Scalable and Extensible Network Monitoring

Sonia Fahmy (Purdue University)
Puneet Sharma (HP Labs)

Leverages prior work joint with: Praveen Yalagandula, Sujata Banerjee, Sujoy Basu, SJ Lee (HP Labs), Ethan Blanton, Sriharsha Gangam, Greg N. Frederickson (Purdue University)

http://networking.hpl.hp.com

http://www.cs.purdue.edu/homes/fahmy/

Goals

- Provide system state in real-time
  - Both network and node state
    - *Active* and passive
    - *E2E* or leverages network element info when available
- Flexible and extensible
  - Easy to add new measurement tools to be developed!
  - Configurable time scales (start time, frequency, number)
  - *Support complex queries*
    - To which node do I have the largest bandwidth?
    - Which game server is within 10ms latency?
- Share measurement info across applications
  - Eliminate redundant expensive measurements
- *Scalable*, secure, and reliable

Challenges

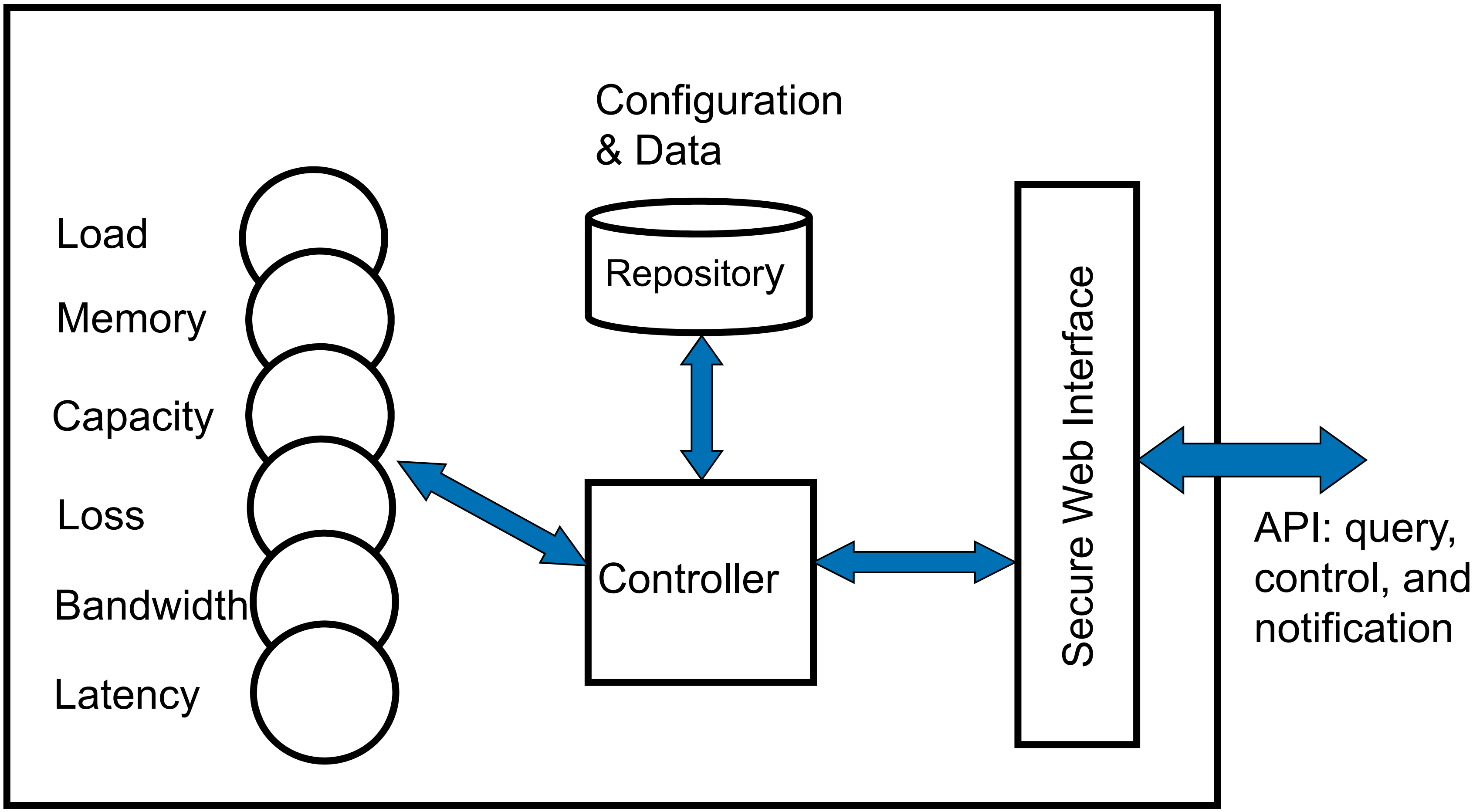
- Tools previously tested only in point-to-point configurations
- Deployment in a large scale setting exposed several issues
  - Hard-coded port numbers leading to port conflicts
  - Need to be started at source and destination simultaneously
  - Large resource requirements leading to end-node crashes
  - Long running times leading to web server timeouts
- On-demand measurements at user defined times, frequencies, and tolerance to error/staleness
- Estimation of load introduced by measurement probes
- Dynamic invocation of inference mechanisms based on measurement request workload

Scalable Sensing Service (S^3)

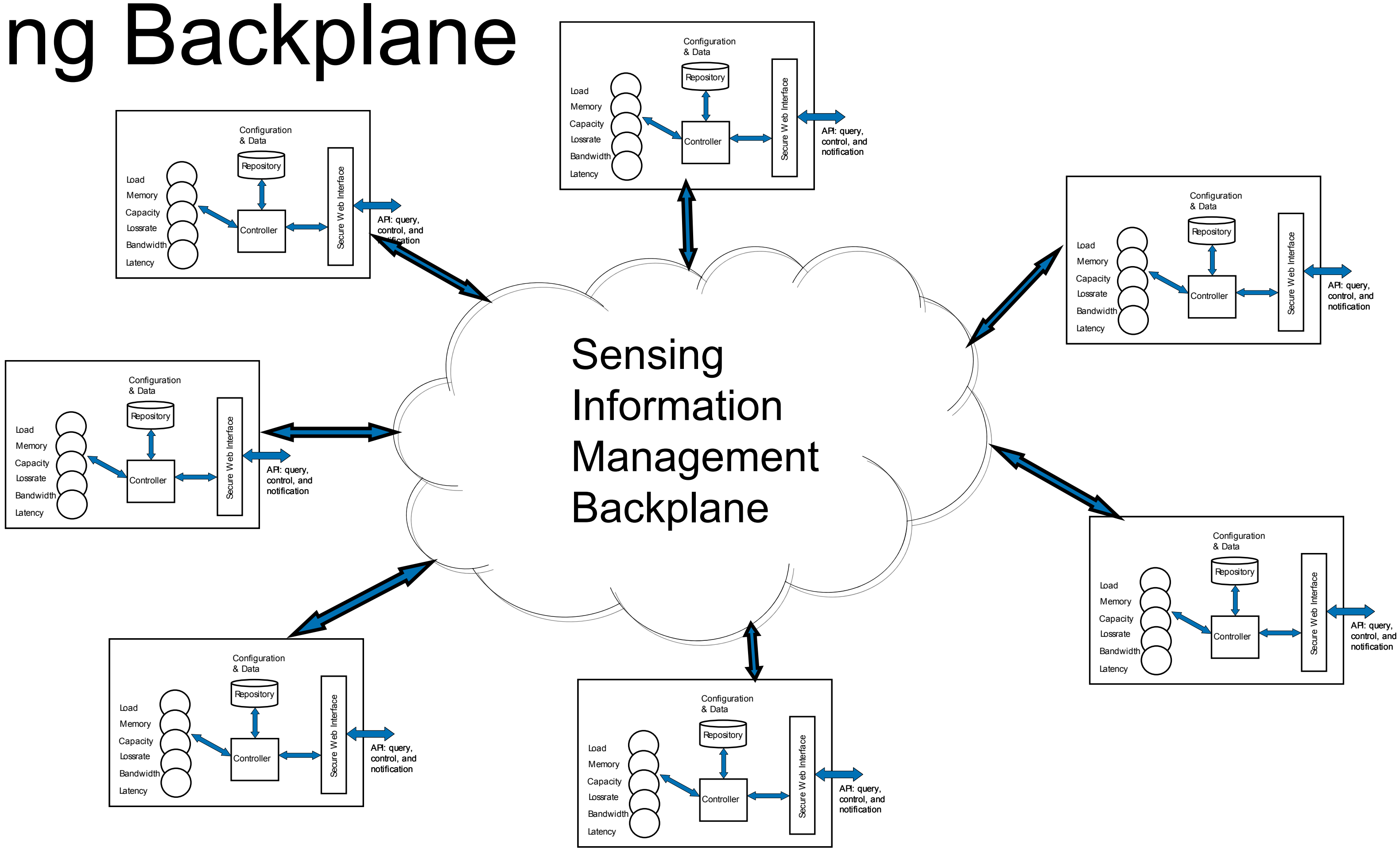
- Sensor pods
  - Measure system state from a node perspective
  - Web-Service enabled collection of sensors
- Backplane
  - Distributed programmable fabric
  - Connects pods, and aggregates measured system state
- Inference engines
  - Infer $O(n^2)$ E2E path info by measuring a few paths
  - Dynamically schedules measurements on pods
  - Aggregates data on backplane

Sensor Pod

Web-Service (WS) enabled collection of sensors



Sensing Backplane



- Aggregate data from end-points
- Configurable and self-managing
- E.g., SDIMS [SIGCOMM 2004]

Scalable Inference Engines

- Large overhead for probing and data exchange
  - $O(N^2)$ measurements in a network of N nodes
  - Dynamically changing ➔ Need frequent probing
- Measurement/Monitoring failures
  - Failed or slow end machines
  - Measurement tool failures
- Inference based on incomplete information
  - Exploit properties such as triangular inequality
  - A coarse estimate may suffice for many applications
- Prediction based on archived information
- Tradeoff between accuracy and overhead
- When and where to use inference? [Blanton et al., ICDCS09]

PURDUE UNIVERSITY

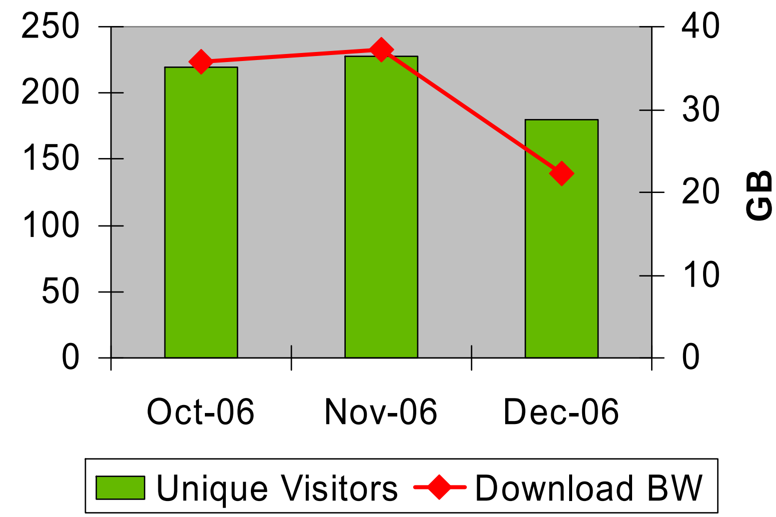
hp invent

Prototype Deployment on PlanetLab

- 700+ nodes scattered across 350+ sites
- Running since January 2006
- All pair network metrics: E2E latency, BW, Capacity, Loss
- Simple backplane: central server
  - Maintains pods, schedules measurements, collects and publishes data
- Stats:~14GB raw data every day, ~1GB compressed

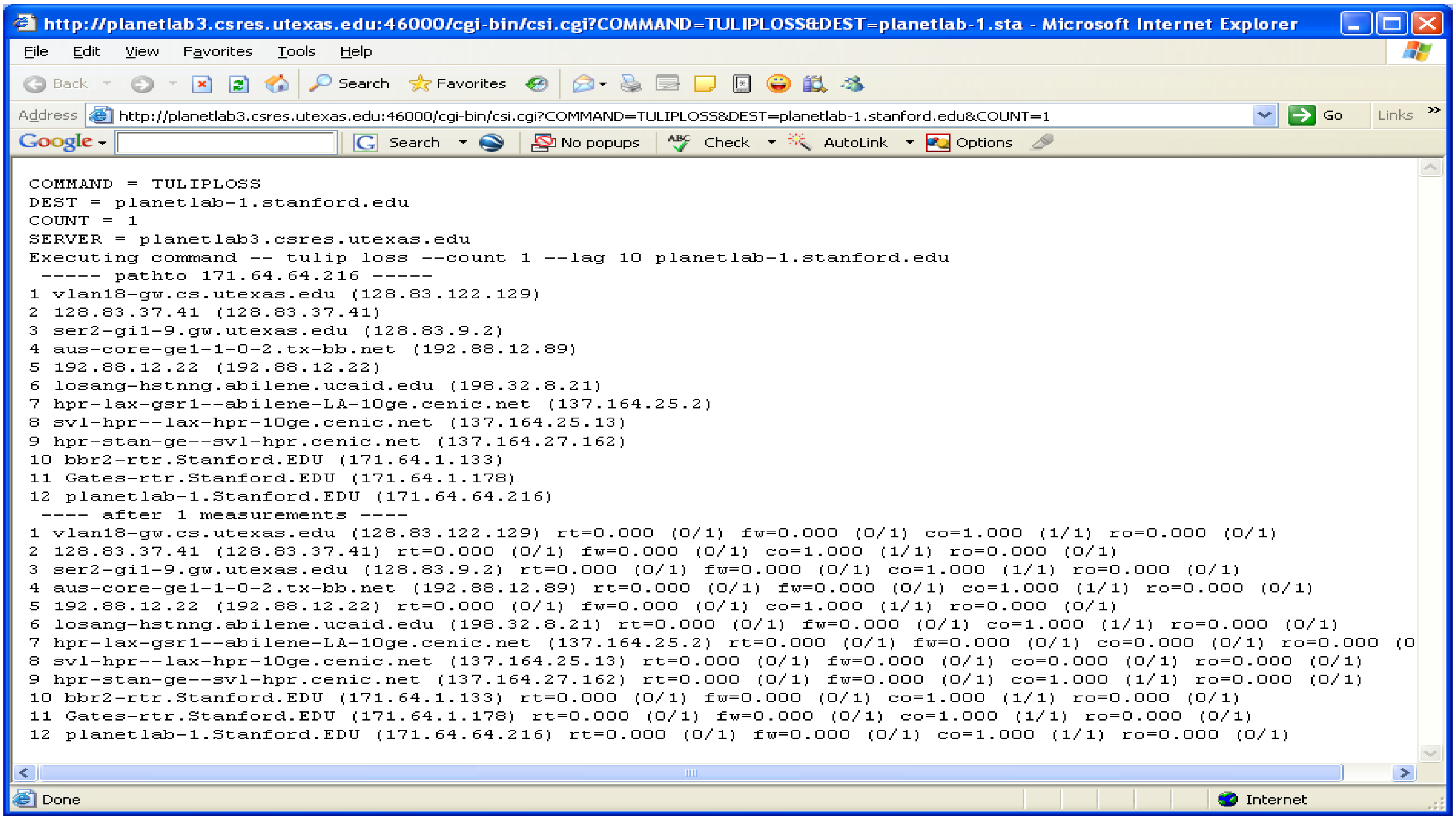
S^3 Data Usage

- Web server stats (2006):
  - ~200 unique visitors/month
  - ~20GB download BW/month



- Projects
  - HP Labs: Bandwidth inference, Resource-aware monitoring, semantic store
  - Others: Purdue University, MSR, U of Washington, Georgia Tech, Harvard, Princeton, Boston University, etc.

Screenshot: Hop-by-hop Loss Sensor




```
COMMAND = TULIPLOSS
DEST = planetlab-1.stanford.edu
COUNT = 1
SERVER = planetlab3.csres.utexas.edu
Executing command -- tulip loss --count 1 --lag 10 planetlab-1.stanford.edu
 ----- pathto 171.64.64.216 -----
1 vlan18-gw.cs.utexas.edu (128.83.122.129)
2 128.83.37.41 (128.83.37.41)
3 ser2-gi1-9.gw.utexas.edu (128.83.9.2)
4 aus-core-ge1-1-0-2.tx-bb.net (192.88.12.89)
5 192.88.12.22 (192.88.12.22)
6 losang-hstnng.abilene.ucaid.edu (198.32.8.21)
7 hpr-lax-gsr1--abilene-LA-10ge.cenic.net (137.164.25.2)
8 svl-hpr--lax-hpr-10ge.cenic.net (137.164.25.13)
9 hpr-stan-ge--svl-hpr.cenic.net (137.164.27.162)
10 bbr2-rtr.Stanford.EDU (171.64.1.133)
11 Gates-rtr.Stanford.EDU (171.64.1.178)
12 planetlab-1.Stanford.EDU (171.64.64.216)
 ---- after 1 measurements ----
1 vlan18-gw.cs.utexas.edu (128.83.122.129) rt=0.000 (0/1) fw=0.000 (0/1) co=1.000 (1/1) ro=0.000 (0/1)
2 128.83.37.41 (128.83.37.41) rt=0.000 (0/1) fw=0.000 (0/1) co=1.000 (1/1) ro=0.000 (0/1)
3 ser2-gi1-9.gw.utexas.edu (128.83.9.2) rt=0.000 (0/1) fw=0.000 (0/1) co=1.000 (1/1) ro=0.000 (0/1)
4 aus-core-ge1-1-0-2.tx-bb.net (192.88.12.89) rt=0.000 (0/1) fw=0.000 (0/1) co=1.000 (1/1) ro=0.000 (0/1)
5 192.88.12.22 (192.88.12.22) rt=0.000 (0/1) fw=0.000 (0/1) co=1.000 (1/1) ro=0.000 (0/1)
6 losang-hstnng.abilene.ucaid.edu (198.32.8.21) rt=0.000 (0/1) fw=0.000 (0/1) co=1.000 (1/1) ro=0.000 (0/1)
7 hpr-lax-gsr1--abilene-LA-10ge.cenic.net (137.164.25.2) rt=0.000 (0/1) fw=0.000 (0/1) co=0.000 (0/1) ro=0.000 (0
8 svl-hpr--lax-hpr-10ge.cenic.net (137.164.25.13) rt=0.000 (0/1) fw=0.000 (0/1) co=0.000 (0/1) ro=0.000 (0/1)
9 hpr-stan-ge--svl-hpr.cenic.net (137.164.27.162) rt=0.000 (0/1) fw=0.000 (0/1) co=1.000 (1/1) ro=0.000 (0/1)
10 bbr2-rtr.Stanford.EDU (171.64.1.133) rt=0.000 (0/1) fw=0.000 (0/1) co=1.000 (1/1) ro=0.000 (0/1)
11 Gates-rtr.Stanford.EDU (171.64.1.178) rt=0.000 (0/1) fw=0.000 (0/1) co=1.000 (1/1) ro=0.000 (0/1)
12 planetlab-1.Stanford.EDU (171.64.64.216) rt=0.000 (0/1) fw=0.000 (0/1) co=1.000 (1/1) ro=0.000 (0/1)
```

S^3 Screenshot

S-Cube: planetlab3.csres.utexas.edu - Internet Explorer configured for HP Labs

File Edit View Favorites Tools Help

Address http://networking.hpl.hp.com/cgi-bin/gennodeinfo-latsorted.cgi?NODE=planetlab3.csres.utexas.edu

S^3 Scalable Sensing Service

|| S^3 home || Proximity/Latency Estimation || Tools || Papers || People ||

Sensing Snapshot for planetlab3.csres.utexas.edu for 589 nodes

This page was generated on 2006-09-07 at 12:55:04 Pacific Time

Destination Node	Estimated Latency(Measured Latency)(msec)	Bottleneck Capacity (Kbps)	Available Bandwidth		Tulip Loss Rate(Lost Probes/Total Probes)
			Pathchirp (Kbps)	Spruce (Kbps)	
planetlab2.csres.utexas.edu	0.412	95000	99328.1121	94436	0.000(0/101)
planetlab1.csres.utexas.edu	0.457 (0.144)	90000	107357.5301	88857	0.000(0/100)
pl1a.pl.utsa.edu	4.679	10100	9509.4717	52	0.010(1/102)
ricepl-1.cs.rice.edu	5.163	10200	9300.7361	8553	0.000(0/101)
planetlab2.tamu.edu	6.833	128000	68786.2855	69043	0.000(0/102)
ricepl-3.cs.rice.edu	10.258	N/A	N/A	N/A	0.000(0/101)
planetlab2.uta.edu	10.545	34000	53187.6058	26746	0.000(0/100)
planetlab1.uta.edu	10.598	31000	47572.455	27892	0.010(1/102)
kupl2.ittc.ku.edu	22.103	363000	91364.2047	0	0.000(0/0)
kupl1.ittc.ku.edu	22.171	90000	90911.5079	90217	0.000(0/0)
plab1.eece.ksu.edu	24.693	20000	47091.8504	15081	0.000(0/101)
planetslug3.cse.ucsc.edu	24.995	102000	92311.3477	90475	0.000(0/102)

http://networking.hpl.hp.com/s-cube Local intranet

Selected Publications

- http://networking.hpl.hp.com/s-cube
- Ethan Blanton, Sonia Fahmy, Greg N. Frederickson, "On the Utility of Inference Mechanisms," In Proceedings of IEEE International Conference on Distributed Computing Systems (ICDCS), 8 pp., June 2009.
- Ethan Blanton, Sonia Fahmy, Sujata Banerjee, "Resource Management in an Active Measurement Service," In Proceedings of the IEEE Global Internet Symposium, 6 pp., April 2008.
- P. Yalagandula, P. Sharma, S. Banerjee, S.-J.Lee, and S. Basu, S3: A Scalable Sensing Service for Monitoring Large Networked Systems," In *Proceedings of the Workshop on Internet Network Measurement 2006,* Pisa, Italy, September 2006.
- Praveen Yalagandula, Sung-Ju Lee, Puneet Sharma, and Sujata Banerjee, "Correlations in End-to-End Network Metrics: Impact on Large Scale Network Monitoring," In *IEEE Global Internet Symposium,* Phoenix, AZ, April 2008.