ITMO UNIVERSITY (www.ifmo.ru)
S.Petersburg, Russia

Transfer of large volume data over Internet with parallel data links and SDN

V. Grudinin, A. Kairkanov, S. Khoruzhnikov, O. Sadov, and A. Shevel (reporter)

Overview

- Intro
- Existing tools
- Idea
- Testbed
- Measurement procedures
- Initial results
- Conclusion

Scientific sources of Large Volume of Data

- **Scientific experiments**
  - http://www.lsst.org - Large Synoptic Survey Telescope
    - ~ 10 PB/year
  - https://www.skatelescope.org/ - Square Kilometre Telescope Array
    - ~ 300-1500 PB/year
  - http://www.cern.ch — CERN
    - ~20-40 PB/year (coming FAIR ~ the same)
  - http://www.iter.org - International Thermonuclear Experimental Reactor
    - ~1 PB/year
  - http://www.cta-observatory.org/ - CTA - The Cherenkov Telescope Array
    - ~20 PB/year

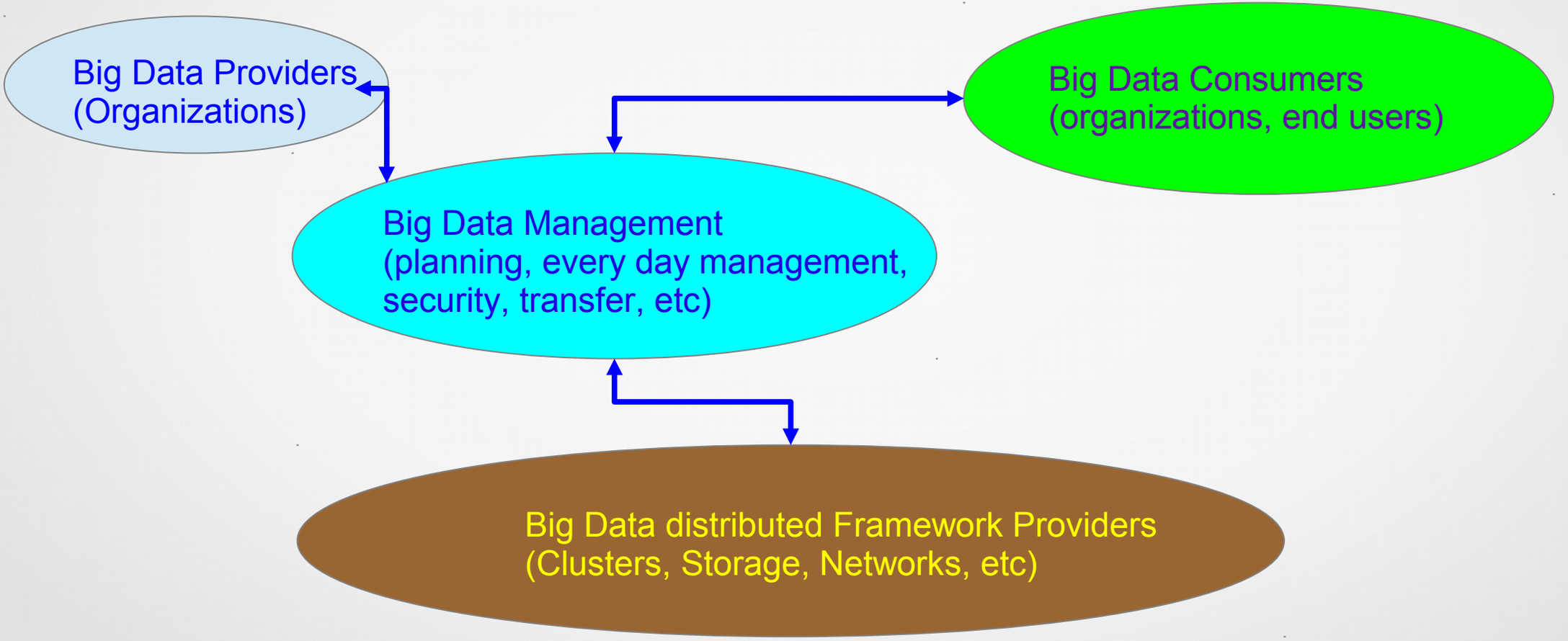
Other generators of Large Volume of Data

- NIST - http://bigdatawg.nist.gov/usecases.php - many examples
  - Almost any communication company.
  - Video from observation video cameras.
  - Data Preservation.

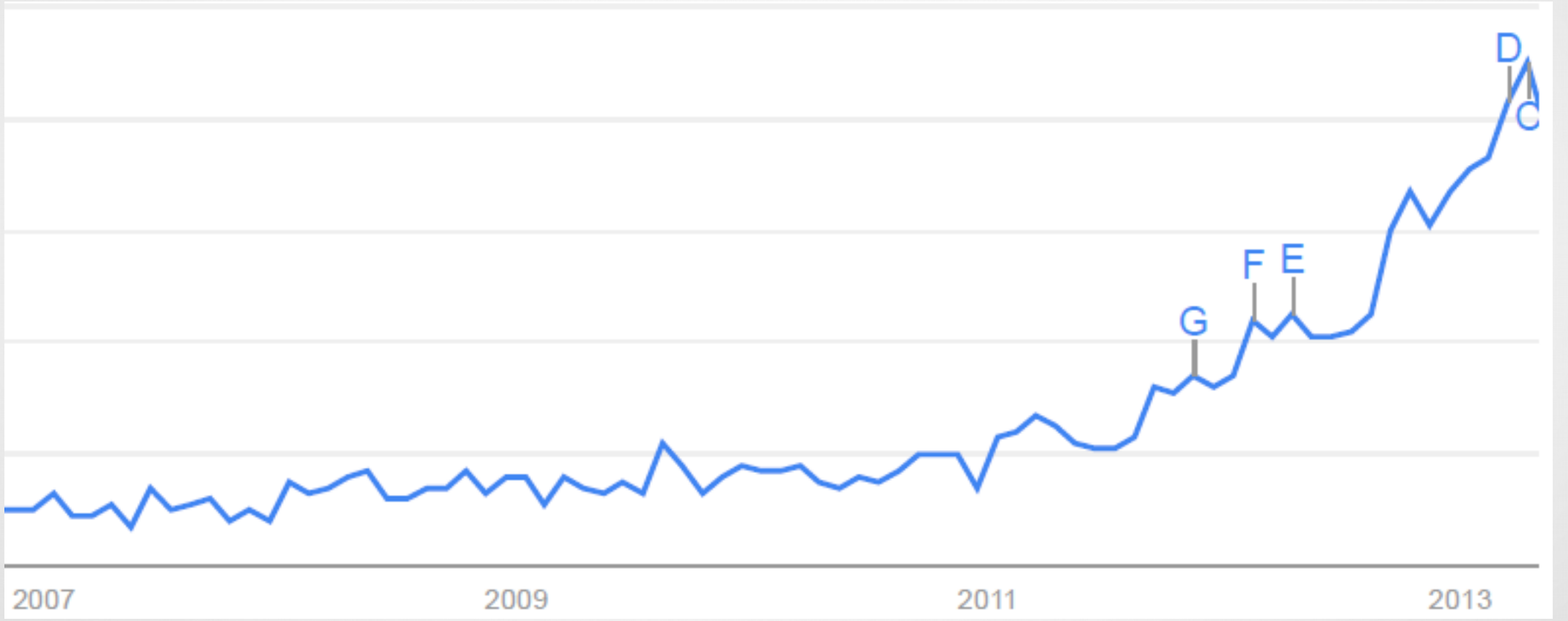
Big Data features and required operations

- Triple «V»
  - Velocity
  - Variety
  - Volume
- Operations
  - Storage
  - Process/Analysis
  - Visualization
  - Transfer

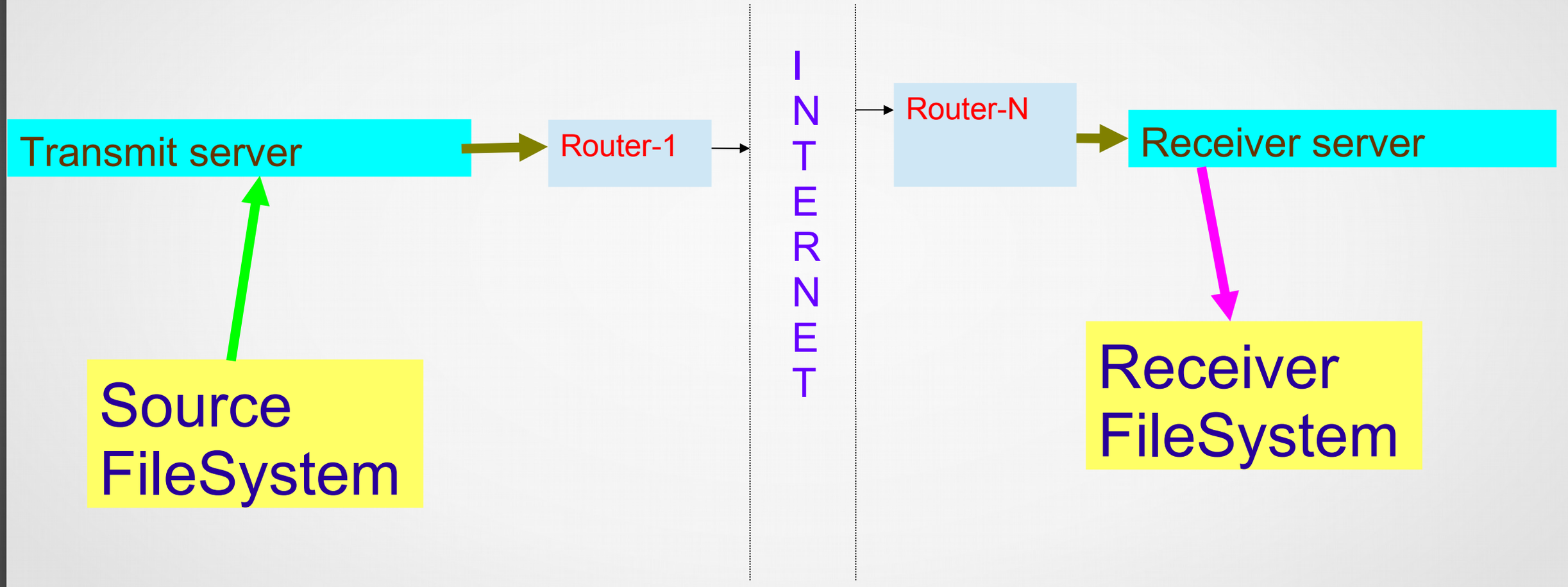
Big Data Architecture



Google trend for «the big data»



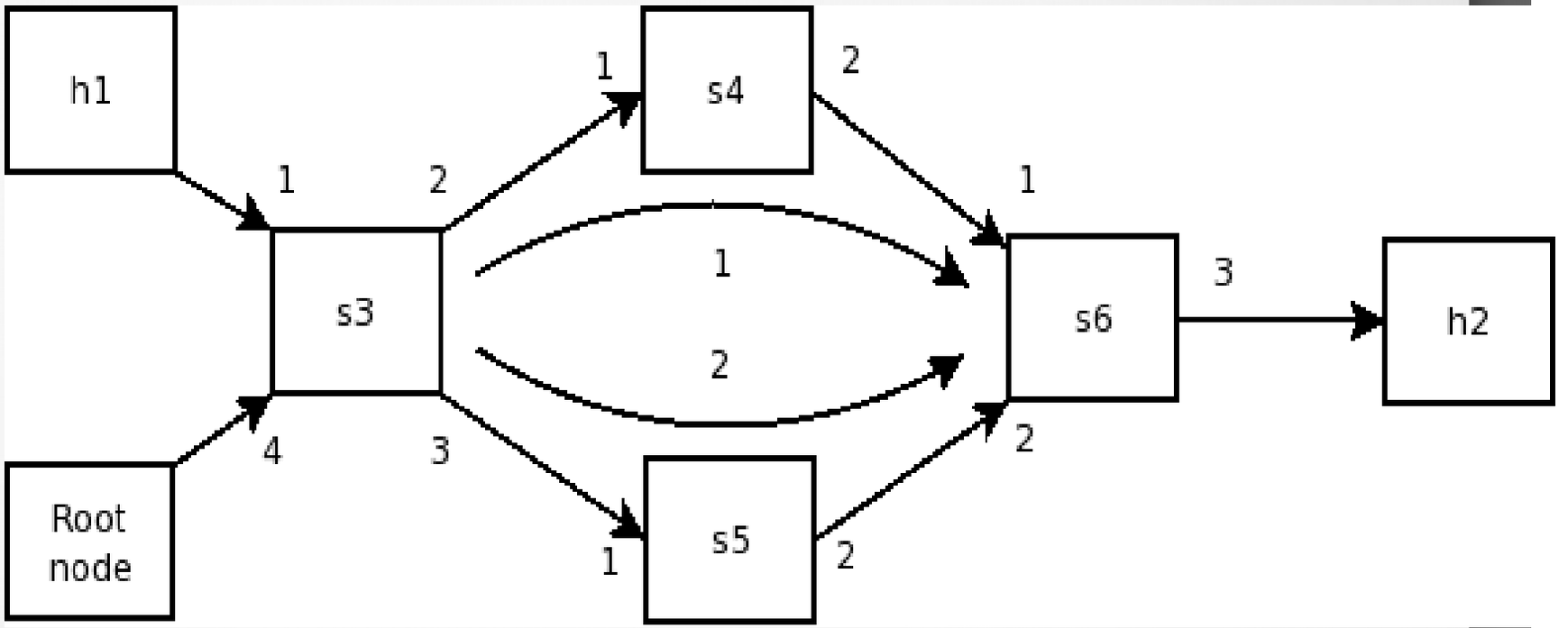
Data transfer process



Tested low level Protocols

1. Link Aggregation Control Protocol - LACP
2. Multi Path TCP - MPTCP
3. Other possibilities?

1 and 2 are quite good for Data Centers

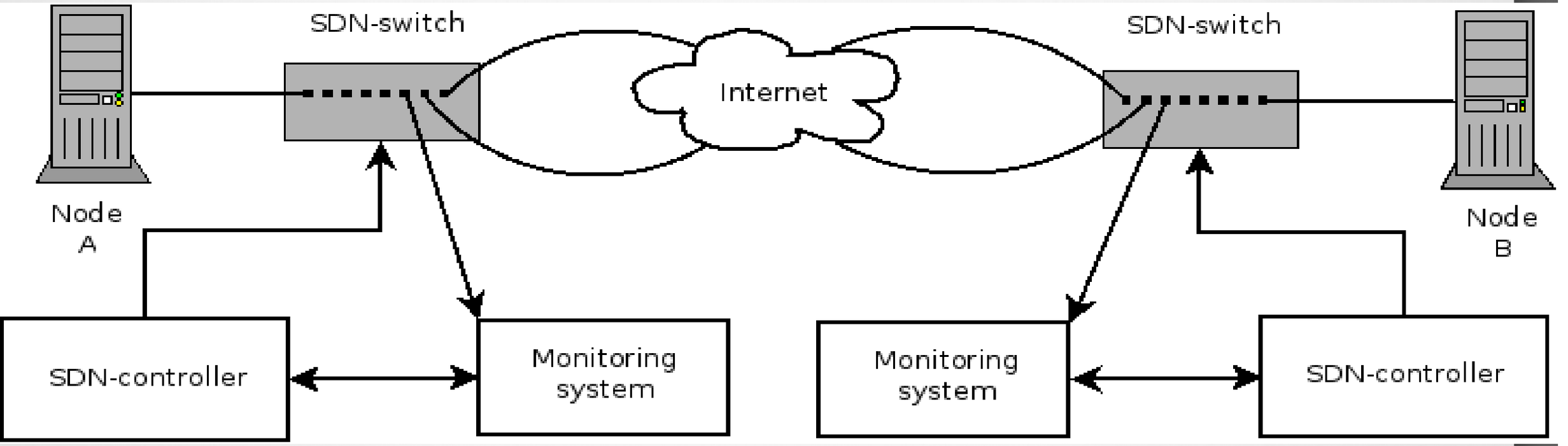


Research

Our goal is to study Big Data transfer over long distant network over parallel data links with SDN approach (protocol Openflow).

Multi Data Links with OpenFlow to implement fine tuning in according to the data links dynamic status (peculiarity of BigData).

The architecture under development



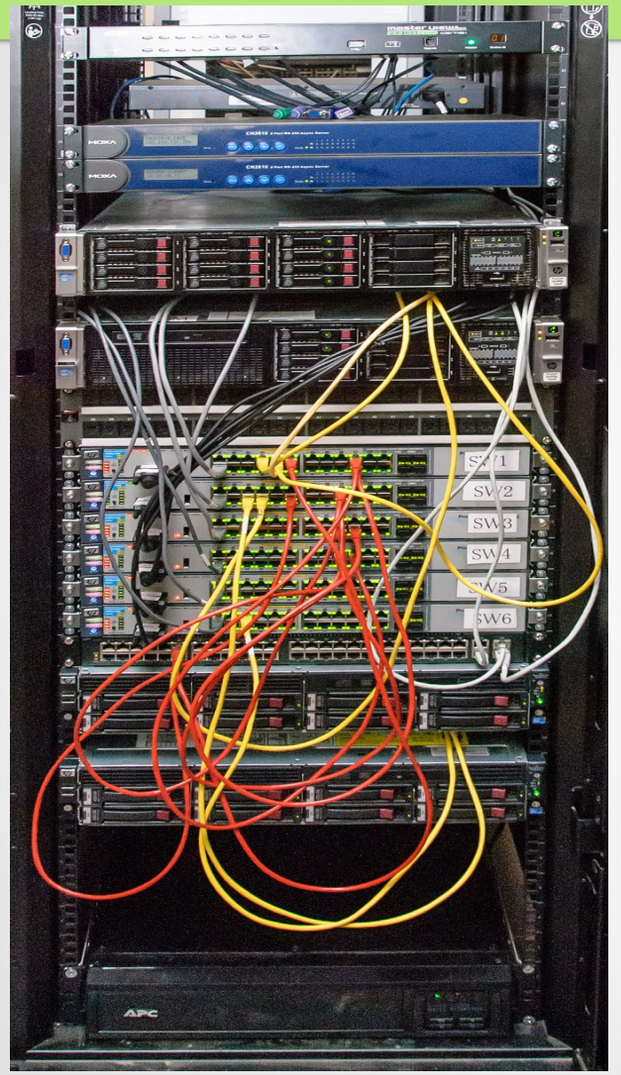
Evaluation of existing programs

- GridFTP - http://www.globus.org/toolkit/data/gridftp/
- BBFTP - http://doc.in2p3.fr/bbftp/
- FDT - http://monalisa.cern.ch/FDT/
- Other tools, including watching for data links, for example, perfSONAR.

What was done until June 2015

- There have been developed, deployed, performed:
  - Two servers (CPU E5-2650v2 @ 2.60GHz, 96 GB, 100TB) under Scientific Linux 6.6 @ ITMO
  - Remote server with the same configuration
  - Required set of Virtual Machines to test a number of data transfer utilities and OpenStack (Icehouse) as appropriate infrastructure.
  - PerfSonar.
  - The set of appliance scripts to perform measurement runs. All scripts are available at *https://github.com/itmo-infocom/BigData*.
  - Initial measurements.

More Testbed details



- OpenFlow switches – Pica8 3290 and 7 HP 3500-24G-PoE yl

Software:

- OpenFlow software switch based on CPqD/of12softswitch and Open vSwitch
  OpenFlow controllers based on CPqD/nox12oflib, NOX classic and NOX
- OpenFlow network emulator Mininet
- A range of other virtualization tools

Measurement logs (catalogs)

- Log directory with name <utility_name>-<source_host>-<target_host>-<YYYY-MM-DD>.
- Abstraction: start date/time, end date/time, average file size, total data transfer size, type of completion, source directory with data, etc.
- Log_file: all messages generated by data transfer utility during measurement.
- Comments: possible comments from person who does measurement.
- Sosreport output: (the utility copies whole directory tree /proc).

perfSONAR-PS bandwi...

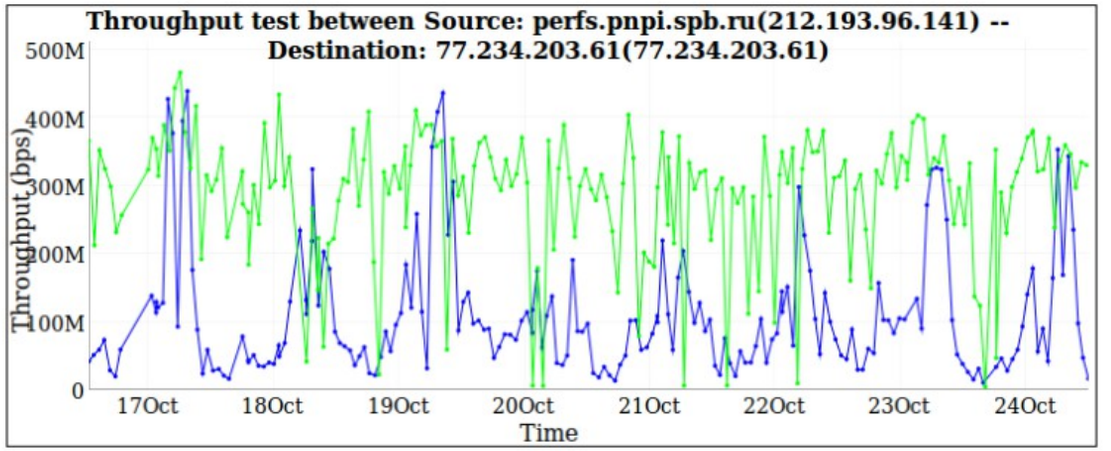
212.193.96.141/serviceTest/bandwidthGraph.cgi?url=http://localhost:8085/perfSONAR_PS/

Google

perfSONAR BWCTL Graph

perfSONAR

Throughput test between Source: perfs.pnpi.spb.ru(212.193.96.141) -- Destination: 77.234.203.61(77.234.203.61)



Graph Key

- Src-Dst throughput
- Dst-Src throughput

<- 1 month

1 month ->

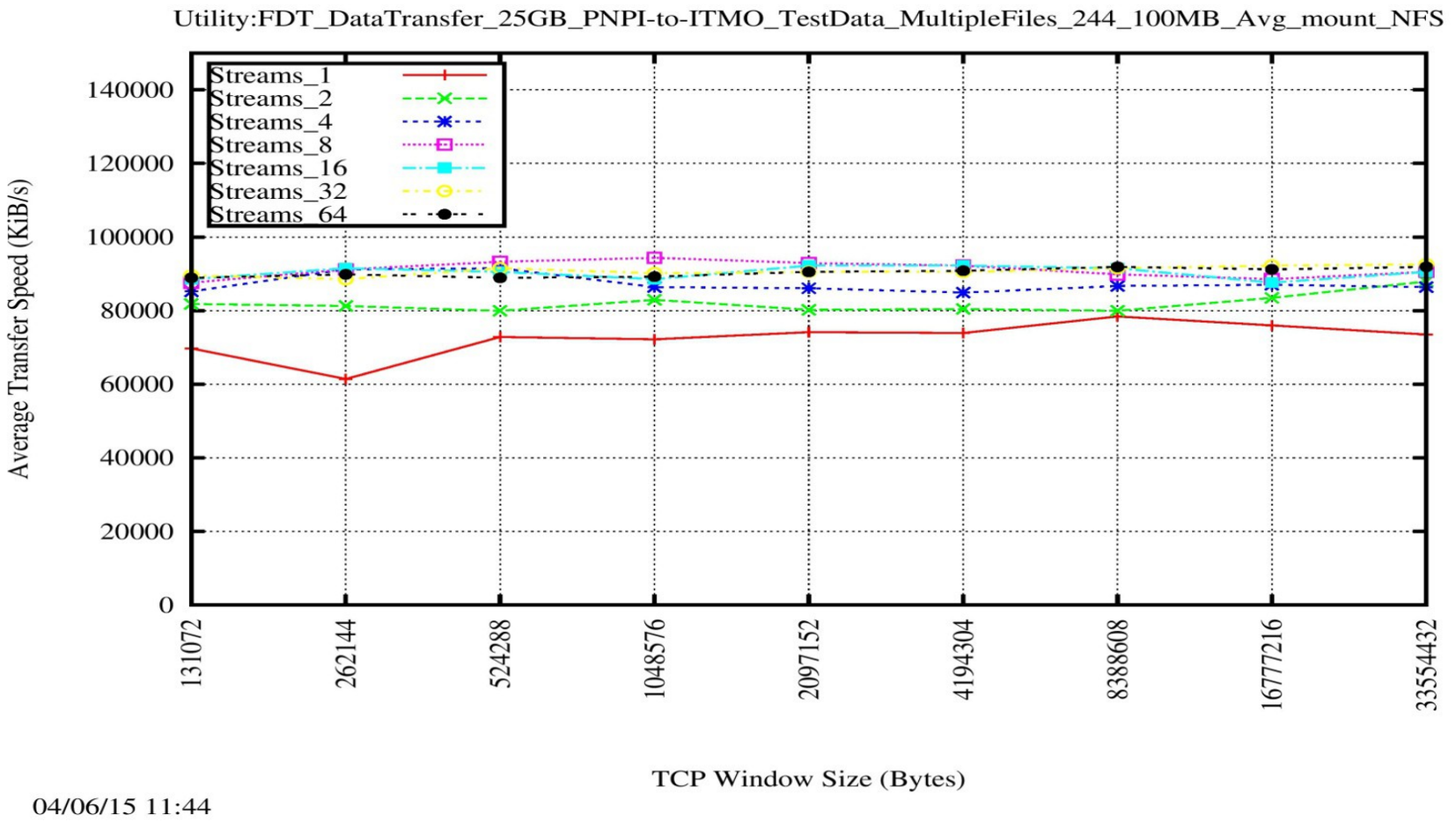
Timezone: GMT+0400 (MSK)

Direction	Max throughput(bps)	Mean throughput(bps)	Min throughput(bps)
Src-Dst	439.13M	109.76M	10.43M
Dst-Src	466.29M	288.91M	4.47M

Show/Hide Link

For help on how to zoom in, zoom out, use the menu options and interact with the graph, click here

Data Transfer from one point in Internet to another one



Conclusion

- Infrastructure consisting of
  - the testbed;
  - the set of scripts to perform the measurements;
  - the system of logs
  - has been implemented.
- This infrastructure will be used for future works.
- Also this infrastructure might be used to compare any other existing tools or future data transfer methods with already been measured.