Network Advances on Smart Cyber Infrastructure for Big Data Processing

Cees de Laat
ATLAS detector @ CERN Geneve

15 Pbyte/year
What Happens in an Internet Minute?

- 20 new victims of identity theft
- 47,000 app downloads
- 61,141 hours of music
- 20 million photo views
- 3,000 photo uploads
- 20 million photo uploads
- 320+ new Twitter accounts
- 100,000 new tweets
- 277,000 logins
- 6 million Facebook views
- 2+ million search queries

639,800 GB of global IP data transferred
135 botnet infections
1,300 new mobile users
100+ new LinkedIn accounts

And Future Growth is Staggering

Today, the number of networked devices = the global population
By 2015, the number of networked devices = 2x the global population
In 2015, it would take you 5 years to view all video crossing IP networks each second.
There is always a bigger fish
Computing vs Data

Computing per unit cost has doubled roughly every 18 months.

Space per unit cost has doubled roughly every 14 months.

So: data becomes exponentially uncomputable.

http://www.mkomo.com/cost-per-gigabyte
GPU cards are disruptive!

- 20,000,000$ in 7 years
- 500$ in 7 years

Top 500

- #1 fastest supercomputer in the world
- #500 nr. 500 supercomputer in the world
- 1 single Graphics Processing Unit

Graph showing the trend of supercomputer performance from 1993 to 2019.
Multiple colors / Fiber

Per fiber: ~ 80-100 colors * 50 GHz
Per color: 10 – 40 – 100 Gbit/s
About 10 Tbit/s per fiber long dist.
BW * Distance ~ 2*10^{17} bm/s

New: Hollow Fiber!
⇒ less RTT!
Mission

Can we create smart and safe data processing systems that can be tailored to diverse application needs?

- **Capacity**
  - Bandwidth on demand, QoS, architectures, photonics, GPU, 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
The GLIF – LightPaths around the World

Amsterdam is a major hub in The GLIF

ExoGeni @ OpenLab – UvA
http://sne.science.uva.nl/openlab/

Installed and up June 3th 2013

Connected via
the new 100 Gb/s
transatlantic
To US-GENI

TNC2013 DEMOS JUNE, 2013

<table>
<thead>
<tr>
<th>DEMO</th>
<th>TITLE</th>
<th>OWNER</th>
<th>AFFILIATION</th>
<th>E-MAIL</th>
<th>A-SIDE</th>
<th>Z-SIDE</th>
<th>PORTS(S)</th>
<th>PORTS(S) TNC2013</th>
<th>DETAILS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Big data transfers with multiqueue OpenFlow and MIPv6</td>
<td>Ronald van der Pol</td>
<td>SURFnet</td>
<td><a href="mailto:renald.wanderspol@surfnet.nl">renald.wanderspol@surfnet.nl</a></td>
<td>TNC/MECC, Maastricht, NL</td>
<td>Chicago, IL</td>
<td>Existing 100G link between Arlington and 85Street</td>
<td>2x100G link - TNC2013 (June/13)</td>
<td>In this demonstration, show how multiplexing, OpenFlow and Multiple VOF (MIPv6) can help in scaling the high bandwidth data between Arlington and Chicago. For enabling gigabit per second bandwidth for demanding traffic scenarios at high speed, this demo uses COTS on 100Gb/s transatlantic link. The network diagram shows the connections, traffic management and some of the delay characteristics.</td>
</tr>
<tr>
<td>2</td>
<td>Visualise 100G traffic</td>
<td>Inder Monga</td>
<td>ISGnet</td>
<td><a href="mailto:inder@isgnet.net">inder@isgnet.net</a></td>
<td>Chicago, IL</td>
<td>TNC showfloor</td>
<td>1x100GE</td>
<td>6x10GE</td>
<td>In this demonstration, we show how the user can view in real-time the traffic coming in from a 100GbE link and the output going out. The network diagram shows the connections and traffic management.</td>
</tr>
<tr>
<td>3</td>
<td>How many modern servers can fit 100Gbps TRANSLAPID Circuit</td>
<td>Inder Monga</td>
<td>ISGnet</td>
<td><a href="mailto:inder@isgnet.net">inder@isgnet.net</a></td>
<td>Chicago, IL</td>
<td>TNC showfloor</td>
<td>1x100GE</td>
<td>1x10GE</td>
<td>In this demonstration, we show how the user can view in real-time the traffic coming in from a 100GbE link and the output going out. The network diagram shows the connections and traffic management.</td>
</tr>
<tr>
<td>4</td>
<td>Post European ExoGENI at TNC2013</td>
<td>Jaap van der Helm</td>
<td>UvA</td>
<td><a href="mailto:vihname@uva.nl">vihname@uva.nl</a></td>
<td>RESCL, Inc</td>
<td>UvA, Amsterdam, NL</td>
<td>1x10GE</td>
<td>1x10GE</td>
<td>In this demonstration, we show how the user can view in real-time the traffic coming in from a 100GbE link and the output going out. The network diagram shows the connections and traffic management.</td>
</tr>
<tr>
<td>5</td>
<td>1up and down North Atlantic 100G</td>
<td>Michael S.</td>
<td>DASIE</td>
<td><a href="mailto:michaels@dasie.net">michaels@dasie.net</a></td>
<td>TNC showfloor</td>
<td>TNC showfloor</td>
<td>1x100GE</td>
<td>1x10GE</td>
<td>In this demonstration, we show how the user can view in real-time the traffic coming in from a 100GbE link and the output going out. The network diagram shows the connections and traffic management.</td>
</tr>
</tbody>
</table>
40Gb/s alien wavelength transmission via a multi-vendor 10Gb/s DWDM infrastructure

Alien light
From idea to realisation!

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

\[ FOM = \sum_{i=1}^{N} \left[ L_i \right] \]

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

\( A \quad B \quad C \)

FoM
5904
5504
1007

Easy-to-use formula that accurately quantifies transmission system performance

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

Test results

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
[3] “ERRORS SAVING ON ALL-OPTICAL CORE NETWORKS”, DAMIANO ILCUCCI AND GABRIELE MOCCHIUTTI, DE MULTIVENDOR TECHNOLOGICAL COMMUNICATION

We are grateful to NORDUnet for providing us with bandwidth on their DWDM line 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.

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

Alien light
From idea to realisation!

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.

\[
FOM = \sum_{i=0}^{N} \left| L_i \right|
\]

where:
- \( FOM \) is the Figure of Merit
- \( N \) is the number of spans
- \( L_i \) is the span losses in dB

Test results

Error-free transmission for 23 hours, 17 minutes → BER < 3.0 \times 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 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] "ONERIAL SOLUTIONS FOR AN ADAPTED NETWORK layE", N. GERSTER ET AL. NETWORKS [2] "A NEW OPTICAL TRANSPORT NETWORK: ADVANCED OPTICAL NETWORK "TRANSLATING THE ALL-OPTICAL CORE NETWORK"" "OPTICAL CORE AND OPTICAL NETWORK" [3] "A NEW OPTICAL TRANSPORT NETWORK: ADVANCED OPTICAL NETWORK. We are grateful to Nortel for providing us with bandwidth on their DWDM line 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.

Acknowledgements
ClearStream @ TNC2011

Setup codename: FlightCees

UvA
- iPerf
  - I7 3.2 GHz Q-core
- Amd Ph II 3.6 GHz HexC

Copenhagen
- iPerf
  - 2* dual 2.8 GHz Q-core

CERN
- CIENA OME 6500

Hamburg
- CIENA OME 6500

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

27 ms RTT

17 ms RTT
We used `numactl` to bind `iperf` to cores.
Yesterday’s Media Transport Method!

8 TByte
Moving Cinegrid Objects Globally

- **Digital Motion Picture for Audio Post-Production**
  - 1 TV Episode Dubbing Reference ~ 1 GB
  - 1 Theatrical 5.1 Final Mix ~ 8 GB
  - 1 Theatrical Feature Dubbing reference ~ 30 GB

- **Digital Motion Picture Acquisition**
  - 4K RGB x 24 FPS x 10bit color: ~ 48MB/Frame uncompressed *(ideal)*
  - 6:1 ~ 20:1 shooting ratios => 48TB ~ 160TB digital camera originals

- **Digital Dailies**
  - HD compressed MPEG-2 @ 25 ~ 50 Mb/s

- **Digital Post-production and Visual Effects**
  - Gigabytes - Terabytes to Select Sites Depending on Project

- **Digital Motion Picture Distribution**
  - Film Printing in Regions
    - Features ~ 8TB
    - Trailers ~ 200GB
  - Digital Cinema Package to Theatres
    - Features ~ 100 - 300GB per DCP
    - Trailers ~ 2 - 4GB per DCP

*UHDTV*(4K) 3840

*2160*
CineGrid Portal

Unified orchestration of distributed CineGrid resources
Motivation:
- previous simulations found >100 times more substructure than is observed!

Simulate large structure formation in the Universe

Method: Cosmological $N$-body code

Computation: Intercontinental SuperComputer Grid

Current (2013) problem:
- 2 PByte data in Oak Ridge!

1.0 Gb/s dedicated network
270 ms RTT
We investigate: for complex networks!
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  Predicate   Object
```

- Location: name, connectedTo
- Device: description, capacity
- Interface: locatedAt, encodingType
- Link: hasInterface, encodingLabel
Network Topology Description

Network topology research supporting automatic network provisioning
• Inter-domain networks
• Multiple technologies
• Generalized to ExoGeni & FED4FIRE

http://redmine.ogf.org/projects/nml-wg
http://redmine.ogf.org/projects/nsi-wg
http://sne.science.uva.nl/ndl
Applications, Data and Networks become aware of each other!
Storage to energy:
• When should you move hot or cold data to a green remote data centra for storage?
• Given different network paths what are the decision boundaries as function of the task complexity.

http://sne.science.uva.nl/bits2energy/  Taal & Grosso
Security of Data in Purpose-Driven Virtual Machines

⇒ Cloud VM’s contain lots of lib’s and features not needed by application
⇒ Everything is a risk for hacks
⇒ Minimize attack surface

I need a Debian VM to run fsl in text mode in the SURFsara HPC Cloud.
The application is `fs1-4.1`. LEFT: before dependency resolution, with all dependency constraints shown and RIGHT: resolved dependencies in a particular setup, `libc` is the center node.
Towards Defining Big Data Architecture Framework

Yuri Demchenko, Marcel Worring, Wouter Los, Cees de Laat

Big Data Paradigm Change: Moving to Data-Centric Models

Current IT and communication technologies are host based or host-centric (service/message centric)
- Any communication or processing are bound to host/computer that runs software
- For security: all security models are host/client based

Big Data requires new data-centric models
- Data location, replication, search, access
- Data lifecycle, transformation, variability
- Data integrity, identification, ownership
- Data centric security and access control

Paradigm changing factors
- Big Data properties: 5+1 V’s
- Data aggregation: multi-domain, multi-format, variability, linkage, referral integrity
- Policy granularity: variety and complex structure, for their access control processing
- Virtualization: Can improve security of data processing environment but cannot solve data security “in rest”
- Mobility of the different components of the typical data infrastructure: data, sensors or data source, data consumer
Research direction

– Information on Infrastructure
– Control & programmability of Infrastructure
– Virtualization
– Networked data processing
– Sustainability & Complexity

Events on the horizon

– I4DW & DSRC
  • Launch Nov 13
– PIRE & OpenScienceDataCloud.org
  • Workshop June 2014 @ UvA
– Research Data Alliance
  • Conference in Amsterdam Sept 2014
Announcement June 2014
PIRE Workshop Amsterdam

• OpenScienceDataCloud.org
• PIRE Fellowship Application (+/- 15)
• The OSDC PIRE Program is six to eight week fully funded fellowship for US graduate student researchers with an information technology background.
• Format:
  • 1 week tutorials and hands on training
  • 2 months research at a participating institute
  • Results in science/tools and papers/posters/
1000 Genomes Project
Human sequence data from populations around the world with the goal of cataloging human genetic variation.
Total Size: 383.5TB  Categories: genomics, biology

ASTER
ASTER Level-1B Registered Radiance at the Sensor
Total Size: 12.7TB  Categories: earth science

Complete Genomics Public Data
Whole human genome sequence data sets provided by Complete Genomics, containing 69 standard, non-
diseased samples as well as two matched tumor and normal sample pairs.
Total Size: 47.1TB  Categories: genomics, biology

Earth Observing-1 Mission
Data gathered by the Advanced Land Imager (ALI) Hyperspectral Imager (Hyperion) instruments on NASA's Earth
Observing-1 Mission (EO-1) satellite.
Total Size: 45.2TB  Categories: earth science, satellite imagery

City of Chicago Public Datasets
Data set from the City of Chicago Data Portal in JSON format for tabular data and the raw files for "blob" data.
Total Size: 9.7GB  Categories: social science

EMDataBank
Unified Data Resource for 3-Dimensional Electron Microscopy
Total Size: 91.3GB  Categories: biology

Enron Emails
Data sets based on the original Enron emails released to the public by the Federal Energy Regulatory
Commission as part of their investigation.
Total Size: 155.9GB  Categories: social science

FlyBase
FlyBase is the leading database and web portal for genetic and genomic information on the fruit fly Drosophila
melanogaster and related fly species.
Total Size: 614.8GB  Categories: biology, genomics
Research Data Alliance

- https://rd-alliance.org

- The Research Data Alliance implements the technology, practice, and connections that make Data Work across barriers.

- The Research Data Alliance aims to accelerate and facilitate research data sharing and exchange.
  - Working groups and interest groups
  - Joining groups and attendance at the twice-yearly plenary meetings is open.

- Plenary Sep 2014 hosted by the Netherlands - Amsterdam
  - Conference Management Team (CMT) Chair: Peter Doorn (DANS)
  - Program Committee (PC): chair Cees de Laat (UvA)
  - Satellite Events Committee (SEC): Jeroen Rombouts (TUD)
Questions?

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