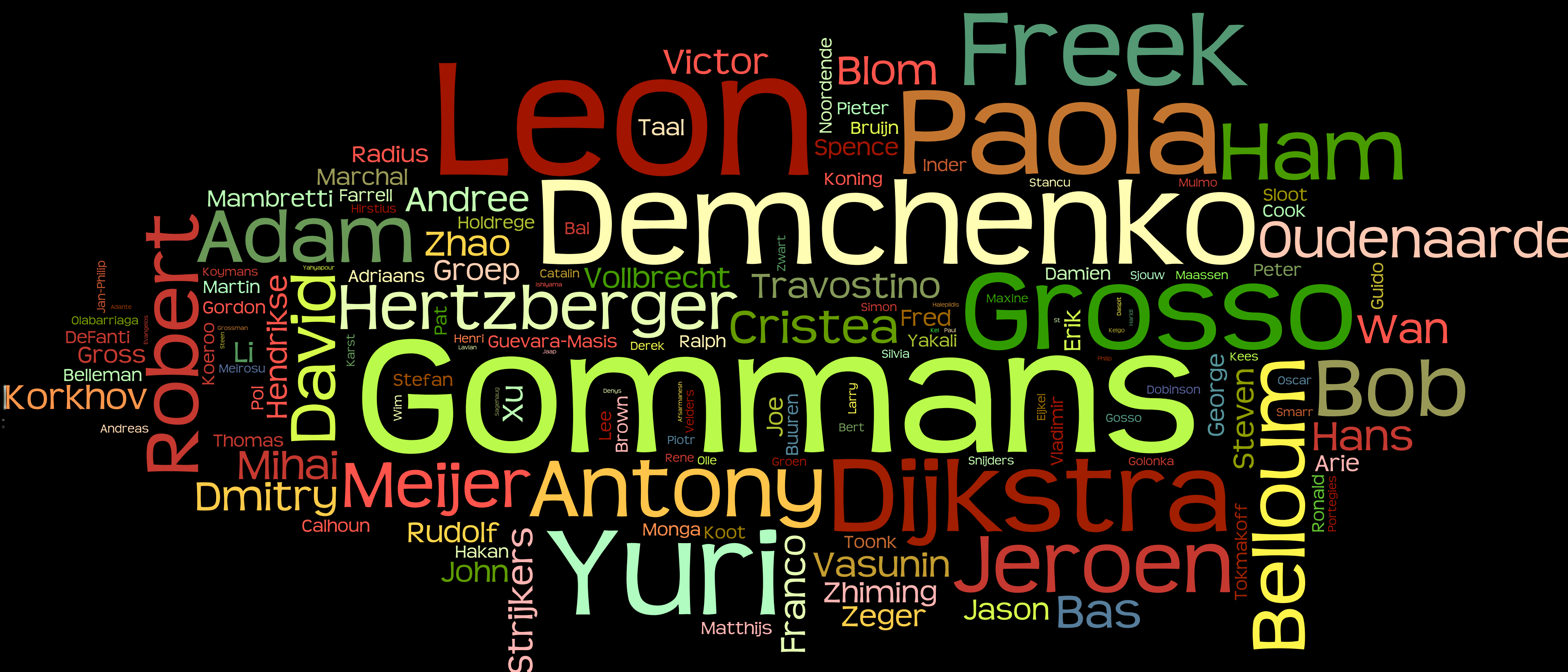
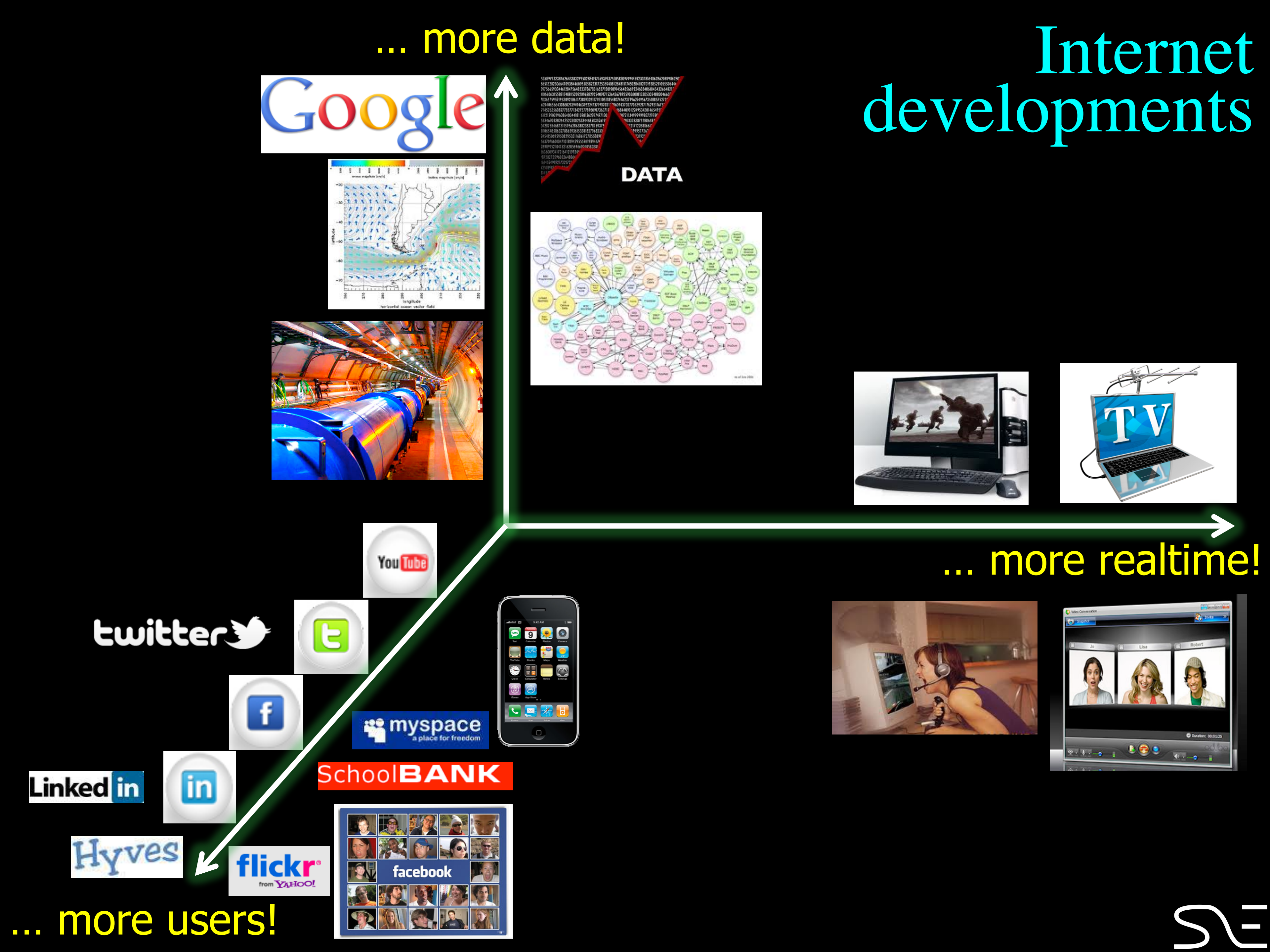
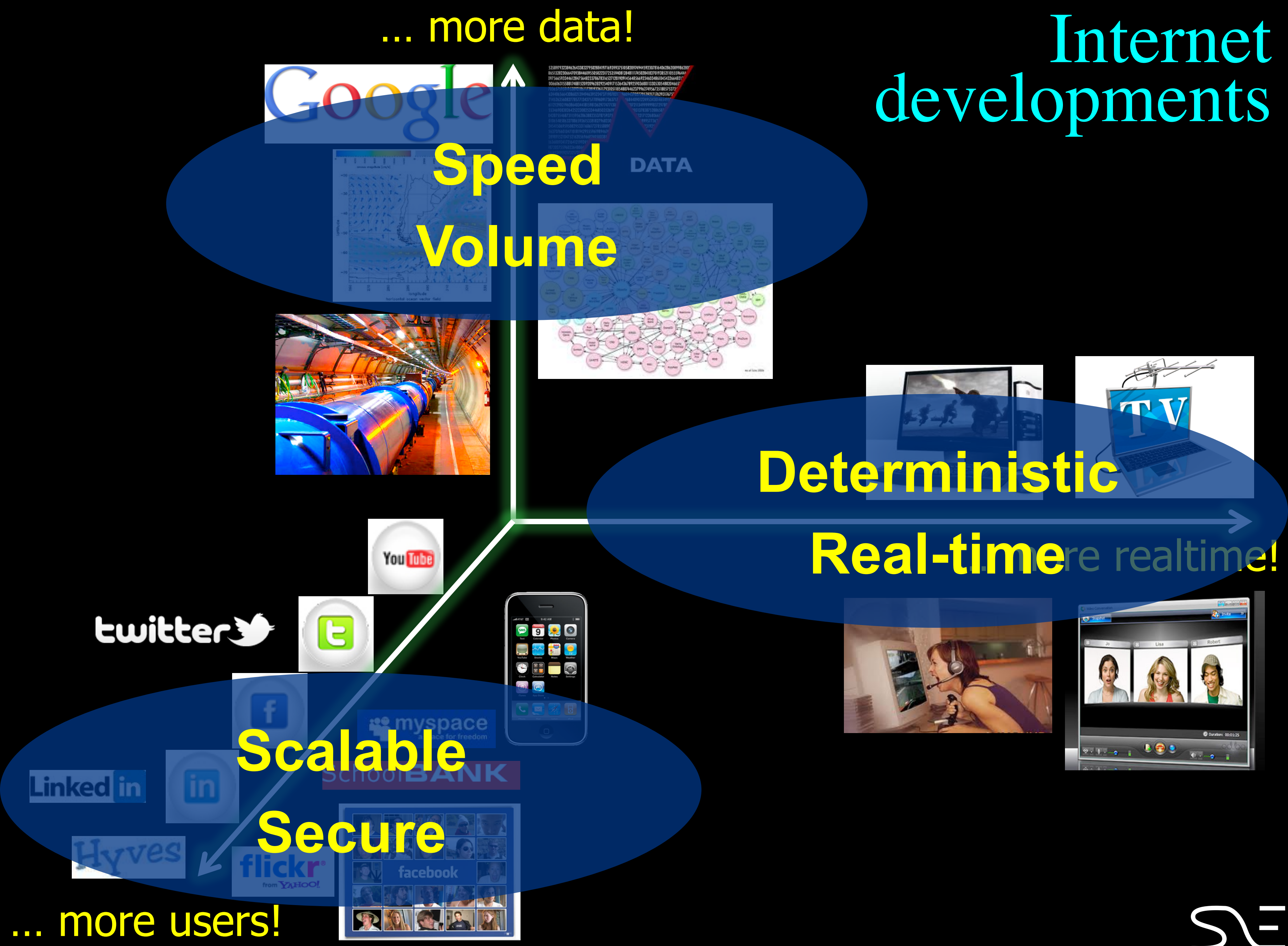
Smart Cyber Infrastructure for Big Data Processing

Cees de Laat




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

Internet developments



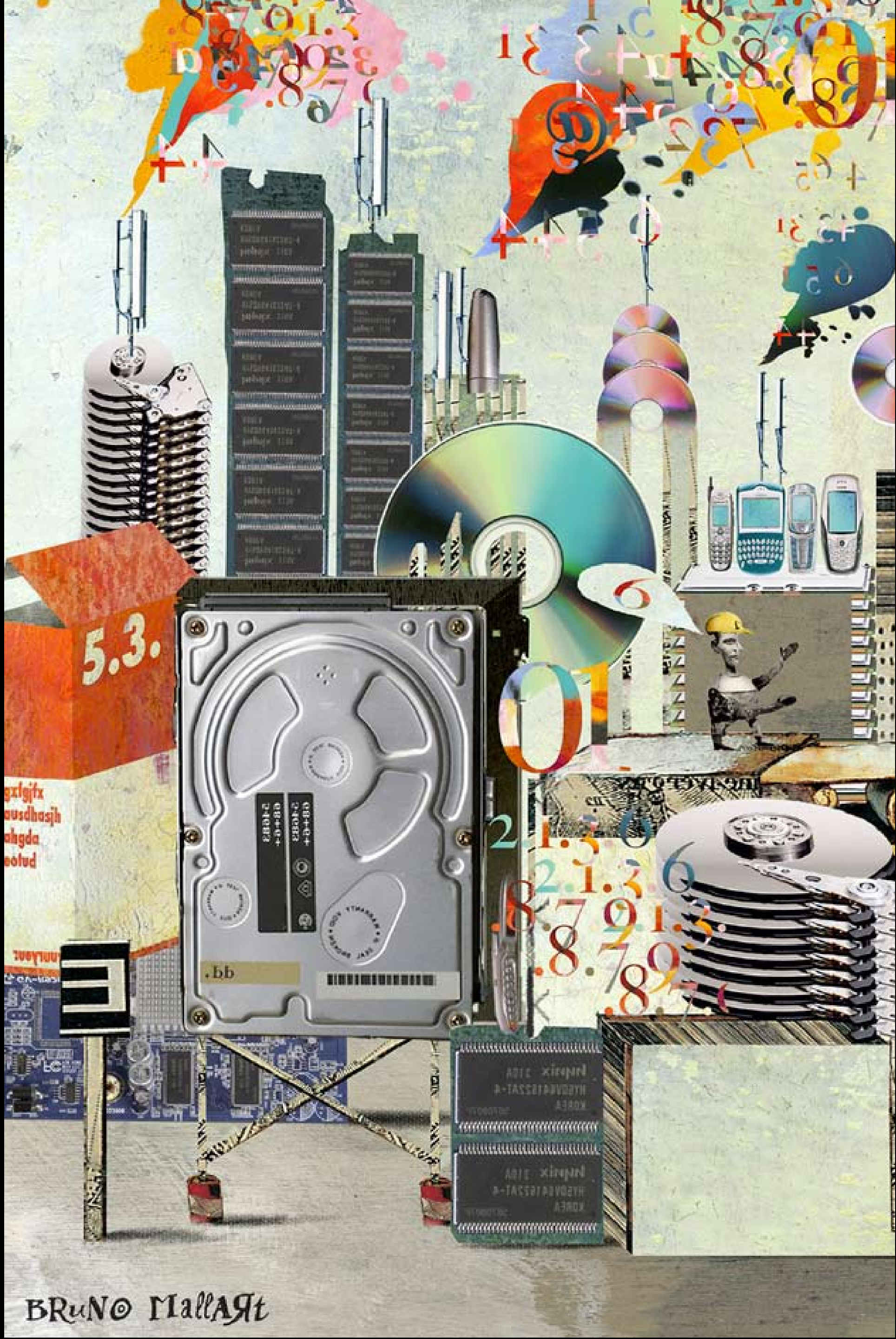
From King's Dutch Academy of Sciences

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!"

https://www.knaw.nl/nl/actueel/publicaties/the-dutch-research-agenda/@@download/pdf_file/20111029.pdf



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?

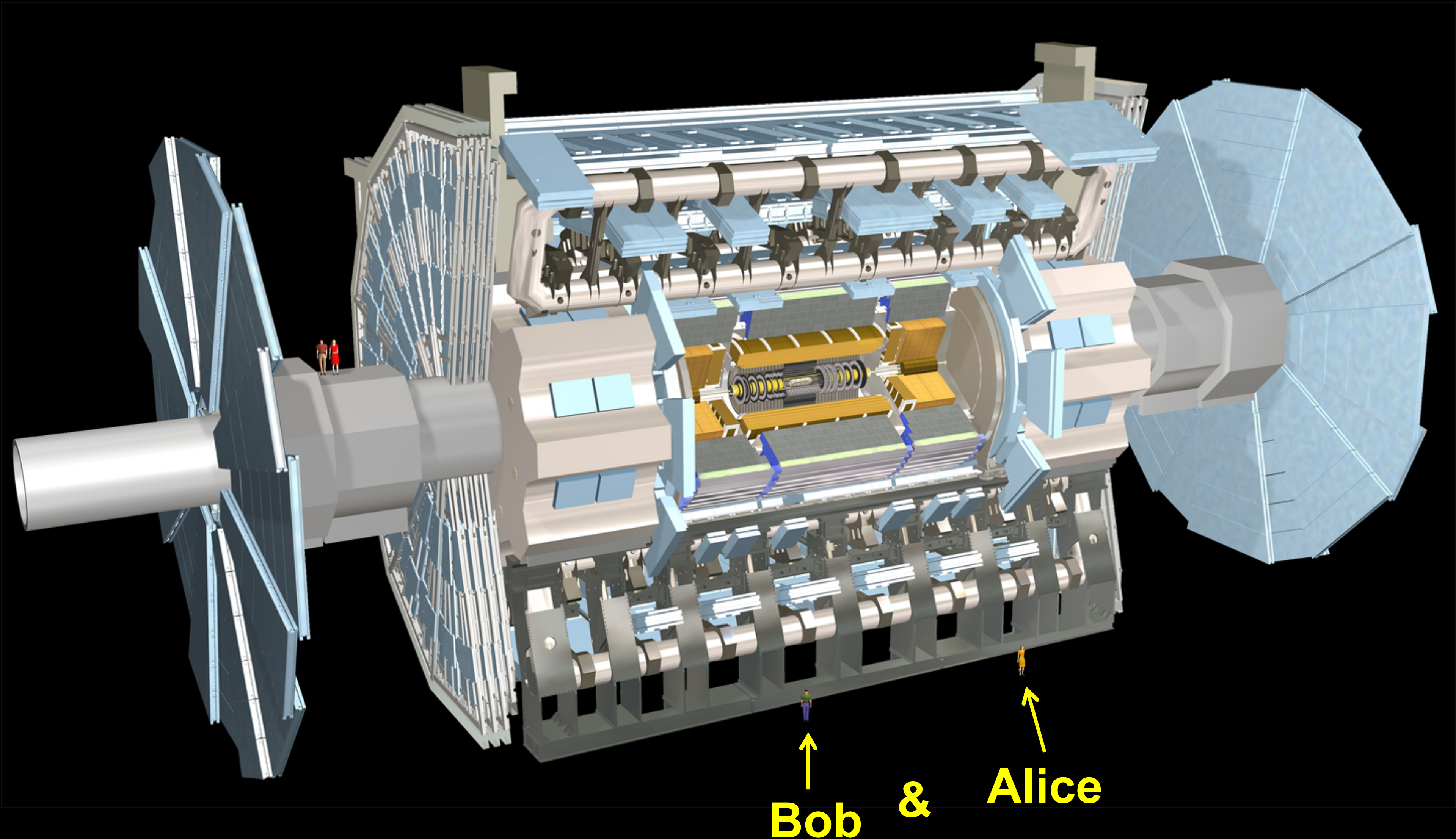
- *Capacity*
  - *Bandwidth on demand, QoS, architectures, photonics, performance*
- *Capability*
  - *Programmability, virtualization, complexity, semantics, workflows*
- *Security*
  - *Anonymity, integrity of data in distributed data processing*
- *Sustainability*
  - *Greening infrastructure, awareness*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

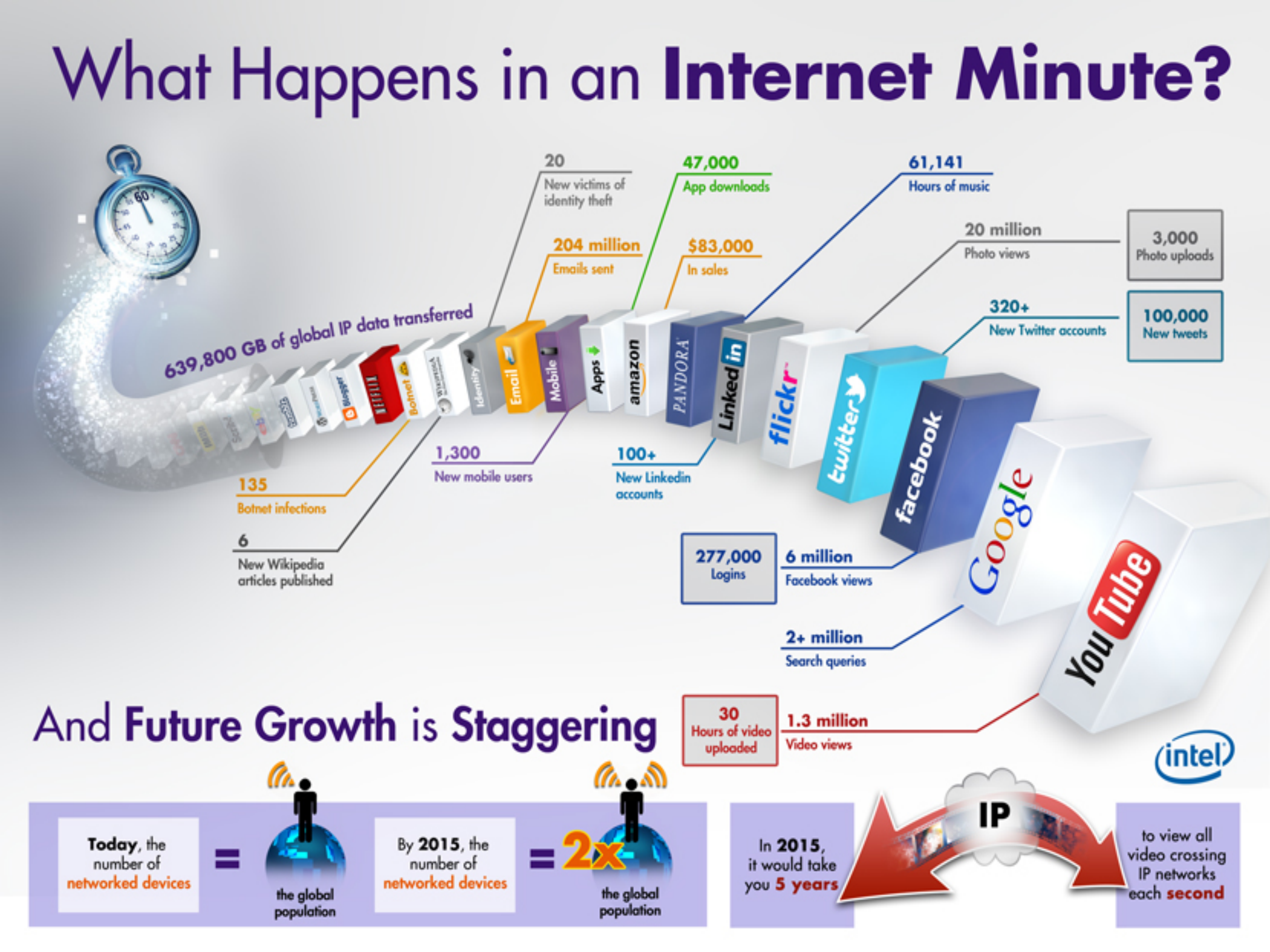
Mission

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

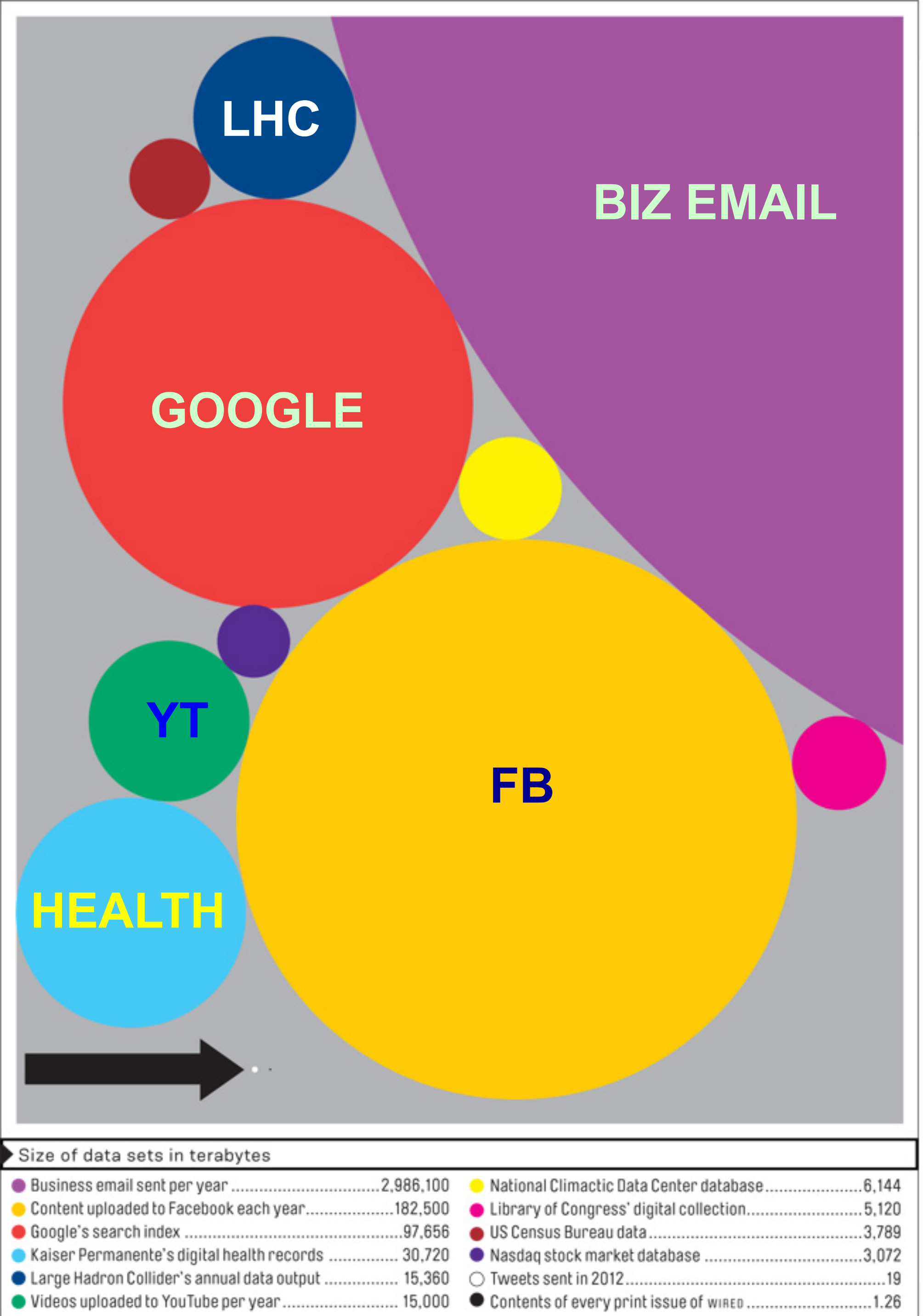
- ***Capacity***
  - ***Bandwidth on demand, QoS, architectures, photonics, performance***
- *Capability*
  - *Programmability, virtualization, complexity, semantics, workflows*
- *Security*
  - *Anonymity, integrity of data in distributed data processing*
- *Sustainability*
  - *Greening infrastructure,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

ATLAS detector @ CERN Geneve



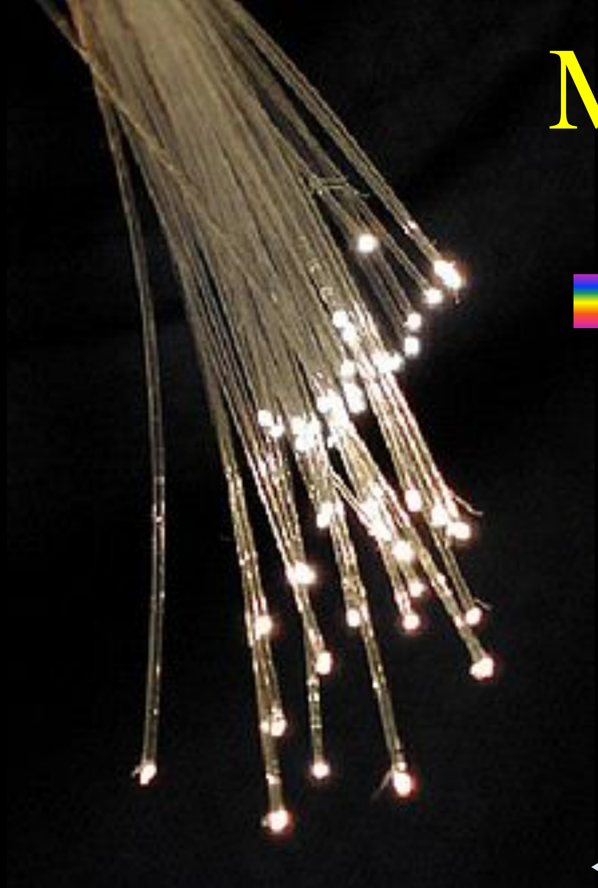


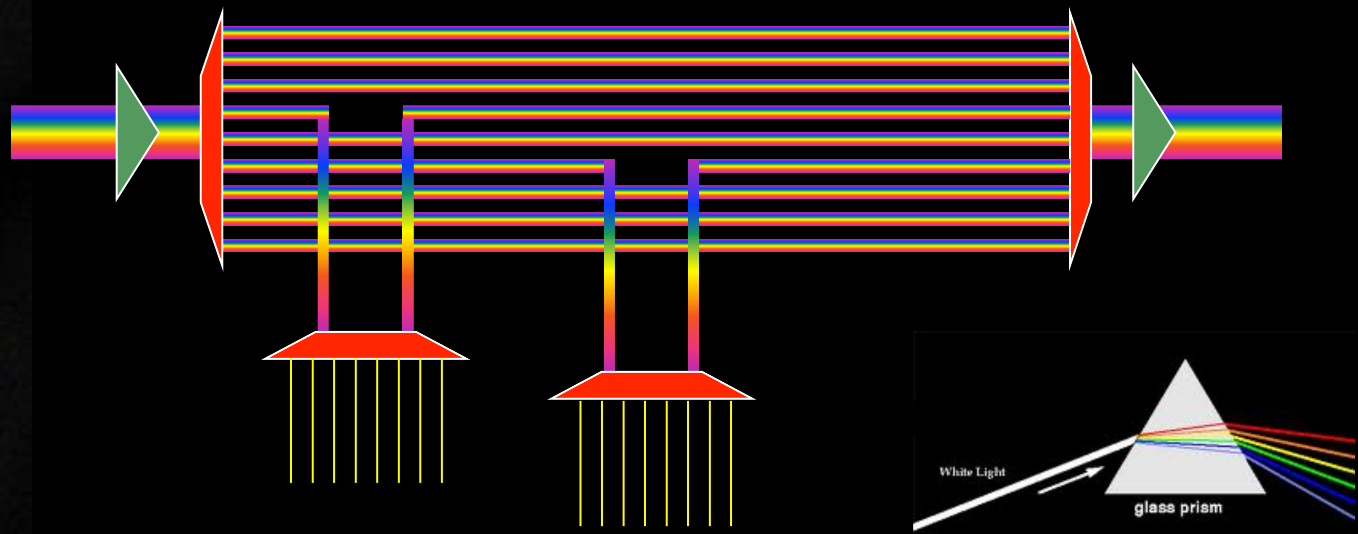
What Happens in an Internet Minute?
639,800 GB of global IP data transferred
20
New victims of identity theft
204 million
Emails sent
47,000
App downloads
$83,000
In sales
61,141
Hours of music
20 million
Photo views
3,000
Photo uploads
320+
New Twitter accounts
100,000
New tweets
135
Botnet infections
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
Linked in
flickr
twitter
facebook
Google
You Tube
intel
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

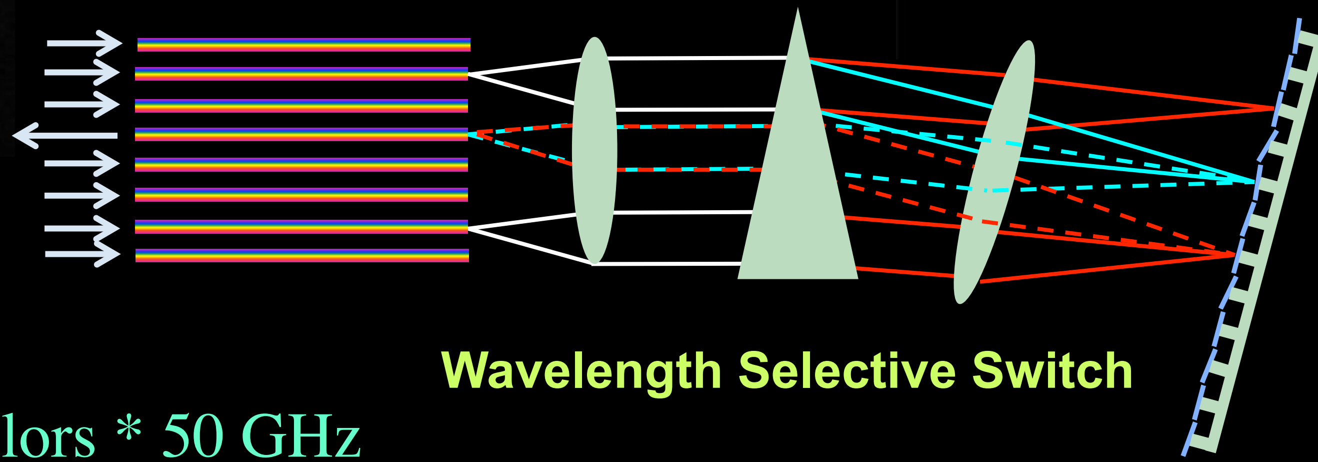


There is always a bigger fish

Multiple colors / Fiber







Wavelength Selective Switch

Per fiber: ~ 80-100 colors * 50 GHz

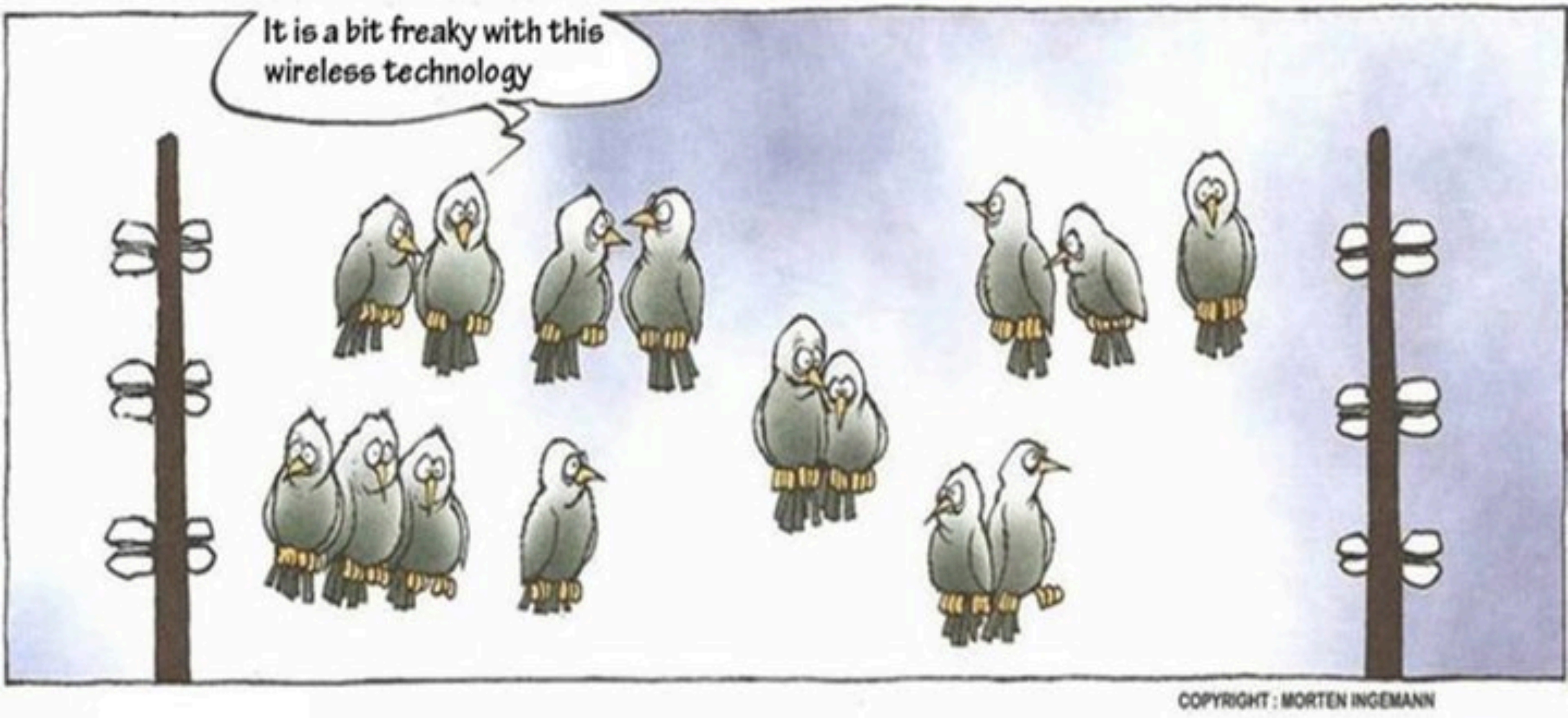
Per color: 10 – 40 – 100 Gbit/s

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.



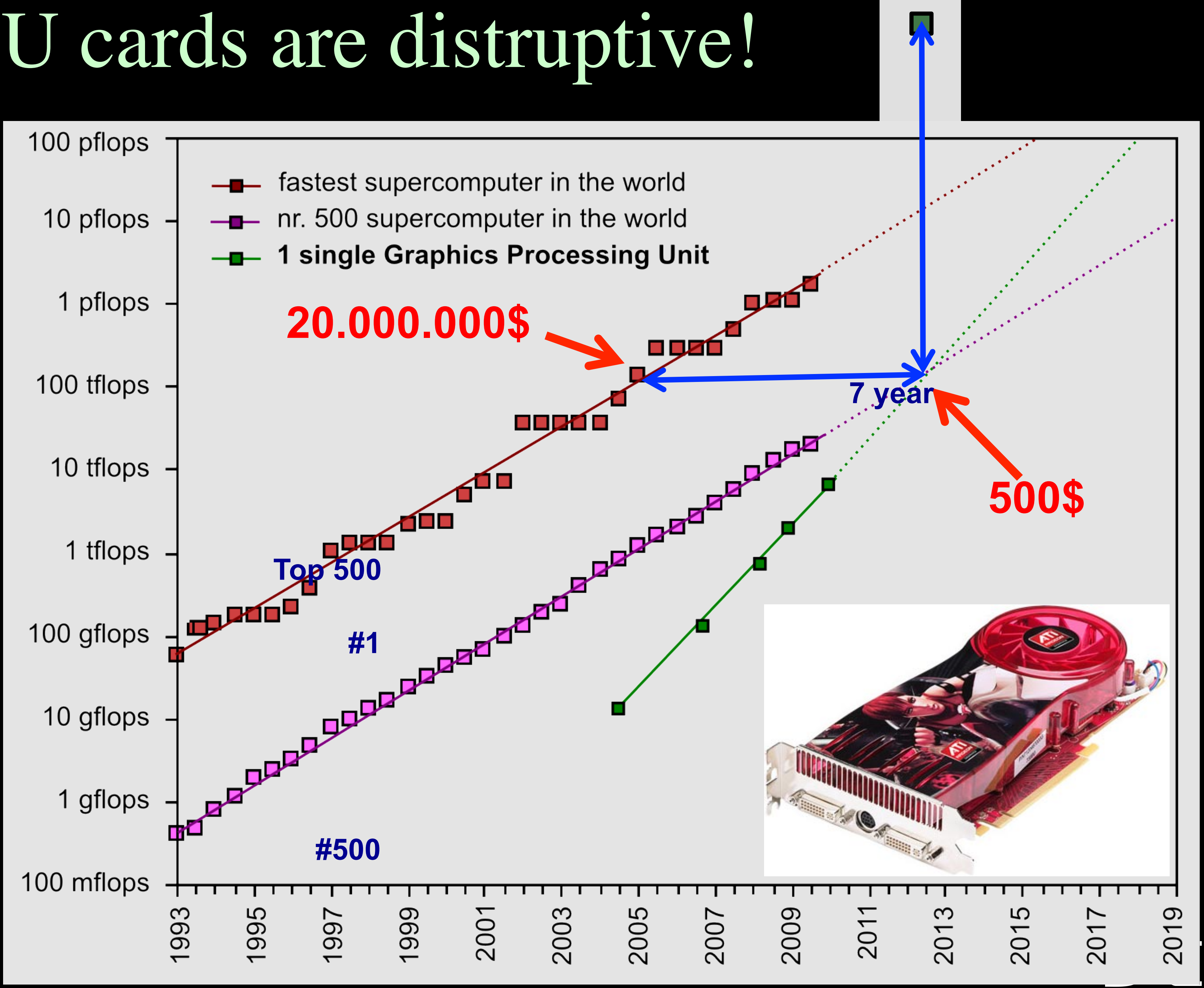

IBM
UPLINK ACTY
TEMP
AUTO
GIMBAL LOCK
PROG
RESTART
NO ATT
TRACKER
STBY
KEY REL
OPR ERR
COMP ACTY
VERB
NOUN
CLR
ENTR
RSET






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

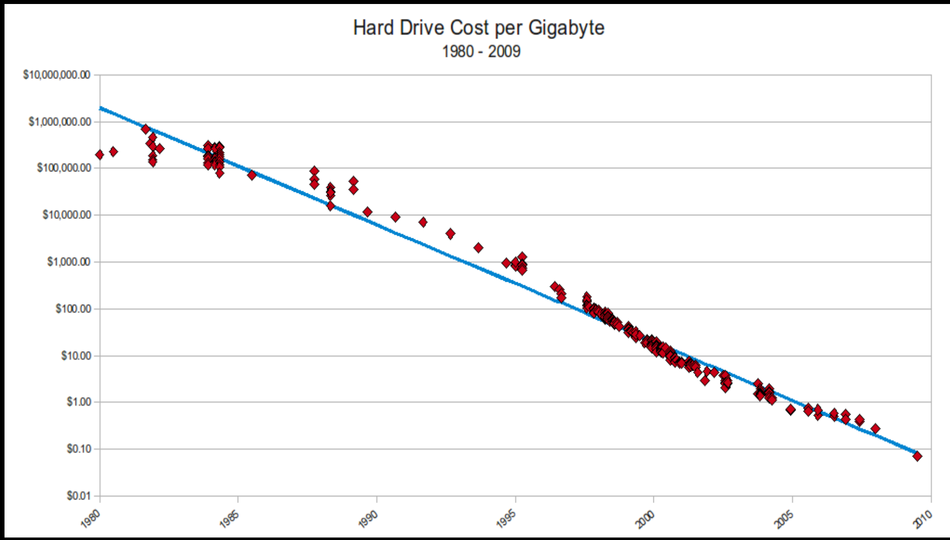

GPU cards are distruptive!



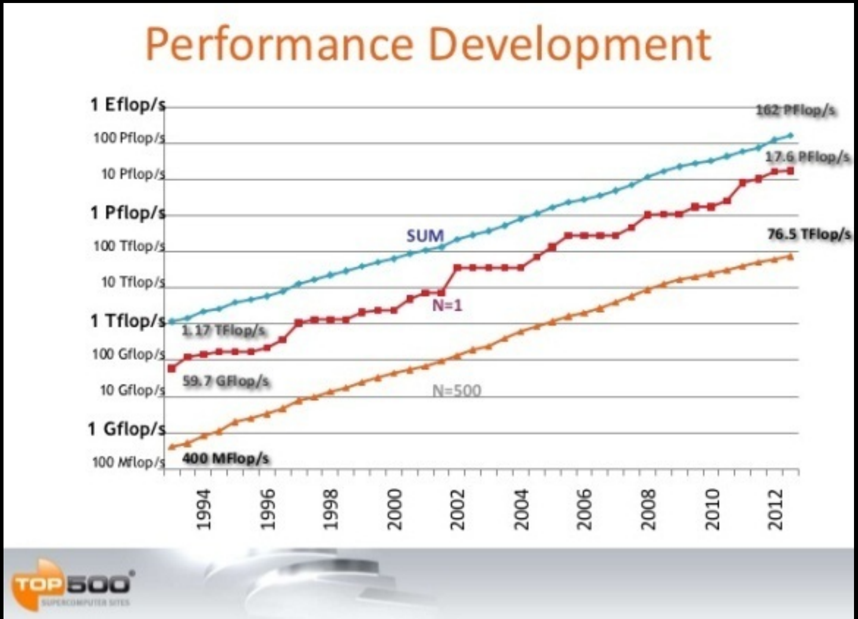
Reliable and Safe!

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.

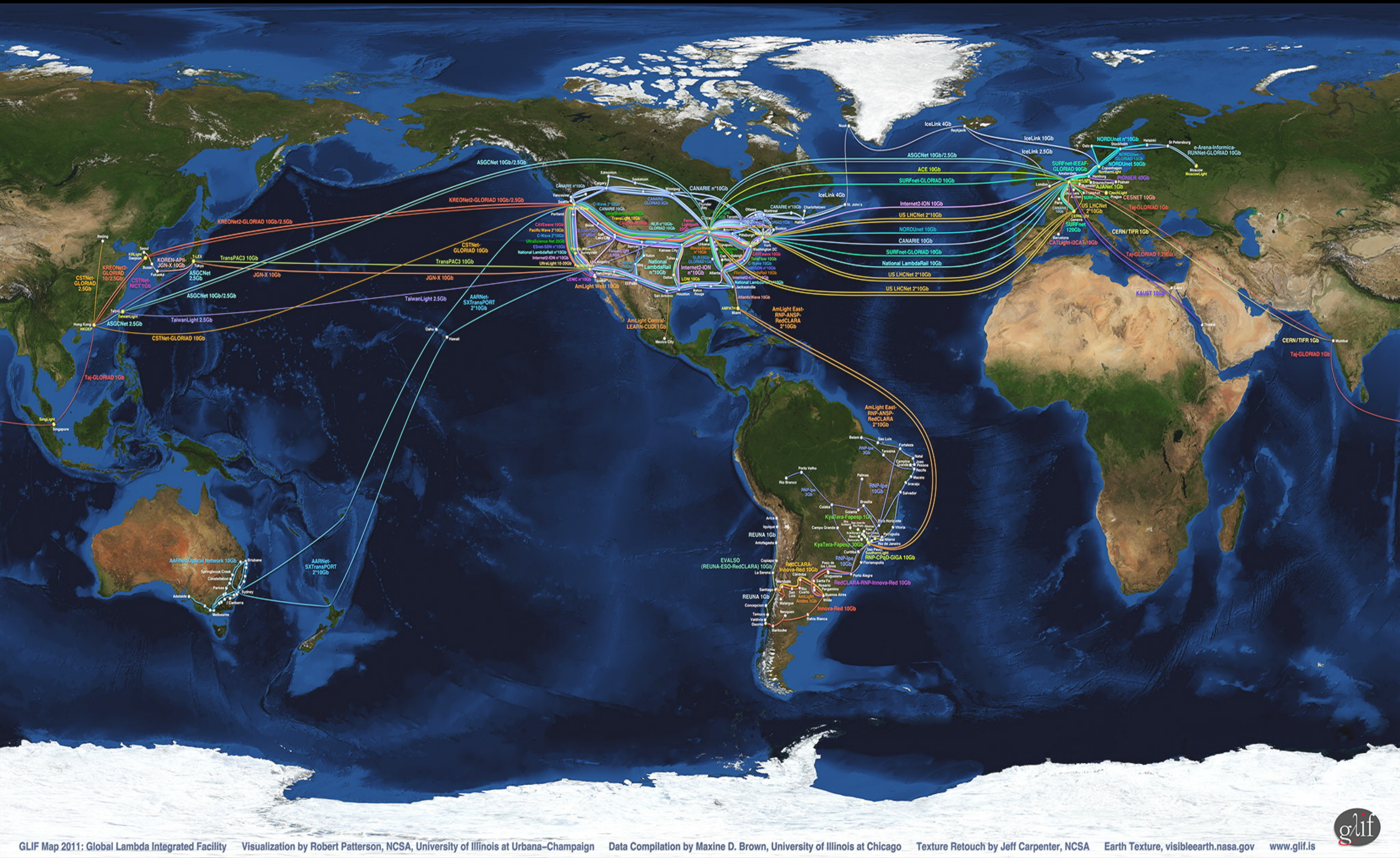


We will soon reach the limits of what is currently feasible and controllable.

https://www.knaw.nl/shared/resources/actueel/publicaties/pdf/20111029.pdf

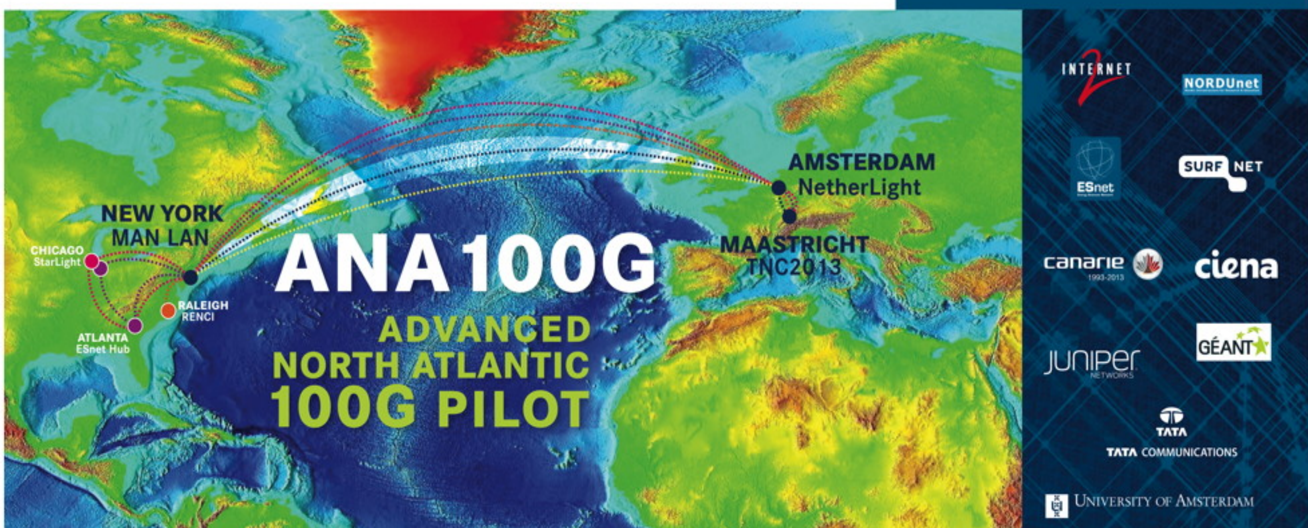
The GLIF – LightPaths around the World

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.



ExoGeni @ OpenLab - UvA

Installed and up June 3th 2013



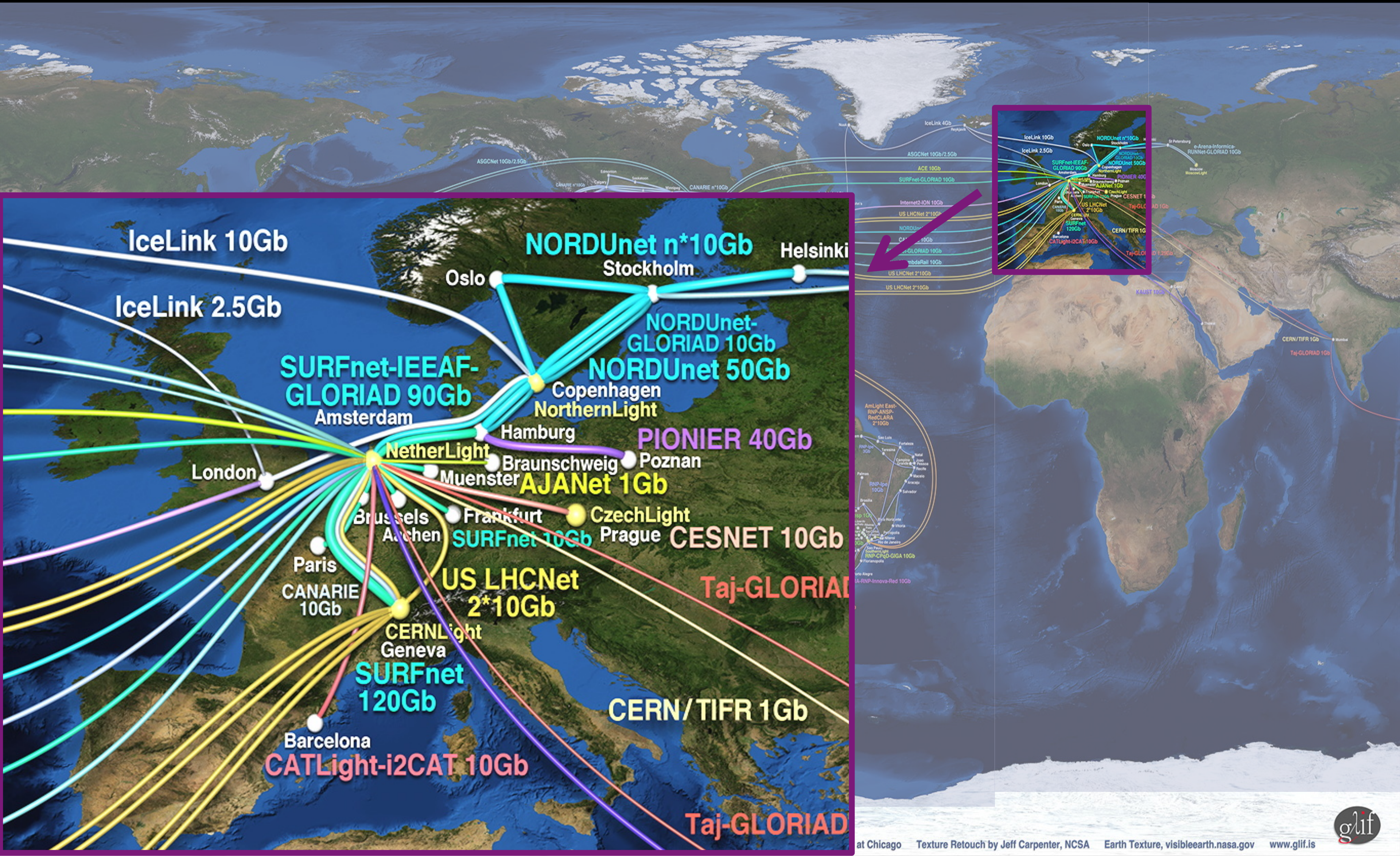
TNC2013 DEMOS JUNE, 2013

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 To US-GENI

Amsterdam is a major hub in The GLIF

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.



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.

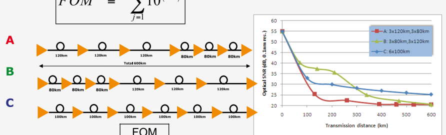
In this demonstration we will investigate the performance of a 40Gb/s PM-QPSK alien wavelength installed on a 10Gb/s DWDM infrastructure.

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

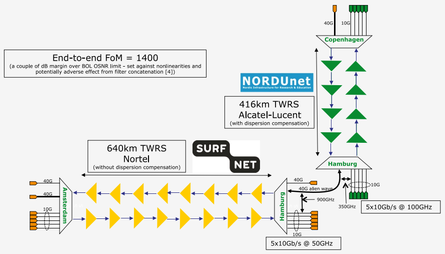


	FOM
A	5504
B	5504
C	1897

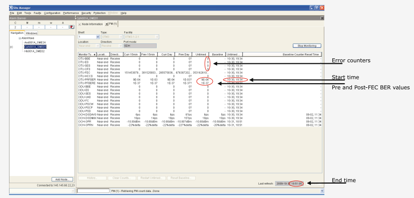
Easy-to-use formula that accurately quantifies transmission system performance

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

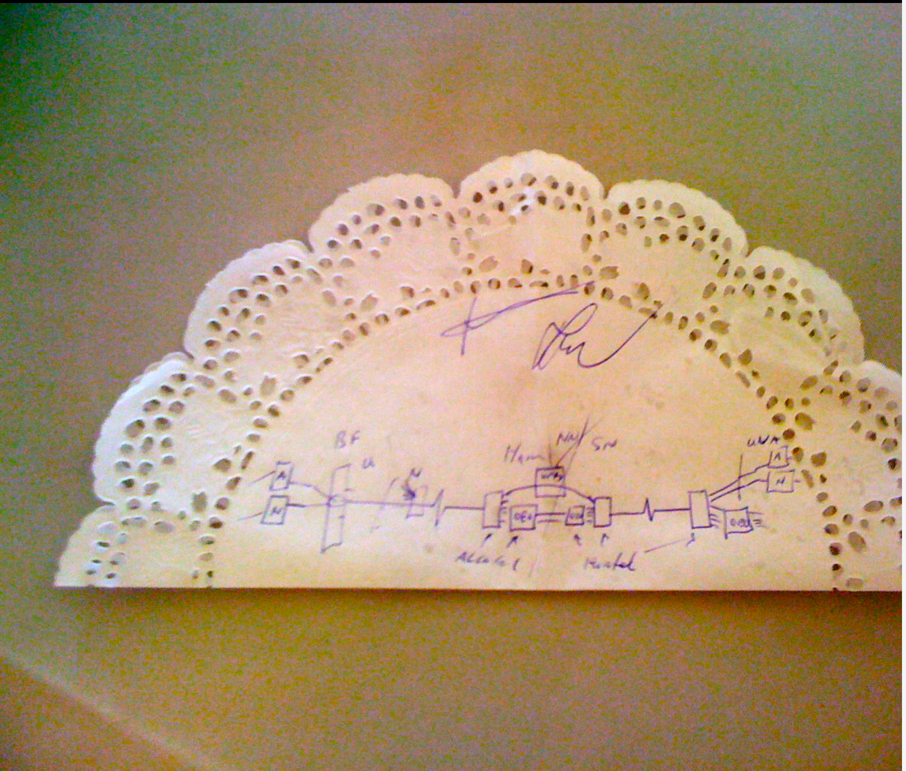
- We have investigated experimentally the all-optical transmission of a 40Gb/s PM-QPSK alien wavelength via a concatenated native and third party DWDM system that both were carrying live 10Gb/s wavelengths.
- The end-to-end transmission system consisted of 1056 km of TWRS (TrueWave Reduced Slope) transmission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

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.

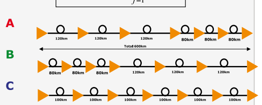
In this demonstration we will investigate the performance of a 40Gb/s PM-QPSK alien wavelength installed on a 10Gb/s DWDM infrastructure.

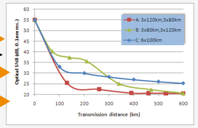
New method to present fiber link quality, FoM (Figure of Merit)

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



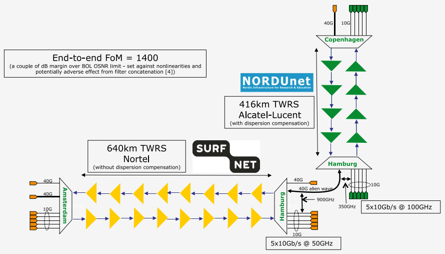


	FOM
A	5504
B	5504
C	1897

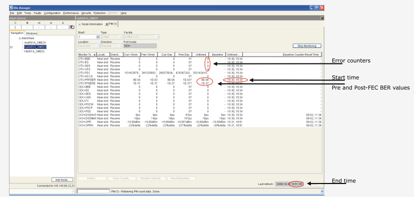
Easy-to-use formula that accurately quantifies transmission system performance

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.

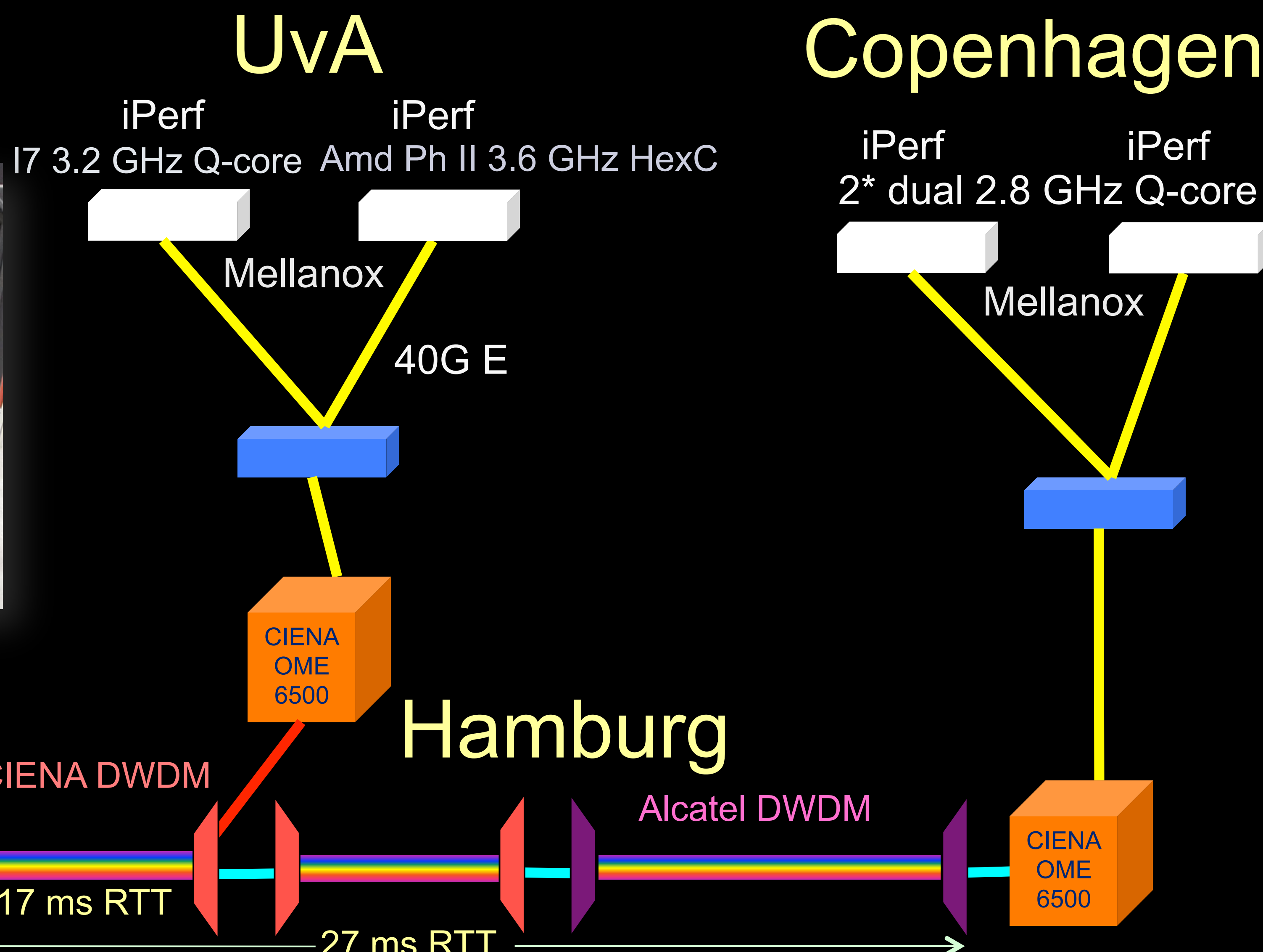
NORTEL · NORDUnet Nordic Infrastructure for Research & Education · telindus belgacom ICT together with ISIT THE STORAGE COMPANY · SURF NET

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



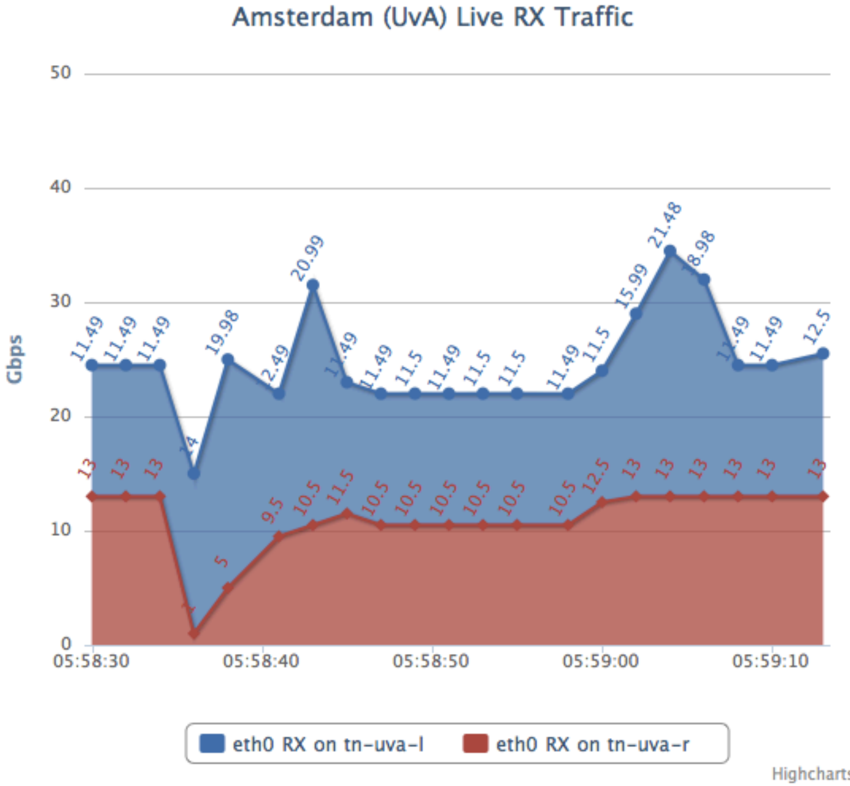
Amsterdam – Geneva (CERN) – Copenhagen – 4400 km (2700 km alien light)

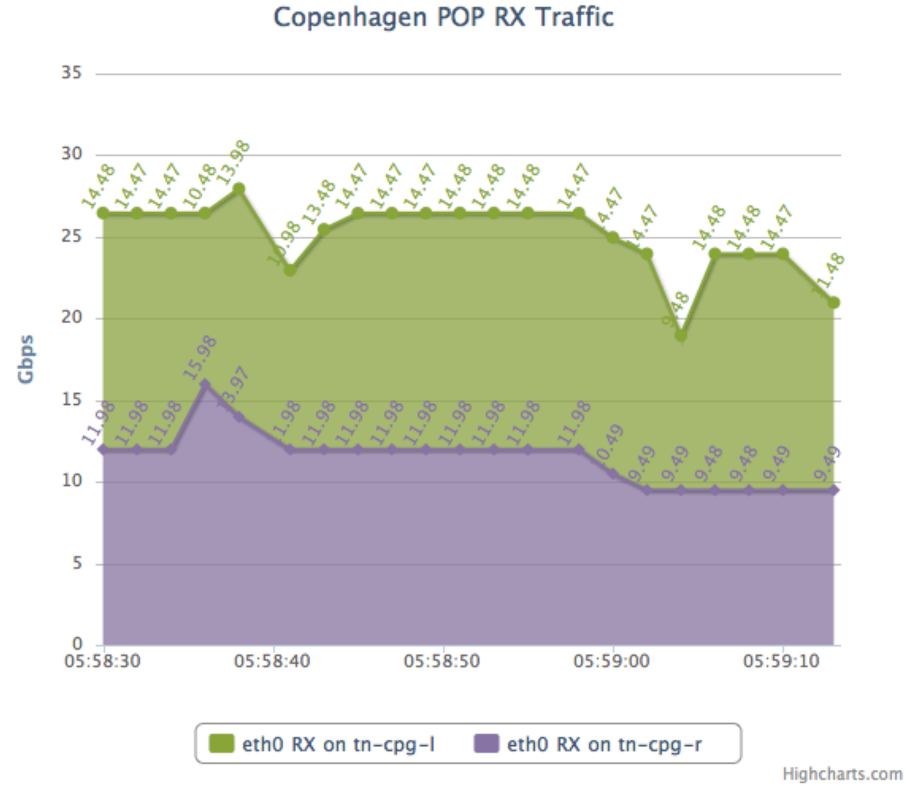
Visit CIENA Booth

surf to http://tnc.delaat.net/tnc11

ClearStream

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





Incoming Amsterdam 25.5 Gbps

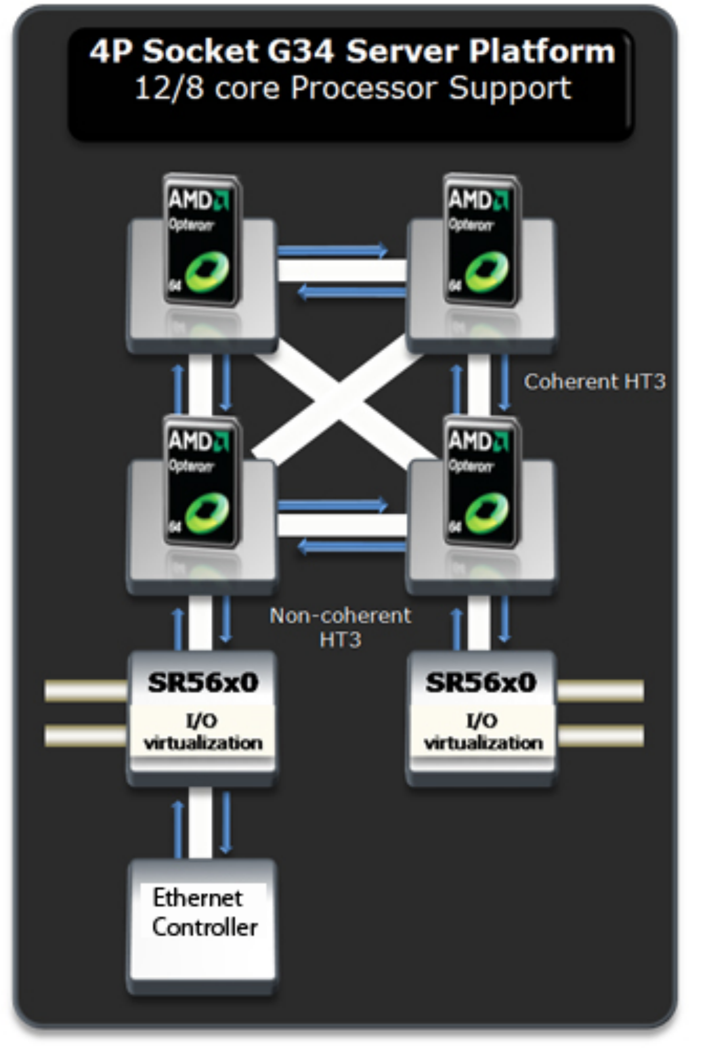
Incoming Copenhagen 20.97 Gbps

Total Throughput 46.47 Gbps RTT 44.032 ms

Results (rtt = 17 ms)

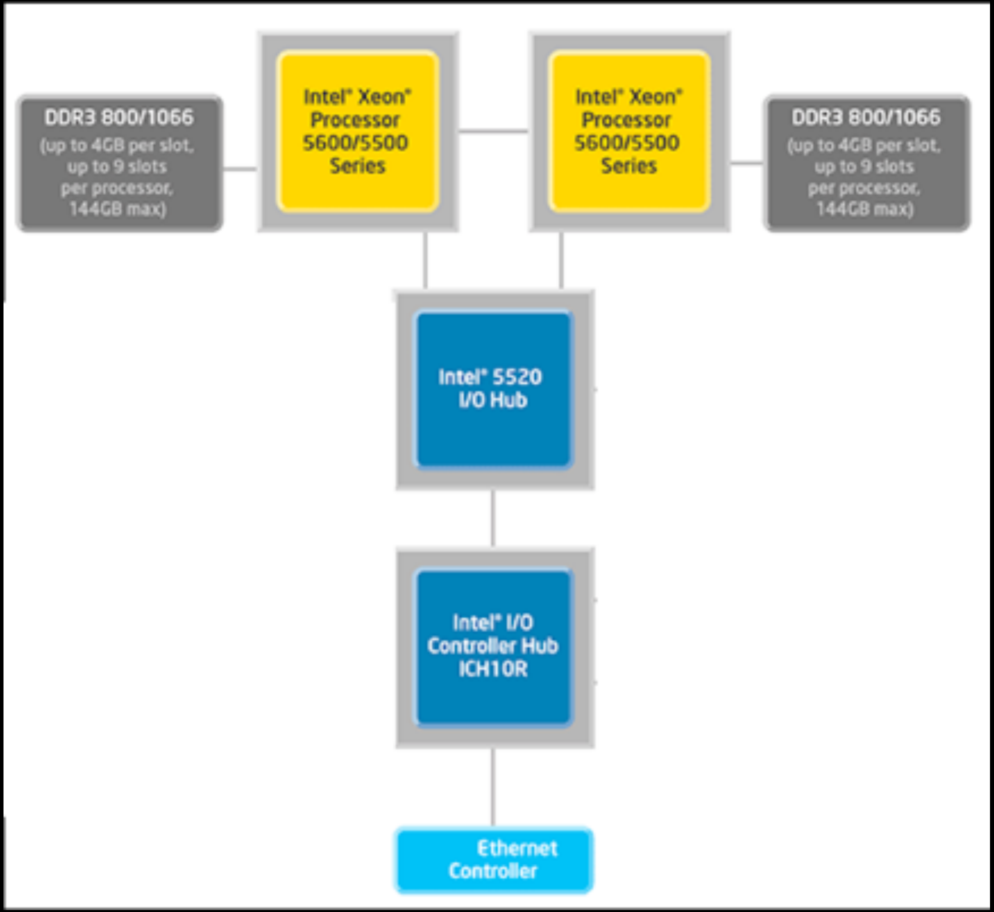
- Single flow iPerf 1 core -> 21 Gbps
- Single flow iPerf 1 core <> -> 15+15 Gbps
- Multi flow iPerf 2 cores -> 25 Gbps
- Multi flow iPerf 2 cores <> -> 23+23 Gbps
- DiViNe <> -> 11 Gbps
- Multi flow iPerf + DiVine -> 35 Gbps
- Multi flow iPerf + DiVine <> -> 35 + 35 Gbps

Server Architecture



DELL R815

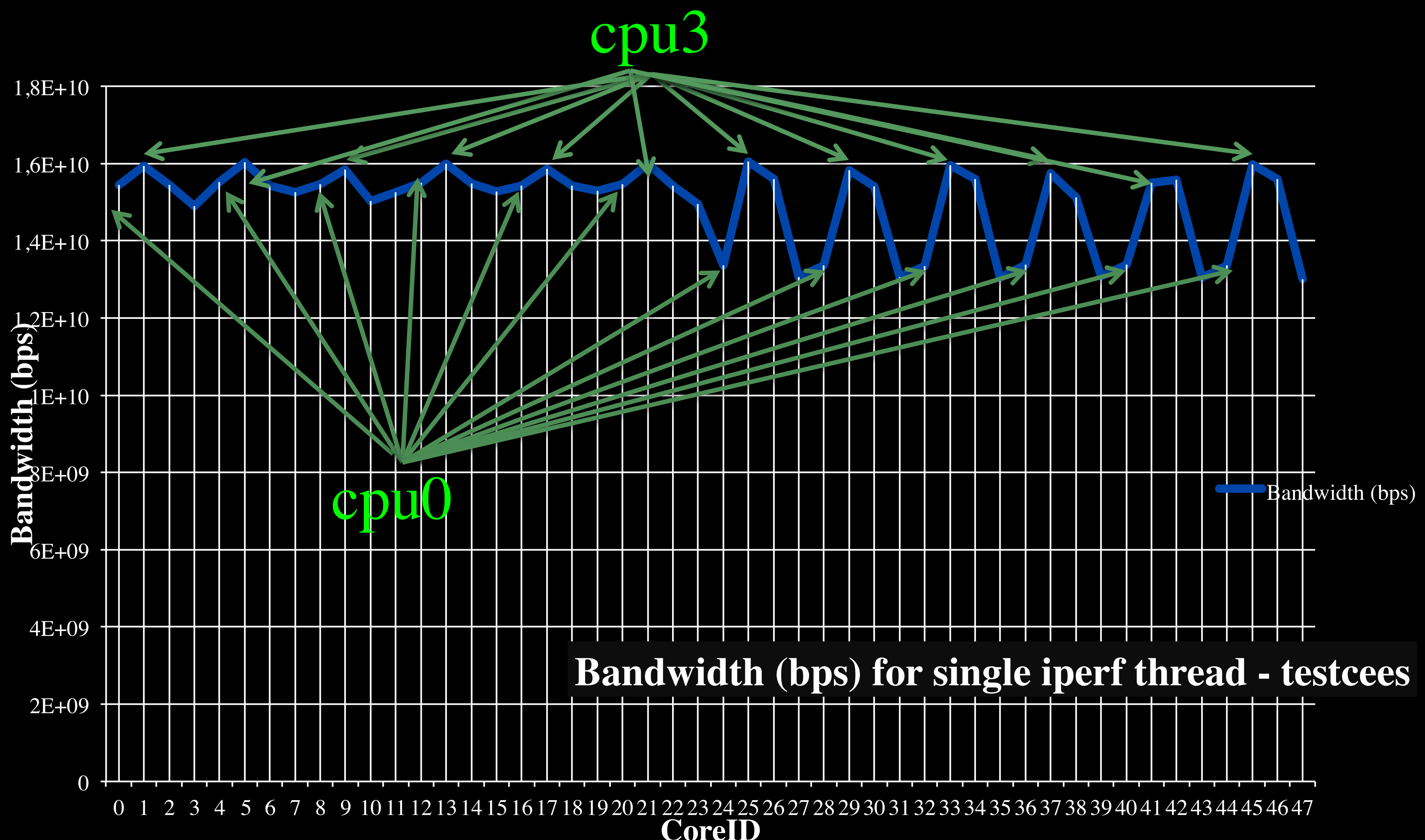
4 x AMD Opteron 6100



Supermicro X8DTT-HIBQF

2 x Intel Xeon

CPU Topology benchmark

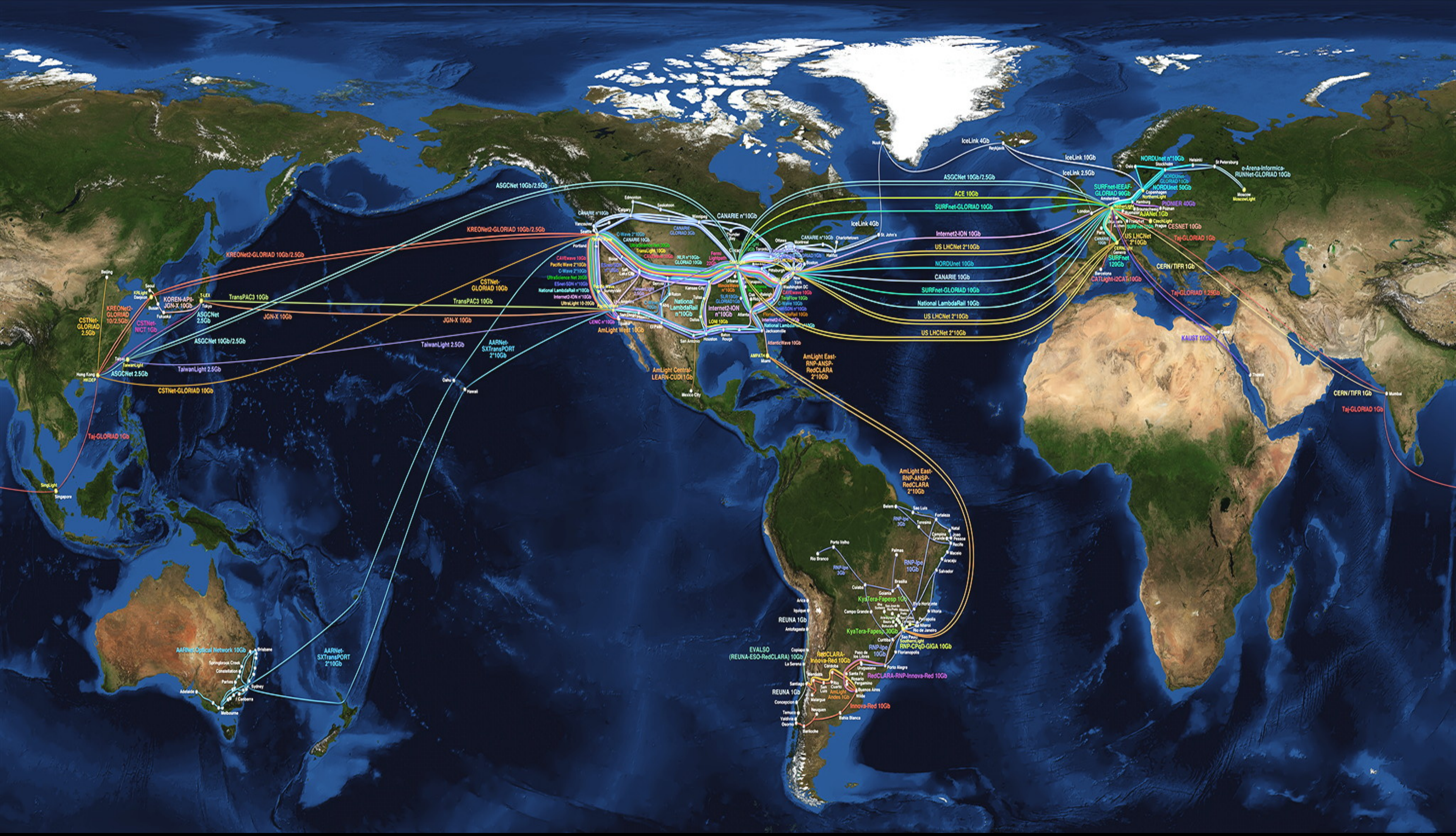


We used numactl to bind iperf to cores

Mission

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

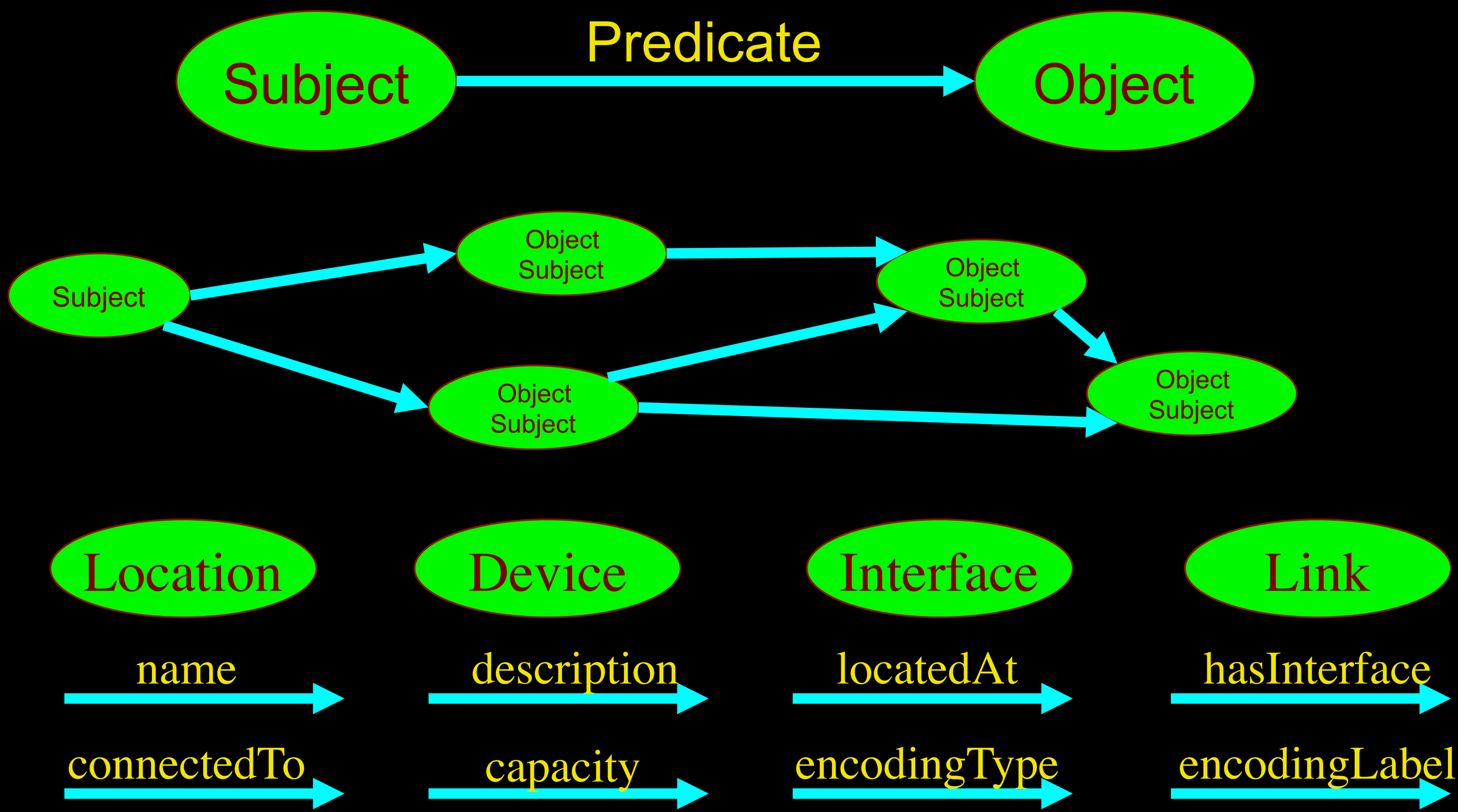
- *Capacity*
  - *Bandwidth on demand, QoS, architectures, photonics, performance*
- ***Capability***
  - ***Programmability, virtualization, complexity, semantics, workflows***
- *Security*
  - *Anonymity, integrity of data in distributed data processing*
- *Sustainability*
  - *Greening infrastructure,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



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):

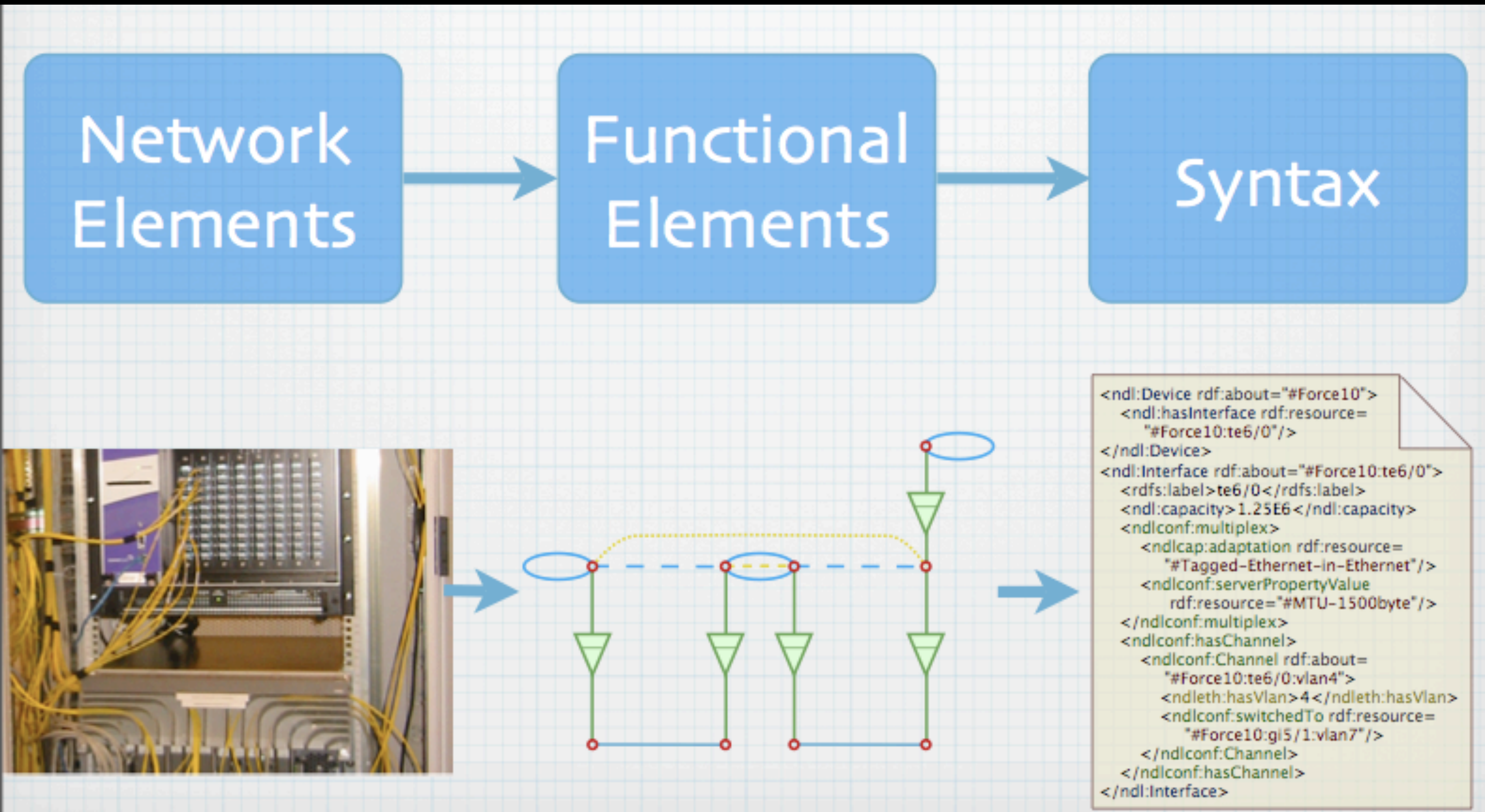


Network Description Language

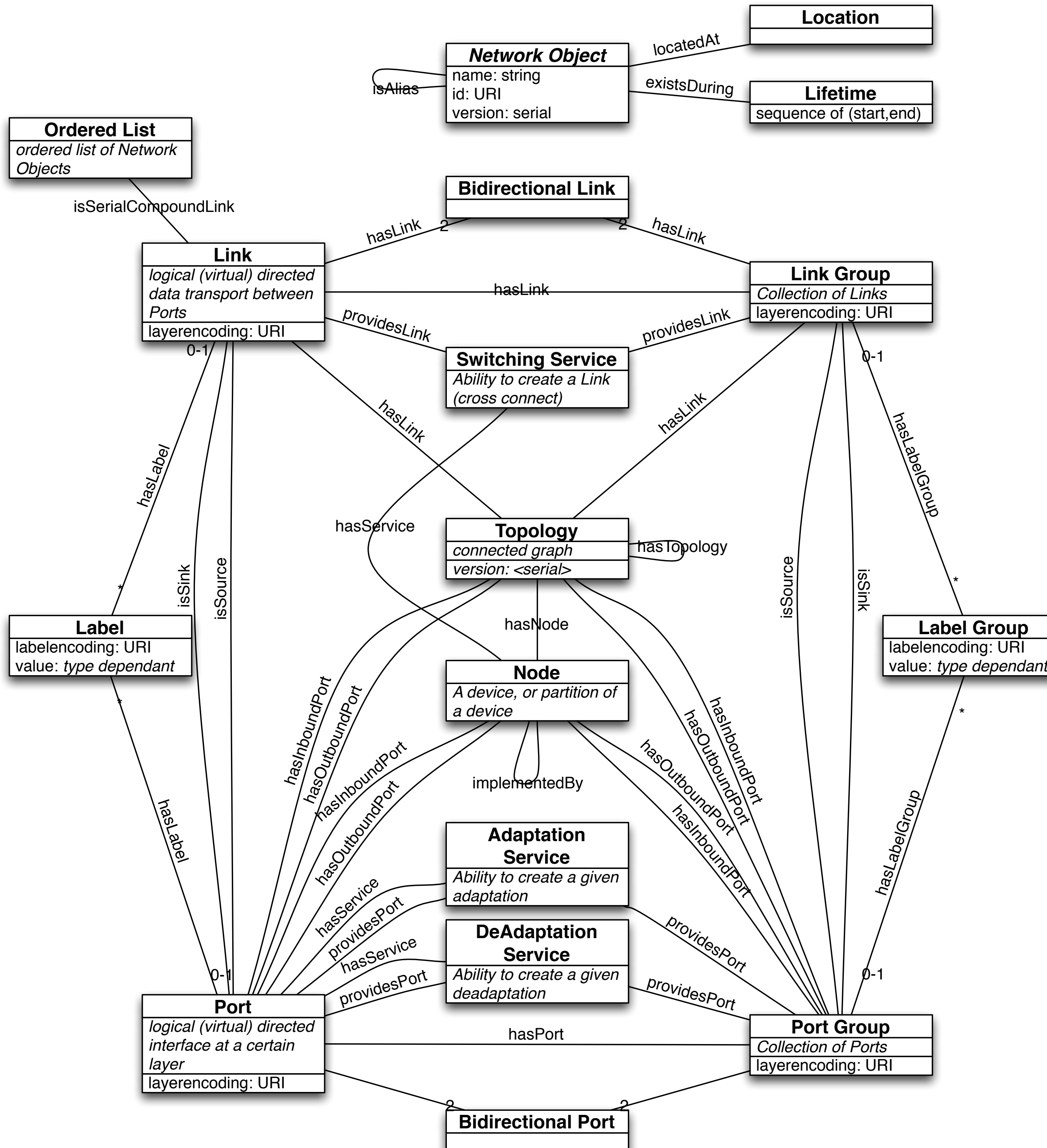
Choice of RDF instead of XML syntax

Grounded modeling based on G0805 description:

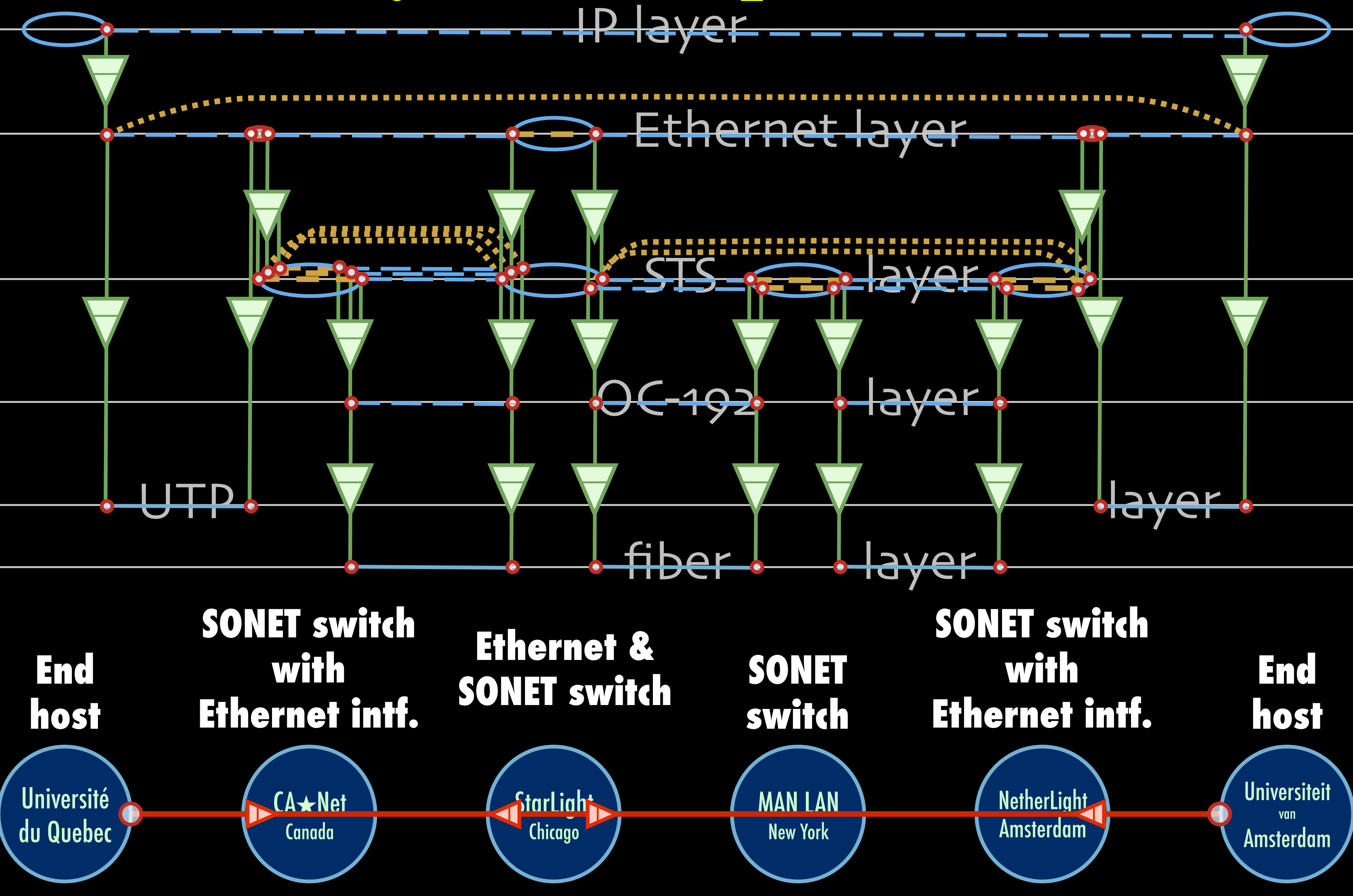
Article: F. Dijkstra, B. Andree, K. Koymans, J. van der Ham, P. Grosso, C. de Laat, *"A Multi-Layer Network Model Based on ITU-T G.805"*



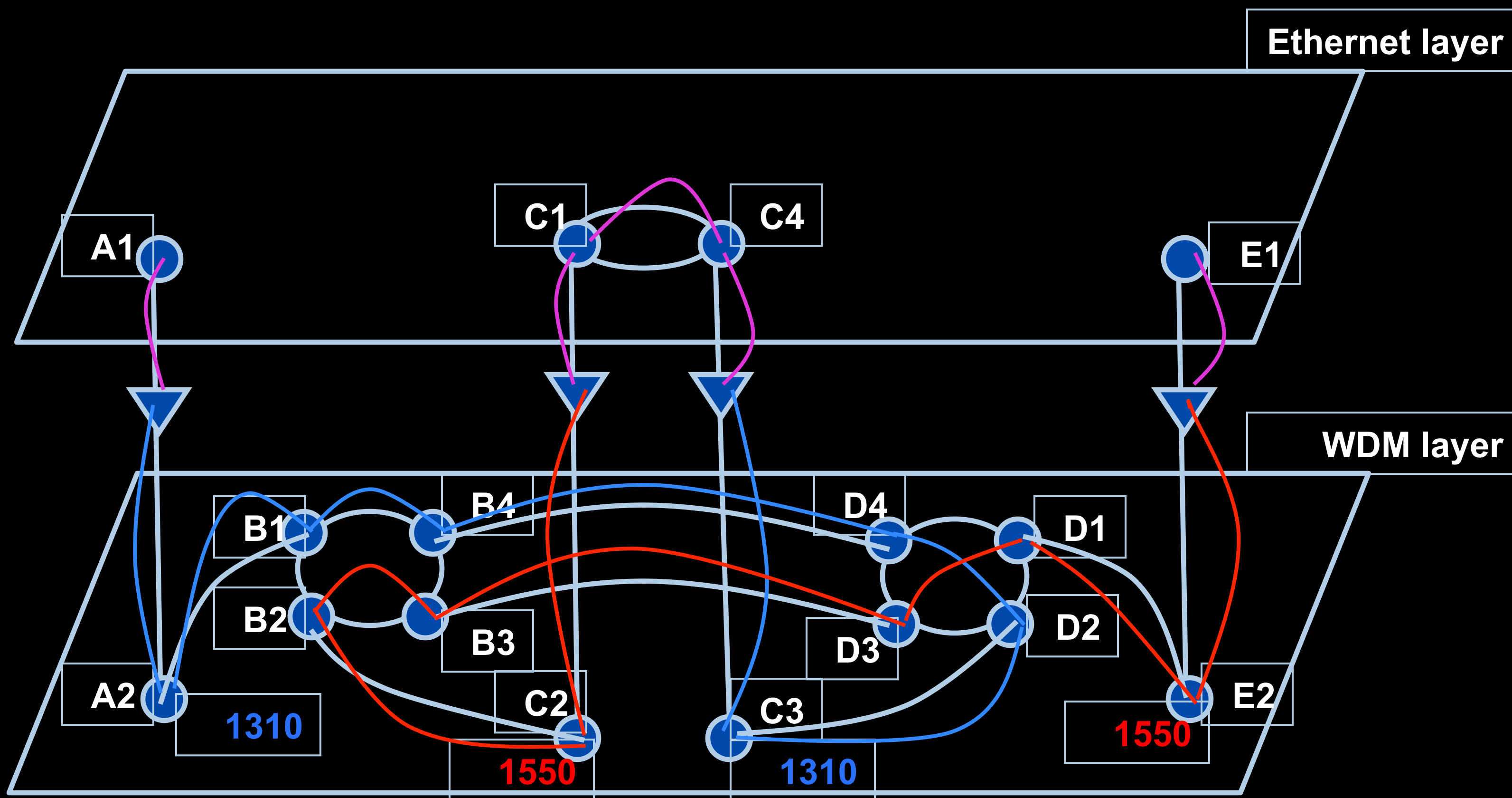
NML OFG spec


Location
Network Object
name: string
id: URI
version: serial
locatedAt
existsDuring
isAlias
Lifetime
sequence of (start,end)
Ordered List
ordered list of Network Objects
isSerialCompoundLink
Bidirectional Link
hasLink
2
Link
logical (virtual) directed data transport between Ports
layerencoding: URI
Link Group
Collection of Links
layerencoding: URI
providesLink
Switching Service
Ability to create a Link (cross connect)
0-1
hasLabel
isSink
isSource
hasLabelGroup
hasService
Topology
connected graph
version: <serial>
hasTopology
hasNode
Label
labelencoding: URI
value: type dependant
Label Group
labelencoding: URI
value: type dependant
Node
A device, or partition of a device
hasInboundPort
hasOutboundPort
implementedBy
Adaptation Service
Ability to create a given adaptation
DeAdaptation Service
Ability to create a given deadaptation
providesPort
Port
logical (virtual) directed interface at a certain layer
layerencoding: URI
hasPort
Port Group
Collection of Ports
layerencoding: URI
Bidirectional Port
*

Multi-layer descriptions in NDL



Multi-layer Network PathFinding

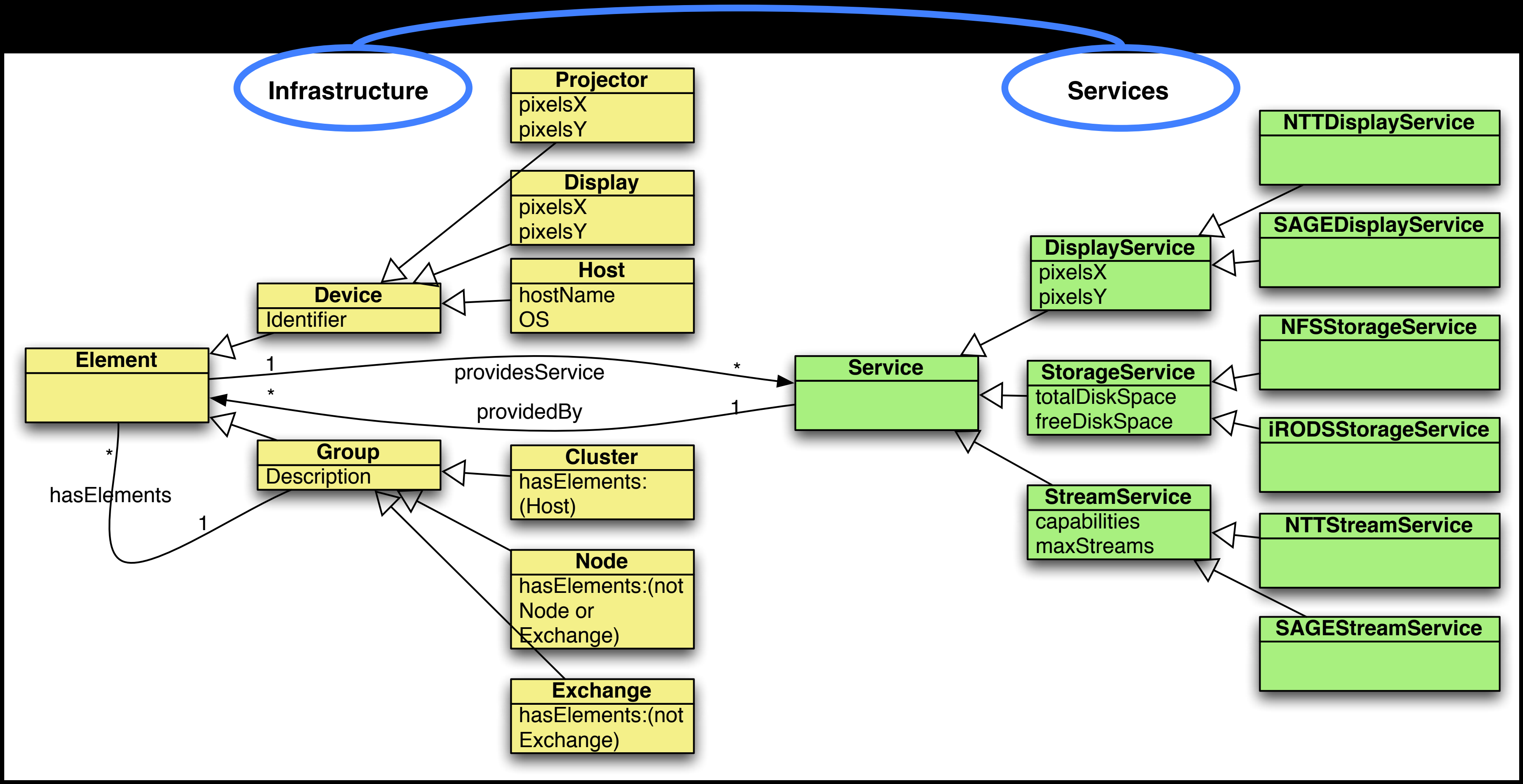


Path between interfaces A1 and E1:
A1-A2-B1-B4-D4-D2-C3-C4-C1-C2-B2-B3-D3-D1-E2-E1

Scaling: Combinatorial problem

Information Modeling

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

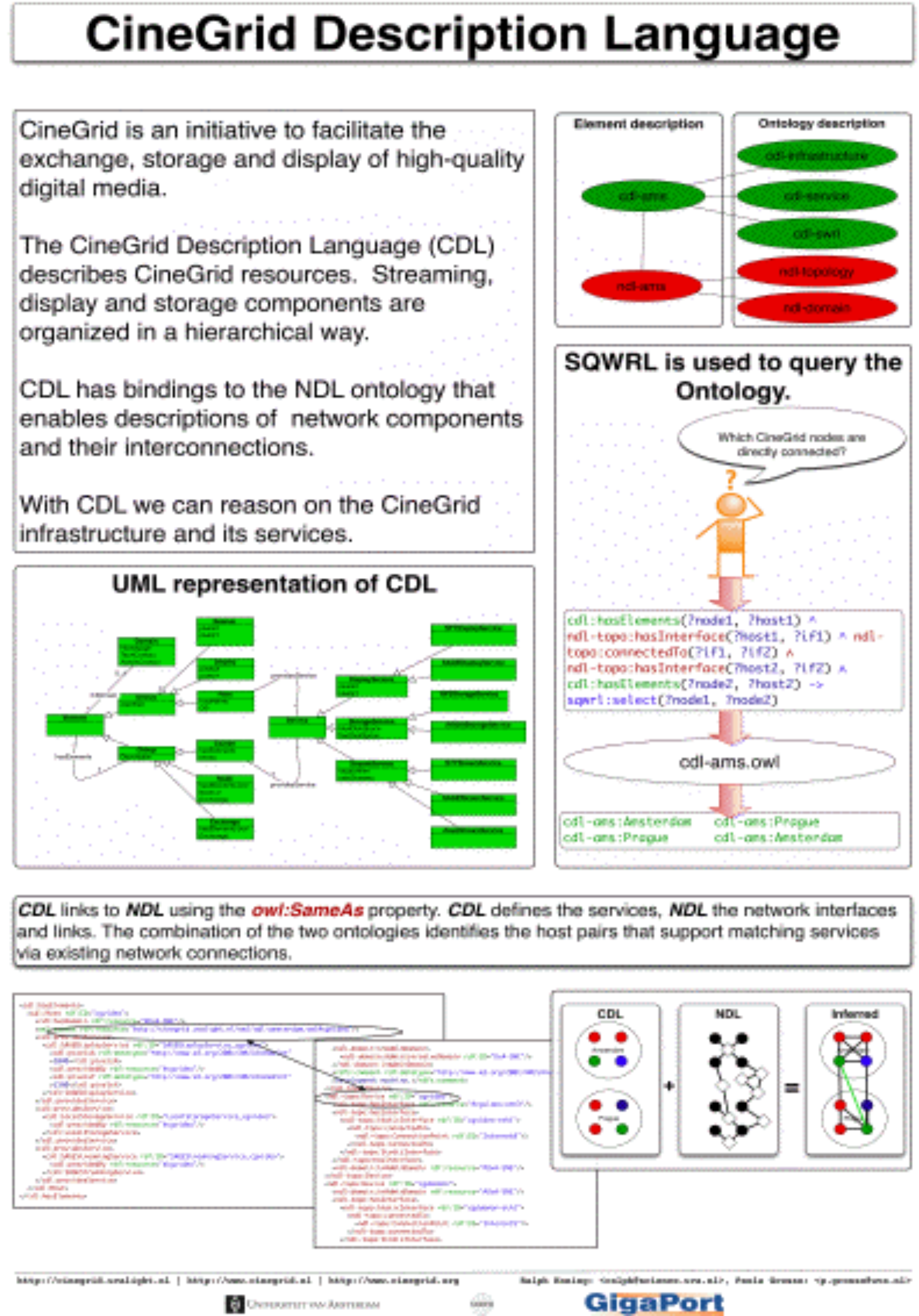


J. van der Ham, F. Dijkstra, P. Grosso, R. van der Pol, A. Toonk, C. de Laat
A distributed topology information system for optical networks based on the semantic web, Elsevier Journal on Optical Switching and Networking, Volume 5, Issues 2-3, June 2008, Pages 85-93

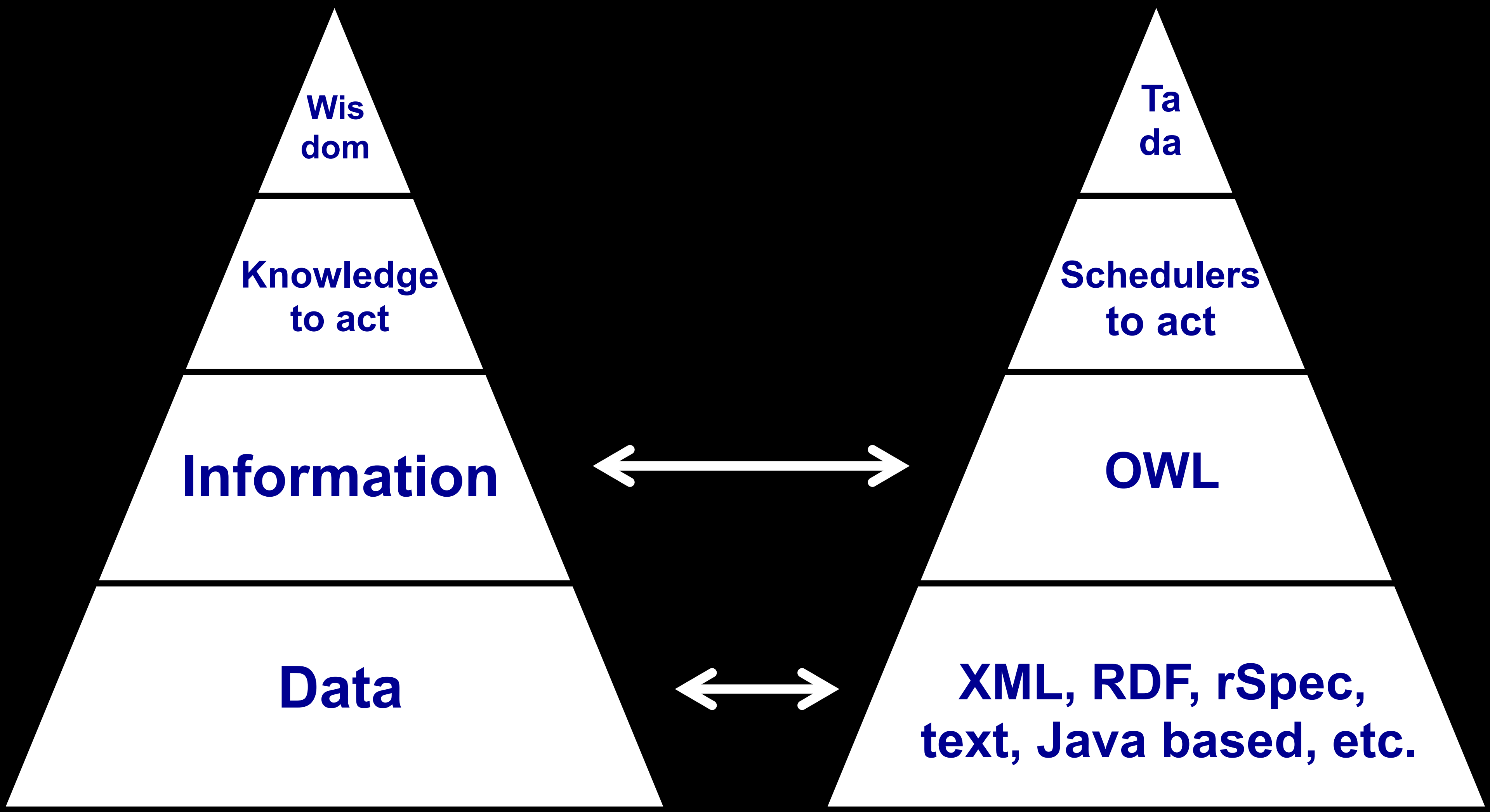
R.Koning, P.Grosso and C.de Laat
Using ontologies for resource description in the CineGrid Exchange
In: Future Generation Computer Systems (2010)

CdL

Applications and Networks become aware of each other!



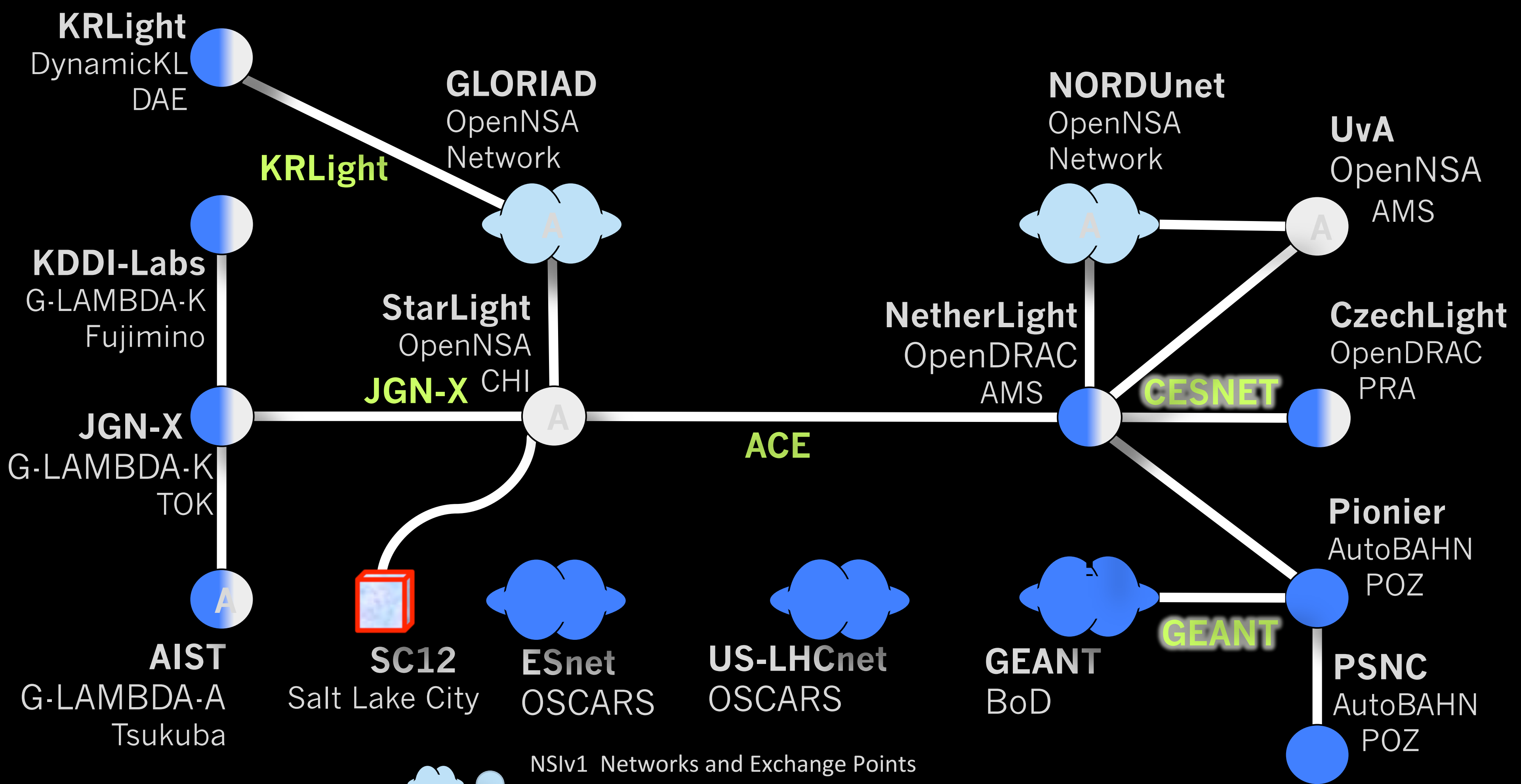
Layers


Wis
dom
Knowledge
to act
Information
Data
Ta
da
Schedulers
to act
OWL
XML, RDF, rSpec,
text, Java based, etc.

Automated GOLE + NSI

NSI v2 Beta Test Fabric   Oct 2013

Ethernet Transport Service



NSIv1 Networks and Exchange Points

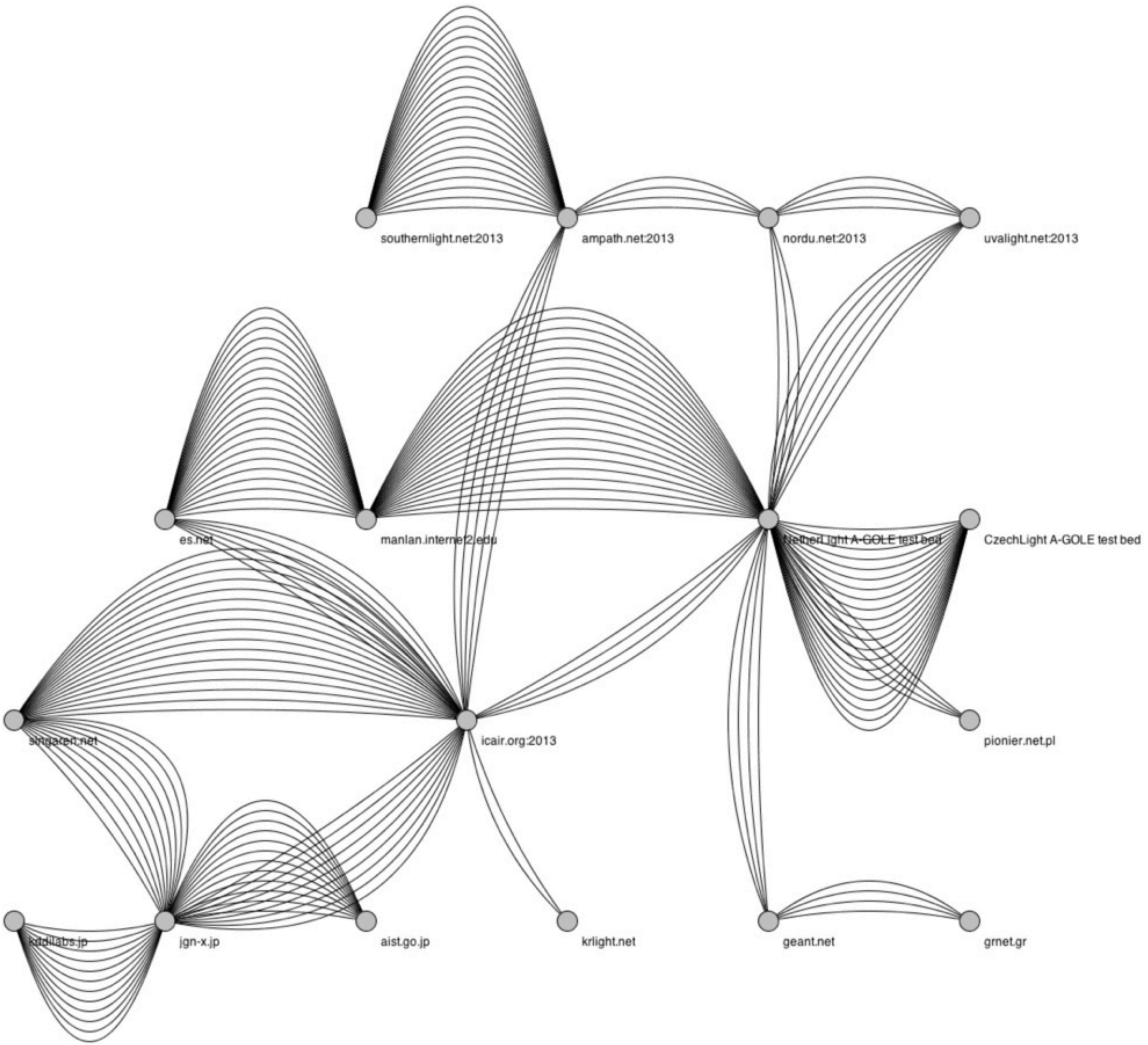
NSIv2 Networks and Exchange Points

NSI peerings (SDPs) unless otherwise indicated these are vlans 1780-1783

Planned/peerings (SDPs)

Thanks Jerry Sobieski

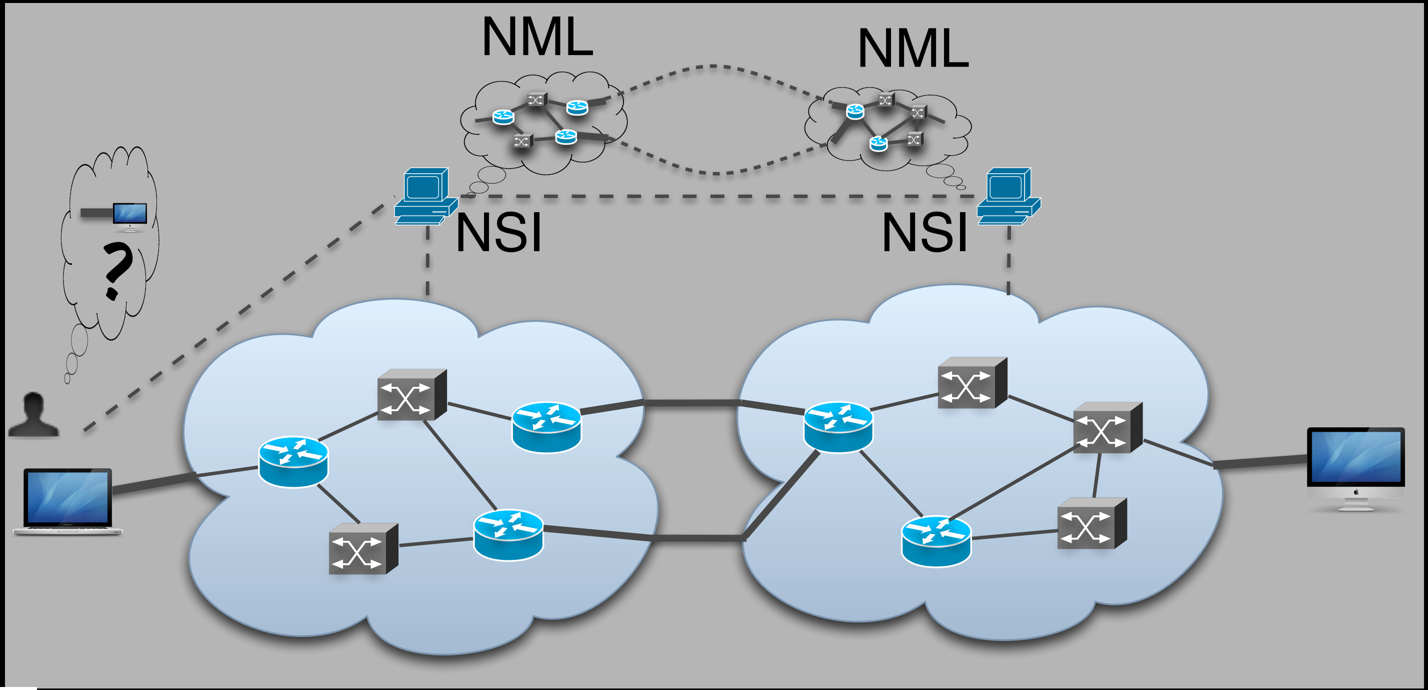
GLIF 2013 in NML



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

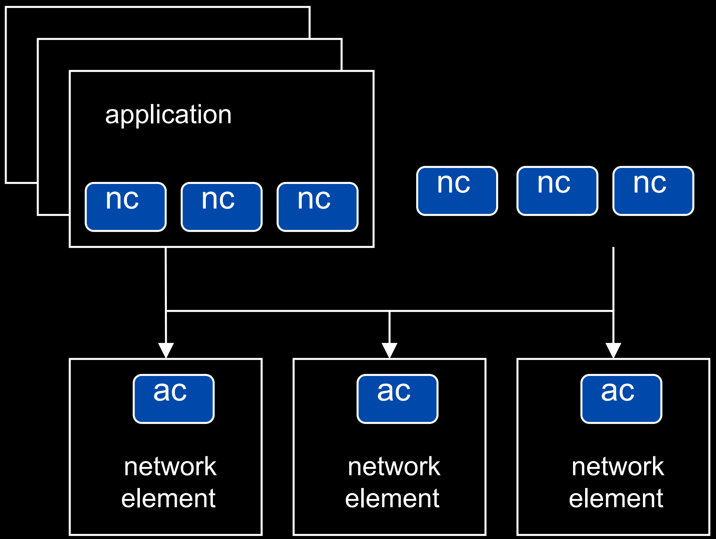
http://sne.science.uva.nl/ndl

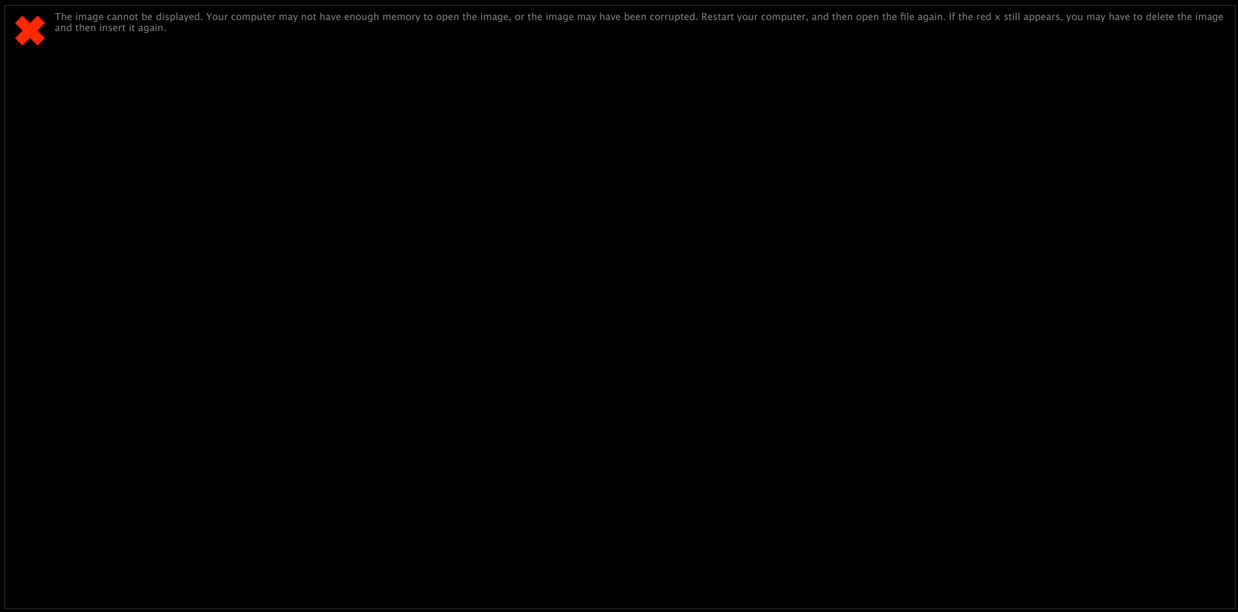
Tera-Thinking

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

User Programmable Virtualized Networks.

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





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

TouchTable Demonstration @ SC08

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*
  - *Anonymity, integrity of data in distributed data processing*
- *Sustainability*
  - *Greening infrastructure, awareness*
- ***Resilience***
  - ***Systems under attack, failures, disasters***

IJKDIJK

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

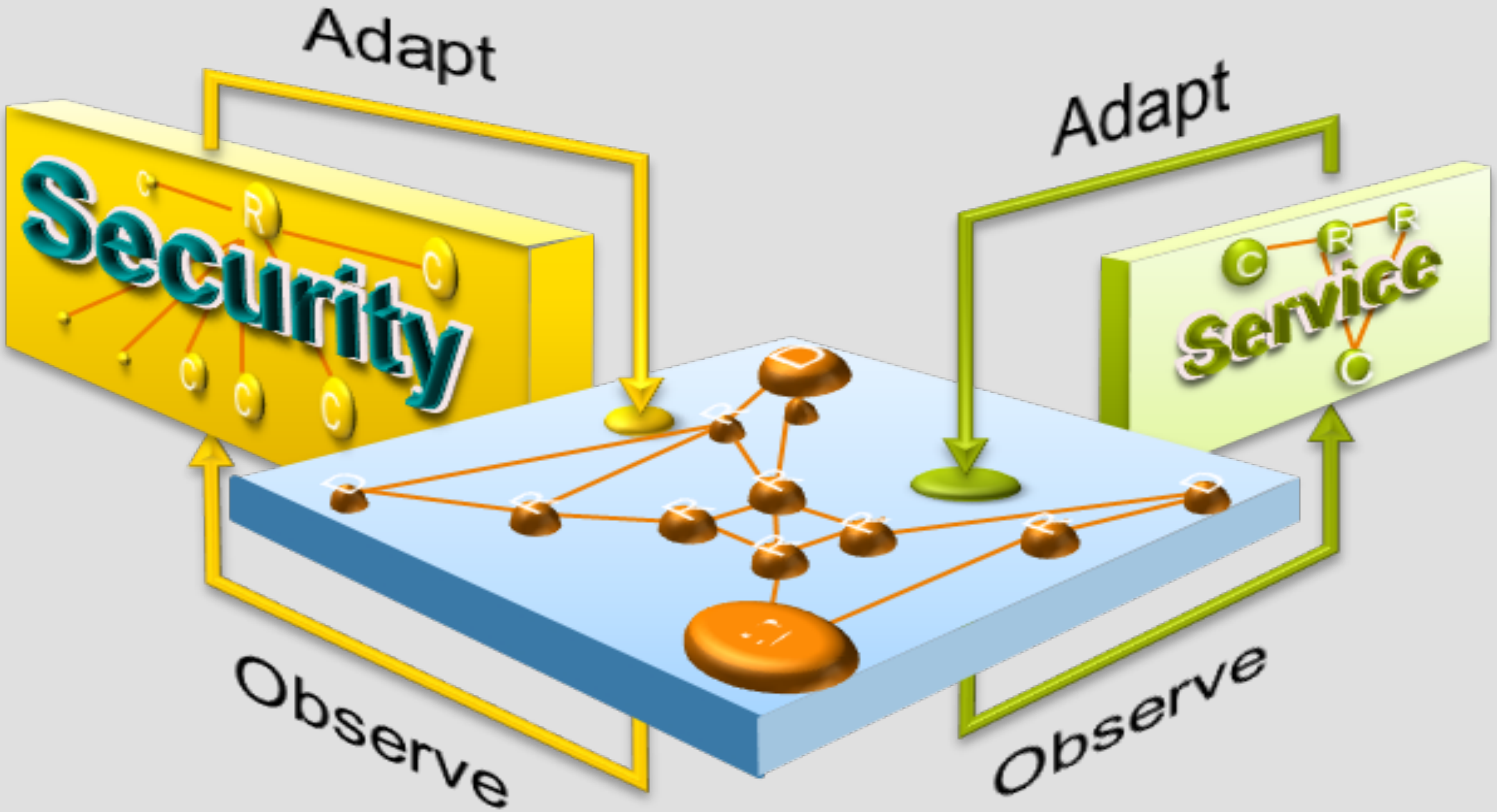
Sensor grid: instrument the dikes

First controlled breach occurred on sept 27th '08:

Many Pflops/s

Many small flows -> 12 Gb/s

SMART Infrastructure



Mission

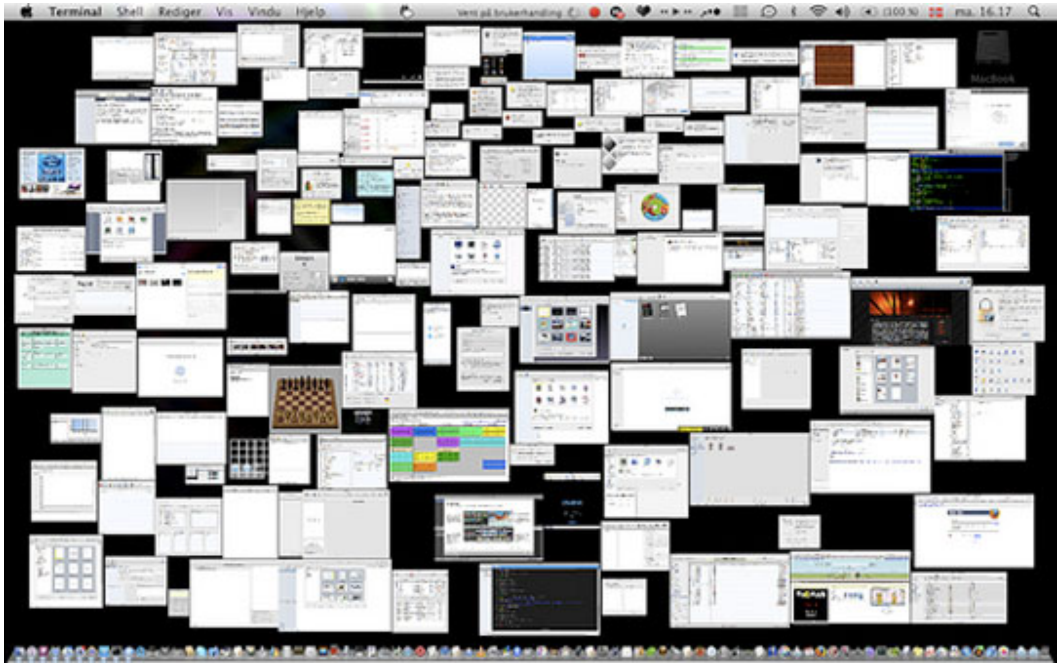
Can we create smart and safe data processing infrastructures that can be tailored to diverse application needs?

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

Physical Machines

N.D. Jebessa

And after a while…



Security

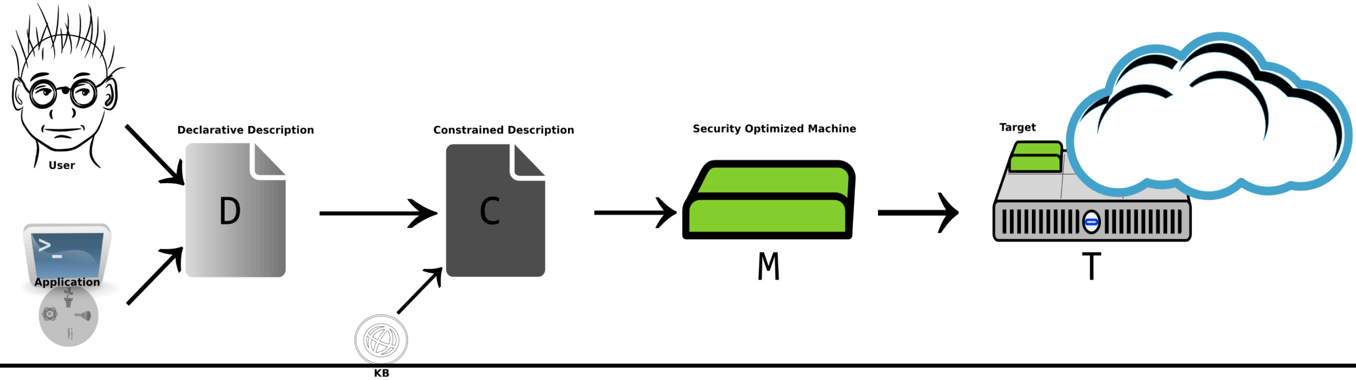
Performance

Management

Naod Duga Jebessa, Guido van 't Noordende, Cees de Laat

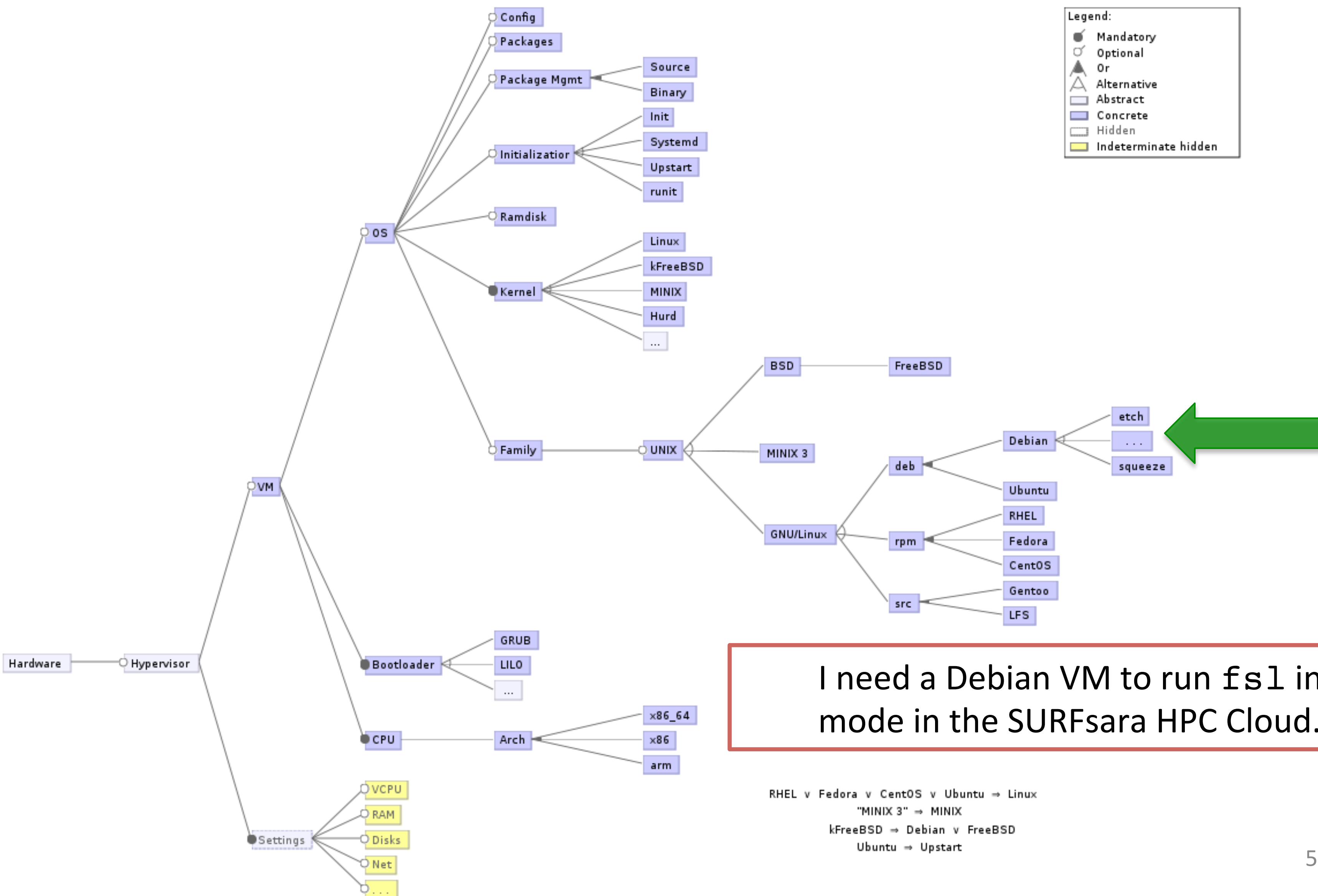
Purpose-Driven Virtual Machines

- ⇒ Virtual machine (**VM**) technology is a key enabler (e.g. clouds, mobility, green IT, …)
- ⇒ Often, a VM serves a **specific purpose** (e.g. host a bioinformatics application)
- ⇒ VM **security** & data **privacy** are very important (e.g. DNA processing in clouds)
- ⇒ **VMs with a general-purpose OS meant for physical machines exhibit redundancy**
- ⇒ **Generic VMs exhibit opacity (e.g. kernel, packages, configurations, …)**
- ⇒ A specific-purpose VM could be optimized for a **minimal TCB** (trusted computing base)
- ⇒ The VM ought to be **transparent** so as to reason about its security and **trustworthiness**



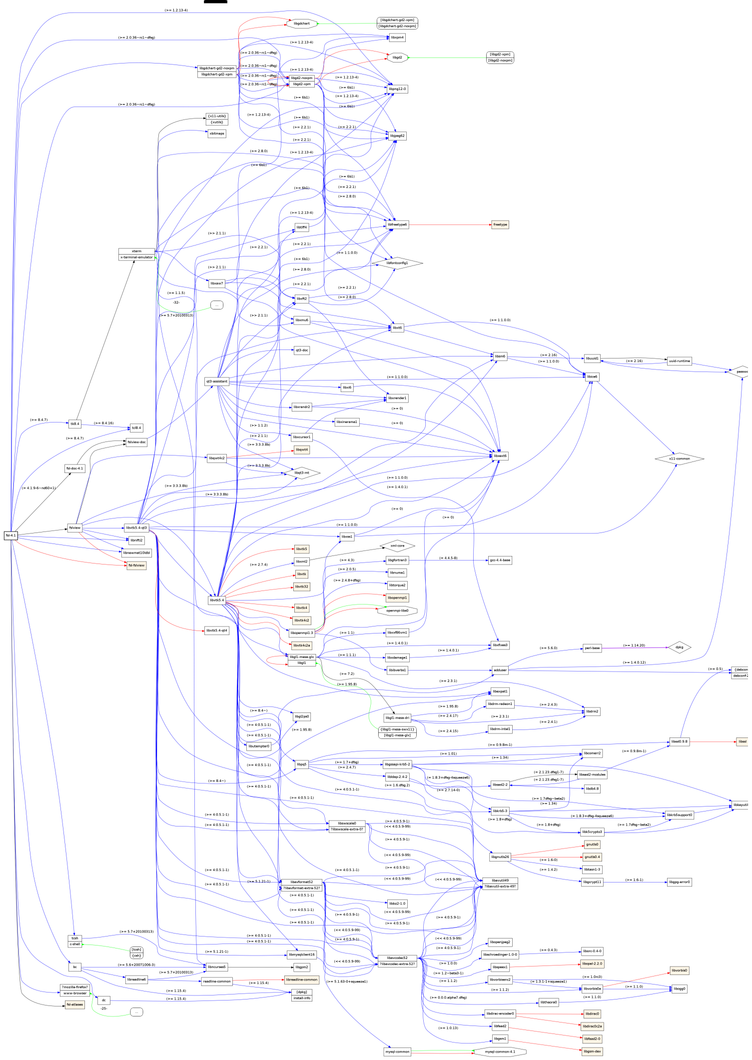
Minimal, transparent and secure VMs optimized for a specific purpose – built automatically from declarative descriptions.

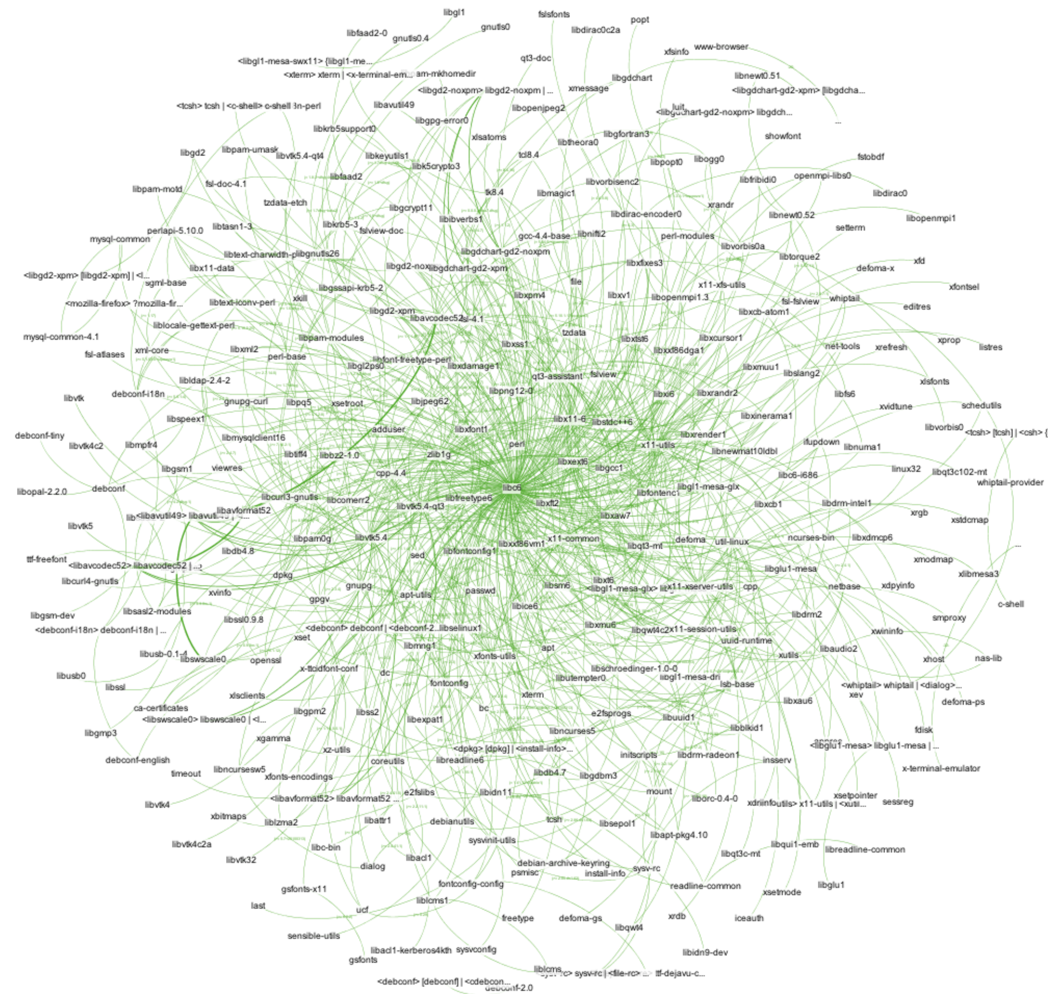
Feature Model for a Generic VM



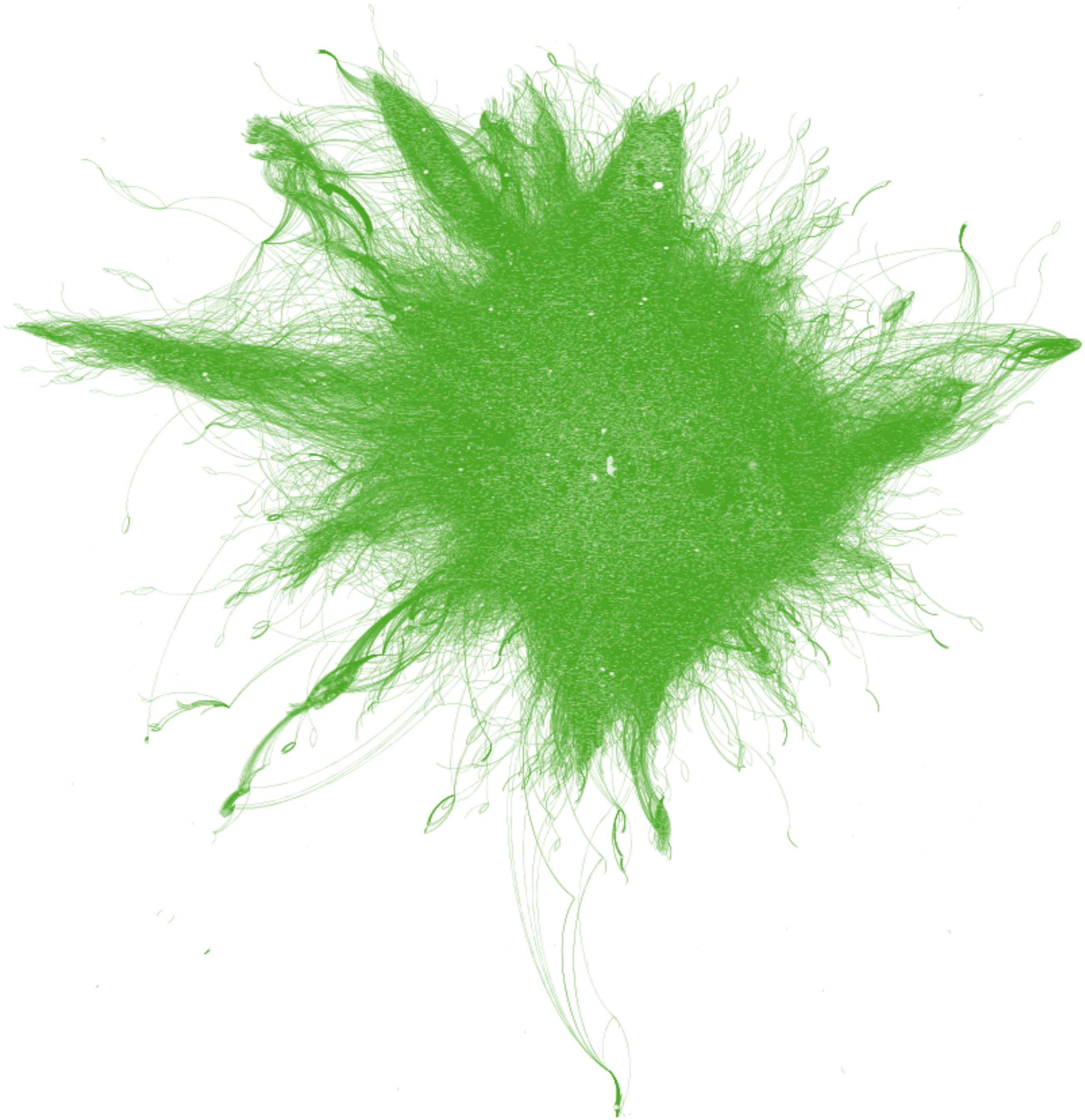
I need a Debian VM to run `fsl` in text mode in the SURFsara HPC Cloud.

Dependencies of an application





The application is `fsl-4.1`. LEFT: before dependency resolution, with all dependency constraints shown and RIGHT: resolved dependencies in a particular setup, `libc` is the center node



Mission

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

- *Capacity*
  - *Bandwidth on demand, QoS, architectures, [illegible]ics, pe[illegible]rmance*
- *Capability*
  - *[illegible]og[illegible]ma[illegible]y, v[illegible]lizat[illegible]n, c[illegible]mple[illegible]y, s[illegible]antics, w[illegible]rkflows*
- *Se[illegible]*
  - *[illegible]nony[illegible]y, [illegible]teg[illegible]y o[illegible] da[illegible] in distributed data processing*
- *Sustainability*
  - *Greening infrastructure, awareness*
- *Resilience*
  - *Systems under attack, failures, disasters*

SMART

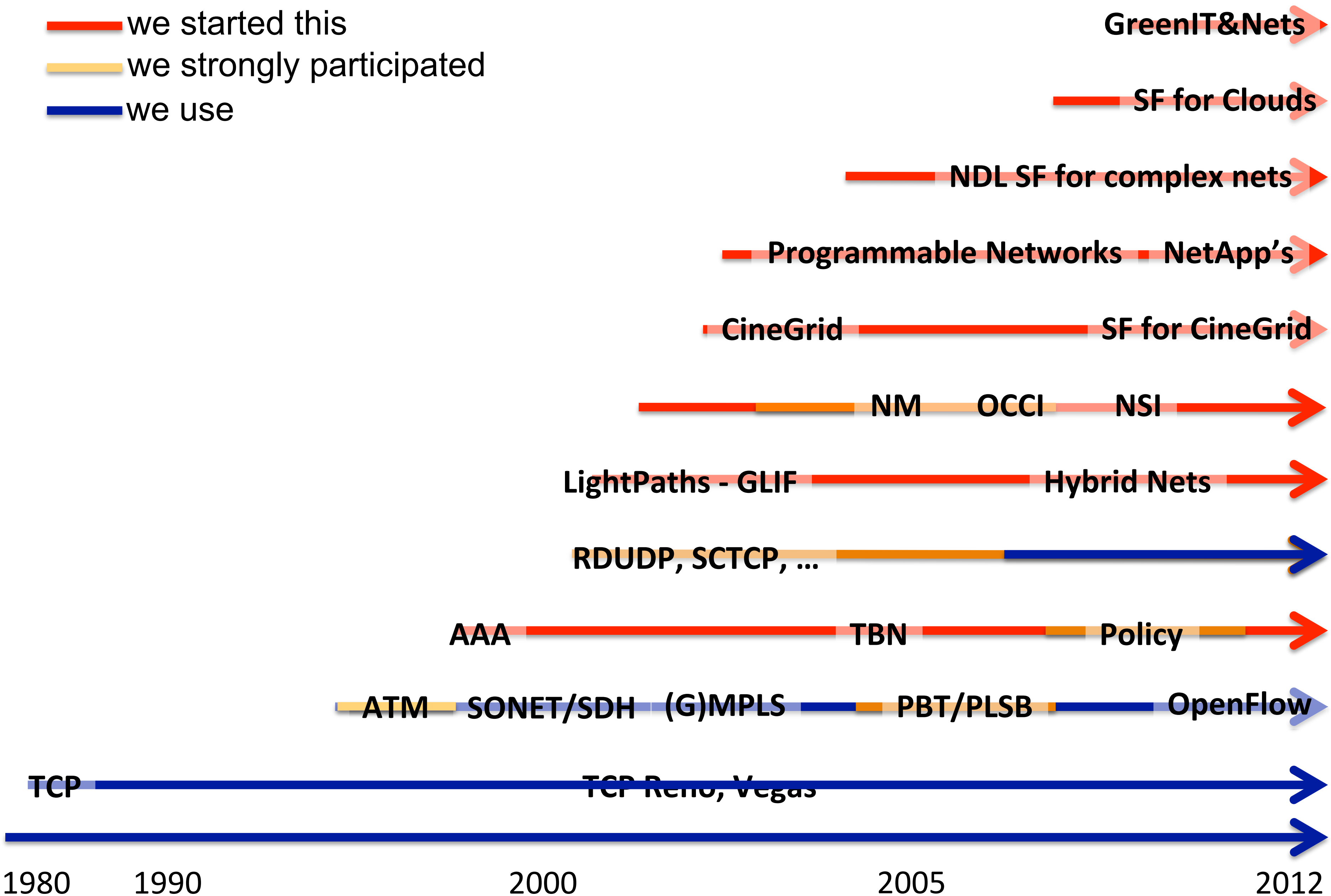
I want to →

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

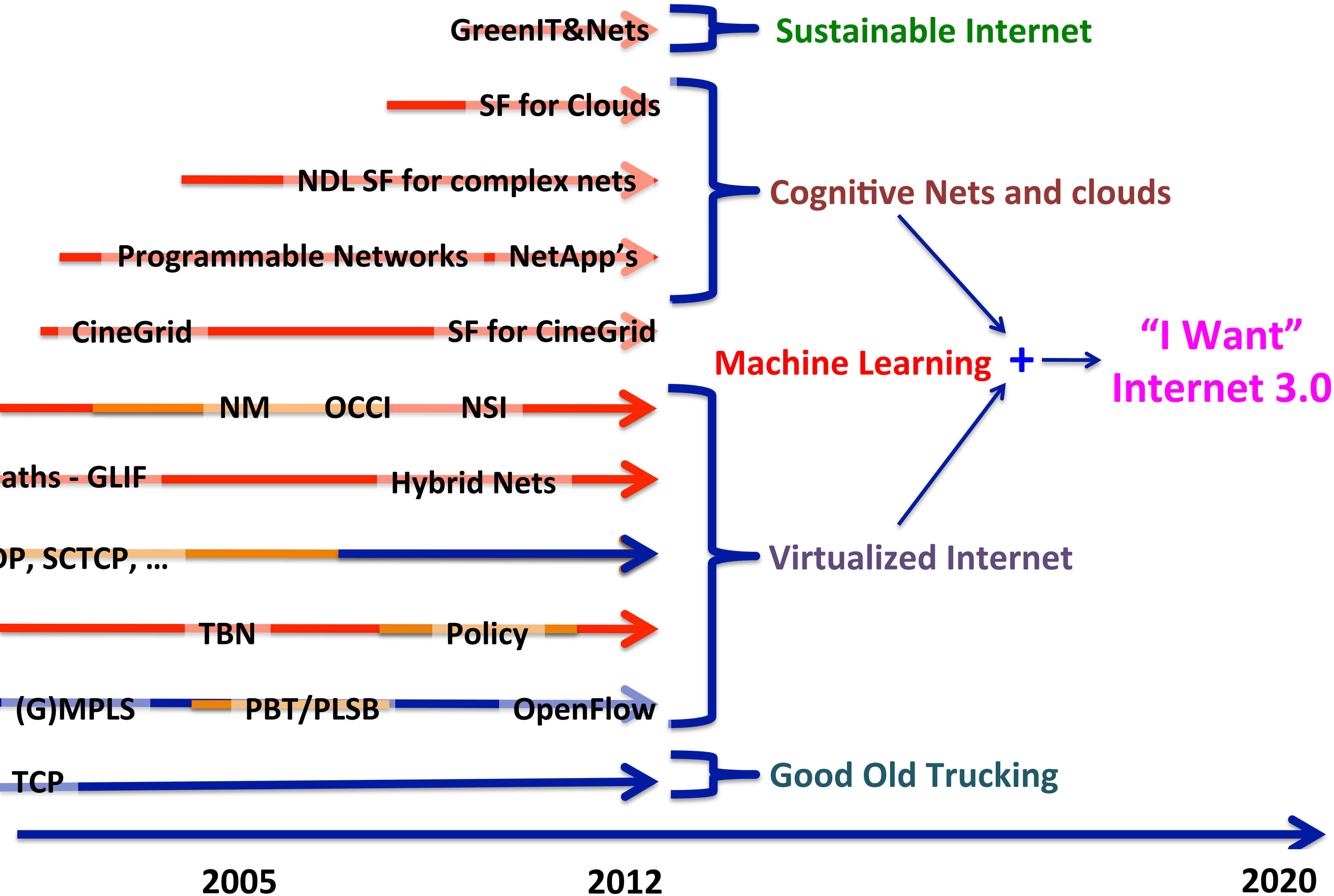
- Big Bugs Bunny can be on multiple servers on the Internet.
- Movie may need processing / recoding to get to 4K for Tiled Display.
- Needs deterministic Green infrastructure for Quality of Experience.
- Consumer / Scientist does not want to know the underlying details.

➔ His refrigerator also just works!

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

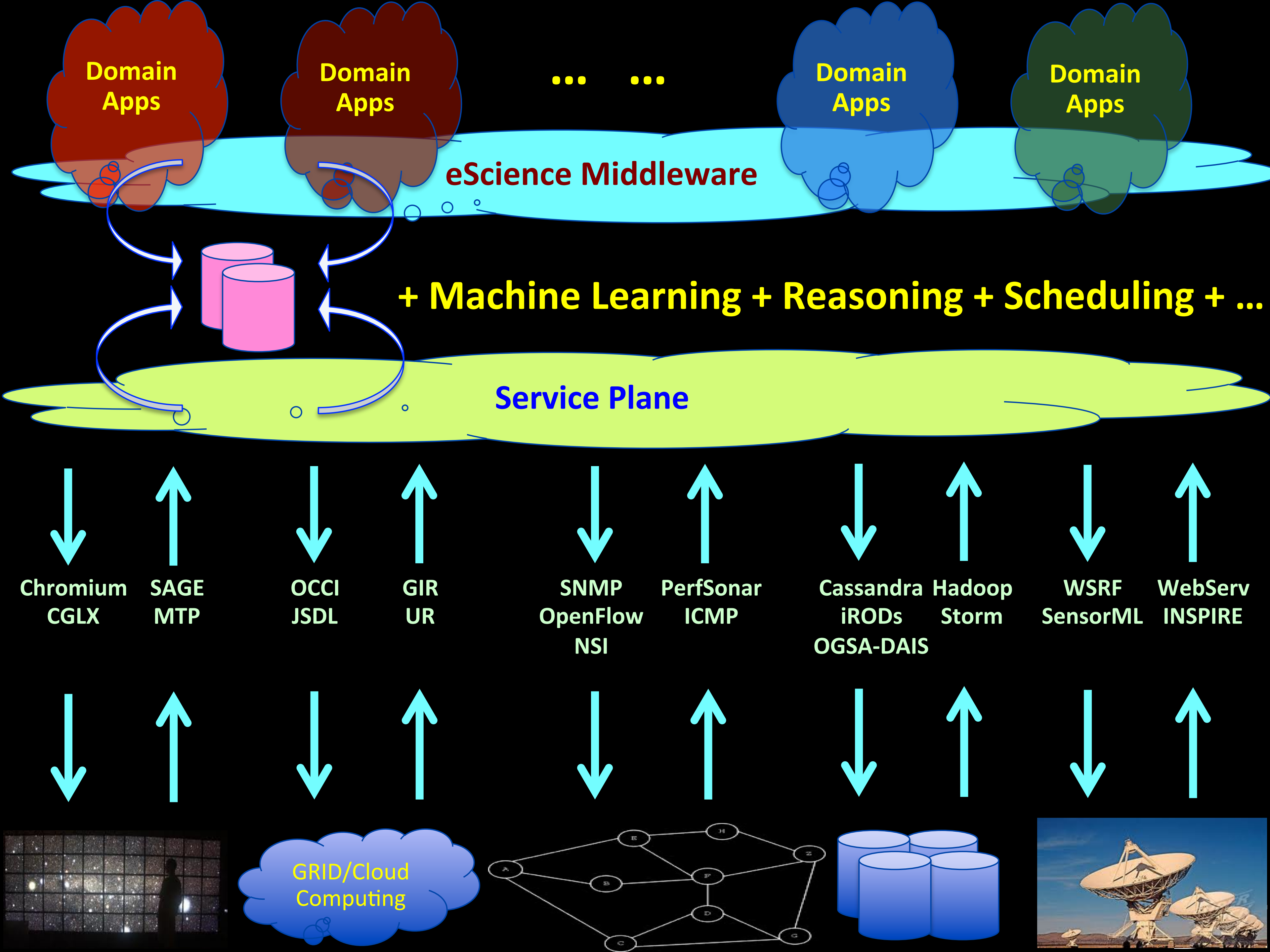
TimeLine

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

I retire

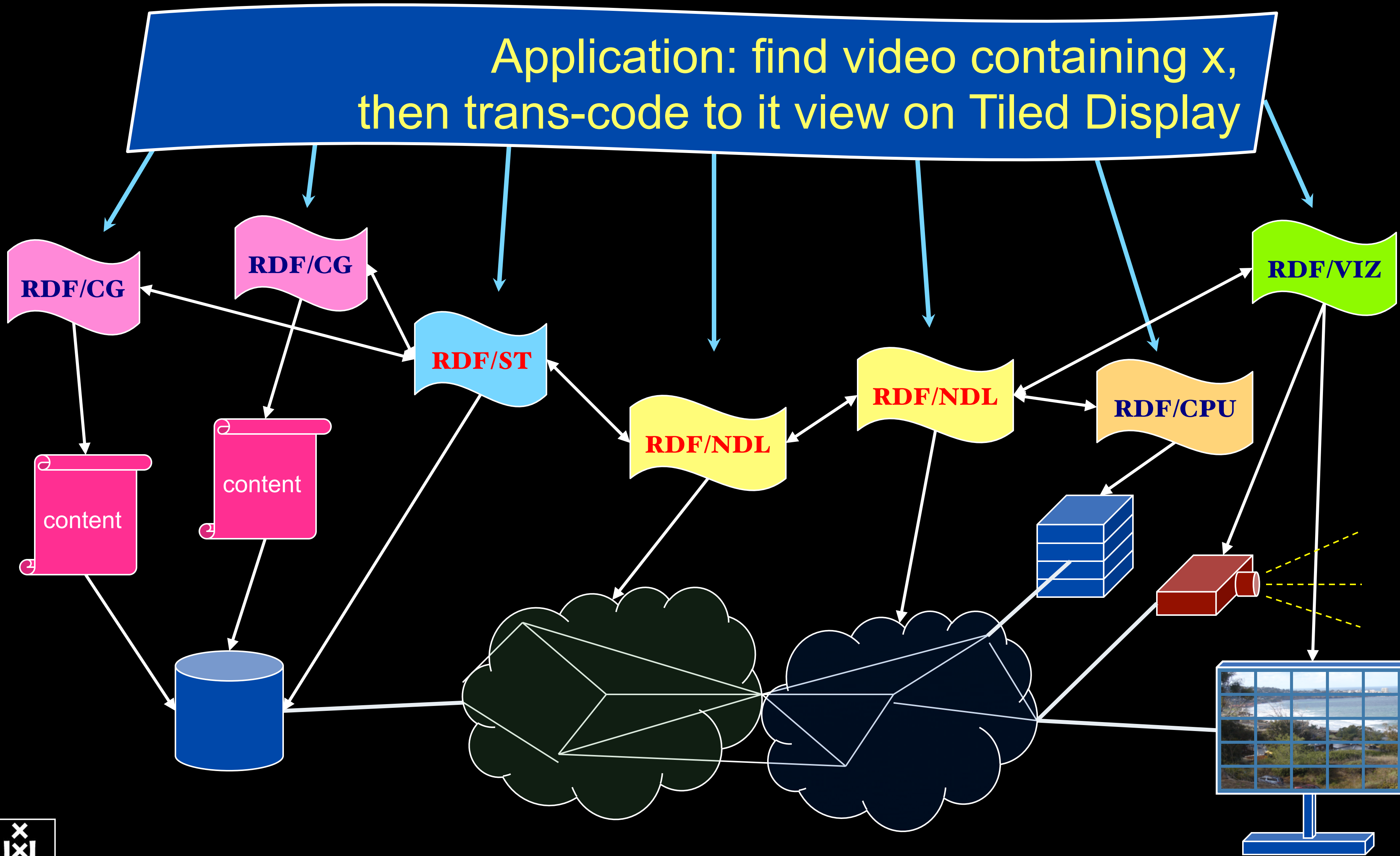
2020

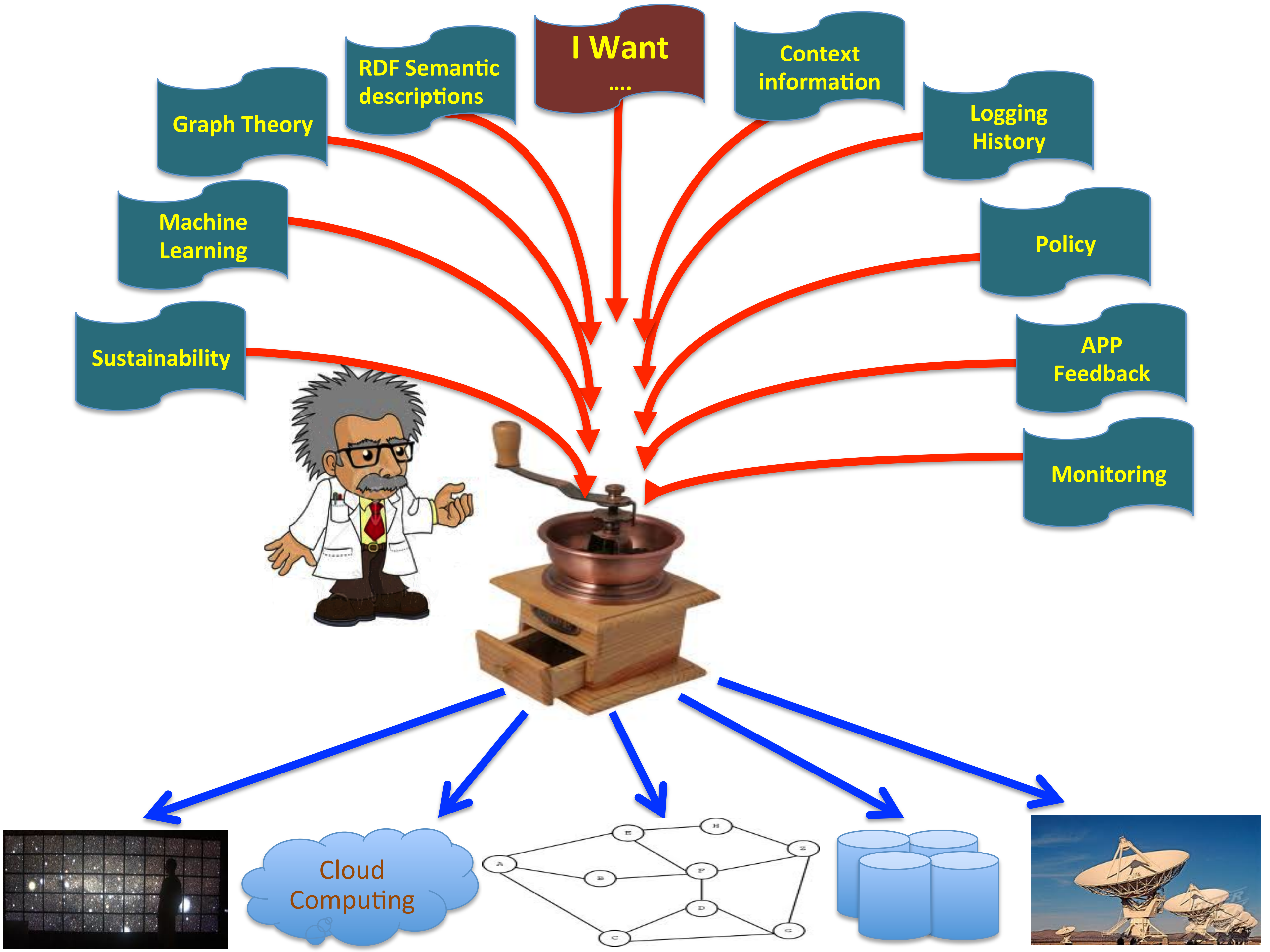
2040



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

RDF describing Infrastructure
"I want"





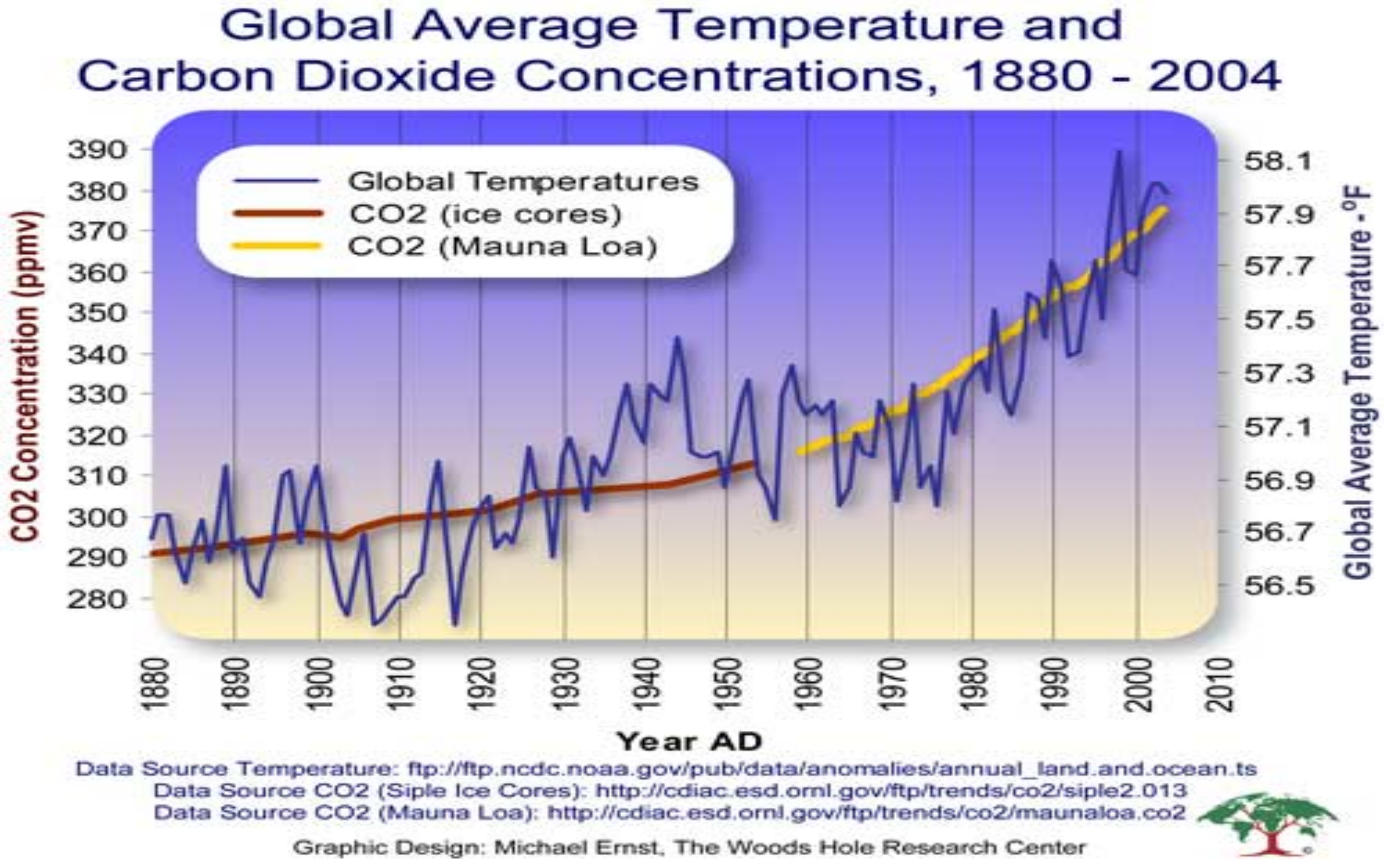
I Want ....
RDF Semantic descriptions
Context information
Graph Theory
Logging History
Machine Learning
Policy
Sustainability
APP Feedback
Monitoring
Cloud Computing

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*
  - *Anonymity, integrity of data in distributed data processing*
- ***Sustainability***
  - ***Greening infrastructure, awareness***
- *Resilience*
  - *Systems under attack, failures, disasters*

Need for GreenIT



Greening the Processing

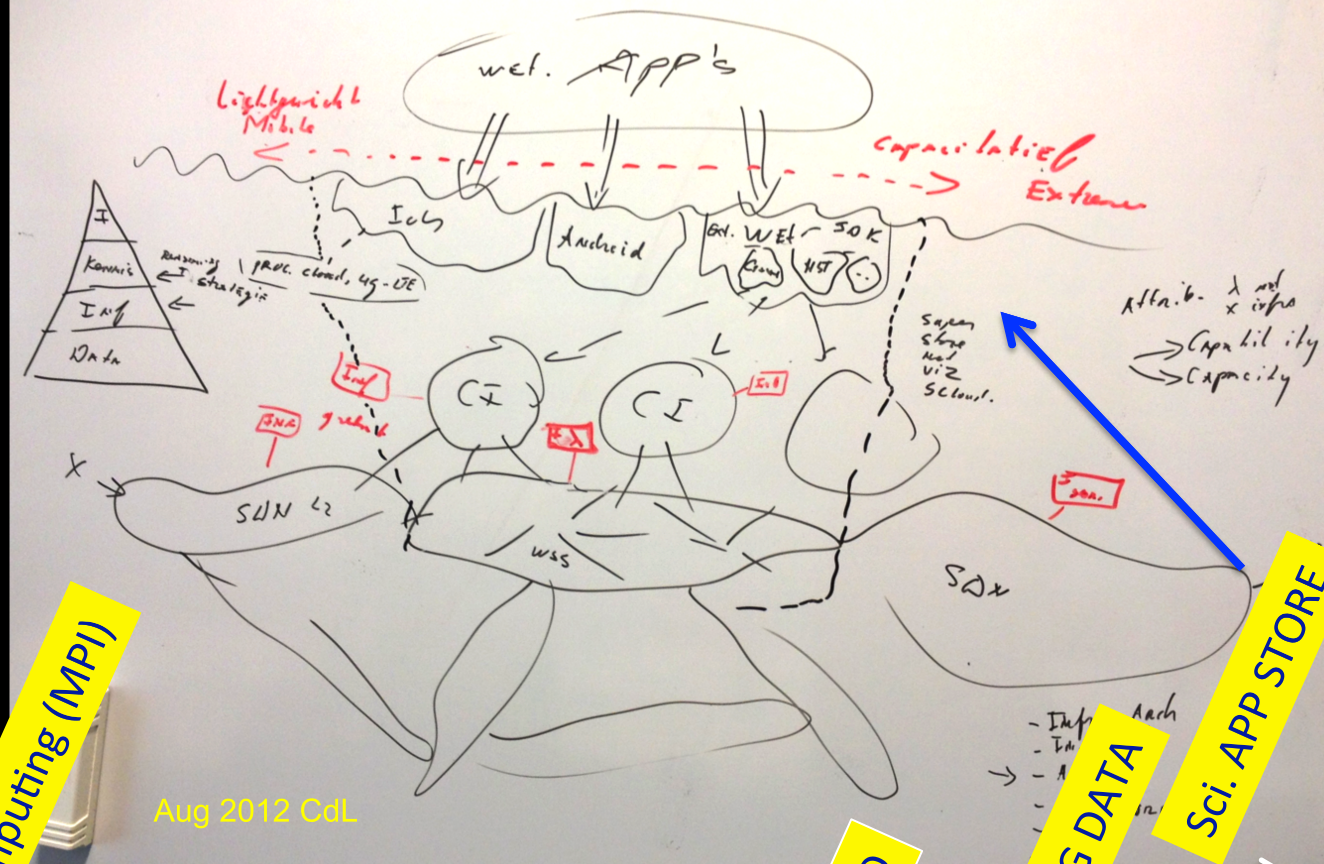
ECO-Scheduling

Conclusion

I want a MiS system!

Catchphrase first used in "Encounter At Farpoint" (28 September 1987) by Gene Roddenberry, and thereafter used in many episodes and films, instructing a crew member to execute an order.

TimeLine



Remote Procedure Call

Distributed Computing (MPI)

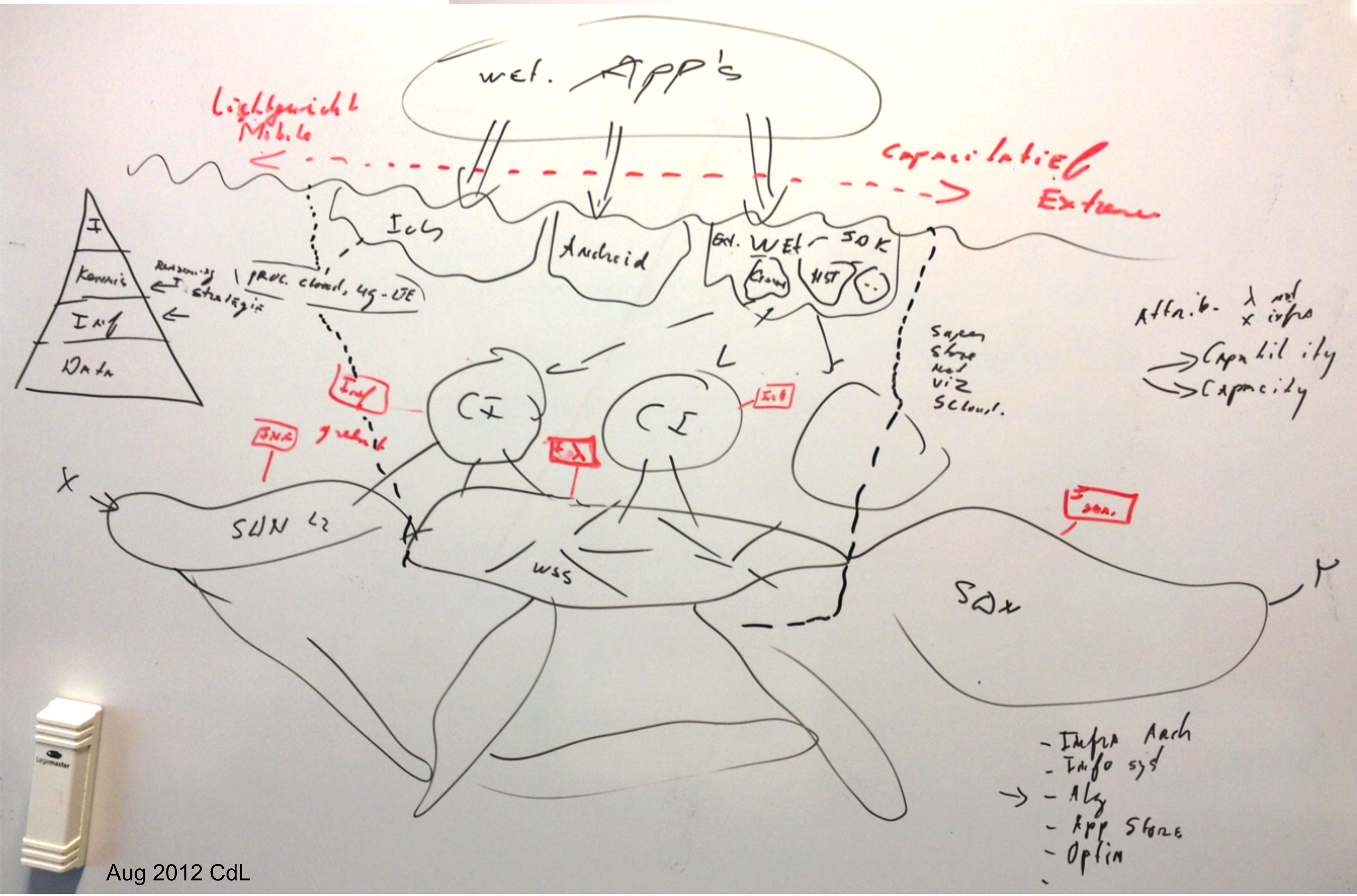
GRID

CLOUD

BIG DATA

Sci. APP STORE

1980 | 1990 | 2000 | 2005 | 2013



1980 1990 2000 2005 2013

Research Data Alliance

- https://rd-alliance.org
- The Research Data Alliance implements the technology, practice, and connections that make Data Work across barriers.
- The Research Data Alliance aims to accelerate and facilitate research data sharing and exchange.
  - Working groups and interest groups
  - Joining groups and attendance at the twice-yearly plenary meetings is open.
- Plenary Sep 2014 hosted by the Netherlands - Amsterdam
  - Conference Management Team (CMT) Chair: Peter Doorn (DANS)
  - Program Committee (PC): co-chairs Cees de Laat & Wouter Los (UvA)
  - Satellite Events Committee (SEC): Jeroen Rombouts (TUD)

Research direction

- Control of Infrastructure
- Information on Infrastructure
- Virtualization
- Networked data processing
- Sustainability & Complexity

Events on the horizon

- PIRE & OpenScienceDataCloud.org
  - Workshop June 2014 @ UvA
- Research Data Alliance
  - Conference in Amsterdam Sept 2014

The constant factor in our field is Change!

The 50 years it took Physicists to find one particle, the Higgs,
we came from:

Assembler, Fortran, COBOL, VM, RSX11, Unix, c, Pascal,
SmallTalk, DECnet, VMS, TCP/IP, c++, Internet, WWW,
ATM, Semantic Web, Photonic networks, Google, Grid,
Phyton, FaceBook, Twitter, Cloud, SDN, Data^3, App's

to:

DDOS attacks destroying Banks and BitCoins!

Conclusion:

Need for Safe, Smart, Resilient Sustainable Infrastructure.

Questions?

Arie Taal
Paola Grosso
Ana Oprescu
Cees de Laat
Marc Makkes
Ralph Koning
Bas Terwijn
Leon Gommans
Fahimeh Alizadeh
Pieter Adriaans
Cosmin Dumitru
Karst Koymans
Yuri Demchenko
Rob Meijer
Karel van der Veldt
Reggie Cushing
Rudolf Strijkers
Miroslav Zivkovic
Jan Sipke van der Veen
Naod Duga Jebessa
Spiros Koulouzis
Hao Zhu
Sander Klous
Jaap van Ginkel
Guido van 't Noordende
Jeroen van der Ham
Mikolaj Baranowski
Steven de Rooij
Paul Klint
Ngo Tong Canh
Souley Madougou
Adianto Wibisono
Magiel Bruntink
Zhiming Zhao
Anna Varbanescu
Marijke Kaat
Niels Sijm
Hans Dijkman
Gerben de Vries
Adam Belloum
Arno Bakker
Marian Bubak
Daniel Romao
Erik-Jan Bos
Peter Bloem

http://delaat.net

http://sne.science.uva.nl

http://www.os3.nl/

http://i4dw.nl/

http://dsrc.nl/

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

http://pire.opensciencedatacloud.org

http://staff.science.uva.nl/~delaat/pire/

https://rd-alliance.org

http://envri.eu

Trip supported by: ENVRI