## Smart Cyber Infrastructure for Big Data Processing Paola Grosso & Cees de Laat



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



- Capacity
- Capability
- Security
- Sustainability
- Resilience



- 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

- 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?





There **i**S always a bigger fish

| Business email sent per year                                   | 2,986,100 |
|--|-----------|
| <ul> <li>Content uploaded to Facebook each year</li> </ul>     |           |
| Google's search index  |           |
| <ul> <li>Kaiser Permanente's digital health records</li> </ul> | 30,720    |
| <ul> <li>Large Hadron Collider's annual data output</li> </ul> | 15,360    |
| Videos uploaded to YouTube per year                            | 15,000    |

| National Climactic Data Center database                    | 6,144 |
|--|-------|
| Library of Congress' digital collection                    |       |
| US Census Bureau data                                      |       |
| <ul> <li>Nasdaq stock market database</li> </ul>           |       |
| O Tweets sent in 2012                                      | 19    |
| <ul> <li>Contents of every print issue of WIRED</li> </ul> | 1.26  |

### more data! . . .





## Multiple colors / Fiber



Wavelength Selective Switch

New: Hollow Fiber! → less RTT! S=



COPYRIGHT : MORTEN INGEMANN

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.

### 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              | AFFILIATIO | N E-MAIL                    | A-SIDE                     | Z-SIDE             | PORTS(S)<br>MAN LAN                                  | PORTS(S)<br>TNC2013                      | DETAILS   |
|------|---|--------------------|------------|-----------------------------|----------------------------|--------------------|--|--|---|
| 1    | Big data transfers with<br>multipathing, OpenFlow<br>and MPTCP          | Ronald van der Pol | SURFnet    | ronald.vanderpol@surfnet.nl | TNC/MECC,<br>Maastricht NL | Chicago, IL        | Existing 100G link<br>between internet2<br>and ESnet | 2x40GE (Juniper)+<br>2x10GE<br>(OME6500) | In this demonstration we show how multipartiting, OpenFlow and Multipath TCP (MPTOP) can help in<br>large the brankers between data centres (Mastanch and Chacago). An OpenFlow application provisions<br>multiple paths between the serving and PCIPC will be used on the environs is an inductionally shed<br>traffic parases all those paths. This forms use 2 XeVO or the transitionities model and flow provides<br>20x00 between MUL And and State(X-E and USIA) refer provides additional paths. |
| 2    | Visualize 100G traffic  | Inder Monga        | ESnet      | imonga@es.net               |                            |                    |  |  | Using an SNMP feed from the Juriper switch at TNC2/013,and/or Enocade AL25 node in MANLAN,<br>this demo would visualize the total traffic on the link, of all demos aggregated. The network diagram<br>will show the transatlantic topology and some of the demo topologies.  |
| 3    | How many modern<br>servers can fill a 100Gbps<br>Transatlantic Circuit? | Inder Monga        | ESnet      | imonga@es.net               | Chicago, III               | 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<br>generate almost BOORp of table. Each server has 4 KO NOS connected to a 400 vitual cruzil, and has<br>gen17 anning to generate traffic. Spectra (new "gen17" through measurement took all in "beta", combines<br>the best features from other tools such as perf, witting, and neight. See: https://mys.net/demos/turc2007/  |
| 4    | First European ExoGENI<br>at Work                                       | Jeroen van der Ham | UvA        | vdham@uva.nl                | RENCI, NC                  | UvA, Amsterdam, NL | 1x 10GE  | 1x 10GE                                  | The ExoGEN racks at RENCI and UvA will be interconnected over a 10G pipe and be on continuously,<br>showing GENI connectivity between Ansterdam and the rest of the GENI nodes in the USA.  |
| 5    | Up and down North<br>Atlantic @ 100G                                    | Michael Enrico     | DANTE      | michael.enrico@dante.net    | TNC showfloor              | TNC showfloor      | 1x 100GE   | 1x 100GE                                 | The DANTE 100EE test set will be placed at the TNC2013 showfloor and connected to the Juniper at<br>1000. When this derno is running a loog (i) MAN LAN's Brocode switch will ensure that the traffic<br>sent to MAN LAN's themse to the showfloor. On display is the throughput and RTT (to show the traffic<br>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



### Alien wavelength advantages

- Direct connection of customer equipment<sup>[1]</sup>
   → cost savings
- Avoid OEO regeneration → power savings
- Faster time to service<sup>[2]</sup> → time savings
- Support of different modulation formats<sup>[3]</sup>
   → 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.



### Transmission system setup

JOINT SURFnet/NORDUnet 40Gb/s PM-QPSK alien wavelength DEMONSTRATION.



### Test results



Error-free transmission for 23 hours, 17 minutes  $\rightarrow$  BER < 3.0  $10^{\text{-16}}$ 

### 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









REFERENCES [1] "OPERATIONAL SOLUTIONS FOR AN OPEN DWOML LAVER", OL GERTEL ET AL, OPE 2009 [2] "ATAT OPTICAL TANSPORT SERVICES", RABBARA E. SANTH, OPE 200 [3] "OPEN SANNES OF ALL-OPTICAL CORE INTROMES", ANDERVICIO DA DACIA ELINISTIER, EL COMUNICATION ACKNOWLEDGEMENTS WE ARE GATEFUL TO NORDUNET FOR PROVIDING US WITH BANDWOTH ON THER DWOML LINK FOR THE SUPERIT AND ASSTANCE DURING THE EXPERIMENTS, WE ALSO ACCIONDULES UTILI BANDWOTH ON THER DWOML LINK FOR THE SUPERIT MITERIAL COMUNICATION DURING THE EXPERIMENTS, WE ALSO ACCIONDULES DE TUTILIDADA DA DORTE FOR THE DIRIGHTATION WORK AND SINULATION SUPPORT DURING THE EXPERIMENTS, WE ALSO ACCIONDULES DE TUTILIDADA DA DORTE FOR THE DIRIGHTATION WORK AND SINULATION SUPPORT

# 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<sup>[1]</sup>
   → cost savings
- Avoid OEO regeneration → power savings
- Faster time to service<sup>[2]</sup> → time savings
- Support of different modulation formats<sup>[3]</sup>
   → 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.

NØRTEL

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.



### Transmission system setup

JOINT SURFnet/NORDUnet 40Gb/s PM-QPSK alien wavelength DEMONSTRATION.



Test results



Error-free transmission for 23 hours, 17 minutes  $\rightarrow$  BER < 3.0  $10^{\text{-16}}$ 

### 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.

NORDUnet





- 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







An effort started in 2010 (in parallel with our involvement in the FP7 projects Geysers and NOVI).

The goal was to capture the concept of virtualization in <u>computing</u> infrastructures and to describe the storage and computing capabilities of the resources.

A key feature is the decoupling of virtualization, connectivity and functionalities.

### It is built upon the NML ontology.

It uses the **nml:node** concept as basic entity to describe resources in computing infrastructures.

It can be used as:

- a stand-alone model (i.e. without any network descriptions),
- in combination with NML by importing the NML ontology into the INDL definition.





### COMMIT/ Our connecting models



'RI



## **NOVI** Federation

ENVRI





lacksquare



## INDL use in NOVI

• Two nodes in the NOVI federation:





# INDL in Geysers



• The virtualization model:



# The OIntEd editor



### http://indl-gui.appspot.com/





COMMIT/

W. Adianto, R.Koning, P. Grosso, A. Belloum, M. Bubak and C.de Laat, OIntEd: online ontology instance editor enabling a new approach to ontology development In: Journal of "Software: Practice and Experience" 2012







## NML and NSI

- NML Network Markup Language and NSI Network Service Interface
- within the OGF.
- See: "Network Markup Language Base Schema version 1"
- The Network Markup Language purpose is to create a functional description of <u>multi-layer</u> and <u>multi-domain</u> networks. It can be used for <u>aggregated</u> or <u>abstracted</u> topologies.
- Under development: the Network Service Interface Topology Extensions (Draft OGF Standard)





## NewQOSPlanner

The NSI – Network Service Interface – creates on the fly connections between domains.



Domains in request



Z. Zhao, J. v/d Ham, A. Taal, R. Koning, P. Grosso and C. de Laat *Planning data intensive workflows on inter-domain resources using the Network Service Interface (NSI)* In: WORKS 2012

## HyperFlow







C. Dumitru, Z. Zhao, P. Grosso and C. de Laat *HybridFlow: Towards Intelligent Video Delivery and Processing Over Hybrid Infrastructures* (In CTS 2013))





### Pareto fronts





- 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

## Green scheduling







COMMIT/



×X×

Q. Chen, P. Grosso, K. van der Veldt, C. de Laat, R. Hofman and H.Bal. *Profiling energy consumption of VMs for green cloud computing* In: International Conference on Cloud and Green Computing (CGC2011), Sydney December 2011

### сомміт/ Energy Efficient Ethernet (802.3az)

23

22

Power savings techniques in hardware can be leveraged in architecturing communication patterns in data centra





D. Pavlov and J. Soert and P. Grosso and Z. Zhao and K. van der Veldt and H. Zhu and C.de Laat *Towards energy efficient data intensive computing using IEEE 802.3az* 

In: DISCS 2012 workshop - Nov 2012

# Networks and CO2

- Take a network (Esnet, working on using SURFnet data)
- Define the traffic model running on it
- Use the energy monitoring information and energy costs data
- Compare path selection strategies : shortest, cheapest and greenest



"A motivation for carbon aware path provisioning for NRENs" (submitted to eEnergy2014)









In region 1 the task should be performed locally, independently of the type of transport network.

In region 2 the task can be performed remotely provided that the connection is a light path.

In region 3 the task should be done remotely for both types of transport networks.

Given different network paths we can identify decision boundaries as function of the task complexity.



## **ECO-Scheduling**



### 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?

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



Ana Oprescu Cees de Laat Marc MakkesRalph Koning Leon Gommans Fahimeh Bas Terwin **Cosmin Dumitri** Pieter Adriaans Rob MeijerKarel van der Veldt ri Demchen Reggie Cushing Jan Sipke van der Veen Miroslav Zivkovic Naod Duga Jebessa Sander Klous Jeroen van der Ham Jaap van Ginke Paul Klint Souley Madougou Ngo Tong Canh Adianto Wibisono Anna banescu Gerben de Vries Hans Dijkman Arno Bakker Marian Bubak

Daniel Romao

Supported by:

Erik-Jan Bos Peter Bloem

Arie Taal



