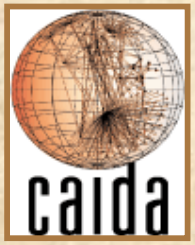
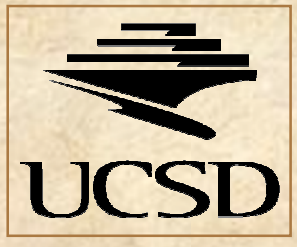
The CAIDA bandwidth estimation testbed



Margaret Murray
CAIDA/SDSC/UCSD



Acknowledgements

- Nevil Brownlee: NeTraMet config
- Johnny Chang: data collection
- Tony Lee, Tuan Le: autotest.pl, autoplot.pl
- Jiri Navratil, Ravi Prasad, Vinay Ribeiro: remote testbed users
- Grant Duvall, Nate Mendoza: router config
- Kevin Walsh: CalNGI, NPRL access
  - Spirent SmartBits 6000 with SmartFlow software
  - Foundry Big Iron router

- Cisco: GSR12008 router
- Juniper: M20 router
- Endace: gigE DAG card for passive monitoring with NeTraMet
- Department of Energy SciDAC grant DE-FC02-01ER25466

The e2e tool grail...

- Users want to push a button to:
  - select the best available e2e path
  - optimize application utilization of bandwidth
- Sponsors expect growth in usage and need for better sharing/planning

...yet many existing high-speed backbones appear to be lightly loaded, and many applications can not take advantage of fatter pipes...

Current e2e Tools

Tool Class	Tool	Authors	Methodology	Tool	Authors	Methodology
Per-hop Capacity	*clink* √	Downey	VPS	*pathchar* √	Jacobson	VPS
	pchar √	Mah	VPS			
End-to-End Capacity	*bprobe*	Carte	pkt pair	*pathrate* √	Dovrolis-Prasad	pkt pairs,train
	nettimer	Lai	pkt pairs	*sprobe* √	Saroiu	pkt pairs
End-to-End Available Bandwidth	*ABw* √	Navratil	unknown	*netest* √	Jin	unknown
	cprobe	Carter	pkt trains	*pathload* √	Jain-Dovrolis	SLoPS
	IGI √	Hu	SLoPs			
Bulk Transfer Capacity	*cap*	Allman	emulate TCP tput			
	treno	Mathis	std TCP tput			
Achievable TCP Throughput	*iperf* √	NLANR	|| TCP connect	*ttcp*	Muuss	TCP connect
	Netperf	NLANR	|| TCP connect			

Why is it so hard to measure e2e performance?

- Are apps tuned for high-speed paths?
- Router slow paths, load balancing
- MTU mismatch
- NIC interrupt coalescence
- Host OS variations
- Router OS variations
- the list goes on…

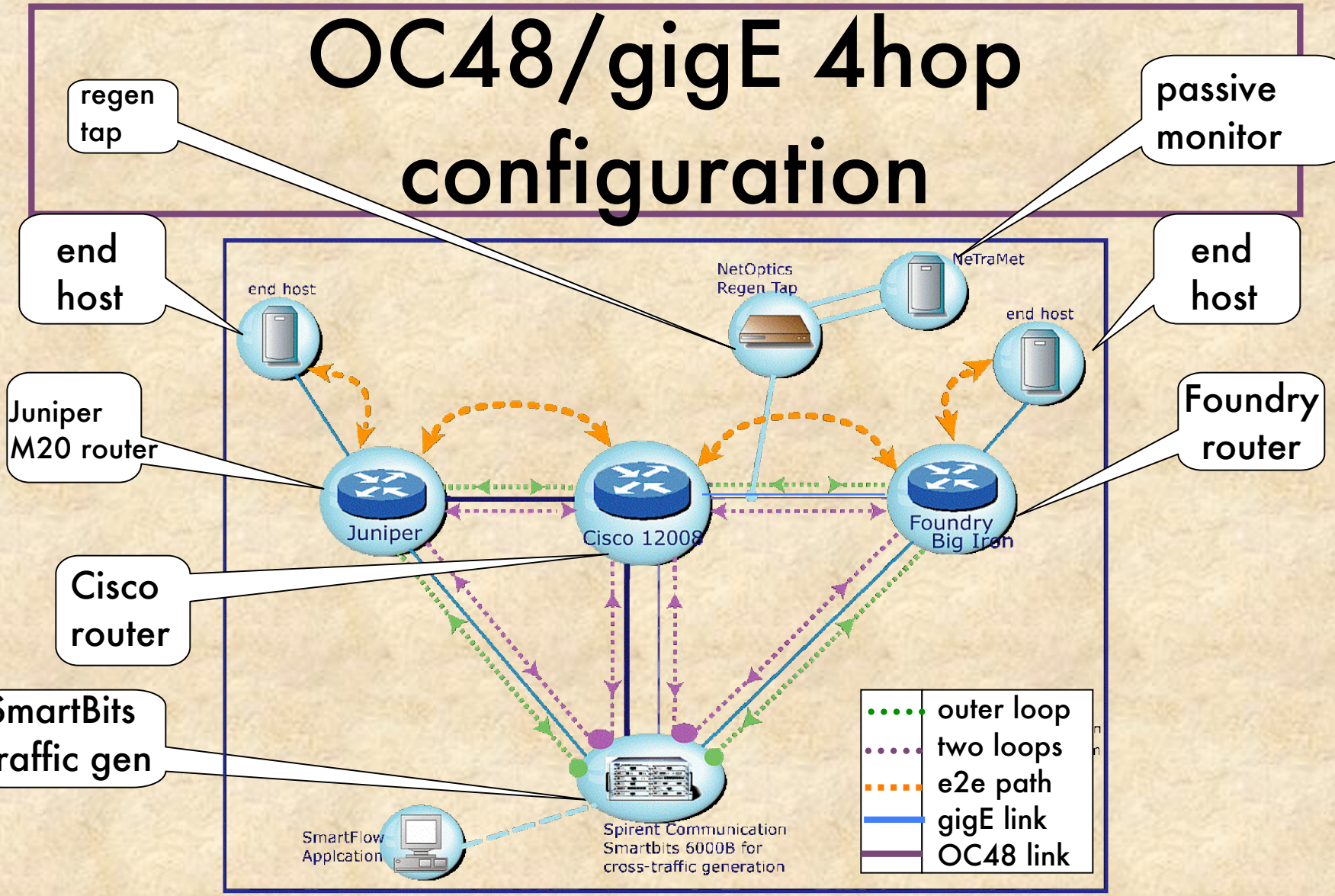
Why use a bwest testbed?

- Use reproducible test conditions
- Test against saturated links
- Test "black box" e2e tools against same scenarios
  - Identify conditions where tools work well
  - Give developers an environment for refining their tools

Why use CAIDA/SDSC testbed?

- Take advantage of CAIDA and CalNGI vendor contacts and equipment
- Integrate with CAIDA passive monitoring technology (NeTraMet, CoralReef)
- Possibilities in the future: connect testbed to specific networks (via SDSC)

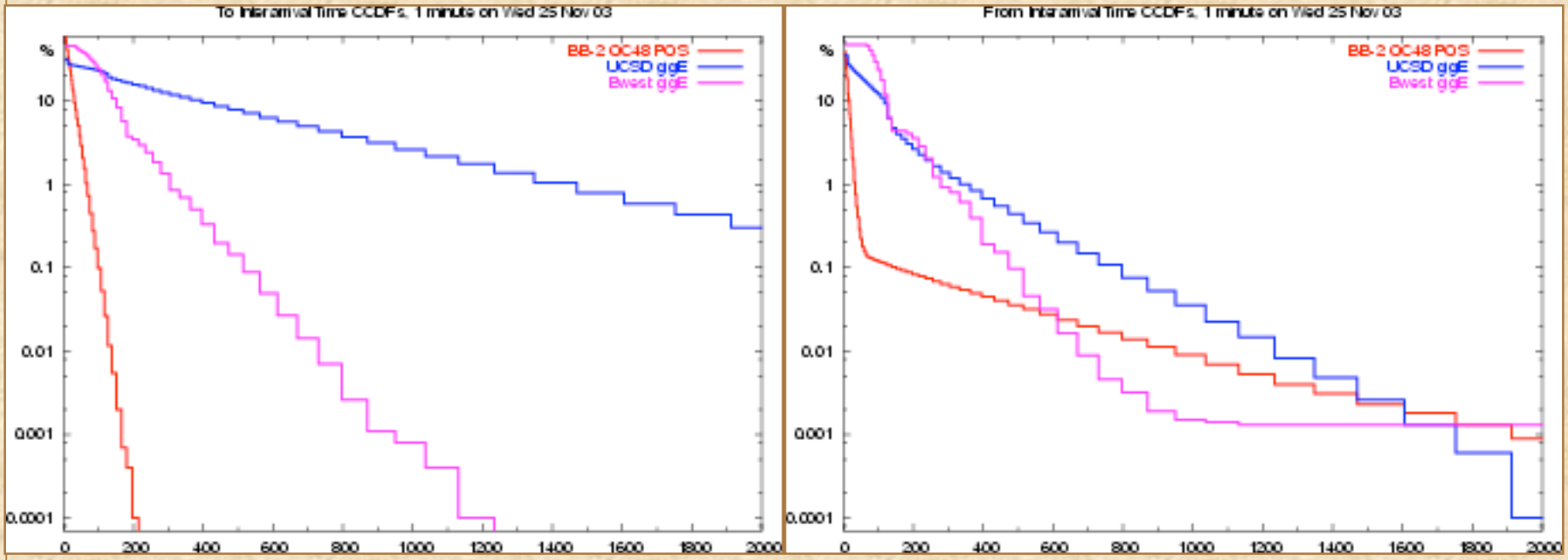
OC48/gigE 4hop configuration



Working with Traffic (real vs. generated)

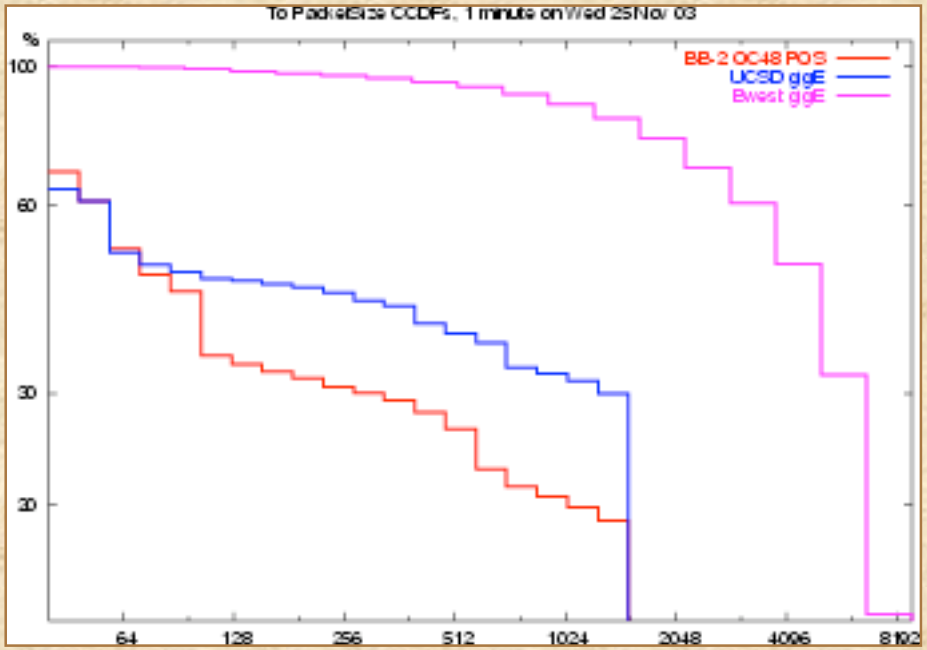
Real Traffic	
Pro: • no question of validity	Con: • ***hard to find, store and use traces of interest***
Generated (Simulated) Traffic	
Pro: • easily reproducible scenarios	Con: • ***how realistic is traffic?***

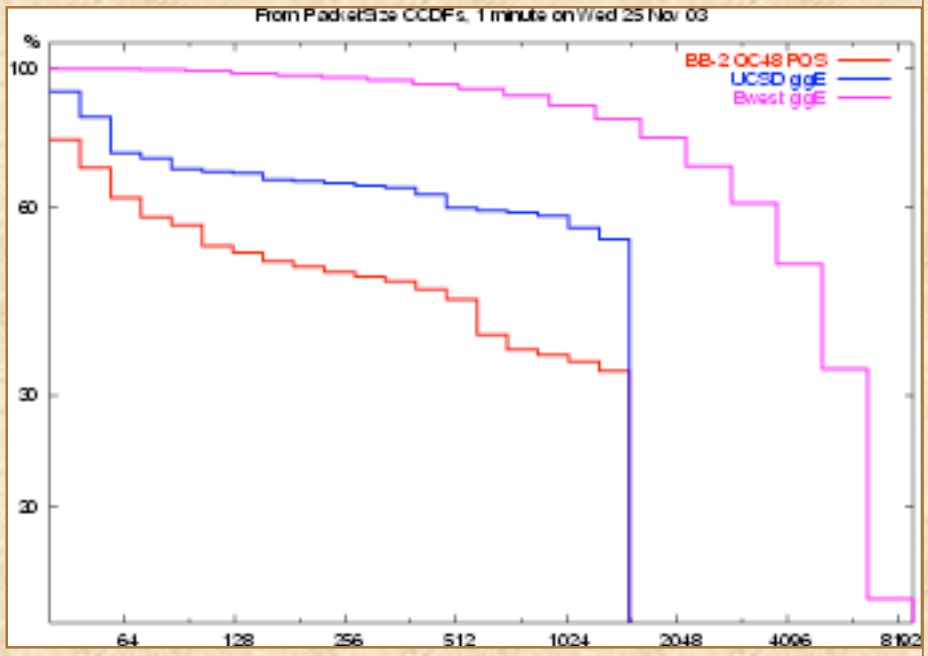
Comparison of Packet IATs (*CCDF: 1min. samples,% vs. IAT usec*)



- **Red** = BB OC48 (Left: 246M, 9%) (Right:686M, 27%)
- **Blue** = UCSD gigE (Left: 40M, 4%) (Right:240M, 24%)
- **Pink** = bwest lab (Left: 593M,59%) (Right:589M, 59%)

Comparison of Pkt Sizes (CCDF:1 min samples,% vs. PktSize b)



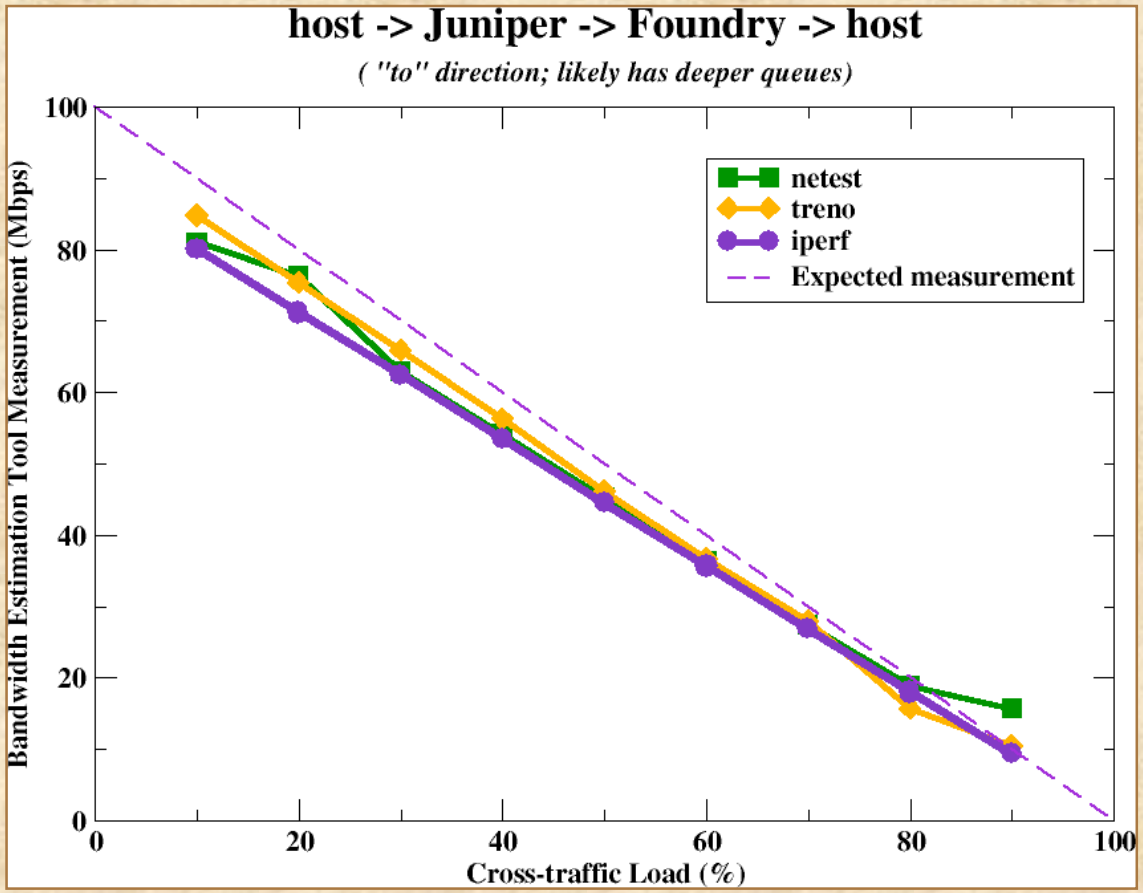


- **Red** = BB OC48 (Left: 246M, 9%) (Right:686M, 27%)
- **Blue** = UCSD gigE (Left: 40M, 4%) (Right:240M, 24%)
- **Pink** = bwest lab (Left: 593M,59%) (Right:589M, 59%)

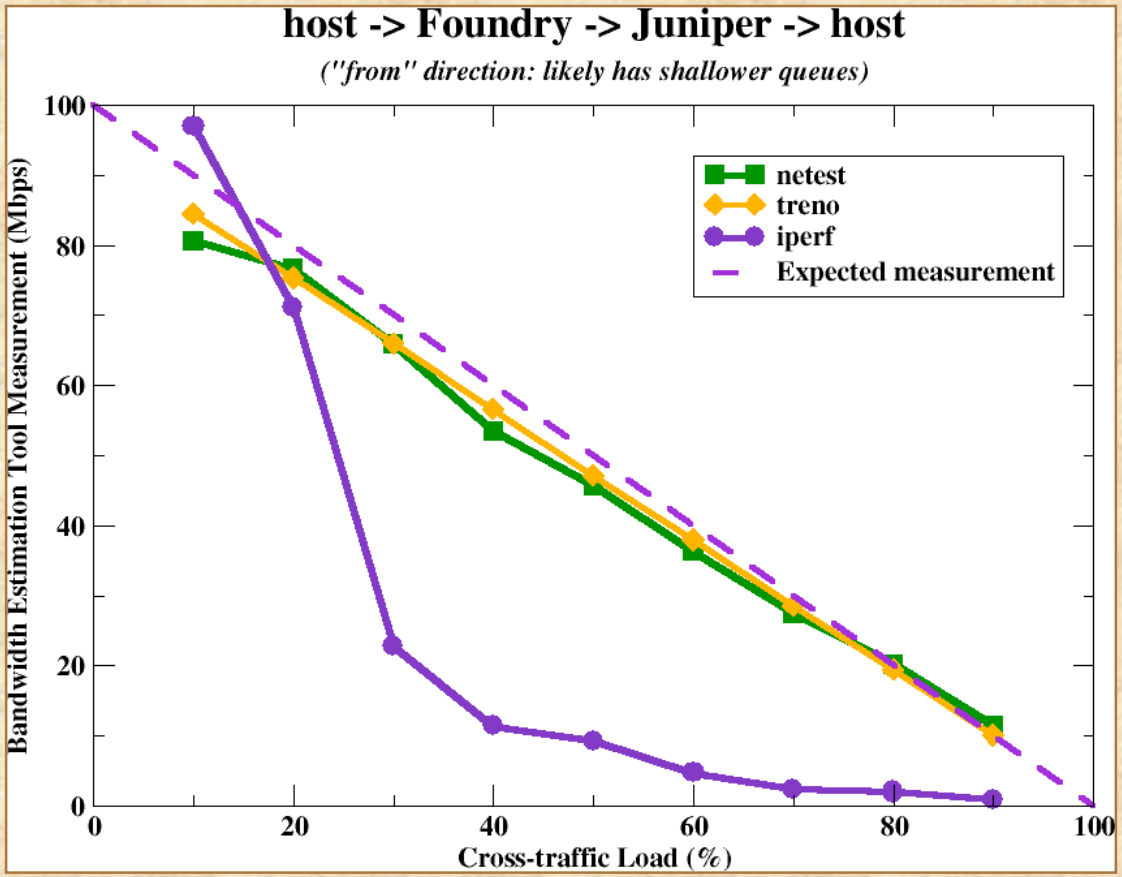
Cross-traffic

- Tool developers told us that generated traffic needs enough packet dispersion to allow dynamics to manifest: We do that.
- SmartFlow traffic has appropriate protocol headers, but does NOT emulate TCP congestion control.
- While most current networks do not implement jumbo MTUs, they will, and we can generate them.

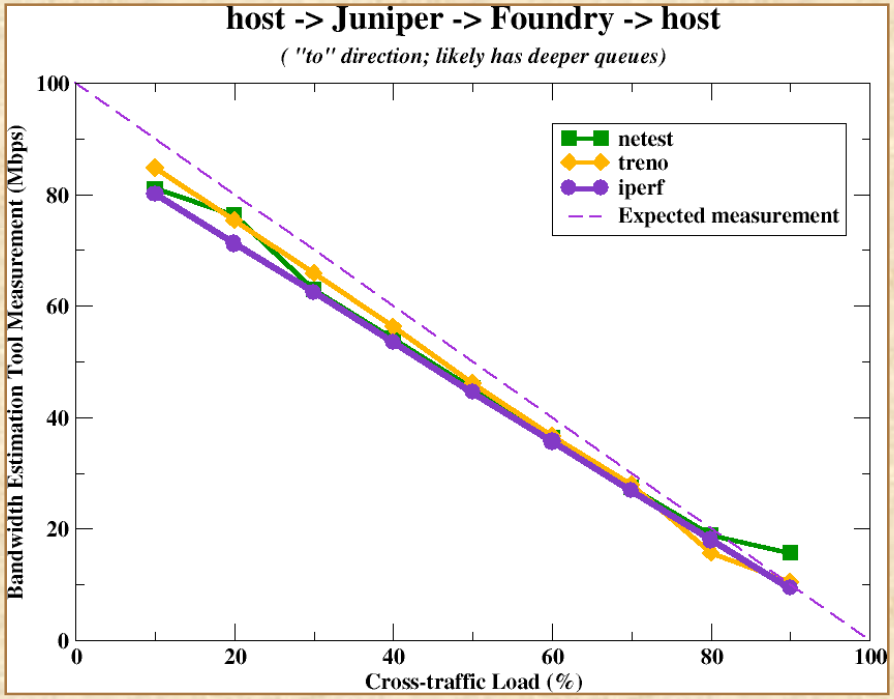
100M 3hop test

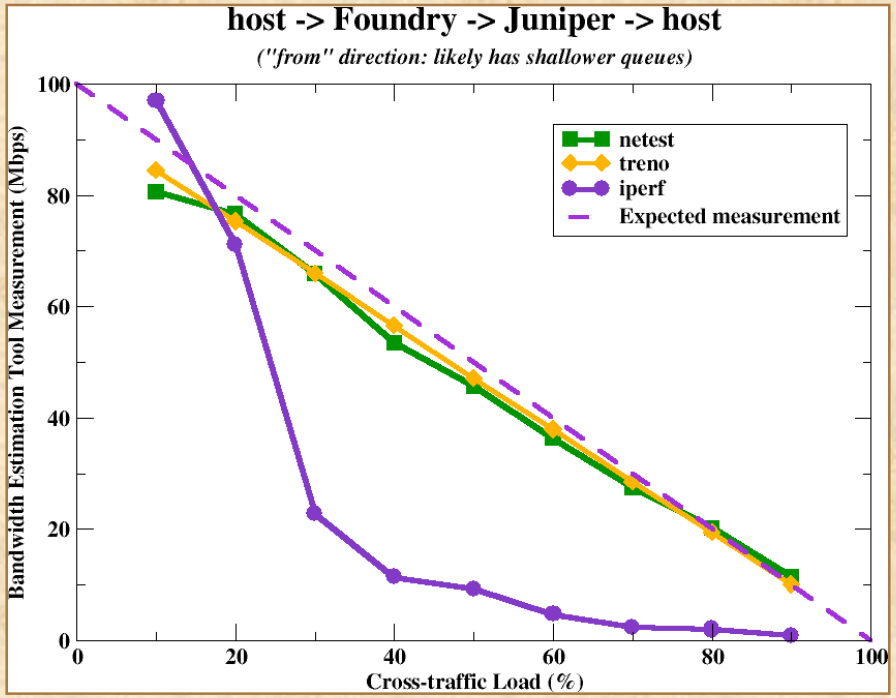


100M 3hop test surprise!



Hypothesis:
router queue depth affects e2e tools



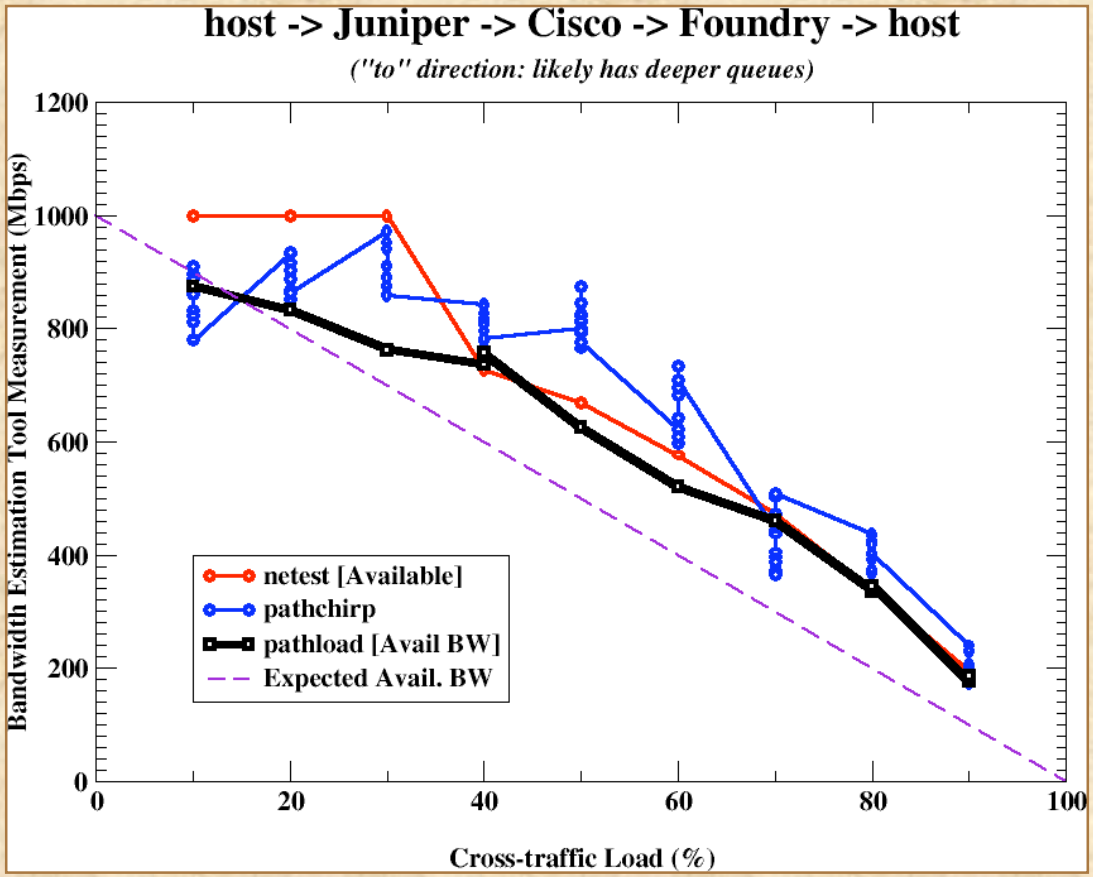


- Juniper M20
  - 9M

...we plan to run single router tests to further investigate why

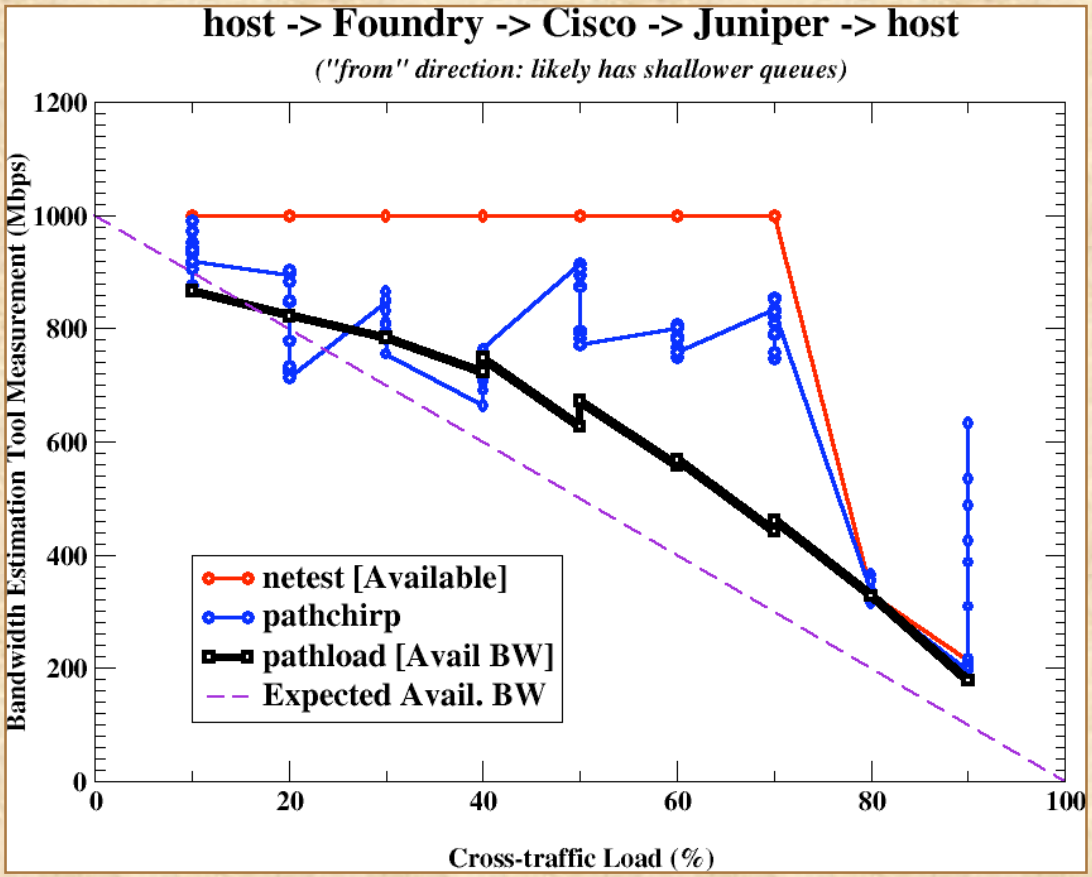
- Foundry Big Iron
  - 1M

OC48/gigE 4hop "to" direction



- Three tools give decreasing measurements as cross-traffic increases.
  - *netest*
  - *pathchirp*
  - *pathload*
- Note: pathchirp and pathload return multiple results

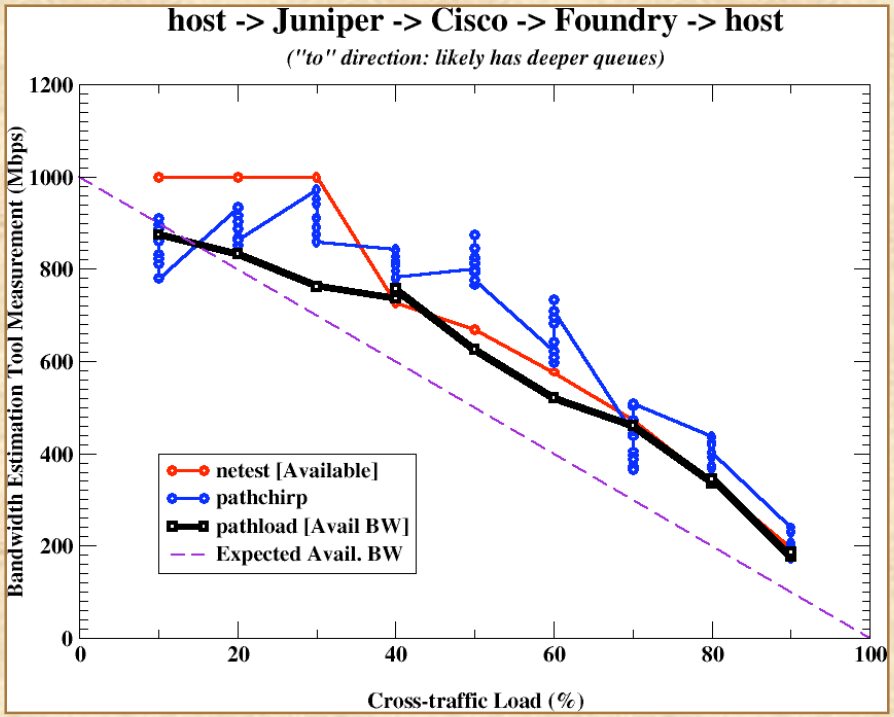
OC48/gigE 4hop "from" direction

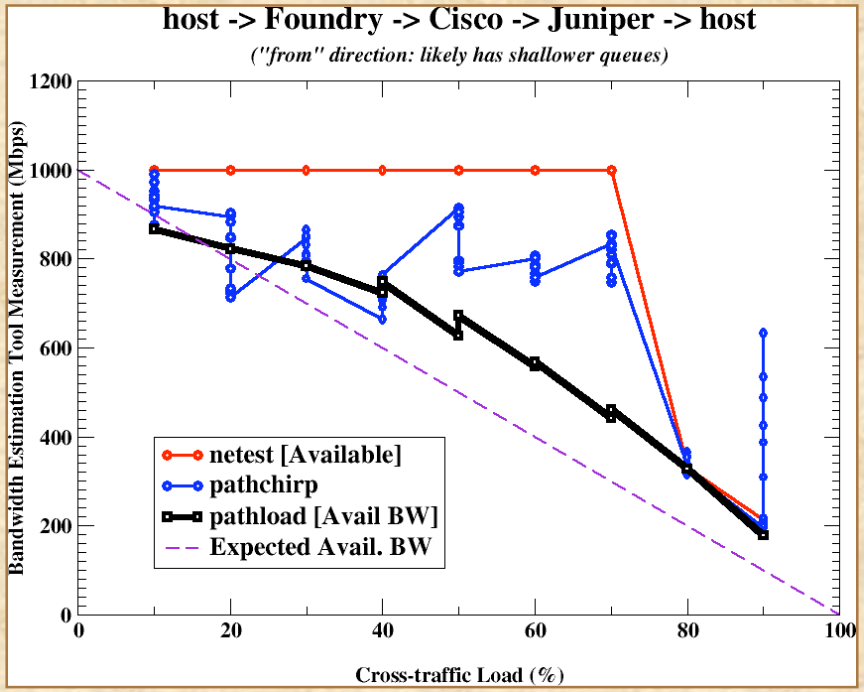


- Different results in opposite direction for
  - netest
  - pathchirp
- Hypothesis: due to *shallower router queues*
- Note multiple results for pathchirp and pathload

gigE/OC48: 4hop test again

only pathload returns similar results in both directions



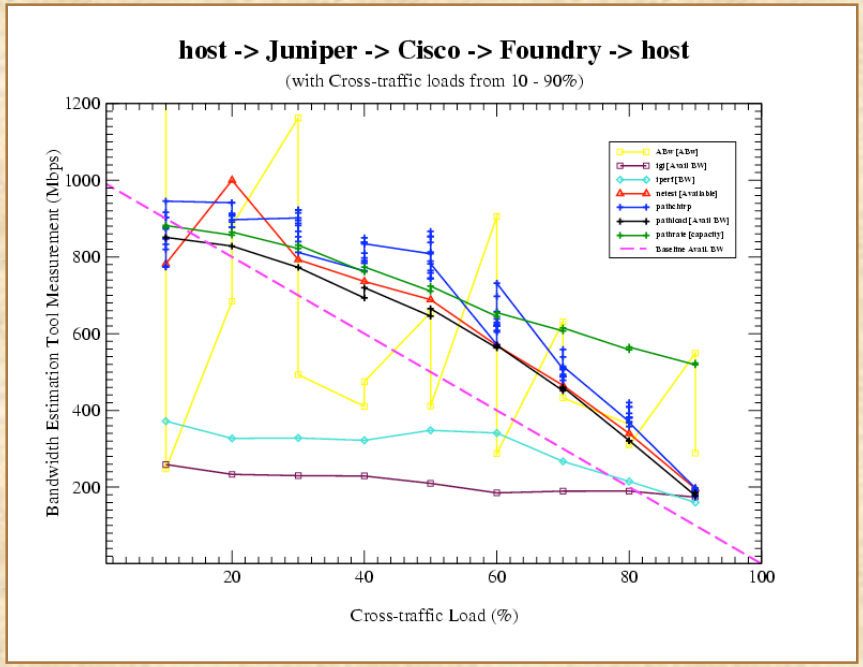


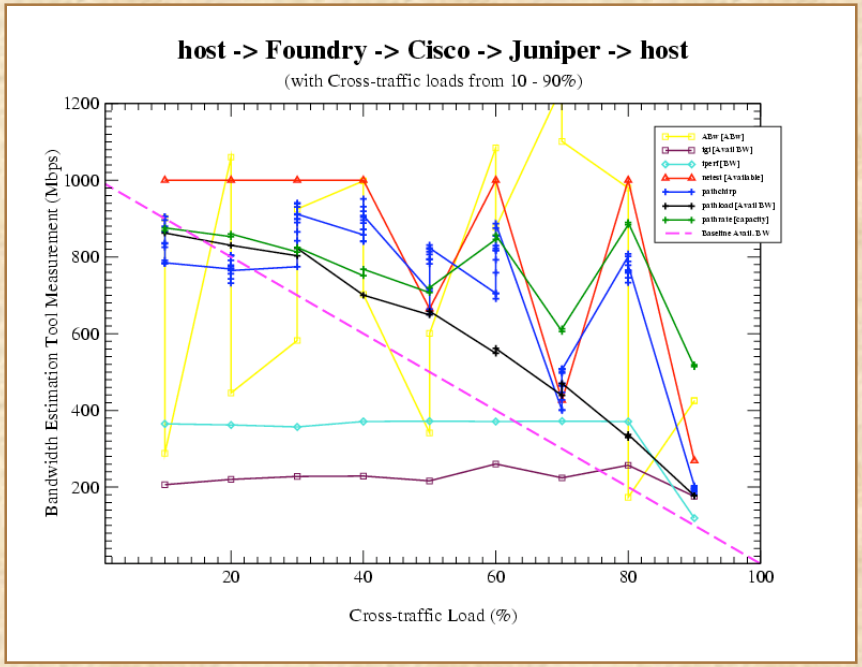
- Juniper M20
  - 9M
- Foundry BigIron
  - 1M

...we plan to run single router tests to further investigate this phenomenon.

gigE/OC48: 4hop test

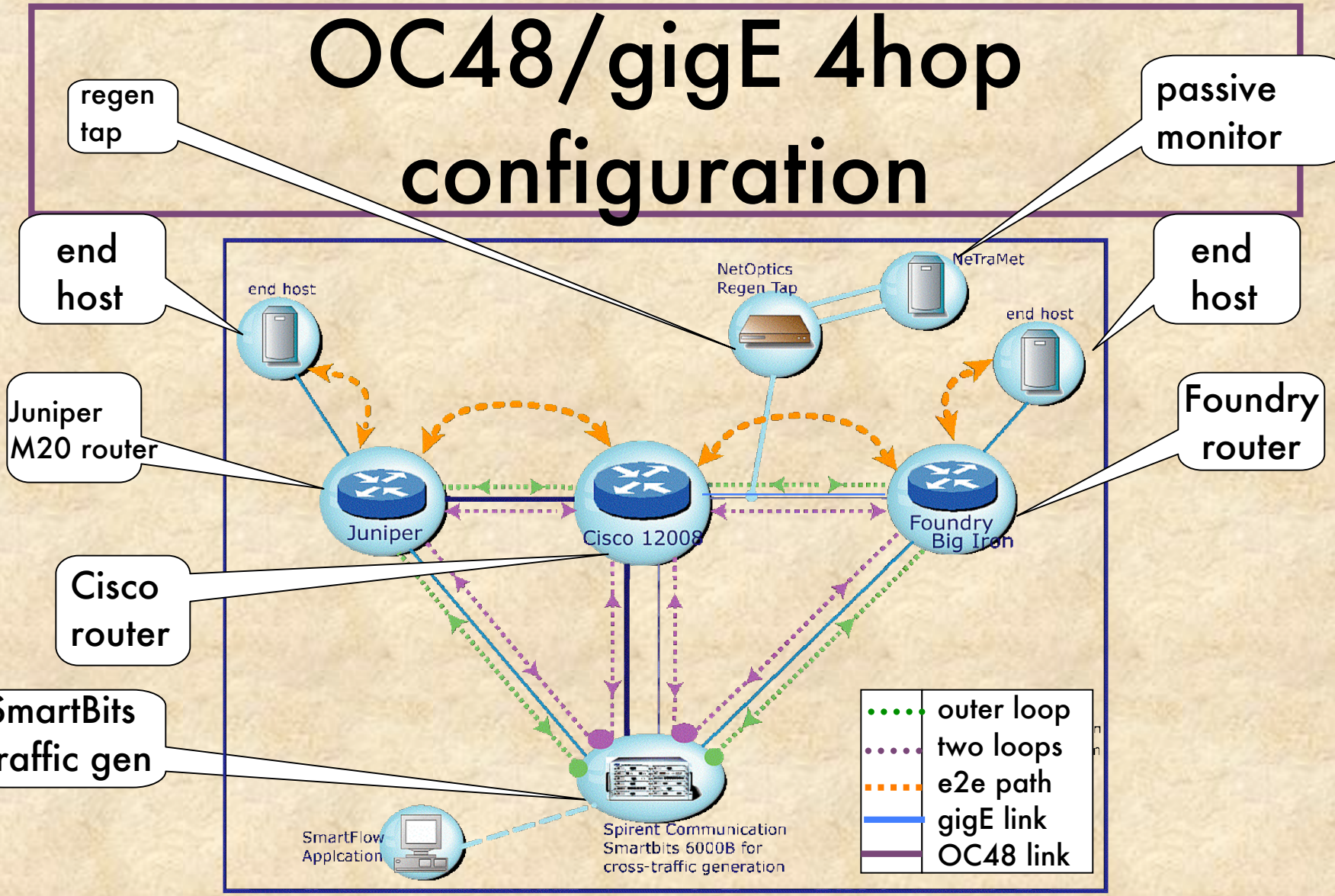
...more investigation needed to understand tool behavior





- Juniper M20
  - 9M
- Foundry Big Iron
  - 1M

OC48/gigE 4hop configuration



Remote Access to Testbed

- ssh wednesday.sdsc.edu, then
  - ssh gomez
  - ssh fester
- VNC control of SmartFlow
  - several config files available, or
  - create your own
- /usr/local/bwest/
  - autotest.pl can run and time one or all tools
  - autoplot.pl can plot IAT and Packet Size distribution

How can the CAIDA bwest testbed address e2e problems?

- generate traffic; saturate high-speed paths
- experiment with different MTU settings
- evaluate different NICs
- try different end host OS configs
- evaluate impact of different routers
- isolate router slow paths, load balancing
- the list goes on…

For more information…

- Contact Marg
  - marg@caida.org
  - 858 534-8928
- Scheduling
  - first come, first served, or ???
- Support
  - DOE grant ends Aug 2004