Using Flow for Realtime Traffic Management in 100G Networks

John Gerth
Johan van Reijendam

Stanford University

Ethernet Speed Evolution

- 1970's "thick" ethernet
  - 3 Mbps over 10Base5 coax
- 1980's "thin" ethernet
  - 10 Mbps over 10BaseT coax
- 1990's "fast" ethernet
  - 100 Mbps over Cat5
- 2000's "gigE" and "ten gig"
  - 1 and 10 Gbps over Cat5e/6 and fiber
- 2010's "QSFP…"
  - 40 and 100 Gbps

STANFORD
COMPUTER SCIENCE

Realtime challenges

- Network bandwidth
  - Now rivals I/O bus speeds
- Processor speeds stagnant
  - Multi-core CPUs
  - Hyperthreading
- Memory
  - Local memory per CPU socket
  - Non-local memory has access penalties

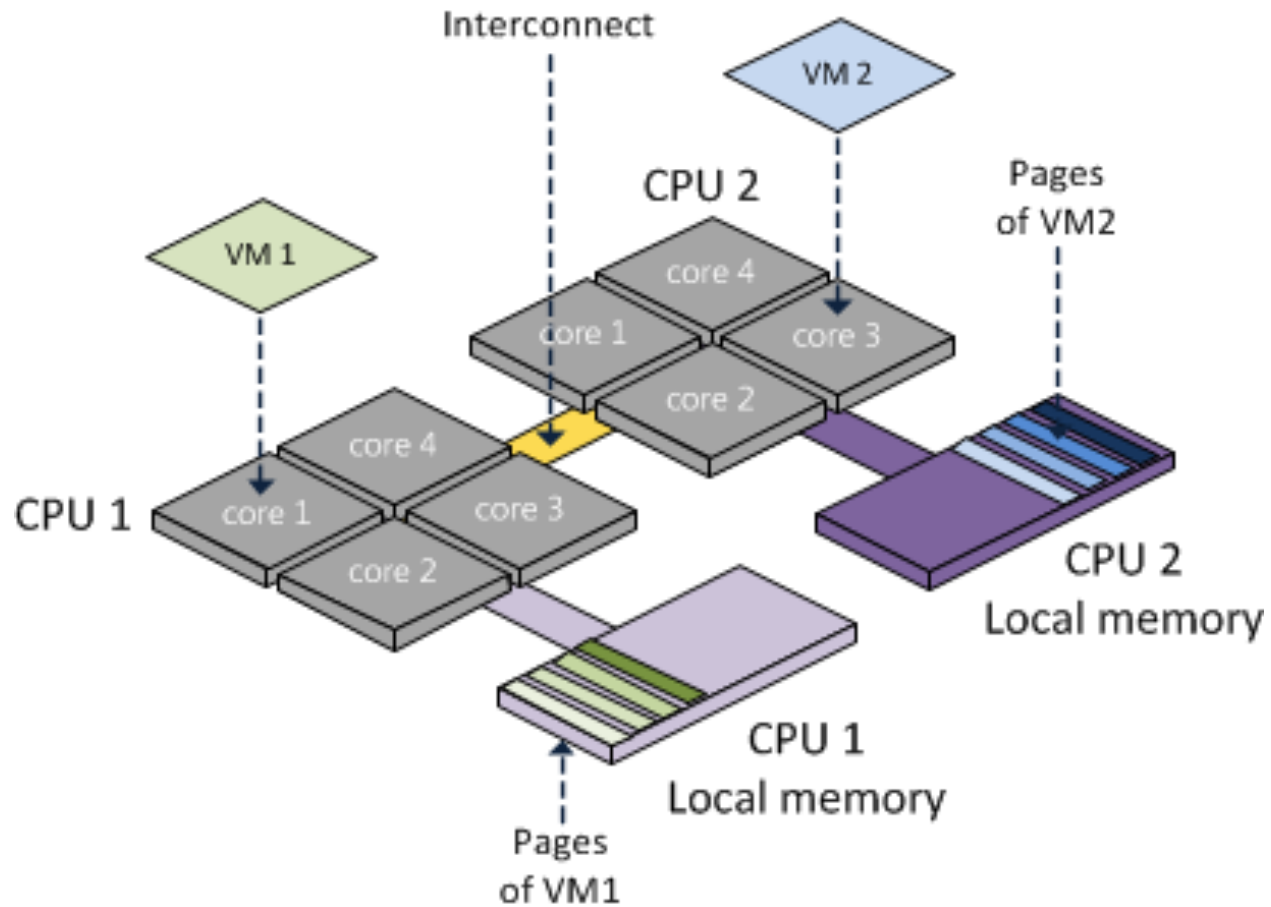STANFORD
COMPUTER SCIENCE

Sensing Design for 100G

- ## NIC
  - Full-duplex link max bandwidth 2x100G
  - PCIe gen3 max bandwidth 115G

- ## Host
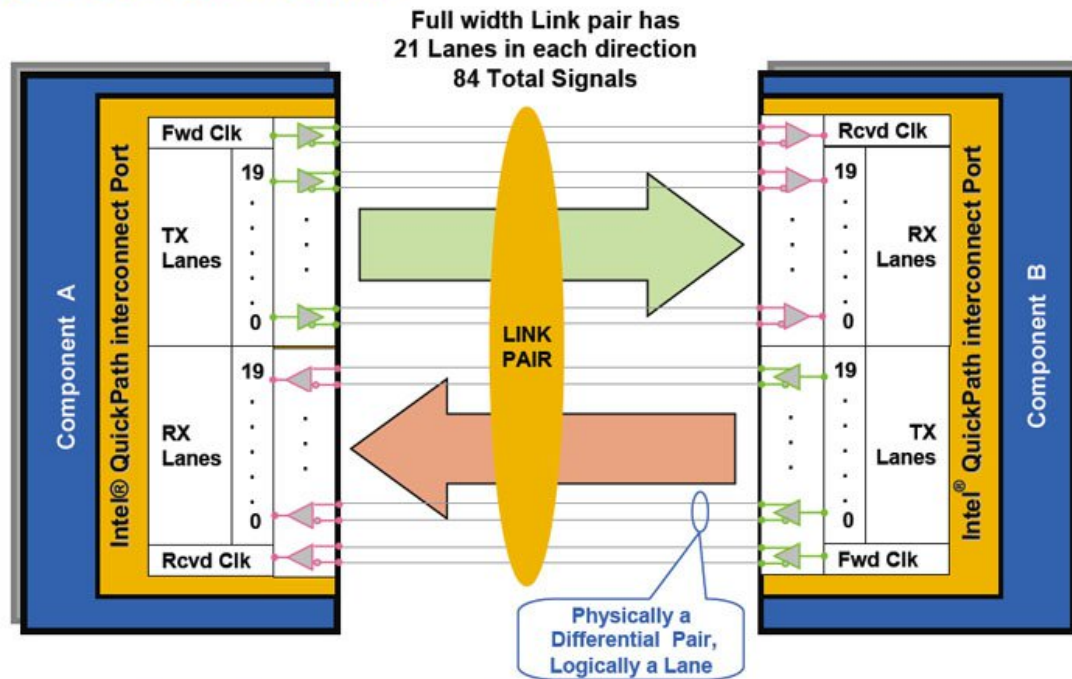 - Assembling flows is multi-core task
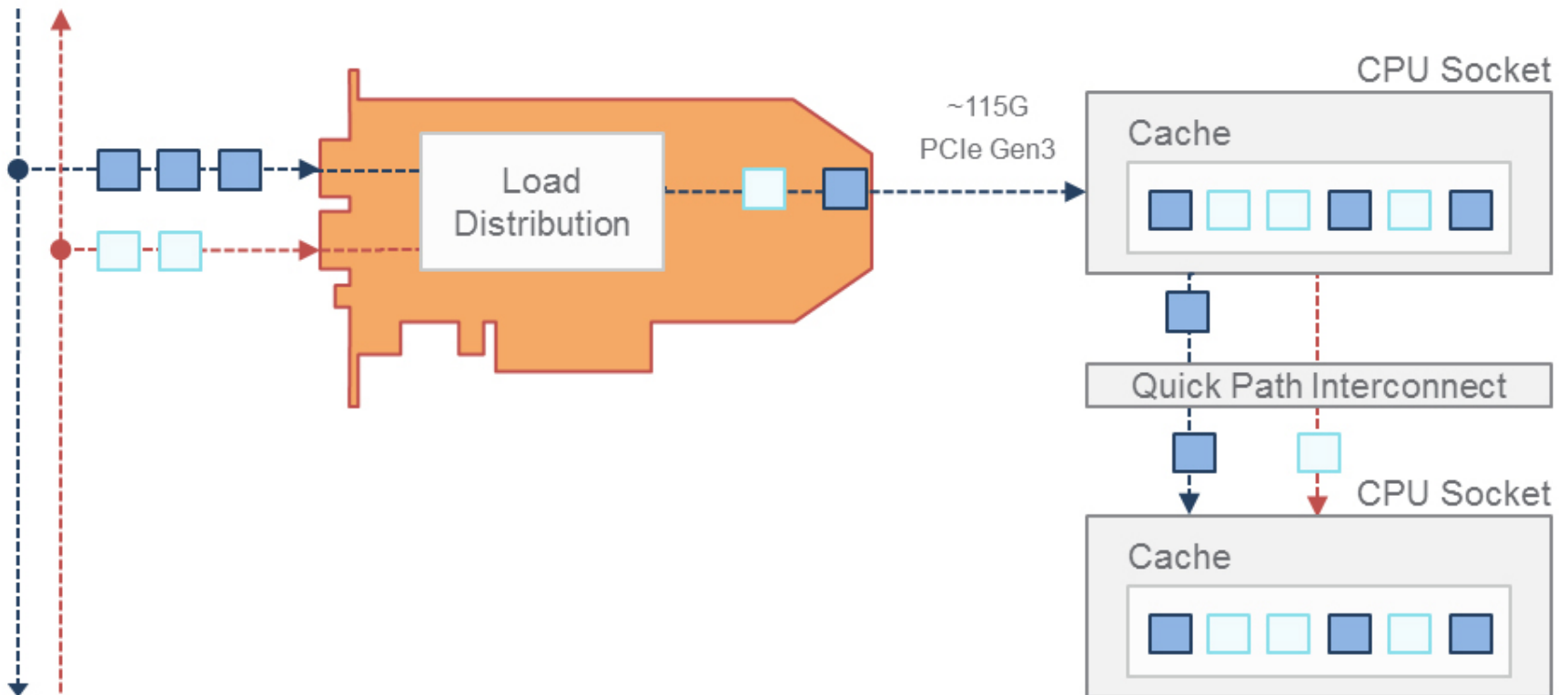 - Minimizing memory latency critical
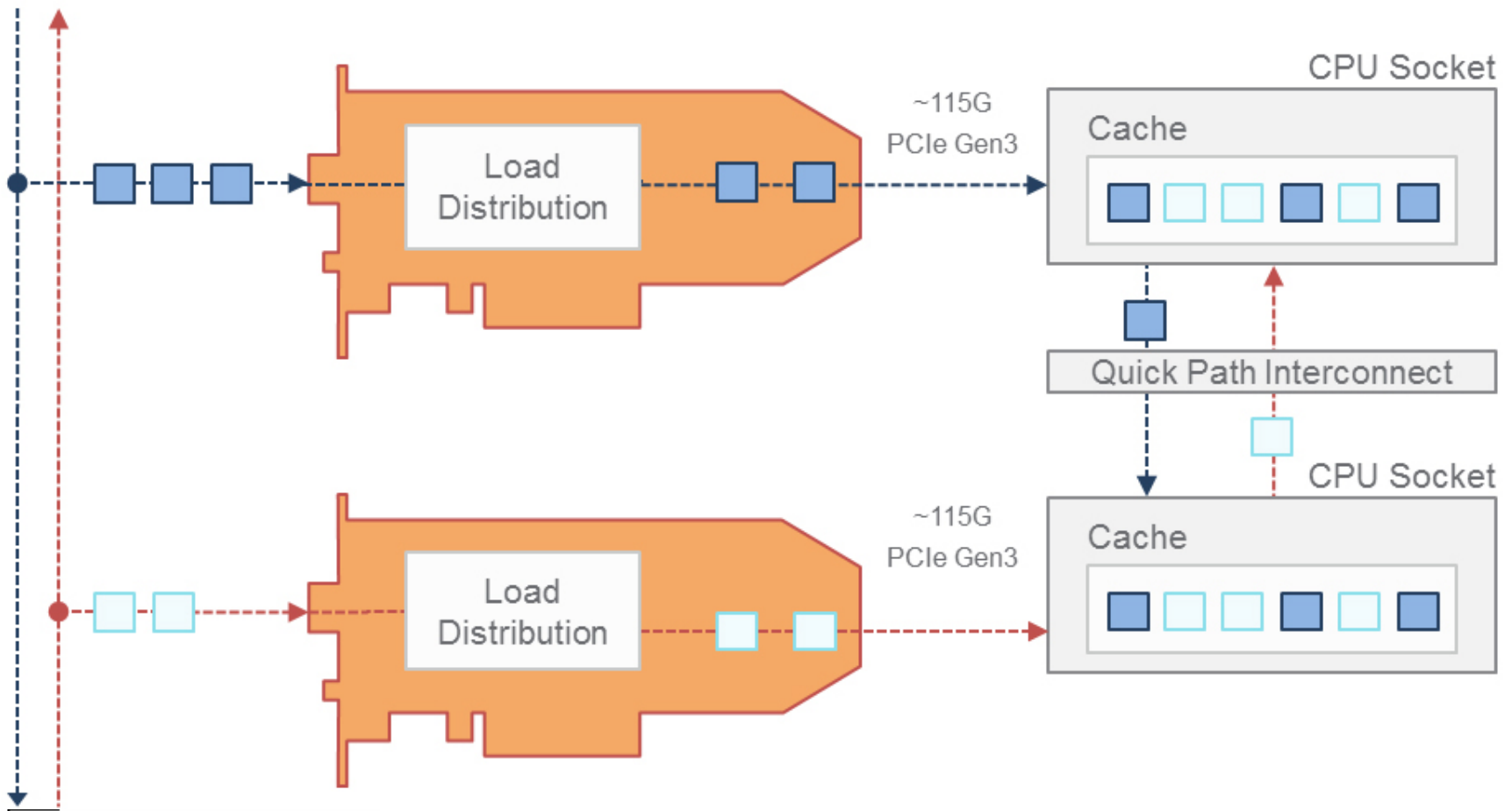
NUMA (Non-Uniform Memory Access)

Intel QPI

Single PCIe Slot

Dual Slots with Intel QPI

Dual with inter-card xfer
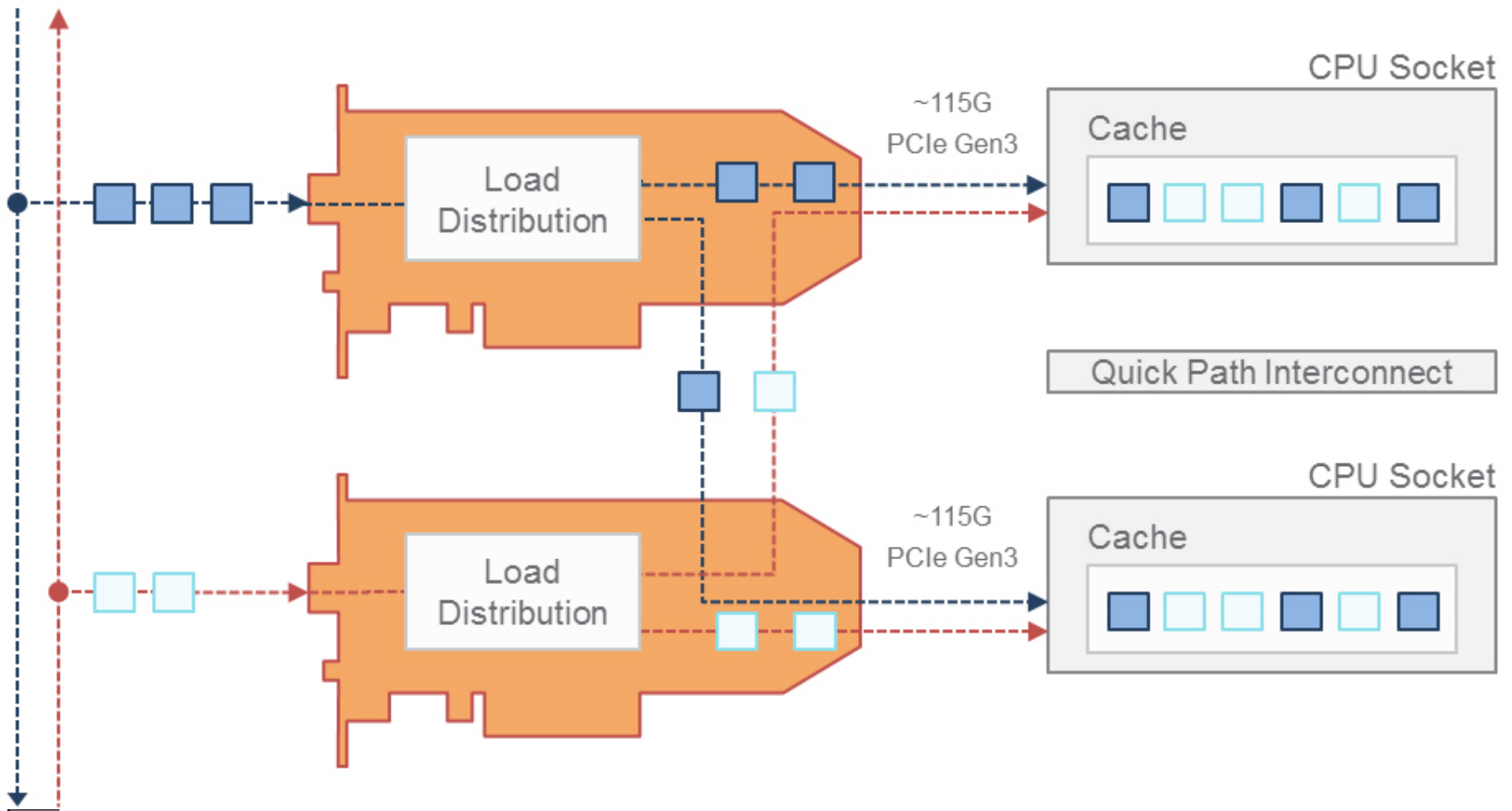
SC16 Demo
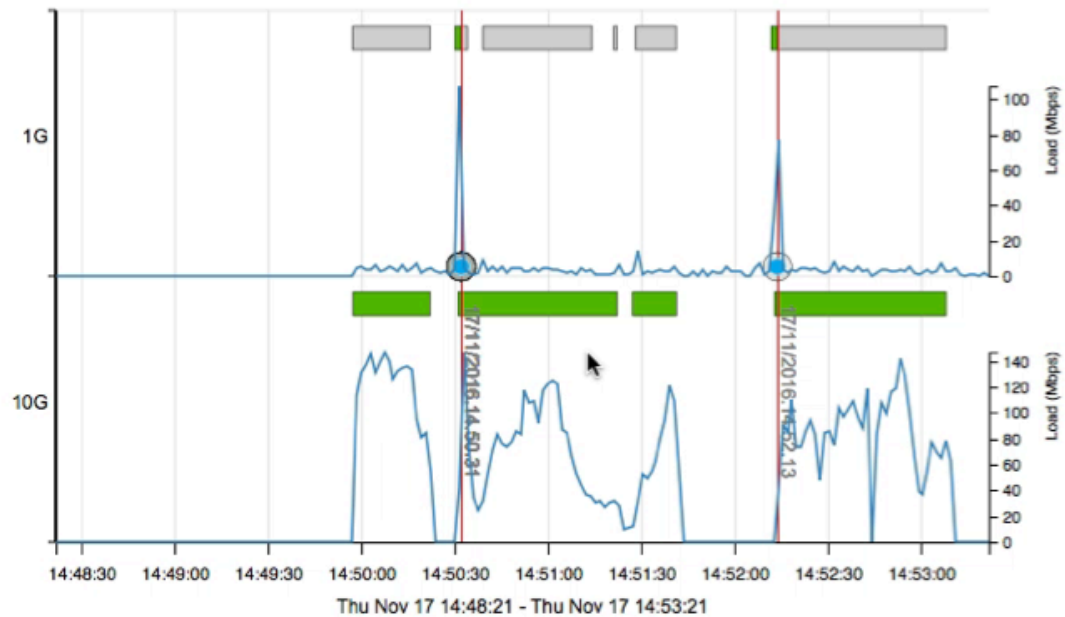
Flow steering

- Goal - optimize bandwidth utilization
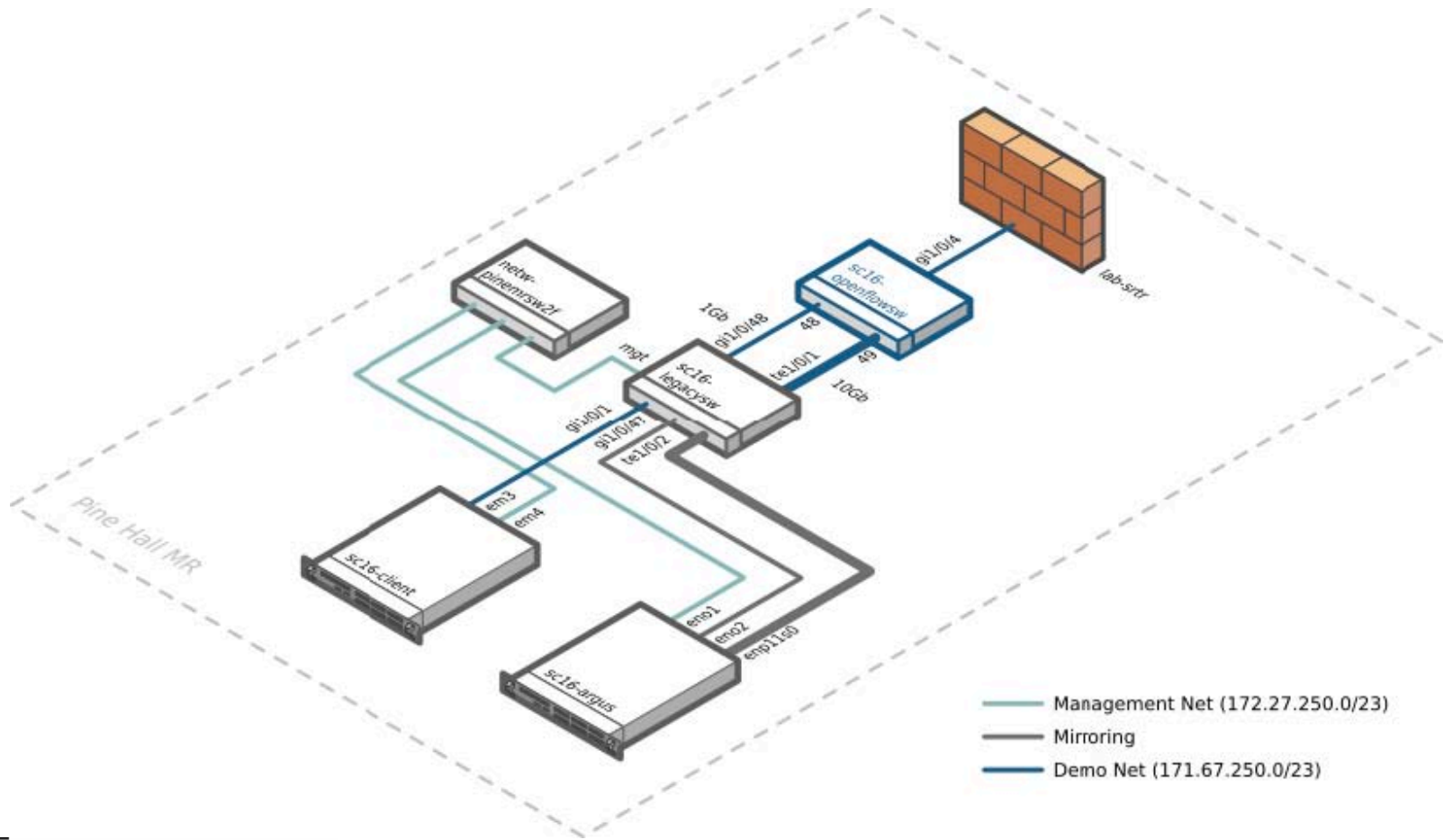 – Don't need or want 100G for all transfers
  – Want to keep pipe full yet uncongested

- Tool – SDN
  – Controller accepts flow-specific commands
  – Can reassign *active* flow paths

- Steering decision
  – Use flow metrics to identify opportunities

Herding elephants

- What are "elephant flows" ?
  - Files in "big data" research environments
  - Video streams
  - Cloud backups
- How can they be identified ?
  - All flow sensors emit periodic records
  - Adjust reporting period
  - Simple byte count thresholds

STANFORD
COMPUTER SCIENCE

Steering Demo Configuration

Management Net (172.27.250.0/23)
Mirroring
Demo Net (171.67.250.0/23)

STANFORD
COMPUTER SCIENCE

Steering

- OpenFlow switch
  - Routes packets based on ACL policies
- OpenFlow controller
  - Pushes ACL policies to switch
- Argus sensor machine
  - Python script tracks elephants in flows
  - Uses REST interface on Controller to add or delete flows from the ACL policy lists

STANFORD
COMPUTER SCIENCE

Steering Demo (place holder)

Beyond Pachyderms

- SDN – software defined networking
  - More than just bandwidth management
- Flow metrics
  - More than packet and byte counts
- Coupling SDN and Flow
  - Realtime audit and validation
  - Fault detection and correction
  - Security monitoring and remediation

STANFORD
COMPUTER SCIENCE

Acknowledgments

- QoSient
  - Argus Pro software with Napatech support
  - Elephant-flow visualization
- Napatech
  - Loan of NT200C01 Network Accelerator card
- Dell
  - Loan of 2x20 core server and
- Stanford Networking
  - OpenFlow controller and 100Gbe link

STANFORD
COMPUTER SCIENCE