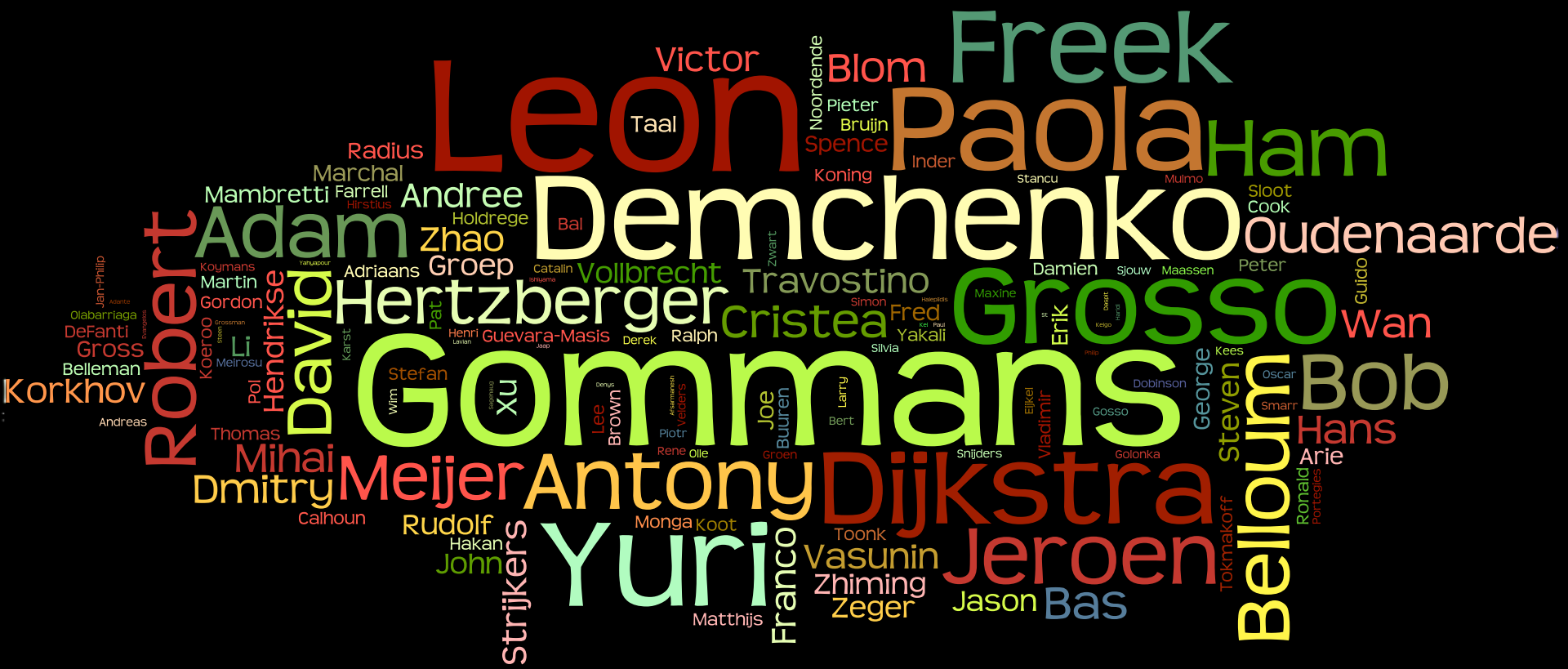
System and Network Engineering Research for Big Data Sciences

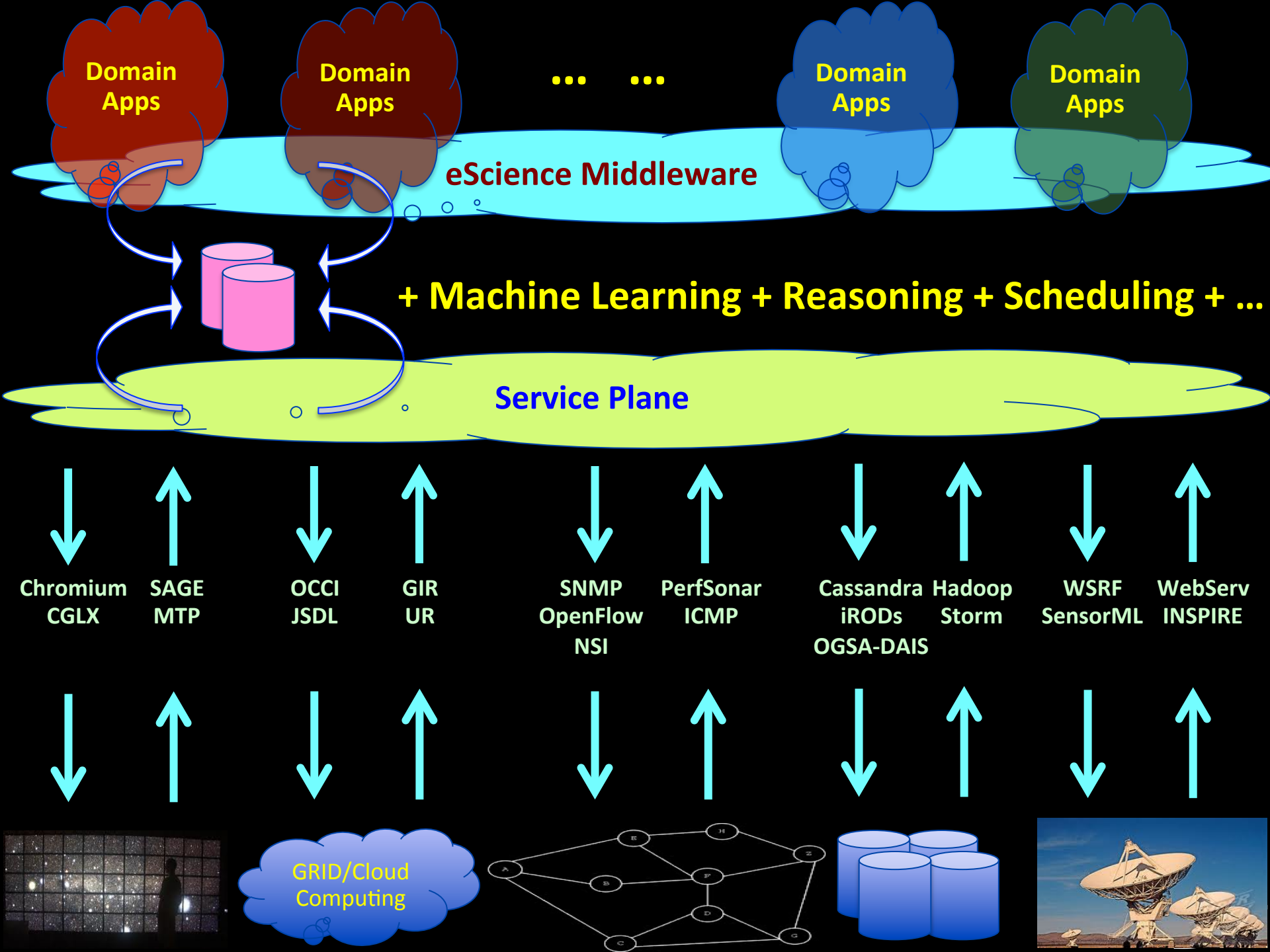
Cees de Laat



Why?

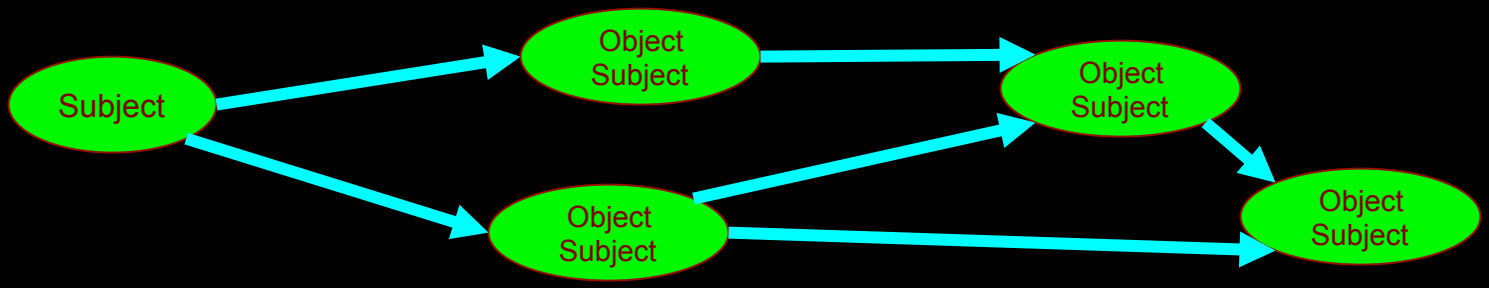
Because we can!

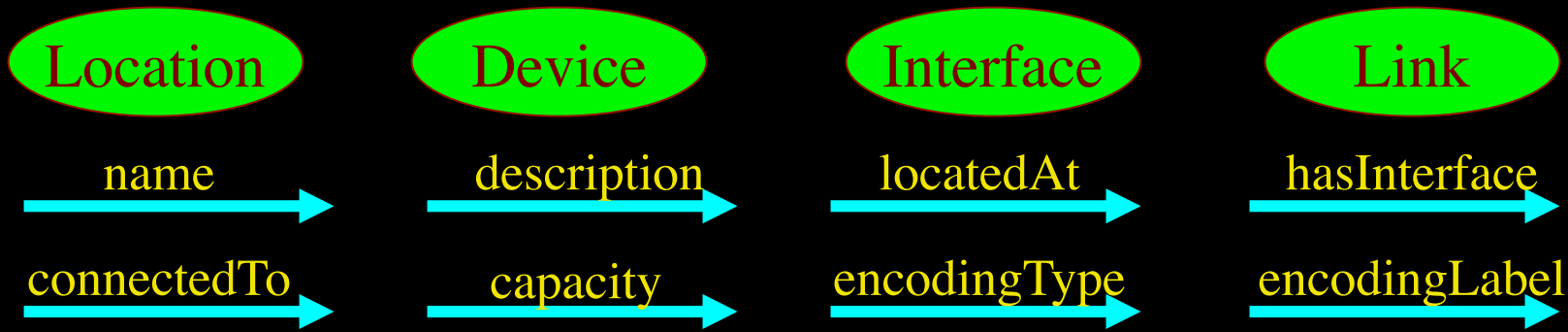
We investigate: for complex networks!


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

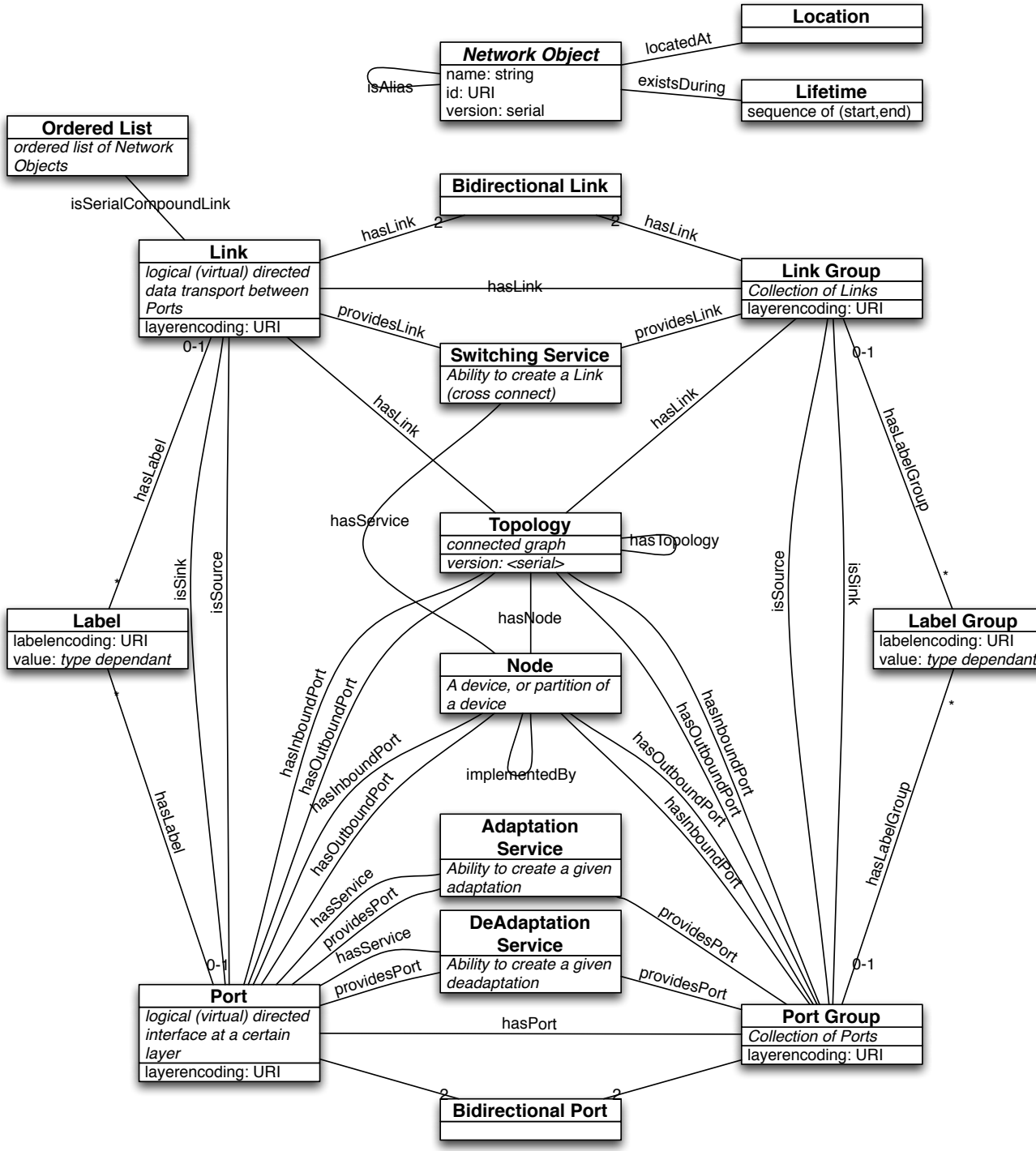
LinkedIN for Infrastructure

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





NML OFG spec


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

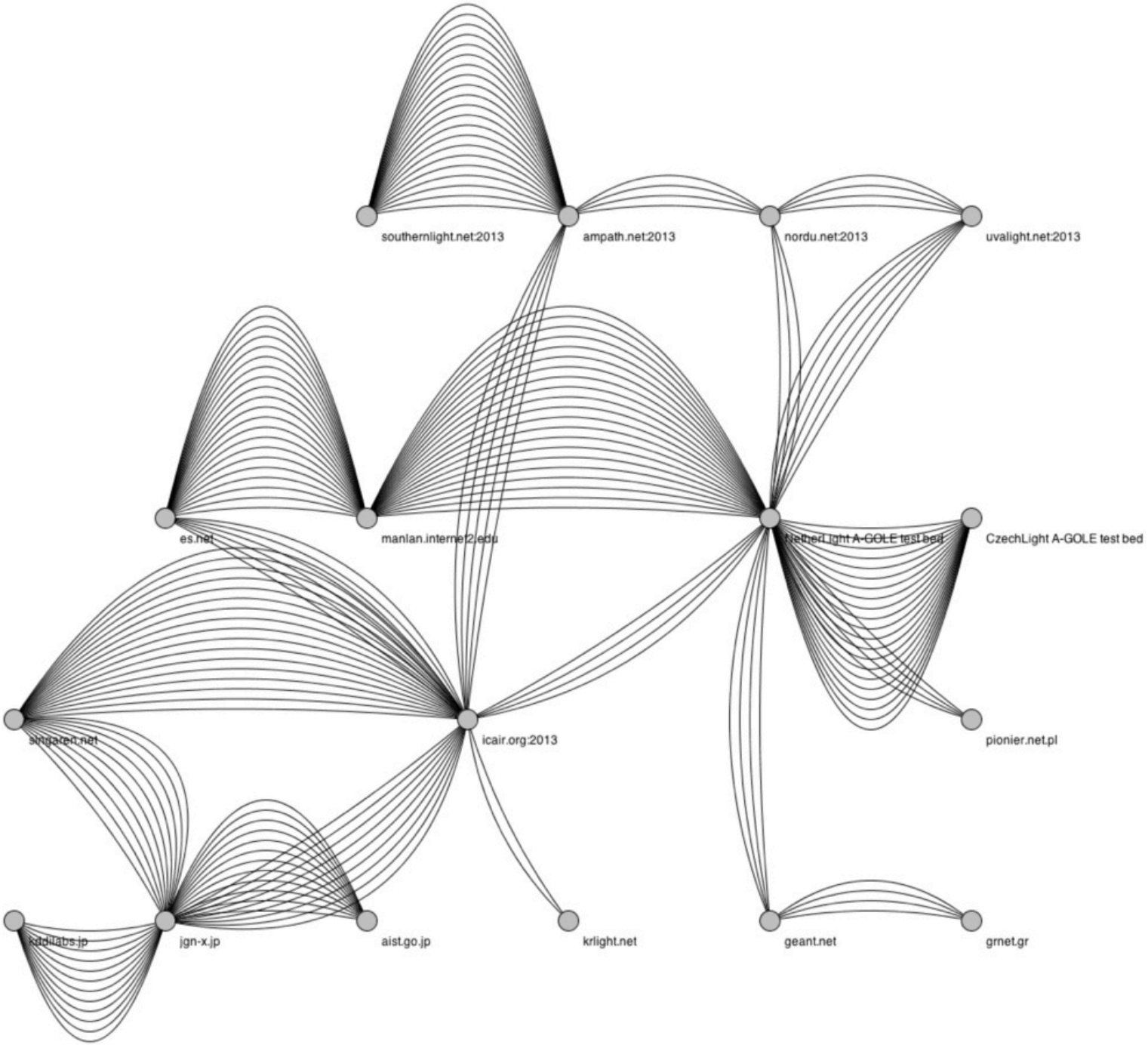
NetherLight in RDF

```
<?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:resource="#tdm3.amsterdam1.netherlight.net:503/2"/>
      <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/
 </ndl:Interface>
```

GLIF 2013

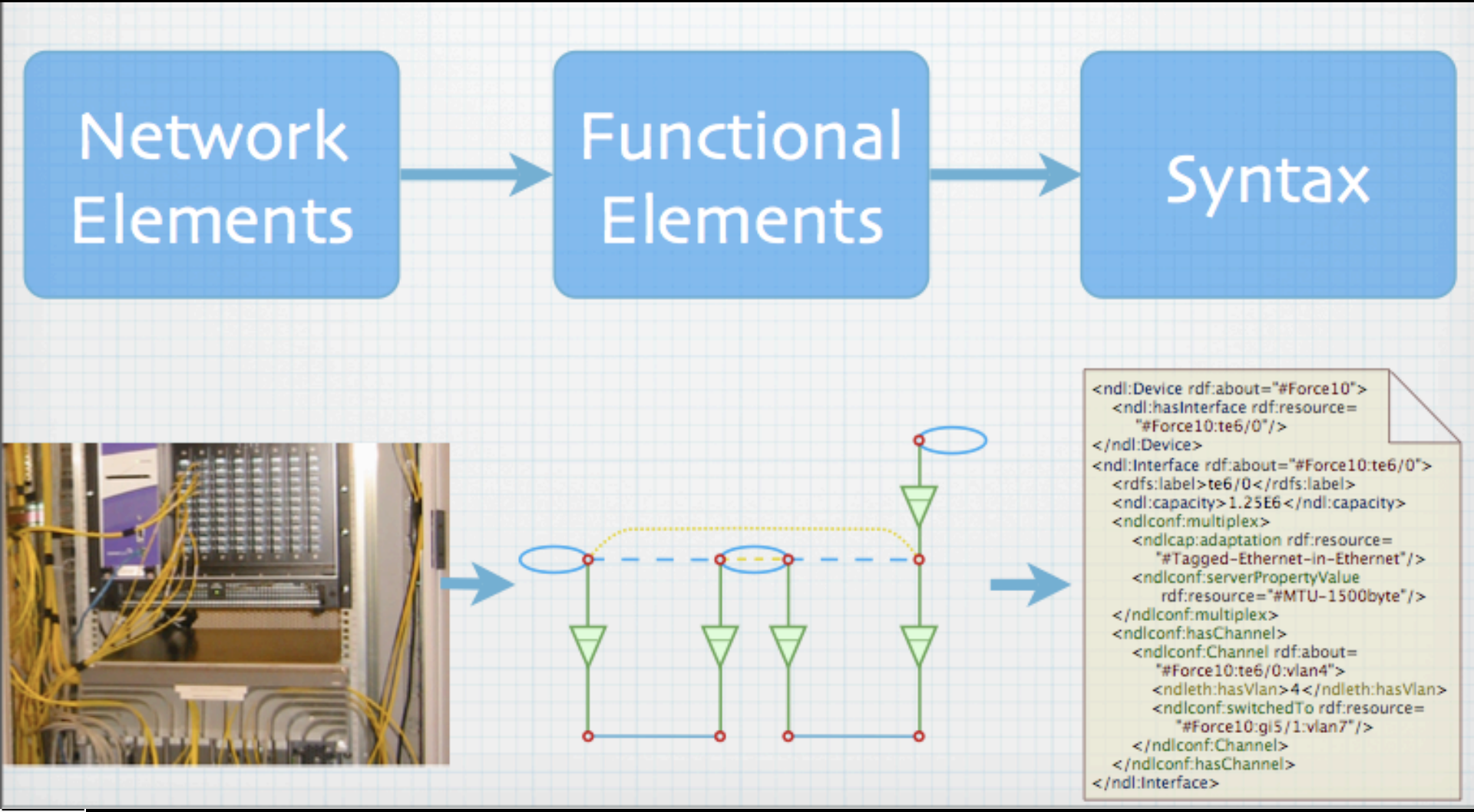

southernlight.net:2013
ampath.net:2013
nordu.net:2013
uvalight.net:2013
es.net
manlan.internet2.edu
NetherLight A-GOLE test bed
CzechLight A-GOLE test bed
singaren.net
icair.org:2013
pionier.net.pl
kddilabs.jp
jgn-x.jp
aist.go.jp
krlight.net
geant.net
grnet.gr

Network Description Language

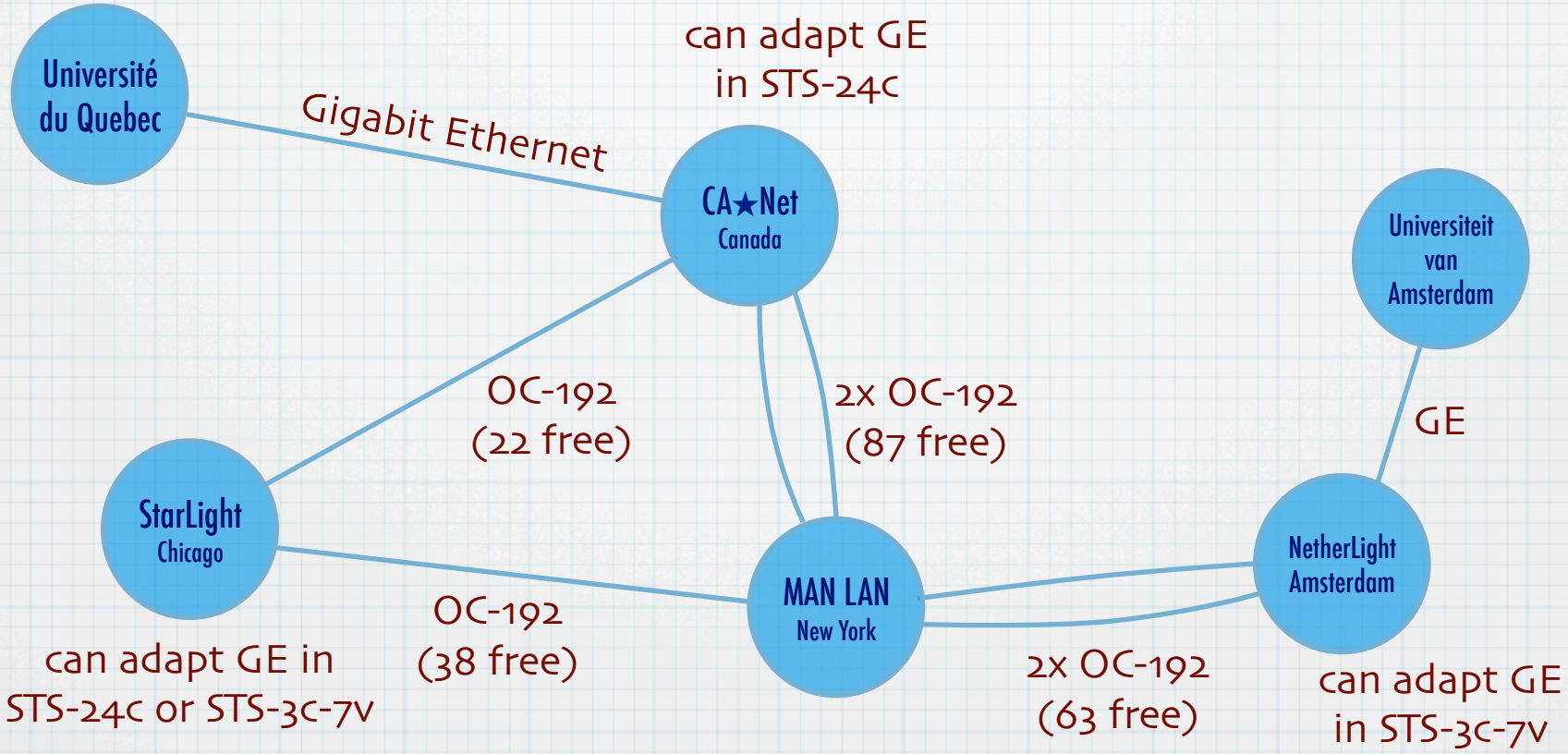
Choice of RDF instead of XML syntax

Grounded modeling based on G0805 description:

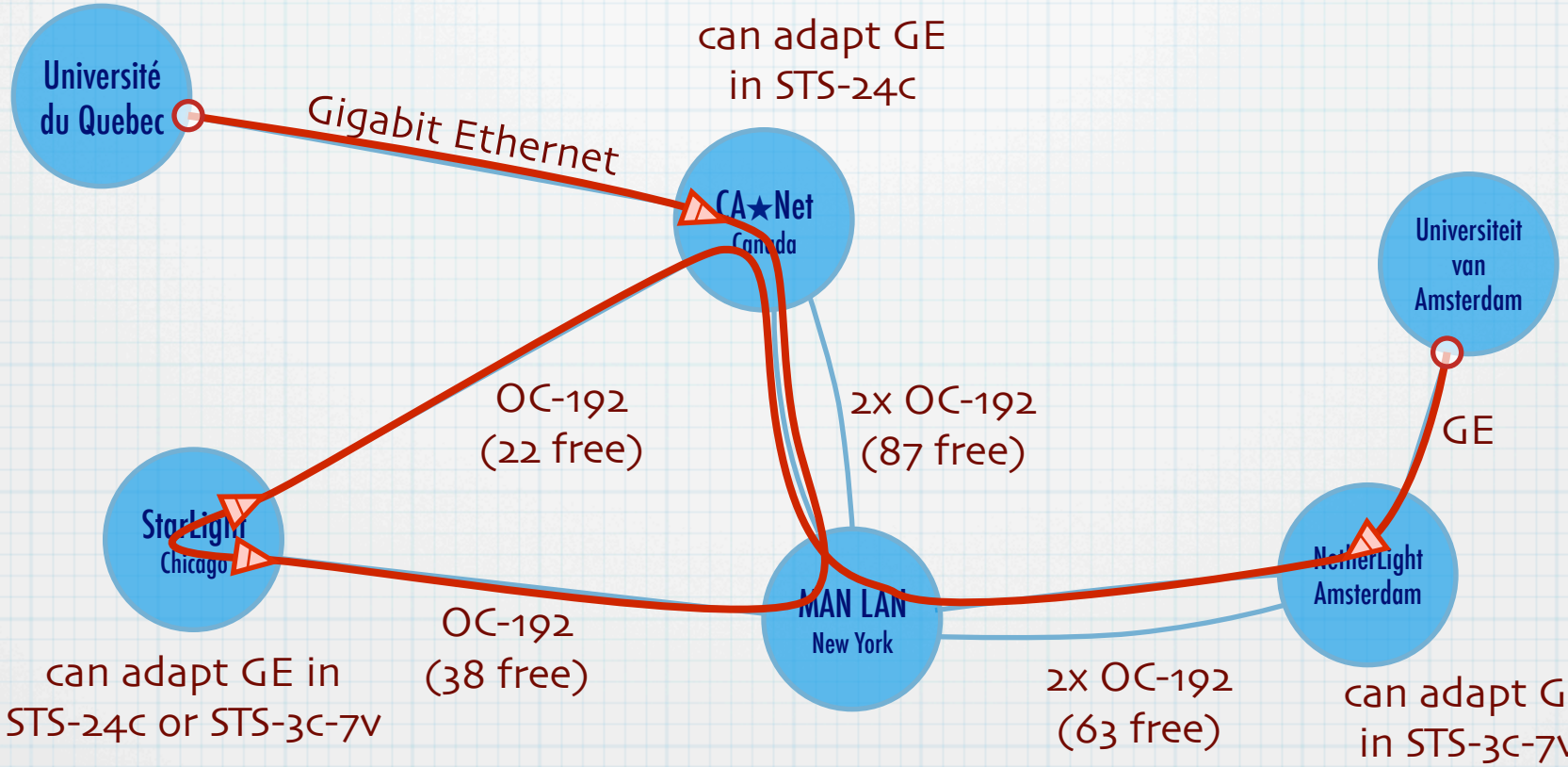
Article: F. Dijkstra, B. Andree, K. Koymans, J. van der Ham, P. Grosso, C. de Laat, *"A Multi-Layer Network Model Based on ITU-T G.805"*



A weird example

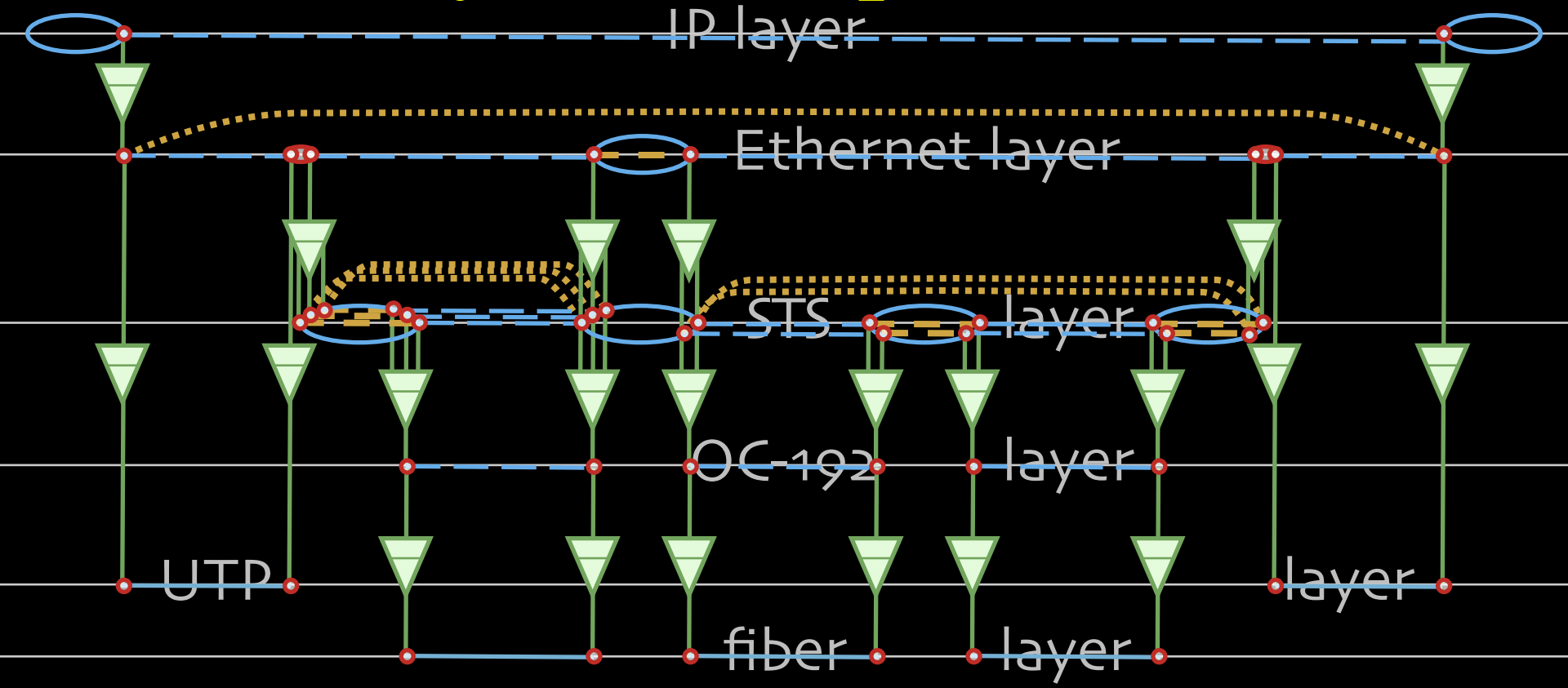


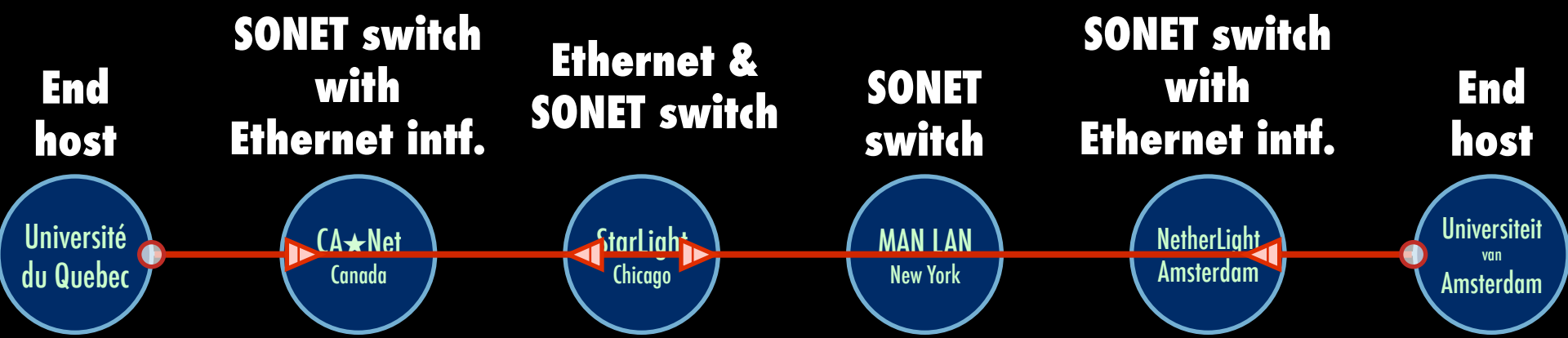
The result :-)



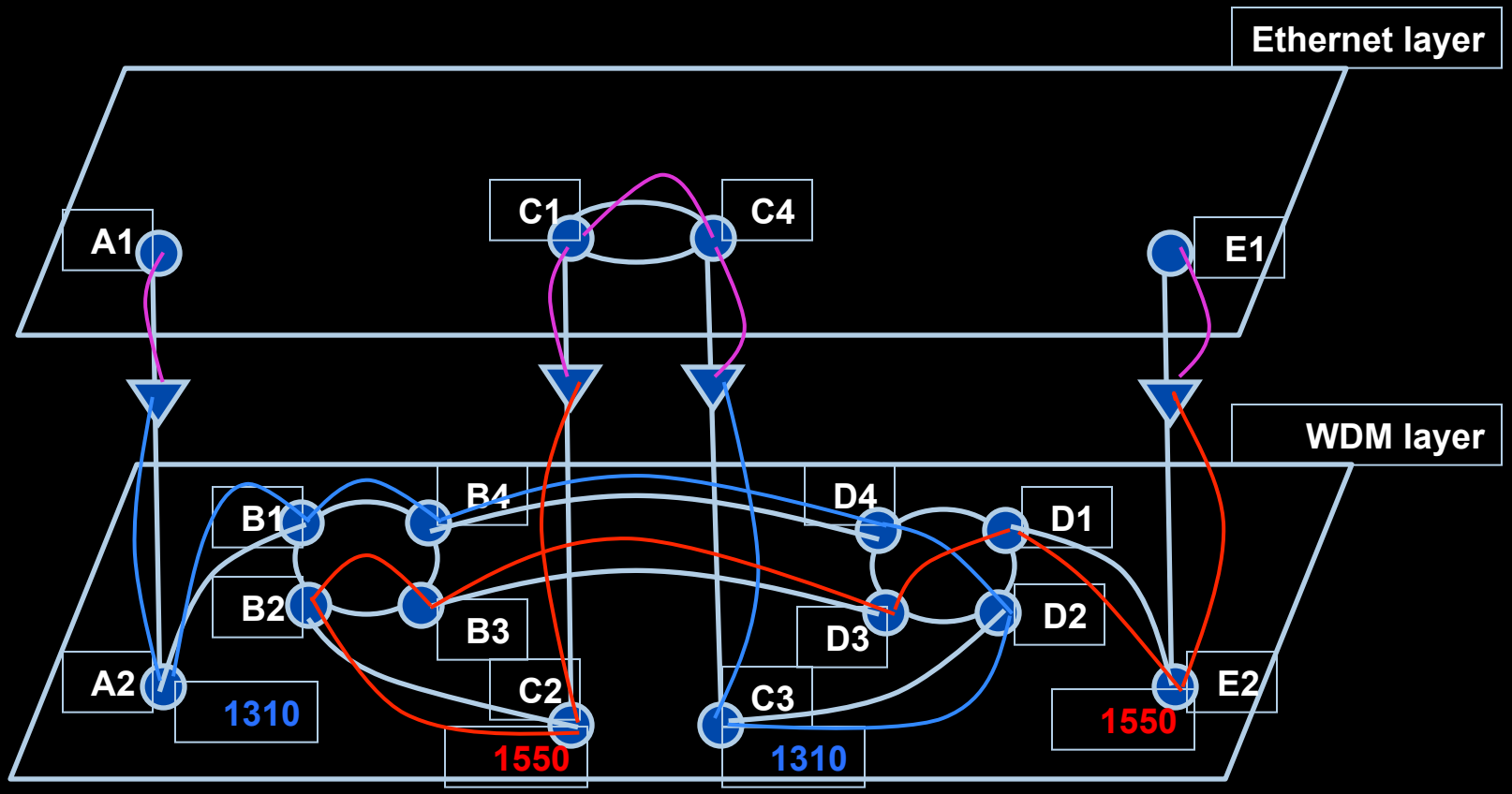
Thanks to Freek Dijkstra & team

Multi-layer descriptions in NDL





Multi-layer Network PathFinding



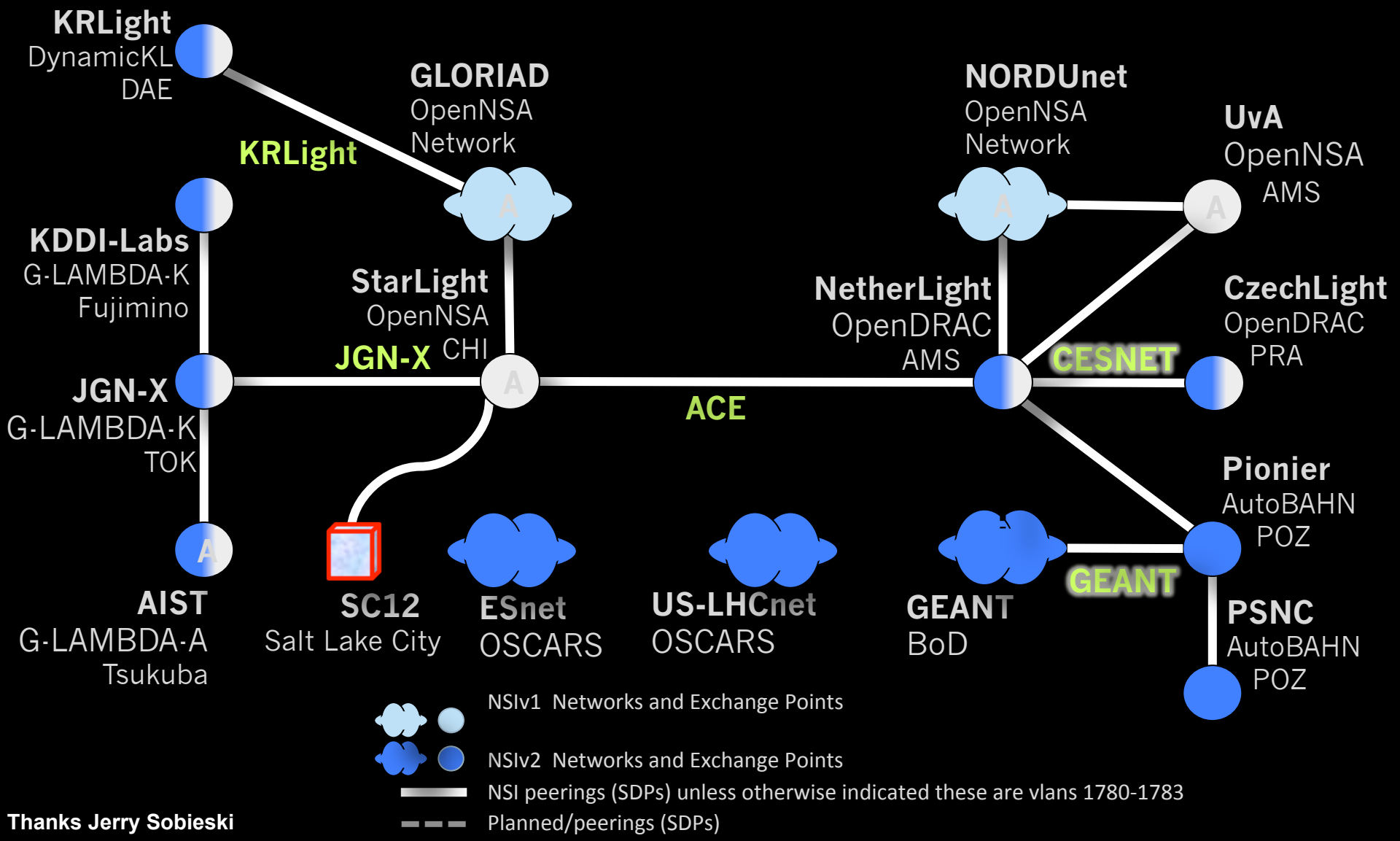
Path between interfaces A1 and E1:
A1-A2-B1-B4-D4-D2-C3-C4-C1-C2-B2-B3-D3-D1-E2-E1

Scaling: Combinatorial problem

Automated GOLE + NSI

Joint NSI v1+v2 Beta Test Fabric Nov 2012

Ethernet Transport Service

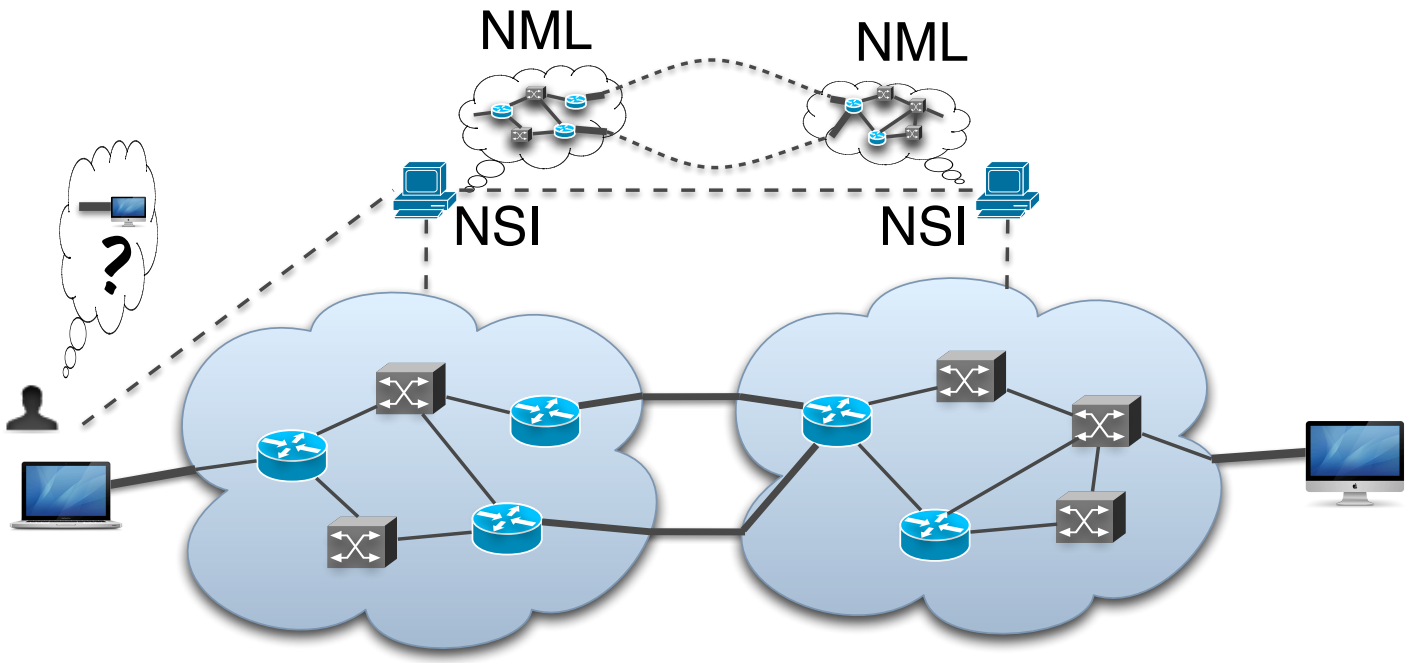


Thanks Jerry Sobieski

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

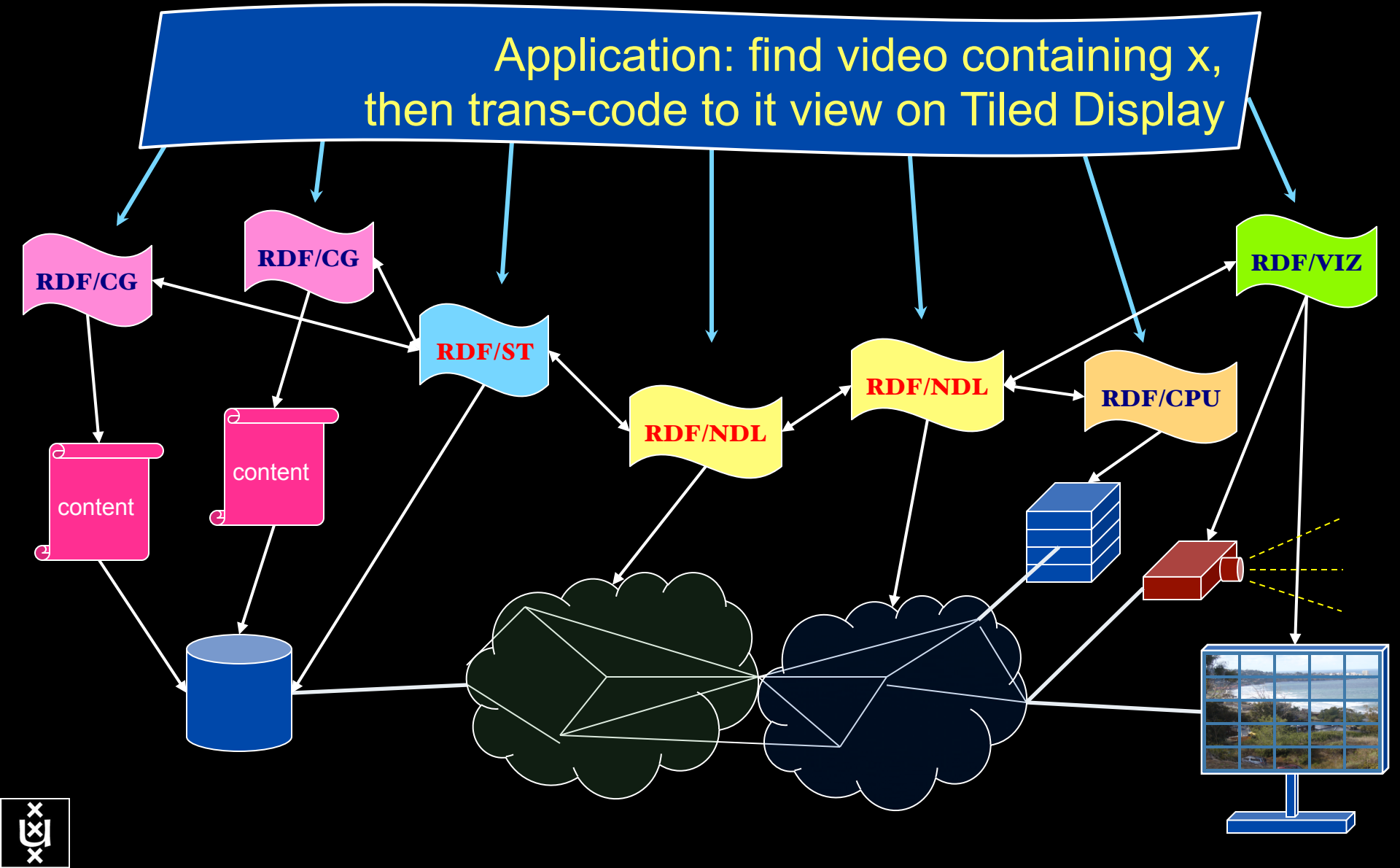
OpenGridForum
OPEN FORUM | OPEN STANDARDS

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.

RDF describing Infrastructure "I want"



CdL

Applications and Networks become aware of each other!

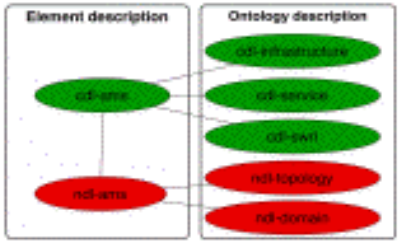
CineGrid Description Language

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

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

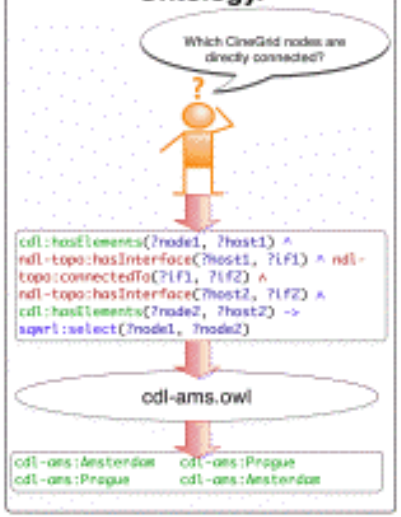
CDL has bindings to the NDL ontology that enables descriptions of network components and their interconnections.

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



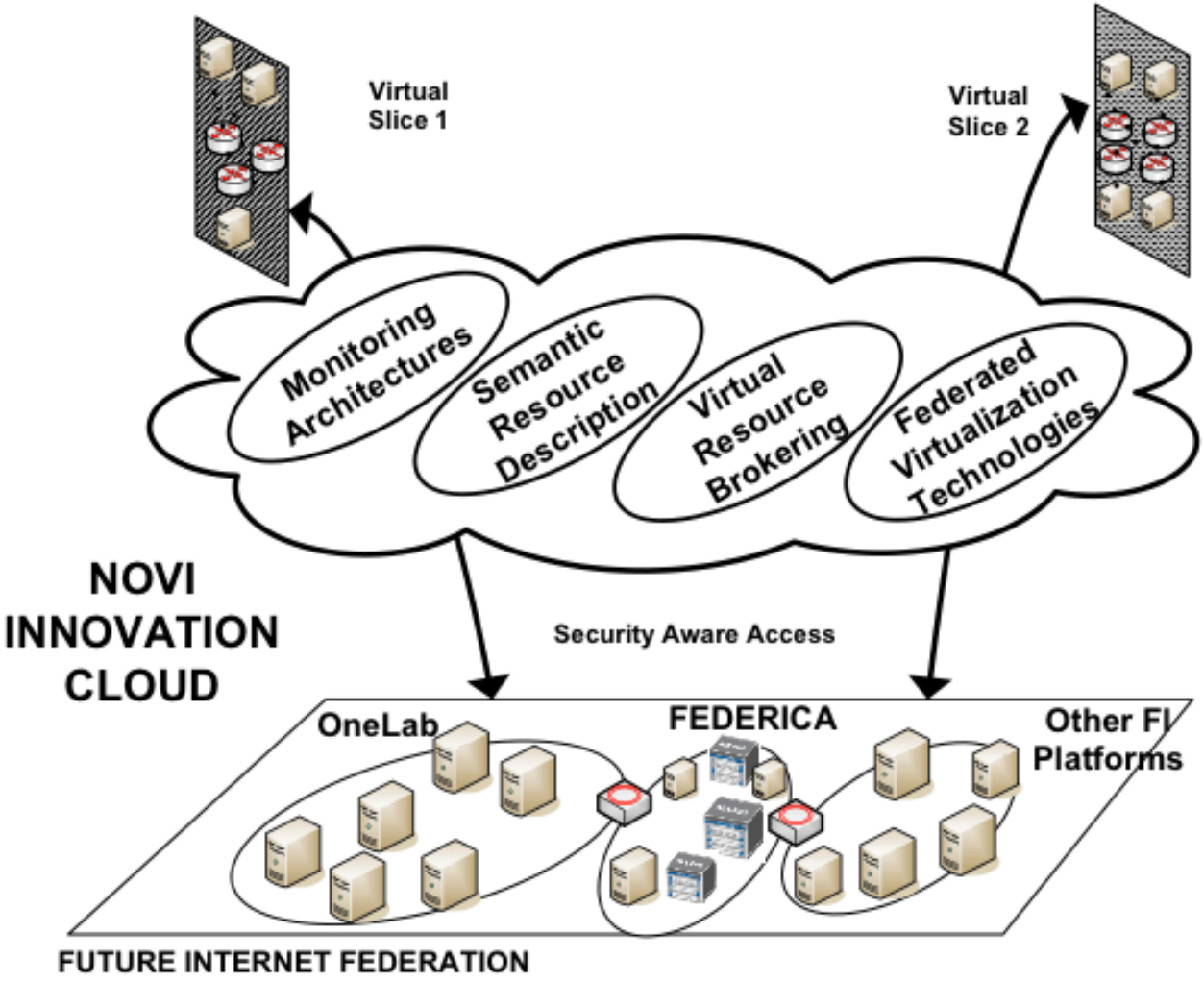
UML representation of CDL

SQWRL is used to query the Ontology.



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

NOVI's mission

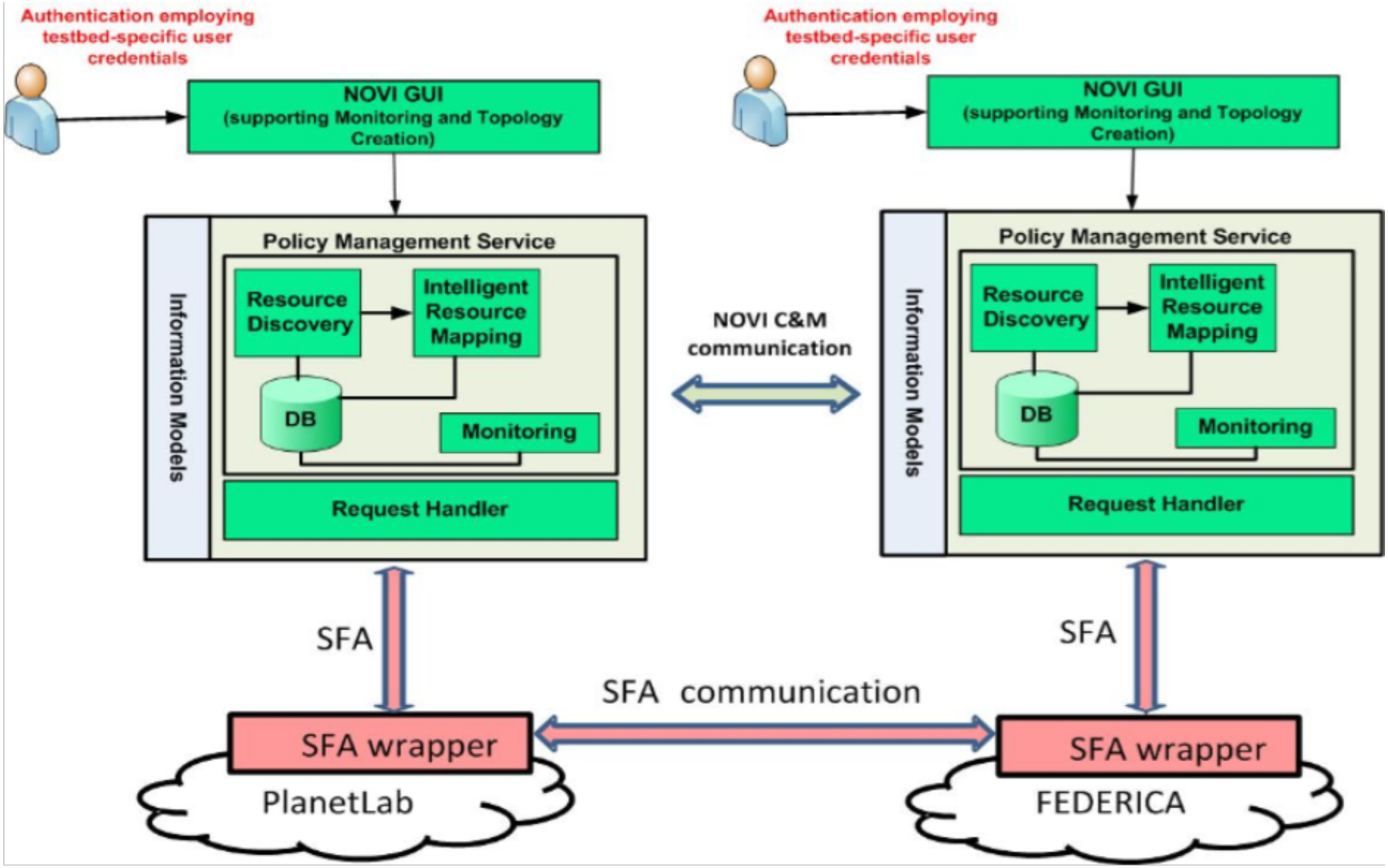


Network Innovation over Virtualized Infrastructures.

- Virtualization
  Virtualization of resources is a main component in these test beds.

- Federation
  Federation of platforms are expected to provide users with richer services.

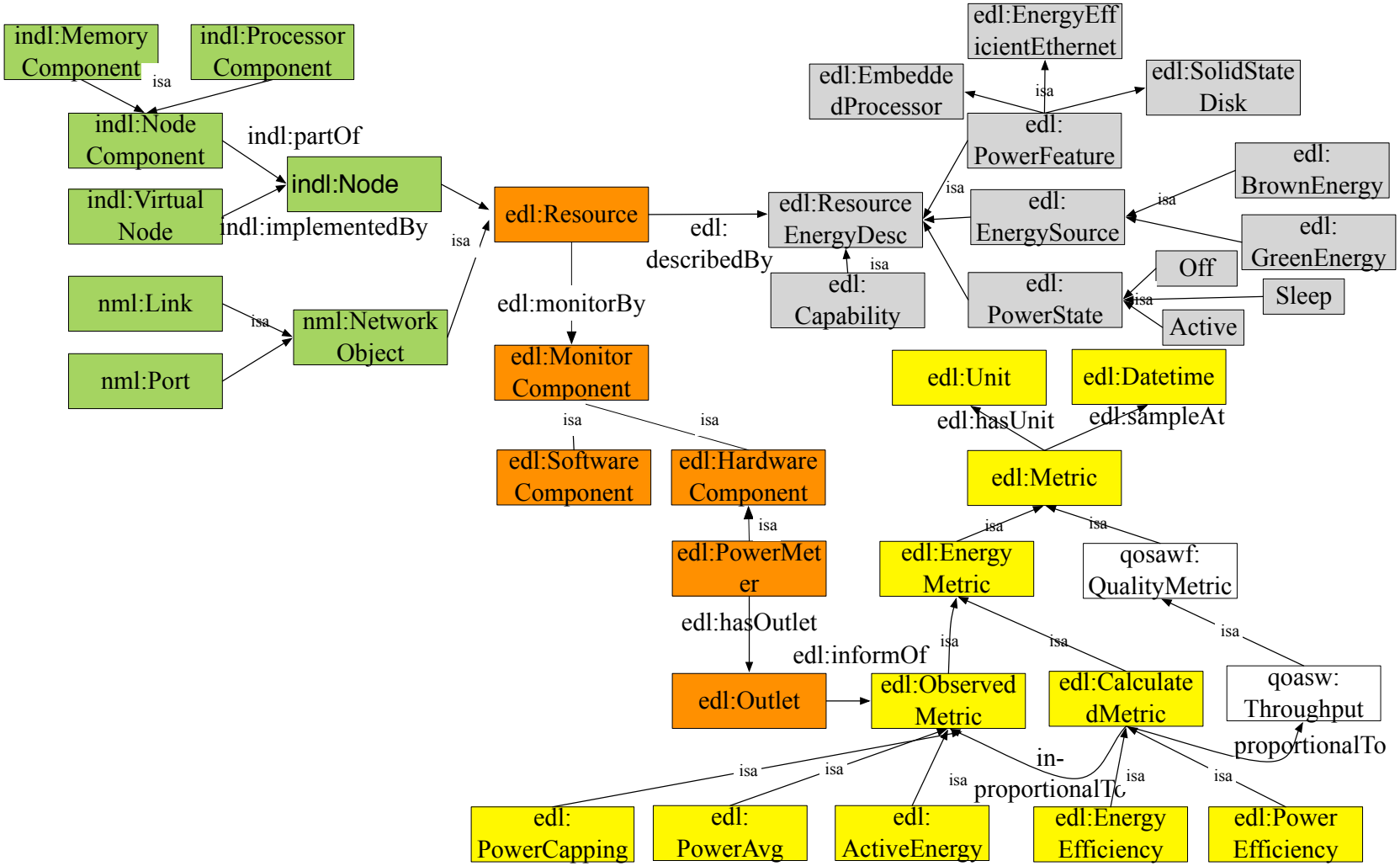
Architecture



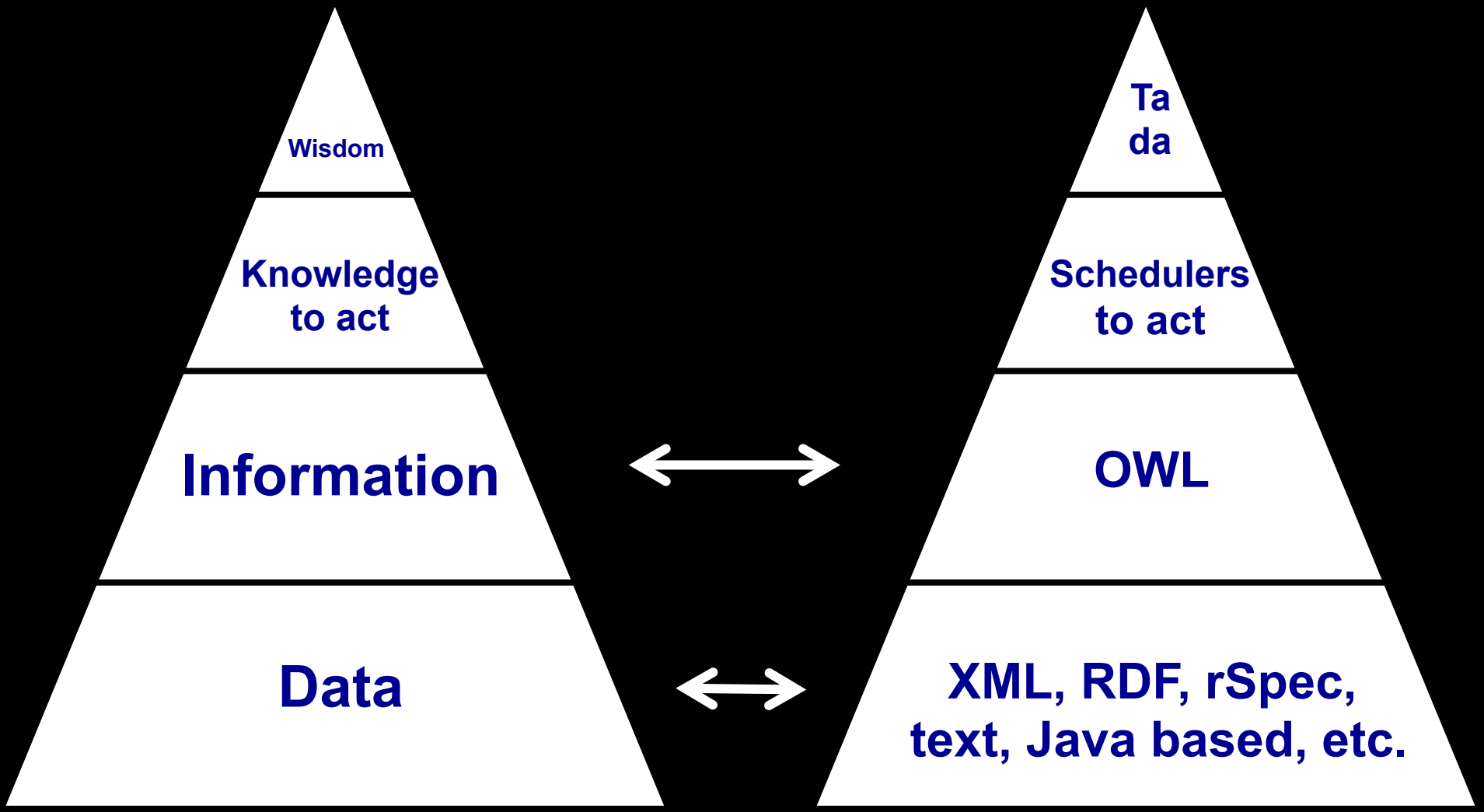
GreenSonar

- Sustainability requires Green-IT in all kinds of resources, compute, network, storage, sensor, …
  - > System wide approach!
- Measure to know; information needed for smart infrastructure.
- Basic idea: been there done that in Networking!
- Why not apply iNDL/NML & PerfSonar methods to provide Green & Energy information?
- Need also application information.
- Big hurdle: energy metrics on heterogeneous resources
- OGF BOF/Charter meeting at OGF36, Chicago.

Energy Description Language – EDL (we have already CdL)

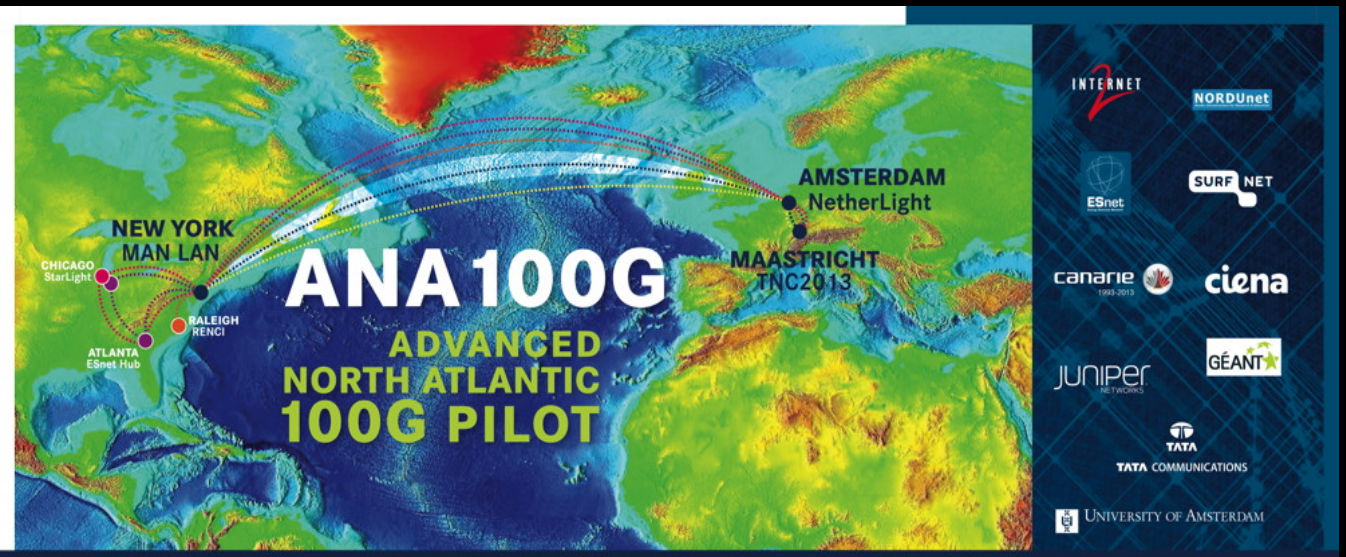


Layers


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

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)+ 2x 10GE (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 ALZS 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

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

Future

- Big Data processing, data centric e-Infrastructure
- PIRE @ UvA (june 2014)
- Research Data Alliance in Amsterdam (september 2014)
- See also:
  - http://ext.delaat.net/
  - http://ext.delaat.net/smartgreen/index.html
  - http://ext.delaat.net/news/2012-03-23/index.html
  - http://wiki.cs.vu.nl/greenclouds/index.php/Main_Page

Trip supported by:
COMMIT/ SURF NET

The constant factor in our field is Change!

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

"Fortran goto", Unix, c, SmallTalk, DECnet, TCP/IP, c++, Internet, WWW, Semantic Web, Photonic networks, Google, grid, cloud, Data^3, App

to:

DDOS attacks destroying Banks and Bitcoins.

Conclusion:

Need for Safe, Smart, Resilient Sustainable Infrastructure.