REDDnet Performance Monitoring

Ezra Kissel
University of Delaware

Performance WG, I2 SMM 2009

Distributed, Data Intensive Collaboration

- **REDDnet:** Research and Education Data Depot Network
- **NSF funded**, additional from Library of Congress
- **"Working storage"** to help manage the logistics of **sharing, moving and staging** large datasets **across wide areas** and **distributed collaborations**.
- **Institutions:** Vanderbilt, Tennessee, Stephen F. Austin, NC State, Nevoa Networks, Delaware

- **Host Sites:** Caltech, Florida, Michigan, ORNL, SDSC, TACC, UC Santa Barbara (Stephen F. Austin, Tennessee, Vanderbilt)

REDDNET

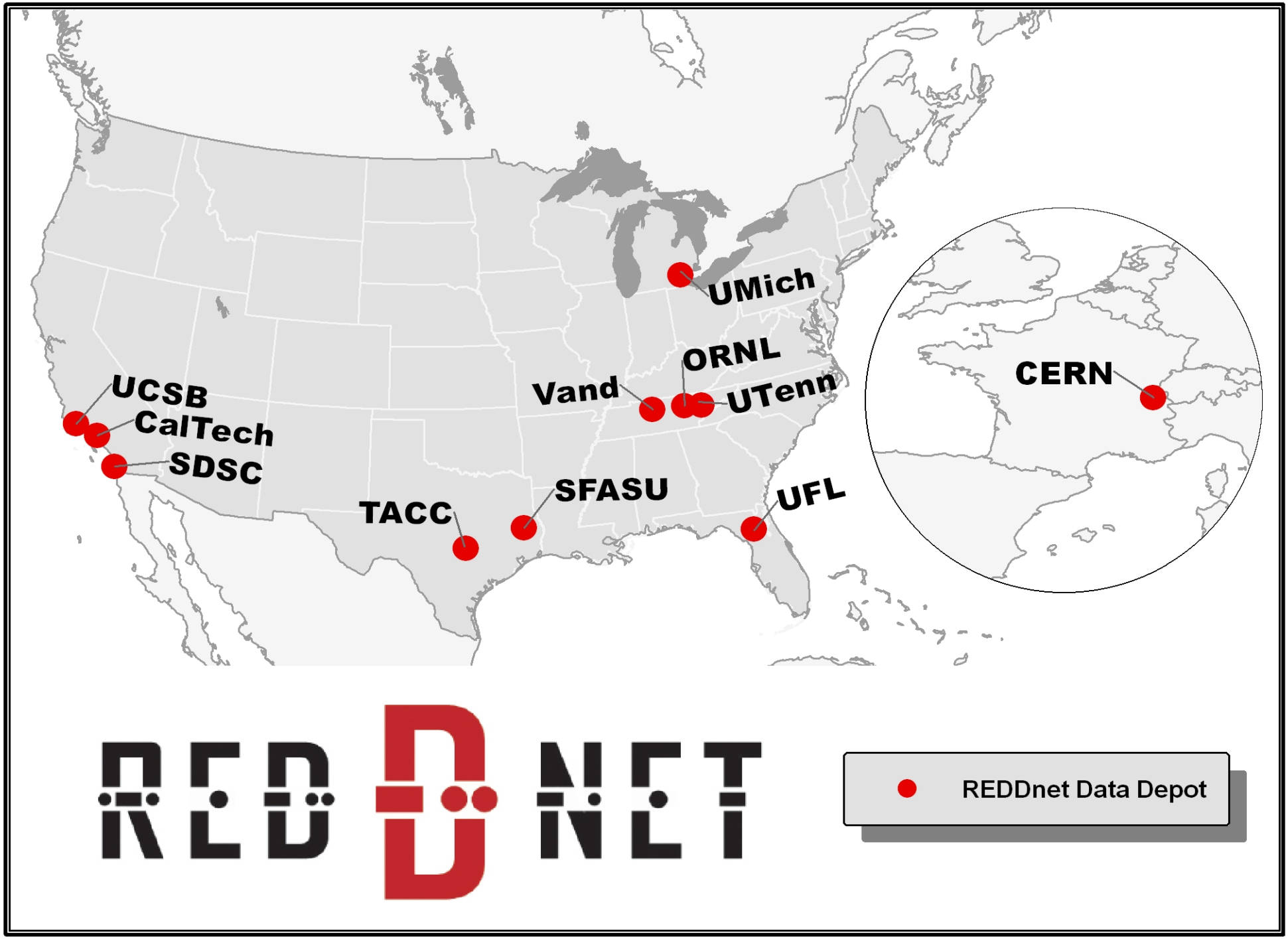
INTERNET2

Working Storage and the Data Pulse Model

- Often collaborations are strongly interested in a data set for a brief period – a month or so – shifting to a new set after.
  - Want data sets to “pulse” – be widely available for period of interest
  - Working storage with a “temporal” element – both space and time are important
- REDDnet is adept at Data Logistics, for example:
  - Moving data across boundaries – such as into or out of the TeraGrid
  - Can take advantage of fast network links such as within TeraGrid
- For more information: http://www.reddnet.org

REDDNET

INTERNET2


UMich
ORNL
UTenn
Vand
UCSB
CalTech
SDSC
TACC
SFASU
UFL
CERN
REDDnet
REDDnet Data Depot

REDDnet Core Elements: IBP, "Data Warming"

- IBP: a highly generic, "best effort" protocol for using storage
- Easy to port and deploy
- Best Effort: low burden on providers
- Reliability, service quality set by middleware sitting on top of IBP
- Common task is augmenting, "warming", data from one site to another
- How to ensure sites perform as expected and make best use of network resources, importance of regular monitoring

REDDNET

INTERNET2

Performance Monitoring Deployment

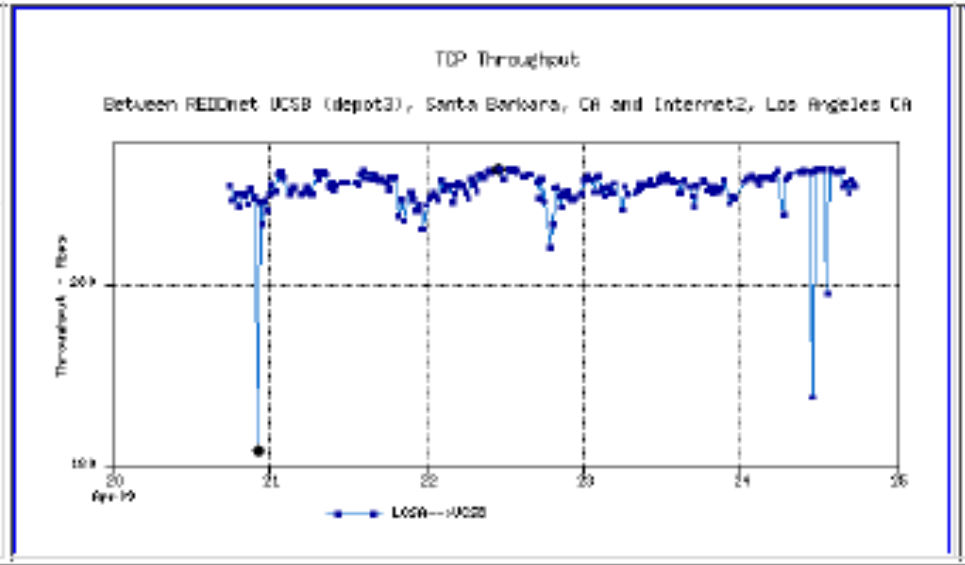
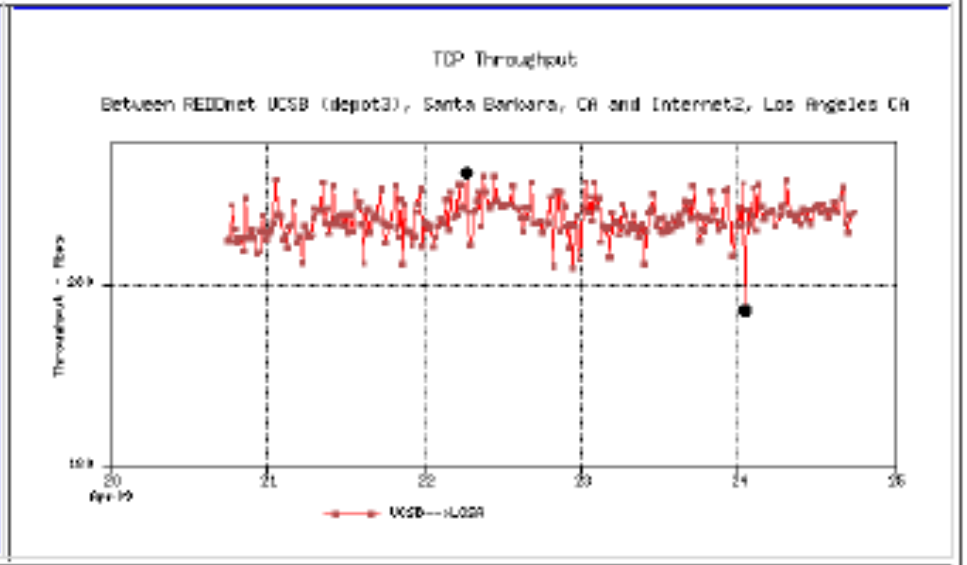
- owamp (3.1)
- bwctl (1.3)
- ndt client (3.5)
- perfSONAR-PS perfSONAR-BUOY (regular testing framework for bwctl)

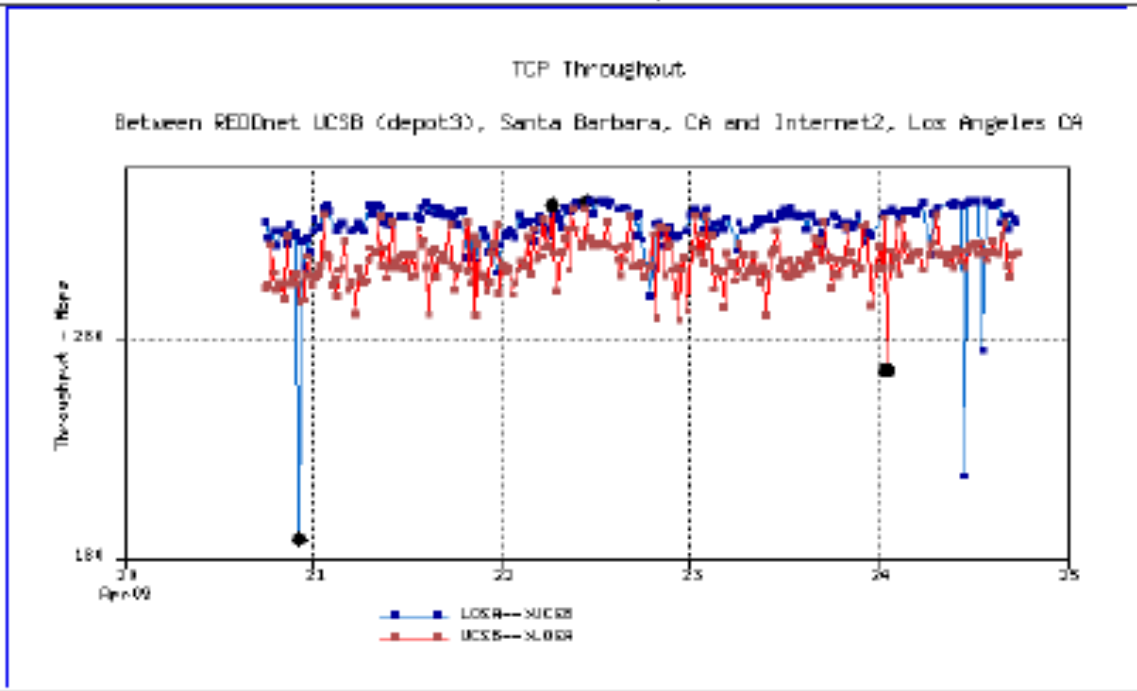
- Established performance baselines using owamp and pSB
  - 9 REDDnet depots to 9 I2 POPs
  - ATLA, CHIC, HOUS, KANS, LOSA, NEWY, SALT, WASH

REDDNET

INTERNET2

BWCTL - REDDnet IPv4 TCP Throughput - Internet2 Los Angeles CA

	Receiving from Internet2, Los Angeles CA	Sending to Internet2, Los Angeles CA
REDDnet UCSB (depot3), Santa Barbara, CA		



Performance Monitoring

- TCP tuning on all hosts – performance improved in places
  - Indication that there were legitimate network issues to resolve
- Picked a set of hosts to investigate from the “worst offenders”
  - owamp to find loss/reordering
  - ndt to evaluate the performance of the path
- Divide and conquer approach
  - Test from depot to POP
  - Divide path (depot to regional, depot to campus edge)
  - Narrow down where the problem “ends”

Example 1

- REDDnet Umich and CHIC I2 POP – 4 hop path
- Initial performance was very poor in one direction
  - running outbound test (client to server) . . . . . 93.44 Mb/s
  - running inbound test (server to client) . . . . . 931.58 Mb/s
- We were able to narrow the problem down to being in the Ultralight domain on a specific Cisco switch.
- The result of the fault was software forwarding of all outbound packets which resulted in the uni-directional problems that were observed.
- After fixing, performance was excellent

REDDNET

INTERNET2

Example 2

- REDDnet Vanderbilt to Atlanta I2 POP
- owamp observed high amounts of loss in one direction, seriously impairing performance
- An unrelated maintenance window independently scheduled the day after detection fixed the problem
- A fortunate coincidence, but fault was detected and brought to attention

REDDNET

INTERNET2

Lessons Learned

- Performance issues can be common and difficult to detect
- Ensure consistent TCP tuning across all hosts as a first step to eliminate easy-to-fix causes
- Asymmetric routing must be accounted for
- Heterogeneous configurations makes consistent baselines a challenge (interface bonding, local network)
- Most importantly, good communication between interested parties is key

Future Steps

- Create a coherent plan for utilizing available tools to most effectively reach immediate monitoring goals
- Enhance regular monitoring between REDDnet depots
  - Establish a mesh of perfSONAR nodes at each site
  - Maintain baselines of available bandwidth, thresholds for “event” notification
  - Determine suitable monitoring intervals to avoid affecting day-to-day traffic
- Maintain a good, synergistic relationship with the perfSONAR community

REDDNET

INTERNET2

Acknowledgments

- REDDnet
  - Paul Sheldon, Bobby Brown, Matthew Binkley, Alan Tackett (Vanderbilt), Terry Moore, Micah Beck, Chris Sellers (UTK), PR Blackwell (SFASU)
- Internet2
  - Jason Zurawski, Aaron Brown, Matt Zekauskas, Jeff Boote
- University of Delaware
  - Ezra Kissel, Martin Swany
- CERN
  - Sandor Rozsa

RED D NET

INTERNET 2