Data Intensive Science Impact on Networks

Eli Dart, Network Engineer

ESnet Network Engineering Group

IEEE Bandwidth Assessment Ad Hoc

December 13, 2011

Outline

Data intensive science – examples

Collaboration and traffic profile

Future landscape

Large Hadron Collider - ATLAS

Large data sets (transfers of tens of terabytes are routine)

Automated data distribution over multiple continents

Large data rates

- ~1 Petabyte per second from the instrument
- Multi-stage trigger farm reduces this to ~200-400MB/sec
- Additional data from event reconstruction
- Large-scale distribution of data to international collaboration
  - 10-40Gbps out to large repositories in Europe, North America, and Asia
  - 5-10Gbps to analysis centers worldwide

This will increase over time as the LHC is upgraded

Genomics

Genome sequencing is in its infancy

Already seeing significant increase in data rates

Increases coming from two directions

- Per-instrument data rate increasing significantly (~10x over 5 years)
- Cost of sequencers plummeting (10x over 5 years)
- Human genome sequencing cost $10,500 in July 2011 from $8.9 million in July 2007 – NYTimes

Wide variety of applications for genomics data as science improves, applications discovered, etc.

Instruments

Many instruments used in basic research are essentially high-resolution digital cameras

The data rates from these instruments are increasing with the capabilities of the instruments

Some instruments in development will be able to collect terabits per second of data

- There are not enough I/O pins on the chip to get all the data out
- On-chip data reduction will be necessary

Transfer or streaming of data to computing resources will be necessary - ~2.5Gbps today, significant growth curve

Futures – Square Kilometer Array

Large radio telescope in the Southern Hemisphere

- Approximately one square kilometer of combined signal collection area
- ~2800 receivers in the telescope array
- Most receivers in a 180km diameter area
- Average 100km run to central correlator
- ~2 Petabytes per second at the central correlator

Distribution of data to international collaborators

- Expected rate of ~100Gbps from correlator to analysis centers
- International collaboration → wide data distribution

There are others (Sensor networks, ITER, etc.)

Collaboration Structures

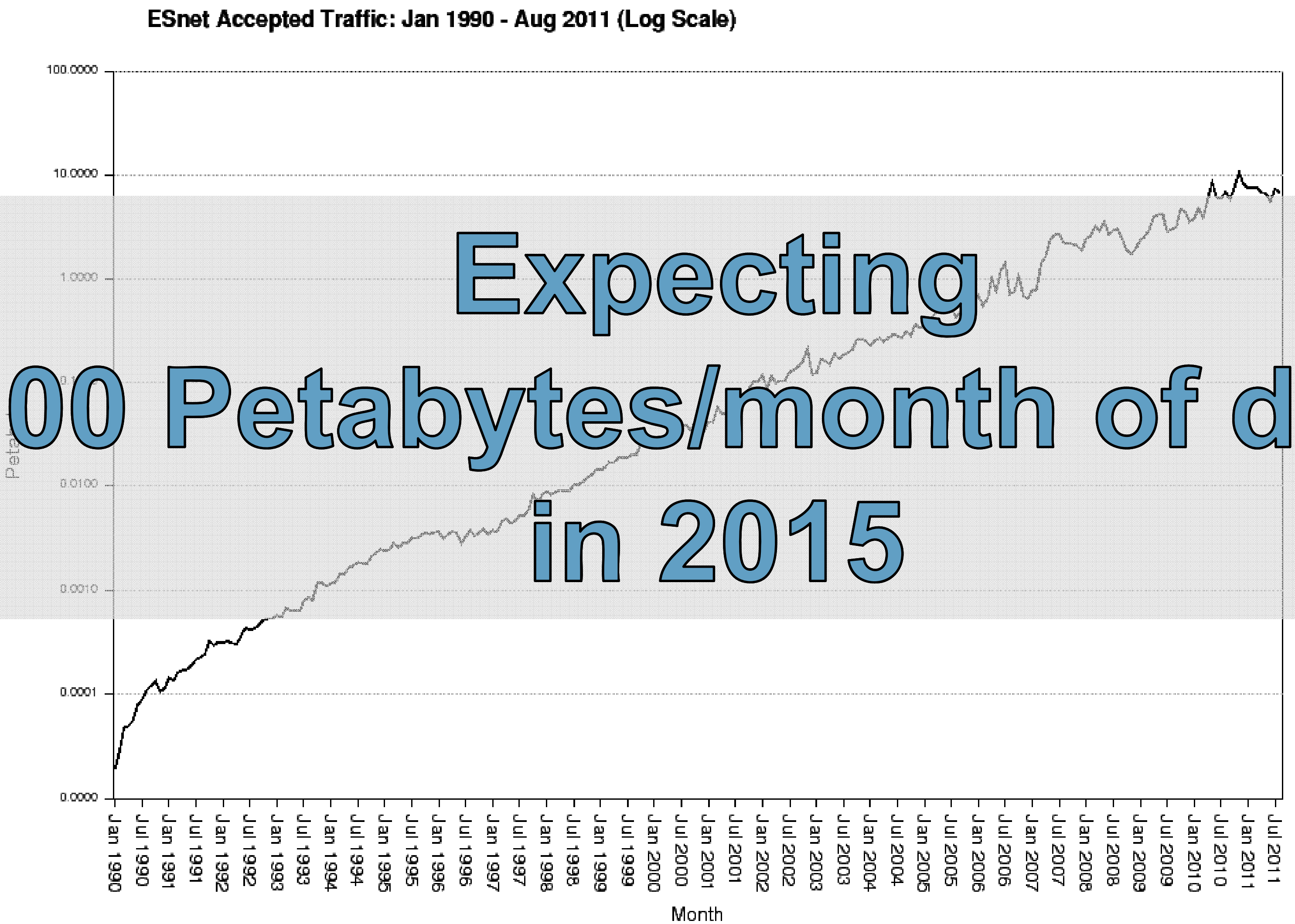
The very structure of modern science assumes there is a network interconnecting all parts of the collaboration

- Large, unique facilities (e.g. LHC) provide a focus for all members of a field
  - Data distributed to scientists
  - Results distributed among collaborators
- Data analysis using local resources also drives data movement
  - Example – large simulation run at supercomputer center, secondary analysis at home institution

Large data sets + increasing scope of collaboration

- Scientific productivity gated on data analysis
- Data moved to analysis, analysis moved to data – both must be supported in the general case

The Science Data Growth



Expecting
100 Petabytes/month of data
in 2015

ESnet

Networks For Data Intensive Science

What does data intensive science traffic look like?

For a given bandwidth, much larger per-flow rates, much smaller flow count

Often a significant fraction (10-50%) of link bandwidth in a single flow (often over intercontinental distances)

TCP is showing its limitations

- Loss sensitivity
- CPU load

IMIX traffic profile is not a good approximation for science traffic

Futures

Data rates will continue to increase

- Sensor data rate scales with semiconductor capabilities (think digital cameras)
- Large facilities will fan out data to large collaborations

Host architecture means new protocols are likely

- Per-flow packet processing (e.g. TCP) is essentially a serial task, and per-core clock rate is essentially flat
- Faced with exponential growth in data rates, what do we do?
- Different means of getting data into host memory (e.g. RDMA over Ethernet) are being tested today
- Again – IMIX traffic is not a good model here

Impact on Networks

Science networks will continue to see a different traffic profile

- Relatively small flow count, relatively large flow rate
- IMIX traffic assumptions won't hold
- Beware LAG-like implementations!

Protocol space is likely to change

- Data mobility requirements will drive new modes of operation – even if it's standard it probably won't be all TCP(Ethernet, or IP, or UDP, with something above for reliability)
- Services are important – predictability and programmability are important (e.g. low loss, OpenFlow)

Questions?

Thanks!

Eli Dart - dart@es.net

http://www.es.net/

http://fasterdata.es.net/