

MPTCP in a Multipath WAN Fabric using OpenFlow

Ronald van der Pol

Ronald.vanderPol@SURFnet.nl

Partners

- *Caltech (booth #809)*
 - Artur Barczyk, Michael Bredel, Azher Mughal, Ramiro Voicu, Harvey Newman
- *iCAIR (booth #501)*
 - Jim Chen, Joe Mambretti
- *SARA (booth #2333)*
 - Freek Dijkstra, Sander Boele, Ronald van der Pol
- *SURFnet (booth #2333)*
 - Gerben van Malenstein

iCAIR



SCinet

ACE
CWAVE

STARLIGHTSM
The Optical STAR TAPSM

PICA



NL Light

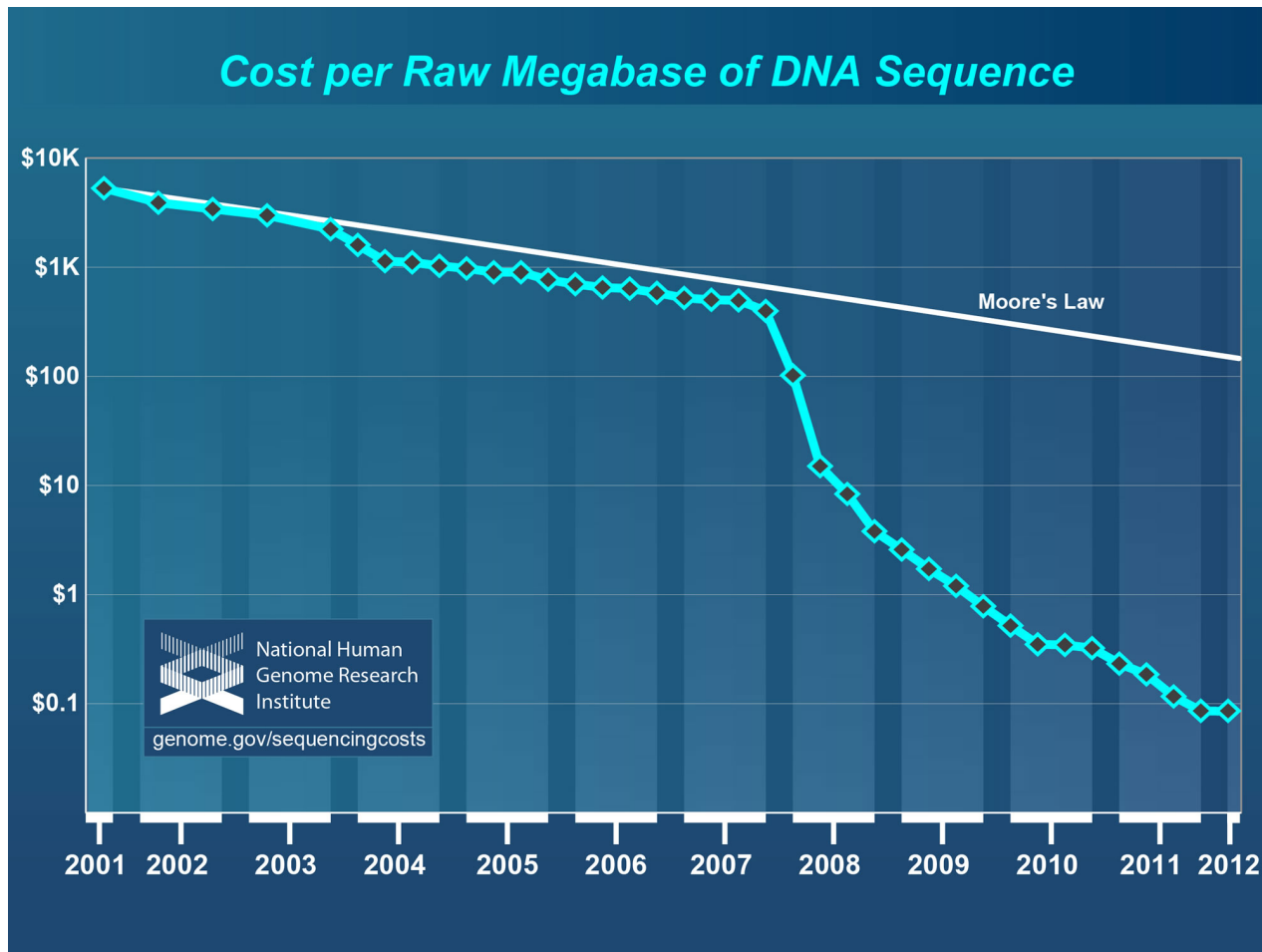
Outline

- Why multipathing?
- MPTCP
- OpenFlow multipathing application
- Demonstration topology
- Results

Why multipathing?

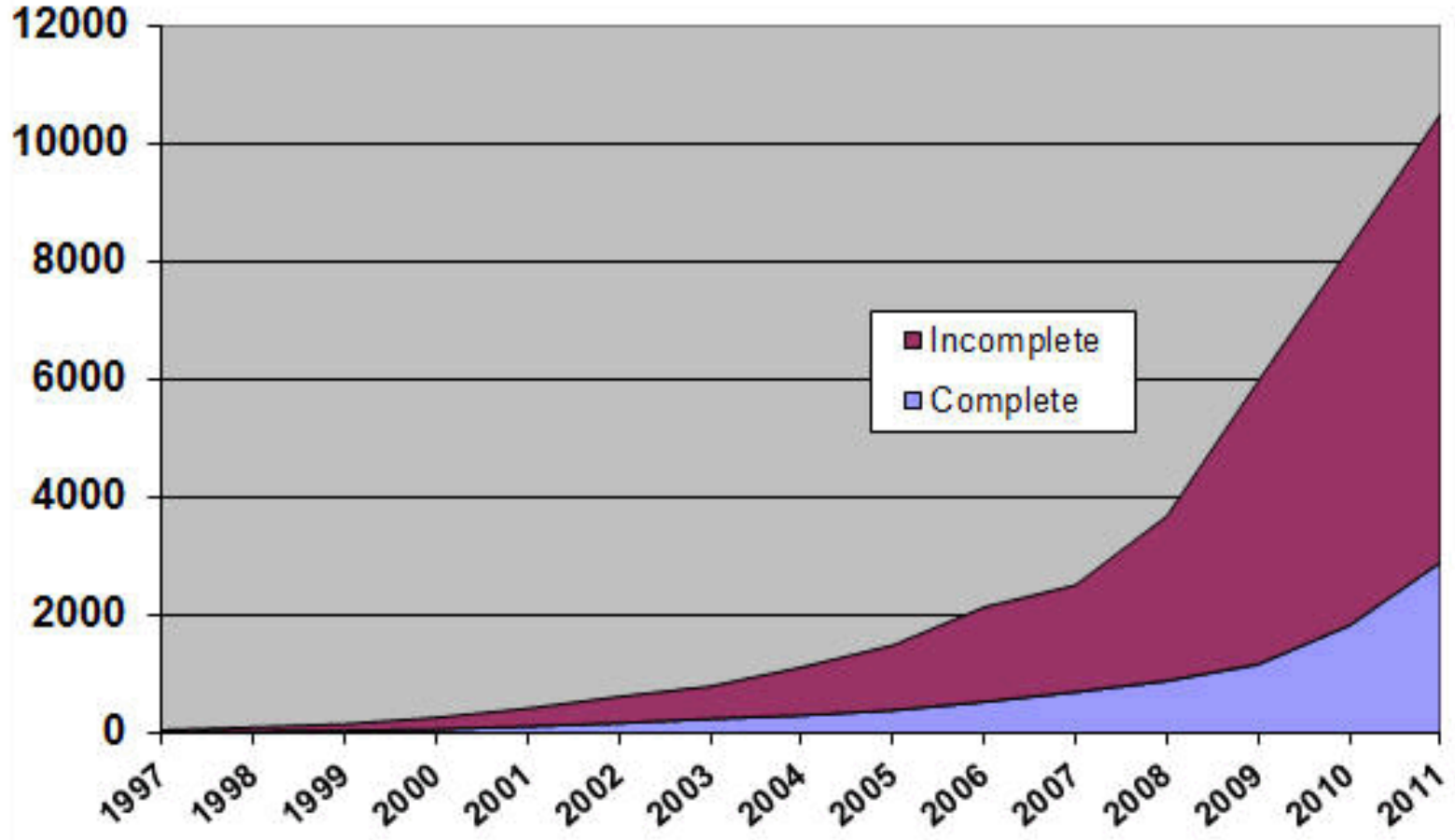
- Data sets are growing exponentially
- Copying these data sets in reasonable time between sites requires a lot of bandwidth
- We are reaching the theoretical limit of fiber throughput
- Next step, like RAID (multiple disks) & multi-core
- Use all available bandwidth in the network
- Use multiple paths simultaneously

DNA Sequencing Price Drop



DNA Sequencing Projects on GOLD

Genomes Online Database



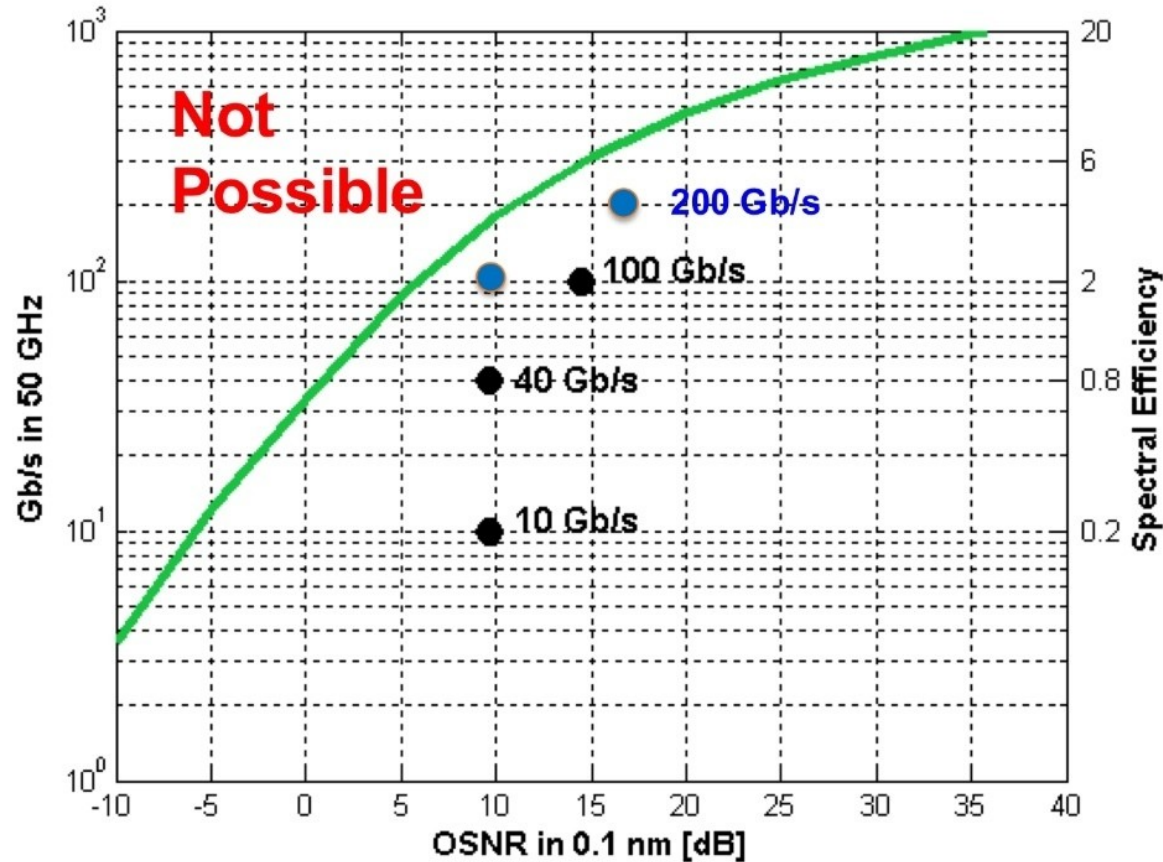
Fast Servers

- Servers can easily fill a 10GE interface
 - Recent servers can fill a 40GE interface
- 40 Gbit/s or 100 Gbit/s end-to-end not always available (e.g. transatlantic) or too costly
- Using multiple links simultaneously is the only option to get high end-to-end throughput

Shannon Limit on Fiber

- Modern modulation technologies and transponders reach the theoretical limit of fiber
- Tradeoff between:
 - Bandwidth
 - Reach
 - Cost
- Either high bandwidth with short reach or lower bandwidth with longer reach

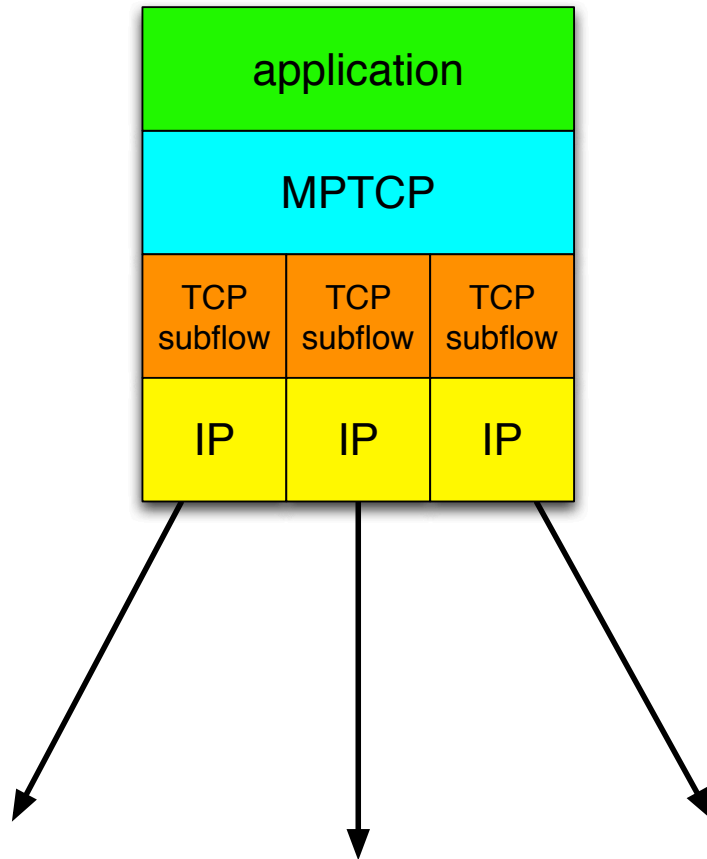
Reaching the Shannon Limit



Multipath TCP (MPTCP)

- Load balancing options:
 - Within the network with e.g. Equal Cost Multipath (ECMP) routing
 - At the end hosts with e.g. MPTCP
- ECMP is hash based, not optimal for small flows
- MPTCP also works for 1 or a few flows
 - MPTCP can also handle links of different speed

MPTCP Stack

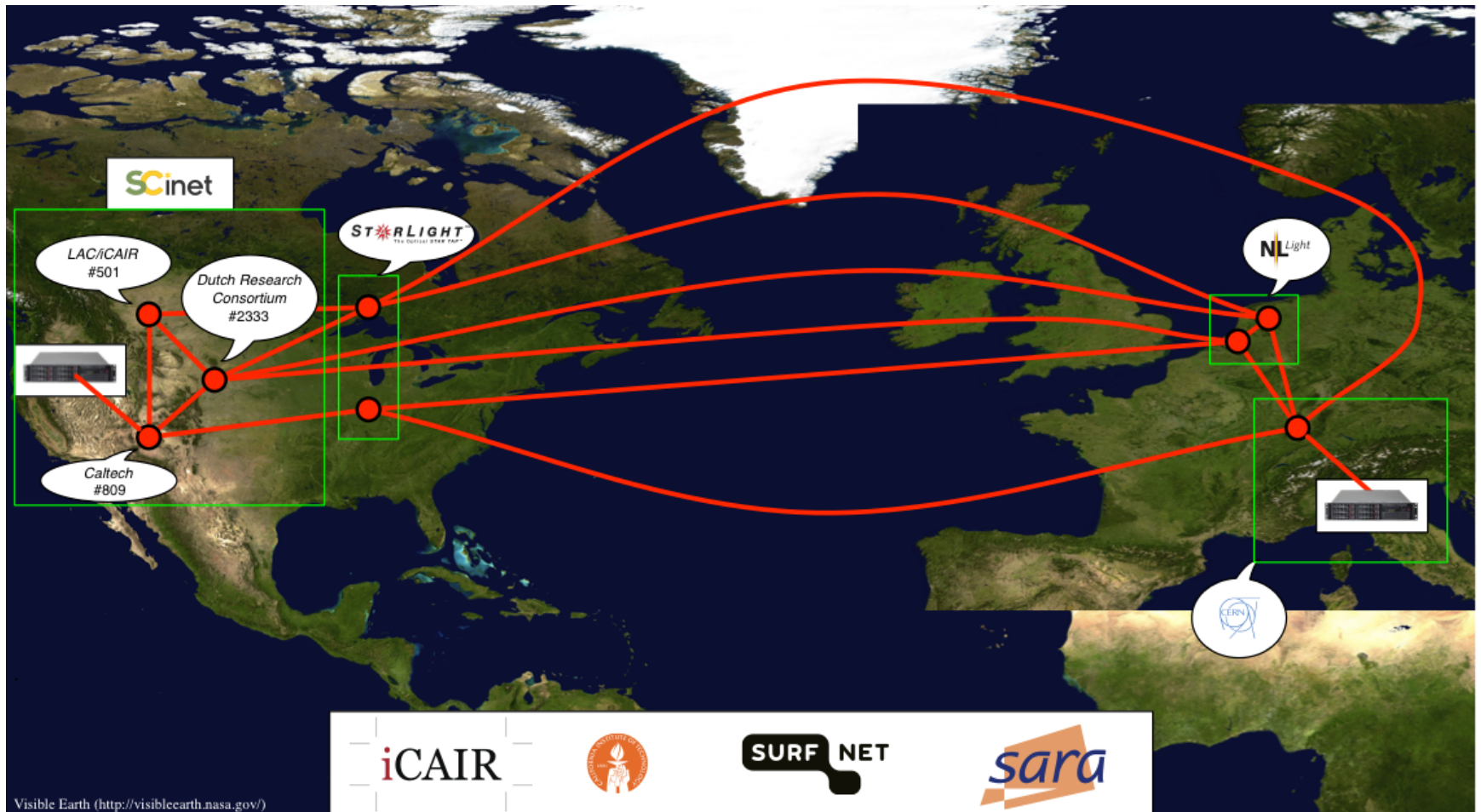


- Normal socket API (no need to change applications)
- MPTCP splits byte stream from application and sends them across multiple subflows
- Each subflow is a normal TCP session to the network
- MPTCP does not setup path, it used paths that are available on a multihomed server

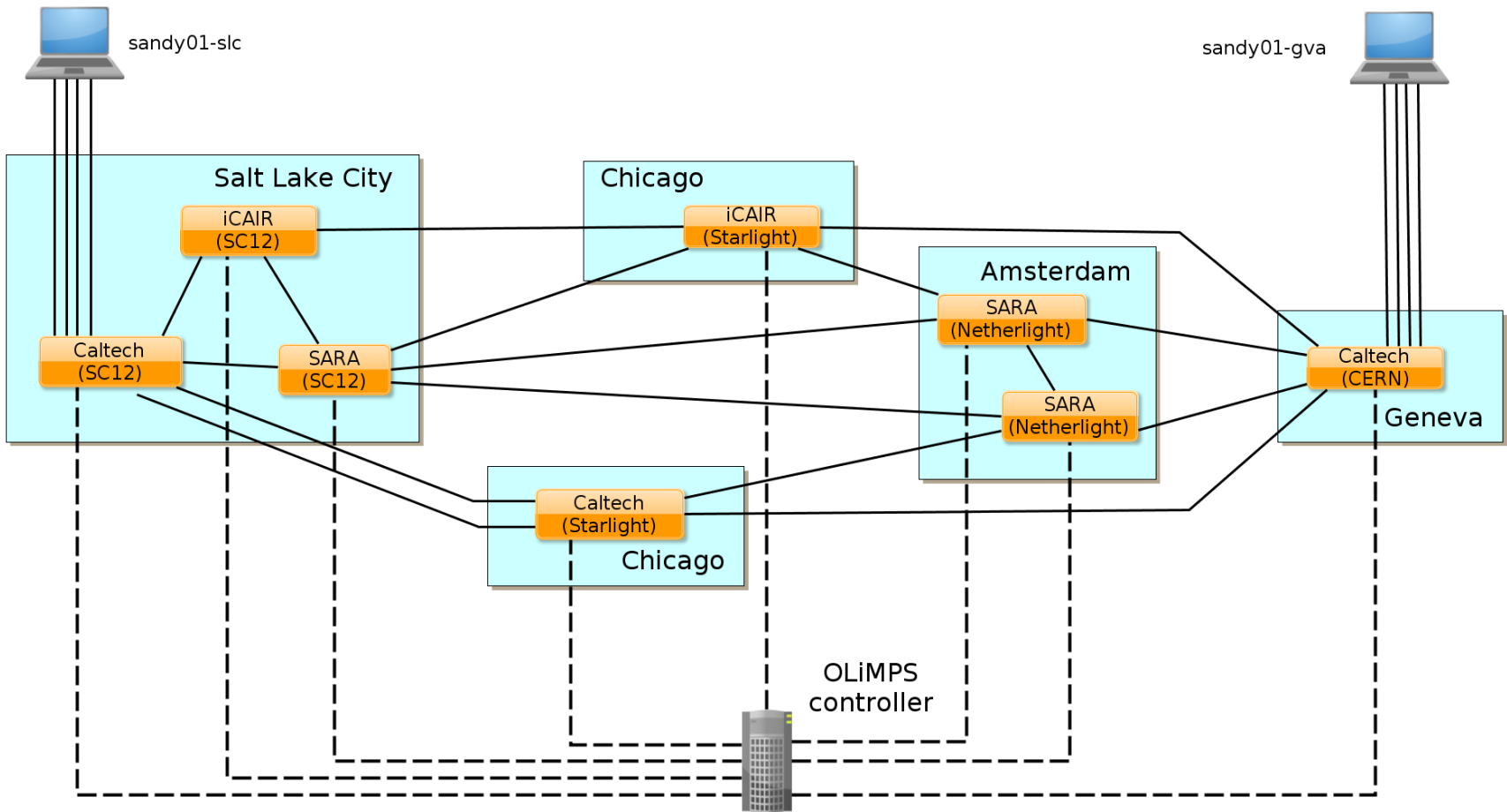
MPTCP Tasks

- Path management
 - Detect and use multiple paths
- Packet scheduling
 - Split byte stream across subflows
- Congestion control
 - Each subflow uses normal TCP congestion control
 - MPTCP automatically moves traffic away from congested subflows to less congested subflows

Intercontinental OpenFlow Network



SC12 Demonstration Topology

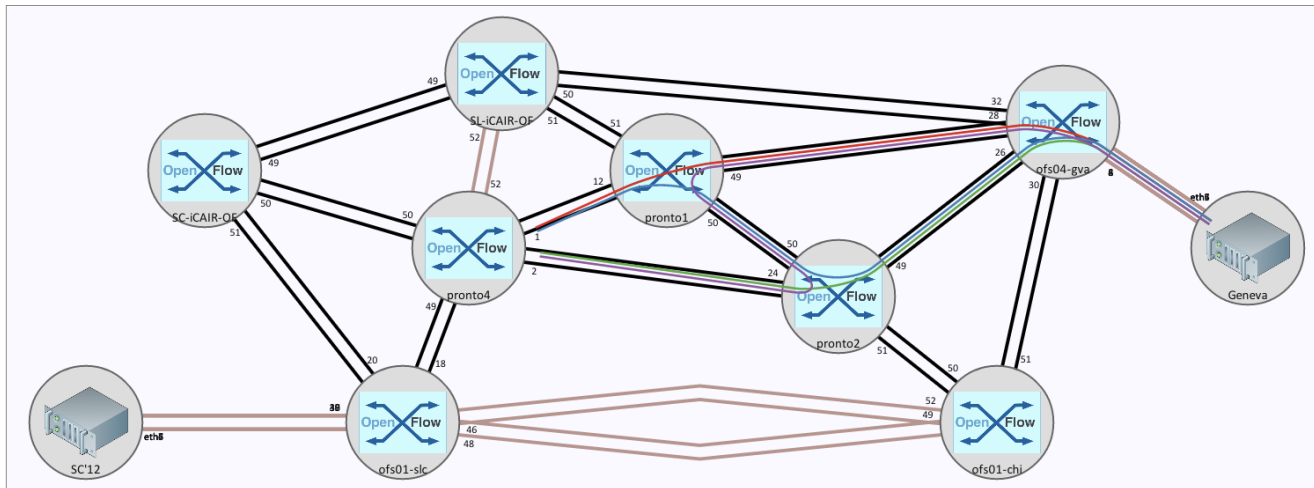


OpenFlow Link Layer Multipath Switching (OLiMPS)

- Floodlight sends & receives LLDP packets on all OpenFlow switch ports → topology
- OLiMPS calculates multiple link disjoint paths from source switch to destination switch
- Flows are mapped to these paths
 - Initially in round robin manner
 - Later e.g. based on real time network load

Results (GVA-AMS)

Multipath TCP streaming from Geneva to Chicago over OpenFlow controlled paths



Demo partners:

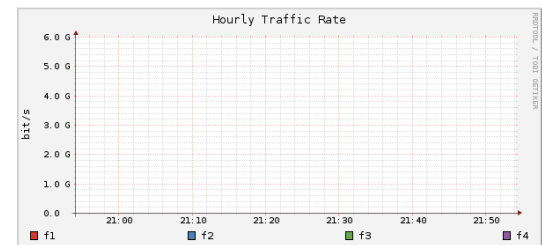
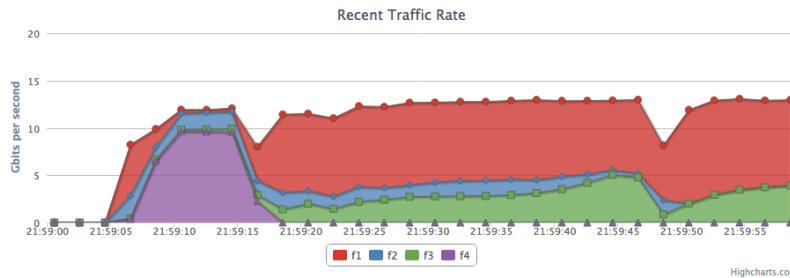


iCAIR

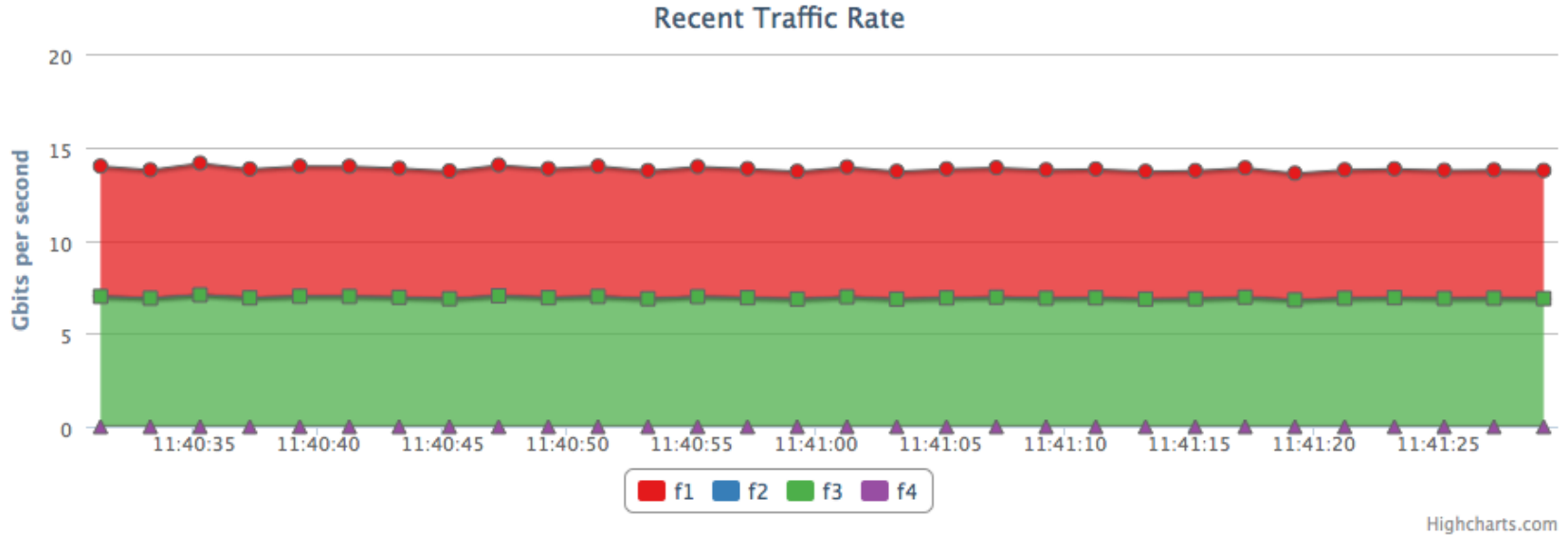


Current Traffic Rate	
Name	Bandwidth
f1	9.04 Gb/s
f2	0.00 b/s
f3	3.91 Gb/s
f4	0.00 b/s
12.95 Gb/s	

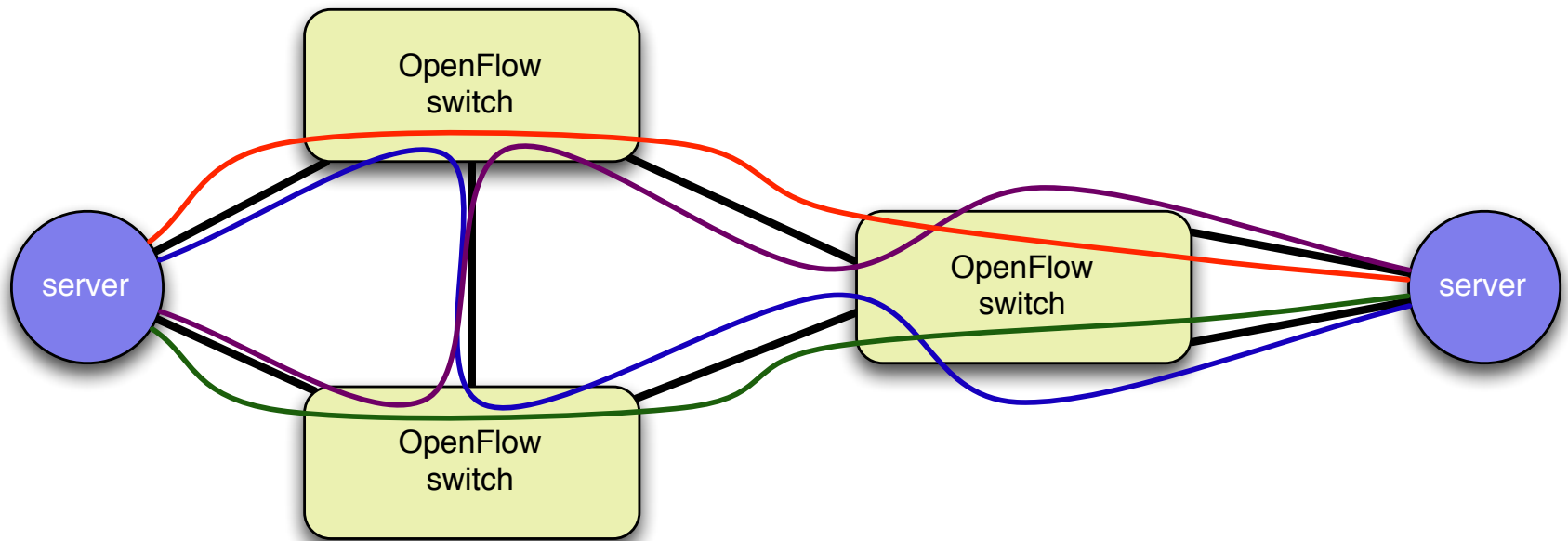
Mode: [view](#) / [debug](#) / [reset](#)



Streaming GVA-AMS (steady state)



OpenFlow Paths



More Information

- <http://mptcp.info.ucl.ac.be/pmwiki.php?n=Main.HomePage>
- <http://www.rvdp.org/publications/SRS-2012-multipath.pdf>
- <http://www.rvdp.org/publications/TNC2013-MPTCP.pdf>

Thank You!

Ronald.vanderPol@SURFnet.nl