Putting the Grid in CineGrid.

Cees de Laat

EU

COMMIT

UvA

NWO

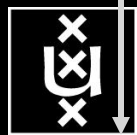
PID/EFRO

SURFnet

NLESC

TNO

NWO/ncf



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.

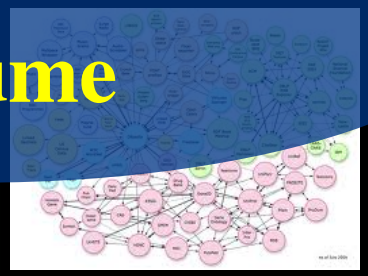
http://www.cinegrid.org/

Internet developments

... more data!

Speed

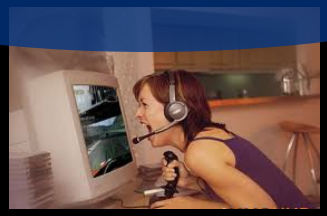
Volume



Deterministic

Real-time

more realtime!

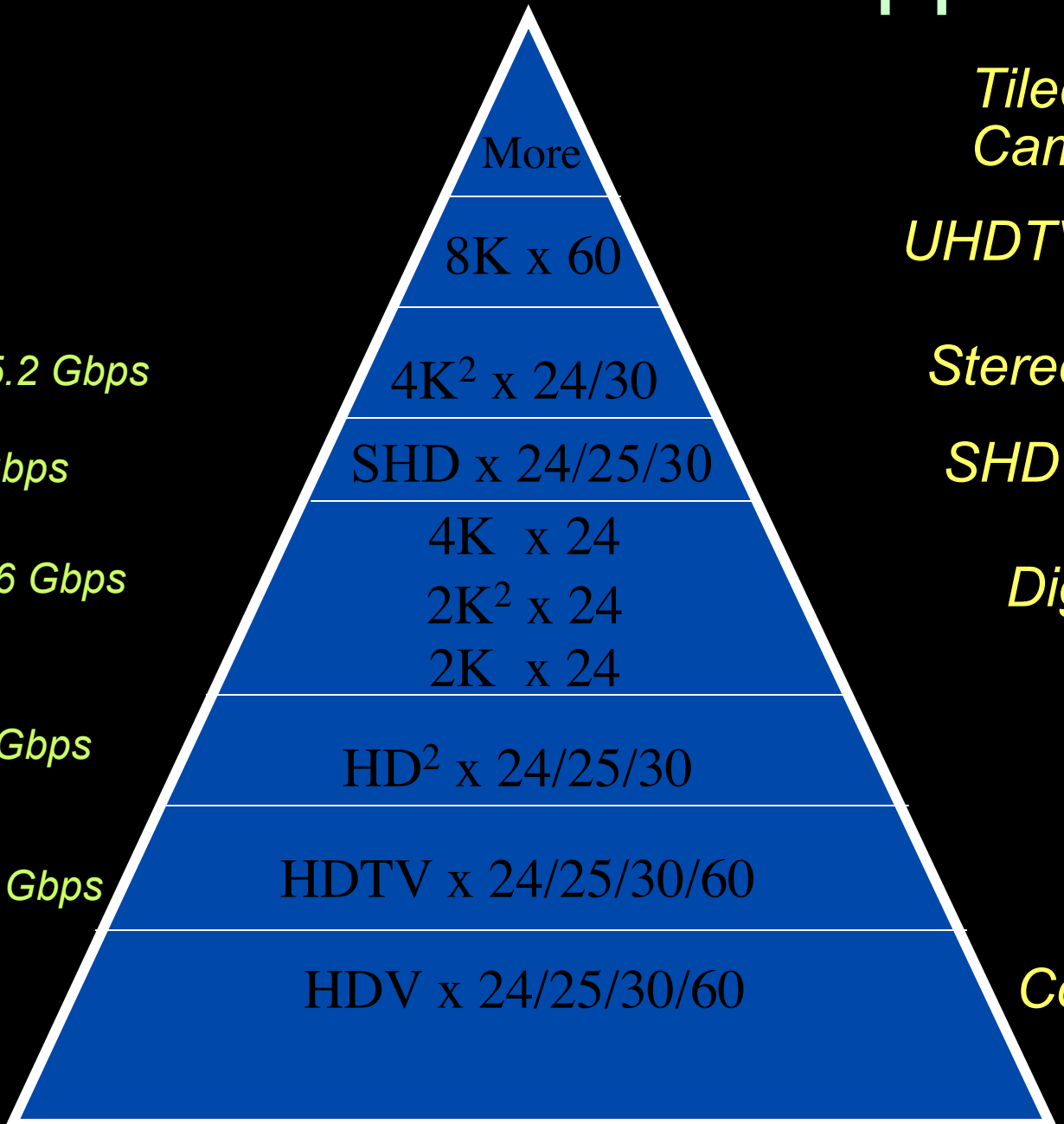


Scalable

Secure

... more users!

CineGrid: A Scalable Approach



Tiled Displays
Camera Arrays

1 - 24 Gbps

UHDTV (far future)

500 Mbps - 15.2 Gbps

Stereo 4K (future)

250 Mbs - 6 Gbps

SHD (Quad HD)

250 Mbps - 7.6 Gbps

Digital Cinema

200 Mbps - 3 Gbps

Stereo HD

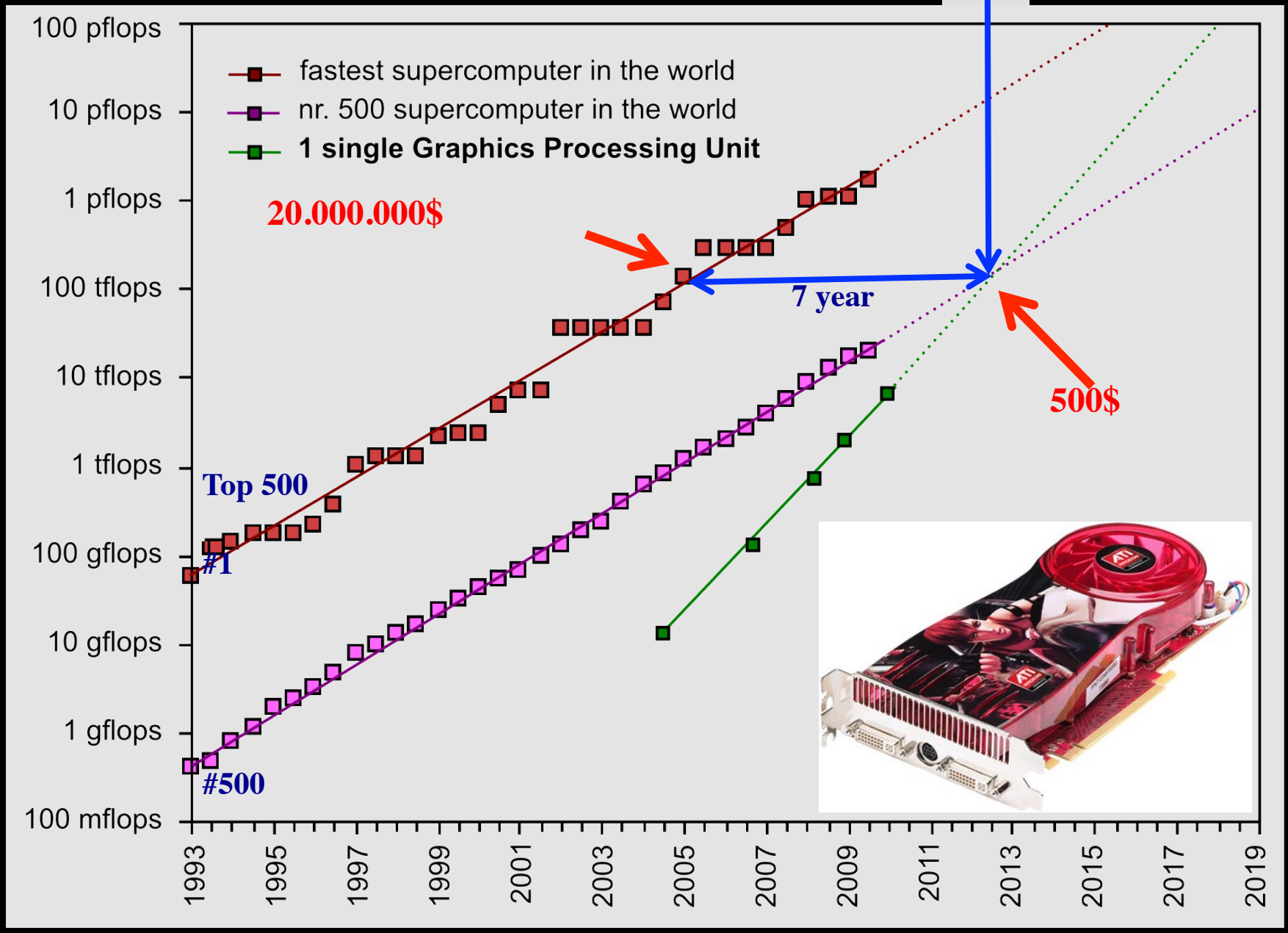
20 Mbps - 1.5 Gbps

HDTV

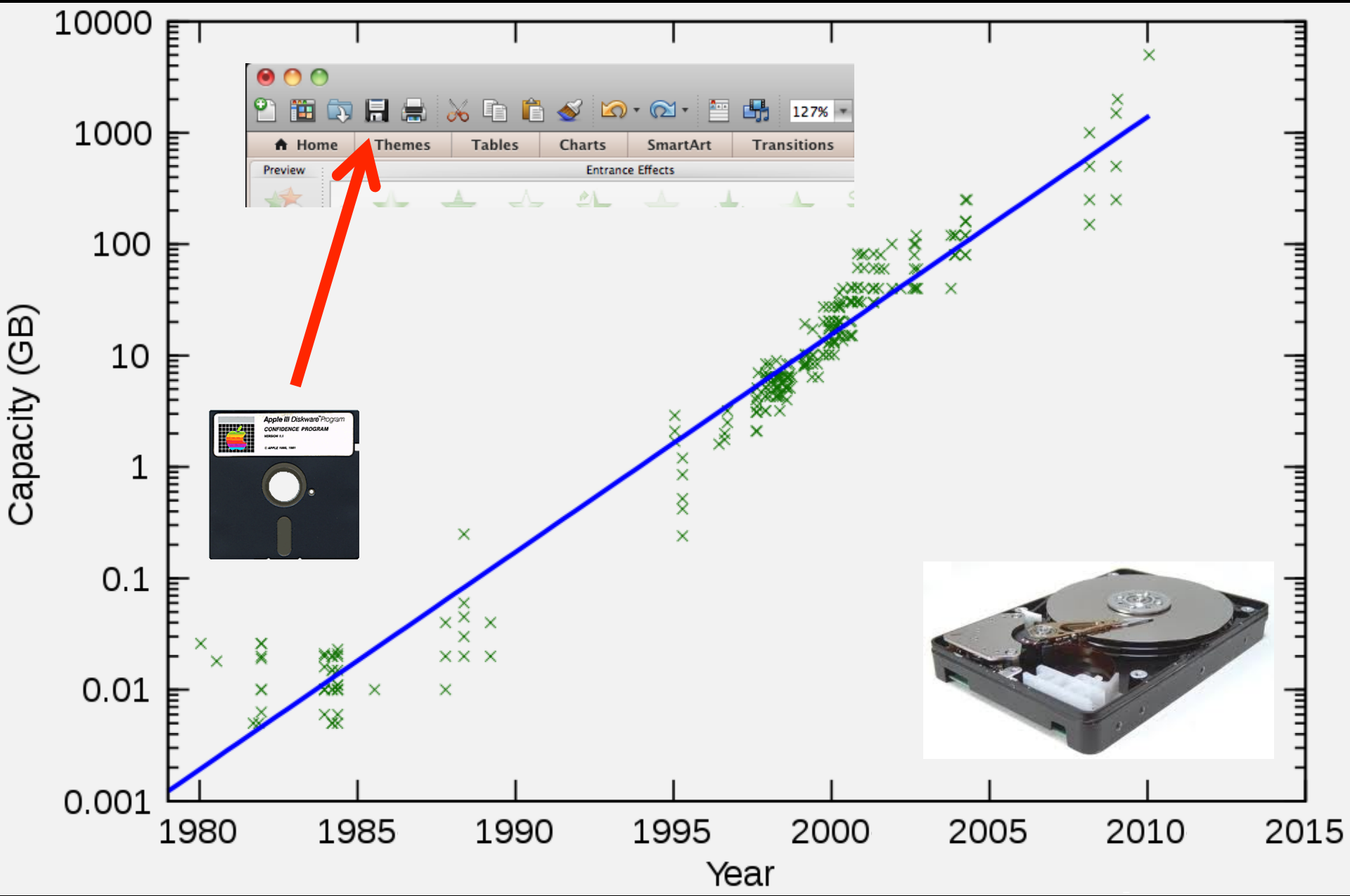
5 - 25 Mbps

Consumer HD

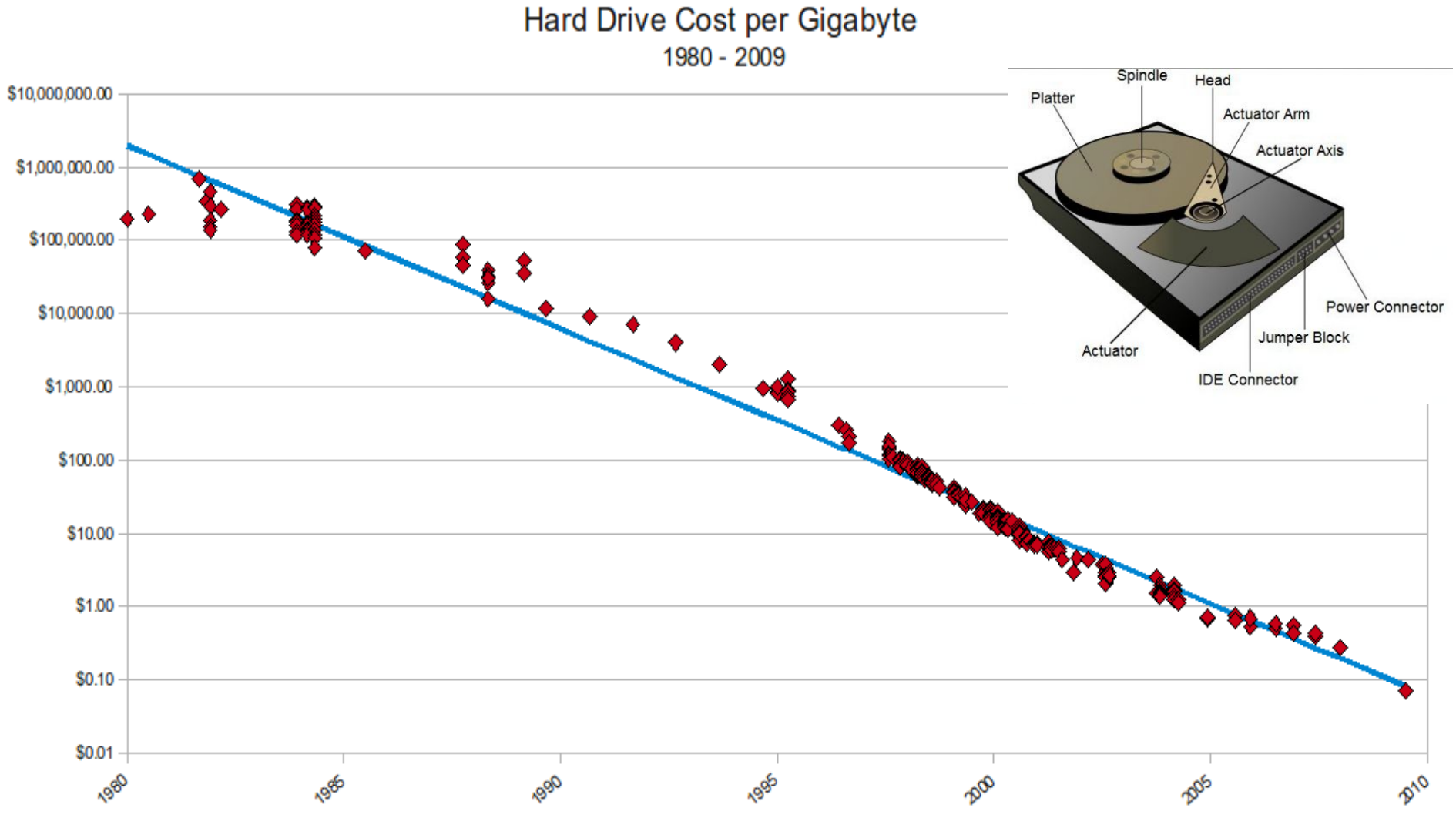
GPU cards are distruptive!



Data storage: doubling every 1.5 year!

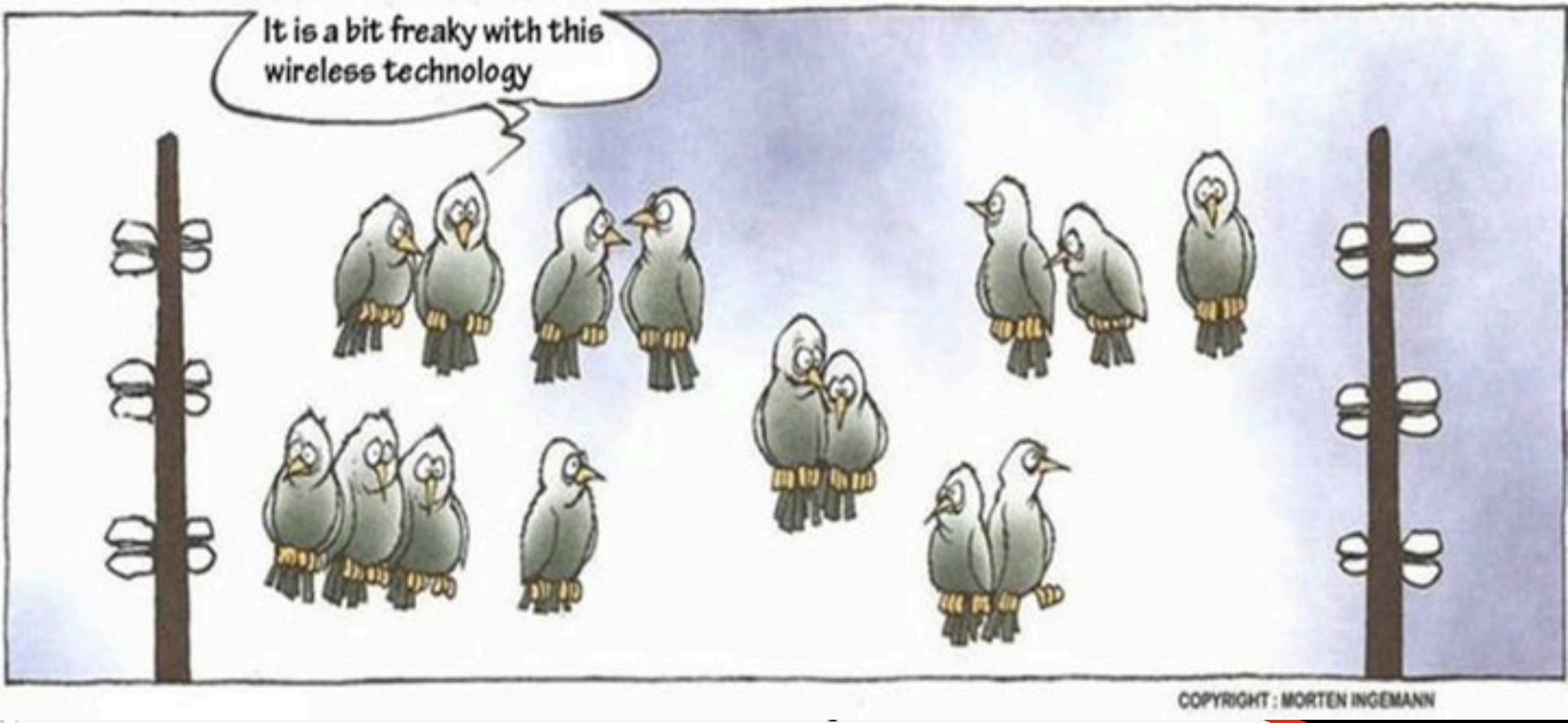


Storage costs less and less



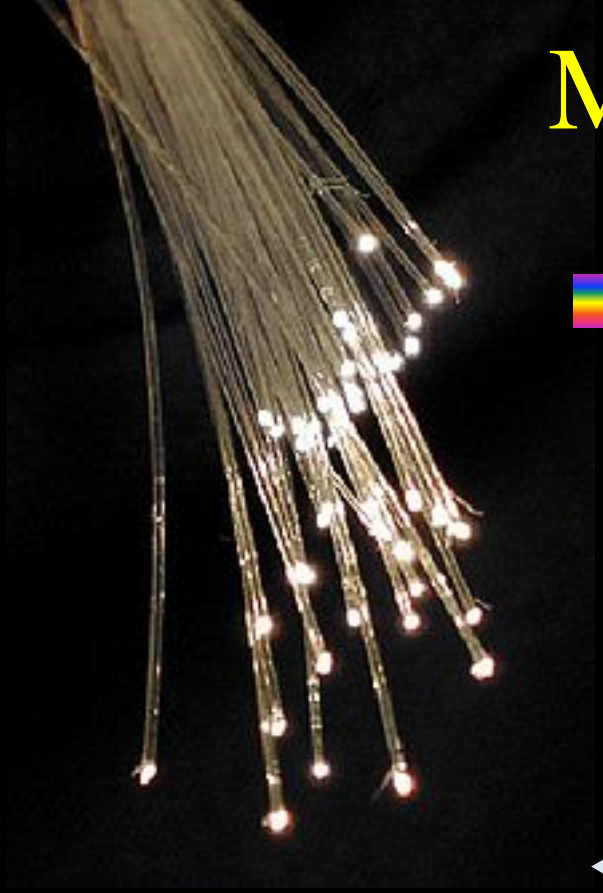
http://www.mkomo.com/cost-per-gigabyte

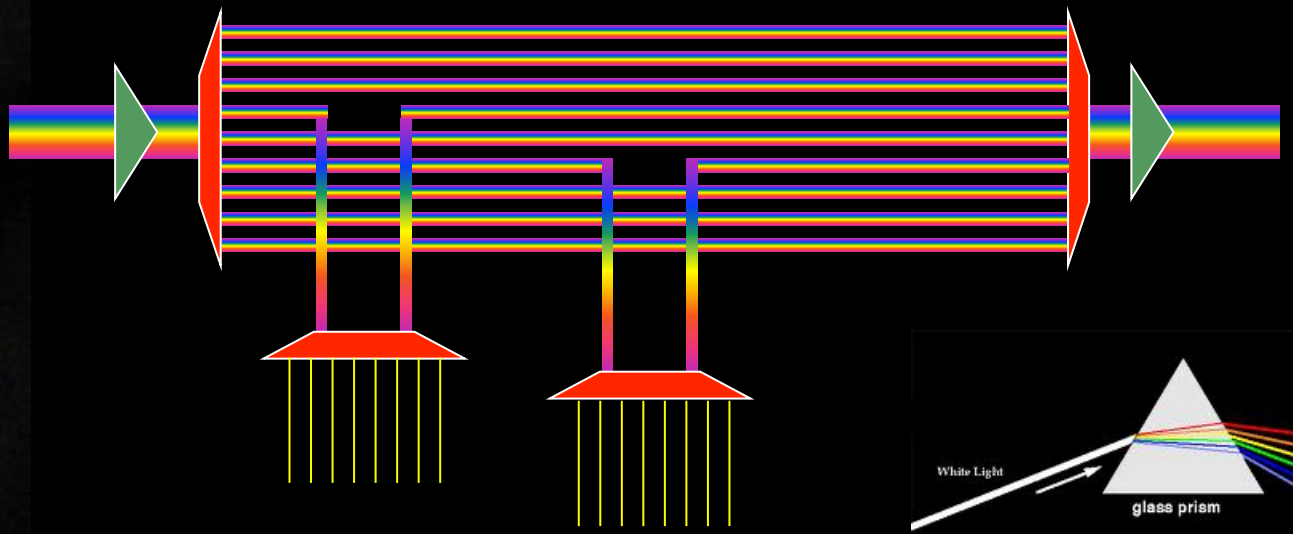
Wireless Networks

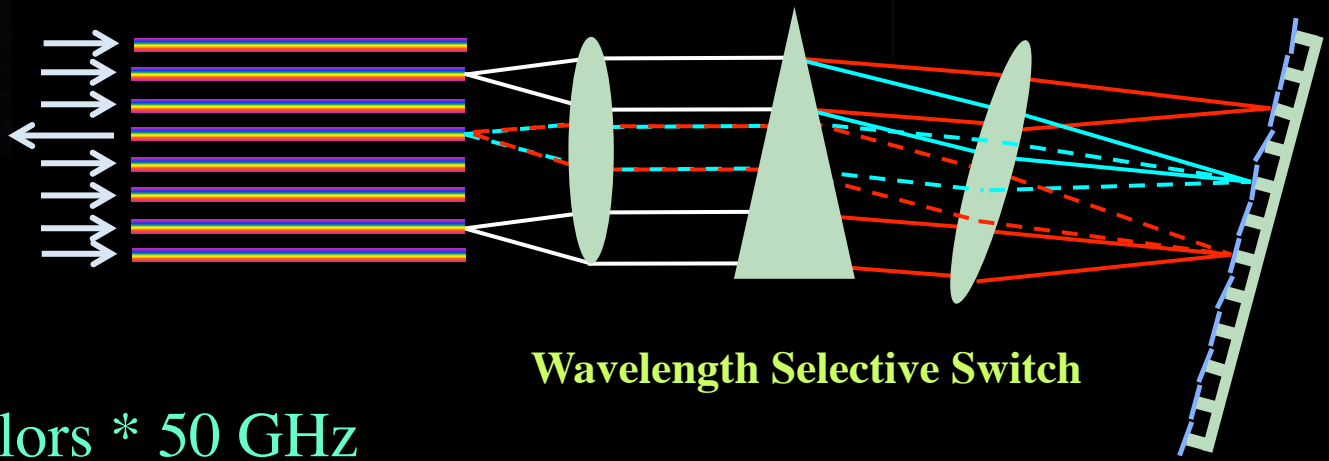


protocol LAN due to the easy comparison and convenience in the digital home. While consumer PC products has just started to migrate to a much higher bandwidth of 802.11n wireless LAN now working on next-generation standard definition is already in progress.

Multiple colors / Fiber







Wavelength Selective Switch

Per fiber: ~ 80-100 colors * 50 GHz

Per color: 10 – 40 – 100 Gbit/s

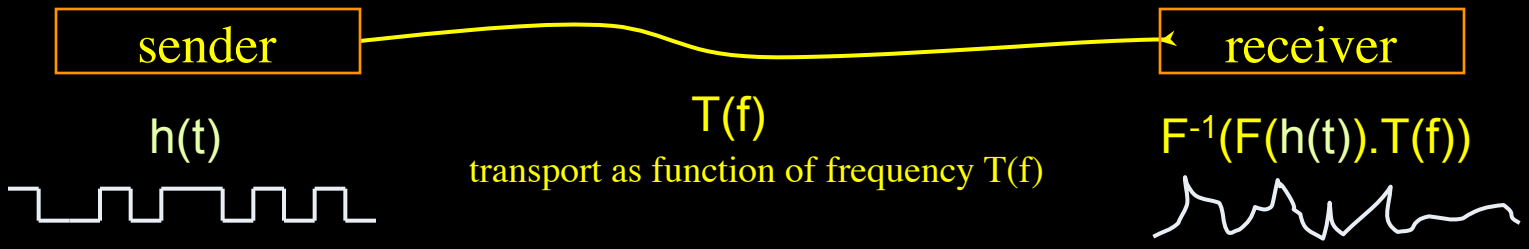
BW * Distance ~ $2*10^{17}$ bm/s

New: Hollow Fiber!

➔ less RTT!

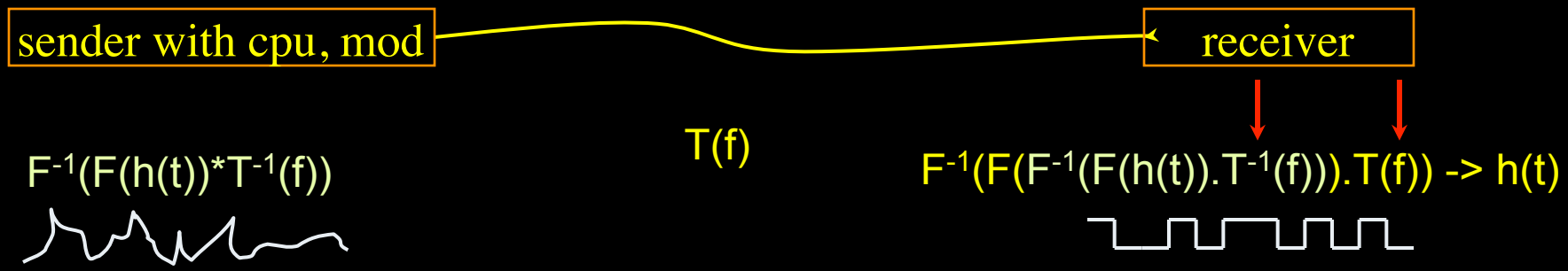
Dispersion compensating modem: eDCO from NORTEL

(Try to Google eDCO :-)

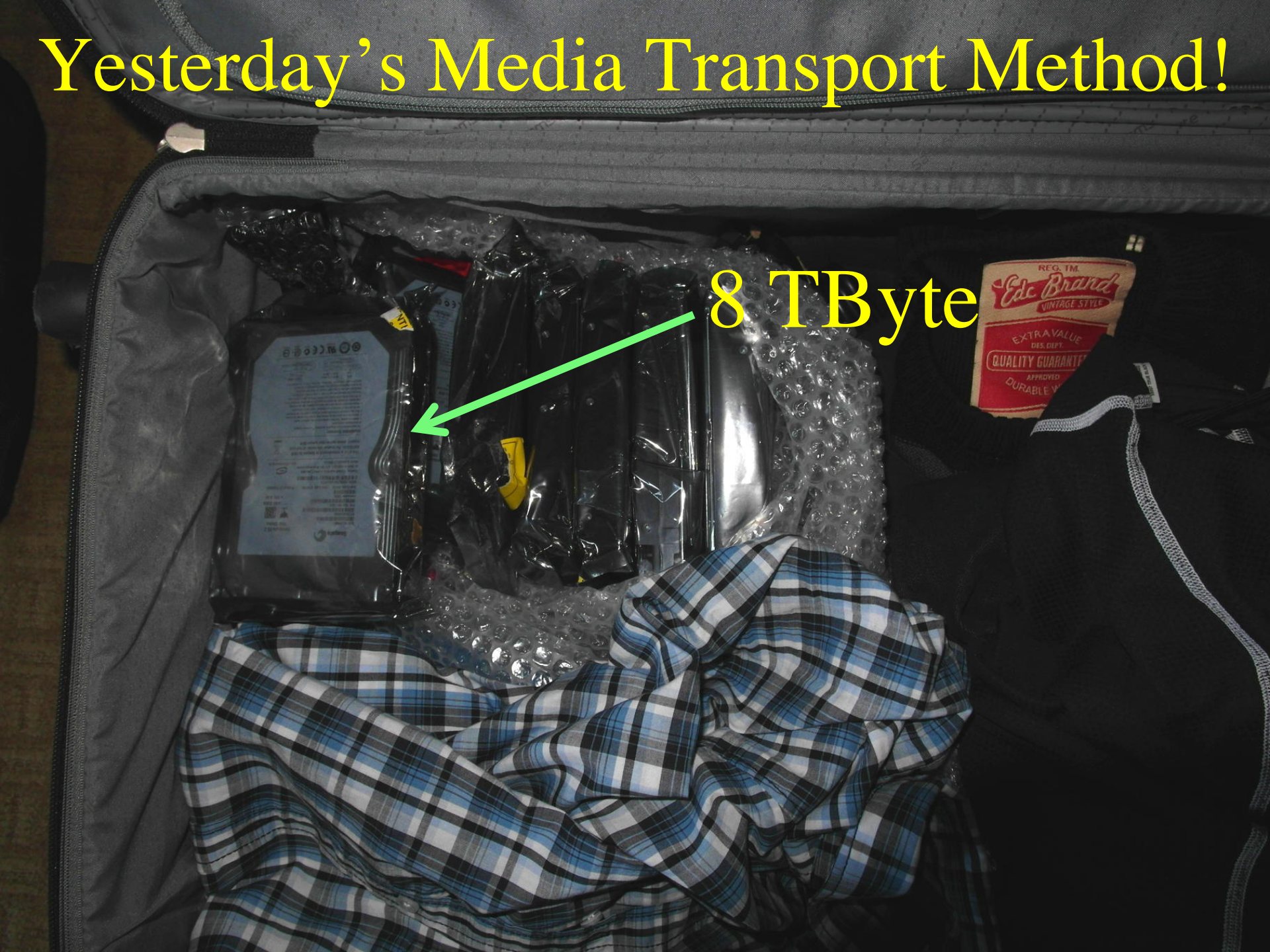


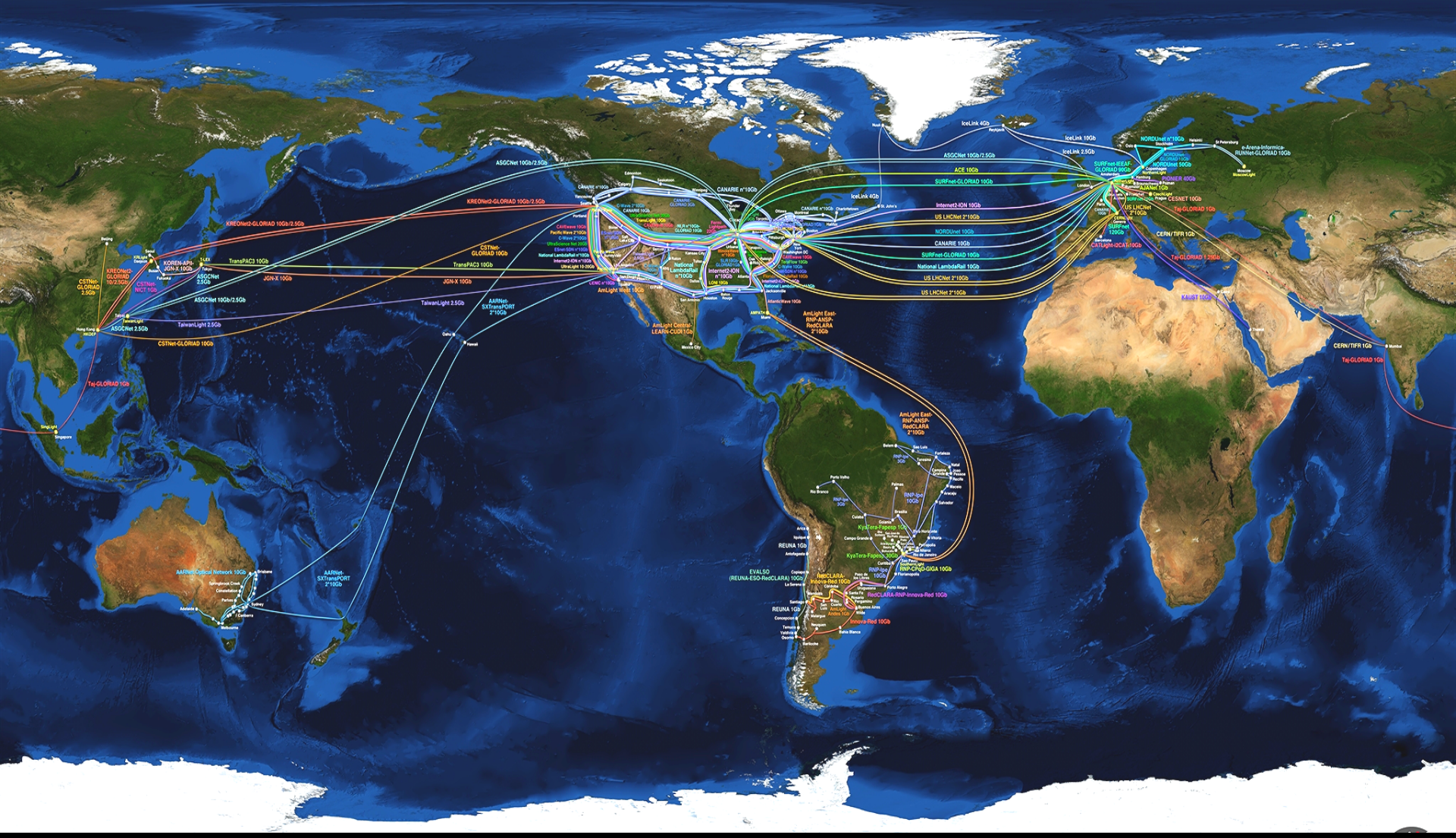
Solution in 5 easy steps for dummy's :

1. try to figure out T(f) by trial and error
2. invert T(f) -> $T^{-1}(f)$
3. computationally multiply $T^{-1}(f)$ with Fourier transform of bit pattern to send
4. inverse Fourier transform the result from frequency to time space
5. modulate laser with resulting h'(t) = $F^{-1}(F(h(t)).T^{-1}(f))$

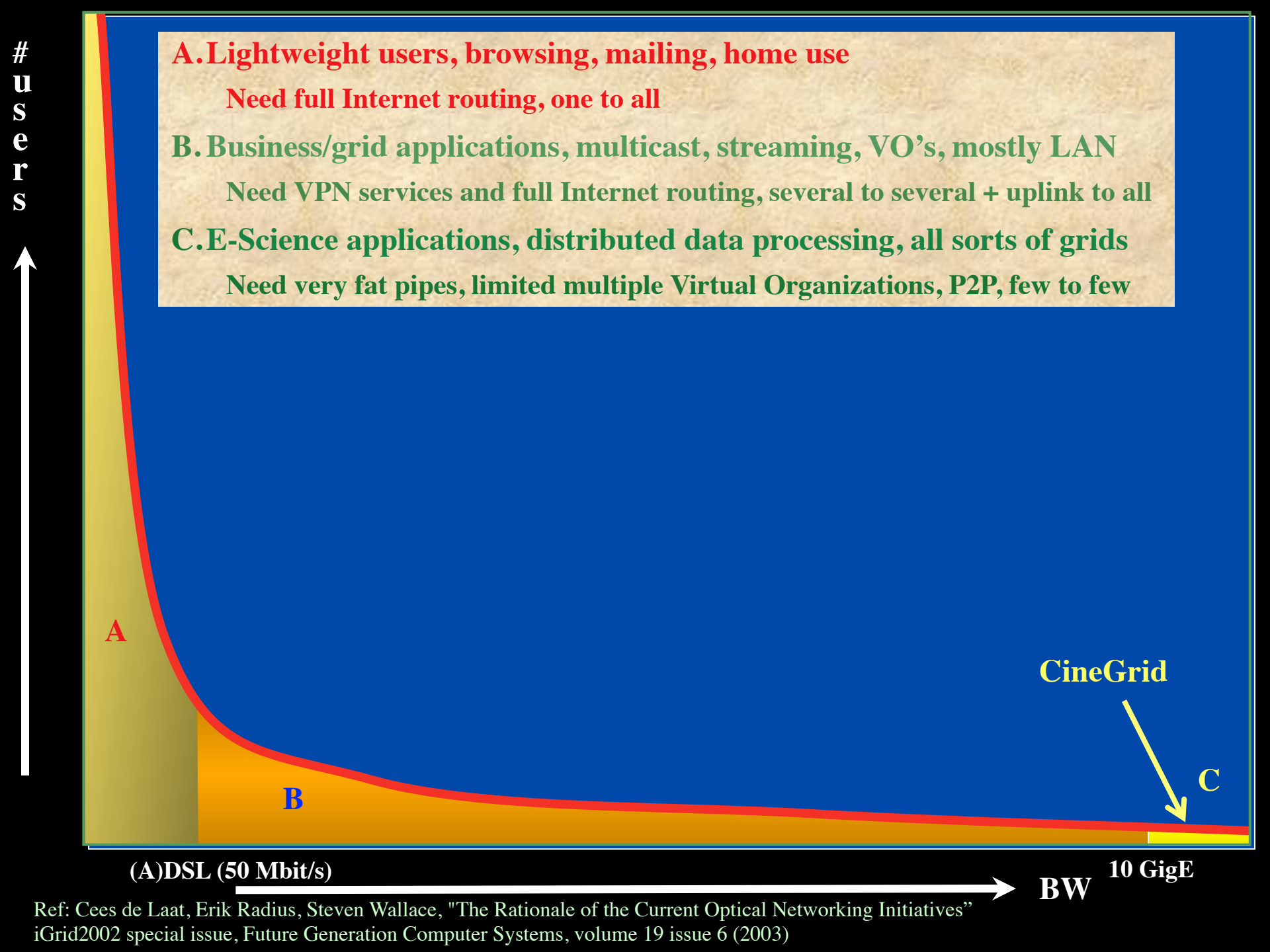


(ps. due to power ~ square E the signal to send **looks** like uncompensated received but is not)


Yesterday's Media Transport Method!
8 TByte
REG. TM.
Ede Brand
VINTAGE STYLE
EXTRA VALUE
QUALITY GUARANTEE
APPROVED
DURABLE



We investigate: TOMTOM for complex networks!

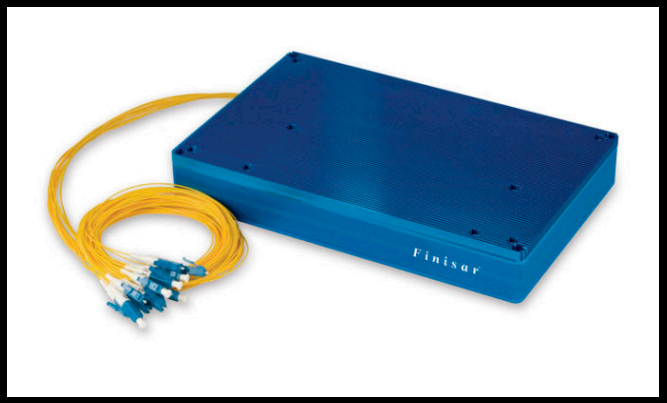


Ref: Cees de Laat, Erik Radius, Steven Wallace, "The Rationale of the Current Optical Networking Initiatives" iGrid2002 special issue, Future Generation Computer Systems, volume 19 issue 6 (2003)

Towards Hybrid Networking!

- Costs of photonic equipment 10% of switching 10 % of full routing
  - for same throughput!
  - Photonic vs Optical (optical used for SONET, etc, 10-50 k$/port)
  - DWDM lasers for long reach are expensive, 10-50 k$
- Bottom line: look for a hybrid architecture which serves all classes in a cost effective way
  - map A -> L3 , B -> L2 , C -> L1 and L2
- Give each packet in the network the service it needs, but no more !

L1 ≈ 1-5 k$/port



L2 ≈ 5-10 k$/port

L3 ≈ 75+ k$/port

Alien light From idea to realisation!

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

NCF

Alien wavelength advantages

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

Alien wavelength challenges

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

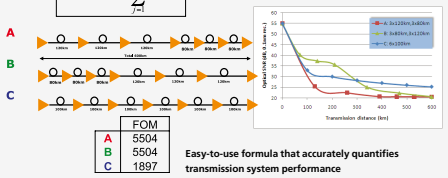
In this demonstration we will investigate the performance of a 40Gb/s PM-QPSK alien wavelength installed on a 10Gb/s DWDM infrastructure.

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

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

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

Lj, span losses in dB
N, number of spans

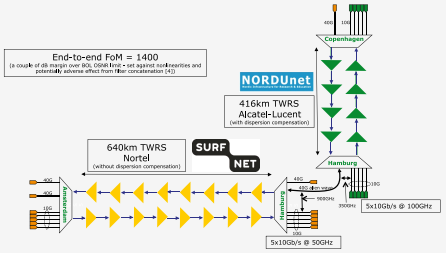


	FOM
A	5504
B	5504
C	1897

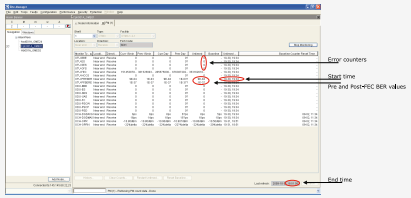
Easy-to-use formula that accurately quantifies transmission system performance

Transmission system setup

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



Test results



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

Conclusions

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

NORTEL NORDUnet telindus together with ISTY SURF NET

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

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

ClearStream @ TNC2011

Setup codename:
FlightCees

UvA

iPerf
I7 3.2 GHz Q-core

iPerf
Amd Ph II 3.6 GHz HexC

Mellanox

40G E

Copenhagen

iPerf
2* dual 2.8 GHz Q-core

iPerf

Mellanox

CIENA
OME
6500

CERN

Hamburg

CIENA DWDM

Alcatel DWDM

CIENA
OME
6500

CIENA
OME
6500

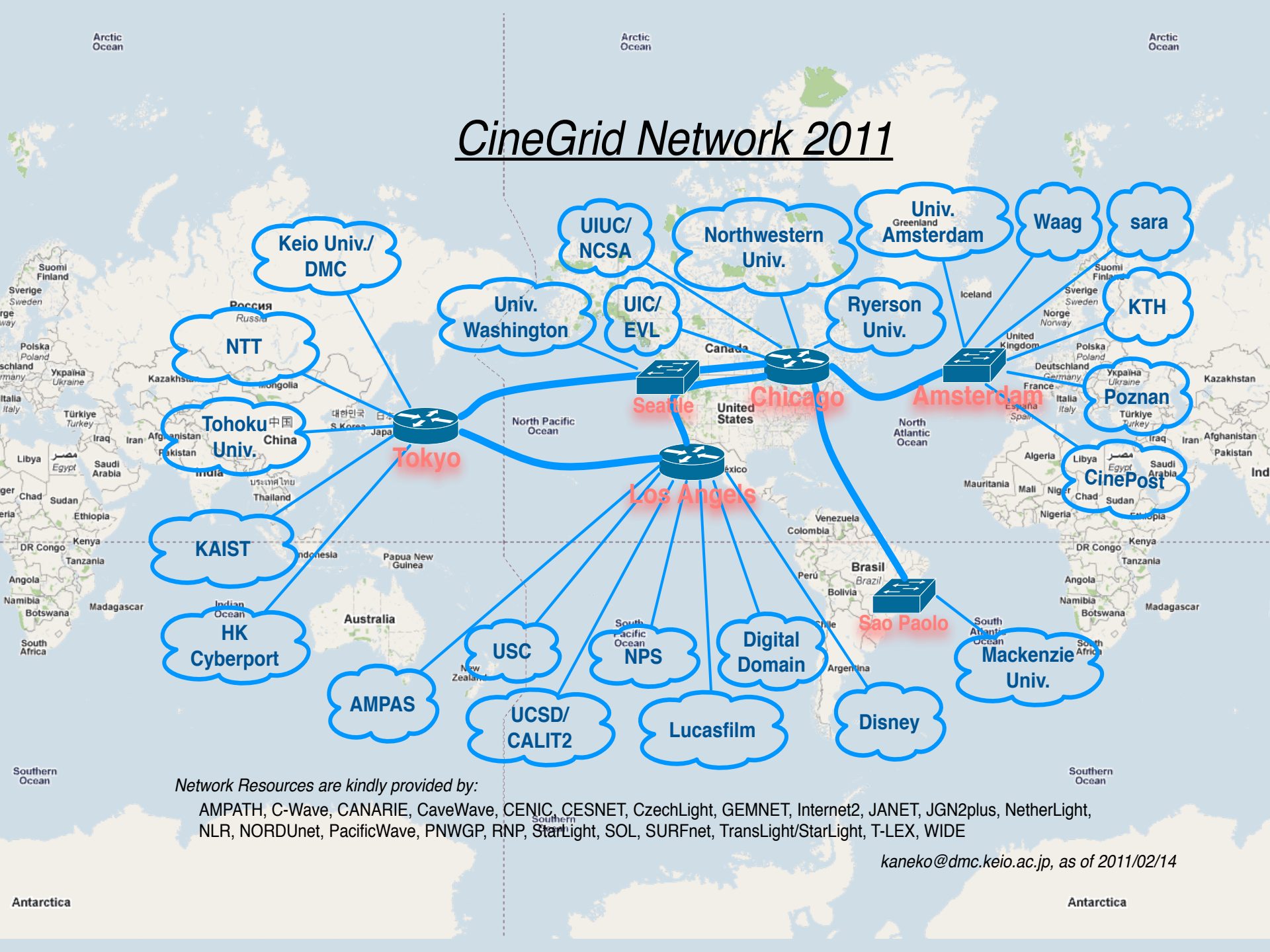
17 ms RTT

27 ms RTT

LH

LH

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

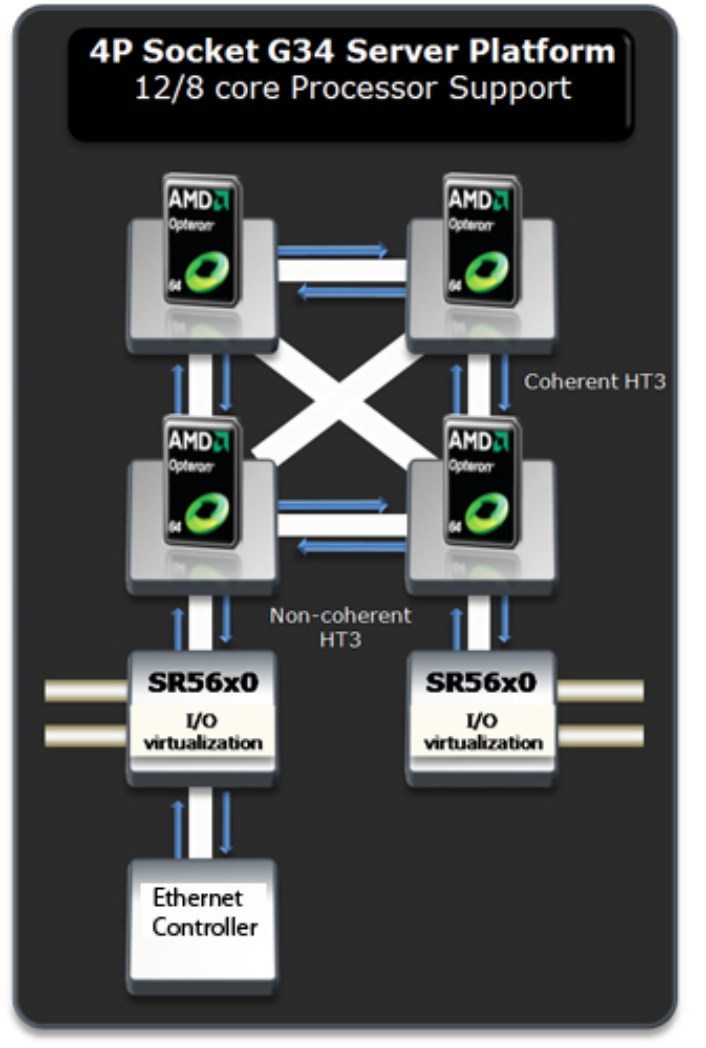

CineGrid Network 2011
Keio Univ./ DMC
NTT
Tohoku Univ.
KAIST
HK Cyberport
AMPAS
USC
UCSD/ CALIT2
NPS
Lucasfilm
Digital Domain
Disney
Mackenzie Univ.
UIUC/ NCSA
Univ. Washington
UIC/ EVL
Northwestern Univ.
Ryerson Univ.
Univ. Amsterdam
Waag
sara
KTH
Poznan
CinePost
Tokyo
Seattle
Chicago
Los Angels
Amsterdam
Sao Paolo
Network Resources are kindly provided by:
AMPATH, C-Wave, CANARIE, CaveWave, CENIC, CESNET, CzechLight, GEMNET, Internet2, JANET, JGN2plus, NetherLight, NLR, NORDUnet, PacificWave, PNWGP, RNP, StarLight, SOL, SURFnet, TransLight/StarLight, T-LEX, WIDE
kaneko@dmc.keio.ac.jp, as of 2011/02/14

100 Gb Network for CineGrid @ TNC12

Results (rtt = 17 ms)

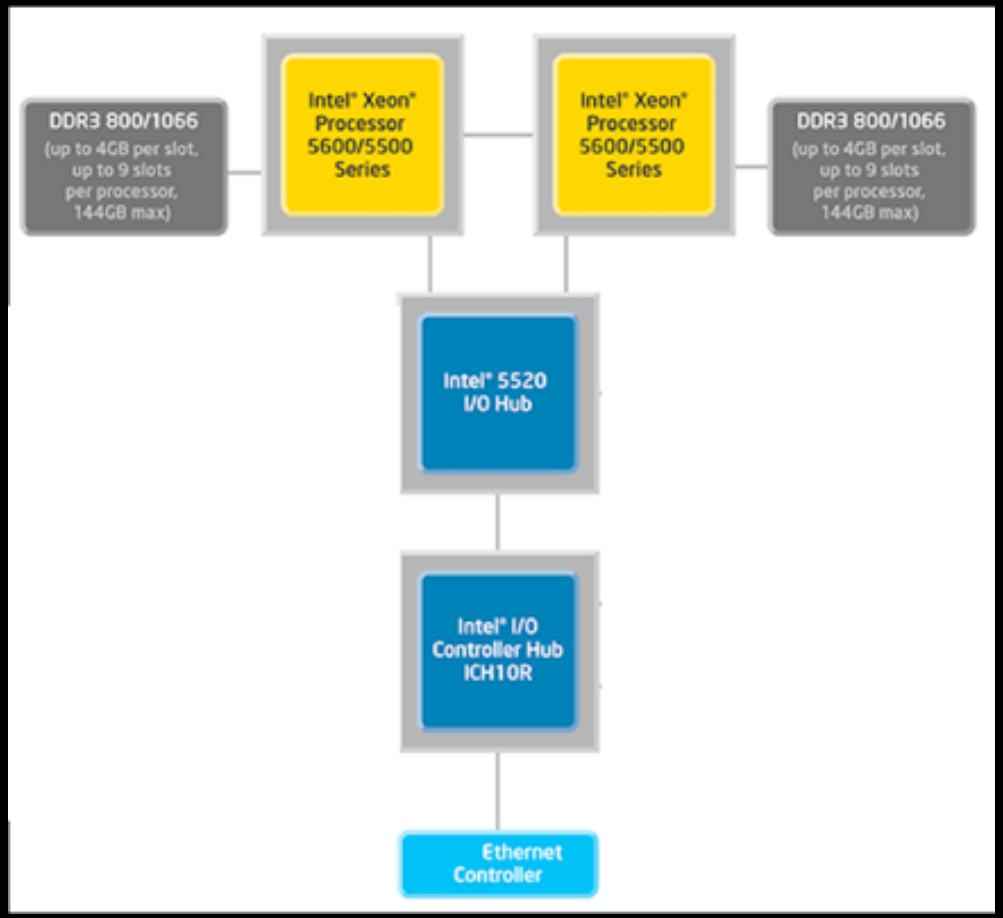
- Single flow iPerf 1 core -> 21 Gbps
- Single flow iPerf 1 core <> -> 15+15 Gbps
- Multi flow iPerf 2 cores -> 25 Gbps
- Multi flow iPerf 2 cores <> -> 23+23 Gbps
- DiViNe <> -> 11 Gbps
- Multi flow iPerf + DiVine -> 35 Gbps
- Multi flow iPerf + DiVine <> -> 35 + 35 Gbps

Server Architecture



DELL R815

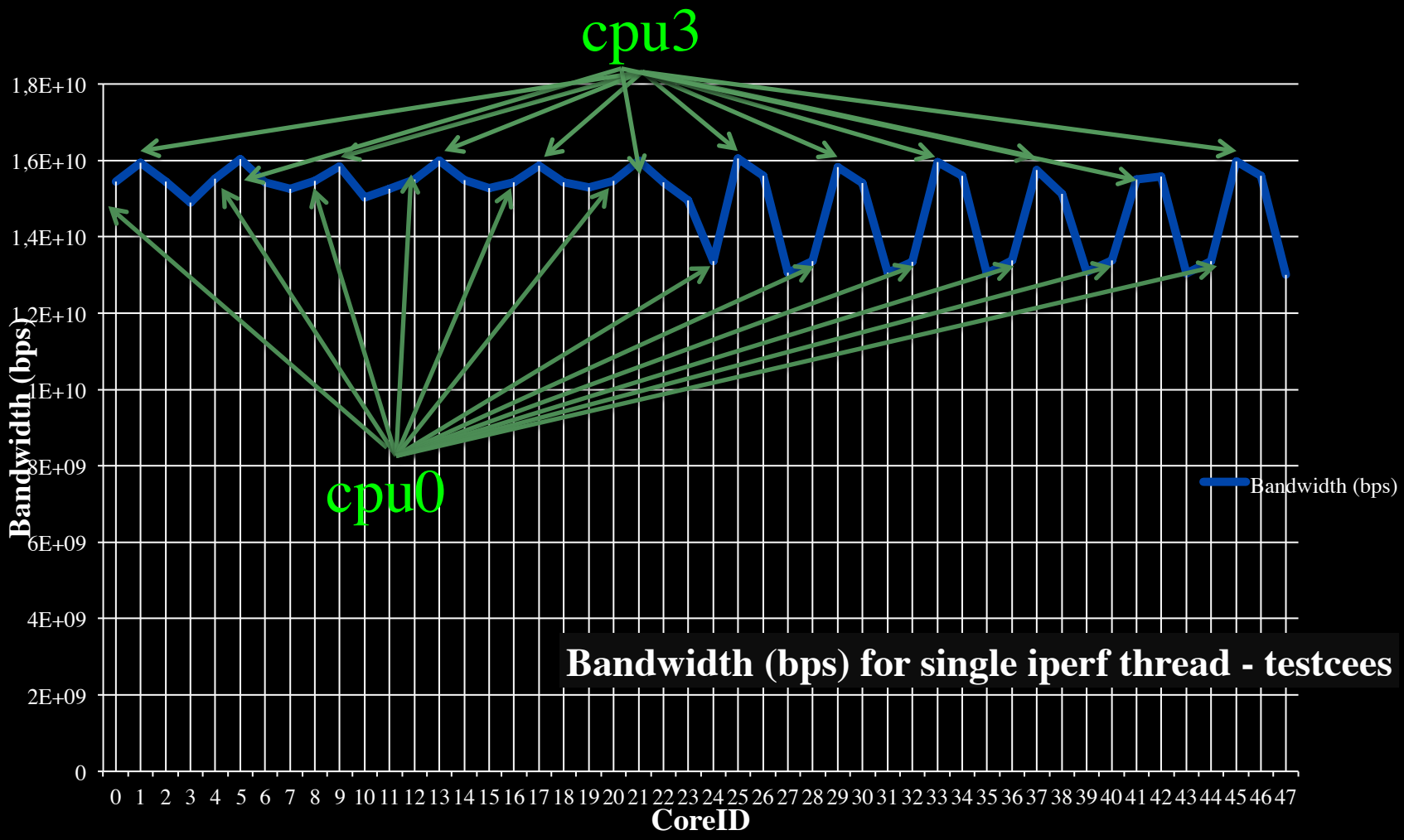
4 x AMD Opteron 6100



Supermicro X8DTT-HIBQF

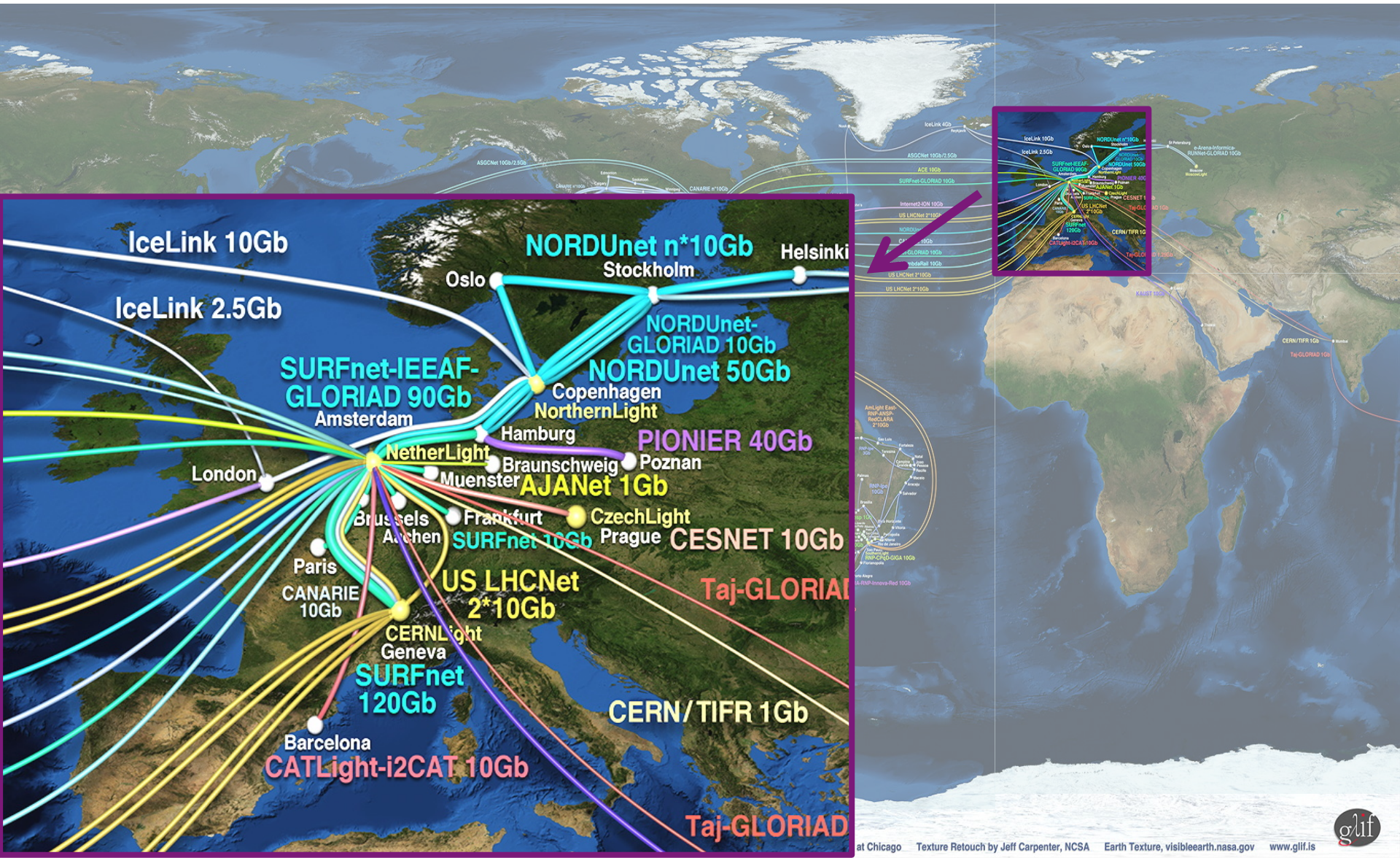
2 x Intel Xeon

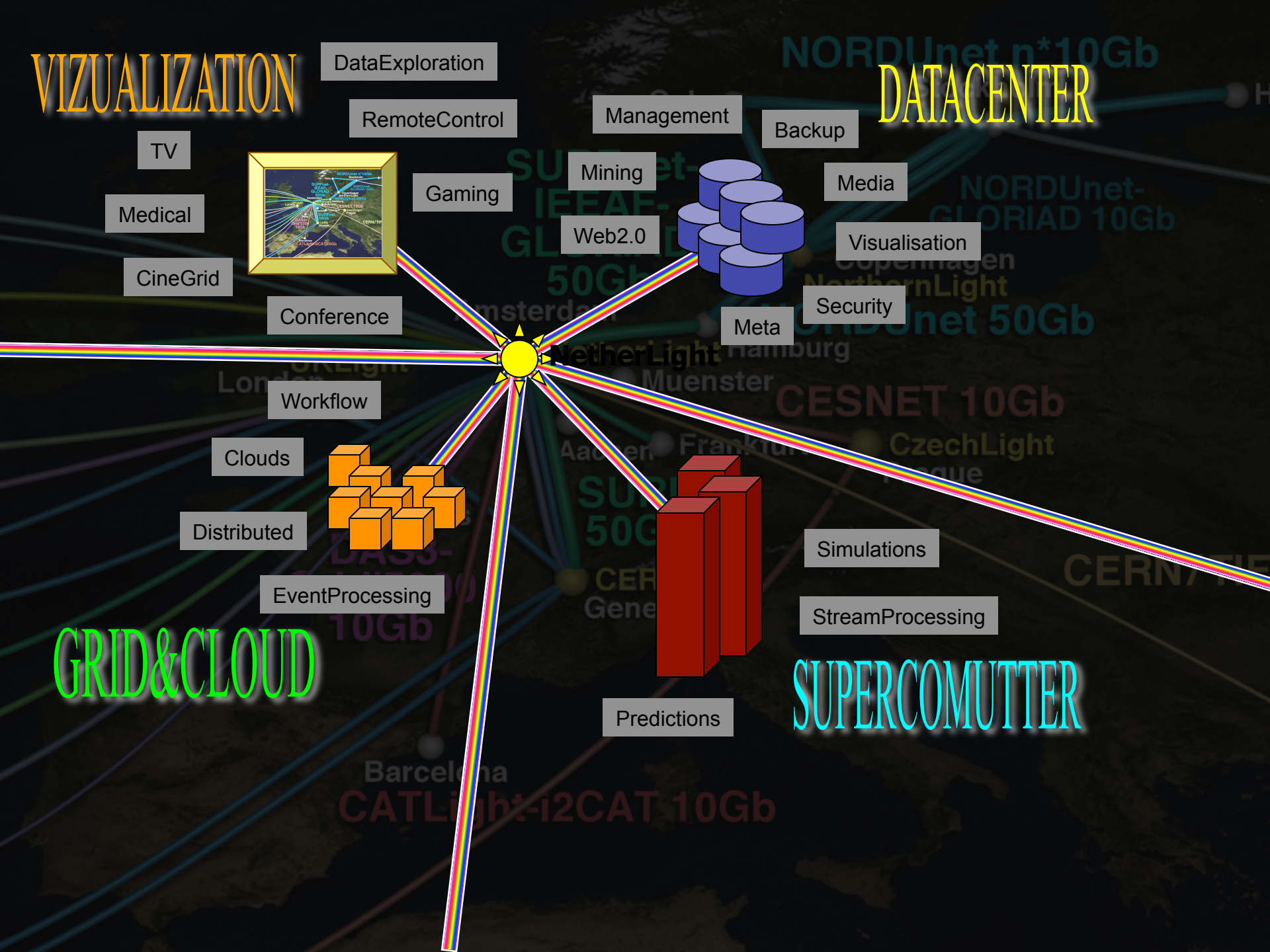
CPU Topology benchmark



We used numactl to bind iperf to cores

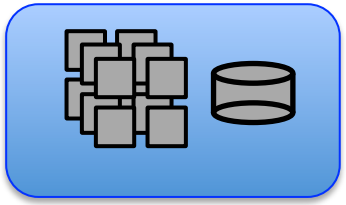
The GLIF – lightpaths around the world



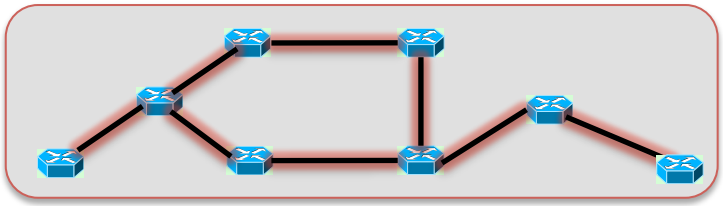

VIZUALIZATION
DataExploration
RemoteControl
TV
Medical
CineGrid
Gaming
Conference
DATACENTER
Management
Backup
Mining
Media
Web2.0
Visualisation
Security
Meta
NetherLight
Workflow
Clouds
Distributed
EventProcessing
GRID&CLOUD
Simulations
StreamProcessing
Predictions
SUPERCOMUTTER

IaaS : Clouds and Network Virtualization

Virtual Compute and Storage Infrastructure



Virtual Network Infrastructure



Cloud APIs (Amazon EC2 ..)

Network Provisioning APIs (NLR Sherpa, DOE OSCARS, Internet2 DRAGON, OGF NSI …)

Cloud Providers

NLR

ESnet
Energy Sciences Network

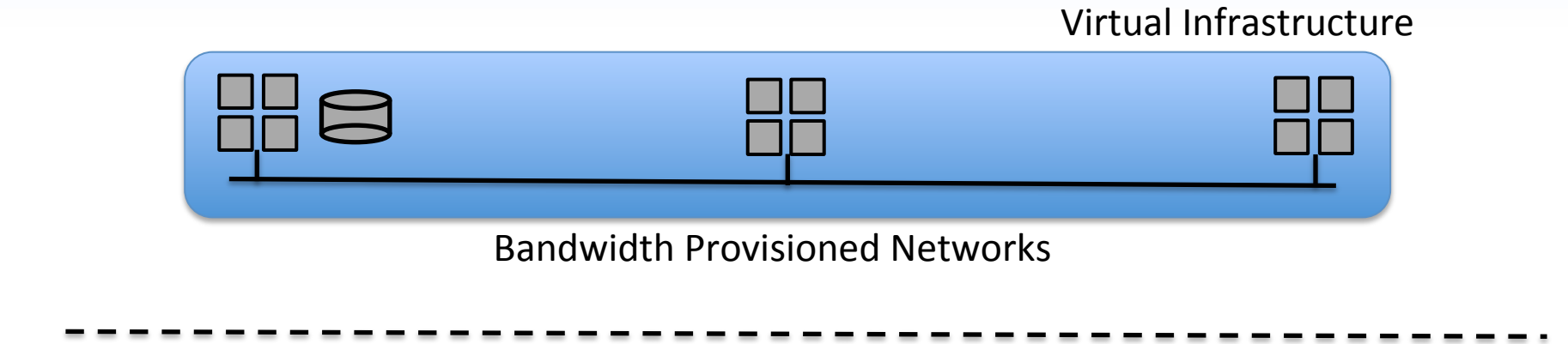
INTERNET2

BEN
Breakable Experimental Network

Network Transit Providers

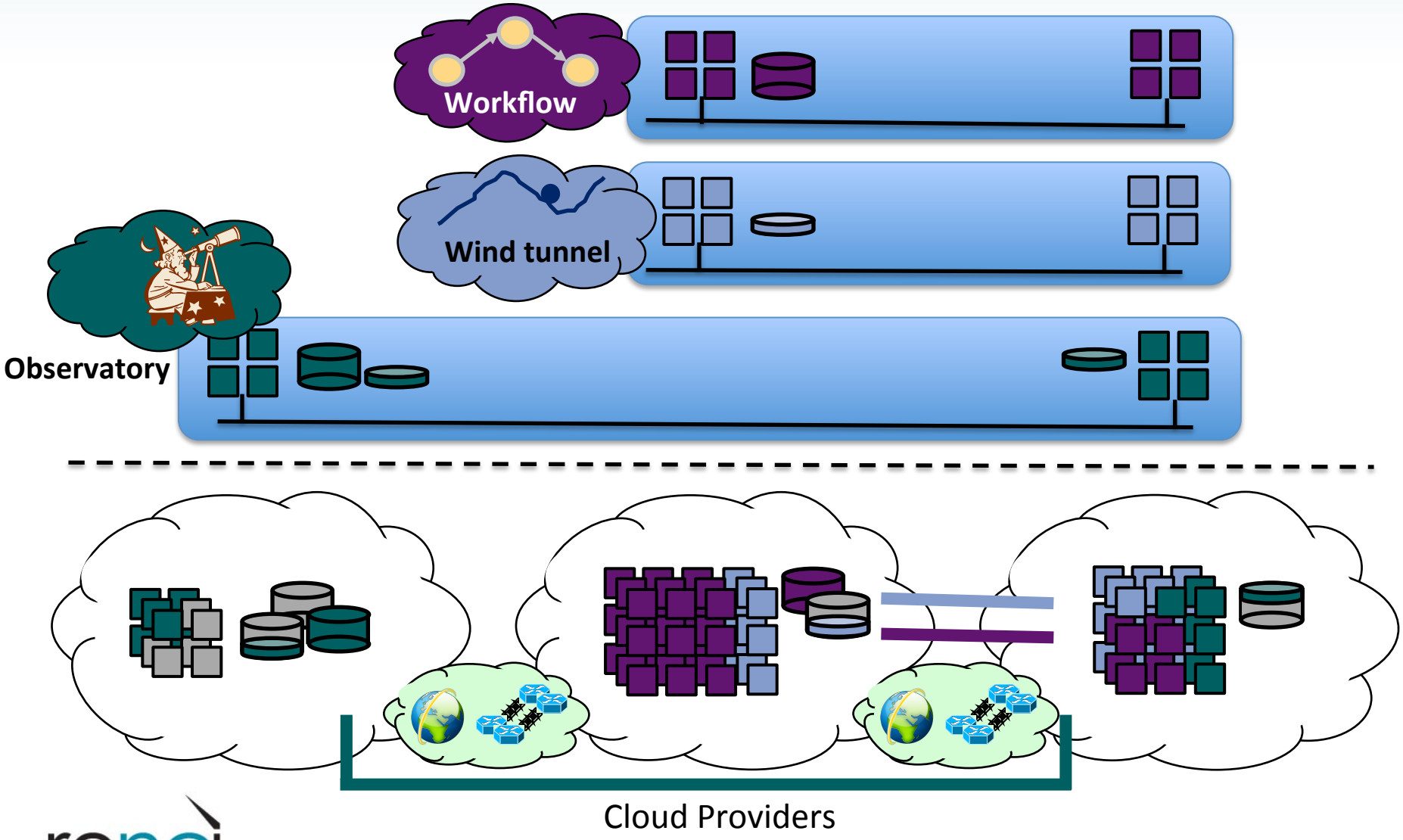
renci

IaaS: Networked Clouds



renci

IaaS: Networked Clouds

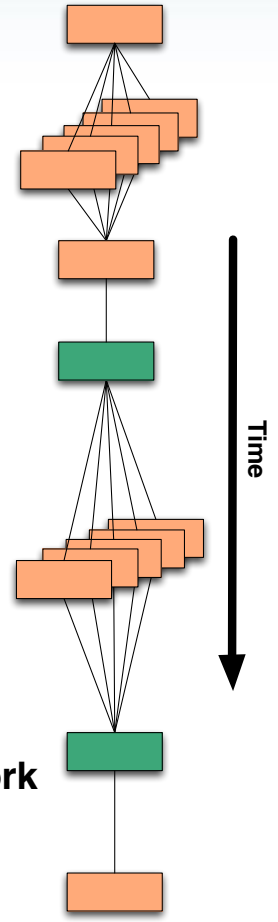


renci

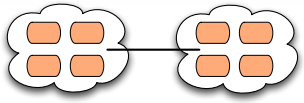
Dynamic Workflow Steps (On-Ramp)

Workflow

1. Start Workflow
2. Compute-intensive Workflow Step
3. Sync Step
4. Dynamically Create High-bandwith Network
5. Data-intensive Workflow Step
6. Dynamically Destroy High-bandwith Network
7. End Workflow

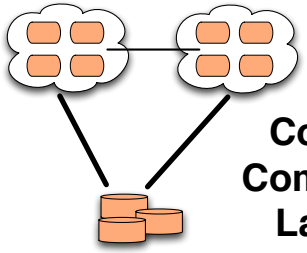


Elastic Slice



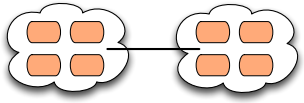
Compute VMs with Network Connection

Data intensive workflow entering a stage of high demand for large data set residing on a remote resource



High-bandwidth Connections between Compute Resources and Large Static Data Set

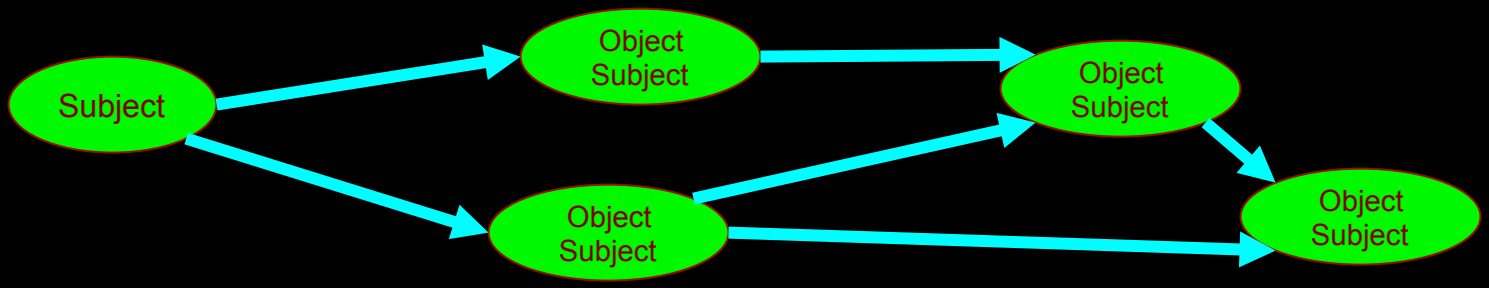
Data intensive workflow leaving a stage of high demand for large data

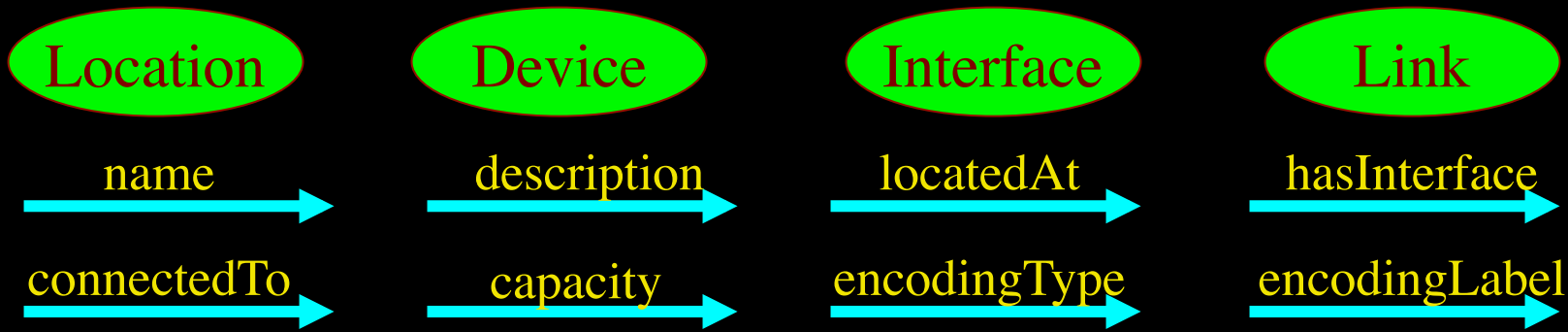


renci

LinkedIN for Infrastructure

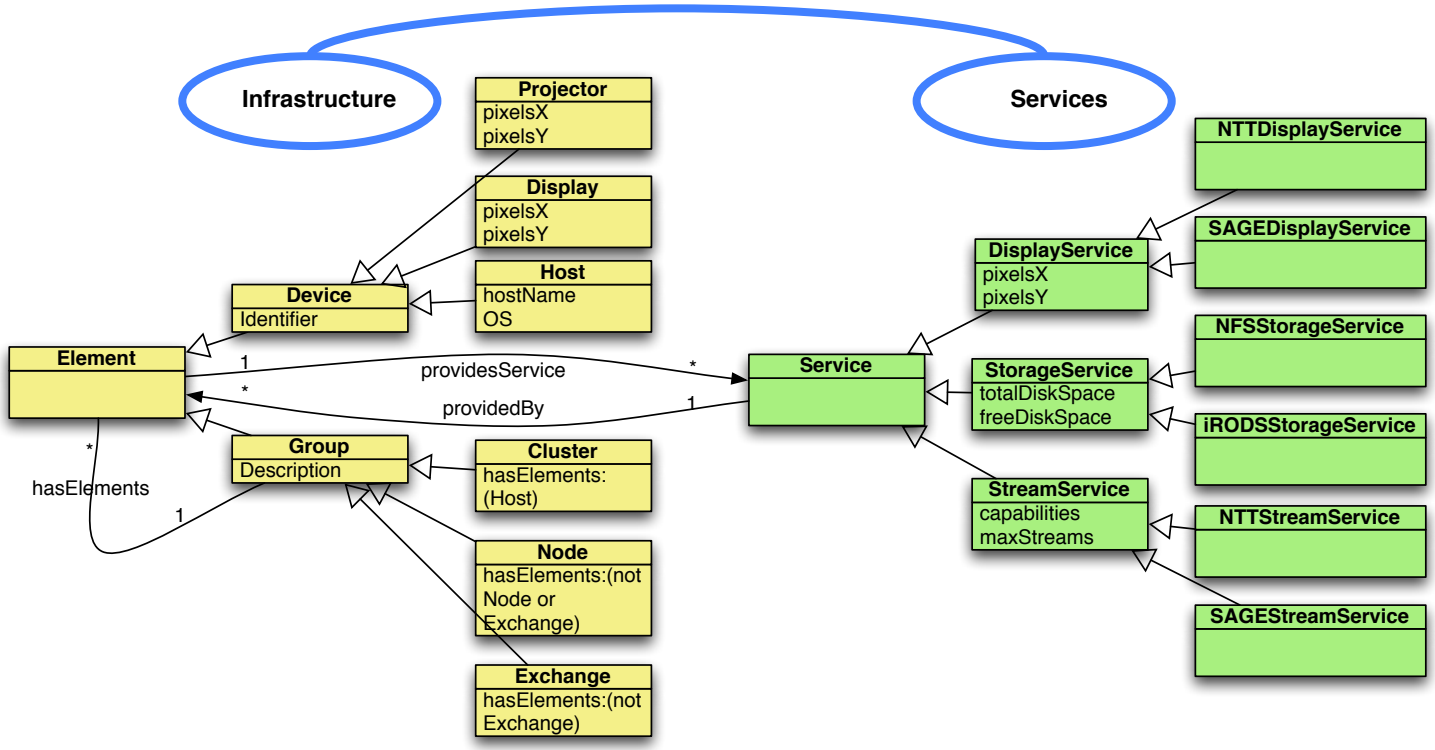
- From semantic Web / Resource Description Framework.
- The RDF uses XML as an interchange syntax.
- Data is described by triplets (Friend of a Friend):





Information Modeling

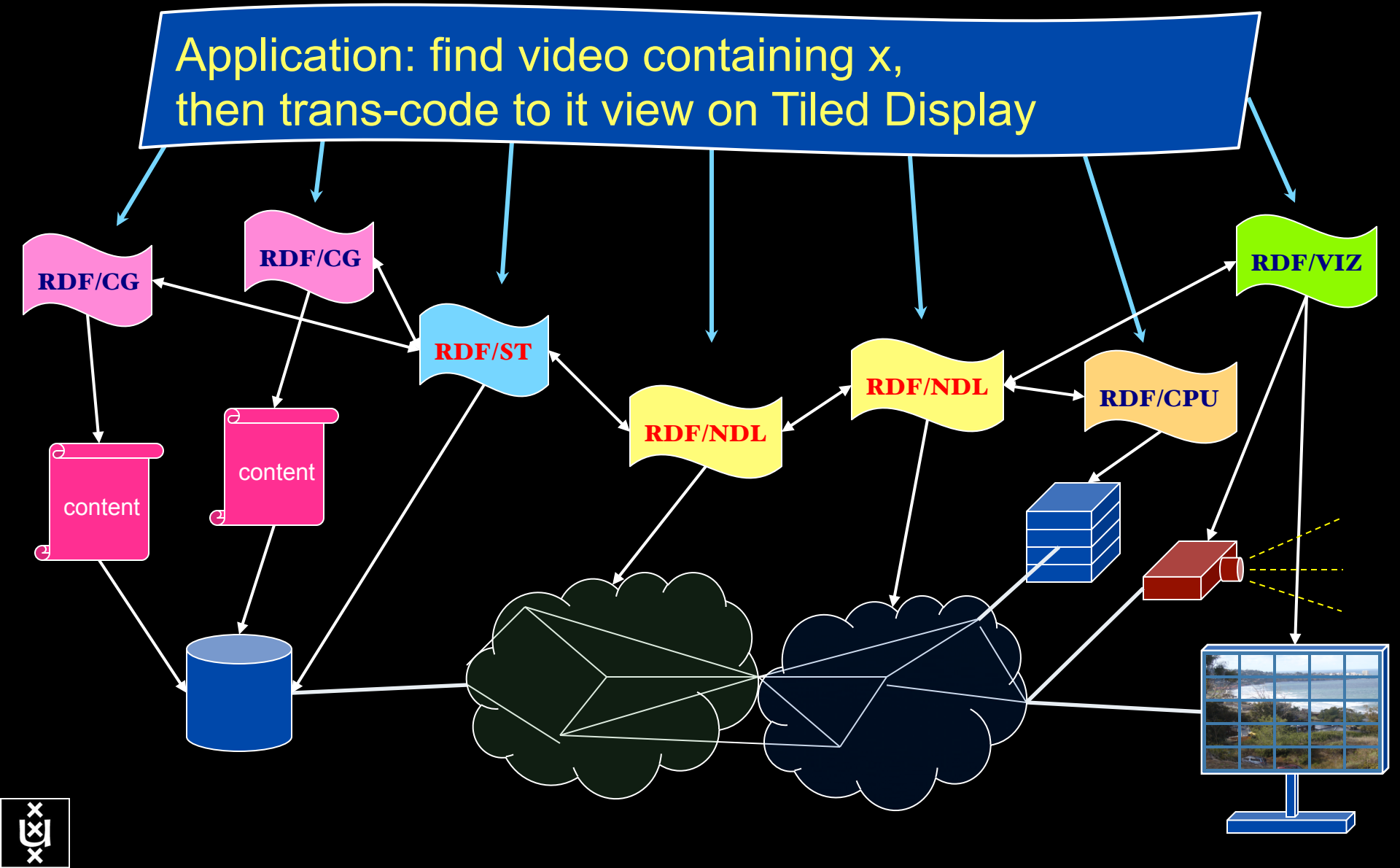
Define a common information model for ***infrastructures*** and ***services***. Base it on Semantic Web.



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"



Direction

- Distributed Comp -> Grid -> Cloud -> Big Data
- What is next? The App…
- Lego Block approach
- Application as a Service
- Elastic Cloud
- Determinism & Real Time?
- CineGrid ToolBox

Variable Rate Music in the streets of Amsterdam!

Why?

I want to:

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

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

The Ten Problems with the Internet

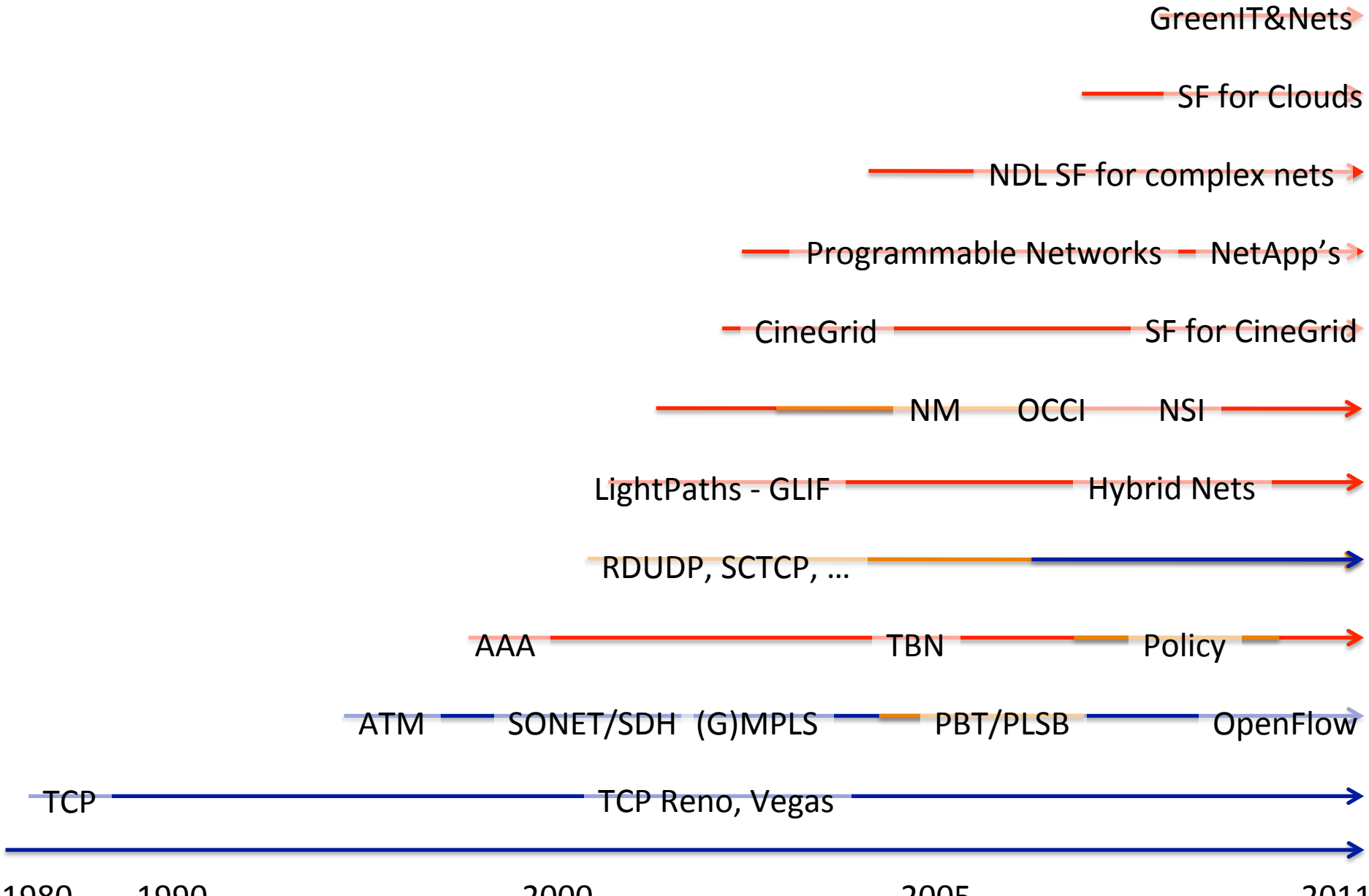
1. **Energy Efficient Communication**
2. Separation of Identity and Address
3. Location Awareness
4. **Explicit Support for Client-Server Traffic and Distributed Services**
5. Person-to-Person Communication
6. Security
7. **Control, Management, and Data Plane separation**
8. **Isolation**
9. Symmetric/Asymmetric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

Nice to have:

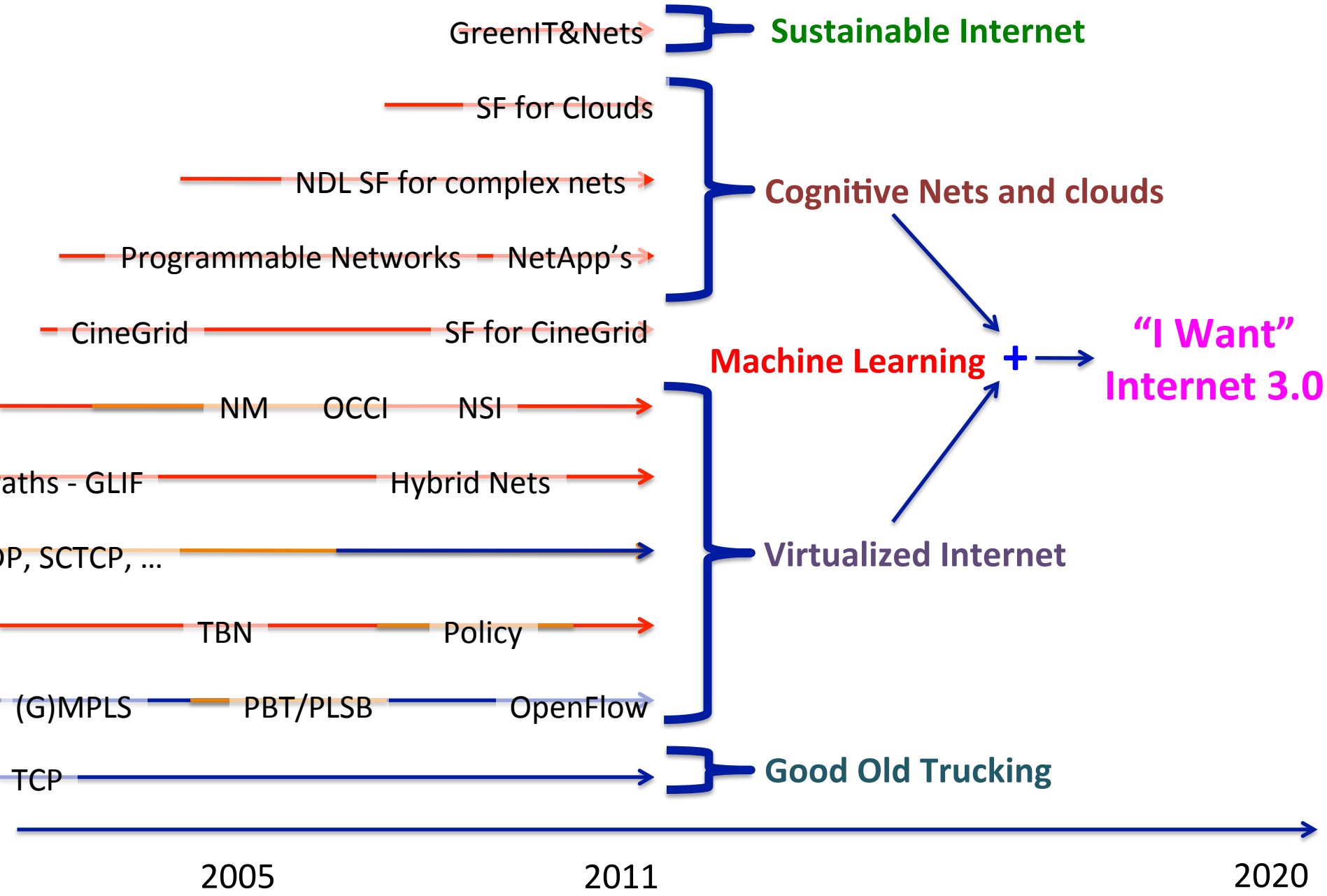
- Global Routing with Local Control of Naming and Addressing
- **Real Time Services**
- **Cross-Layer Communication**
- Manycast
- Receiver Control
- Support for Data Aggregation and Transformation
- **Support for Streaming Data**
- **Virtualization**

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

TimeLine


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

TimeLine


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

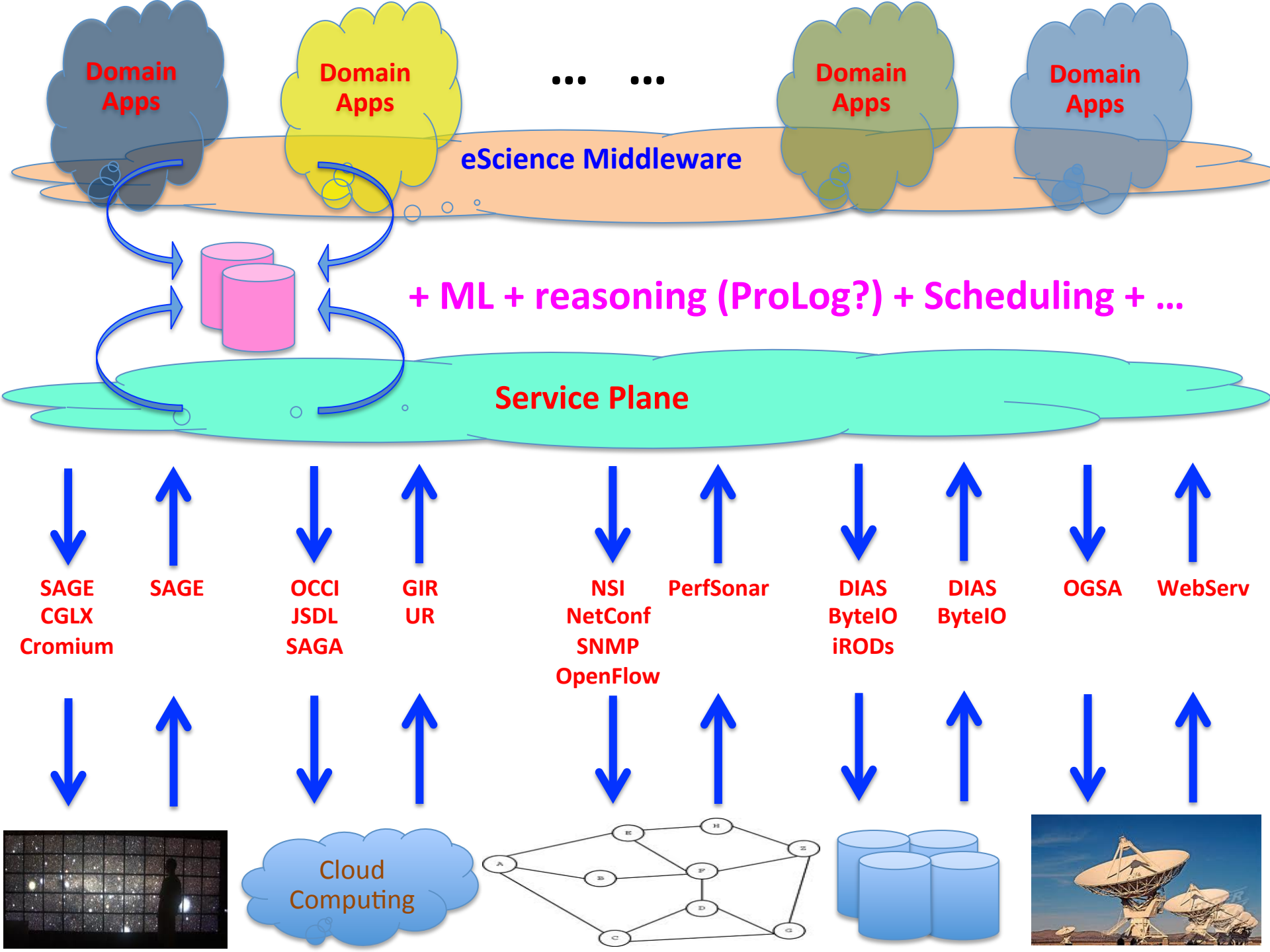
TimeLine

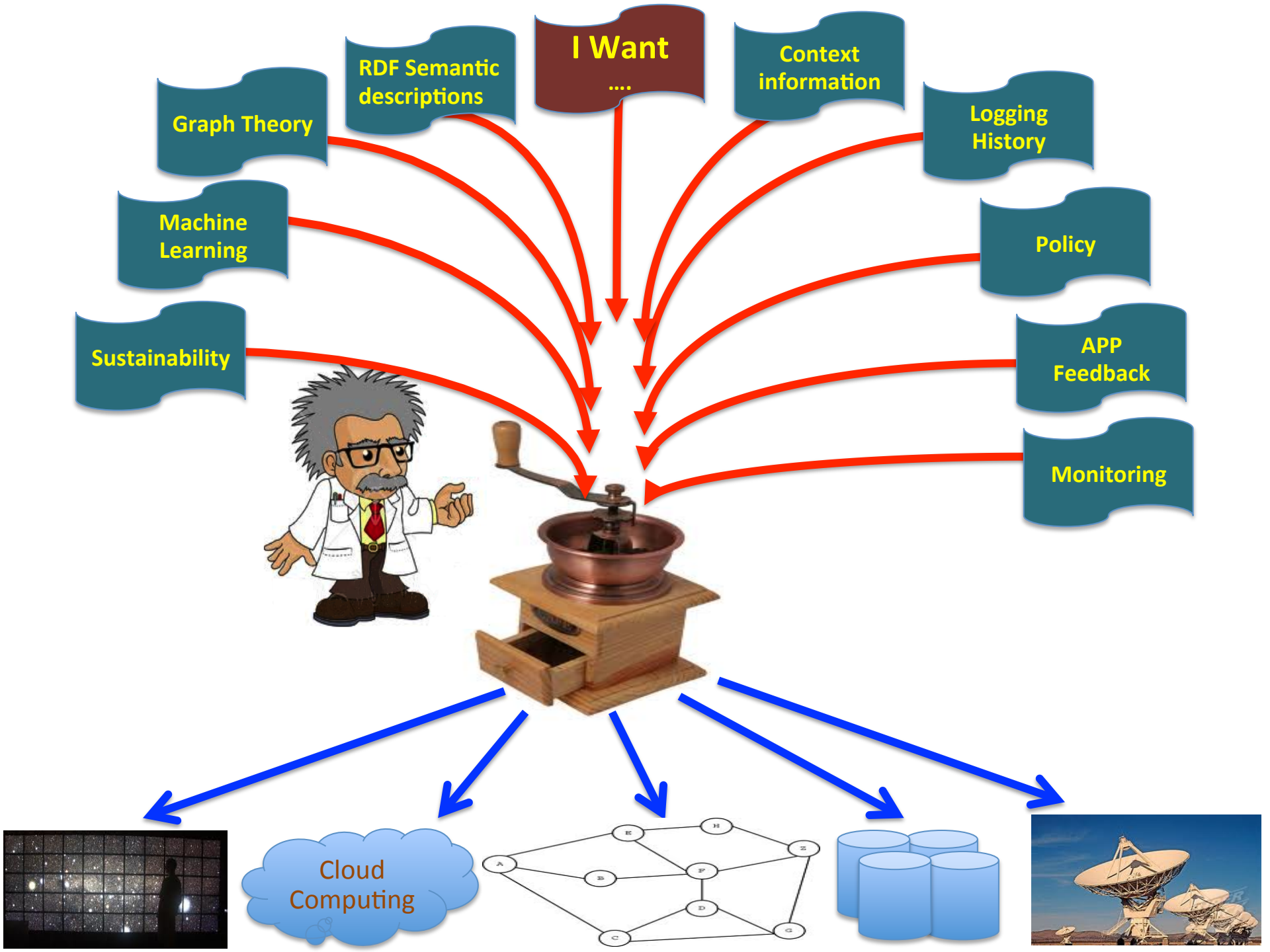
- Sustainable Internet
- Cognitive Nets and clouds
- Machine Learning + → "I Want" Internet 3.0
- Virtualized Internet
- Good Old Trucking

I retire

2020

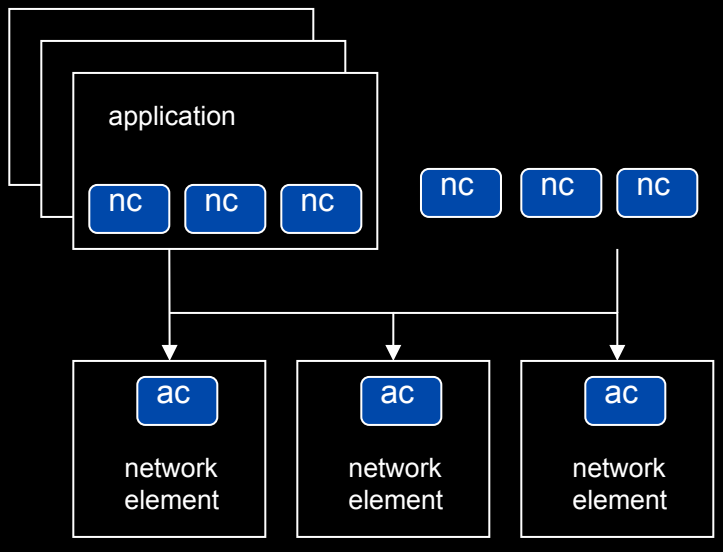
2040

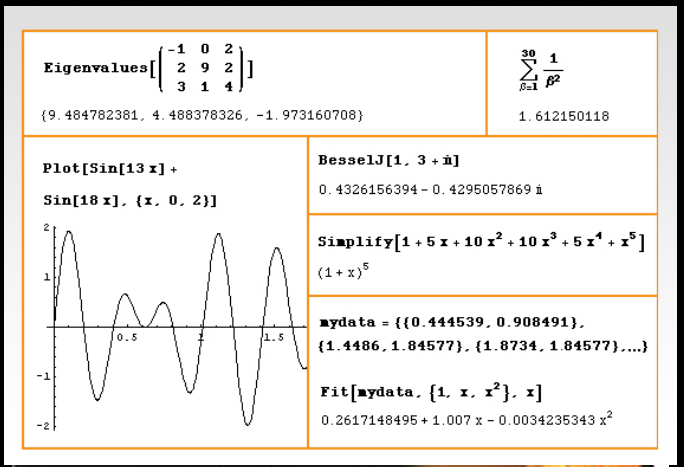

Domain Apps
Domain Apps
... ...
Domain Apps
Domain Apps
eScience Middleware
+ ML + reasoning (ProLog?) + Scheduling + …
Service Plane
SAGE
CGLX
Cromium
SAGE
OCCI
JSDL
SAGA
GIR
UR
NSI
NetConf
SNMP
OpenFlow
PerfSonar
DIAS
ByteIO
iRODs
DIAS
ByteIO
OGSA
WebServ
Cloud Computing

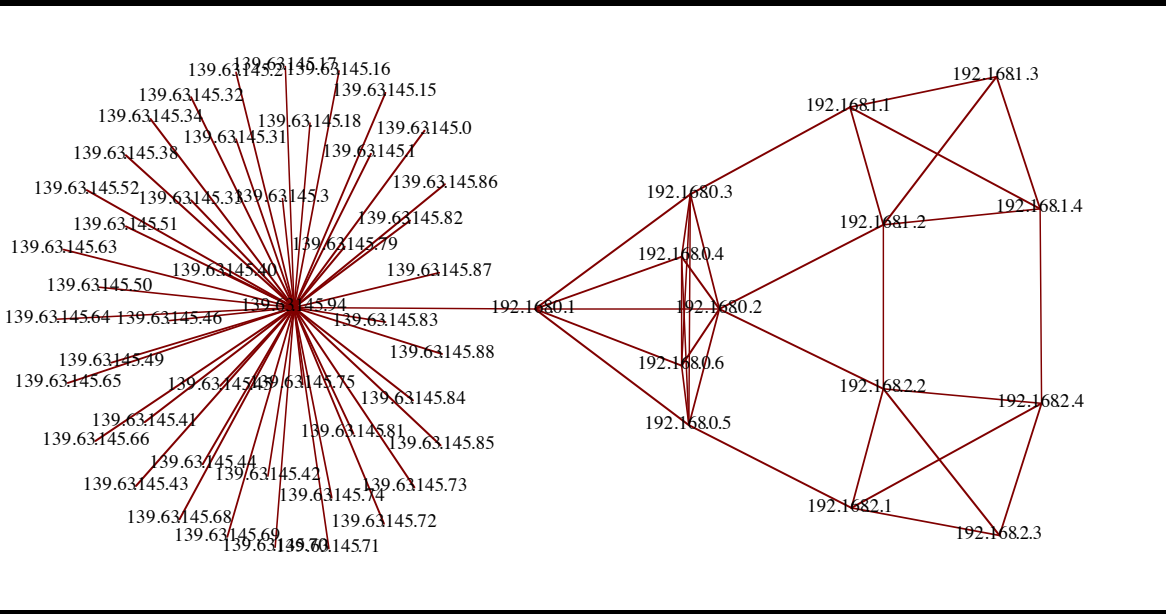

I Want
....
Graph Theory
RDF Semantic descriptions
Context information
Logging History
Machine Learning
Policy
Sustainability
APP Feedback
Monitoring
Cloud Computing

User Programmable Virtualized Networks.

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







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

Greening the Processing System

ECO-Scheduling

Scientific Publications: FGCS Special Issue on CineGrid!

Volume 27, Issue 7,june 2011

Guest Editors: Naohisa Ohta & Paul Hearty & Cees de Laat

Editorial: CineGrid: Super high definition media over optical networks.

1. Real-time long-distance transfer of uncompressed 4K video for remote collaboration.
2. Media Network (HPDMnet): An advanced international research initiative and global experimental testbed.
3. Producing and streaming high resolution digital movies of microscopic subjects.
4. Enabling multi-user interaction in large high-resolution distributed environments.
5. Tri-continental premiere of 4K feature movie via network streaming at FILE 2009.
6. A collaborative computing model for audio post-production.
7. Design and implementation of live image file feeding to dome theaters.
8. Beyond 4K: 8K 60p live video streaming to multiple sites.
9. Using ontologies for resource description in the CineGrid Exchange.
10. CineGrid Exchange: A workflow-based peta-scale distributed storage platform on a high-speed network.
11. CSTP: A parallel data transfer protocol using cross-stream coding.
12. Multi-point 4K/2K layered video streaming for remote collaboration.

Why?

Because we can!

Q & A

I did not talk about:

- Astronomy, LifeSciences on CI
- Knowlegde complexity
- Security & privacy
- AAA
- …

http://ext.delaat.net/

Slides thanks to:

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