

# Big Data transfer over Internet

*Sergey Khoruzhnikov, Vladimir Grudinin, Oleg Sadov, Andrey Shevel, Arsen Kairkanov*

ITMO University, St.Petersburg (Russia)  
HEPIX Spring 2014, Annecy



Presenter: Andrey Shevel

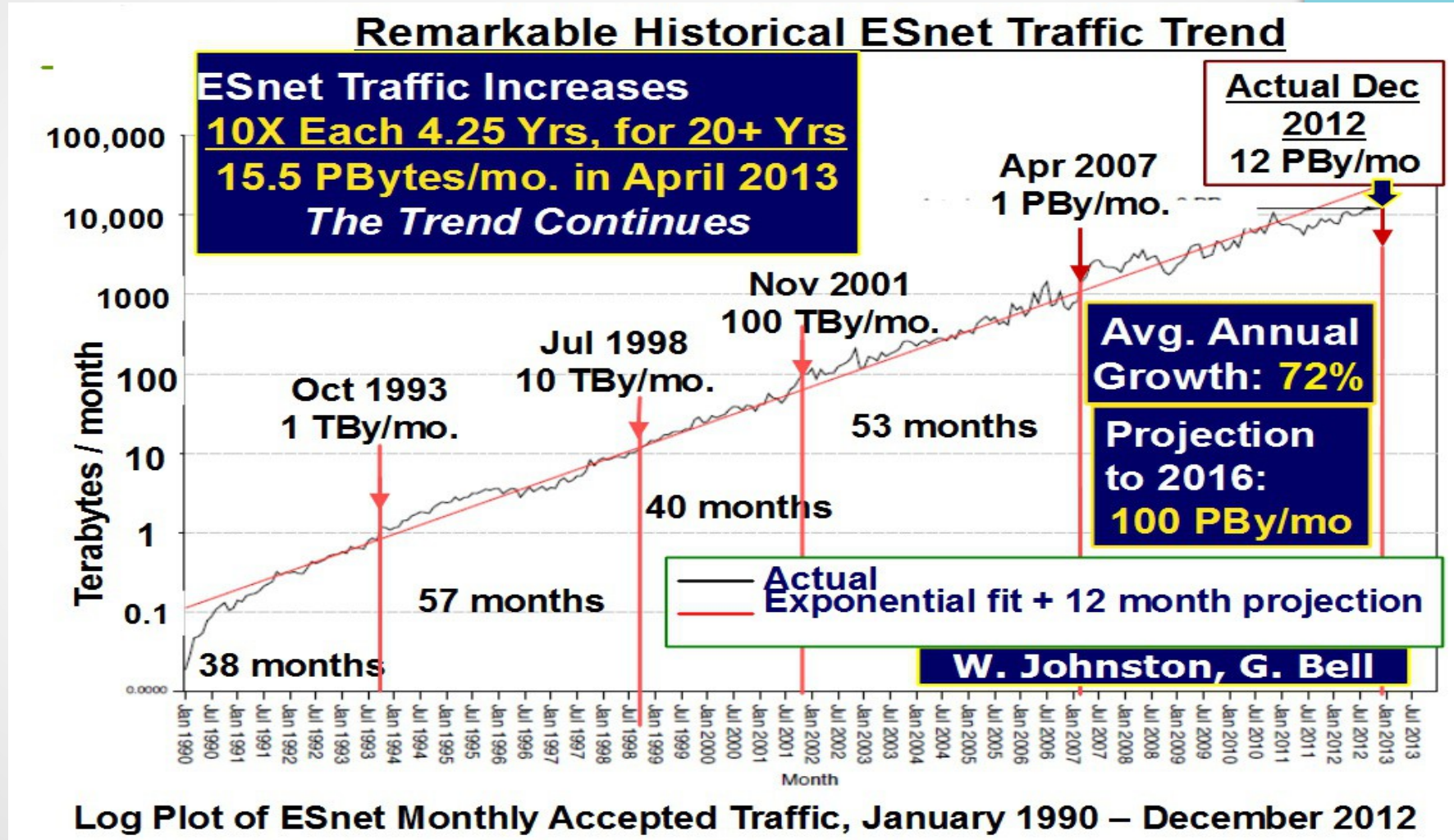
# Outline

- Sources of the Big Data.
- Ecosystem of the Big Data.
- Technology of the Big Data transfer.
- Our recently started research.

# Scientific sources of Big Data

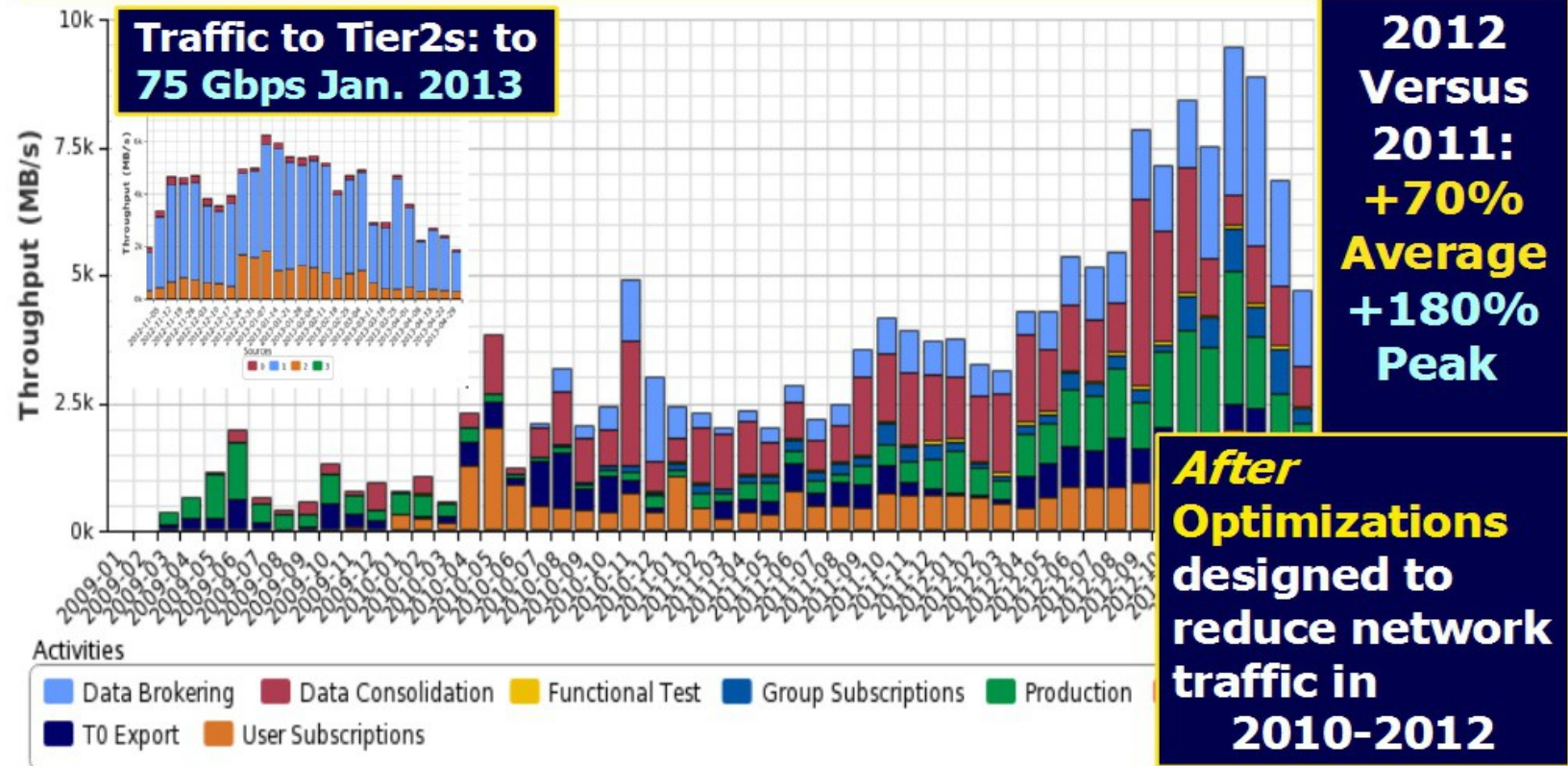
- **Scientific experimental installations**

- <http://www.lsst.org> - Large Synoptic Survey Telescope
  - **15 TB per night (may be 10 PB/year)**
- <https://www.skatelescope.org/> - Square Kilometre Array
  - **300-1500 PB/year**
- <http://www.cern.ch> — CERN
  - **~20PB/year (FAIR ~ 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**



## ATLAS Data Flow: 2009- April 2013

**2012-13: >50 Gbps Average, 112 Gbps Peak**  
**171 Petabytes Transferred During 2012**



# Which are the future demands

Paul Sheldon & Alan Tackett  
Vanderbilt University



## 10 Year Projections

### ■ Requirements in 10 years? Look 10 years back...

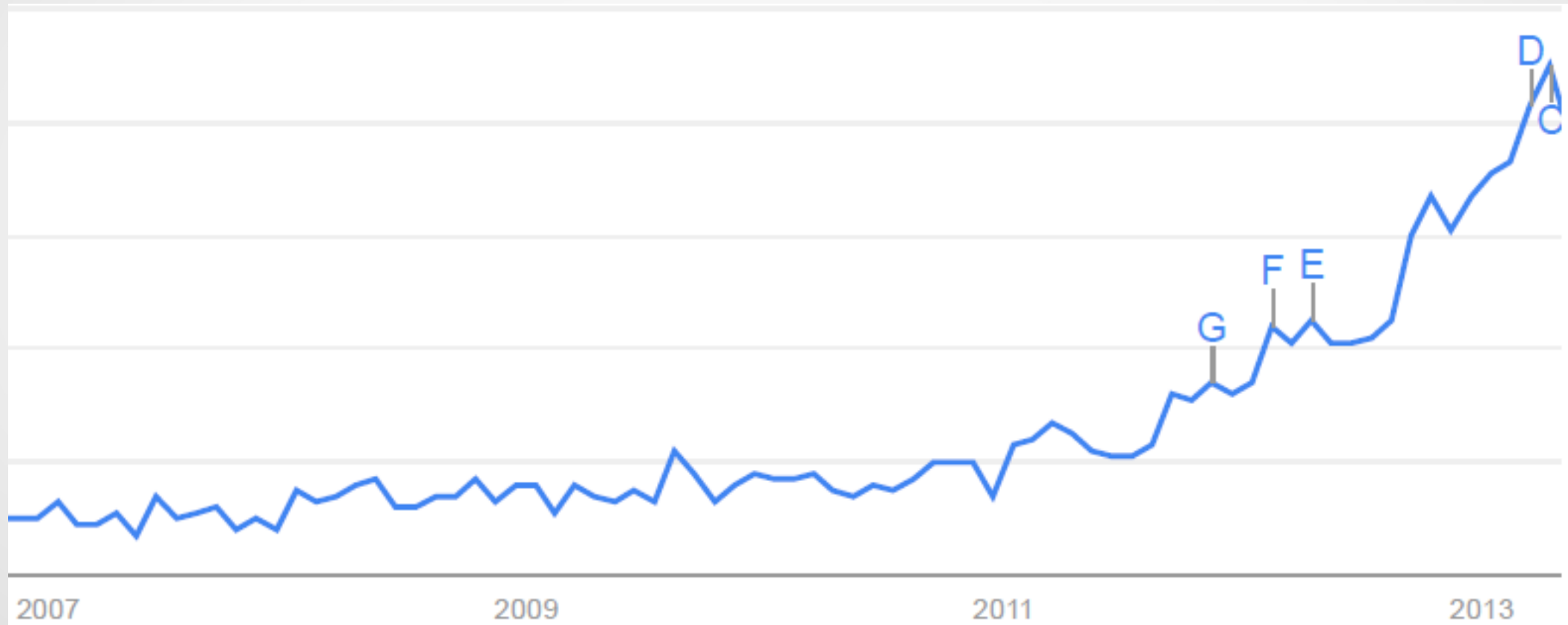
Source: Ian Fisk and Jim Shank

Metric	<u>Tevatron(2003)</u>	<u>LHC(2012)</u>
Remote Computing Capacity	15kHS06 (DZero Estimated)	450kHS06 (CMS)
User Jobs launched per day	10k per day	200-300k jobs per day
Disk Capacity per experiment in PB	0.5PB	60PB
Data on Tape per experiment	400TB	70PB
MC Processing Capacity per month for Full Simulation	3M	300M
Data Served from <u>dCache</u> at FNAL per day	25TB per day	10PB per day
Wide Area networking from host lab	200Mb/s	20000Mb/s
Inter VO transfer volume per day	6TB ( <u>DZero</u> SAM)	546TB (ATLAS)

# Other Big Data providers and what was done to study

- NIST - <http://bigdatawg.nist.gov/usecases.php> - many examples and use cases
  - Any communication company.
  - Video streams from tens of thousands of watching cameras.
  - Data Preservation.

# Google trend for «the big data»

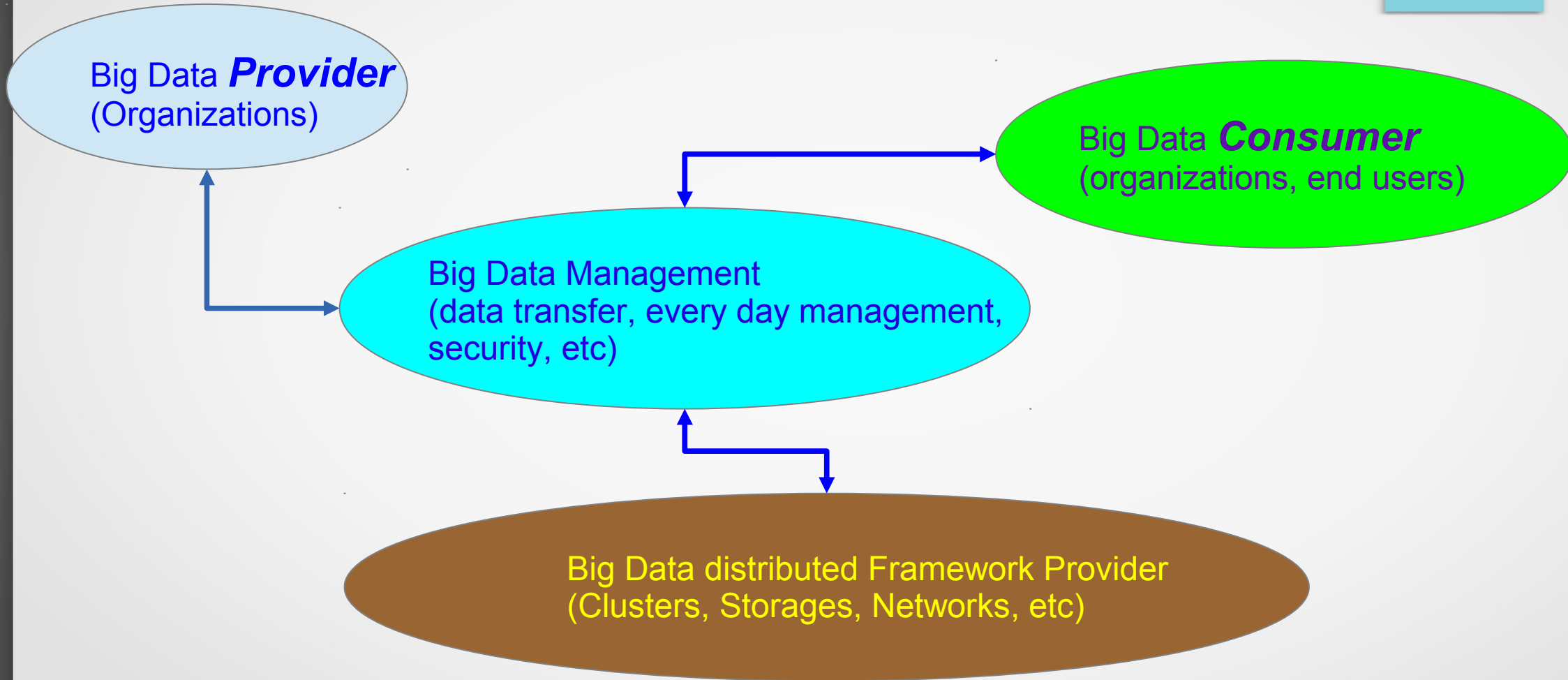




# Big Data

- Big Data have the set of features: *triple V*
  - Data flow Velocity
  - Data Volume
  - Data Variety

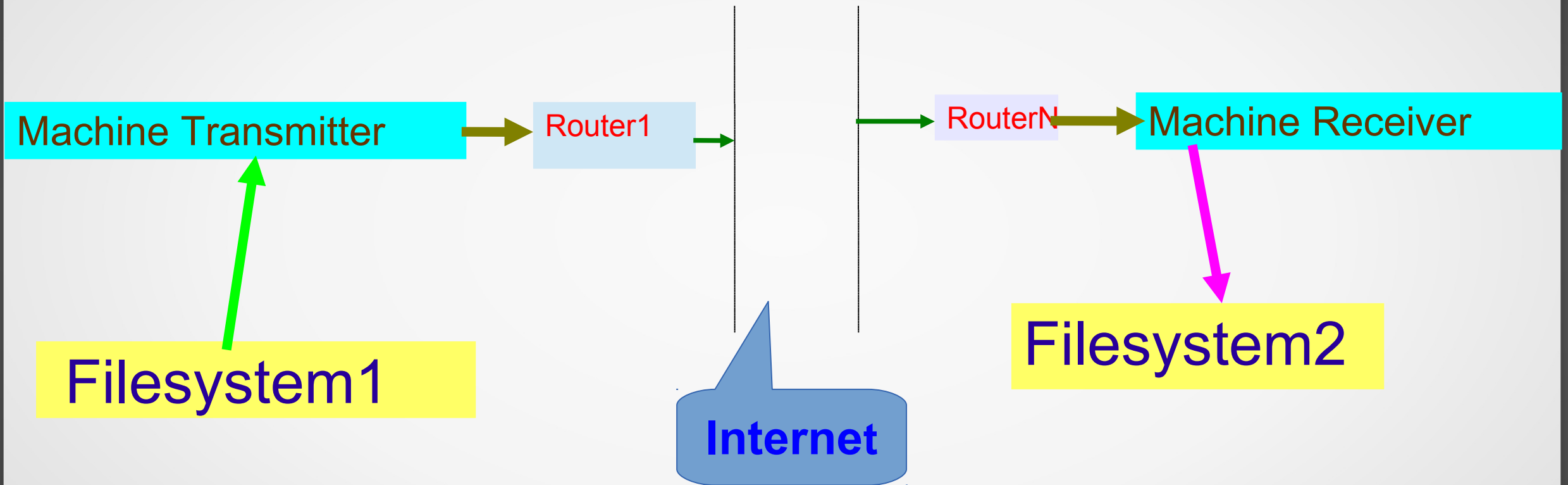
# Big Data ecosystem



# Peculiarity of the Big Data transfer

- Big Data transfer might consume many hours or days.
- The situation in data link might be changed: RTT, % of lost network packages, data link bandwidth).
- Finally, it might occur the interruption (*minutes?*, *hours?*, *days?*) in operation of data link .
- Obviously it is useful to have access to two or more independent data links.

# Process of the Data transfer



# Research topic at ITMO University: the transfer of Big Data

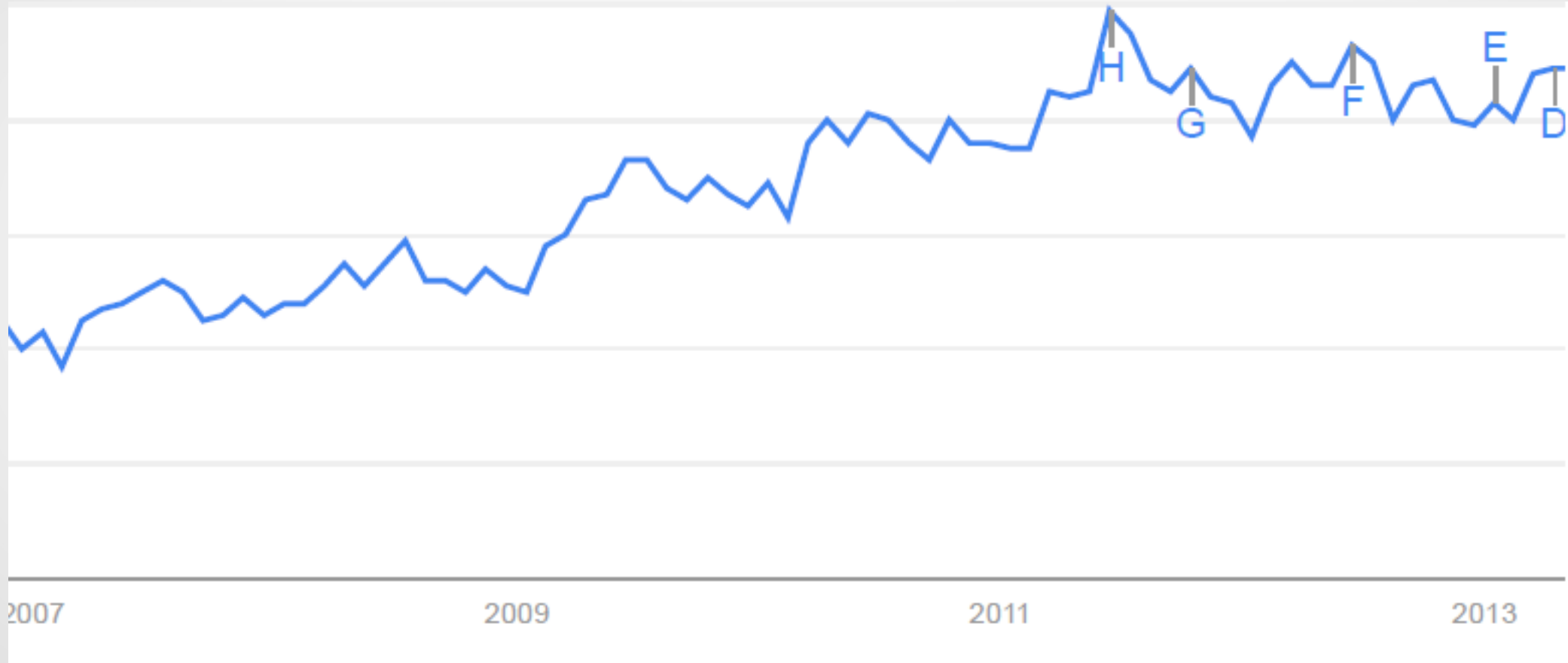
In laboratory of network technologies <http://sdn.ifmo.ru/> at ITMO University <http://www.ifmo.ru/> the new research «Big Data transfer over Internet» has been formed .

- It is planned to implement the special testbed (100 TB of disk storage + server 96 GB of main memory under OS RedHat/ScientificLinux on each side).
  - Comparative study of the existing tools of the data transfer (testing and measurements).
  - To use the testbed as instrument to compare various tools (tracking for the measurements + results).
  - Extended automatic tracking information about measurements is under development.

# Planned measurements

- Local and long distant sites with existing data links (**not** only most advanced links).
- The use of more than one data link in parallel.
- Recently we obtained some experience with Software Defined Networks (SDN) approach (*protocol Openflow*) and now we plan to use it in the Big Data transfer.

# Google trend for SDN



# Technology peculiarities with Big Data transfer

- Still main protocols — stack of TCP/IP.
  - Number of network parameters in Linux (around ½ thousand).
    - Important parameters: e.g. size of block, size of TCP Window, etc.
  - Main method to decrease the transfer time (even over one data link) is using the multi-stream data transfer.



## Testing on the first stage (program tools)

- GridFTP - <http://www.globus.org/toolkit/data/gridftp/>
- BBFTP - <http://doc.in2p3.fr/bbftp/>
- FDT - <http://monalisa.cern.ch/FDT/>
- FTS3 - <http://fts3-service.web.cern.ch/>
- Also technology components to watch the data links status, e.g. perfSONAR.

# Partners (information exchange)

- On Big Data

Laboratory of Information Technology (LIT) <http://lit.jinr.ru/>

@ Joint Institute for Nuclear Research (JINR.ru)

- On Software Defined Networks

- The Application Research Center for Computer Networks at Moscow University <http://arccn.ru/>

- We are starting to collaborate with GENI (<http://www.geni.net/>)

# Ideas to compare the data transfer tools

- Availability.
- API.
- Performance.
- Reliability.
- Operation tracking.
- Ability to predict the time to transfer the data on the base of existing tracking records.
- Required resources: memory, CPU time, etc.
- Others.

# Main goals

- Combining the developed contemporary components and methods with ideas, developments, experience to achieve maximum speed for Big Data transfer on existing links.
- To create the testbed which would be used as place where researchers might compare theirs (new) tools for data transfer with earlier recorded measurement results.
- To suggest the collaboration with ... (suggestions?)
- To invite students from ... (suggestions?)

# What was done until now

- There were deployed:
  - two servers (HP DL380p Gen8 E5-2609, Intel(R) Xeon(R) CPU E5-2640 @2.50GHz, 64 GB) under Scientific Linux 6.5.
  - **OpenStack** (Havana) with appropriate set of Virtual Machines to test a number of mentioned utilities.
  - PerfSonar.
- Well underway the development of the script to perform measurement run.

Test random data generation has been prepared: it is generated the directory with randomly distributed file sizes; it is possible to define average size of files and dispersion (<https://github.com/itmo-infocom/BigData>).

# Questions ?

## Decimal

<b>Value</b>		<b>Metric</b>
1000	KB	kilobyte
$1000^2$	MB	megabyte
$1000^3$	GB	gigabyte
$1000^4$	TB	terabyte
$1000^5$	PB	petabyte
$1000^6$	EB	exabyte
$1000^7$	ZB	zettabyte
$1000^8$	YB	yottabyte