Mastering Complex Cyber Infrastructure

Cees de Laat

EU
COMMIT
UvA
NWO
PID/EFRO
SURFnet
TNO
Science Faculty @ UvA

Informatics Institute

- CSA: Computer Systems Architecture (dr. A.D. Pimentel)
- FCN: Federated Collaborative Networks (Prof. dr. H. Afsarmanesh)
- IAS: Intelligent Autonomous Systems (Prof. dr. ir. F.C.A. Groen)
- ILPS: Information and Language Processing Systems (Prof. dr. M. de Rijke)
- ISIS: Intelligent Sensory Information Systems (Prof. dr. ir. A.W.M. Smeulders)
- SCS: Section Computational Science (Prof. dr. P.M.A. Sloot)
- SNE: System and Network Engineering (Prof. dr. ir. C.T.A.M. de Laat)
- TCS: Theory of Computer Science (Prof. dr. J.A. Bergstra)
Mission

The System and Network Engineering research group (SNE) focuses its research on the complexity of emerging hybrid System and Network architectures and the associated models, protocols and system approaches for data processing in science.

The group builds tools and proof of concept applications that promote optimal use of high speed networks. Security of the required mechanisms, infrastructure, middleware, applications and the privacy of data in distributed processing environments is an essential aspect of the research.
... more users!

... more data!

Internet developments

... more realtime!
Internet developments

... more data!

Speed

Volume

Deterministic

Real-time

Scalable

Secure

... more users!
GPU cards are disruptive!

- Fastest supercomputer in the world
- Nr. 500 supercomputer in the world
- 1 single Graphics Processing Unit

Top 500

#1

#500

20,000,000$

7 year

500$

100 pflops

10 pflops

1 pflops

100 tflops

10 tflops

1 tflops

100 gflops

100 gflops

100 mflops

100 mflops

1993-2019
Data storage: doubling every 1.5 year!
Multiple colors / Fiber

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!
Complex e-Infrastructure!
Complex e-Infrastructure!
Why?

I want to:

“Show Big Bug Bunny in 4K on my Tiled Display using green Infrastructure”
Why?

I want to:

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

*TimeLine*

**Good Old Trucking**
- TCP

**Virtualized Internet**
- (G)MPLS
- PBT/PLSB
- OpenFlow
- TBP, SCTCP, ...
- Policy
- TBN
- Hybrid Nets
- GLIF

**“I Want” Internet 3.0**
- OCCI
- NSI
- NM

**Sustainable Internet**
- Cognitive Nets and clouds
  - NDL SF for complex nets
  - SF for Clouds
  - Programmable Networks
- CineGrid

**GreenIT&Nets**
- GreenIT, NDL SF for complex nets

**Machine Learning**
- “I Want” Internet 3.0
- NDL SF for Clouds
- Programmable Networks
- CineGrid
- SF for CineGrid

**2005** **2011** **2020**
Timeline

- Sustainable Internet
- Cognitive Nets and clouds
- Virtualized Internet
- Good Old Trucking

“I Want” Internet 3.0

I retire

2020

2040
<table>
<thead>
<tr>
<th>Green-IT</th>
<th>X</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>Privacy/Trust</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Authorization/policy</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Programmable networks</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>40-100Gig/TCP/WF/QoS</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Topology/Architecture</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Optical Photonic</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Speed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volume</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deterministic Real-time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scalable Secure</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Legend:
- Green-IT
- Privacy/Trust
- Authorization/policy
- Programmable networks
- 40-100Gig/TCP/WF/QoS
- Topology/Architecture
- Optical Photonic

SNE @ UvA
Green-IT
Privacy/Trust
Authorization/policy
Programmable networks
40-100Gig/TCP/WF/QoS
Topology/Architecture
Optical Photonic

Speed
Volume
Deterministic
Real-time
Scalable
Secure

Ijkdijk/Urban Flood
LifeWatch/ENVRI
CosmoGrid/eVLBI
EU-GN3/NOVI/Geyesers
SURFnet/GLIF/Cloud

Speed
Volume
Deterministic
Real-time
Scalable
Secure
Organizational news: Grant e-Science Center

Henri Bal (VU)  
Cees de Laat (UvA)

Project title:  
Generic e-Science initiative for the Netherlands e-Science Center

Grant: 500 k€ (3 y) -> total including matching 3 pd’s + 1 AIO

Currently two subprojects:

1. Big Data, Distributed Data Processing for LOFAR  
   Rob van Nieuwpoort (VU, ASTRON) - Paola Grosso (UvA), Ralph Wijers (UvA)

2. Information Management  
   Frank van Harmelen (VU) – Pieter Adriaans (UvA)

Some other grants at I VL – System & Network Engineering research in the past year:  
GigaPort3-2012 318 k€ (1y), EU-ENVRI (3y) 248 k€, SURF (1y) 240 k€  
GigaPort3-2011 238 k€ (1y), COMMIT (5y) 1480 k€, NWO GreenClouds (4y) 205 k€
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 small flows -> 12 Gb/s

Many Pflops/s
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 Virtualized Networks, accepted for publication to the IEEE e-Science 2006 conference Amsterdam.
In the Intercloud, virtual servers and networks become software.

- Virtual Internets adapt to the environment, grow to demand, iterate to specific designs.
- Network support for application specific interconnections are merely optimizations: Openflow, active networks, cisco distributed switch.
- But how to control the control loop?

**Diagram:**

- Application
- Controller
- Reference
- Model
- $NC_x$
- $NC_y$
- AC Properties
- AC Actions
- $\delta$
- NE
Interactive Networks

Rudolf Strijkers 1, 2
Marc X. Makkes 1, 2
Mihai Christea 1
Laurence Muller 1
Robert Belleman 1
Cees de Laat 1
Robert Meijer 1, 2

1 University of Amsterdam, Amsterdam The Netherlands
2 TNO Information and Communication Technology, Groningen, The Netherlands
We investigate: for complex networks!
The GLIF – lightpaths around the world
Alien light
From idea to realisation!

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

Alien wavelength advantages
- Direct connection of customer equipment\(^1\) \(\rightarrow\) cost savings
- Avoid OEO regeneration \(\rightarrow\) power savings
- Faster time to service\(^2\) \(\rightarrow\) time savings
- Support of different modulation formats\(^3\) \(\rightarrow\) 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.

Test results
Error-free transmission for 23 hours, 17 minutes \(\rightarrow\) BER \(<\) 3.10^-15

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 wave lengths.
- 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.

References
\(^1\) "TRANSMISSION SOLUTIONS FOR 40-GIGABIT RATES", L. GERTEL ET AL., OFC/NFOEC 2009
\(^2\) "NEXT OPTICAL TRANSPORT NETWORKS"; ABRAMS & DIGNAT, OSA TECHNICAL DIGEST
\(^3\) "TWIN WAVES OF ALL-OPTICAL CORE NETWORKS"; ABRAMS & DIGNAT, OSA TECHNICAL DIGEST

Acknowledgements
We acknowledge Nortel 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.

Transmitter Module Test

End-to-end Path = 1056 km TWRS (TrueWave Reduced Slope) transmission fiber

License for use: NCF

Communication Systems

Networking Technologies

Alien Wavelength

R&D Network

Optical Systems

Innovations in Network Architecture

Automation and Management

Optimized Services

Optimization of Network Performance
Alien light
From idea to realisation!

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

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.

\[
FOM = \sum_{i=1}^{N} L_i \quad \text{[dB]}
\]

\(L_i\): span losses in dB
\(N\): number of spans

Test results

Error-free transmission for 23 hours, 17 minutes → BER \(\times 3.0 \times 10^{-14}\)

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.

REFERENCES
\[1\] "OPERATIONAL SOLUTIONS FOR 40G WAVELENGTHS LAYER", O. GERSTE, ET AL., IEC'2009
\[2\] "40 Gb/s TRANSPORT NETWORKS" MAURICE EBEL, DIGITAL OPTICAL
\[3\] "TRANSMISSIOM OF ALL-OPTICAL CORE NETWORKS" ANWAR AL-MASRI, LUIGI GACERI, ET AL., ECOC'2009

ACKNOWLEDGMENTS
We are grateful to NORDUnet for providing us with bandwidth on their testbed for this experiment and also for their support and assistance during the experiments. We also acknowledge TELEFIS and NORTTEL for their infrastructure 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
- Mellanox

CERN
- CIENA OME 6500

Hamburg
- CIENA OME 6500

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

  Subject \( \rightarrow \) Predicate \( \rightarrow \) Object

  - Location
    - name
    - connectedTo
  - Device
    - description
    - capacity
  - Interface
    - locatedAt
    - encodingType
  - Link
    - hasInterface
    - encodingLabel
RDF describing Infrastructure

Application: find video containing x, then trans-code to it view on Tiled Display

RDF/CG → RDF/CG
RDF/CG → RDF/ST
RDF/ST → RDF/NDL
RDF/NDL → RDF/NDL
RDF/NDL → RDF/CPU
RDF/CPU → RDF/VIZ

content

PG&CdL
Virtualisatie van infrastructuur & QoS

Currently, testbeds are usually local, each featuring a specific technology and oriented to a specific set of tests. Virtualization is a keyword having a profound influence in technology. In its network-oriented meaning, it corresponds to the creation of multiple virtual networks from a physical infrastructure. The network virtualization can be performed at most layers of the ISO/OSI stack, from the data link to the network layer, e.g., through the creation of multiple logical router instances in the same physical router. In its system-oriented sense the process is even faster and more intriguing, leading to multiple operating systems with very varying tasks running on the same hardware platform. The importance of virtualisation becomes more evident if the virtual entity is allowed to move itself from one host to another in the network.

Last, but not least, there is a rising need to join forces between Research Institutes, NREN, vendors, in order to face the challenge and to reach consensus on the way forward to the new Internet.

The FEDERICA Concept

The issues in the previous section provide the background on which the FEDERICA project is rooted. The issues listed cannot be faced in a production network environment and require the evolution of the current research activities. The activity should leverage a broad confrontation of ideas through experimental verification and involve as many researchers as possible. To be effective, the research and support should be extended to the European scale.

The best way to face this challenge is to create a dedicated infrastructure to allow disruptive research and be capable of hosting different types of network researches and experiments. FEDERICA will deploy, support and promote a federated virtual e-Infrastructure, over existing NREN-GEANT2 substrates, along with a tool-bench to manage the virtualized environment and to facilitate disruptive experiments by network research communities. The following Figure 1 provides a pictorial description of the planned FEDERICA infrastructure.
The laboratory of environmental research infrastructures

Deep Earth, land and sea, the atmosphere

Living and dead environments
A cottage industry in the data desert

- Interdisciplinary challenges
  - Data generator
  - User
  - Data generator
  - User
  - Data generators
  - User

- Support services
  - Tool
  - Tool
  - Tool
  - Tool

- Data infrastructure
  - Data storage
  - Data storage
  - Data storage
  - Data storage
Data generators

Users

Data Services
+ processed data and workflows

Community Support Services

Data transfer
Fast data transmission
Operation at remote sites

User functionalities
Virtual Environments & Collaborative organisations
Security & Protection

Data discovery & Navigation
Data submission tools (meta) data tagging tools
Operational Semantic Interoperability

Workflow Generator
Knowledge management
Virtualisation

Persistant storage capacity
24/7 operation
Preservation & Sustainability (digital asset management)

Authenticity
Certification & Integrity
GUIDs
Challenges

• Data – Data – Data
  – Archiving, publication, searchable, transport, self-describing, DB innovations needed, multi disciplinary use

• Virtualisation
  – Another layer of indeterminism

• Greening the Infrastructure
  – e.g. Department Of Less Energy: [link to PDF]

• Disruptive developments
  – BufferBloath, Revisiting TCP, influence of SSD’s & GPU’s
  – Multi layer Glif Open Exchange model
  – Invariants in LightPaths (been there done that 😊)
    - X25, ATM, SONET/SDH, Lambda’s, MPLS-TE, VLAN’s, PBT, OpenFlow, ….
  – Authorization & Trust & Security and Privacy
The Way Forward!

- Nowadays scientific computing and data is dwarfed by commercial & cloud, there is also no scientific water, scientific power.
  - Understand how to work with elastic clouds
  - Trust & Policy & Firewalling on VM/Cloud level
- Technology cycles are 3 – 5 year
  - Do not try to unify but prepare for diversity
  - Hybrid computing & networking
  - Compete on implementation & agree on interfaces and protocols
- Limitation on natural resources and disruptive events
  - Energy becomes big issue
  - Follow the sun
  - Avoid single points of failure (aka Amazon, Blackberry, …)
  - Better very loosely coupled than totally unified integrated…
Hybrid Networking <-> Computing

Routers <-> Supercomputers

Ethernet switches <-> Grid & Cloud

Photonic transport <-> GPU’s

What matters:

Energy consumption/multiplication
Energy consumption/bit transported
ECO-Scheduling
Education- Master SNE

- Open Source approach
- Based on open and non-discriminatory standards
- Privacy and Security
- Digital security & forensics
- Internet infrastructure
- Master closely related to researchgroup!