Beyond Hybrid Networking

Cees de Laat

SURFnet

EU

BSIK

NWO

University of Amsterdam

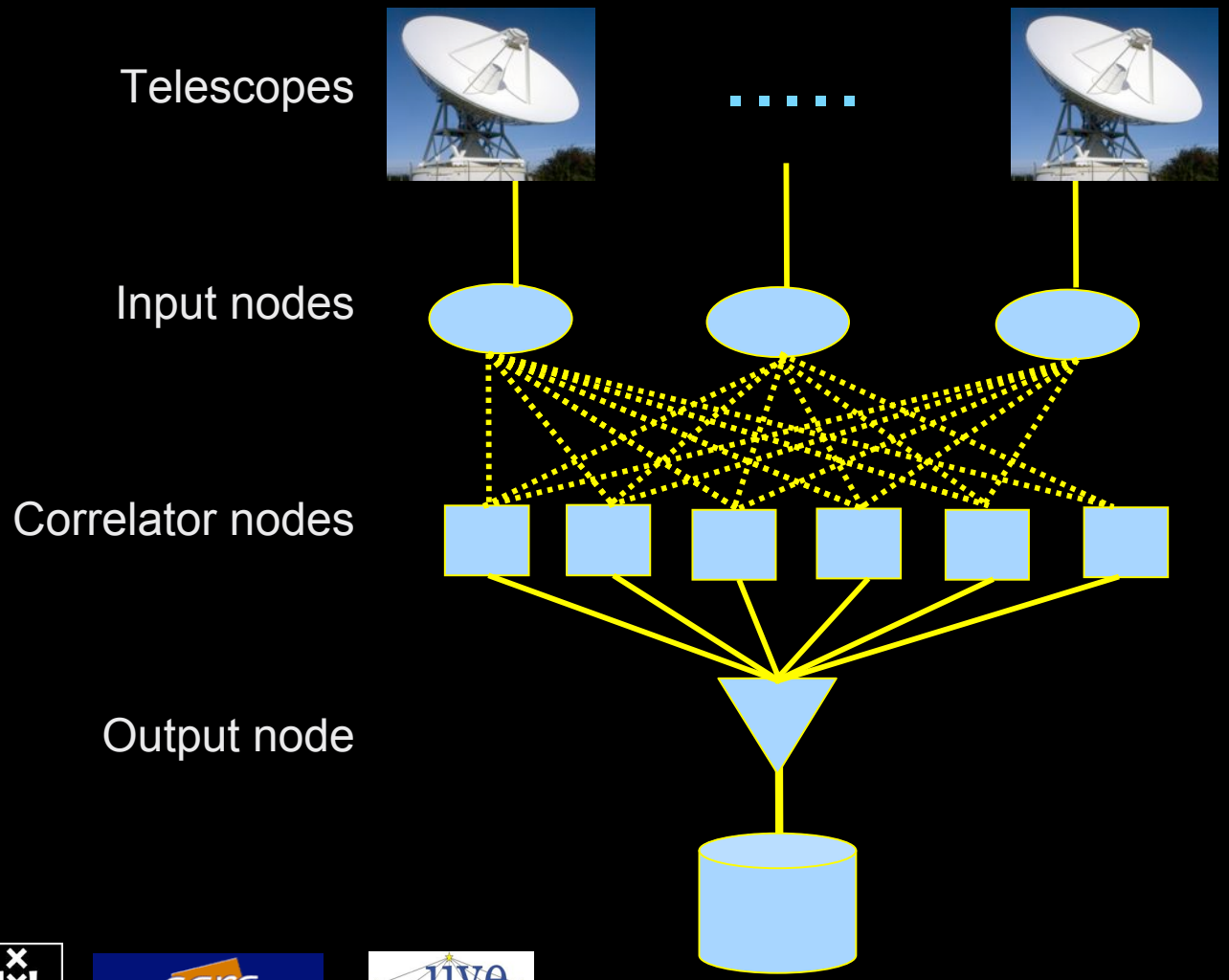
TNO
NCF

Trends

- We have made baby-steps on the path to optical networking
  - Still many mails and phone calls
- See several trends:
  - lambda's get fatter and cheaper
  - photonic technology cheap per bandwidth
  - embedded computation capacity increasing
  - latency and high bandwidth congestion avoidance conflict
  - ethernet is getting circuit properties (PBT)
  - applications need more and more predictable behaviour

The SCARIe project

SCARIe: a research project to create a Software Correlator for e-VLBI.
VLBI Correlation: signal processing technique to get high precision image from spatially distributed radio-telescope.



To equal the hardware correlator we need:

16 streams of 1Gbps

16 * 1Gbps of data

2 Tflops CPU power

2 TFlop / 16 Gbps =

1000 flops/byte

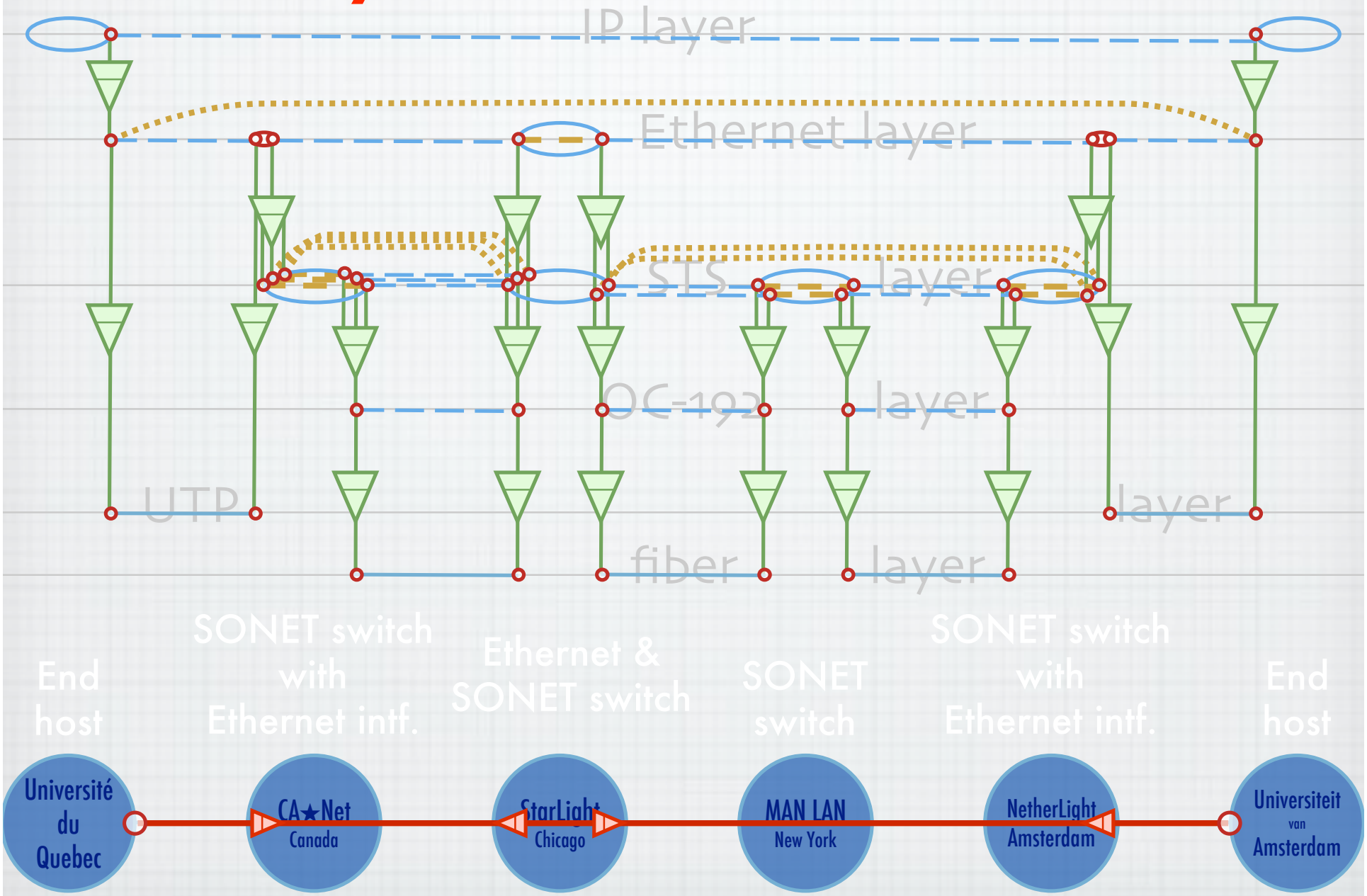
THIS IS A DATA FLOW PROBLEM !!!

sara Computing & Networking Services

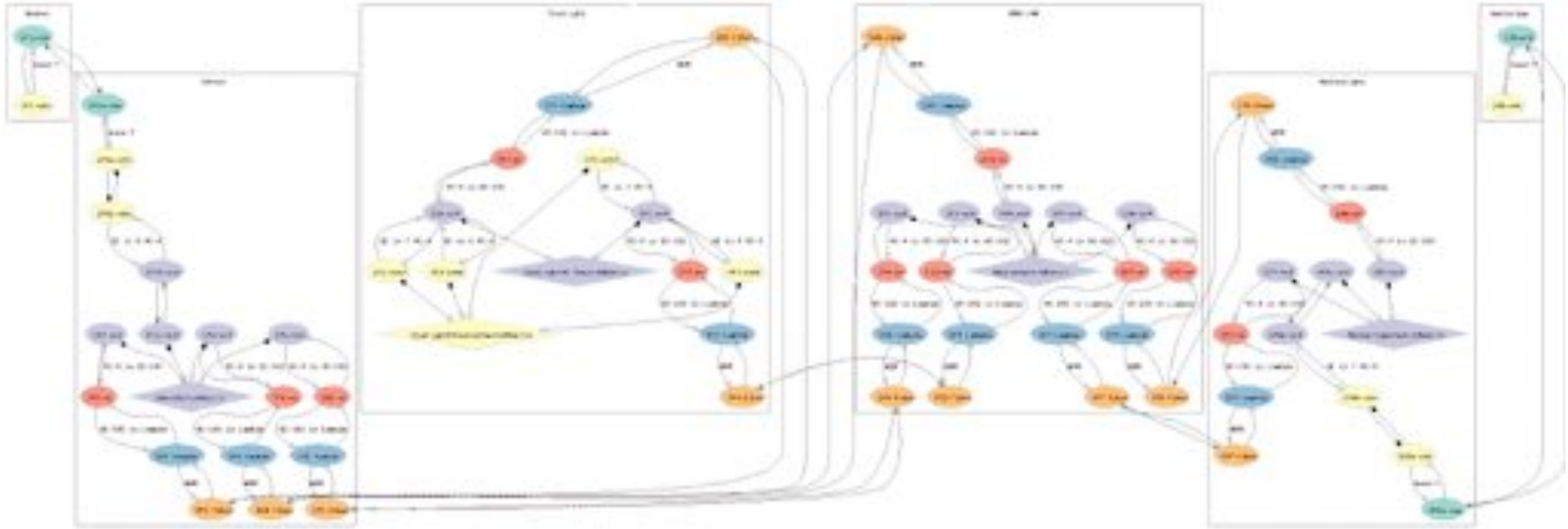
jive JOINT INSTITUTE FOR VLBI IN EUROPE

CineGrid@SARA

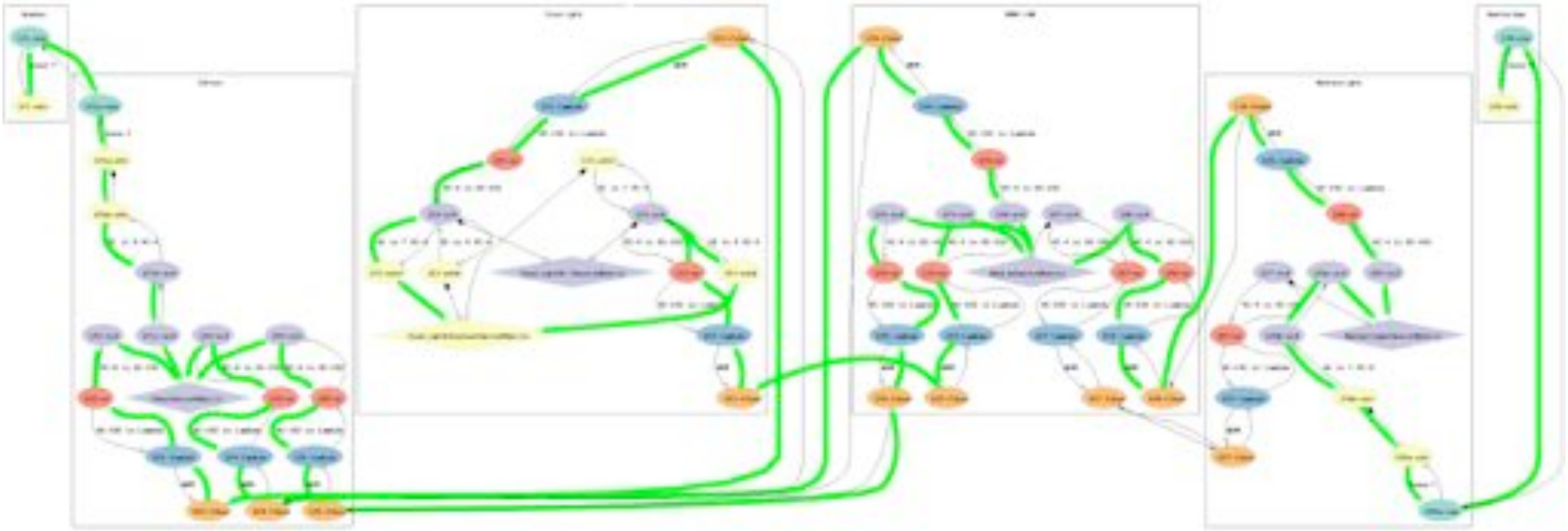
Multi-layer extensions to NDL



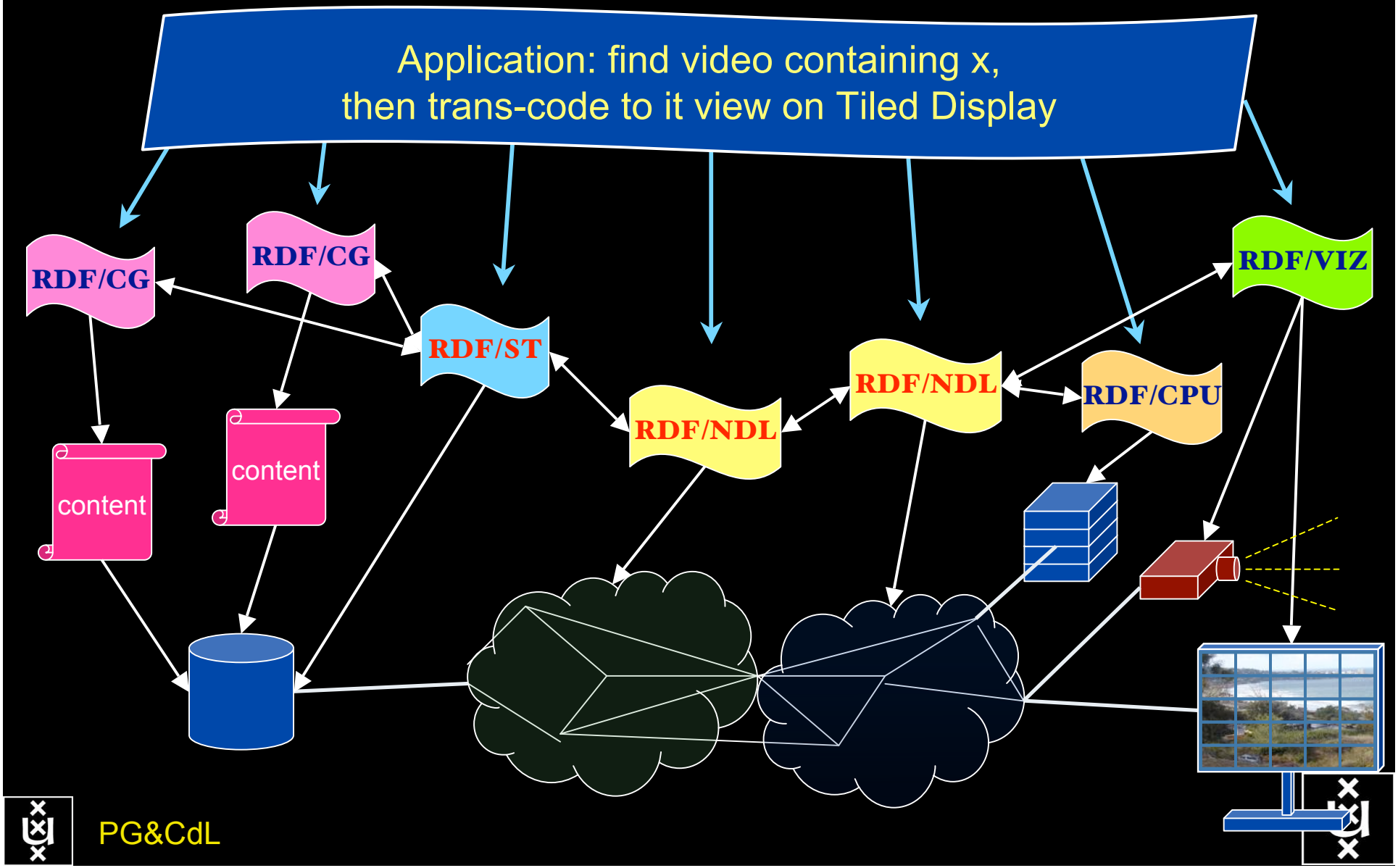
MultiDomain MultiLayer pathfinding in action



MultiDomain MultiLayer pathfinding in action

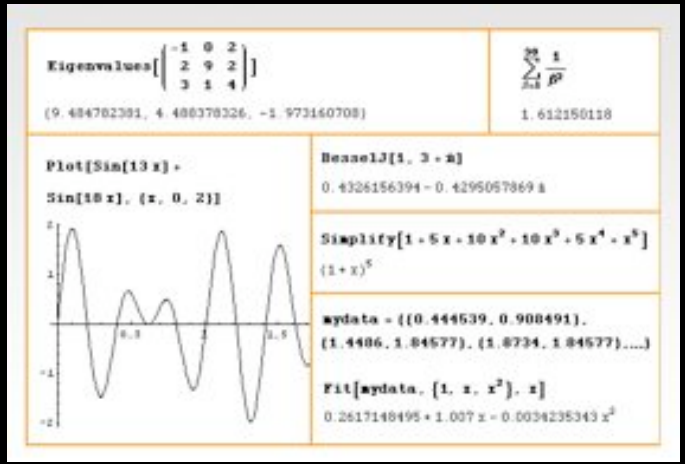


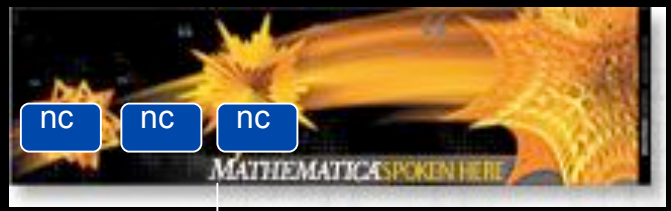
RDF describing Infrastructure

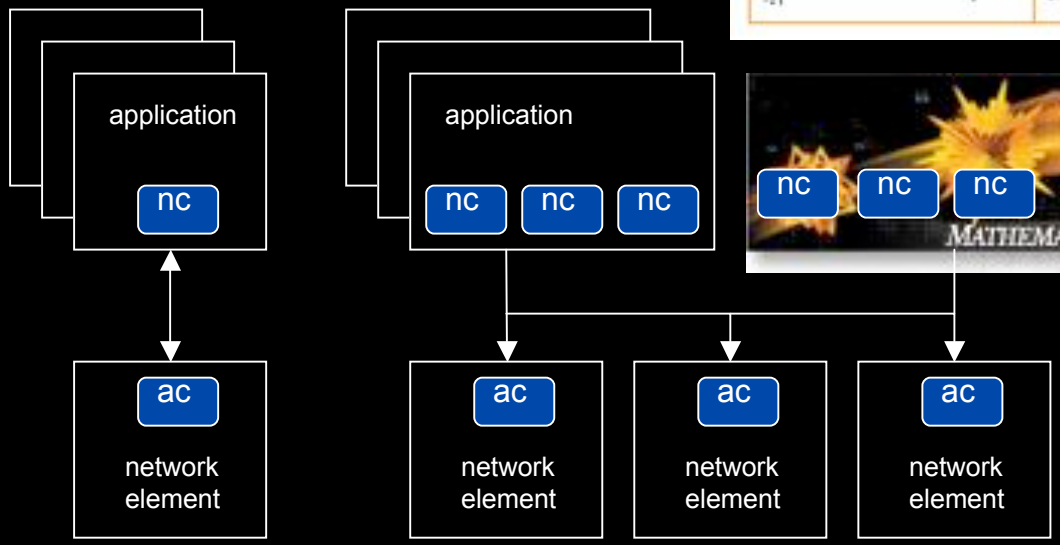


User Programmable Virtualized Networks allows the results of decades of computer science to handle the complexities of application specific networking.

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







StarPlane

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

Topology matters can be dealt with algorithmically

Results can be persisted using a transaction service built in UPVN

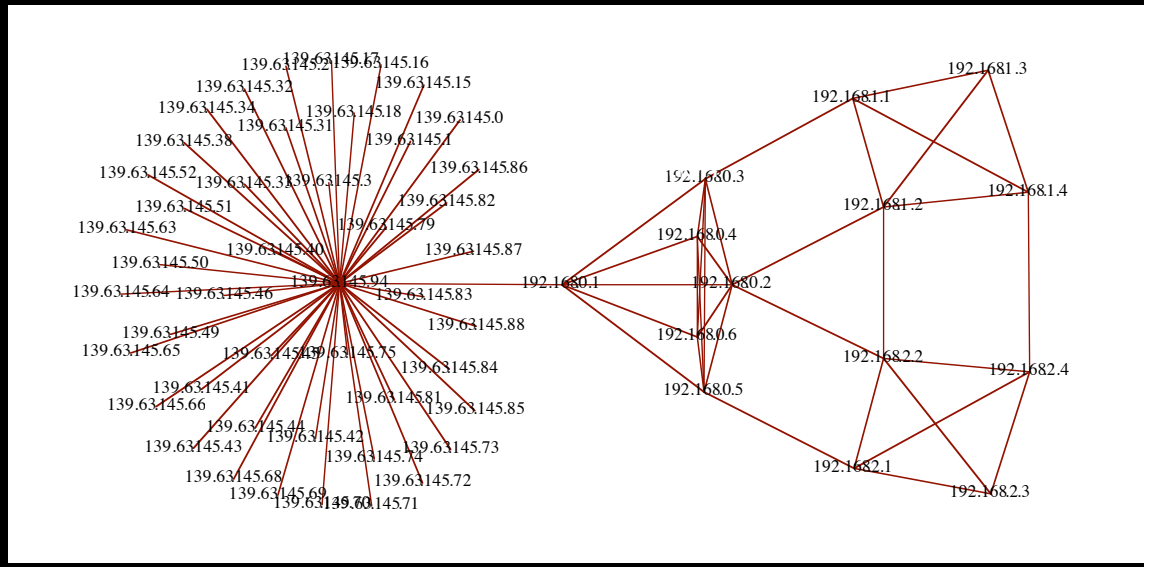
Initialization and BFS discovery of NEs

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

Available methods:
{DiscoverNetworkElements,GetLinkBandwidth,GetAllIpLinks,Remote,
NetworkTokenTransaction}

Global`upvnverbose = True;
AbsoluteTiming[nes = BFSDiscover["139.63.145.94"];][[1]]
AbsoluteTiming[result = BFSDiscoverLinks["139.63.145.94", nes];][[1]]

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

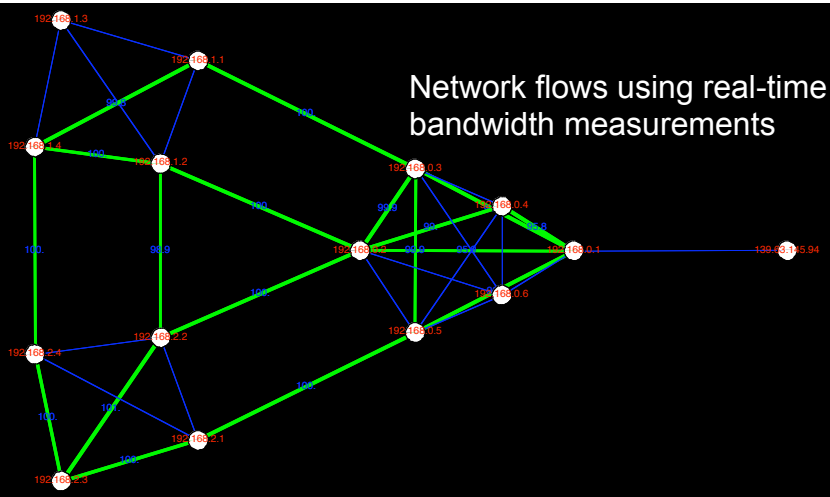


Transaction on shortest path with tokens

```
nodePath = ConvertIndicesToNodes[
                ShortestPath[  g,
                               Node2Index[nids,"192.168.3.4"],
                               Node2Index[nids,"139.63.77.49"]],
                               nids];
Print["Path: ", nodePath];
If[NetworkTokenTransaction[nodePath, "green"]==True,
        Print["Committed"], Print["Transaction failed"]];

Path:
{192.168.3.4,192.168.3.1,139.63.77.30,139.63.77.49}

Committed
```



Network flows using real-time bandwidth measurements

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

StarPlane

TeraThinking

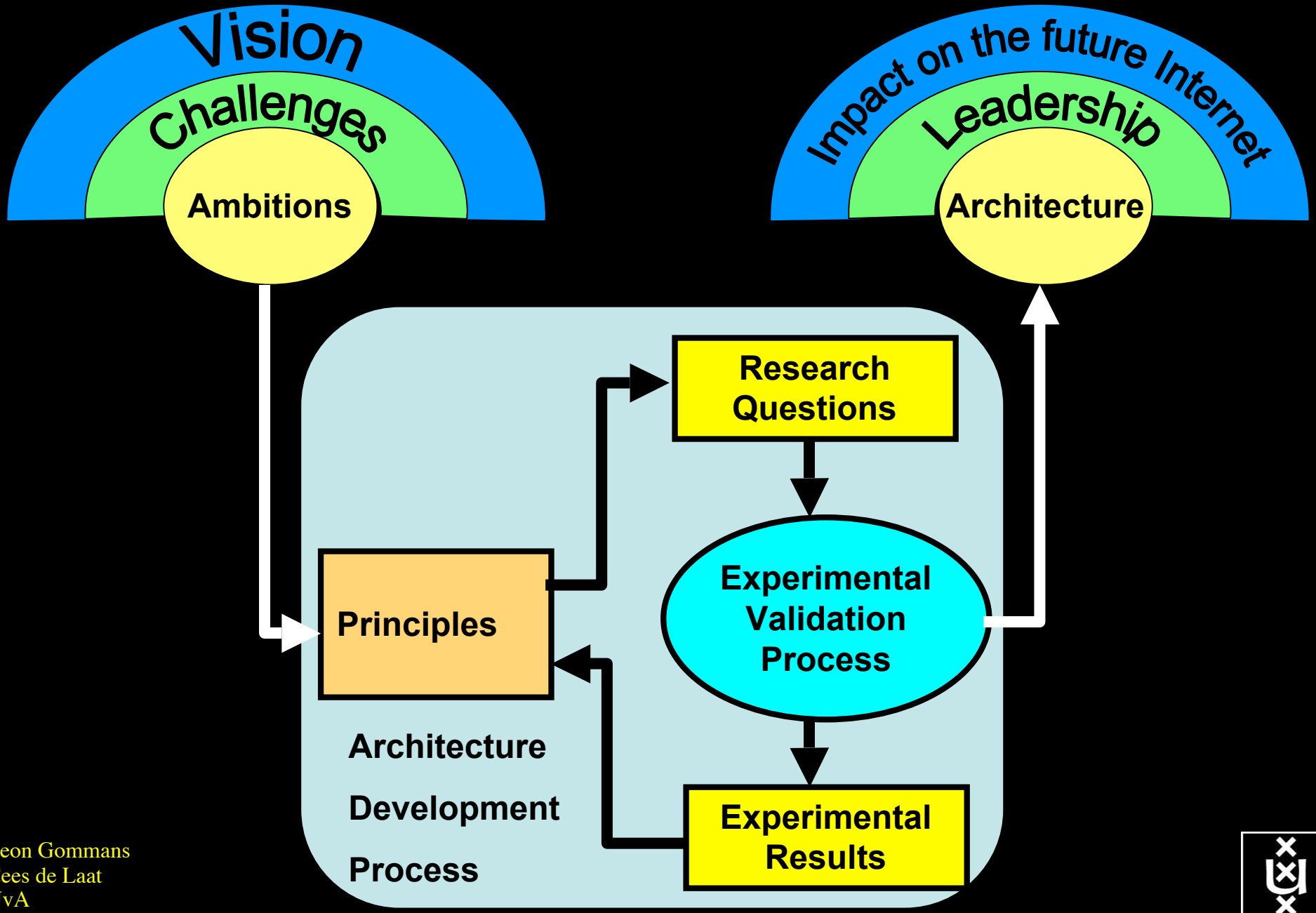
- What constitutes a Tb/s network?
- CALIT2 has 8000 Gigabit drops ?->? Terabit Lan?
- look at 80 core Intel processor
  - cut it in two, left and right communicate 8 TB/s
- think back to teraflop computing!
  - MPI makes it a teraflop machine
- massive parallel channels in hosts, NIC's
- TeraApps programming model supported by
  - TFlops -> MPI / Globus
  - TBytes -> OGSA/DAIS
  - TPixels -> SAGE
  - TSensors -> LOFAR, LHC, LOOKING, CineGrid, ...
  - Tbit/s -> ?

ref Larry Smarr & CdL

Need for discrete parallelism

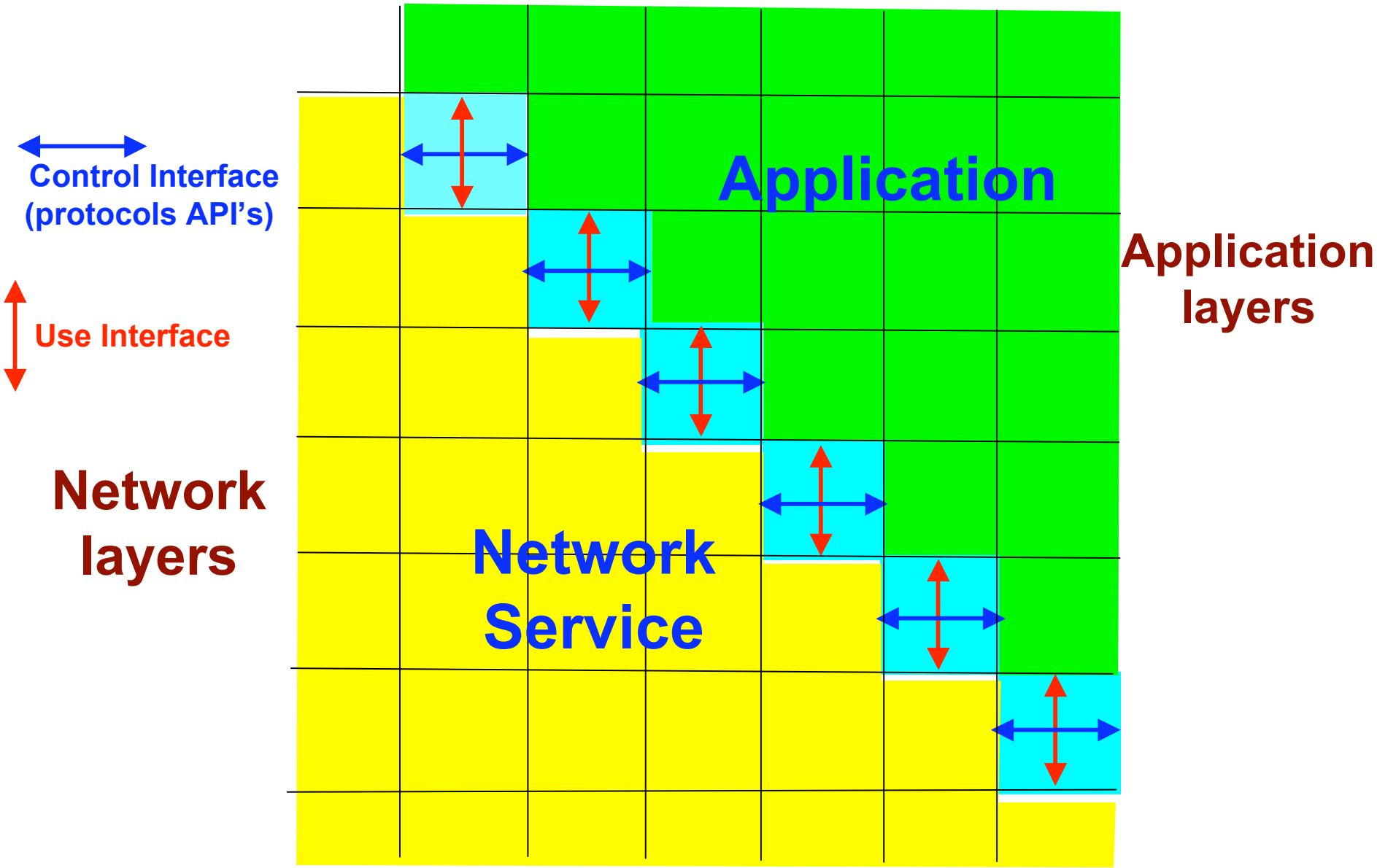
- it takes a core to receive 1 or 10 Gbit/s in a computer
- it takes one or two cores to deal with 10 Gbit/s storage
- same for Gigapixels
- same for 100's of Gflops
- Capacity of every part in a system seems of same scale
- look at 80 core Intel processor
  - cut it in two, left and right communicate 8 TB/s
- massive parallel channels in hosts, NIC's
- Therefore we need to go massively parallel allocating complete parts for the problem at hand!

Process



Leon Gommans
Cees de Laat
UvA

Functional building blocks



The playfield => GLIF

Top 2 recommendations to NSF

1a) keep GLIF populated with Lambda's, CBDF's

1b) support the new generation of exchanges (GOLE's) that are going to augment the current internet exchanges

1c) populate the infrastructure with programmable L1, L2 and L3 objects

2a) keep GLIF populated with People

2b) make sure adequate Systems and Networks engineering research is funded

2c) try to couple US and EU funding so that we can participate at each others projects (e.g. OptIPuter)

Questions ?

I did not talk about StarPlane

AAA

…