On Terabit Flow Analysis
FloCon 2008, Savannah

Jonathan M. Smith
CIS Department, U. Penn
Terabit Network Applications

• Full-fidelity remote visualization and interactive simulation for 80fps HD / 3D HD and beyond, support for holographic visualization

• High-speed sensor data from science experiments

• Immersive simulations and high-fidelity massively multiplayer virtual worlds

• Receive and analyze many concurrent high-fidelity streams of video and/or sensor data - multiple uses in public safety, financial services and other domains
Challenges for Flow Analysis?

• New kinds of traffic:
  – Extremely High Data Rates
  – Long flows
  – New patterns with P2P and sensors

• Correlation - obtaining the “high ground”
  – Rare events vs. attenuated sampling?

• New analysis possible with DPI

• Goal: ingest, record and analyze it all!
Tradespace: data rates vs. analysis

The “high ground”: high aggregation plus high data processing rates
The Terabit Chokepoint

Problem/Challenge: Network chokepoint (I/O and memory) between fibers and CPUs.
Today’s Single-Core PC Performance Measurements
(Using UBUNTU Linux “MEMTEST” utility)

- **L1 Cache:** 180Gb/s
- **L2 Cache:** 100Gb/s
- **DRAM:** 16Gb/s
Challenge of Dense Wavelength Division Multiplexing (DWDM)

- Fiber bandwidth is serial bit rate multiplied by number of wavelengths
- E.g., 128*40Gbps in deployed systems (128 lambdas of OC768c SONET)
Processing Must Scale with Fiber Capacity

- Parallel processing seems necessary
- Memory/processing elements to track line rates and number of wavelengths?
Many-Core CPU/GPU Future

• Parallelism floodgate unleashed
  – GPUs and CPUs converging
• Teraflop+ performance in 2009
  – E.g., 32 cores @ 2Ghz
  – 16-element “short” vectors
  – 100 terabit/sec aggregate register bandwidth
  – 1 terabit/sec GDDR3 memory bandwidth
• How do we feed it?

80-core Intel test chip
Technical Approach

- Constraints: pins, power, cost
- Switch-based interconnects, parallel paths
  - Direct network/processor interface?
- Stream/graphics engines, banked memories
  - Special high-end pool of DRAMs for NICs?
- New software structures for multicore
Components looking good - architecture needed

- 1 TB (8 Tbps) memory technologies announced. Fiber good to >10Tb/sec
- 80-1000 cores @ 1-10 Gbps each
- Major challenges: fiber/electronic boundary, data distribution, interconnection network architectures (see, e.g., Dally+Towles)
Even *more* processing to scale with fiber capacity?

- *Parallel processing* at both multicore (perhaps NPUs?) and “box” level
- Cores track line rates, while degree of “box” parallelism matched against grosser units of wavelengths, *e.g.*, 8:
Help architects to help you

- Computer architects (see Proc. ISCA, Micro, ASPLOS, HPCA, …) evaluate proposals with benchmarks
- Media benchmarks are being developed
  http://euler.slu.edu/~fritts/mediabench/
- Flow analysis needs benchmarks for flow analysis tasks - input side, not just netflow outputs (this is after the fact)
Summary

• The future is in parallelism
  – Dense Wavelength Division Multiplexing (DWDM)
  – On-chip networks for multicore
  – Trees for “box”-scale parallelism

• Huge challenges remain
  – Software for new parallelism / media stream analysis; topological choices (e.g., Batcher-Banyan + Crossbar?); load-balancing algorithms

• Need to get flow analysis workloads on computer architecture radar
Acknowledgments

• “Terabit Edge Research Activity” (TERA), joint work with Milo M. K. Martin of U. Penn, supported by DARPA/IPTO

• “Advanced Broadband Intrusion Detection Engine” (ABIDE), joint work with M. B. Greenwald and E. Lewis, supported by ARO