Network Measurement: From Collection to Visualization

Jon M. Dugan
Network Engineer, ESnet
Lawrence Berkeley National Laboratory

2014 Technology Exchange
Indianapolis, IN
27 October 2014
Intro

• What is ESnet?
• Why measure? Why visualize?
• ESnet’s Measurement, Analysis and Visualization Architecture
• Time Series data
• Demo
• Analysis & Reporting
• The MyESnet Portal
• Reusable Components and Open Source
What is ESnet?

High-speed national network, optimized for DOE science missions:

- connecting 40 labs, plants and facilities with >100 networks
- $32.6M in FY14, 42FTE
- older than commercial Internet, growing twice as fast

$62M ARRA grant for 100G upgrade:

- transition to new era of optical networking
- fiber assets + access to spectrum shared with Internet2
- world’s first 100G network at continental scale

Culture of urgency:

- 4 awards in past 3 years
- R&D100 in FY13
- “5 out of 5” for customer satisfaction in last review
• abundant capacity
• programmability
• end-to-end performance
Why measure?

“If you can’t measure something, you can’t understand it.
If you can’t understand it, you can’t control it.
If you can’t control it, you can’t improve it.”
― H. James Harrington

“The most important things cannot be measured.”
― W. Edwards Deming
Why visualize?

**Insight**  “There is a magic in graphs. The profile of a curve reveals in a flash a whole situation — the life history of an epidemic, a panic, or an era of prosperity. The curve informs the mind, awakens the imagination, convinces.” —Henry D. Hubbard

**Discovery**  “The greatest value of a picture is when it forces us to notice what we never expected to see.” —John Tukey

**Understanding**  “Visualizations act as a campfire around which we gather to tell stories.” —Al Shalloway
Architecture
esmond

- Collection and storage of large sets of time series data
- Calculates summaries but retains all data
- Initially focused on SNMP measurements of routers
- Recently started to store perfSONAR measurements
  - More details at perfSONAR talk in this room 2:15 on Wednesday
- Interested in exploring storing flow data
- Basis for our monthly reporting
- Uses the Cassandra database engine
- Open source: http://software.es.net/esmond/
Flow data

- Flow data gives more detail about traffic patterns
- Statistically sampled
- We currently have two systems that handle flow
  - Arbor PeakFlow SP: great if you know what you’re looking for ahead of time. Aggregates and discards raw data.
  - Home grown system: Saves raw data and allows asking questions after the fact.
Events

- Additional data can help in the interpretation of visualizations
- Events can be added by humans or harvested from log data
Time Series

• SNMP, Flow and Event data are all time series
• Our visualizations were built as needed (and often rapidly)
• Users want to explore more interactively, but current viz not quite flexible enough
• We need a cohesive time series abstraction
• Visualization components expect data in this unified format
• Analysis, reporting and potentially data collection will use this format
• Simplify code and make it more reusable
Demo

Live demo of current time series work
Analysis

• Using the raw counters to produce other interesting measurements

• Examples
  – Instantaneous total traffic in/out of ESnet
    • Traffic map and chart on portal homepage: https://my.es.net/
  – Ranking of busiest interfaces in the network
    • http://stats.es.net/top.html
  – Daily and monthly rollup of traffic volume
Reporting

Monthly traffic reports
- Traffic growth trends
- Overlay reports (LHCONE)
- OSCARS reports

Future
- Busy interfaces
- Capacity planning
- Sector breakdowns
- ESnet Site breakdowns
Metadata: Databases and Curated Data

• Making effective use of this data requires understanding how the pieces fit together

• ESDB
  – Source of truth
    • Contacts
    • Organizations
    • Locations
    • Circuits
  – Otherwise homeless data
  – Developed circuit components useful in other contexts
  – Beyond the scope of this talk, but see me for a demo
MyESnet

A data-driven website for displaying information and visualizations about ESnet, its sites and community.
MyESnet: SC13 Visualization
Reusable Components

- Visualizations are built as reusable components
- Each visualization is made to be as general as possible
- Separation of client side code and server side code
- React enables composable, reusable components
  - Modularity can be a challenge in web apps
  - Many approaches (Backbone, Marionette, Angular, Ember)
Open Source

Announcement
- Starting in Q1 2015 ESnet will start to release parts of the portal as open source
- One component at a time – focus on sharing the reusable bits
- Not a turn key solution

Motivation
- Develop a common toolkit for visualizing networks
- Lower barrier to entry
- Allow people to use what we’ve built so they don’t have to build their own
- Encourage others to share their visualizations
- Give back to the community
Questions?

Jon M. Dugan <jdgman@es.net>
https://my.es.net/
http://software.es.net/
http://fasterdata.es.net/
https://my.es.net/demos/techex14

Come talk to us in the exhibit hall between sessions!

Thanks for listening!