

# SC24 Scitags Demo

Tristan Sullivan  
LHCONE Meeting Manchester  
March 19/25

Randall Sobie  
HEPNET, University of Victoria

# SC24

- SC24 was held in Atlanta Nov. 17-22
- Scinet: Global network built specifically for SC
- Networking Research Exhibitions (NREs): Demonstrations of new network features, cutting-edge performance, etc.
- See **my talk** at Beijing LHCON Meeting Oct. 2024



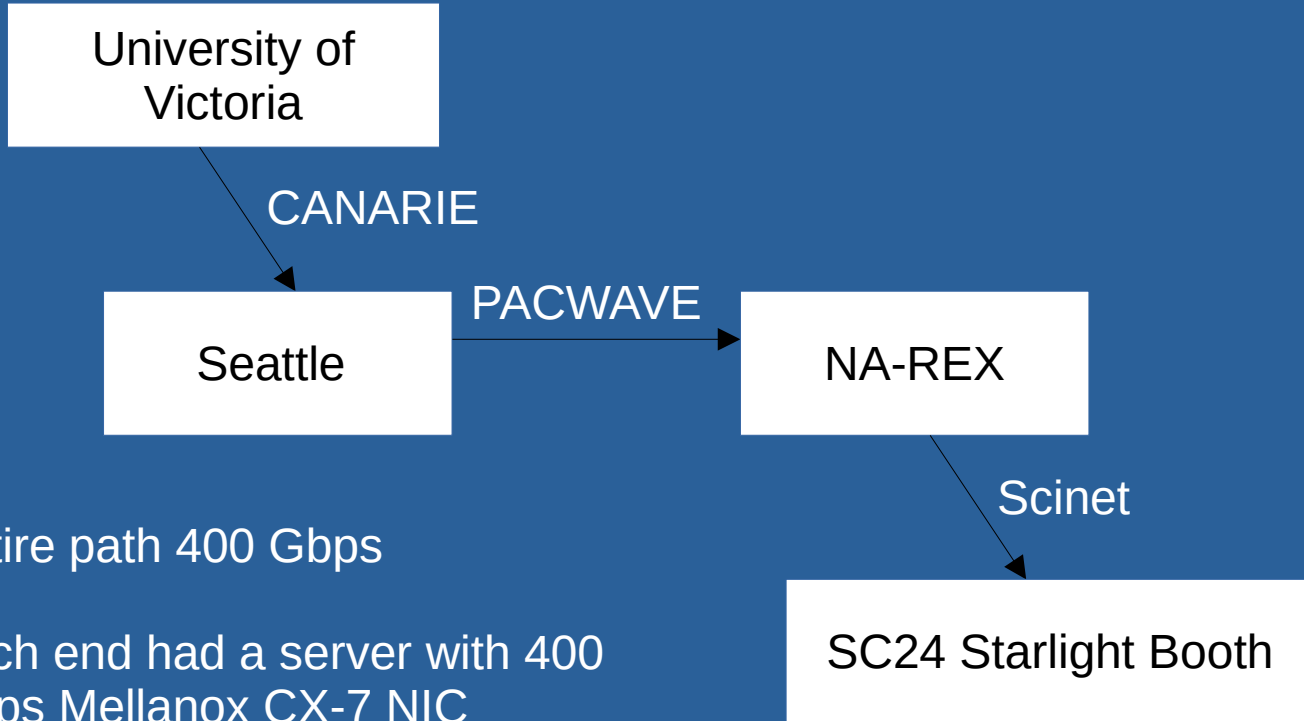
SC23 Network Map

# Scitags

Offset	Octet	0								1								2								3							
Octet	Bit	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
0	0	Version				Traffic class				Flow label																							
4	32	Payload length																Next header								Hop limit							

- Scitag added by eBPF when the kernel sends a packet to the NIC, not at the time of packet creation
- Overwriting flow label lightest-weight option
- Adding extension header requires changing next header field and inserting new header
- Next header values:
  - 0: Hop-By-Hop
  - 6: TCP
  - 60: Destination Options

# Network Path and Demo Plan



Entire path 400 Gbps

Each end had a server with 400 Gbps Mellanox CX-7 NIC

UVic received equipment donated from Dell, Lenovo, and Nvidia

Demo plan was three-fold:

- Try to reach 400 Gbps memory-to-memory transfers with iperf3
- Test performance with flow-label marking with eBPF
- Test performance with extension-header marking with eBPF

Two-hour time window for demo: forty minutes for each case

# Results, Day One



Monitoring bandwidth, CPU usage, and memory usage on UVic server  
Effect on bandwidth of the two marking strategies is visible in bandwidth plot (top-left)  
CPU usage impact is small, memory negligible

# Results, Day Two



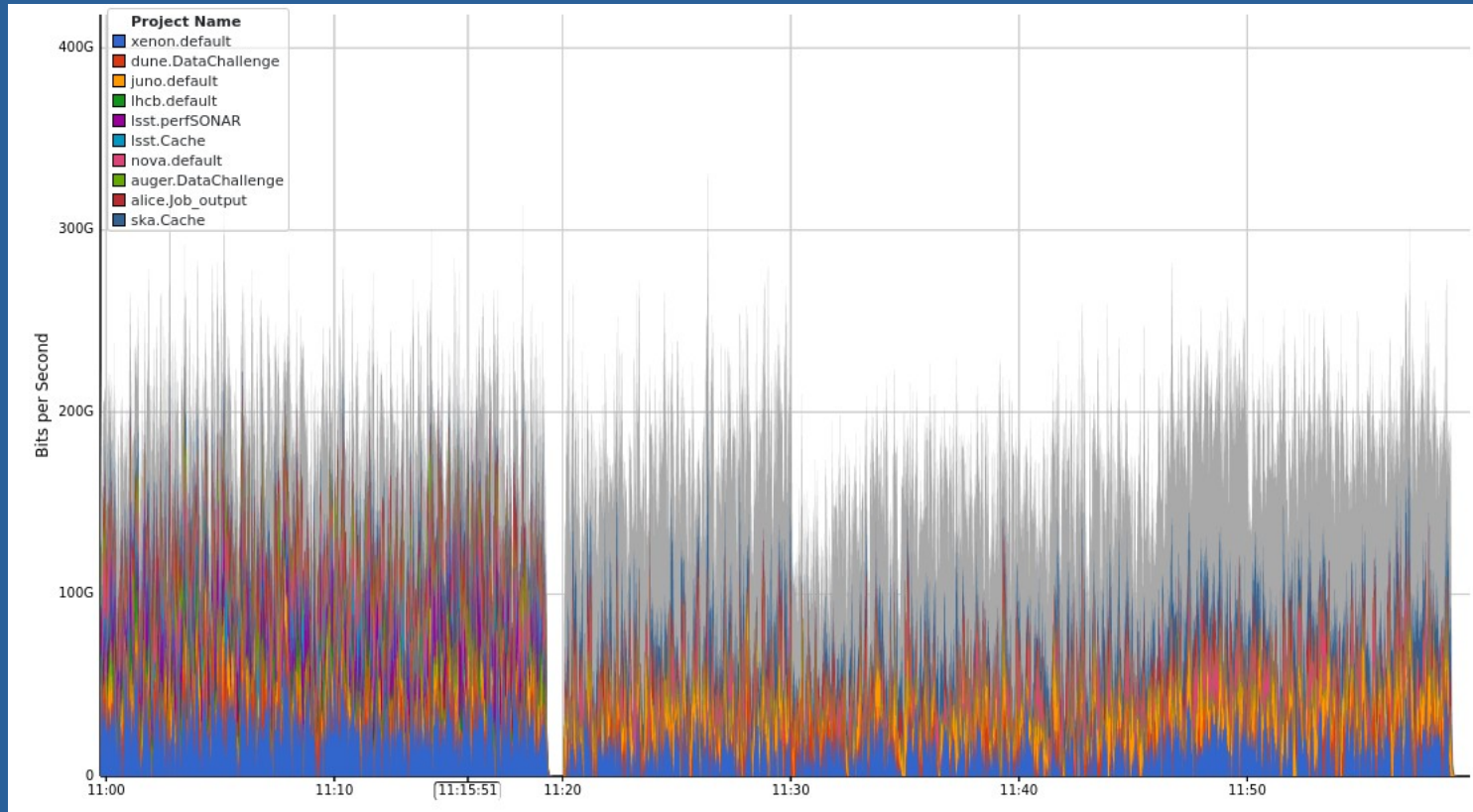
Try capping bandwidth at 192 Gbps to see if EH marking has the same choppy bandwidth; turns out it does

# Results, Day Three



Same as first day, but try switching directions for first five minutes (yellow in top-left)  
SC to UVic gives higher bandwidth than UVic to SC  
Plan to investigate with Starlight 400G servers – sometime before SC25...

# Sflow Monitoring



Sflow monitoring sent by UVic server to a collector, courtesy of InMon

Transfers used sixteen iperf3 streams, each of which had a random Scitag assigned

# Recent Progress and Conclusion

- Flowd-go developed by Pablo Soto; very nice reimplementaion of packet and flow marking daemon in Go
- I have deployed flowd-go on a perfsonar machine I manage at UVic for Belle-II, and configured it to use flow label marking with eBPF. As far as I know, first actual marked packets in any kind of production use
- Room for improvement in SC24 demo bandwidth performance, but sufficient to show that eBPF marking is viable at high rate