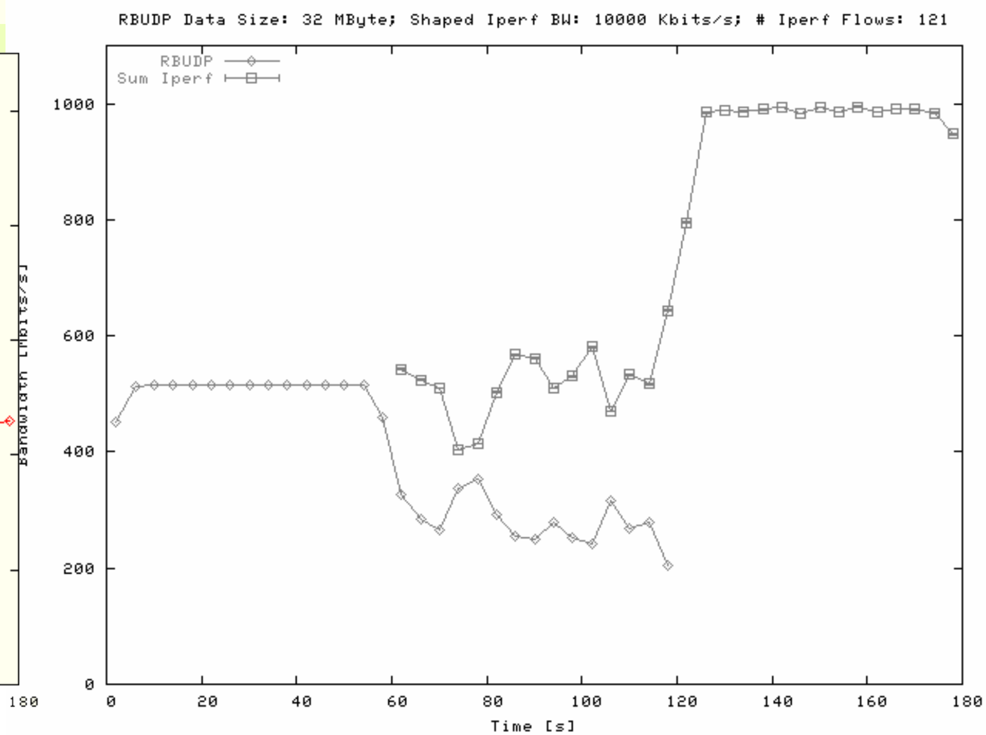


Protocol tests

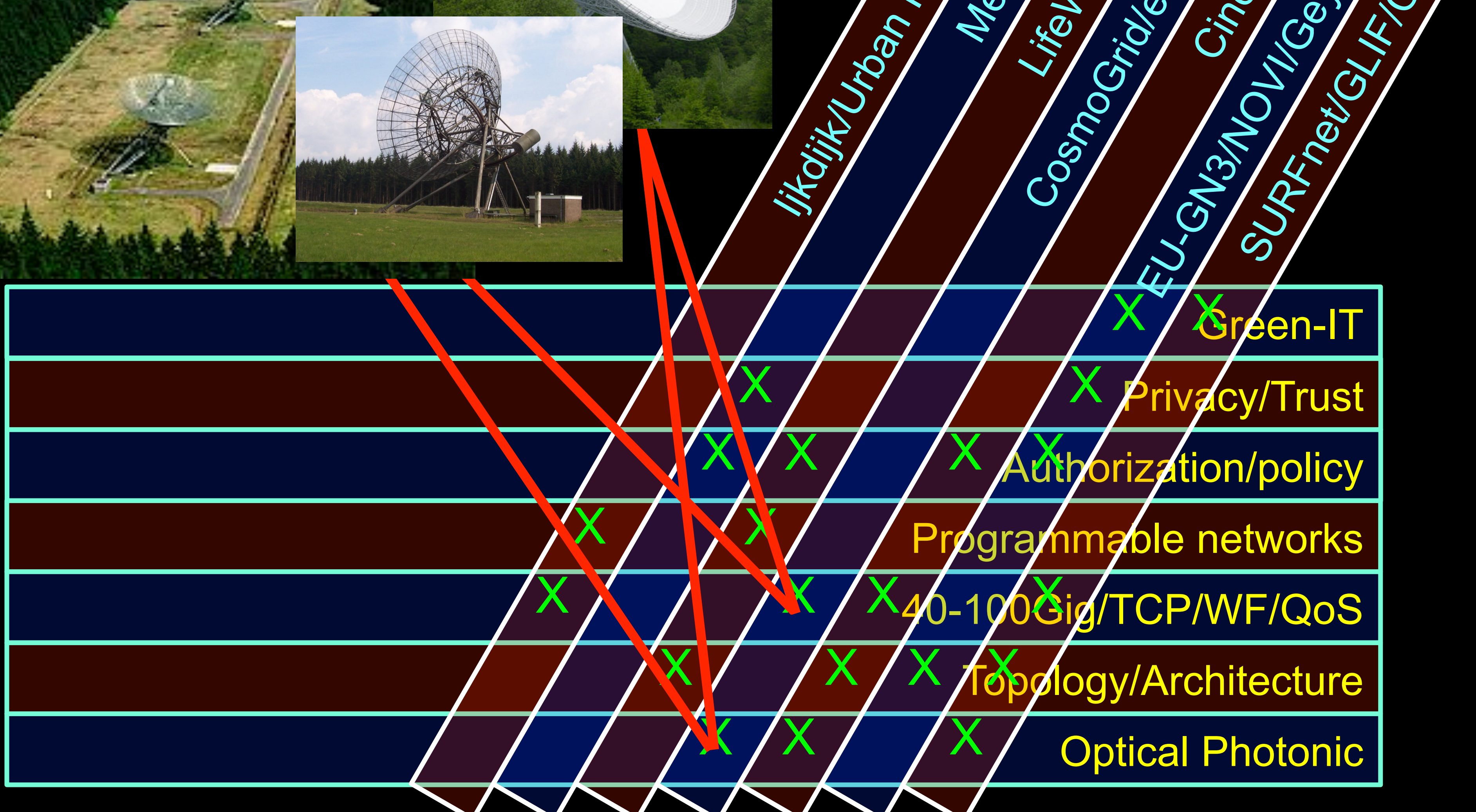






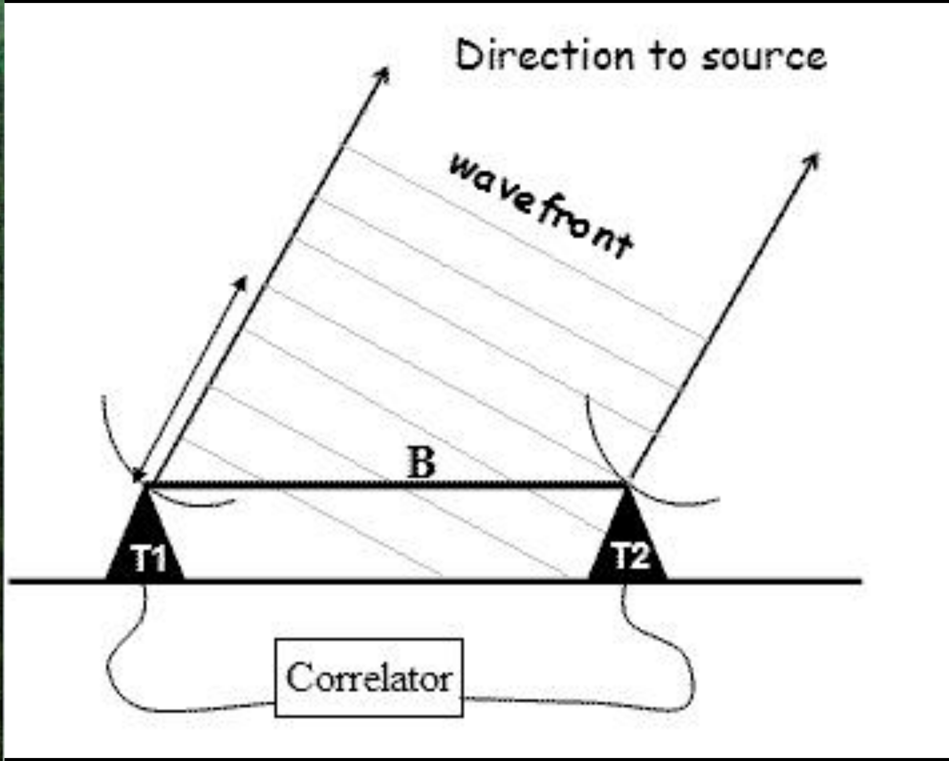


SNE @ UvA

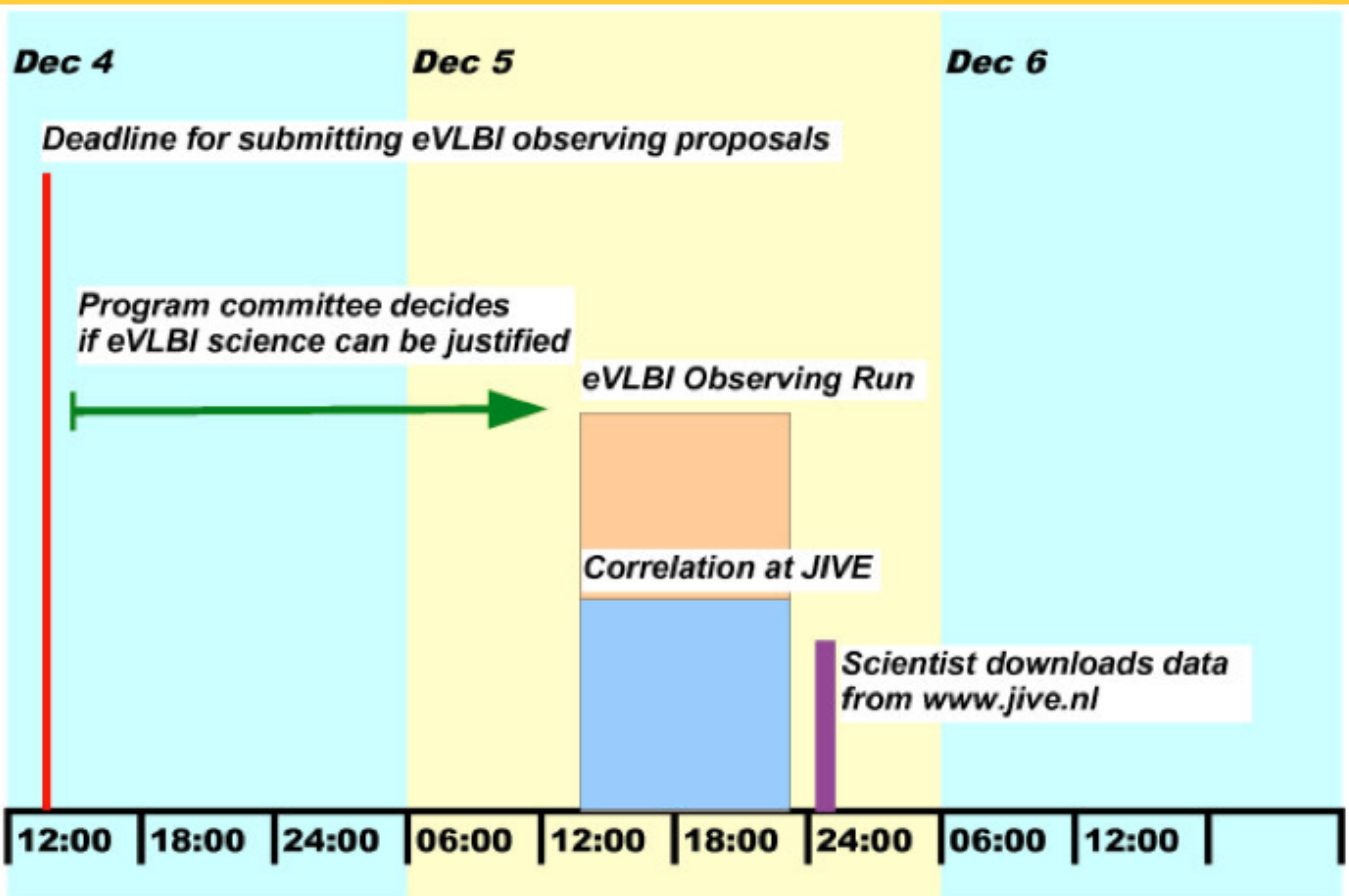


Ijkdijk/Urban Flood
Medical
LifeWatch
CosmoGrid/eVLBI
CineGrid
EU-GN3/NOVI/Geysers
SURFnet/GLIF/Cloud
Green-IT
Privacy/Trust
Authorization/policy
Programmable networks
40-100Gig/TCP/WF/QoS
Topology/Architecture
Optical Photonic

e -Very Large Base Interferometer



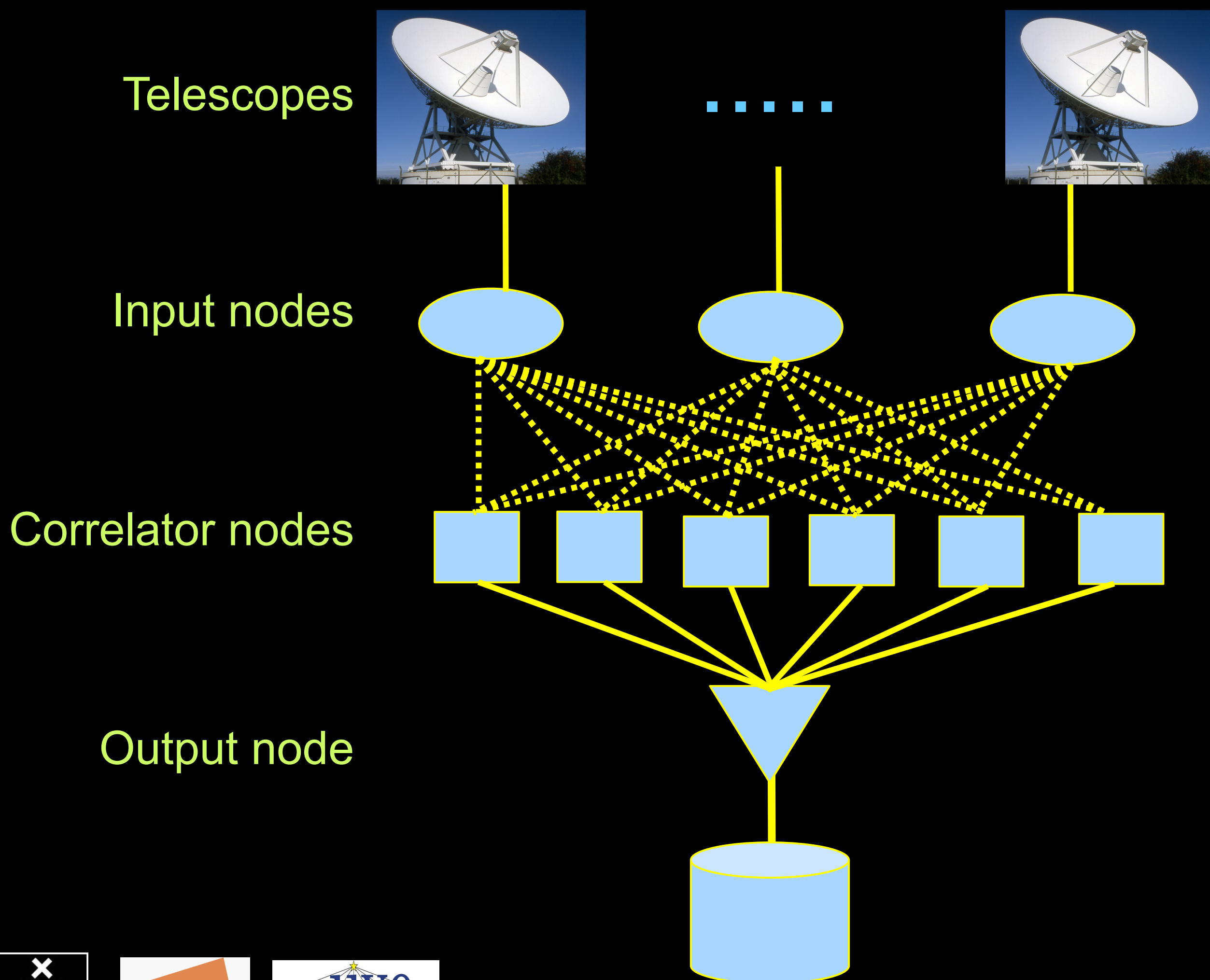
eEVN: European VLBI Network



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:

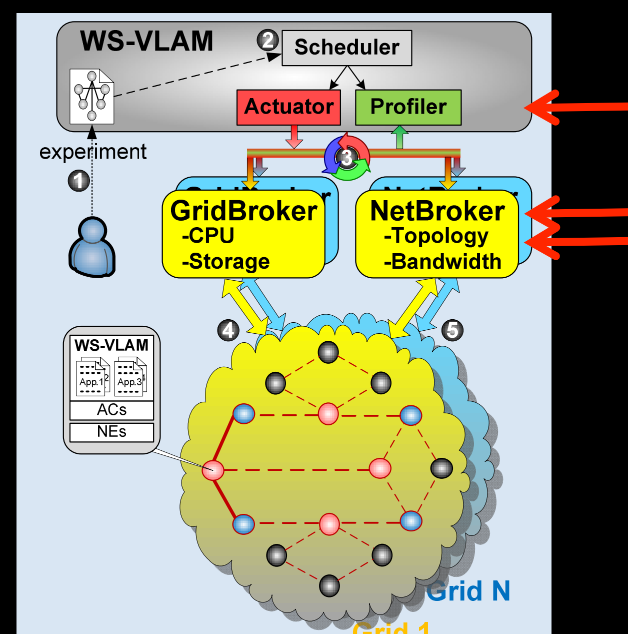


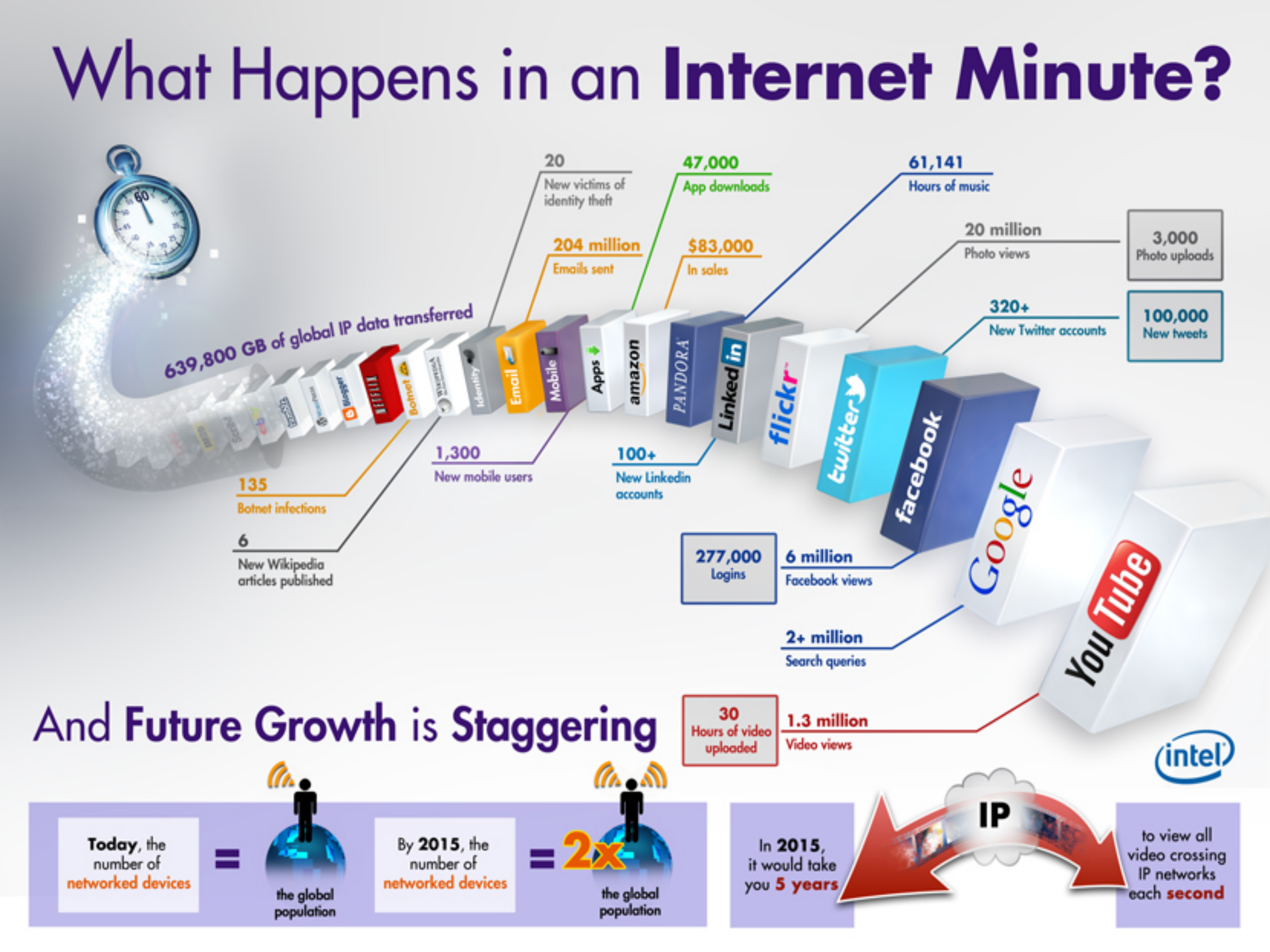
Figure 2. Grid architecture that includes programmable network services.

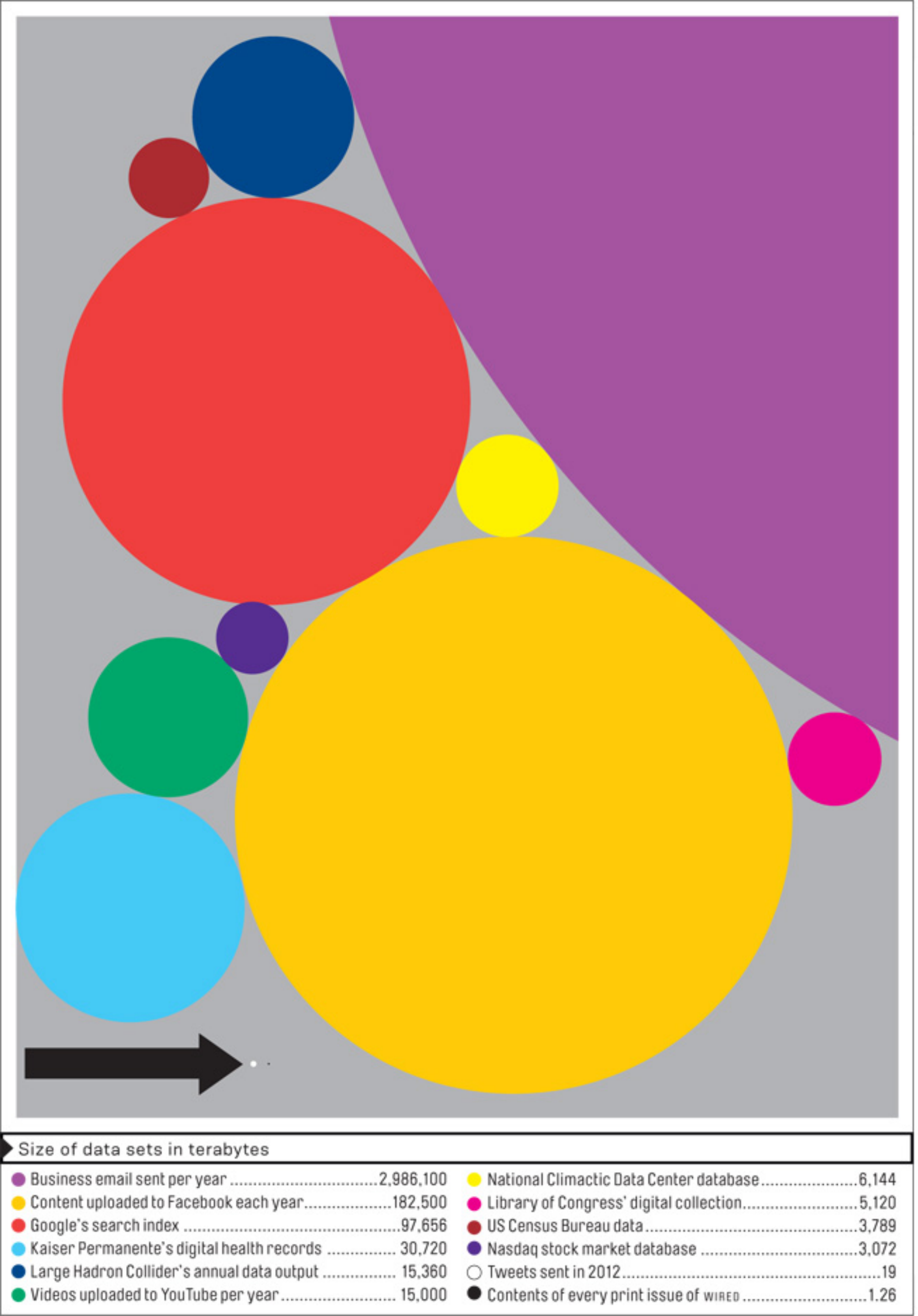
LOFAR as a Sensor Network

20 flops/byte

2 Tflops/s

- LOFAR is a large distributed research infrastructure:
  - 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

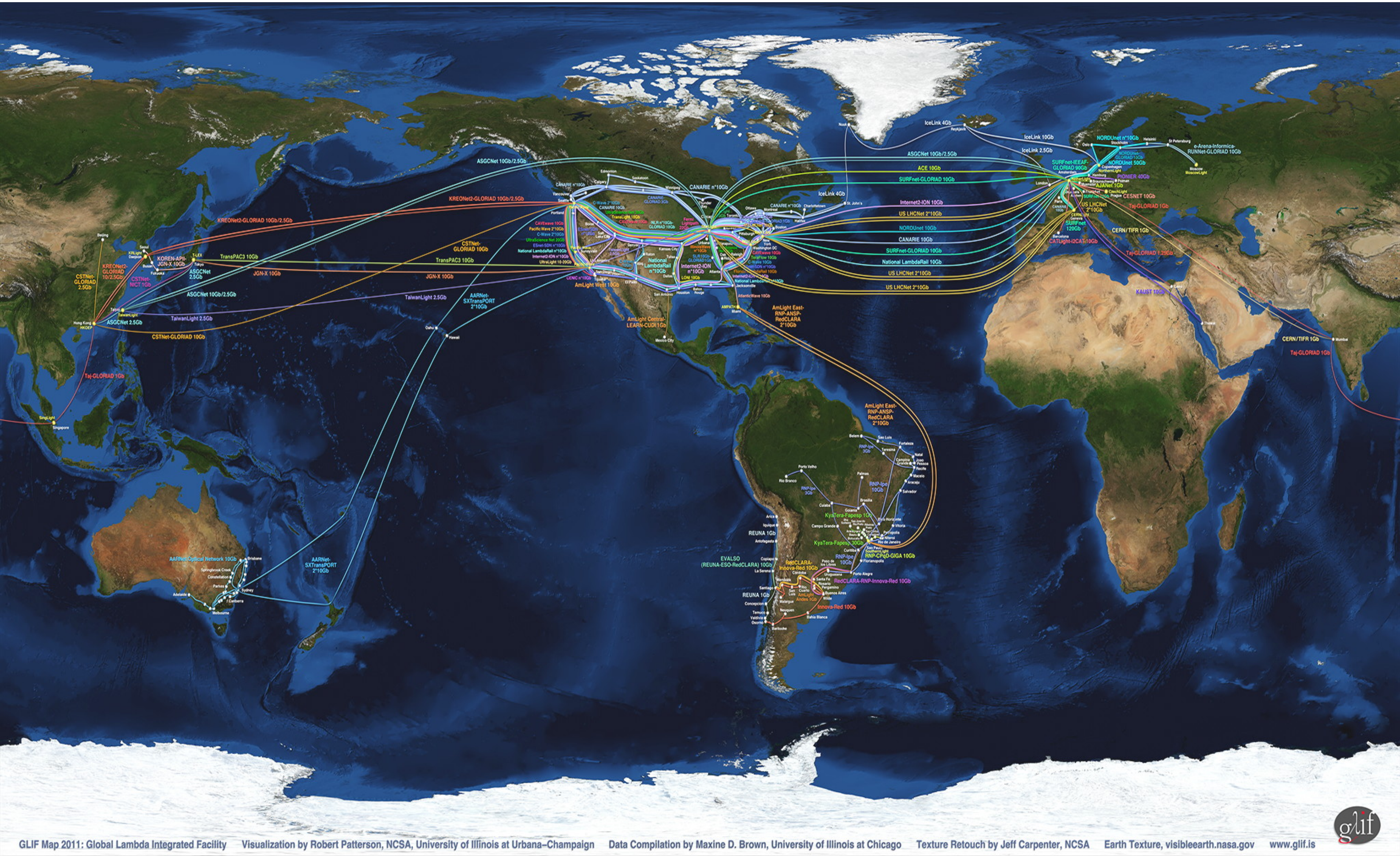

What Happens in an Internet Minute?
639,800 GB of global IP data transferred
20
New victims of identity theft
204 million
Emails sent
47,000
App downloads
$83,000
In sales
61,141
Hours of music
20 million
Photo views
3,000
Photo uploads
320+
New Twitter accounts
100,000
New tweets
135
Botnet infections
6
New Wikipedia articles published
1,300
New mobile users
100+
New Linkedin accounts
277,000
Logins
6 million
Facebook views
2+ million
Search queries
30
Hours of video uploaded
1.3 million
Video views
Botnet
Identity
Email
Mobile
Apps
amazon
PANDORA
Linkedin
flickr
twitter
facebook
Google
YouTube
And Future Growth is Staggering
Today, the number of networked devices
=
the global population
By 2015, the number of networked devices
= 2x
the global population
In 2015, it would take you 5 years
IP
to view all video crossing IP networks each second
intel



There is always a bigger fish

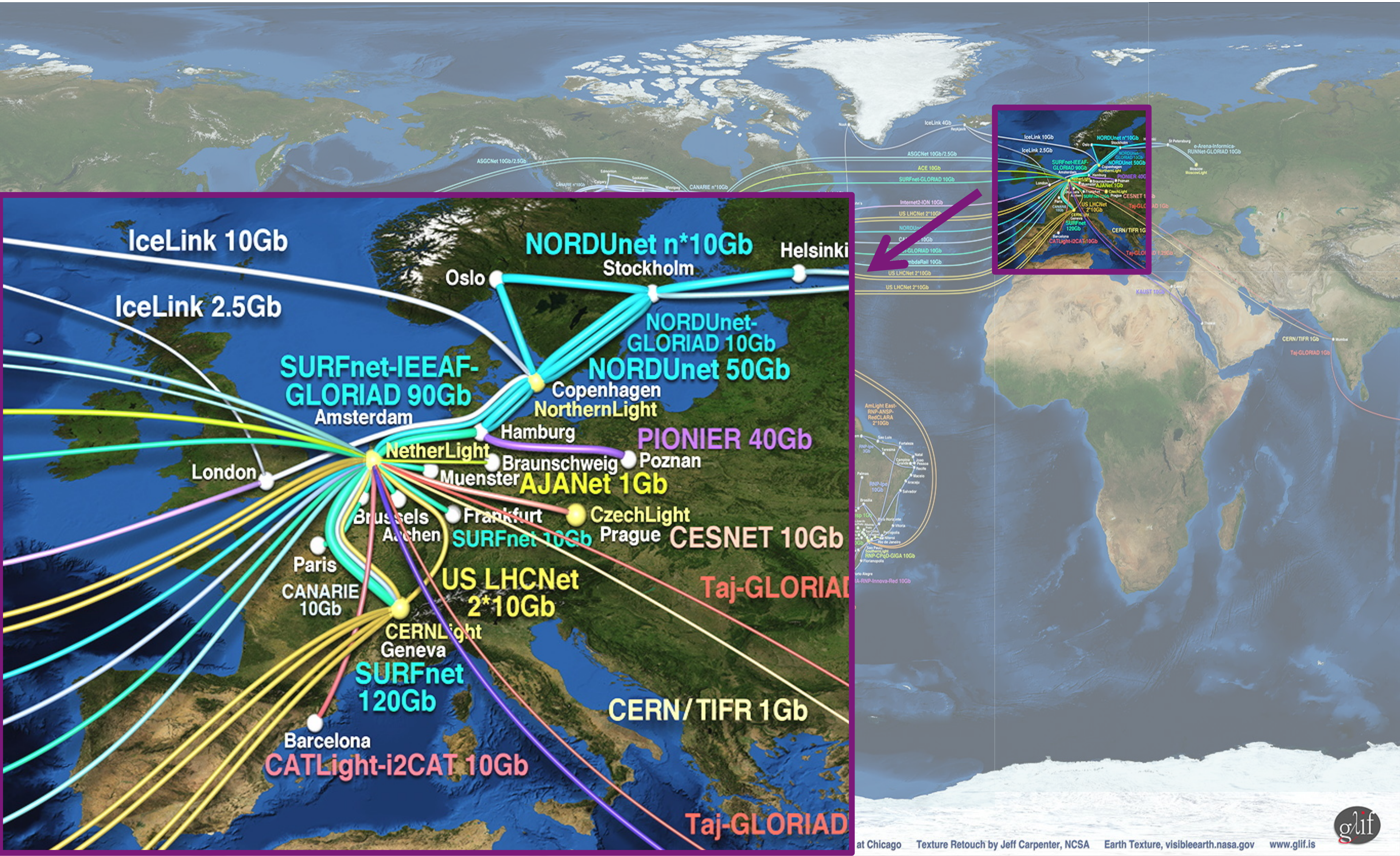
The GLIF – LightPaths around the World

F Dijkstra, J van der Ham, P Grosso, C de Laat, “A path finding implementation for multi-layer networks”, Future Generation Computer Systems 25 (2), 142-146.

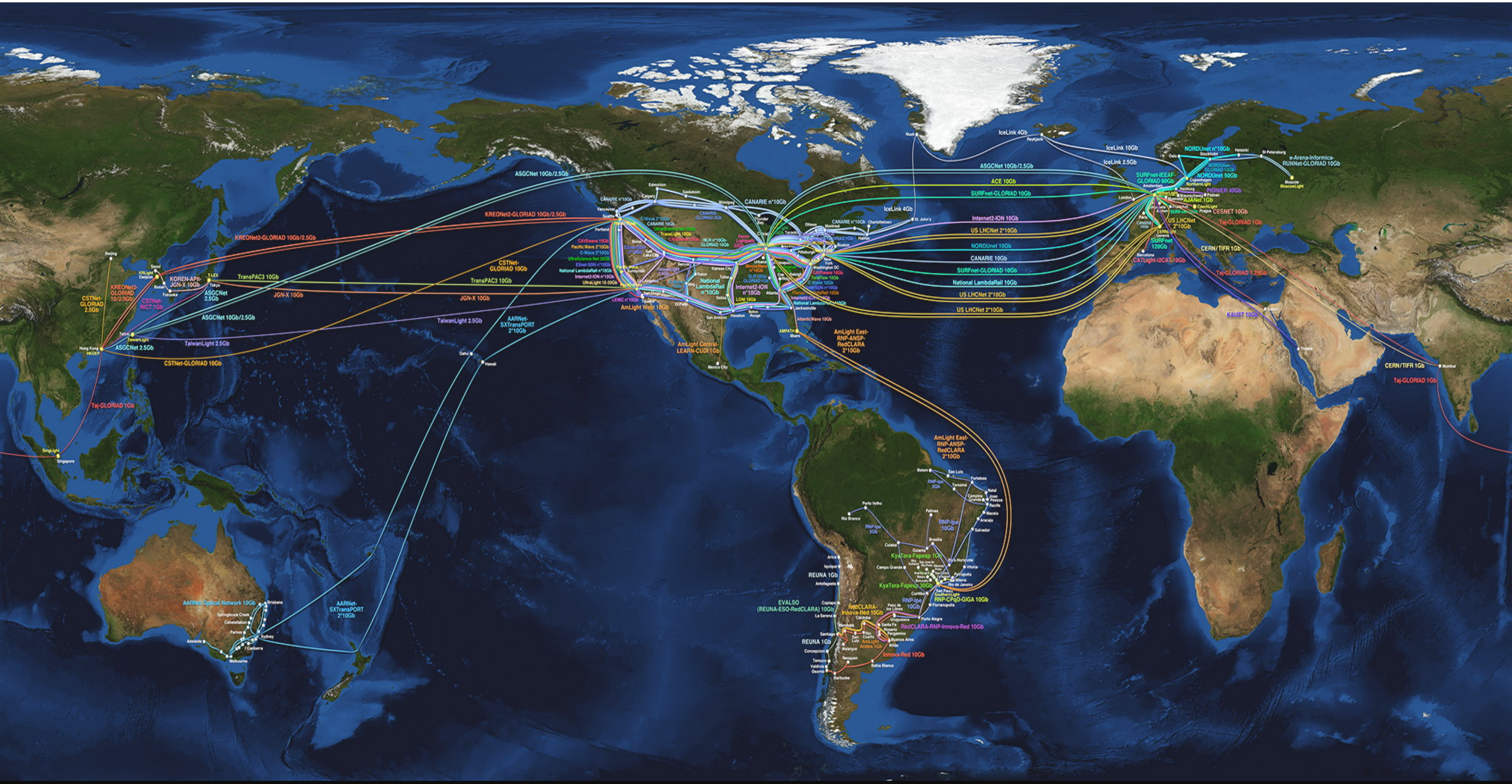


The GLIF – LightPaths around the World

F Dijkstra, J van der Ham, P Grosso, C de Laat, “A path finding implementation for multi-layer networks”, Future Generation Computer Systems 25 (2), 142-146.



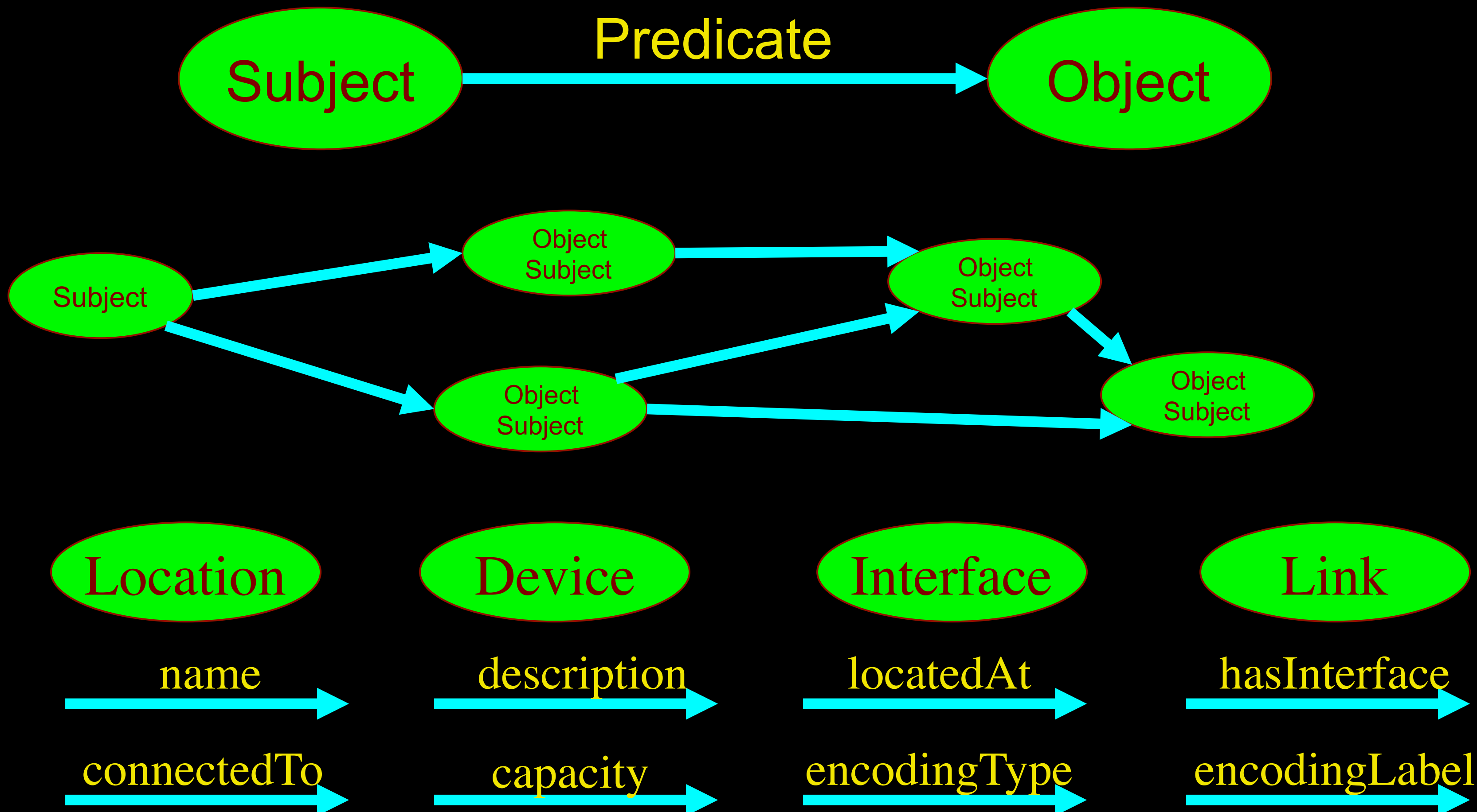
The GLIF – LightPaths around the World



We investigate: TOMTOM 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):

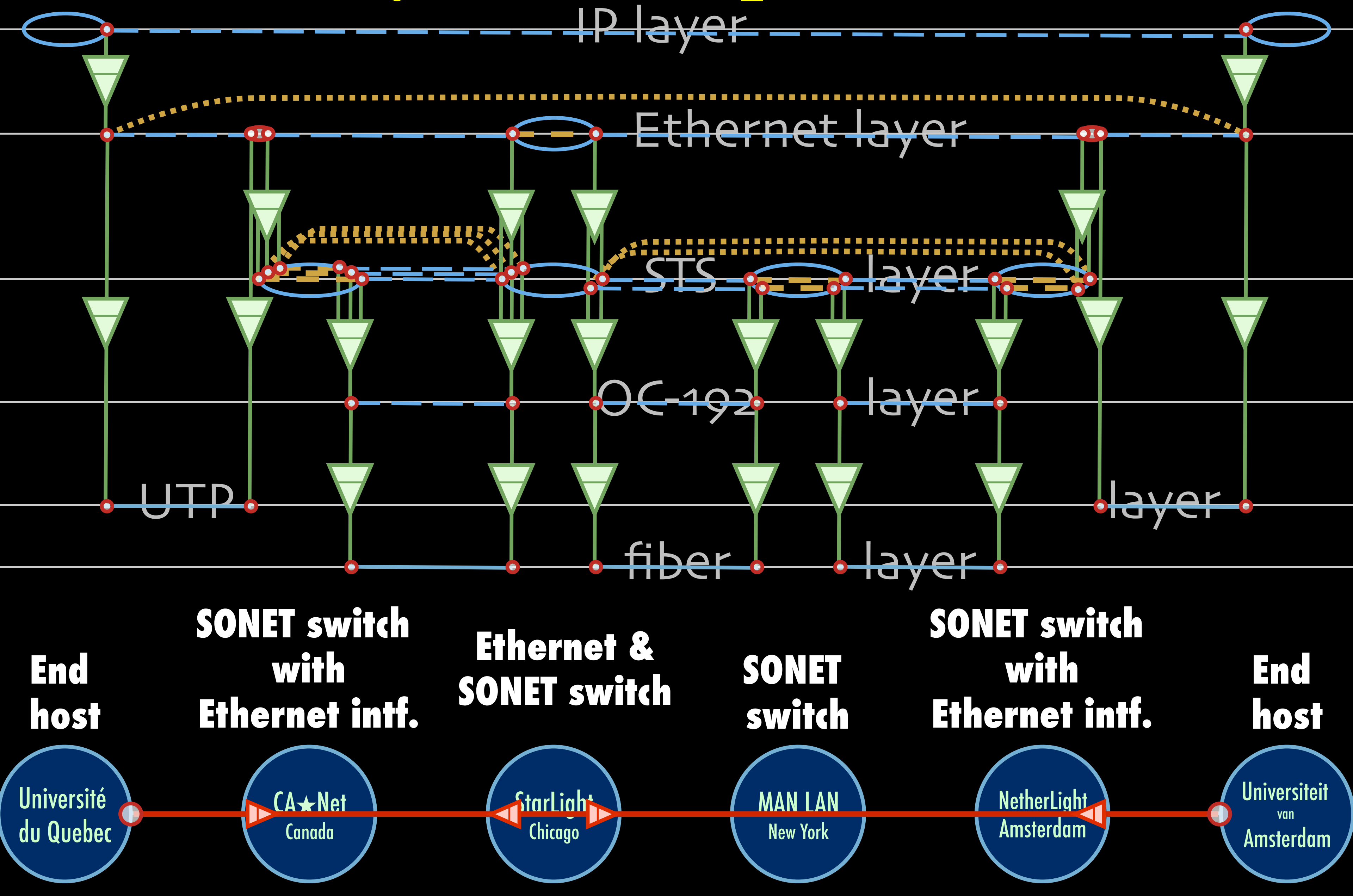


NetherLight in RDF

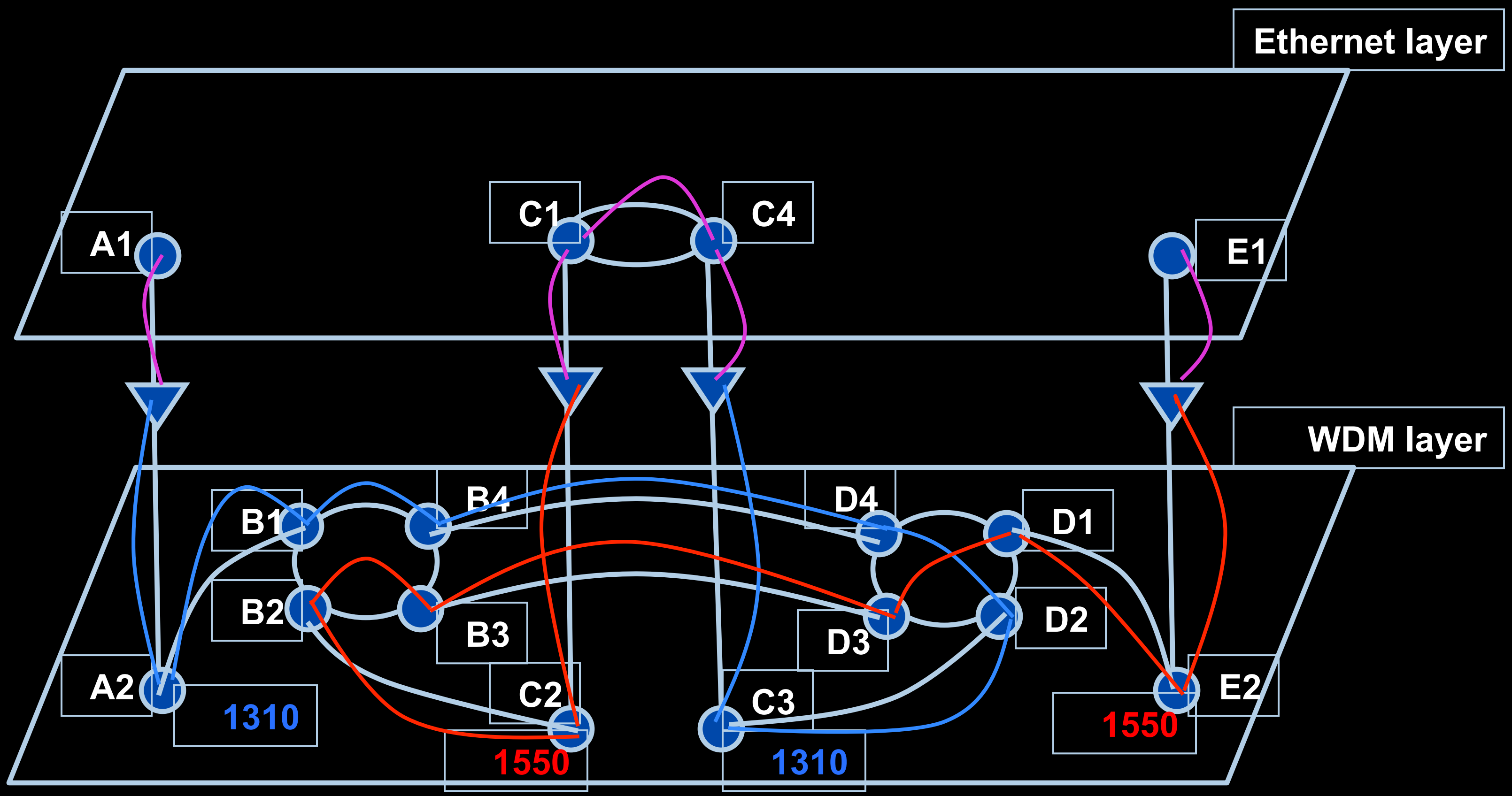
```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
      xmlns:ndl="http://www.science.uva.nl/research/air/ndl#">
<!-- Description of Netherlight -->
<ndl:Location rdf:about="#Netherlight">
      <ndl:name>Netherlight Optical Exchange</ndl:name>
</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
      <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



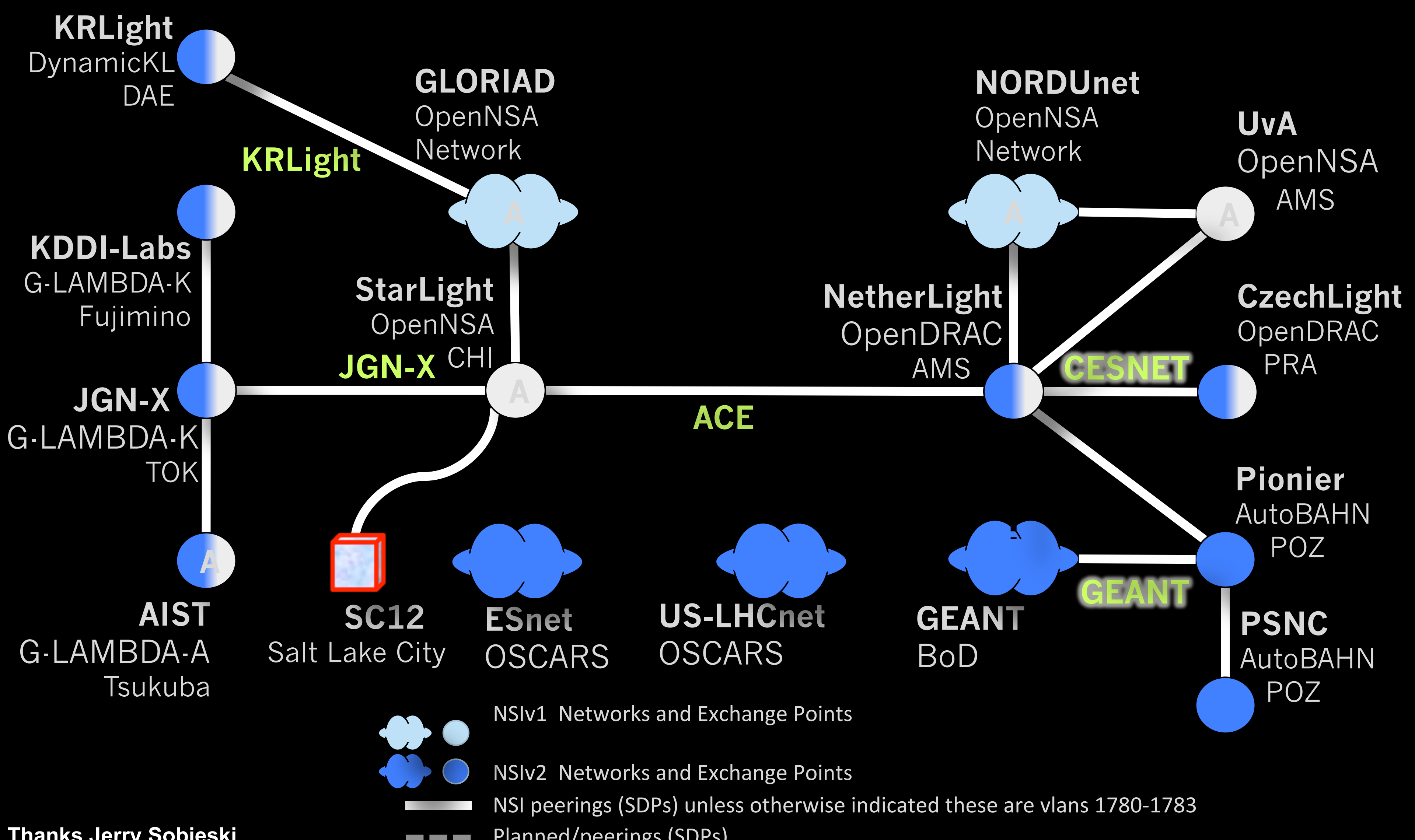
Path between interfaces A1 and E1:
A1-A2-B1-B4-D4-D2-C3-C4-C1-C2-B2-B3-D3-D1-E2-E1

Scaling: Combinatorial problem

Automated GOLE + NSI

Joint NSI v1+v2 Beta Test Fabric Nov 2012

Ethernet Transport Service

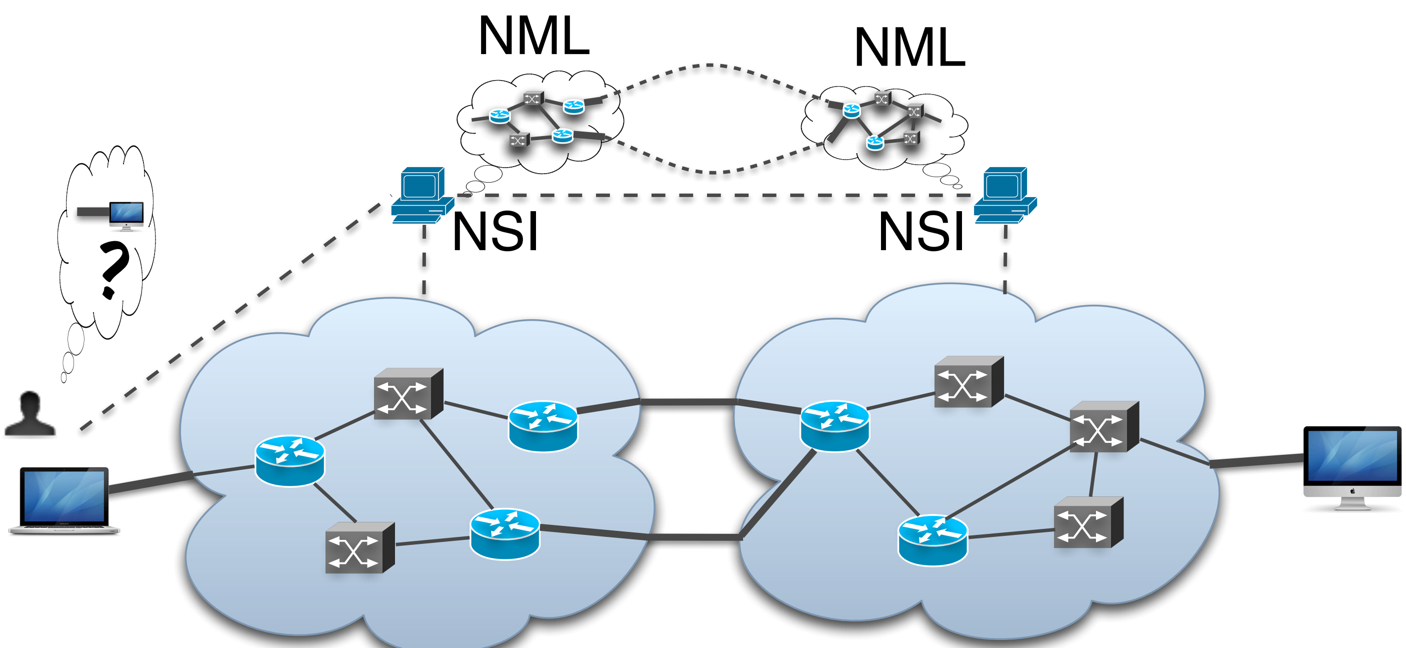


Thanks Jerry Sobieski

Network Topology Description

Network topology research supporting automatic network provisioning

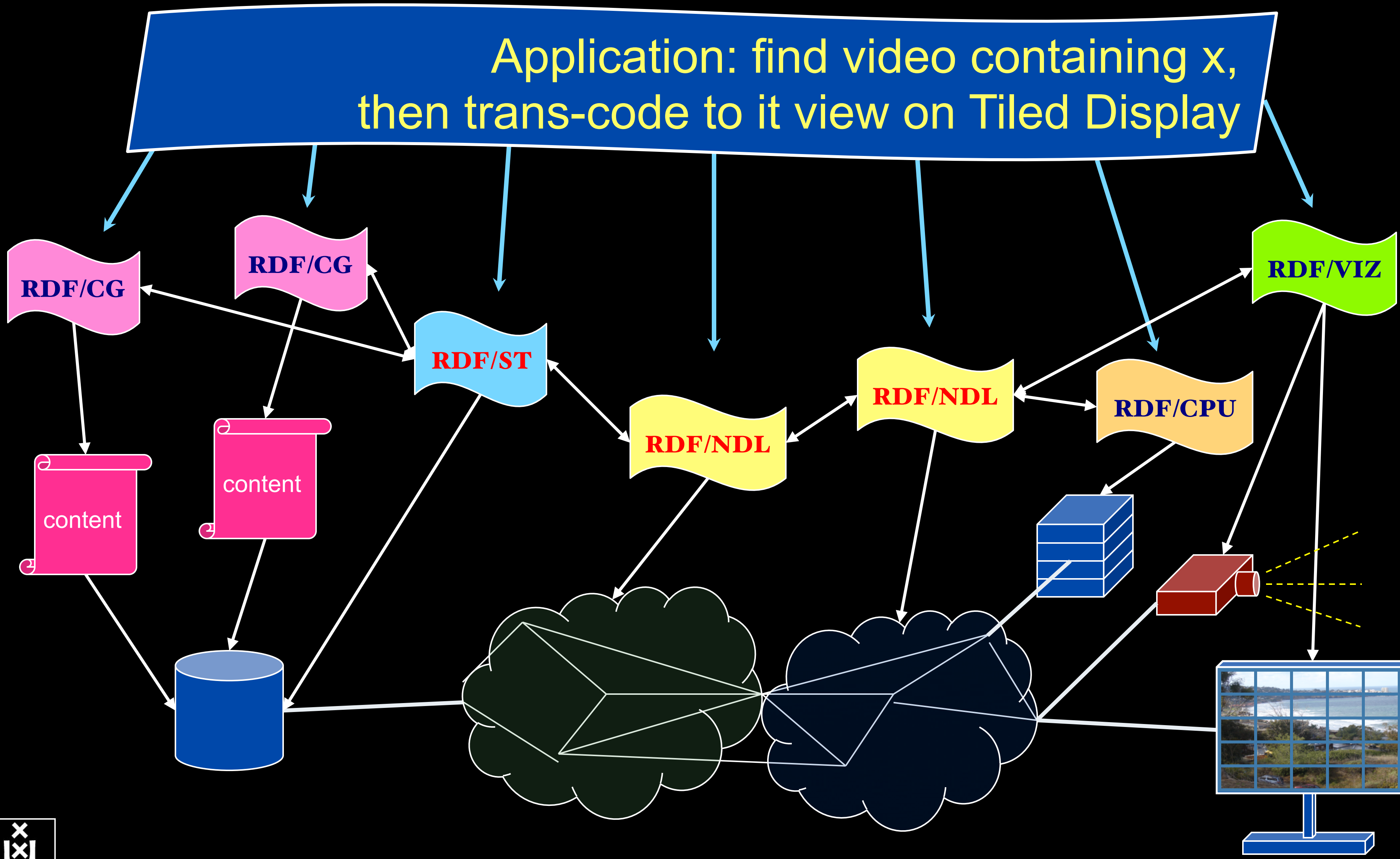
- Inter-domain networks
- Multiple technologies
- Based on incomplete information
- Possibly linked to other resources



http://redmine.ogf.org/projects/nml-wg
http://redmine.ogf.org/projects/nsi-wg

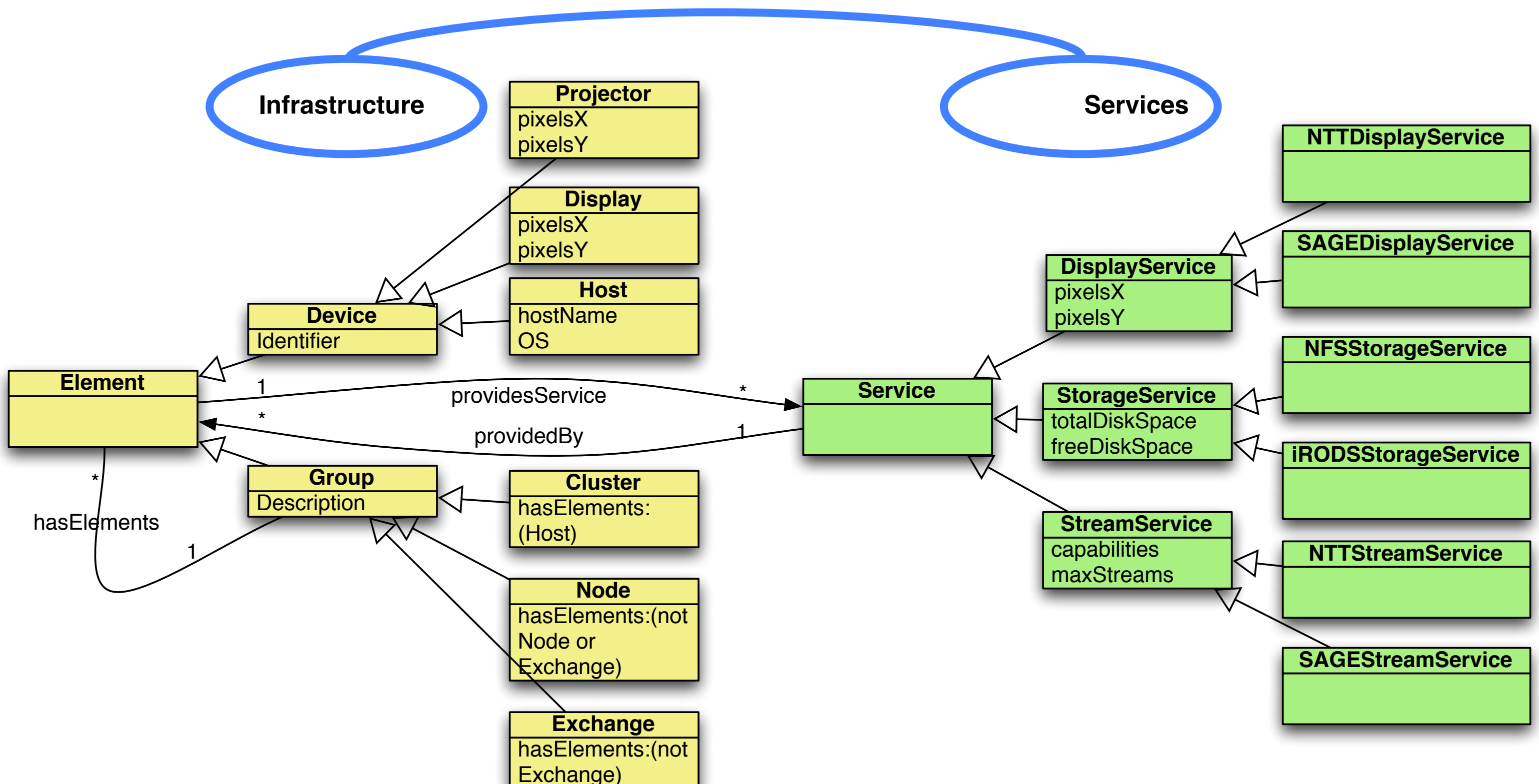
http://sne.science.uva.nl/ndl

RDF describing Infrastructure
“I want”



Information Modeling

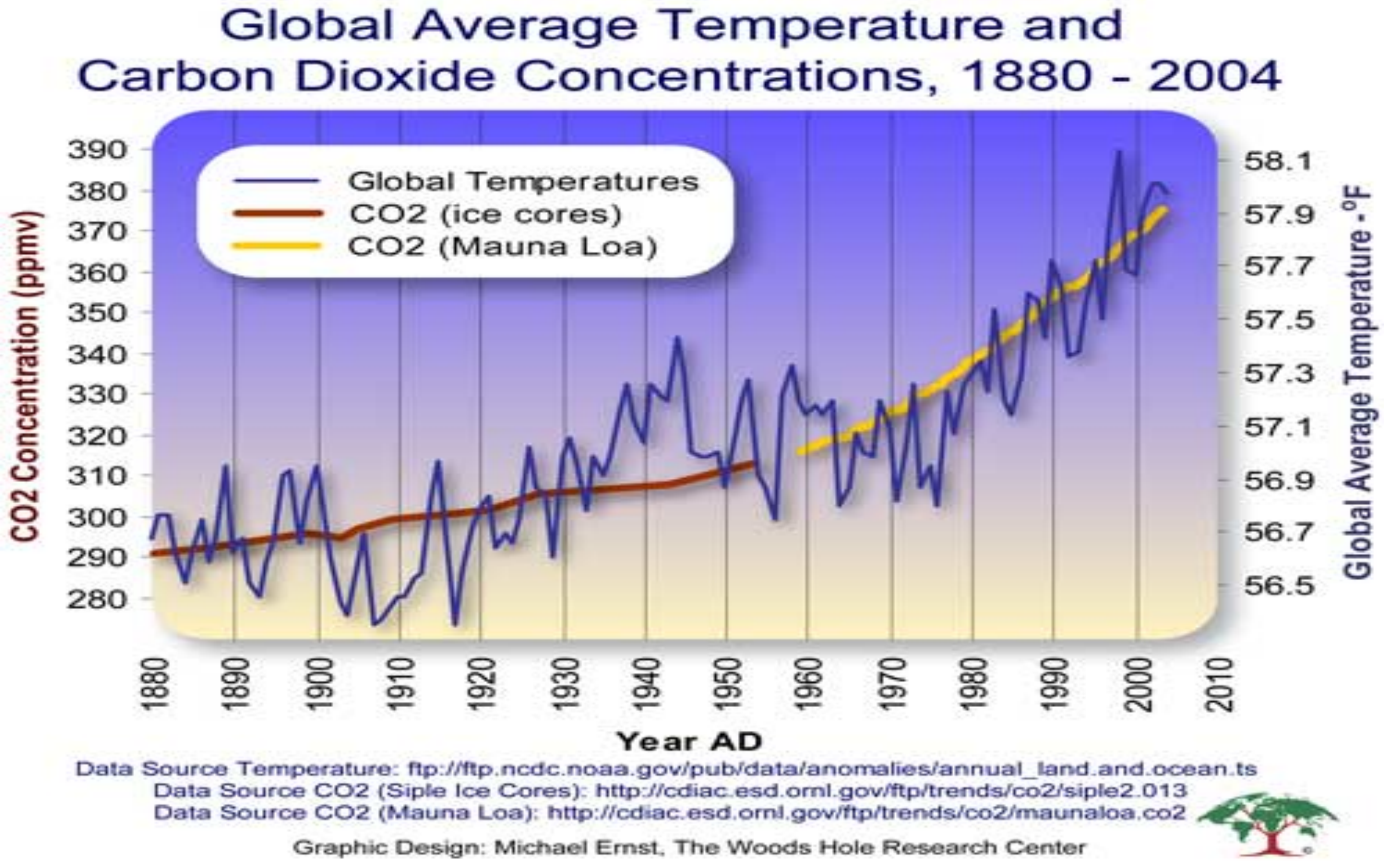
Define a common information model for ***infrastructures*** and ***services***.
Base it on Semantic Web.



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, 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)

Need for GreenIT



Greening the Processing System

ENERGY STAR

Turn Green Tech into Greenbacks

IT Certifications for Jobs That Make a Difference

Uptime Institute Accredited Tier Designer

The Uptime Institute has long been a proponent for green data center design and implementation. Its certification course on data center design embeds green principles into the curriculum.

CERTIFIED GREEN COMPUTING USER SPECIALIST

GCI™ Green Computing Initiative

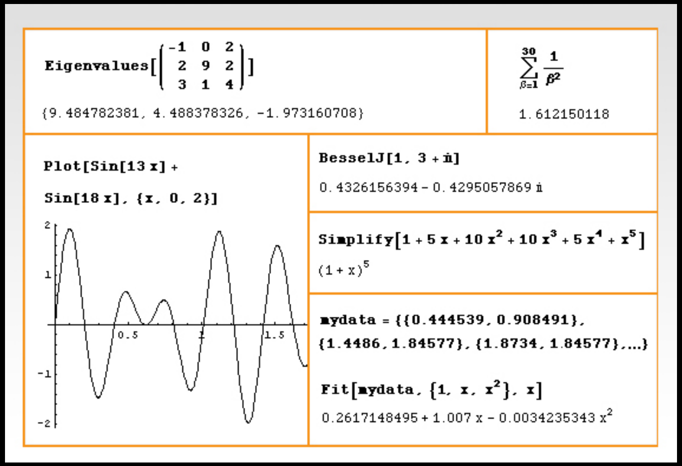
SUSTAINABILITY

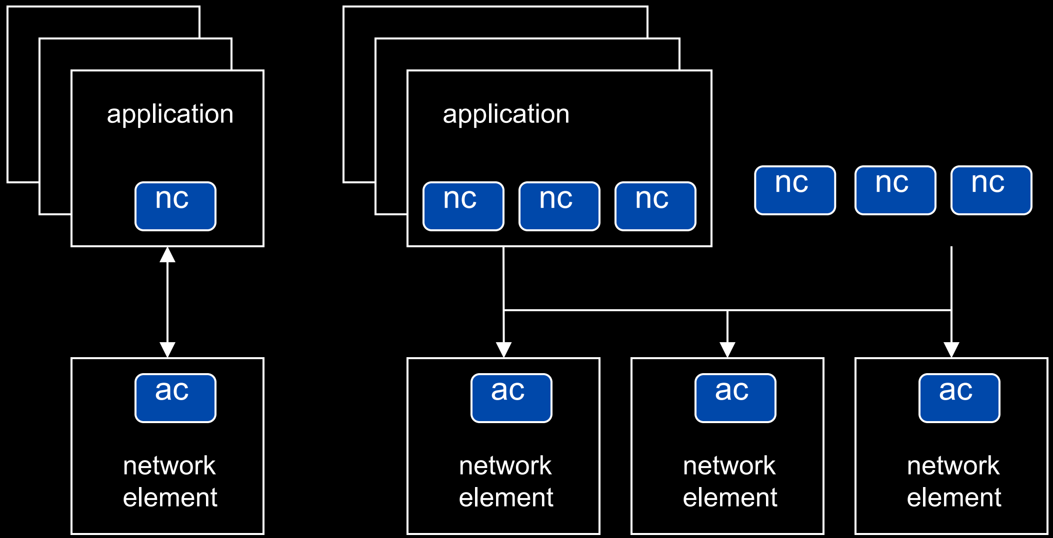
Your Career

ECO-Scheduling

User Programmable Virtualized Networks allows the results of decades of computer science to handle the complexities of application specific networking.

- The network is virtualized as a collection of resources
- UPVNs enable network resources to be programmed as part of the application
- Mathematica, a powerful mathematical software system, can interact with real networks using UPVNs





Mathematica enables advanced graph queries, visualizations and real-time network manipulations on UPVNs

Topology matters can be dealt with algorithmically
Results can be persisted using a transaction service built in UPVN

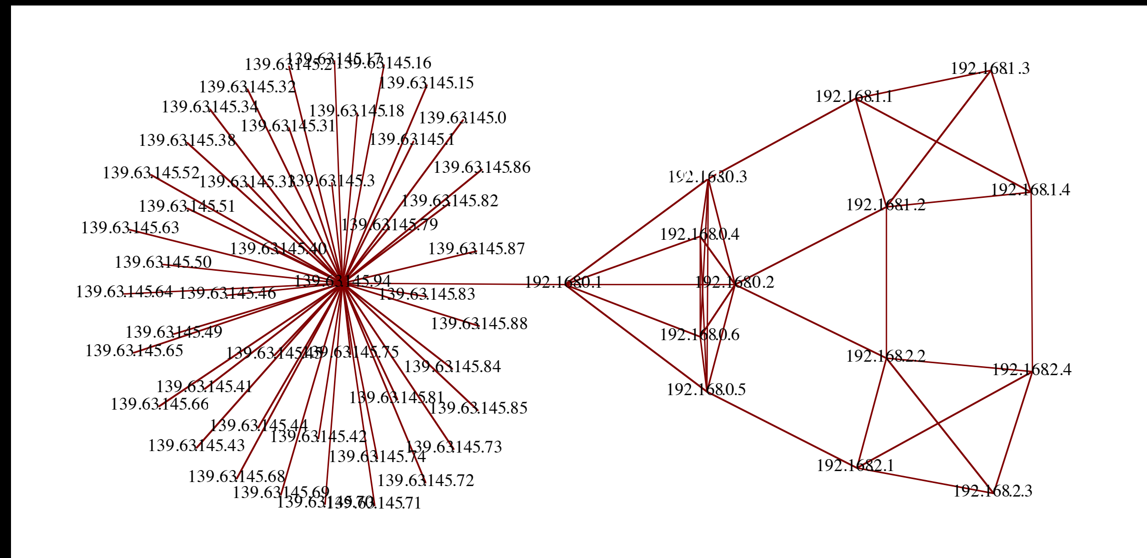
Initialization and BFS discovery of NEs

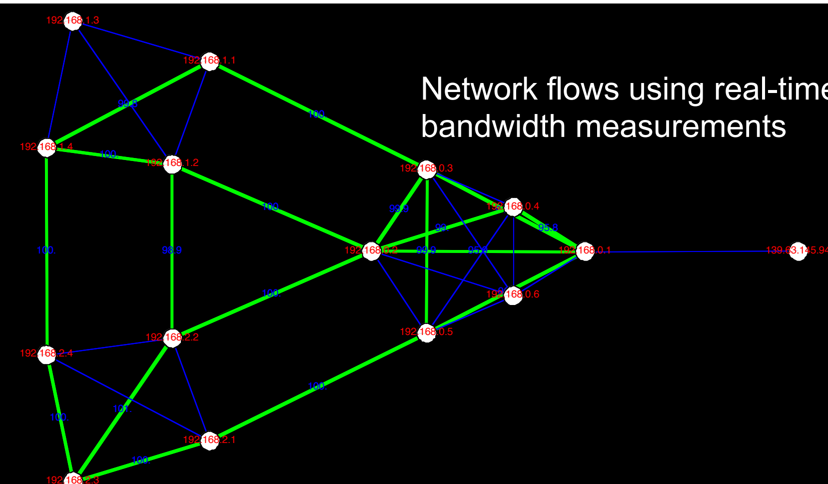
```
Needs["WebServices`"]
<<DiscreteMath`Combinatorica`
<<DiscreteMath`GraphPlot`
InitNetworkTopologyService["edge.ict.tno.nl"]

Available methods:


{DiscoverNetworkElements,GetLinkBandwidth,GetAllIpLinks,Remote,
NetworkTokenTransaction}

Global`upvnverbose = True;

AbsoluteTiming[nes = BFSDiscover["139.63.145.94"];][[1]]

AbsoluteTiming[result = BFSDiscoverLinks["139.63.145.94", nes];][[1]]

Getting neigbours of: 139.63.145.94
Internal links: {192.168.0.1, 139.63.145.94}
(...)
Getting neigbours of:192.168.2.3
```

Transaction on shortest path with tokens

```
Internal links: {192.168.2.3}
```

```
nodePath = ConvertIndicesToNodes[
           ShortestPath[     g,
                      Node2Index[nids,"192.168.3.4"],

      Node2Index[nids,"139.63.77.49"]],
                      nids];
Print["Path: ", nodePath];
If[NetworkTokenTransaction[nodePath, "green"]==True,
      Print["Committed"], Print["Transaction failed"]];

Path:
{192.168.3.4,192.168.3.1,139.63.77.30,139.63.77.49}

Committed
```





Network flows using real-time bandwidth measurements

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.

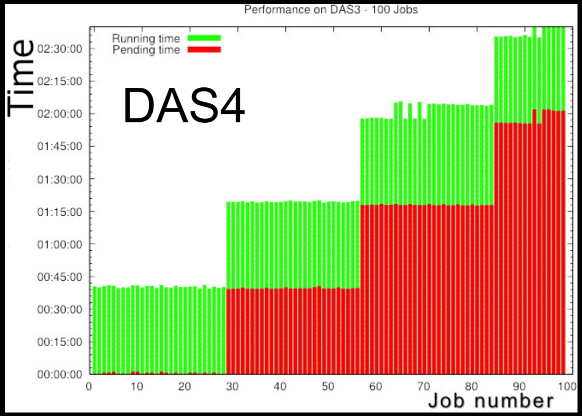
StarPlane

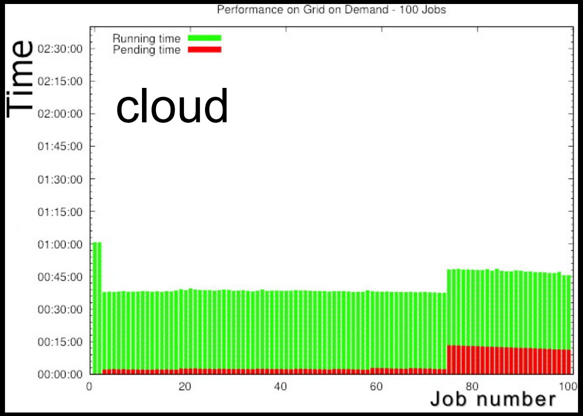
Demonstration of *optimizing the computing problem* ("Clouds")

If computing is 'infinite' and movable, then workflows and applications can ***program*** the network.

You can introduce new metrics when creating and optimizing these infrastructures (*e.g* power consumption)

Grid-on-demand



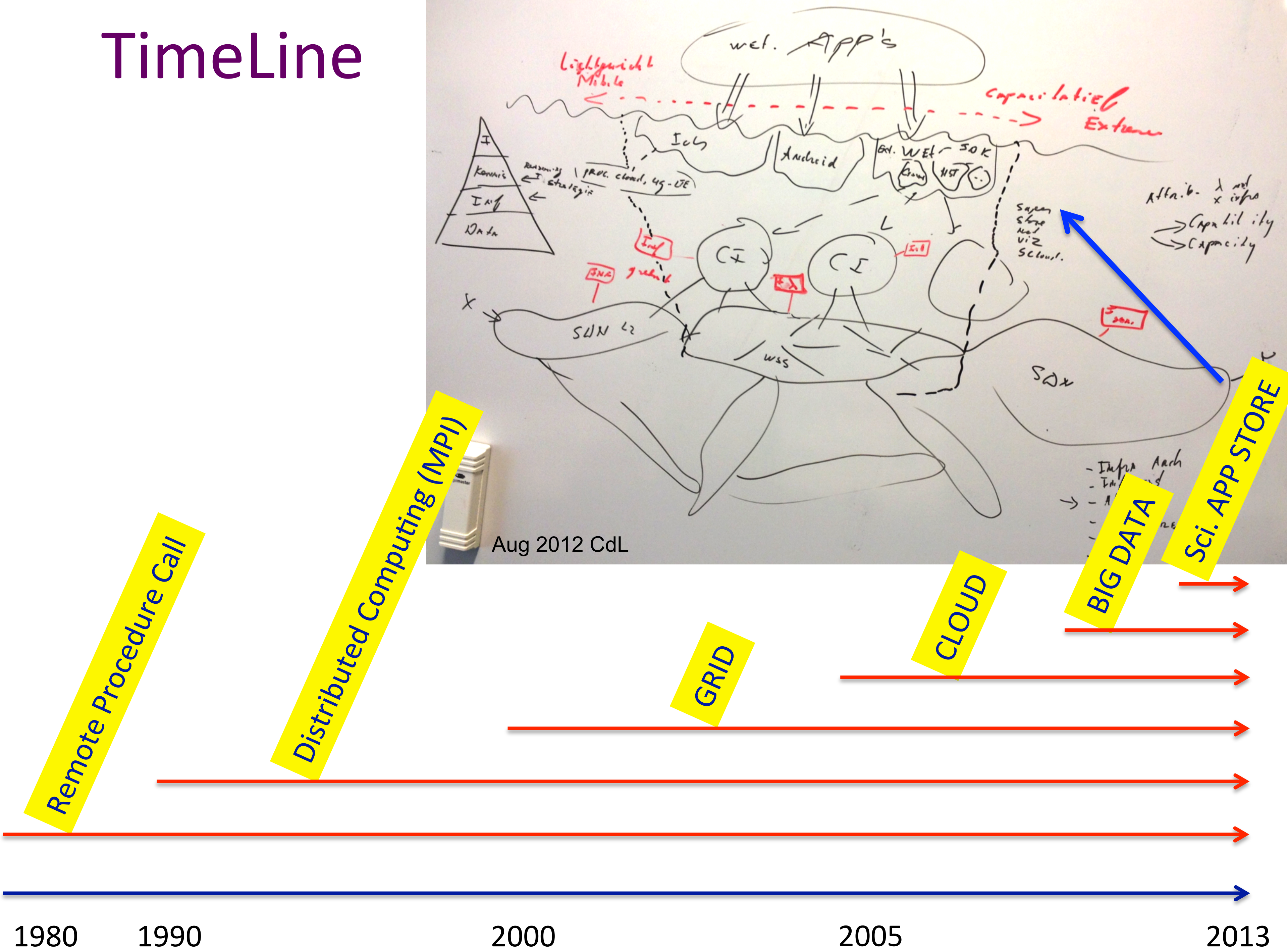


User programmable networks

R.Strijkers, W.Toorop, A.van Hoof, P.Grosso, A.Belloum, D.Vasuining, C. de Laat, R. Meijer, "AMOS: Using the Cloud for On-Demand Execution of e-Science Applications", In: Proc. eScience2010 conf, Dec. 2010.

Y. Demchenko, C.Ngo, M.Makkes, R.Strijkers, C. de Laat, "Defining Inter-Cloud Architecture for Interoperability and Integration.", 3th intl conf on Cloud Computing, GRIDs, and Virtualization (CLOUDCOM 2012), July 22-27, 2012, Nice, France. BEST PAPER AWARD

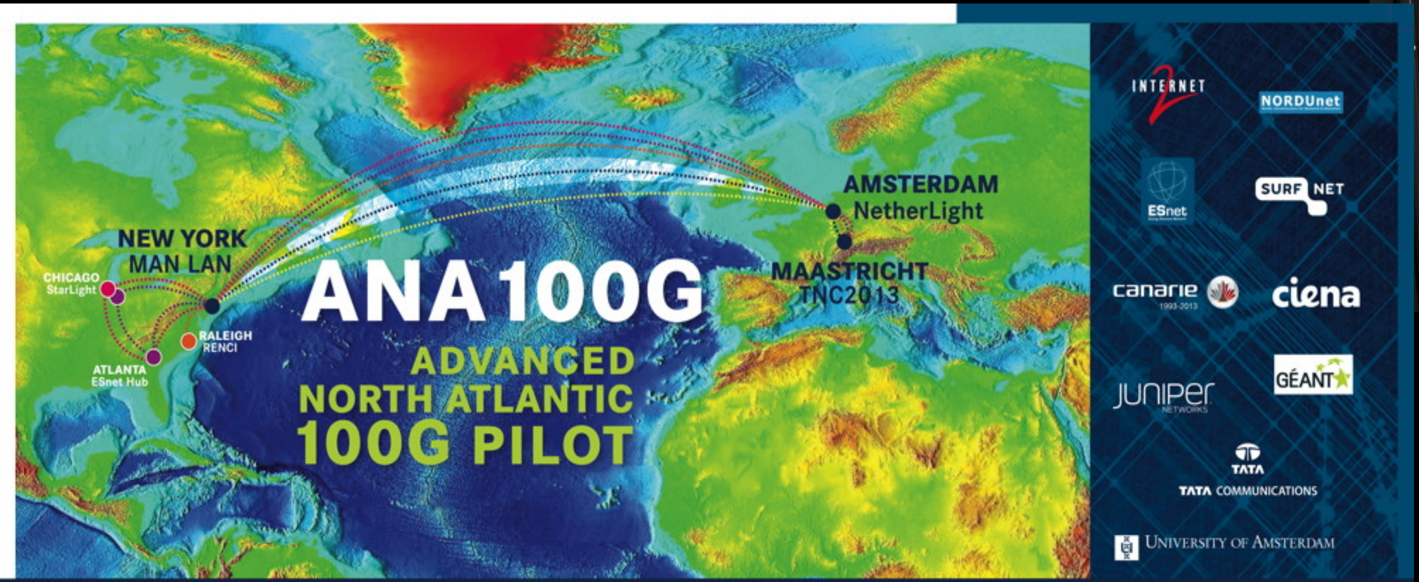
TimeLine



Aug 2012 CdL
Remote Procedure Call
Distributed Computing (MPI)
GRID
CLOUD
BIG DATA
Sci. APP STORE
1980
1990
2000
2005
2013

ExoGeni @ UvA

Installed and up June 3th 2013



TNC2013 DEMOS JUNE, 2013

DEMO	TITLE	OWNER	AFFILIATION	E-MAIL	A-SIDE	Z-SIDE	PORTS(S) MAN LAN	PORTS(S) TNC2013	DETAILS
1	Big data transfers with multipathing, OpenFlow and MPTCP	Ronald van der Pol	SURFnet	ronald.vanderpol@surfnet.nl	TNC/MECC, Maastricht NL	Chicago, IL	Existing 100G link between internet2 and ESnet	2x40GE (Juniper)+ 2x10GE (OME6500)	In this demonstration we show how multipathing, OpenFlow and Multipath TCP (MPTCP) can help in large file transfers between data centres (Maastricht and Chicago). An OpenFlow application provisions multiple paths between the servers and MPTCP will be used on the servers to simultaneously send traffic across all those paths. This demo uses 2x40G on the transatlantic 100G link. ESnet provides 2x40G between MAN LAN and StarLight, ACE and USLHCnet provide additional 10GEs.
2	Visualize 100G traffic	Inder Monga	ESnet	imonga@es.net					Using an SNMP feed from the Juniper switch at TNC2013,and/or Brocade AL2S node in MANLAN, this demo would visualize the total traffic on the link, of all demos aggregated. The network diagram will show the transatlantic topology and some of the demo topologies.
3	How many modern servers can fill a 100Gbps Transatlantic Circuit?	Inder Monga	ESnet	imonga@es.net	Chicago, Ill	TNC showfloor	1x 100GE	8x 10GE	In this demonstration, we show that with the proper tuning and tool, only 2 hosts on each continent can generate almost 80Gbps of traffic. Each server has 4 10G NICS connected to a 40G virtual circuit, and has iperf3 running to generate traffic. ESnet's new "iperf3" throughput measurement tool, still in 'beta', combines the best features from other tools such as iperf, nuttcp, and netperf. See: https://my.es.net/demos/tnc2013/
4	First European ExoGENI at Work	Jeroen van der Ham	UvA	vdham@uva.nl	RENCI, NC	UvA, Amsterdam, NL	1x 10GE	1x 10GE	The ExoGENI racks at RENCI and UvA will be interconnected over a 10G pipe and be on continuously, showing GENI connectivity between Amsterdam and the rest of the GENI nodes in the USA.
5	Up and down North Atlantic @ 100G	Michael Enrico	DANTE	michael.enrico@dante.net	TNC showfloor	TNC showfloor	1x 100GE	1x 100GE	The DANTE 100GE test set will be placed at the TNC2013 showfloor and connected to the Juniper at 100G. When this demo is running a loop @ MAN LAN's Brocade switch will ensure that the traffic sent to MAN LAN returns to the showfloor. On display is the throughput and RTT (to show the traffic traveled the Atlantic twice)

Connected via the new 100 Gb/s transatlantic wavelength

ExoGeni @ UvA

- Part of UvA's OpenLab ➔ Open for everyone!
- Installed and up June 3th 2013
- Connected via the new 100 Gb/s transatlantic
- To study programmability on all layers
- To study computing to data vs data to computing
- To study GreenSonar & objective based networking
- Study multi service exchange & DMZ features
- To study Big Data processing algorithms on mixed latency
- PIRE project with Grossman and Alvares
- Give students access to try out their bright and stupid ideas!
- DAS4/5, CineGrid exchange node, pure photonic TUE

CineGrid Mission

To build an interdisciplinary community that is focused on the research, development, and demonstration of networked collaborative tools to enable the production, use and exchange of very-high-quality digital media over photonic networks.

http://www.cinegrid.org/

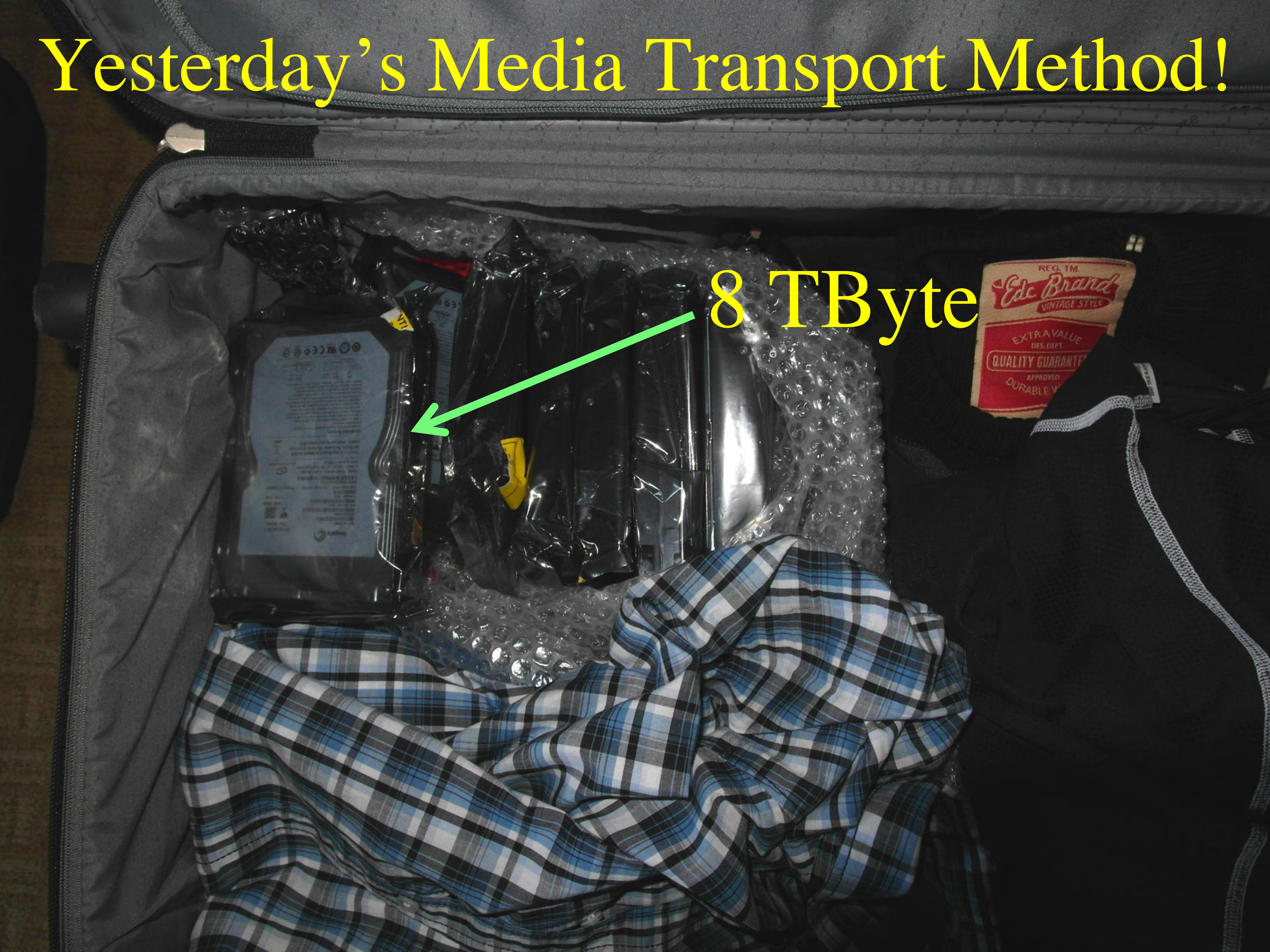
Why is more resolution is better?

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

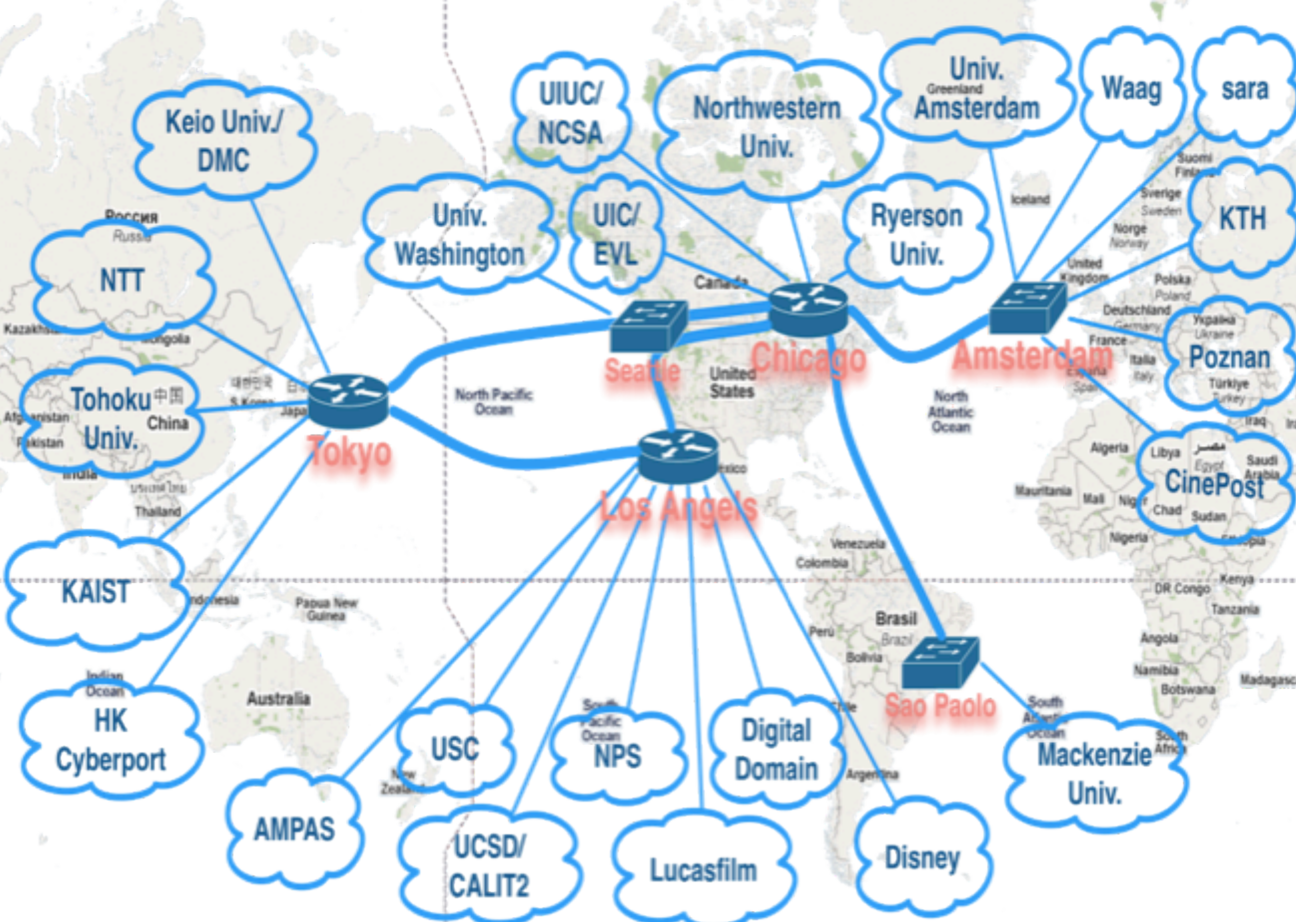
Yutaka TANAKA
SHARP CORPORATION
Advanced Image Research Laboratories

Moving Big Data Objects Globally

- **Digital Motion Picture for Audio Post-Production**
  - 1 TV Episode Dubbing Reference ~ 1 GB
  - 1 Theatrical 5.1 Final Mix ~ 8 GB
  - 1 Theatrical Feature Dubbing reference ~ 30 GB
- **Digital Motion Picture Acquisition**
  - 4K RGB x 24 FPS x 10bit/color: ~ 48MB/Frame uncompressed *(ideal)*
  - 6:1 ~ 20:1 shooting ratios => 48TB ~ 160TB digital camera originals
- **Digital Dailies**
  - HD compressed MPEG-2 @ 25 ~ 50 Mb/s
- **Digital Post-production and Visual Effects**
  - Gigabytes - Terabytes to Select Sites Depending on Project
- **Digital Motion Picture Distribution**
  - Film Printing in Regions
    - Features ~ 8TB
    - Trailers ~ 200GB
  - Digital Cinema Package to Theatres
    - Features ~ 100 - 300GB per DCP
    - Trailers ~ 2 - 4GB per DCP


Yesterday's Media Transport Method!
8 TByte
REG. TM.
Edc Brand
VINTAGE STYLE
EXTRA VALUE
QUALITY GUARANTE

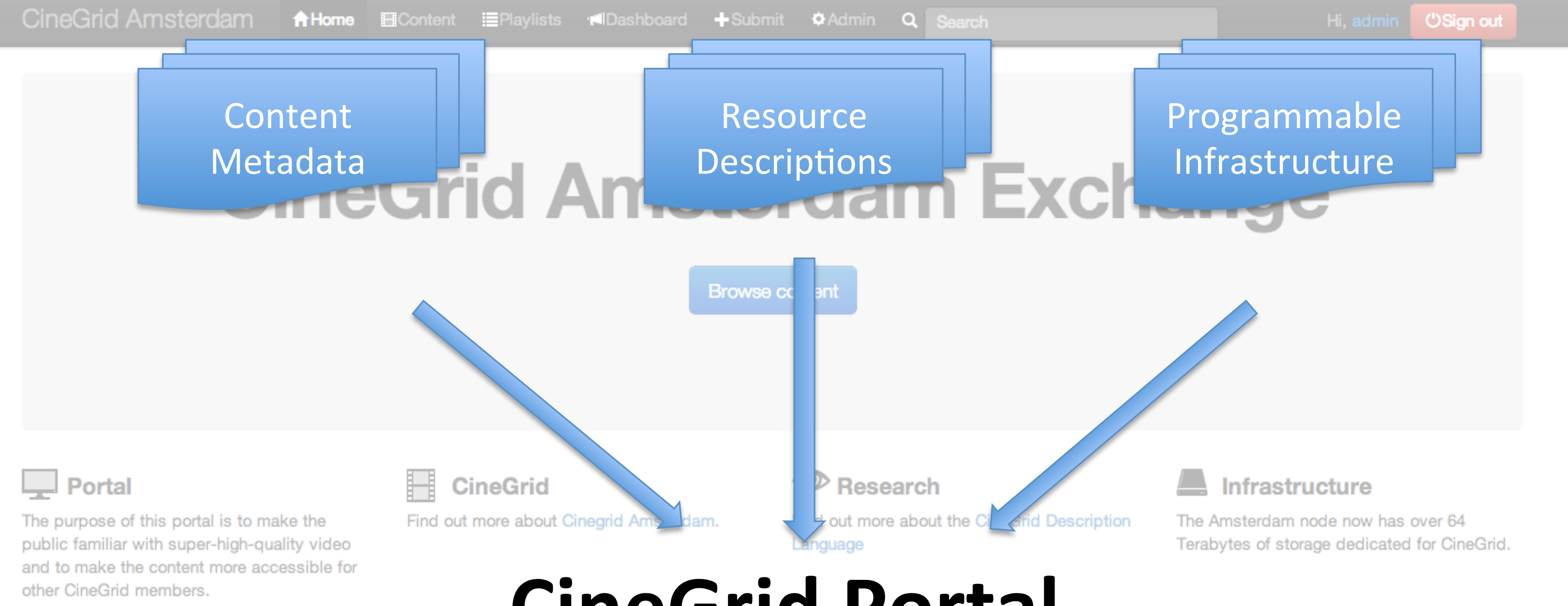
CineGrid Network 2012



Network Resources are kindly provided by:

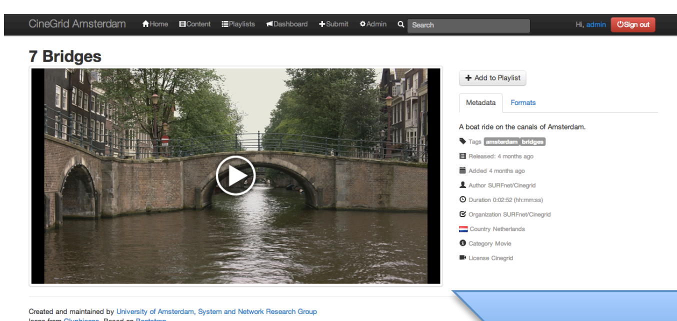
AMPATH, C-Wave, CANARIE, CaveWave, CENIC, CESNET, CzechLight, GEMNET, Internet2, JANET, JGN-X, NetherLight, NLR, NORDUnet, PacificWave, PNWGP, RNP, StarLight, SOL, SURFnet, TransLight/StarLight, T-LEX, WIDE

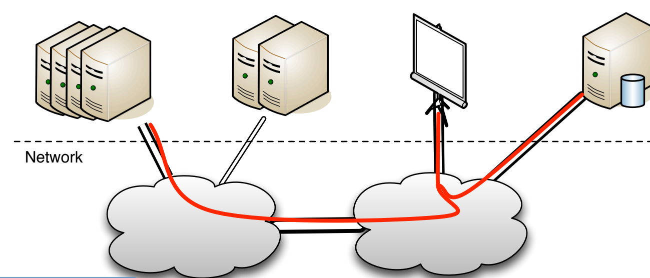
kaneko@dmc.keio.ac.jp, as of 2012/07/27



CineGrid Portal

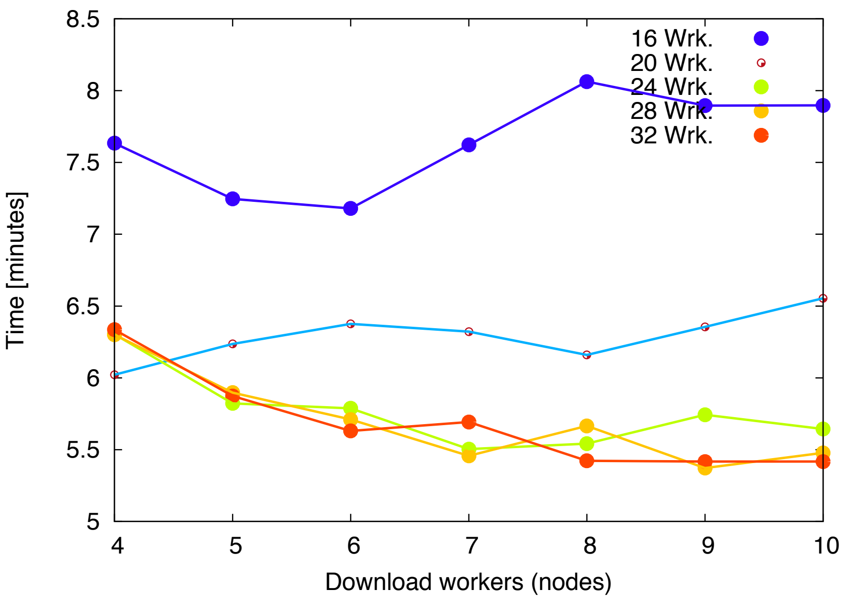
Unified orchestration of distributed CineGrid resources

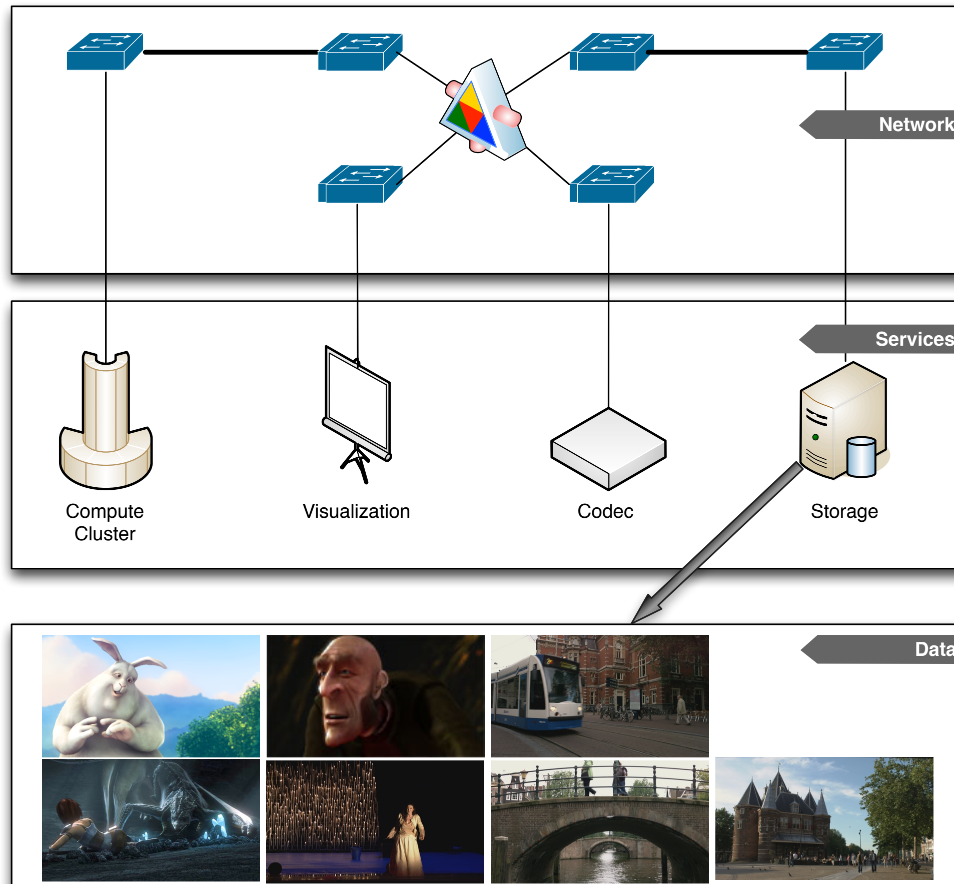




HyperFlow

Encoding times improve as the end nodes are connected via dynamic lightpaths



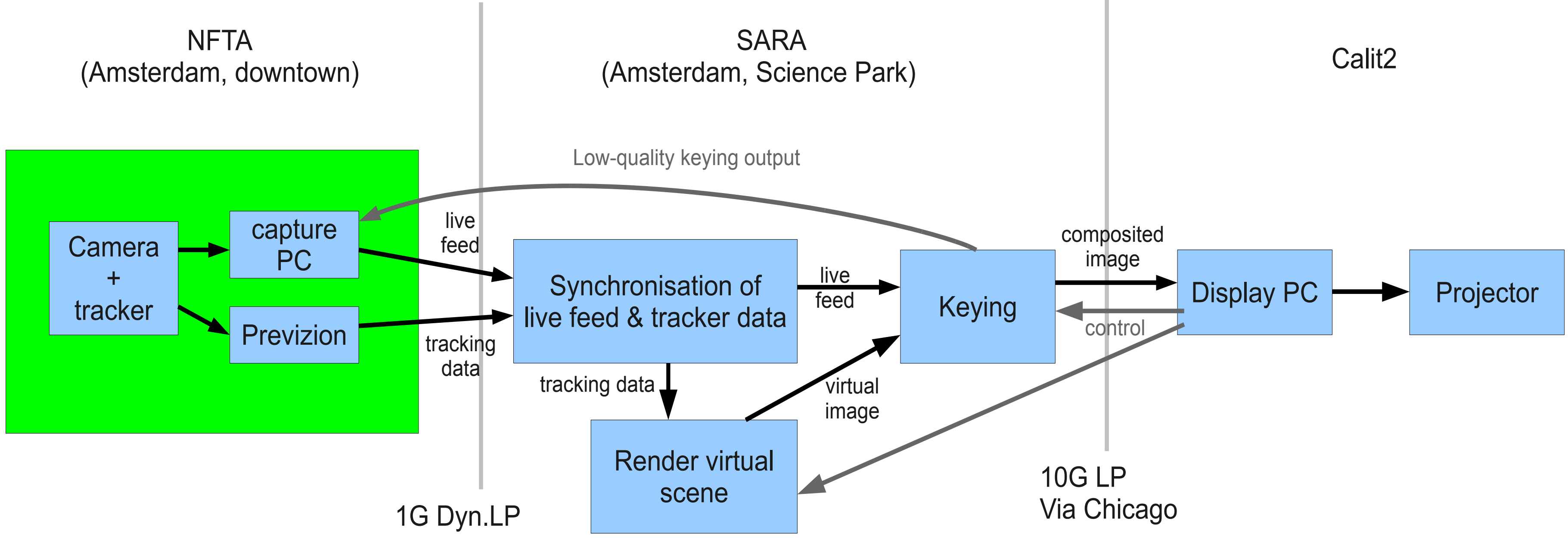


C. Dumitru, Z. Zhao, P. Grosso and C. de Laat
HybridFlow: Towards Intelligent Video Delivery and Processing Over Hybrid Infrastructures
(In CTS 2013))

Real Time Rendering Workflow

Demo setup

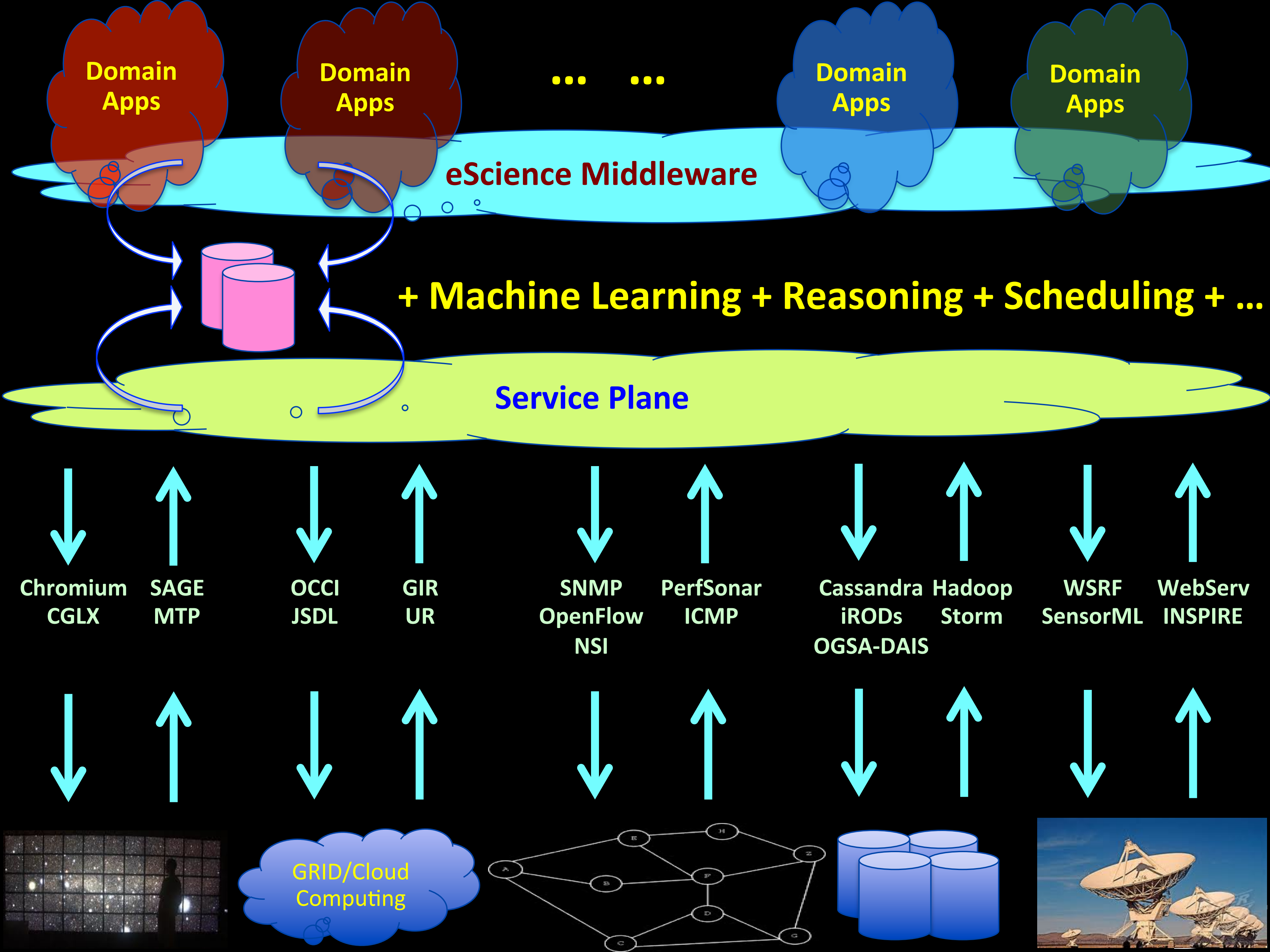
- Three locations
  1) NFTA: greenscreen studio, Previzion, camera(+man), actress (+ dress)
  2) SARA: render node for keying, virtual scene rendering
  3) Calit2: keying controls, projection of final output, director
- Two lightpaths in between
- Video-conferencing for communication + low quality keying output back to NFTA



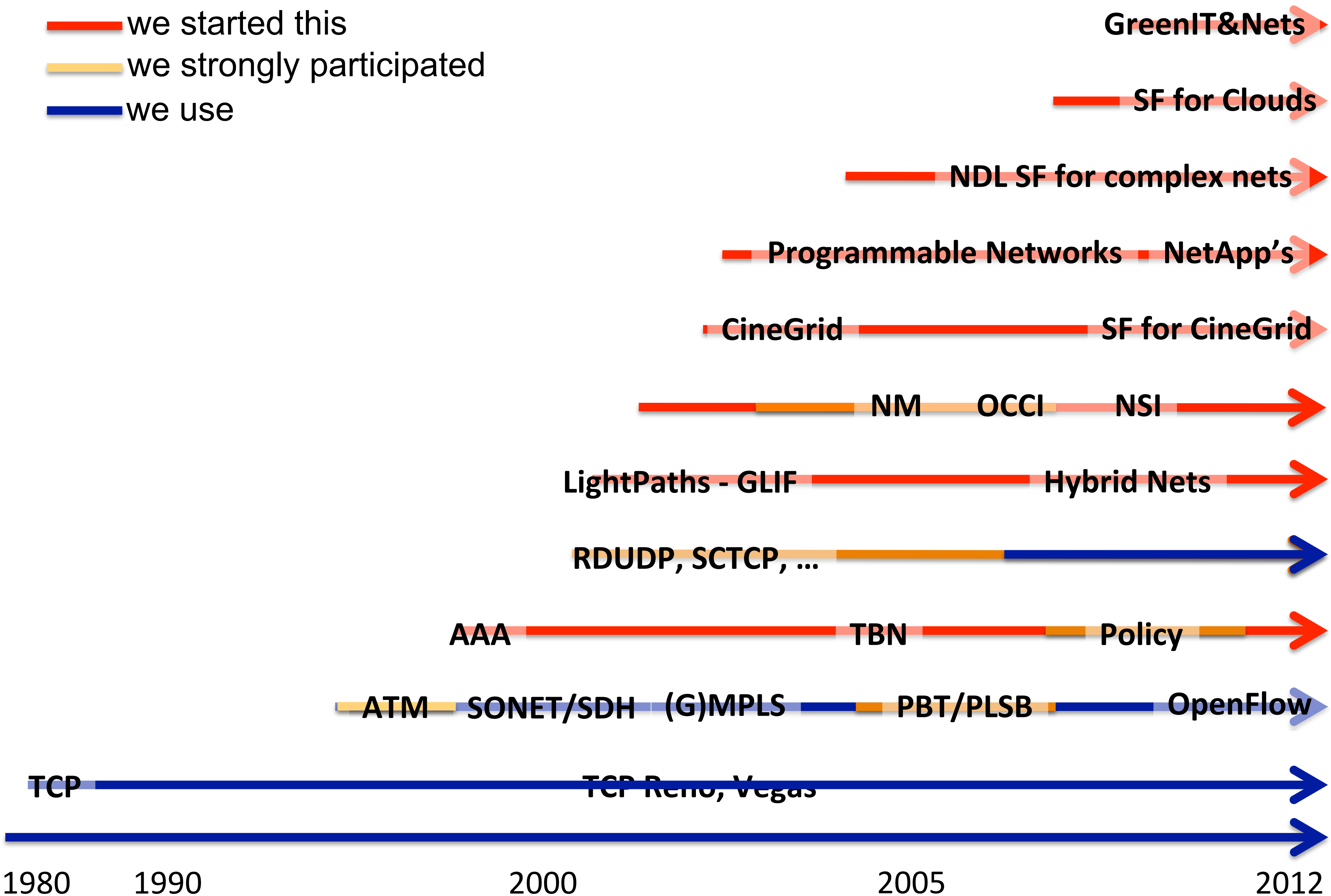
I want to

"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.

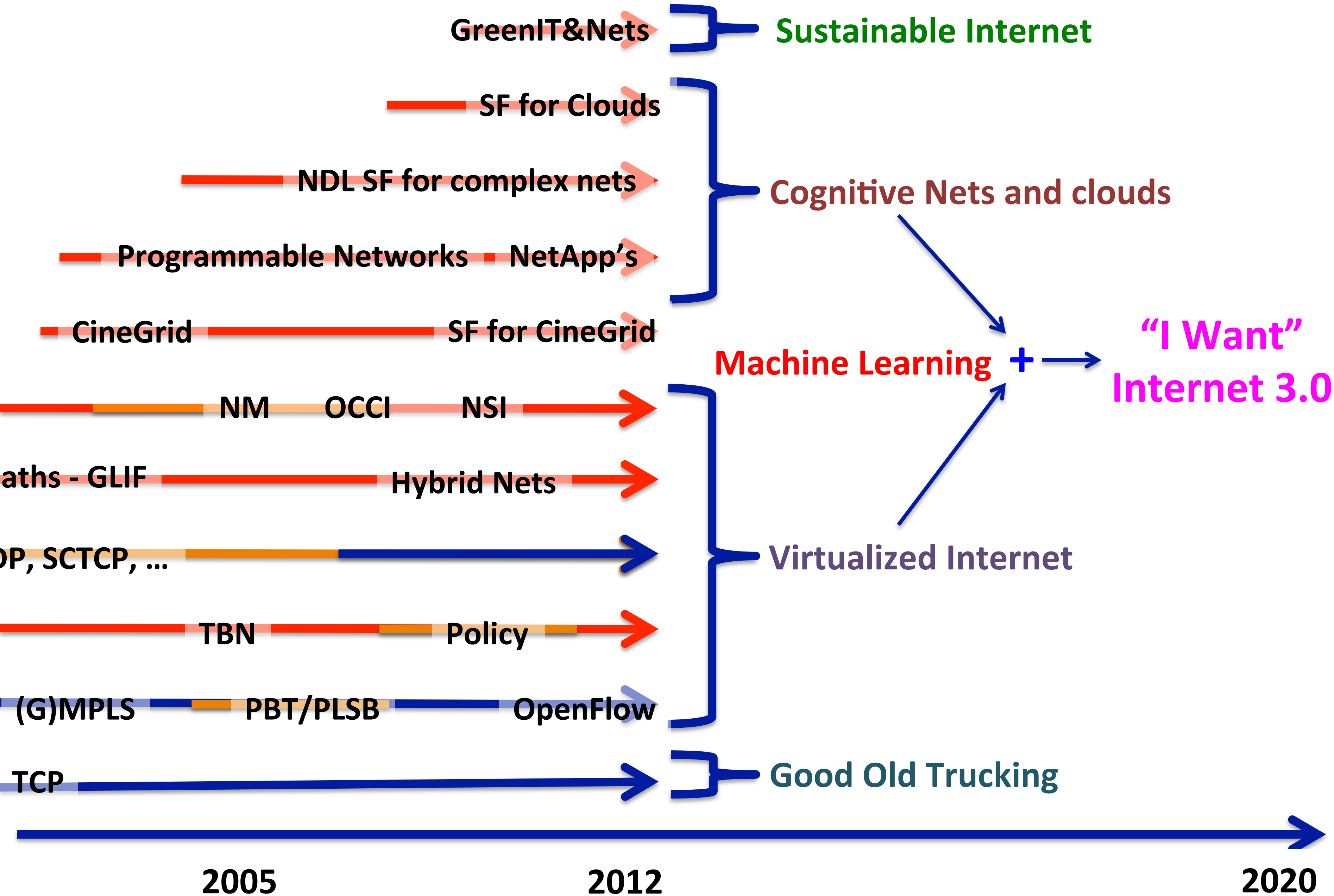

Domain Apps
Domain Apps
... ...
Domain Apps
Domain Apps
eScience Middleware
+ Machine Learning + Reasoning + Scheduling + ...
Service Plane
Chromium CGLX
SAGE MTP
OCCI JSDL
GIR UR
SNMP OpenFlow NSI
PerfSonar ICMP
Cassandra iRODs OGSA-DAIS
Hadoop Storm
WSRF SensorML
WebServ INSPIRE
GRID/Cloud Computing

TimeLine



we started this
we strongly participated
we use
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
2012

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
2012
2020

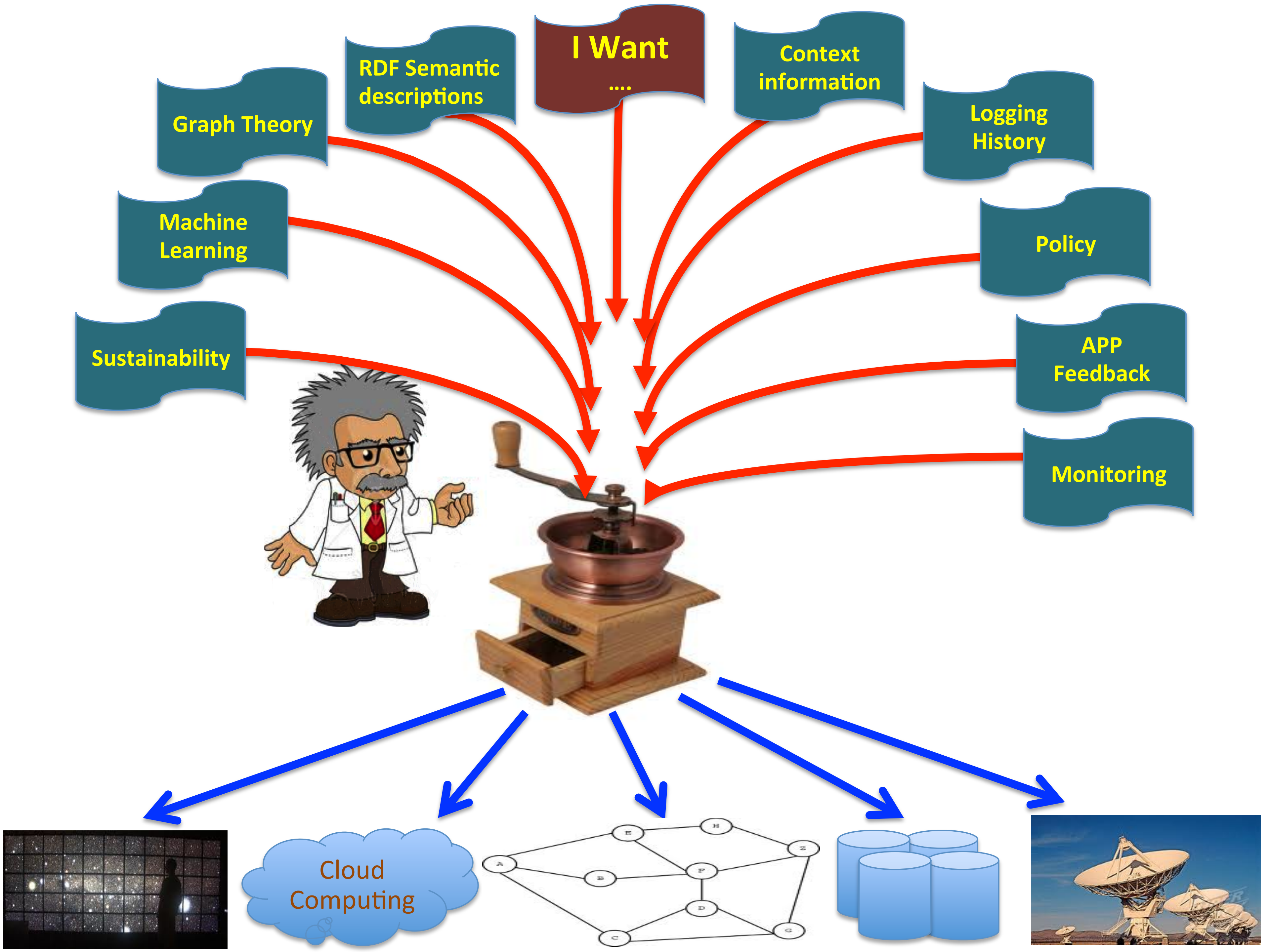
TimeLine

- Sustainable Internet
- Cognitive Nets and clouds
lachine Learning + → "I Want" Internet 3.0
- Virtualized Internet
- Good Old Trucking

I retire

2020

2040

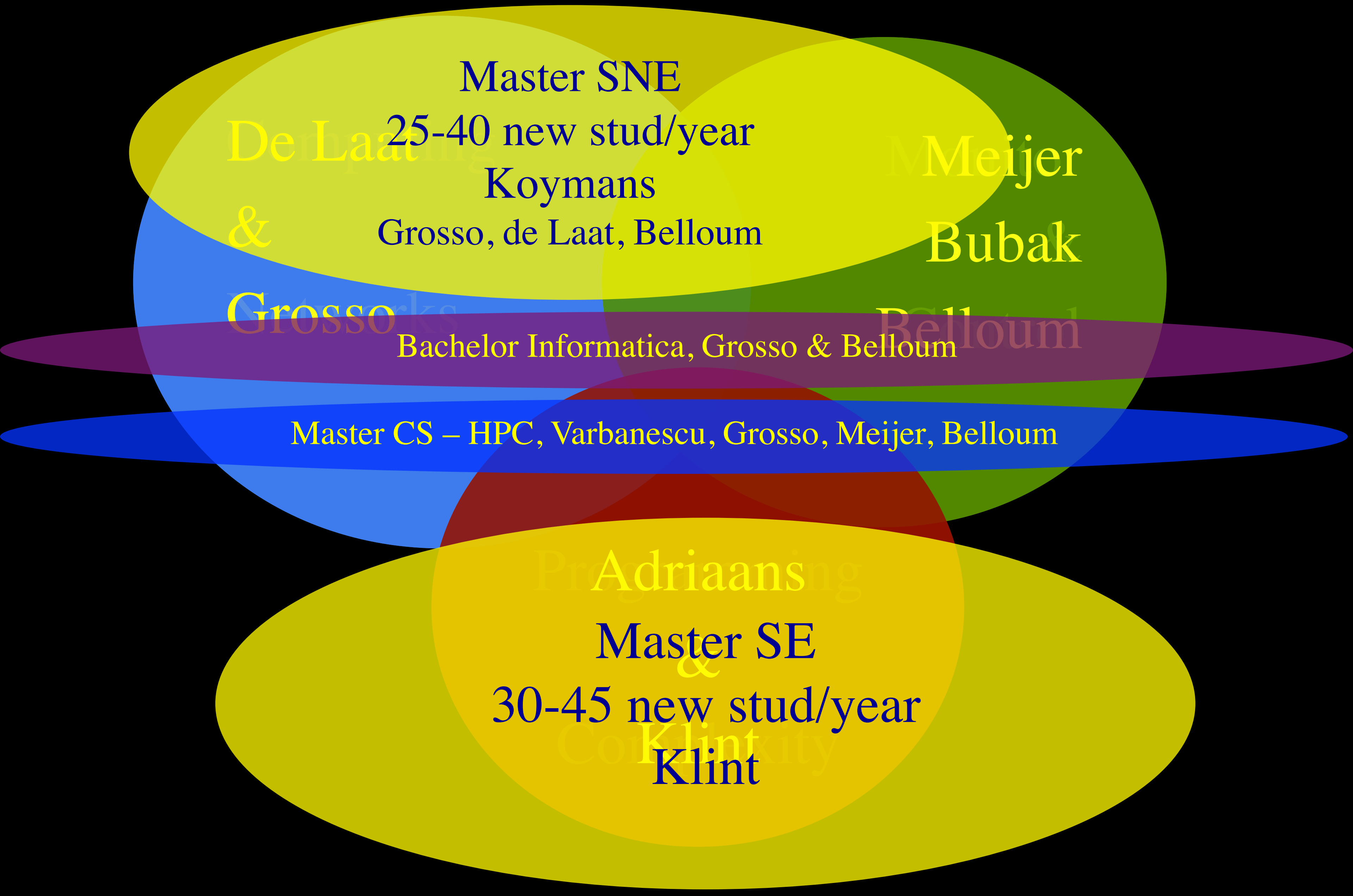


I Want
....
RDF Semantic descriptions
Context information
Graph Theory
Logging History
Machine Learning
Policy
Sustainability
APP Feedback
Monitoring
Cloud Computing
A
B
C
D
E
F
G
H
Z

Why?

Because we can!

SNE


Master SNE
25-40 new stud/year
Koymans
Grosso, de Laat, Belloum
De Laat
&
Grosso
Meijer
Bubak
Belloum
Bachelor Informatica, Grosso & Belloum
Master CS – HPC, Varbanescu, Grosso, Meijer, Belloum
Adriaans
&
Klint
Master SE
30-45 new stud/year
Klint

EU

SARA

SURF-eSRC

Pieken-in-de-Delta

SURFnet

NWO-NCF

FES

UVA

NWO