UvA-SNE activities in Gigaport

StarPlane, NGE, DAS4, NDL

Cees de Laat

Arie Taal

Jeroen van der Ham

Fred Wan

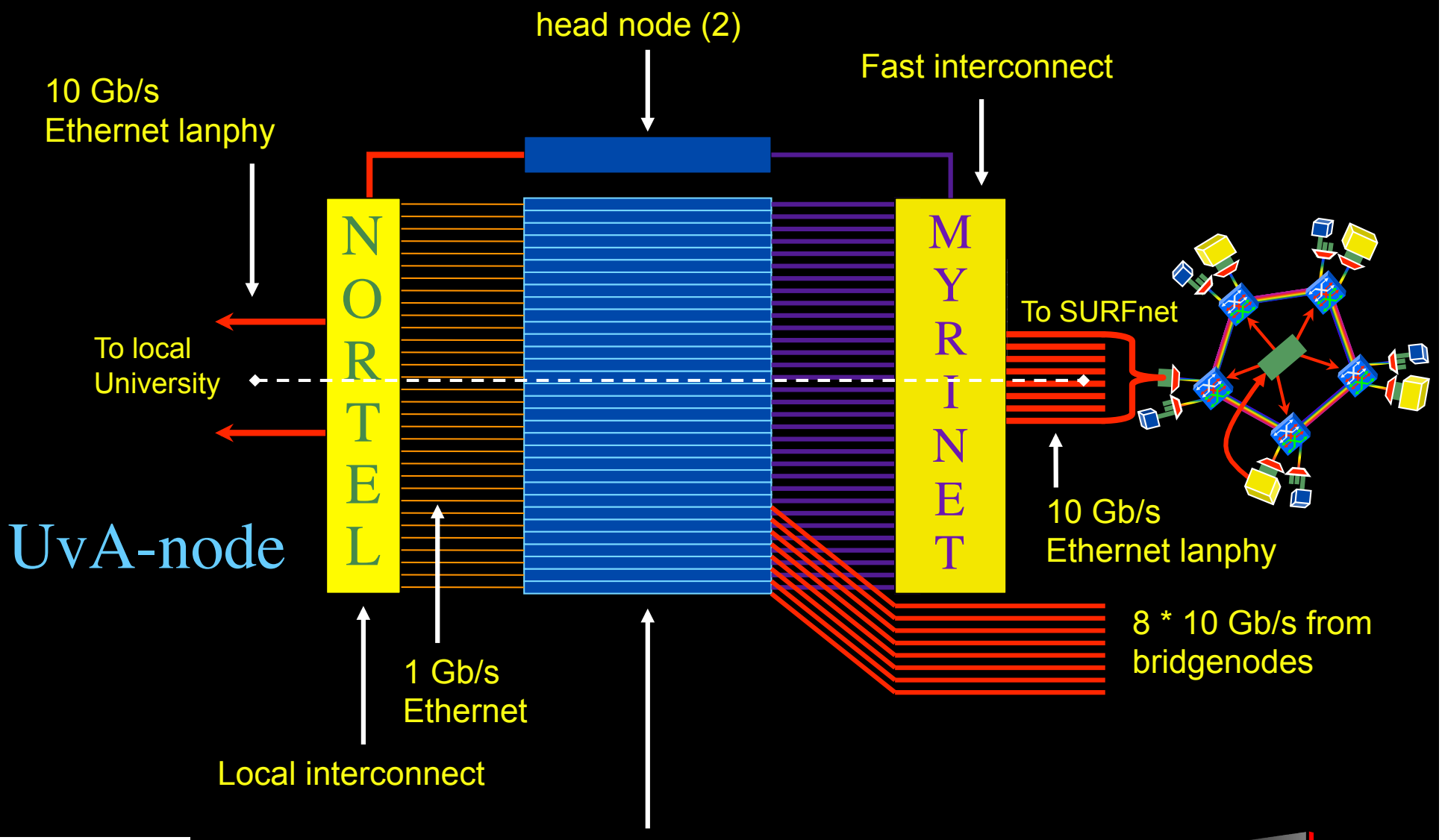
Mihai Christai

Ralph Koning

Paola Grosso

University of Amsterdam

DAS-3 Cluster Architecture



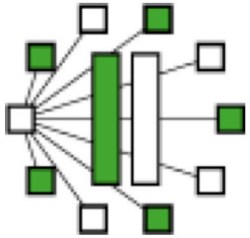
Advanced School for Computing and Imaging

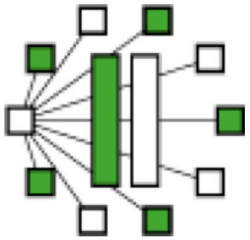
DAS-3

Games and Model Checking

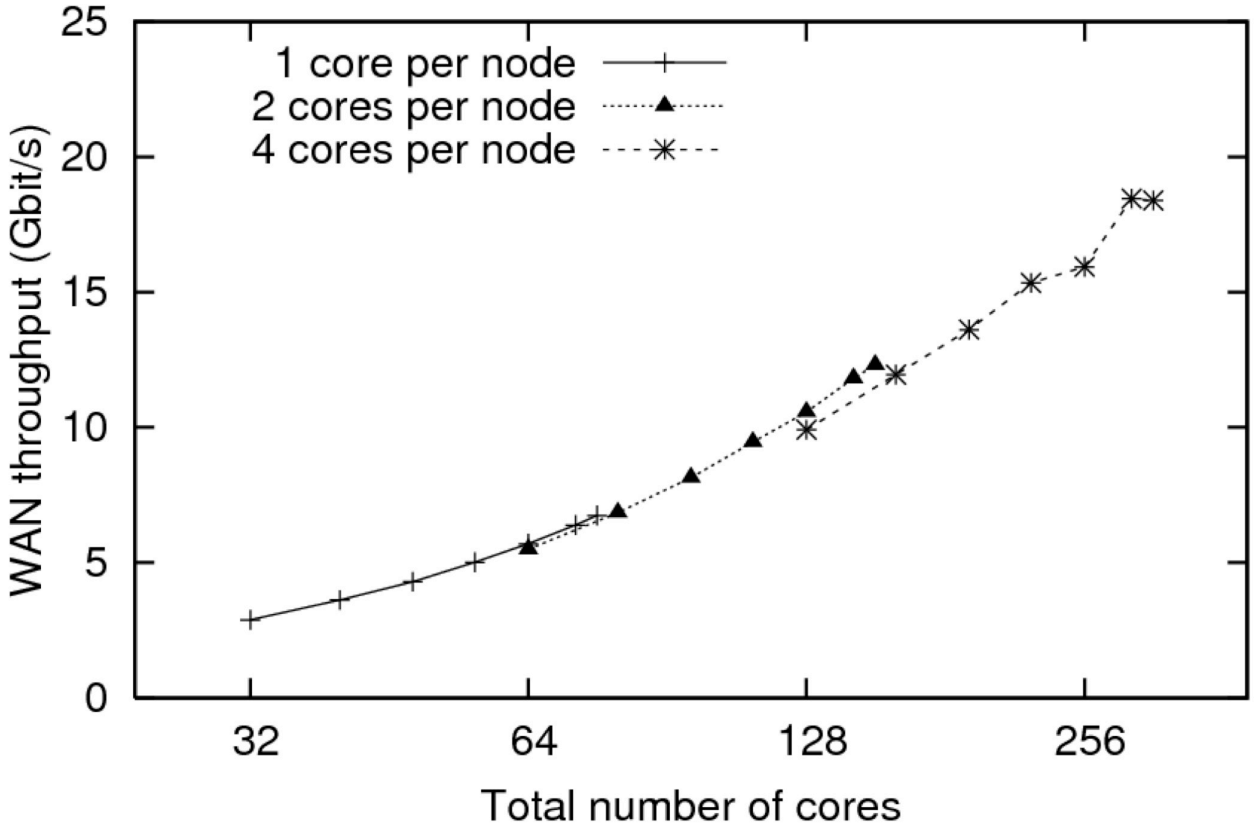
- Can solve entire Awari game on *wide-area* DAS-3 (889 B positions)

  - Needs 10G private optical network [CCGrid'08]
- Distributed model checking has very similar communication pattern
  - Search huge state spaces, random work distribution, bulk asynchronous transfers
- Can efficiently run DeVinE model checker on wide-area DAS-3, use up to 1 TB memory [IPDPS'09]





Required wide-area bandwidth



vrije Universiteit

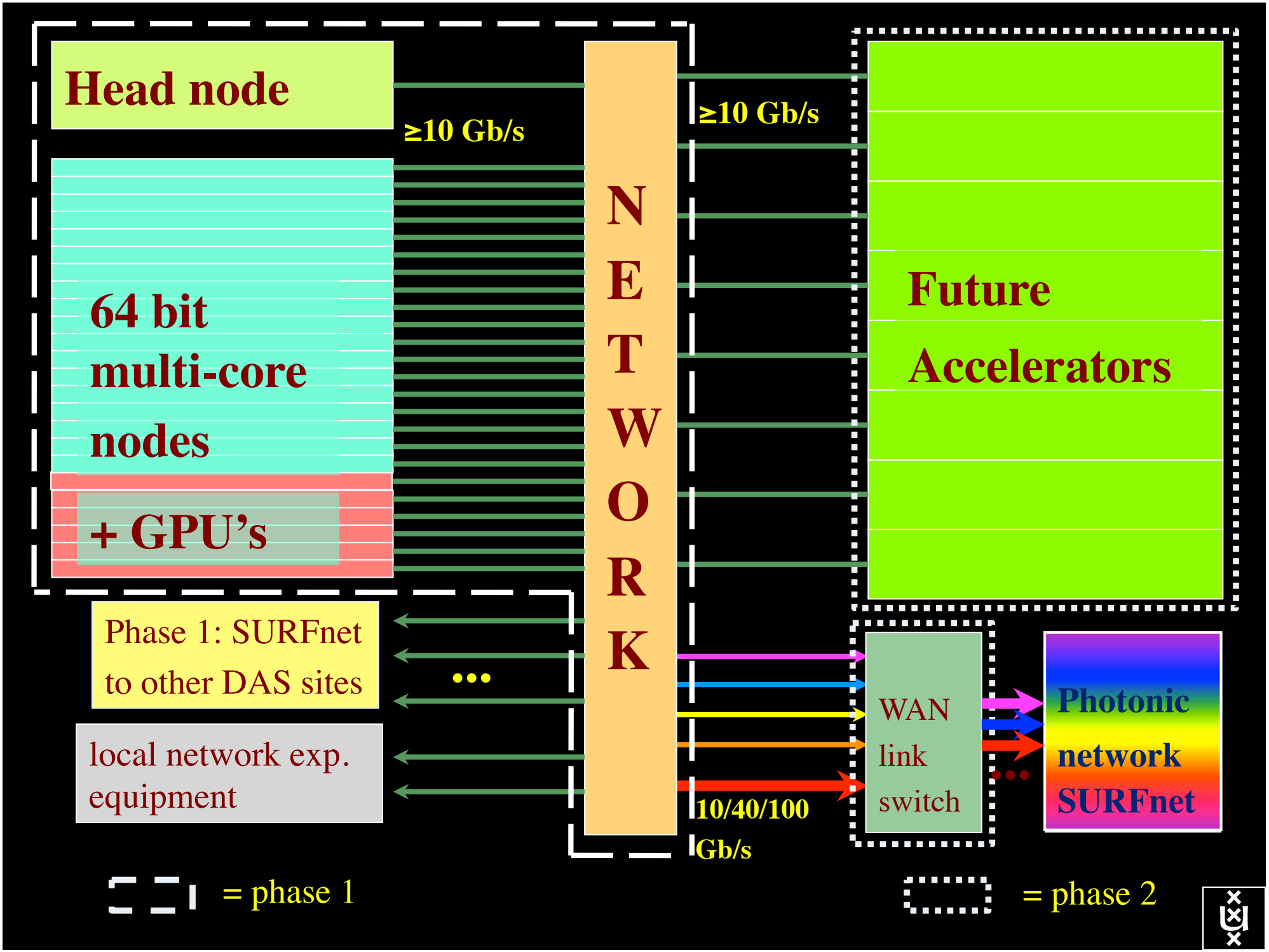
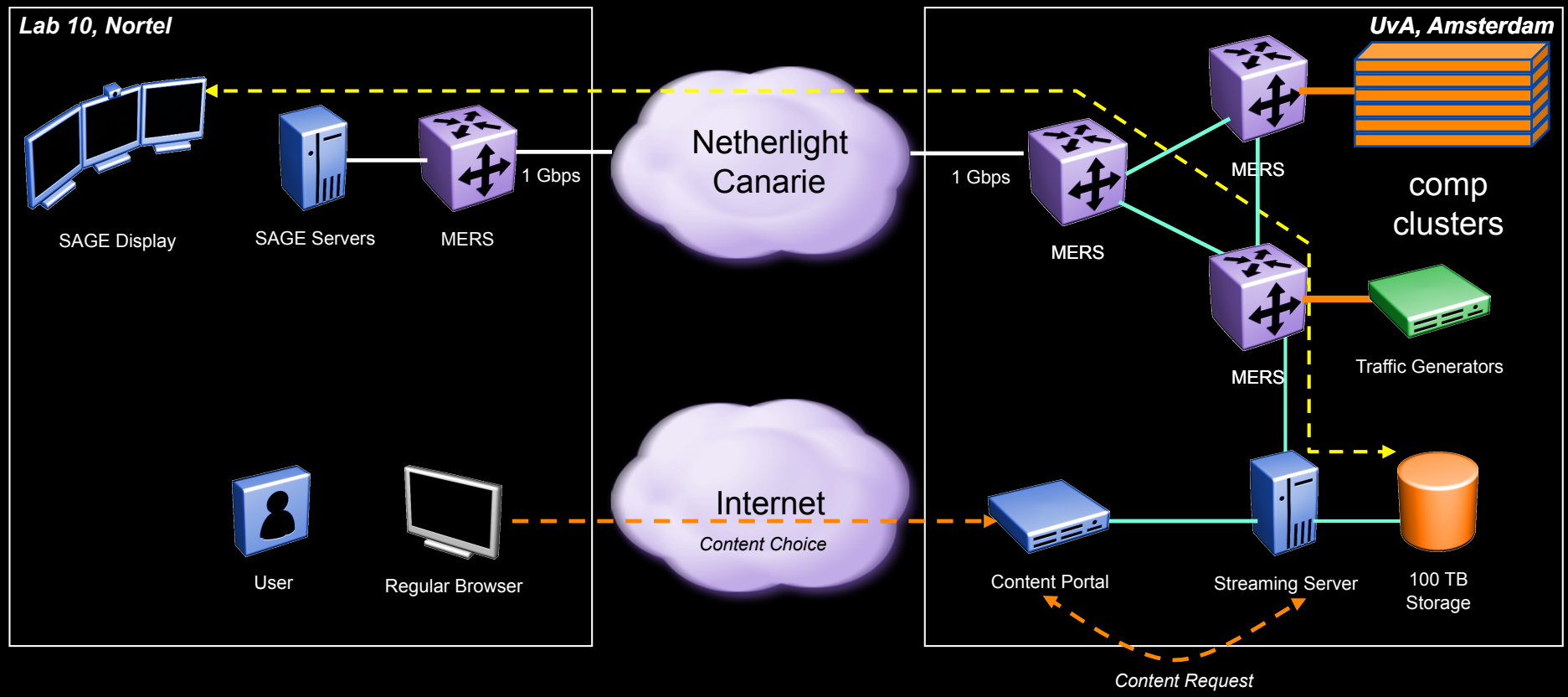
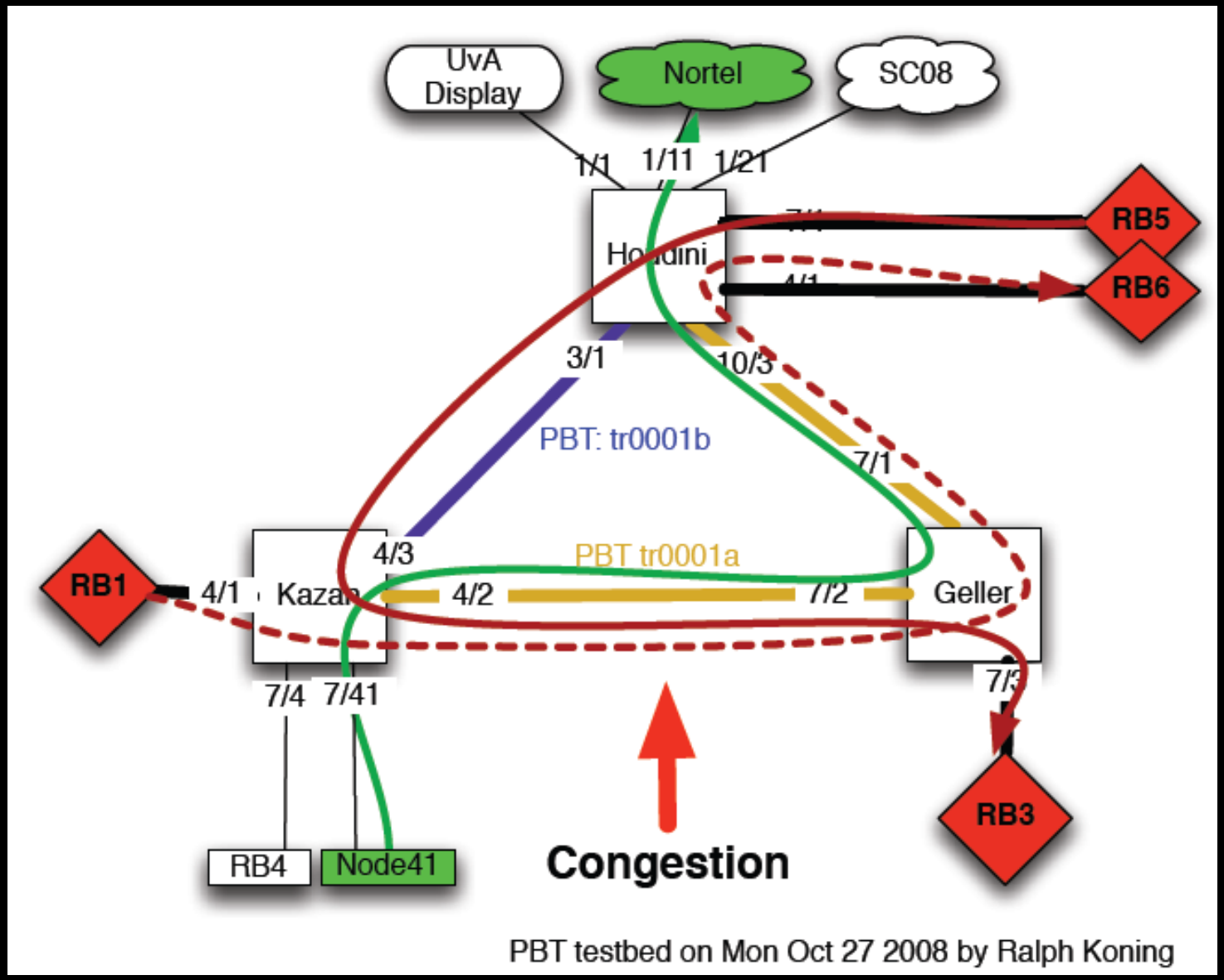

Head node
≥10 Gb/s
64 bit
multi-core
nodes
+ GPU's
N
E
T
W
O
R
K
≥10 Gb/s
Future
Accelerators
Phase 1: SURFnet
to other DAS sites
local network exp.
equipment
WAN
link
switch
10/40/100
Gb/s
Photonic
network
SURFnet
= phase 1
= phase 2

Diagram for SAGE video streaming to ATS



UvA Testbed

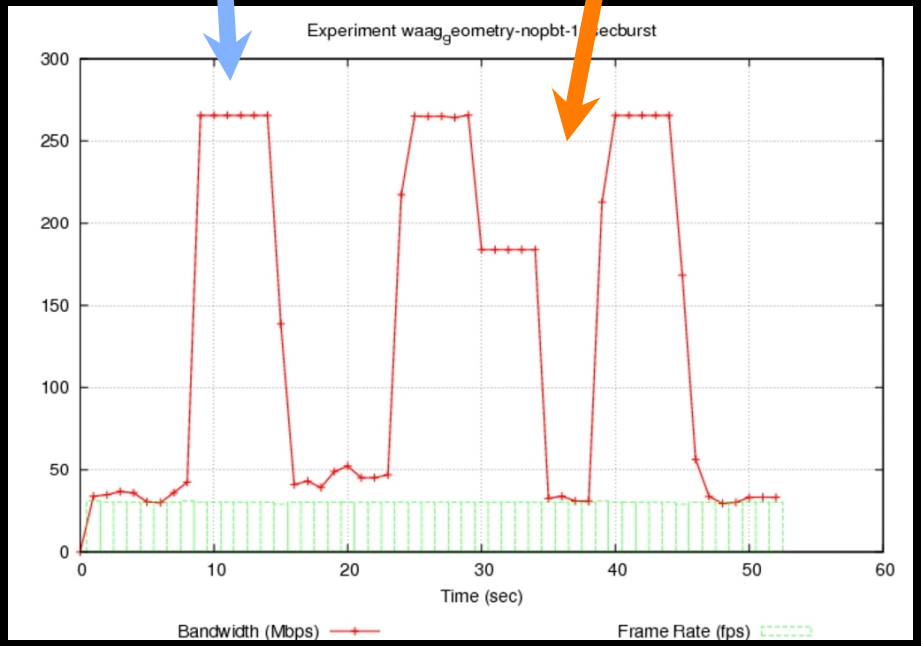


Congestion introduced in the network with multiple PBT paths carrying streamed SHD Content

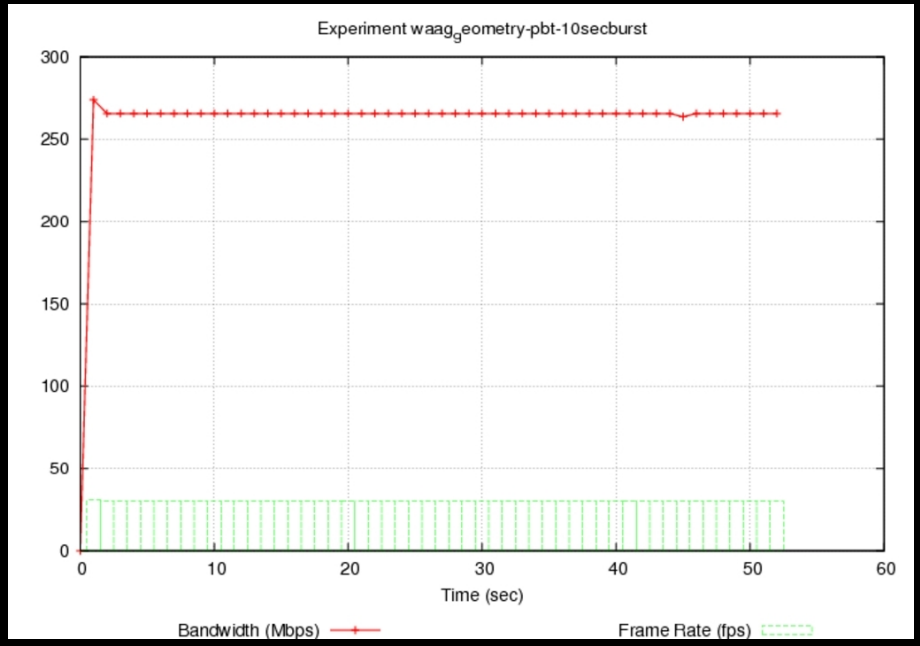
Experimental Data

Sage without background traffic

Sage with background traffic



10 Second Traffic bursts with No PBT

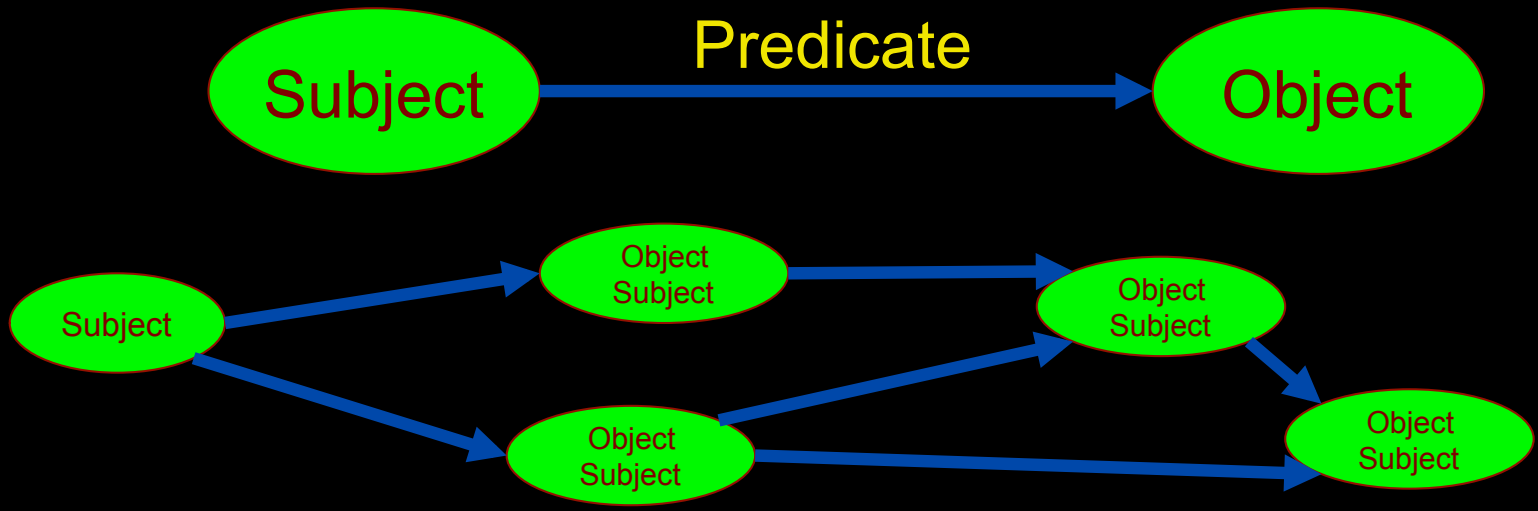


10 Second Traffic bursts with PBT

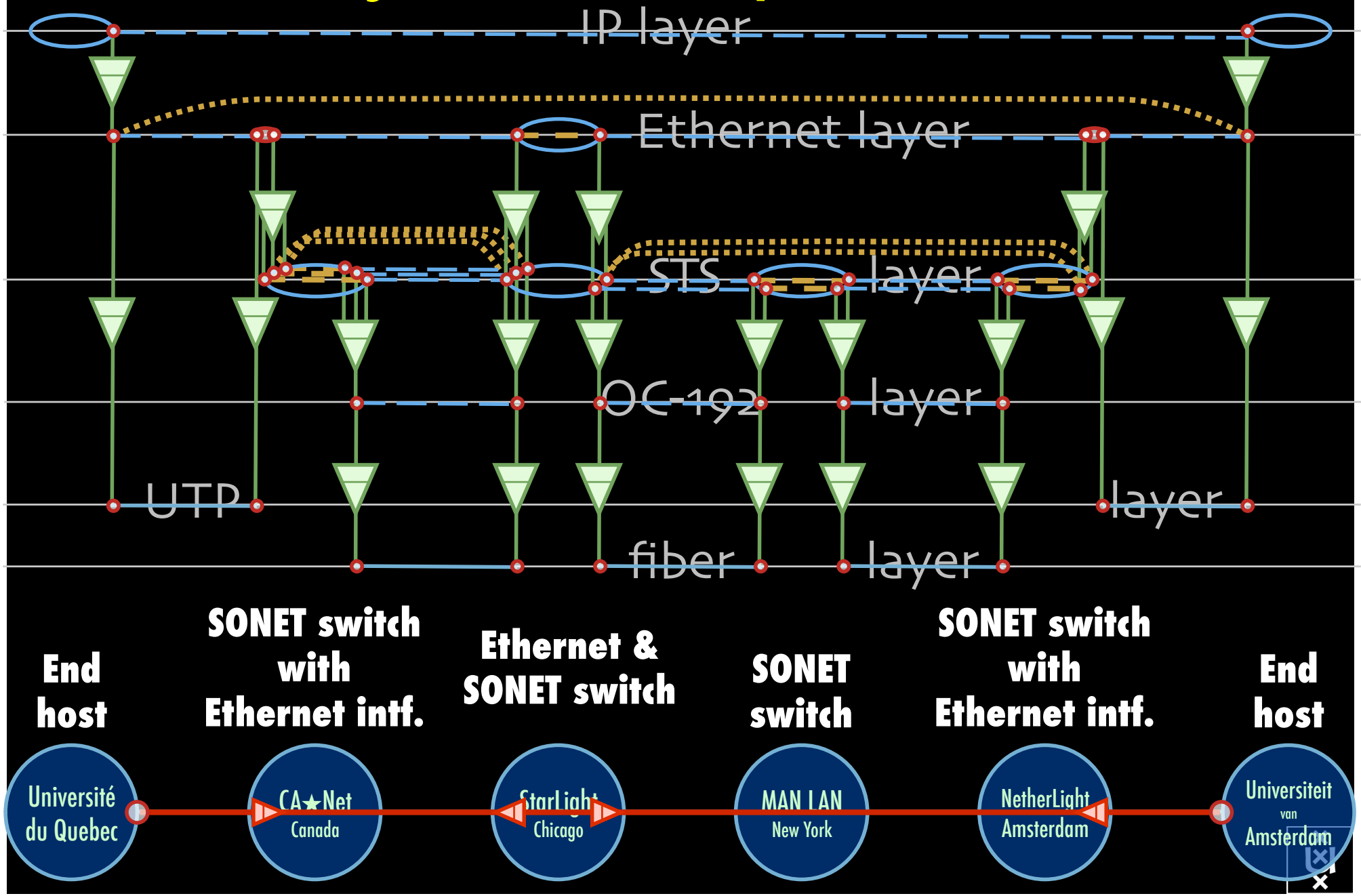
PBT is *SIMPLE* and *EFFECTIVE* technology to build a shared Media-Ready Network

Network Description Language

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



Multi-layer descriptions in NDL



Activities

- RDF Infrastructure
- Integration of NDL and Fenius
- Cooperation with ESnet on OSCARS
- Standardization

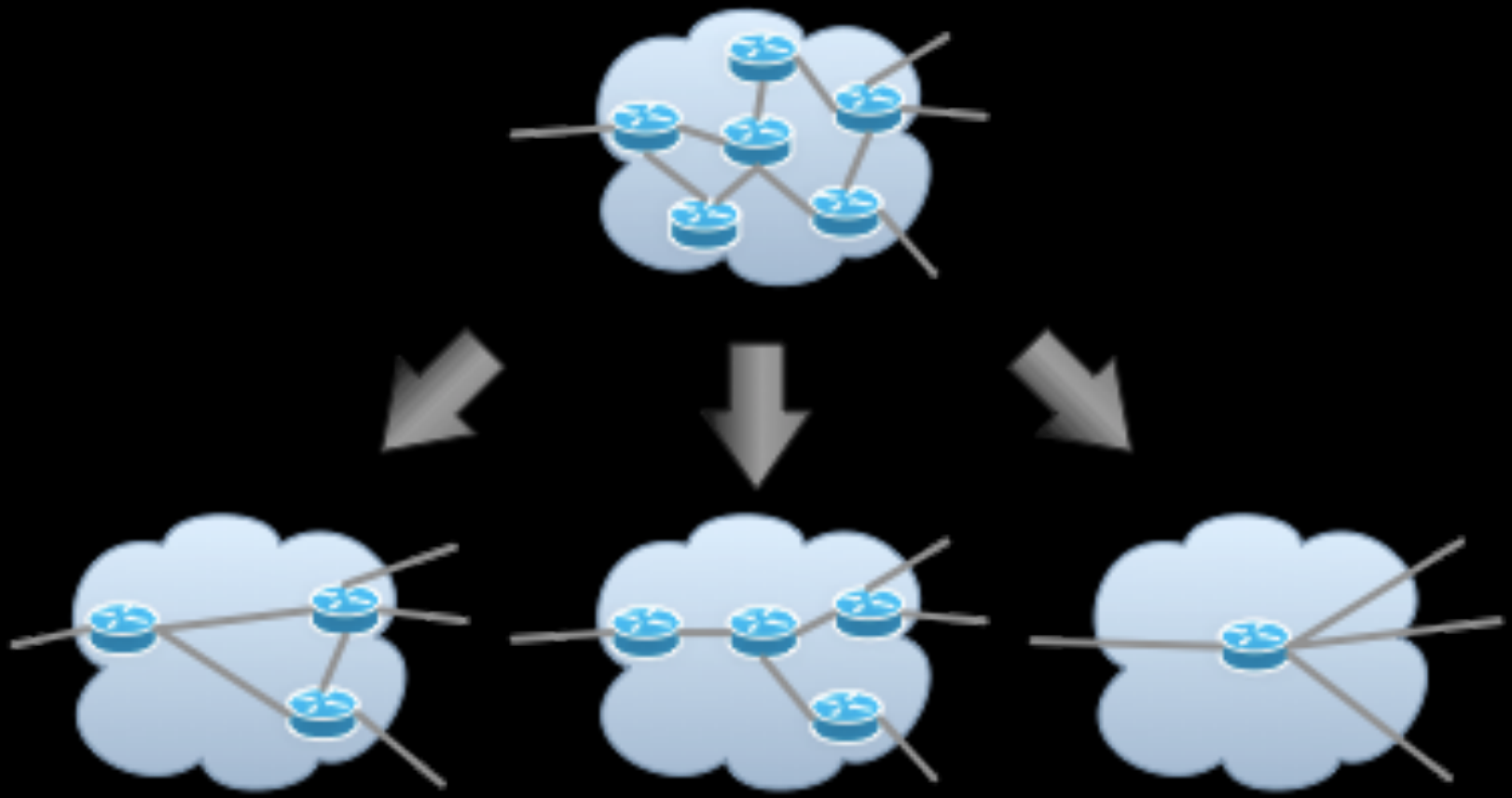
NDL and Fenius

- Fenius
  - is about inter-domain provisioning
  - does not do topology exchange
- New GLIF Task-Force
  - Distributed Topology Exchange (DTOX)

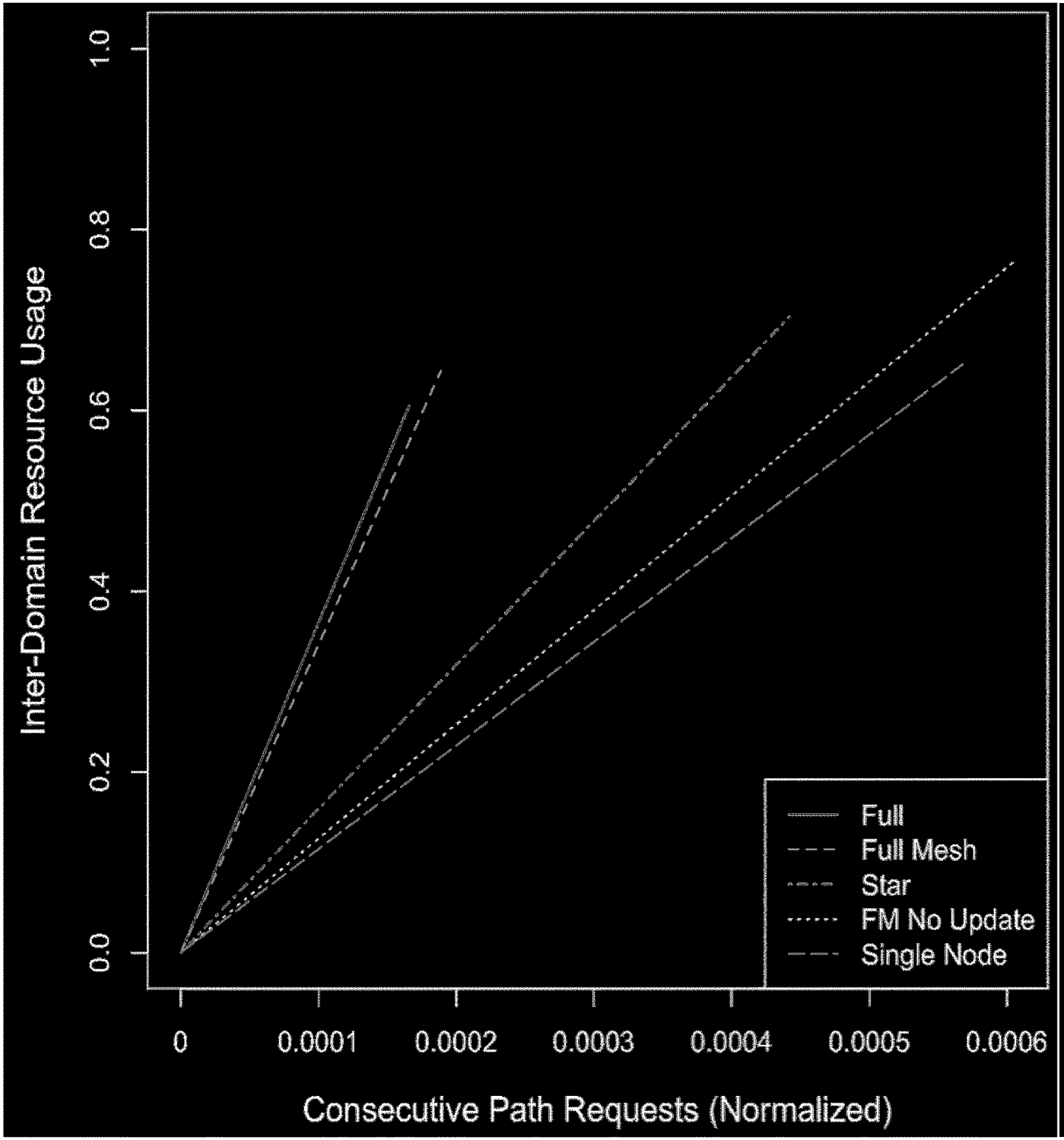
Standardization

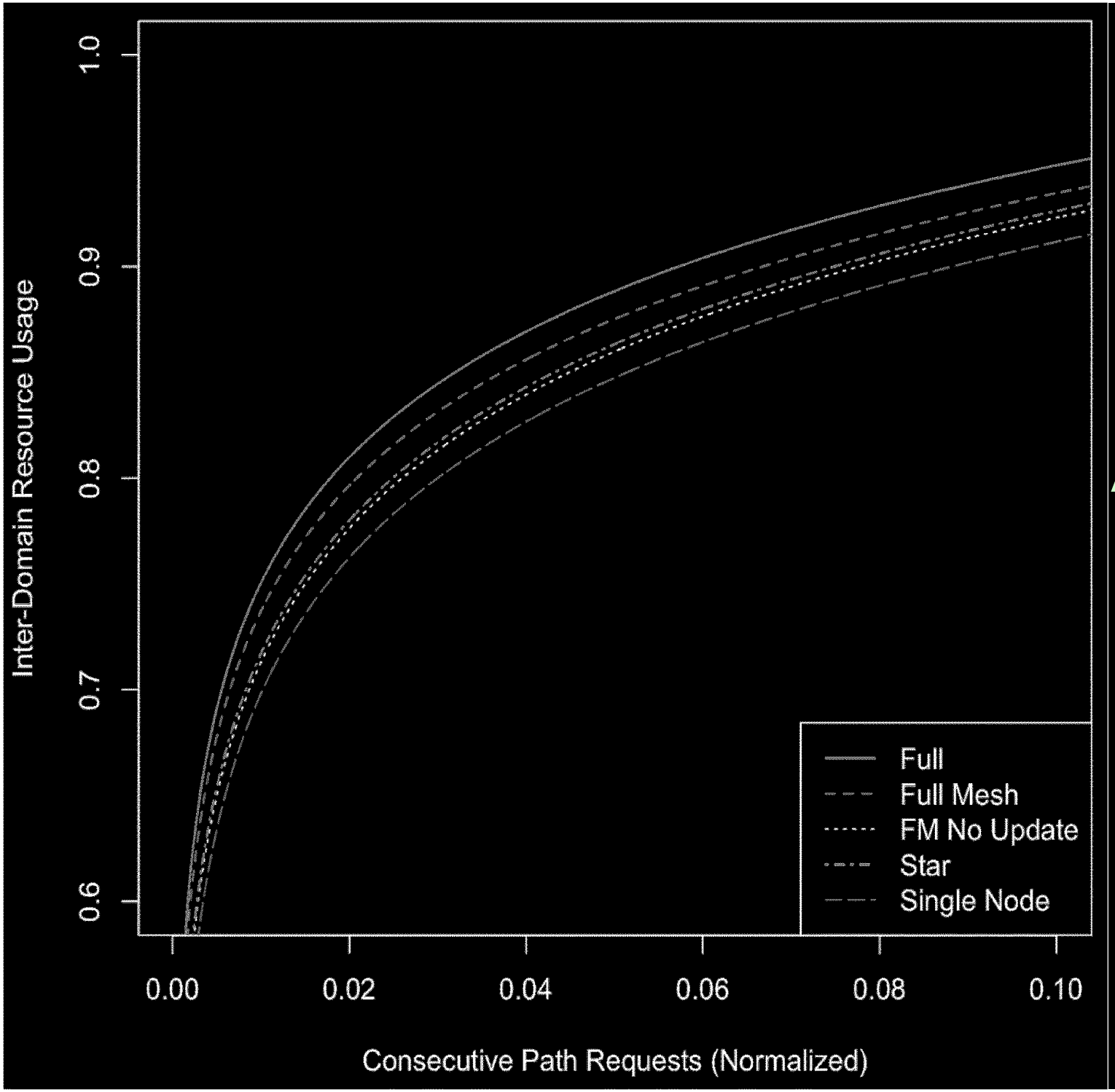
- NML is slowly progressing
  - Schema Document
- NSI is working frantically
  - Terminology Glossary
  - Architecture Document
  - NSI Protocol Document

Topology Aggregation



Topology Aggregation - Initial



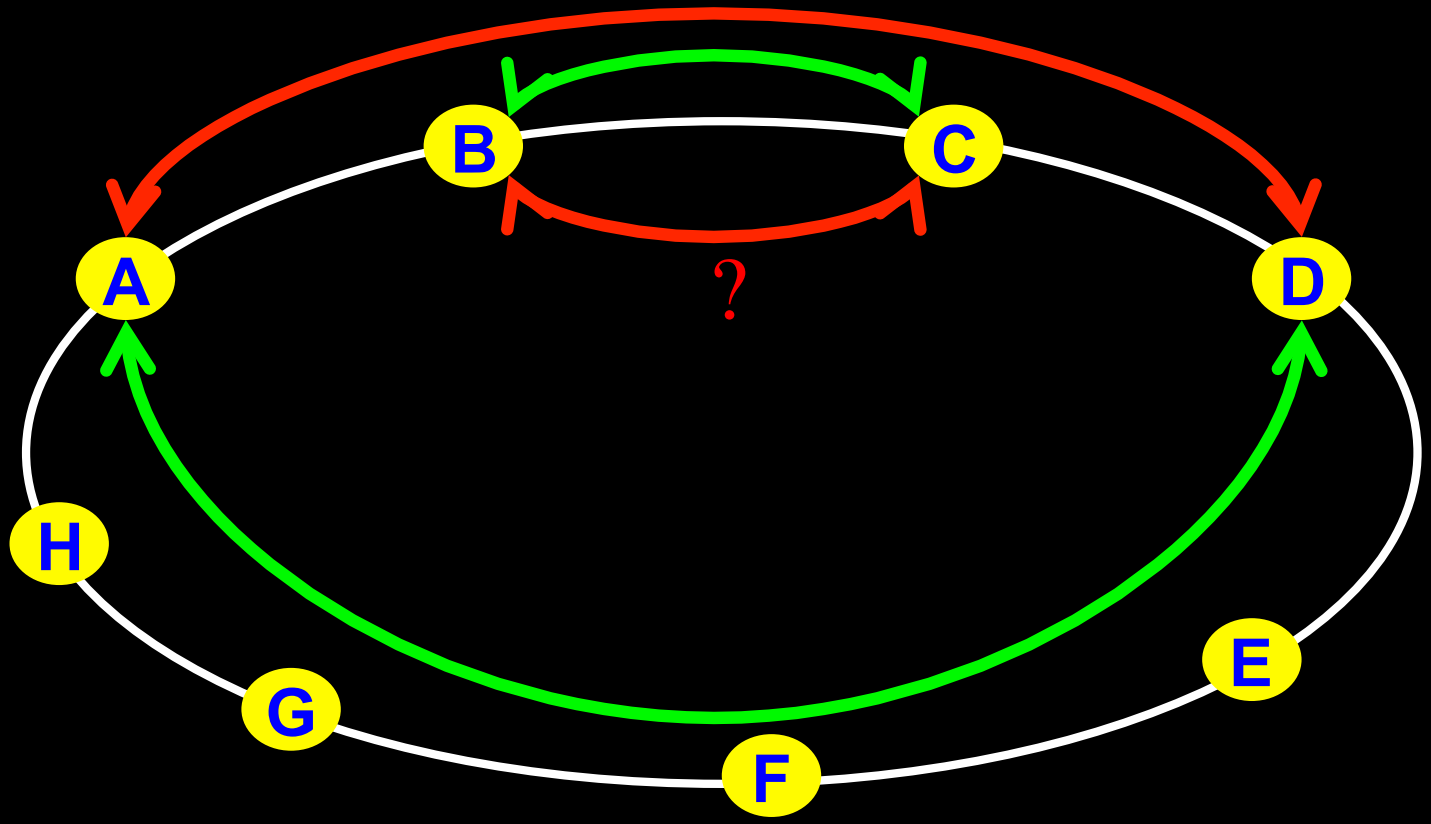


Topology Aggregation - Saturation

The Problem

I want AD and BC

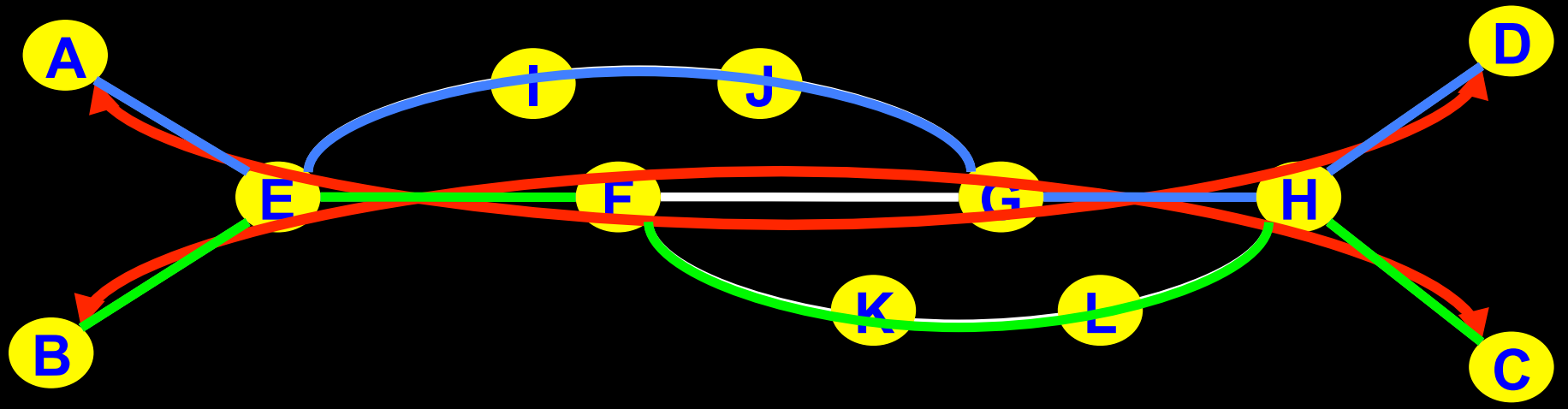
Success depends on the order



Another one ☺

I want AD and BC

Success does not even depend on the order!!!

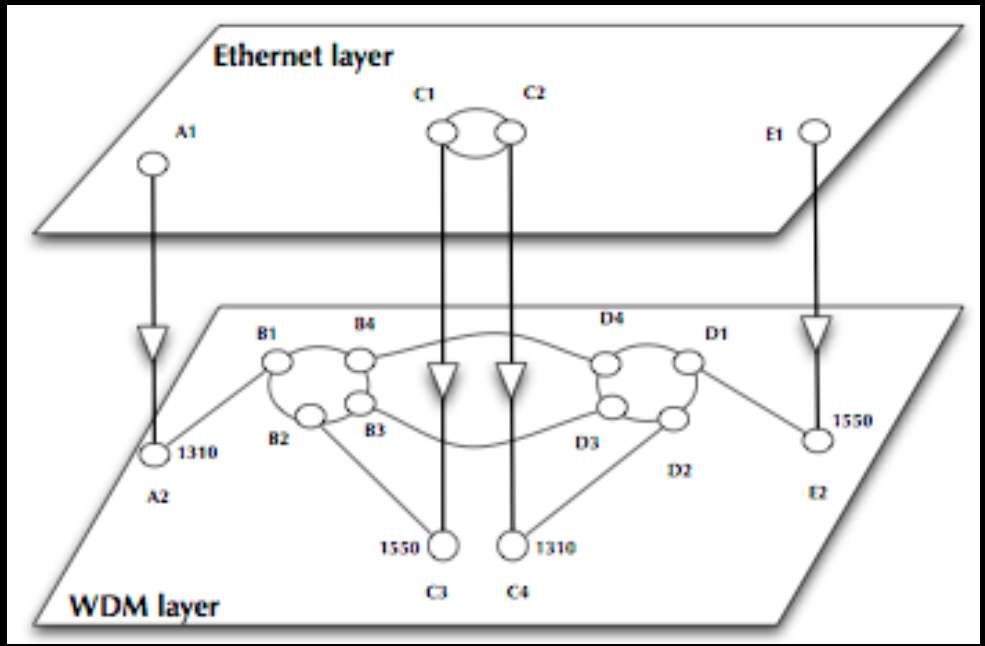


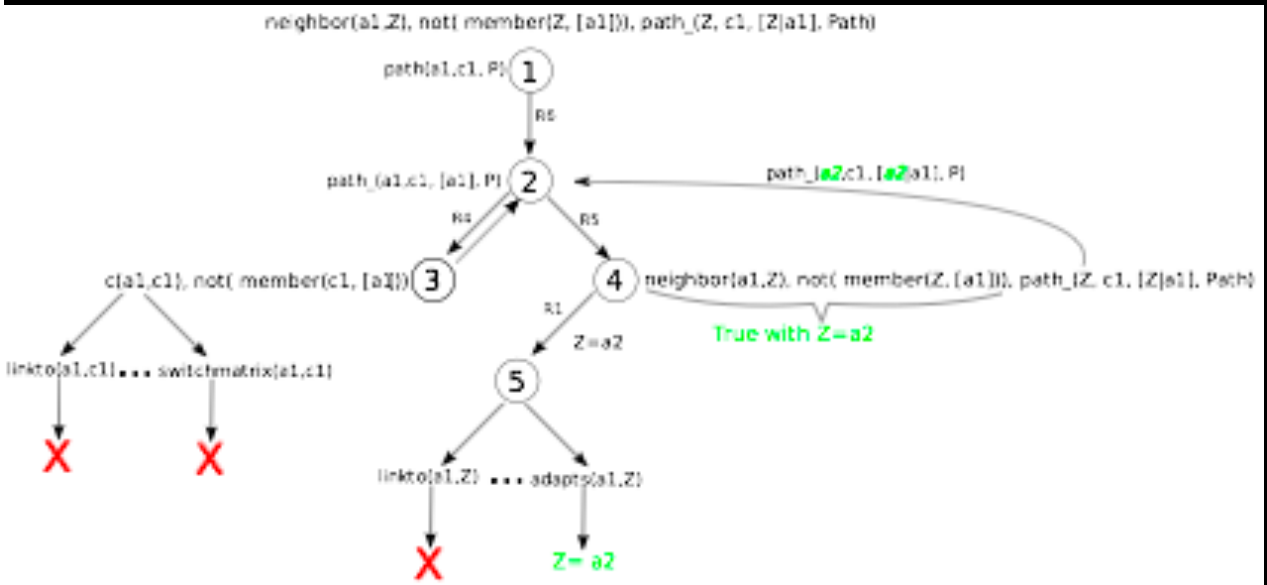
Wouldn’t it be nice if I could request [AD, BC, ...]

NDL + PROLOG

Research Questions:

- order of requests
- complex requests
- usable leftovers





- Reason about graphs
- Find sub-graphs that comply with rules

Multi-domain 2-layer networks

How do multi-domain 2-layer networks look like?

Guess: Projection algorithm (2-layer: Ethernet /WDM)

Steps:

1. Generate a multi-domain graph by BA-algorithm
2. Generate a graph for each domain by BA-algorithm
3. For each domain graph project random nodes onto WDM layer
4. Connect the domains at each layer according to the multi-domain graph
5. Assign random wavelengths to the adaptation links

Advantage:

- Number of adaptations determined by the degree of the projected nodes
- Multi-domain Ethernet-layer as well as the multi-domain WDM-layer graph are not necessarily connected.

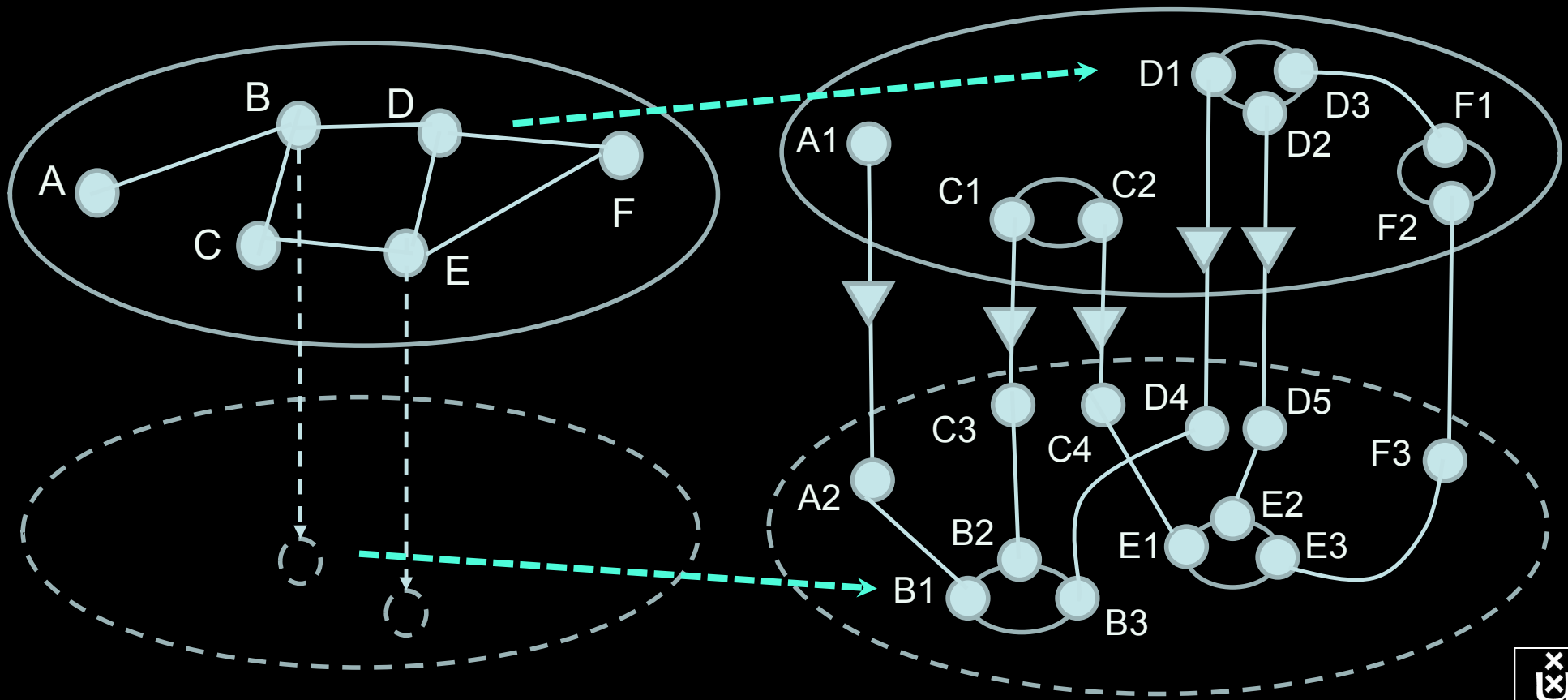
Input parameters:

- Number domains, number of nodes(devices) per domain
- Ratio of Ethernet-devices over WDM-devices per domain
- Distribution of wavelength

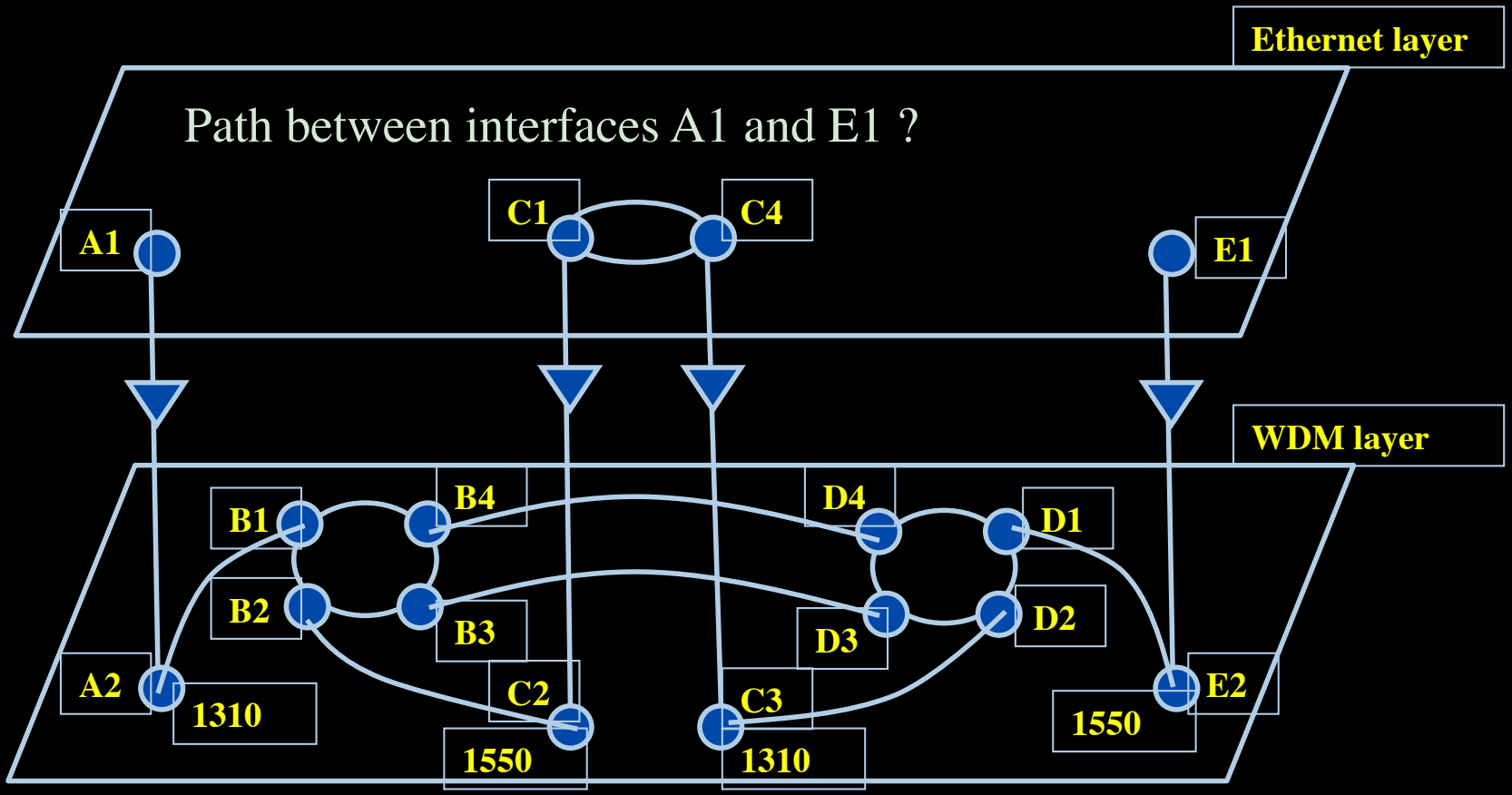
Multi-domain 2-layer networks

Projection algorithm

BA-algorithm to generate a graph for each domain
Project random nodes onto WDM layer



Multi-layer Network PathFinding



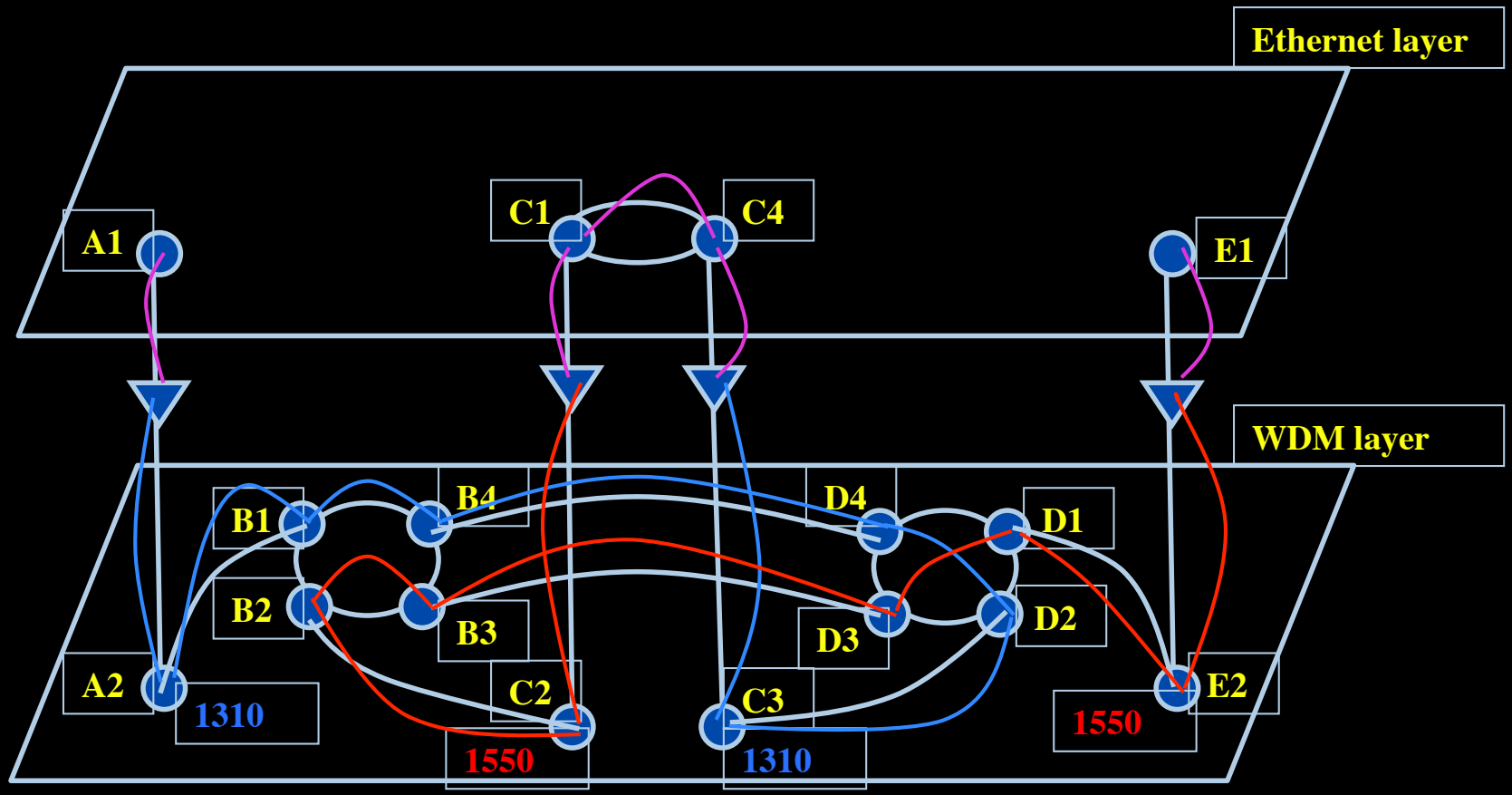
Prolog rule:

```
linkedto( Intf1, Intf2, CurrWav ):-                    %-- is there a link between  Intf1 and Intf2 for wavelength CurrWav ?
    rdf_db:rdf( Intf1, ndl:'layer', Layer ),           %-- get layer of interface Intf1 → Layer
    Layer == 'wdm#LambdaNetworkElement',               %-- are we at the WDM-layer ?
    rdf_db:rdf( Intf1, ndl:'linkedTo', Intf2 ),        %-- is Intf1 linked to Intf2 in the RDF file?
    rdf_db:rdf( Intf2, wdm:'wavelength', W2 ),         %-- get wavelength of Intf2 → W2
    compatible_wavelengths( CurrWav, W2 ).             %-- is CurrWav compatible with W2 ?
```

linkedto(B4, D4, CurrWav) is true for any value of CurrWav

linkedto(D2, C3, CurrWav) is true if CurrWav == 1310

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

Prolog pathfinding results

DFS path constraints: Number of different wavelength

#Domains (#Ether:#WDM) (<#Intf>)(<#Adap>)	Prolog time [ms] μ(σ)	Timeouts	Success %	
3 (9:6)(55)(11)	20(4)	0	100	No max #wav
4 (48:32)(377)(73)	2620(8245)	74	92.6	
4 (96:64)(771)(147)	6592(11802)	207	79.3	
3 (9:6)(55)(11)	20(4)	0	100	#wav ≤3
4 (48:32)(377)(73)	1303(5052)	22	97.8	
4 (96:64)(771)(147)	3910(10045)	51	94.9	
3 (9:6)(55)(11)	20(4)	0	100	#wav ≤2
4 (48:32)(377)(73)	755(3210)	8	98.9	
4 (96:64)(771)(147)	3240(9052)	38	96.1	

Prolog pathfinding results

Parallel calls: A->B and B->A

Projection: A->B

#Domains (#Ether:#WDM) (<#Intf>)(<#Adap>)	Prolog time [ms] μ(σ)	Timeouts	Success %
3 (9:6)(55)(11)	20(4)	0	100
4 (48:32)(377)(73)	755(3210)	8	98.9*
4 (96:64)(771)(147)	3240(9052)	38	96.1*

#wav ≤2

Projection: first of A->B and B->A

3 (9:6)(55)(11)	19(1)	0	100
4 (48:32)(377)(73)	144(486)	0	100
4 (96:64)(771)(147)	601(2722)	2	99.6*

#wav ≤2

*false negatives also taken into account

How random are Projection-algorithm networks?

Input parameters:

1) Number of domains, number of devices per domain
2) Ratio of Ethernet-devices over WDM-devices per domain
3) Distribution of wavelengths

Data set:

50 graphs for each topology
20 random <src,dst> pairs for each graph

Input parameter:

(#dom, #nodes/dom): (3,5), (3,10), (3,20), (3,40), (4,5), (4,10), (4,20), (4,40)
#nodes/dom, ratio): (5, 3:2), (10, 6:4) , (20, 12:8), (40, 24:16)
Draw random wavelenghts and assign to random group of 4 adaptations

Conclusions

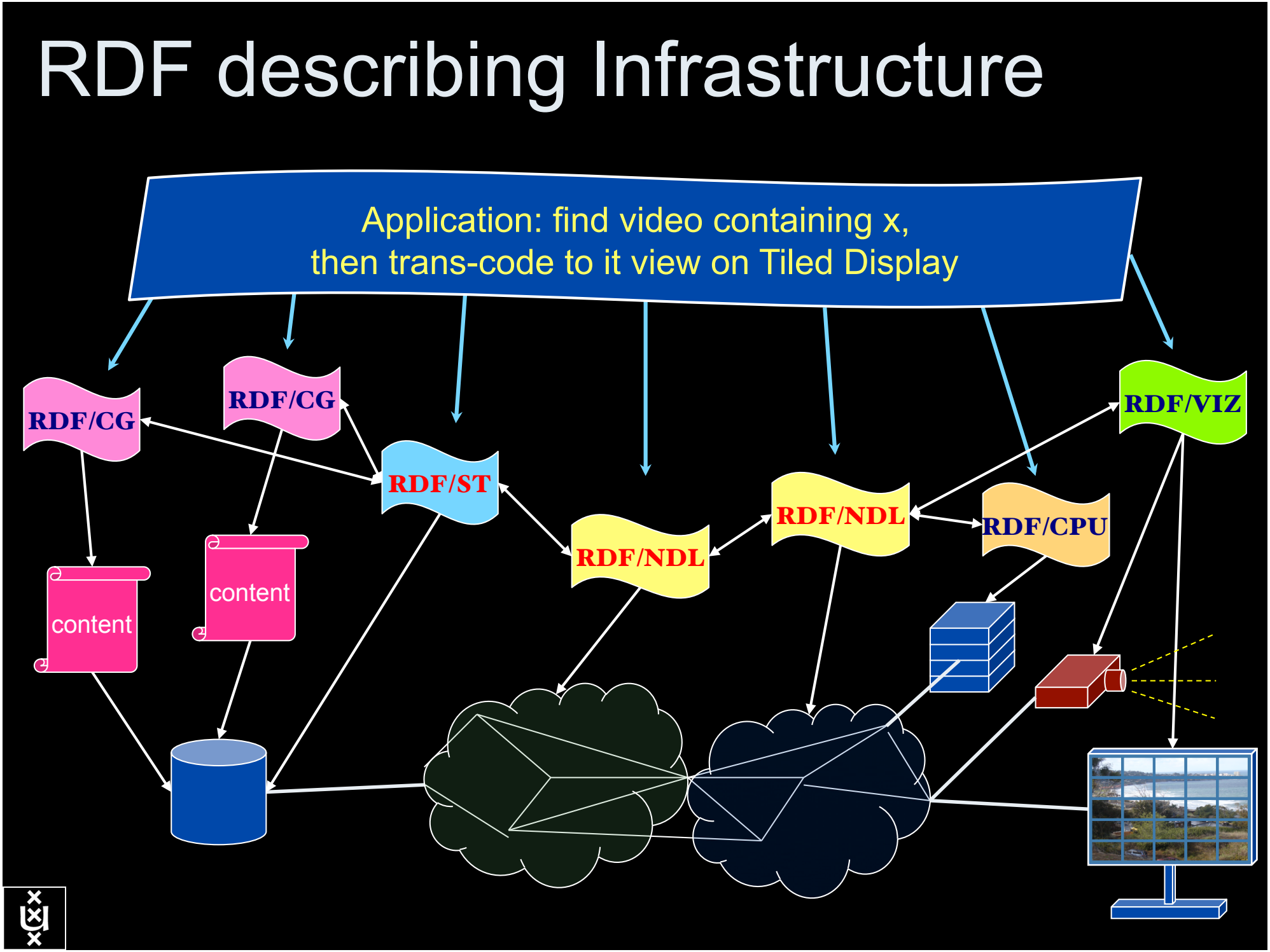
We expect that Projection algorithm generates networks that are more random than real networks if input parameters are not properly chosen.

Nevertheless, with suitable constraints and parallel calls a performance > 99.6 % could be achieved by a Prolog DFS path finding algorithm for graphs up to 800 interface nodes.

Study of real networks is needed to determine:

- The ratio of devices among layers
- The distribution of wavelengths
- Suitable constraints

RDF describing Infrastructure



Applications and Networks become aware of each other!

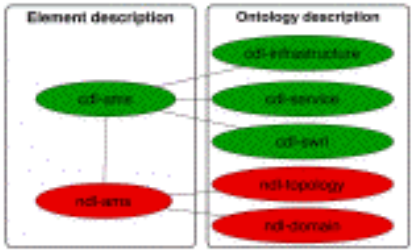
CineGrid Description Language

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

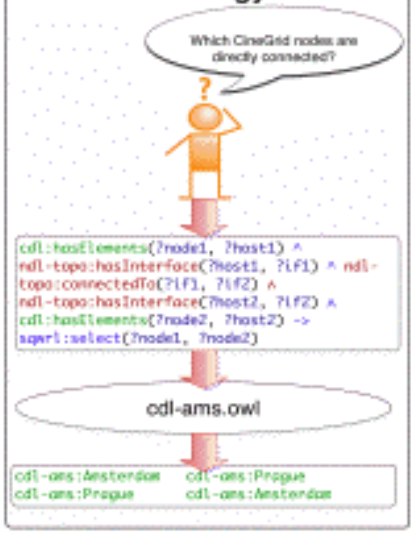
The CineGrid Description Language (CDL) describes CineGrid resources. Streaming, display and storage components are organized in a hierarchical way.

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

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



SQWRL is used to query the Ontology.



UML representation of CDL

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.


Video
Service
UK Scho

Questions ?