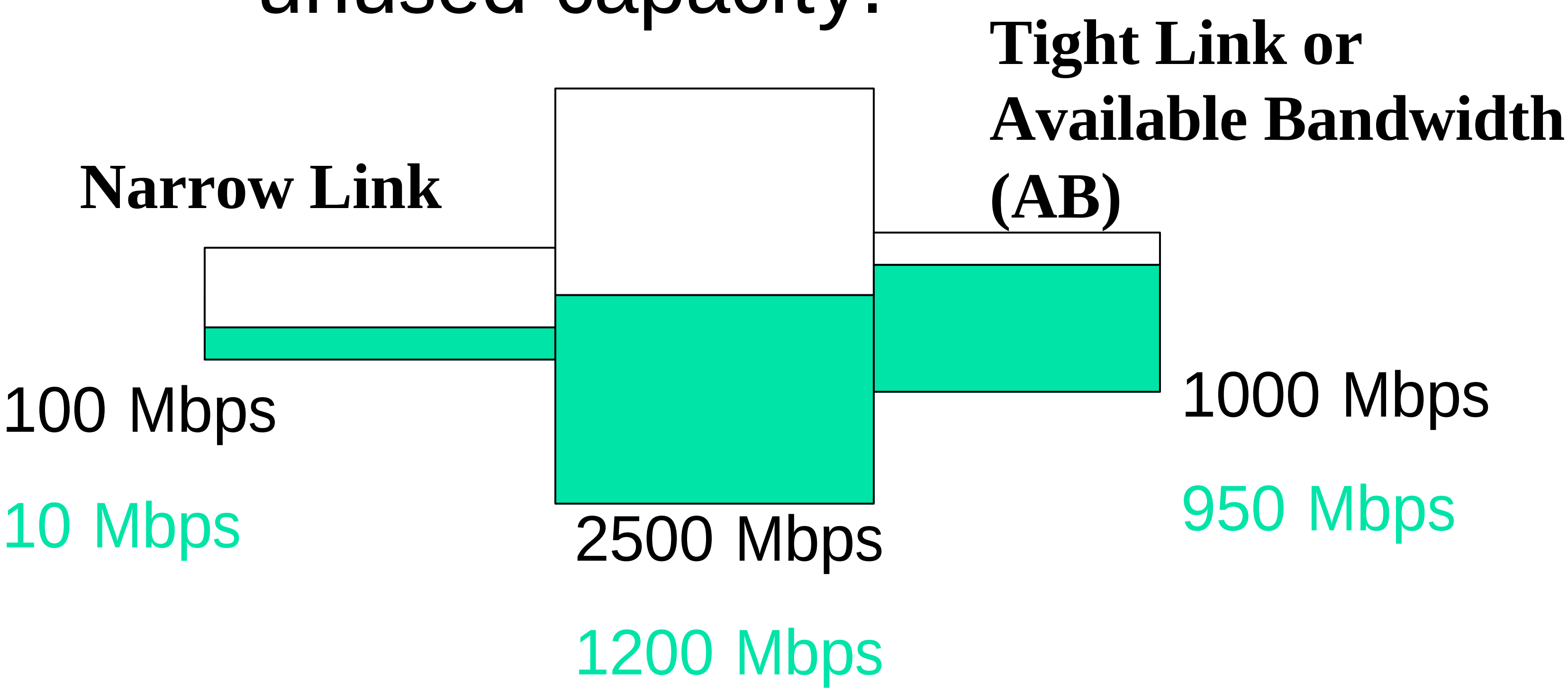
Comparison of Public End-to-End Bandwidth Estimation tools on High-Speed Links

Alok Shriram, Margaret Murray, Young Hyun, Nevil Brownlee, Andre Broido, Marina Fomenkov and kc claffy

What is Available Bandwidth ?

- Available bandwidth of an end-to-end path is the link with the minimum unused capacity.



Tools Under Consideration for this Study

Available Bandwidth Tools

- Pathload [Dovrolis]
- Pathchirp [Ribeiro]
- Abing [Navaratil]
- Spruce [Strauss]

Bulk Transfer Capacity Tool

- Iperf [Dugan] (Unofficial Standard)
- Not Considering tools like
  - IGI/PTR[Hu]
  - Cprobe[Crovella]
  - Pipechar[Jin]
  - Netest [Jin]

Tool properties and Metrics

- Tool Accuracy
- Operational Characteristics
  - Measurement Time
  - Intrusiveness, Overhead

Why would we want to do this?

- Perform comprehensive, cross-tool validation.
  - Previous validation limited to low speed paths.
  - No comprehensive bias free evaluation.
  - Cross-traffic scenarios

Goal

- Discover insights about tool usability and deployment.
- Compare tool methodologies.

Where are we doing this?

- Study in two parts

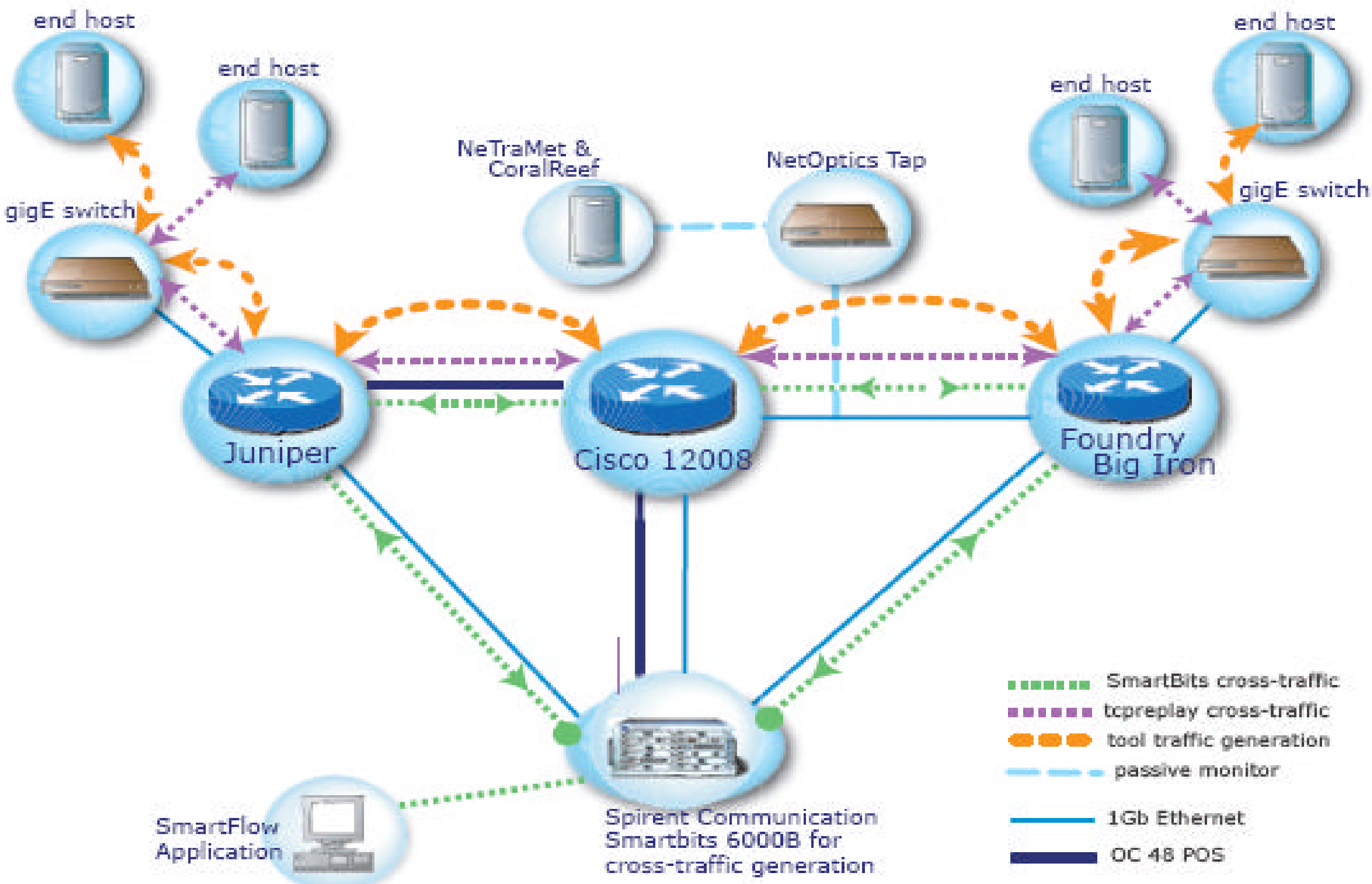
- First part in a laboratory setting where we can set most parameters.

- Second part on actual internet paths with access to SNMP counters.

Outline

- **Laboratory Setup**
- Part 1: Laboratory Experiments
  - SmartBits
  - Tcpreplay
- Part 2: Internet Experiments
  - Abilene Network
  - SDSC ORNL

Our Lab Topology



Methods of Generating Cross-Traffic

- Prior results criticized because of “unrealistic” cross-traffic scenarios.
- Two Methods of Cross-Traffic generation
  - SmartBits
  - TCPreplay
- We attempt to recreate as realistic cross-traffic as possible
- We analyze the cross-traffic using two separate monitoring utilities.
  - NeTraMet
  - CoralReef

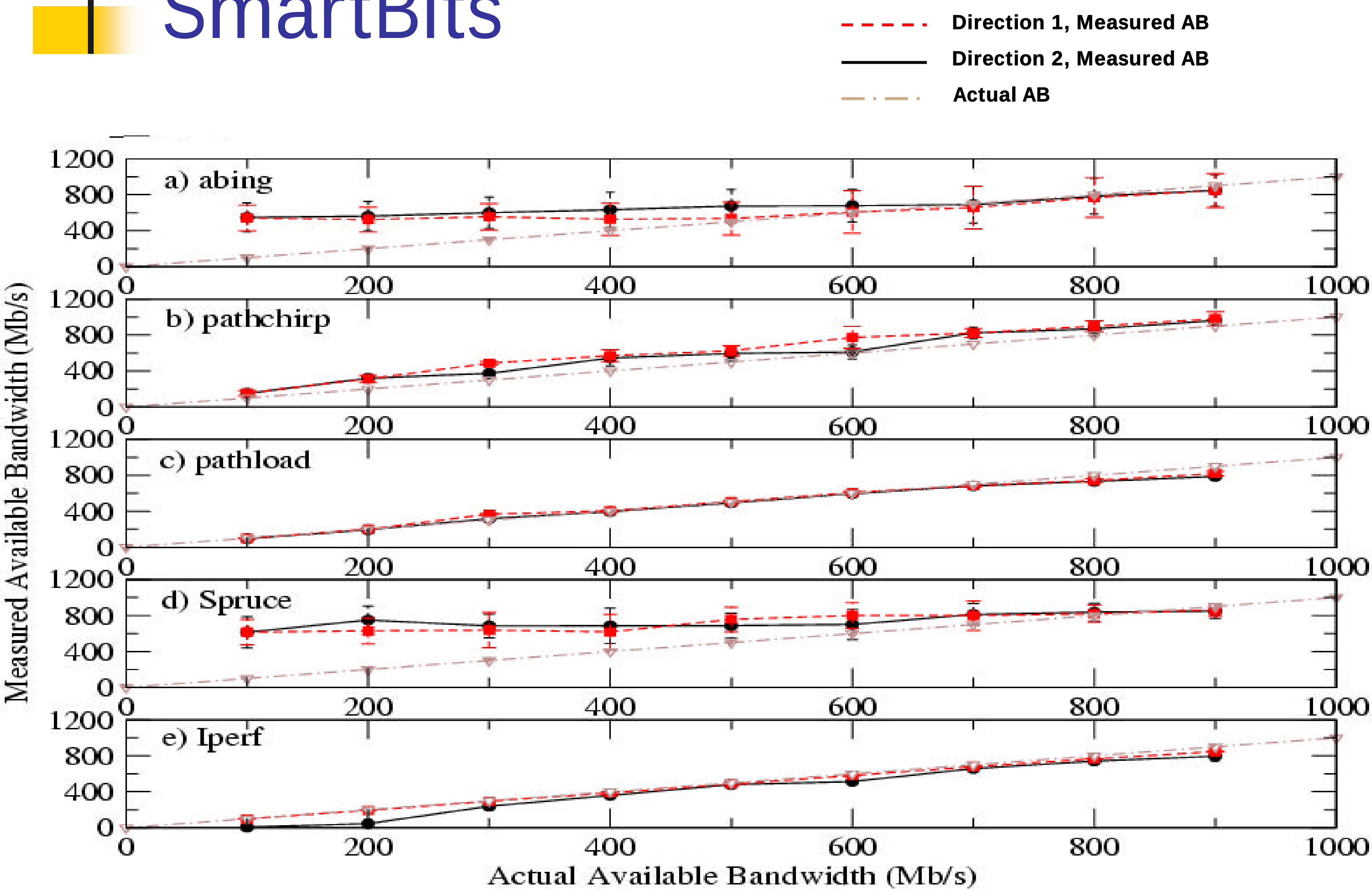
Outline

- Laboratory Setup
- **Part 1: Laboratory Experiments**
  - **SmartBits**
  - Tcpreplay
- Part 2: Internet Experiments
  - Abilene Network
  - SDSC ORNL

Experiments with SmartBits

- SmartBits generates a known load
- Running in both directions of the shared path.
  - Range from 100 to 900 Mb/s
  - Increments of 100 Mb/s
- SmartBits cross-traffic for **6** minutes
- AB tools back-to-back for **5** minutes.
- Average the results.

Accuracy of Tools Using SmartBits



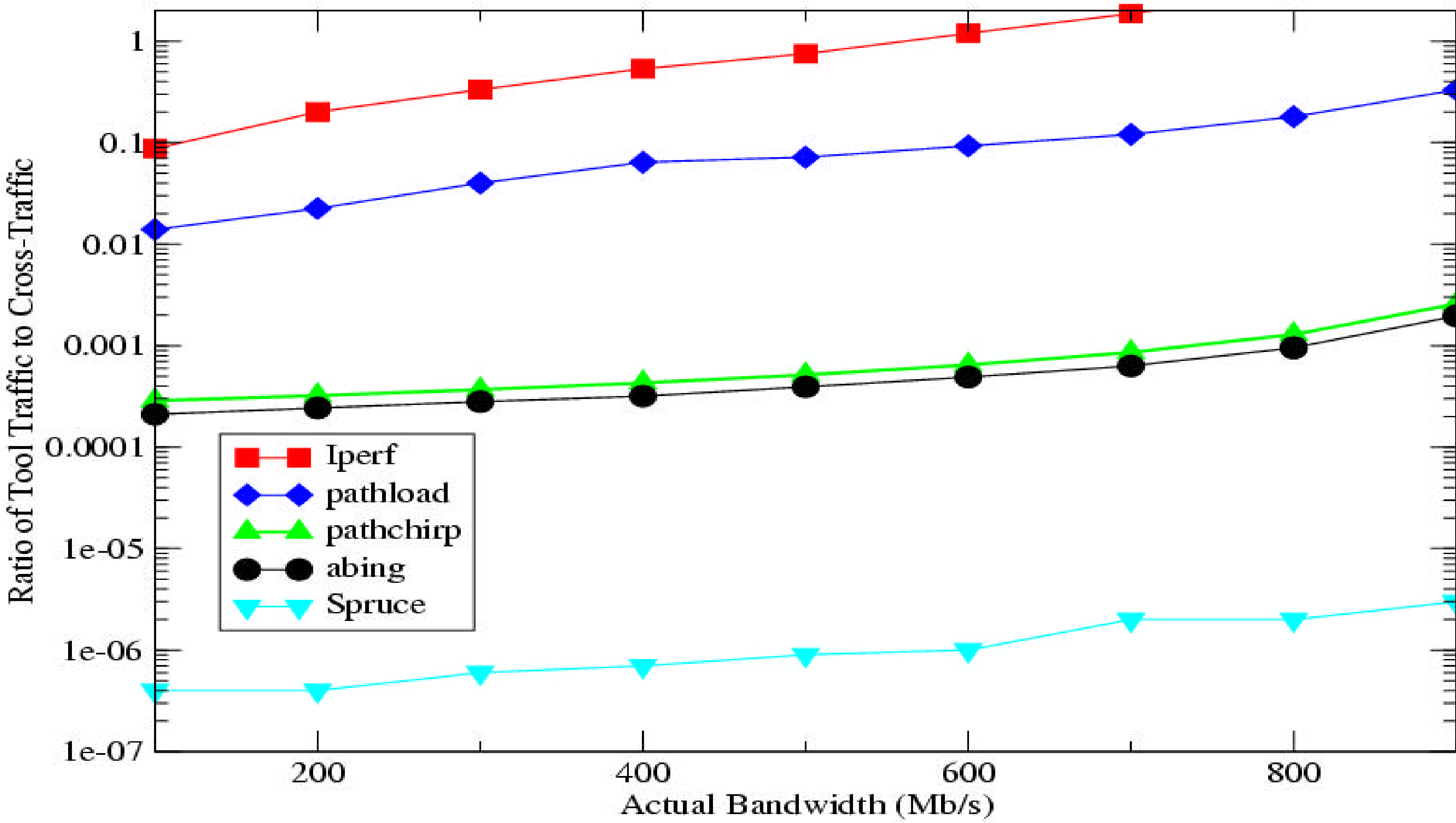
Why do Spruce and Abing perform poorly?

- Both send 1500 byte packet pairs with some interval **t** between packet pairs
- Compute AB by averaging the IAT between all the packet pairs
- Normal IAT should be 11-13 µs.
- Interrupt coalescence or delay quantization causes IAT jumps to 244 µs in some samples
- These delays throw off estimates.

Measurement Time

- Abing: 1.3 to 1.4 s
- Spruce: 10.9 to 11.2 s
- Pathload: 7.2 to 22.3 s
- Patchchirp: 5.4 s
- Iperf: 10.0 to 10.2 s

Probe Traffic Overhead Injected by tool



Outline

- Laboratory Setup
- **Part 1: Laboratory Experiments**
  - SmartBits
  - **Tcpreplay**
- Part 2: Internet Experiments
  - Abilene Network
  - SDSC ORNL

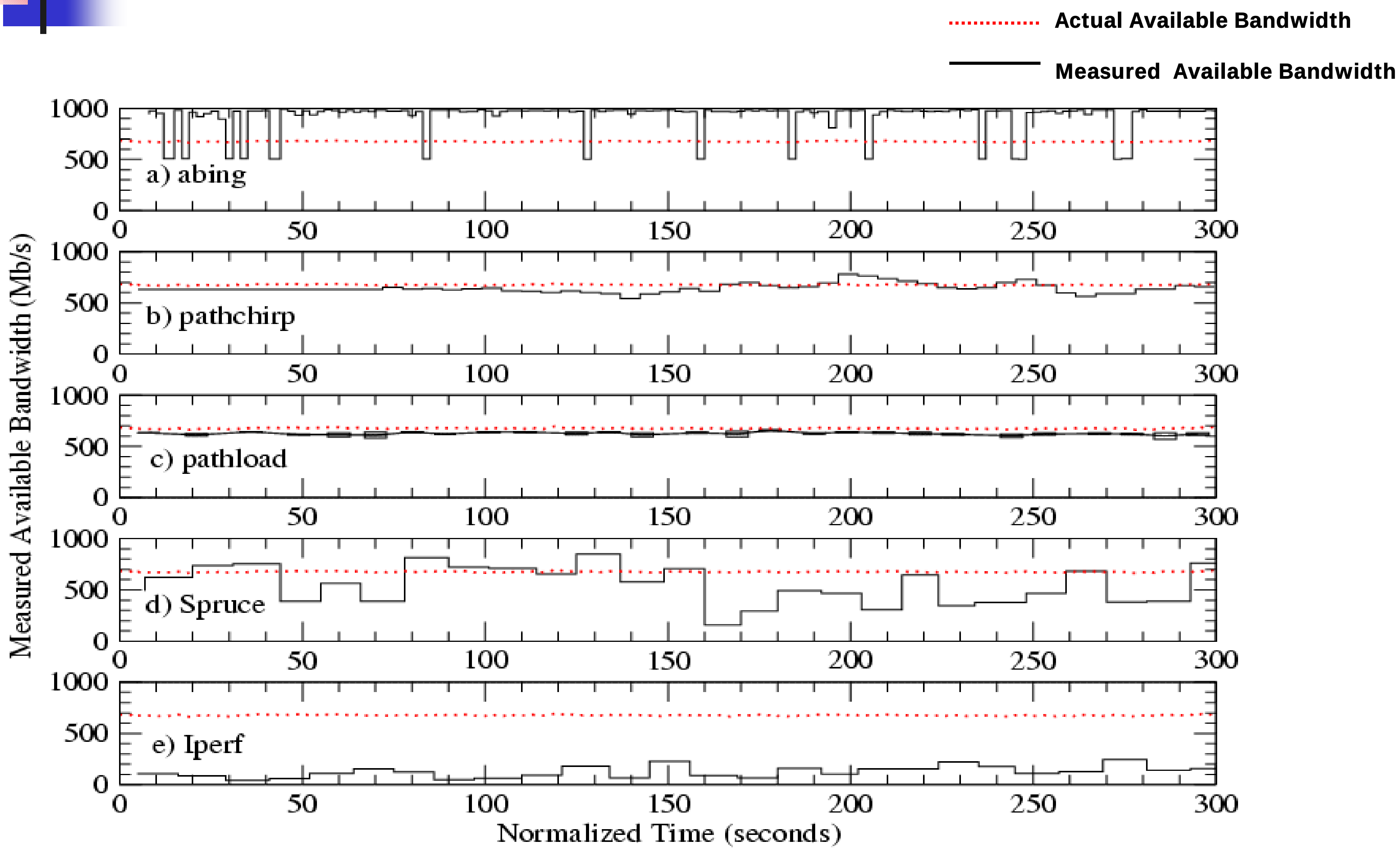
Tests with Tcpreplay

- Tcpreplay: Tool to replay pcap trace
  - IAT and Packet Size distributions identical to real traffic
  - Not congestion aware.
- Used two traces (Sonet & Ethernet)
  - Sonet: Avg Load -102Mb/s
  - Ethernet: Avg Load -330Mb/s
- Cross-Traffic flowing in one direction.

Tests with TCPreplay

- TCPreplay to regenerate trace traffic
  - One direction of the shared path
- TCPreplay cross-traffic for **6** minutes
- Run AB measuring tools back-to-back
- Plot a time-series of the measurements against the actual values of AB.

Accuracy with TCPreplay



Why Does Iperf perform Poorly?

- Iperf encounters approx 1% packet loss
- Caused by
  - Small buffers on the switches
  - Long retransmit timer 1.2 s
- Performance Improved by
  - Reducing retransmit timer
  - Bypassing the bottleneck buffer

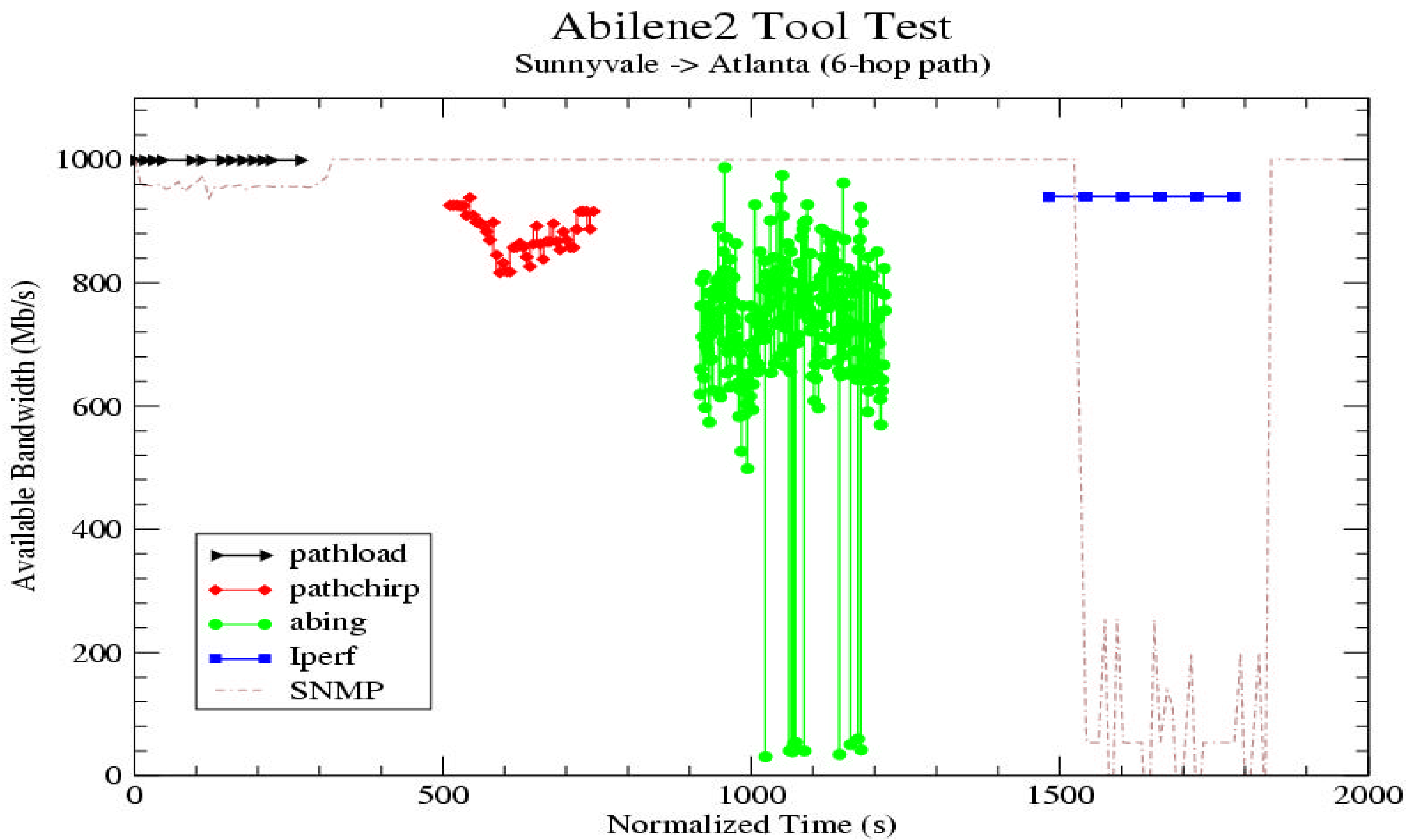
Outline

- Laboratory Setup
- Part 1: Laboratory Experiments
  - SmartBits
  - Tcpreplay
- **Part 2: Internet Experiments**
  - **Abilene Network**
  - SDSC ORNL

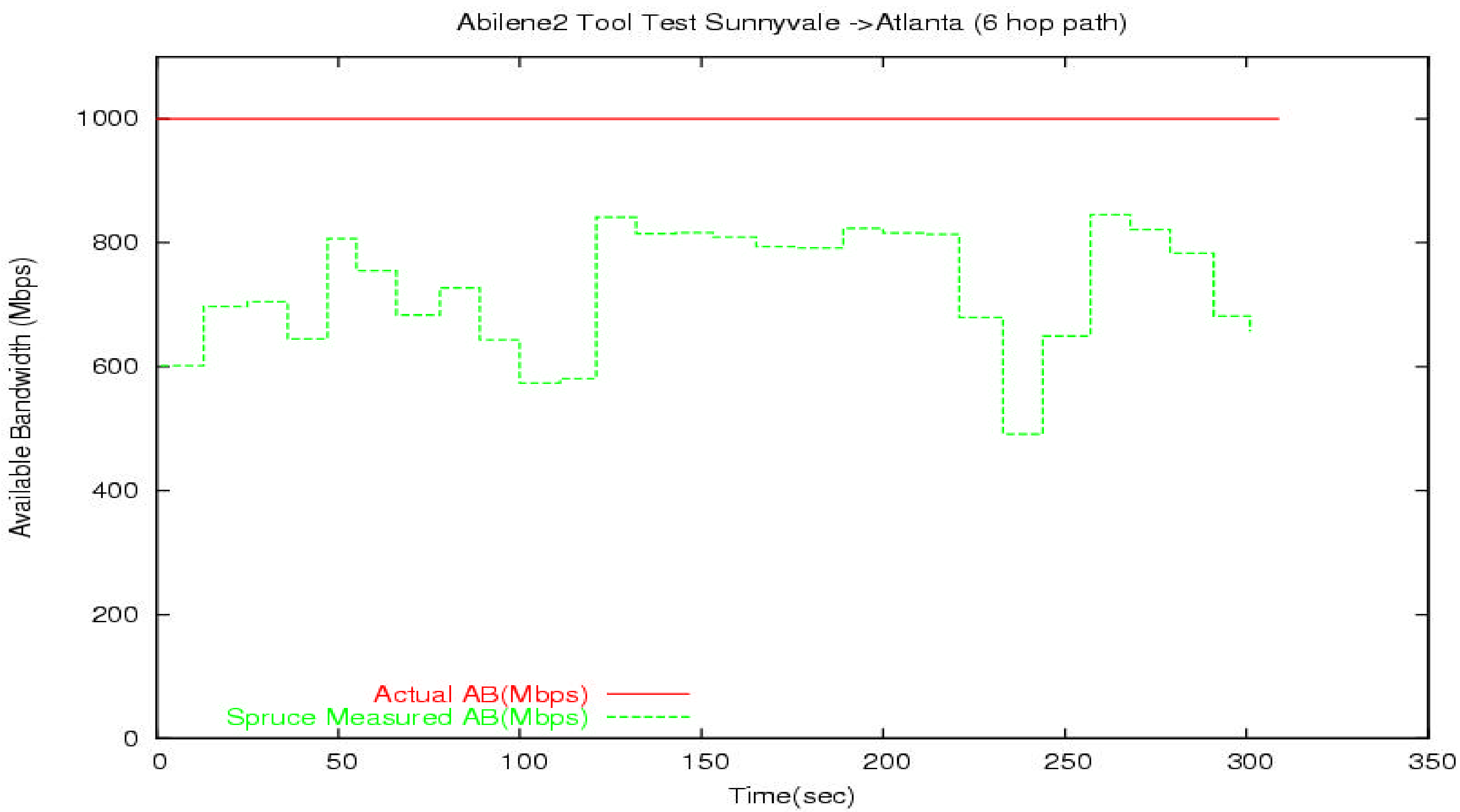
Abilene Experiment (SNVA-ATLA)

- End-to-End path on Abilene from Sunnyvale to Atlanta (5pm EST)
- 6 hop path
- Access to 64 bit InOctets for all the routers along the path
- Tight and Narrow link was the end host 1Gb/s access link.
- All other links 2.5 Gb/s and 10 Gb/s.

Abilene Experiments



Spruce run on the Abilene Path



Outline

- Laboratory Setup
- Part 1: Laboratory Experiments
  - SmartBits
  - Tcpreplay
- **Part 2: Internet Experiments**
  - Abilene Network
  - **SDSC ORNL**

SDSC-ORNL experiments

- SDSC->ORNL
  - 622 Mb/s Narrow Link
  - 1500 Byte MTU
- ORNL->SDSC
  - 1 Gb/s Narrow link
  - 9000 Byte MTU
- Assume that narrow link is the tight link
- No access to SNMP information

SDSC-ORNL path

Direction	Path Capacity, MTU		Probe Packet Size	Tool Reading Abing (Mb/s)	Tool Reading Pathchirp (Mb/s)	Tool Reading Pathload (Mb/s)	Tool Reading Spruce (Mb/s)
SDSC to ORNL	622 Mb/s, 1500		1500 9000	178/241 f/664	543 f	>324 409-424	296 0
ORNL to SDSC	1000 Mb/s, 9000		1500 9000	727/286 f/778	807 816	>600 846	516 807

Conclusions

- Pathload and Pathchirp are the most accurate
- Iperf requires maximum buffer size and is sensitive to small packet loss.
- 1500B packets and μs time resolution are insufficient for accurate measurement on high speed paths
- Delay quantization negatively affects tools using packet pair techniques like Abing and Spruce.

Future Work

- Impact of responsive cross-traffic on Available Bandwidth estimates
  - Spirent Avalanche traffic generator
- Impact of packet sizes on bandwidth estimation robustness.
- Impact of router buffer sizes on available bandwidth and achievable TCP throughput measurement .

Thank You !!!

Acknowledgements

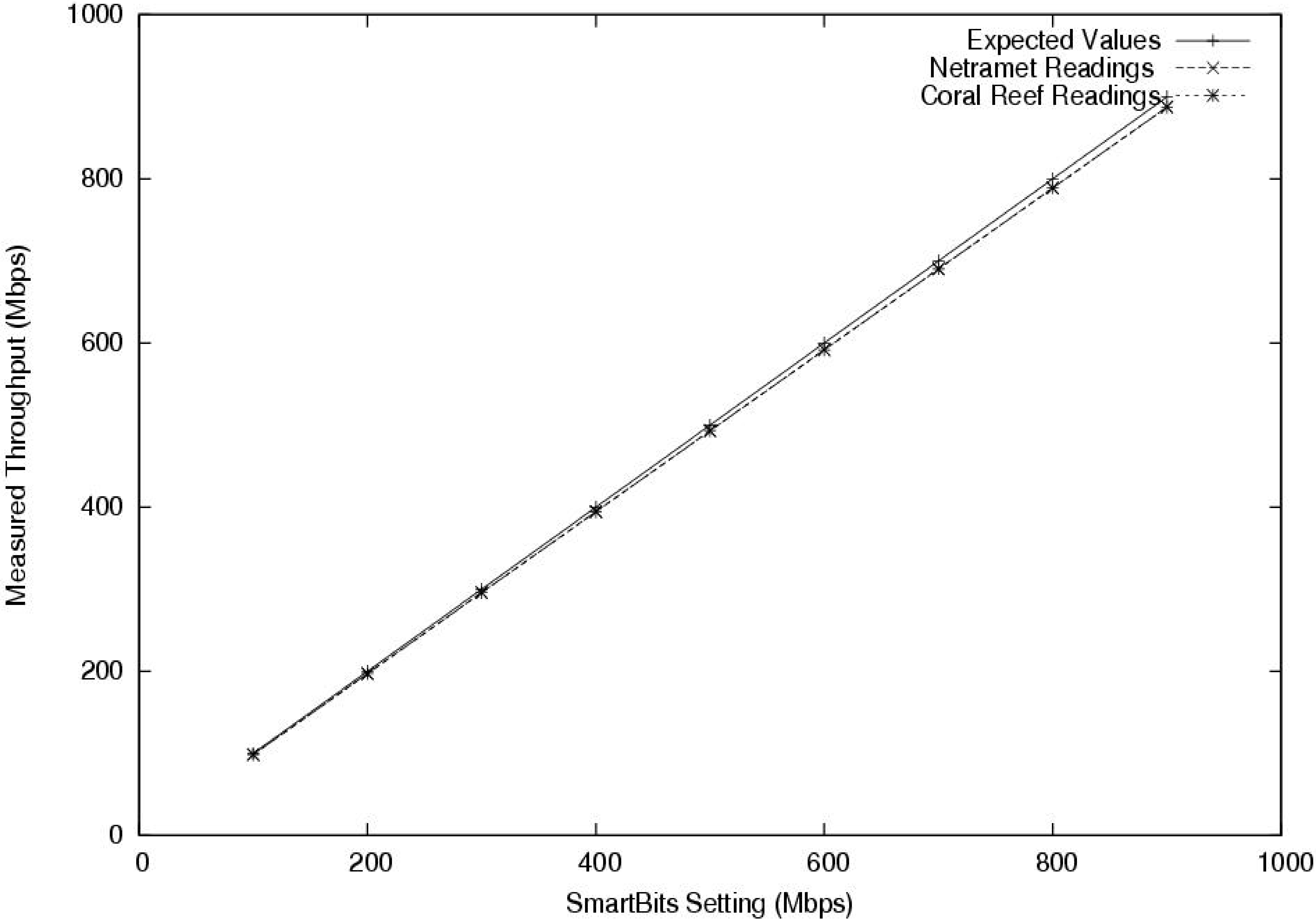
Matthew Zekauskas,

Aaron Turner,

Brendan White,

Dan Anderson,

Ken Keys


Expected Values
Netramet Readings
Coral Reef Readings
Measured Throughput (Mbps)
SmartBits Setting (Mbps)
0
200
400
600
800
1000

Cross Traffic Characteristics of SmartBits