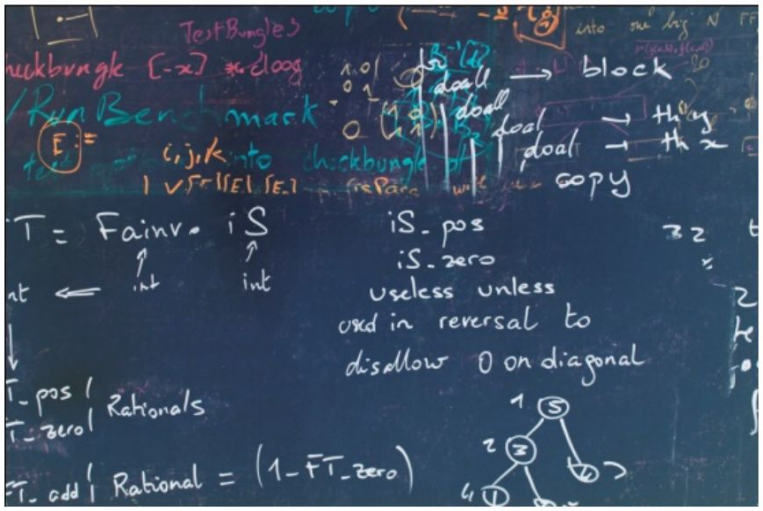
High-Performance Many-Core Networking: Design and Implementation

Jordi Ros-Giralt, PhD

giralt@reservoir.com

Innovating the Network for Data-Intensive Science (INDIS)

Nov 16th 2015



Reservoir Labs

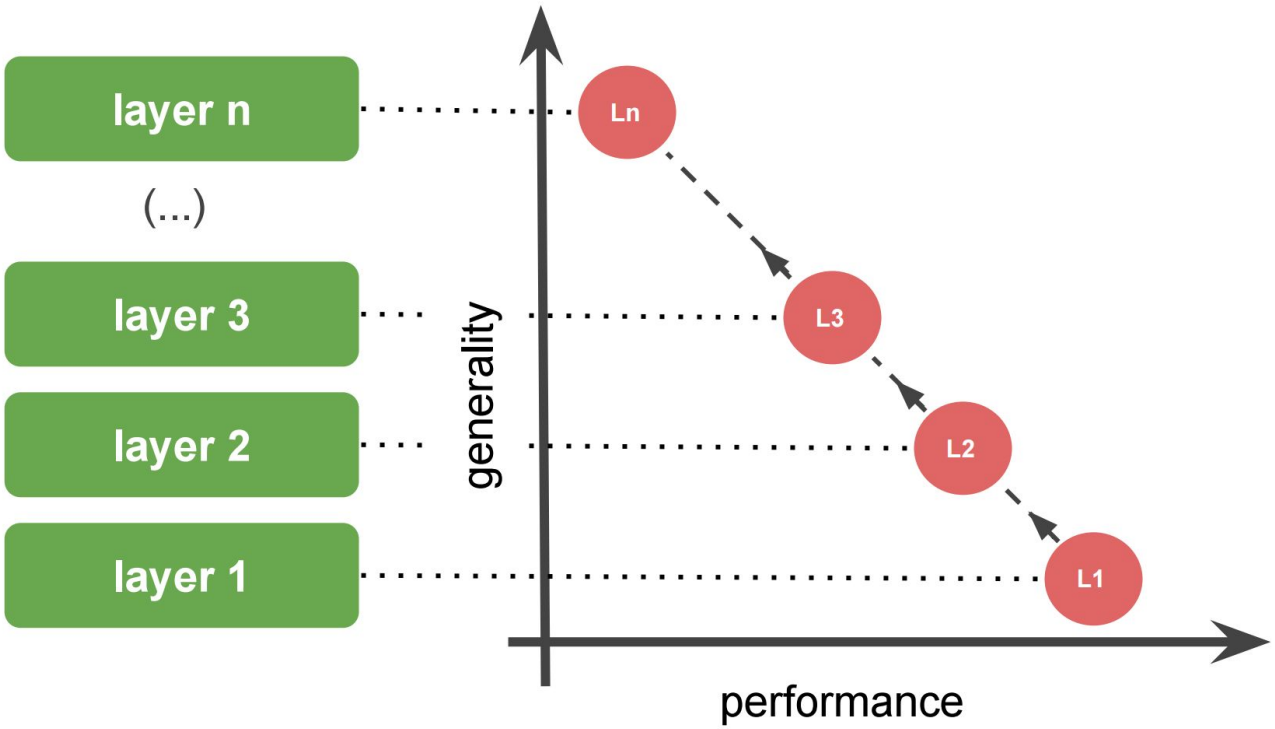
632 Broadway
Suite 803
New York, NY 10012

High-Performance Many-Core Networking: Design and Implementation

TOC:

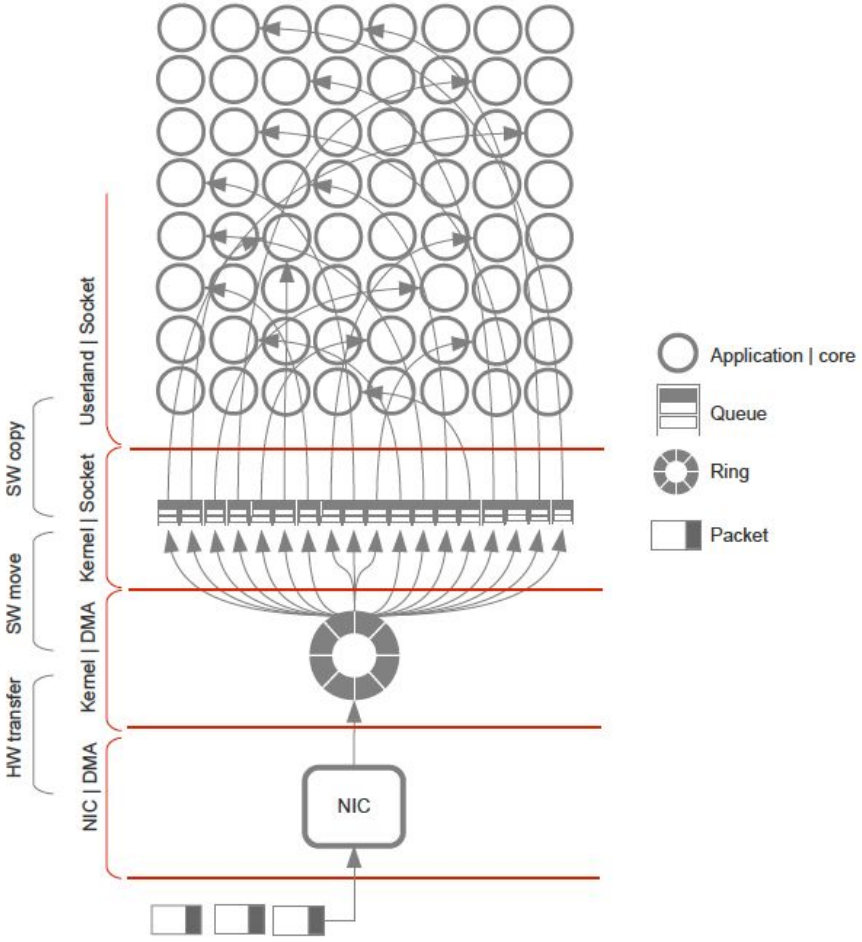
- Trade-Off Generality versus Performance
- Problem space: performance | algorithms
- DNAC: Dynamic Network Acceleration for many-Core
- Performance benchmark
- Providing HPC network visibility at SC15 (SCinet)

Trade-Off Generality Versus Performance



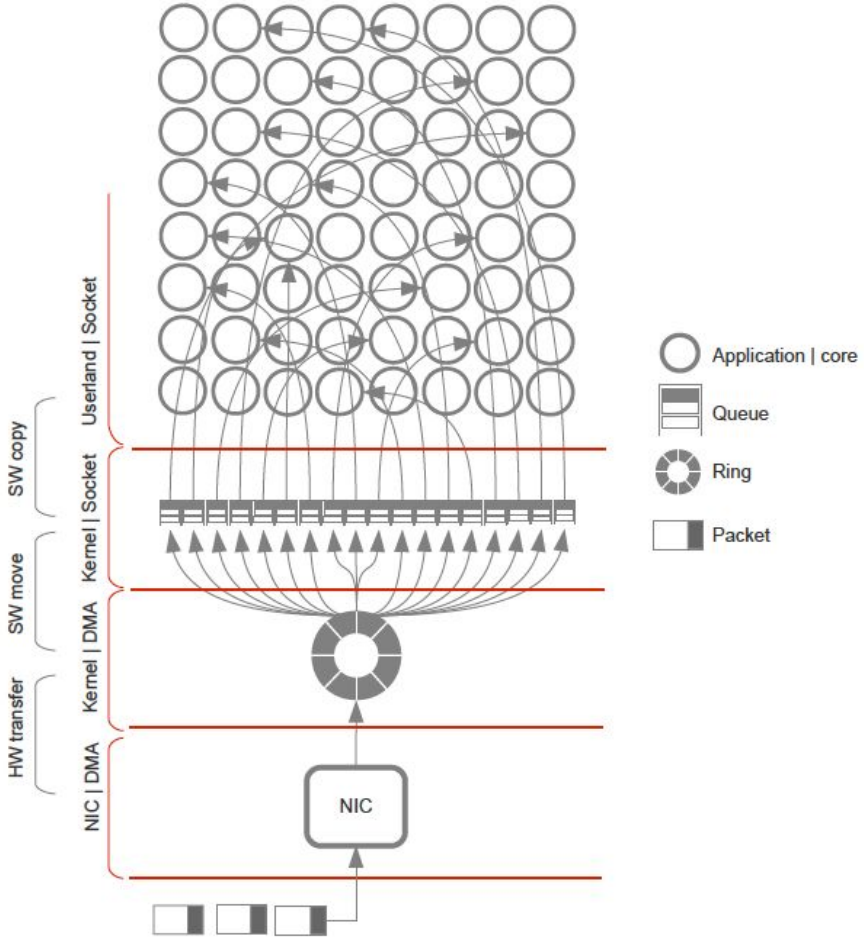
Network Stack Packet Flow

Standard Network Stack:

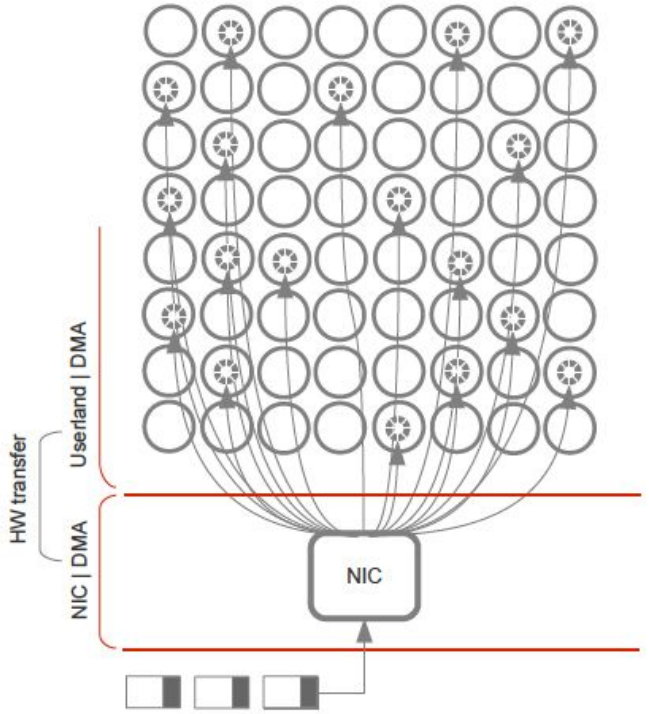


Network Stack Packet Flow

Standard Network Stack:



HPC Optimized Network Stack:

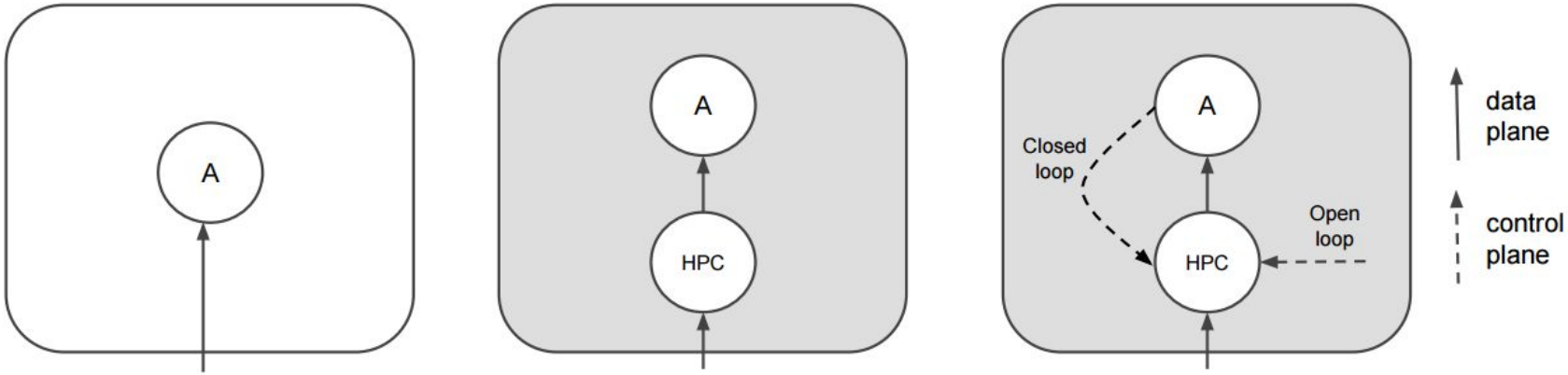


Problem Space

- Two orthogonal problems:
  - High performance computing (performance)
  - Dynamic packet forwarding (algorithms)

Problem Space

- Two orthogonal problems:
  - High performance computing (performance)
  - Dynamic packet forwarding (algorithms)



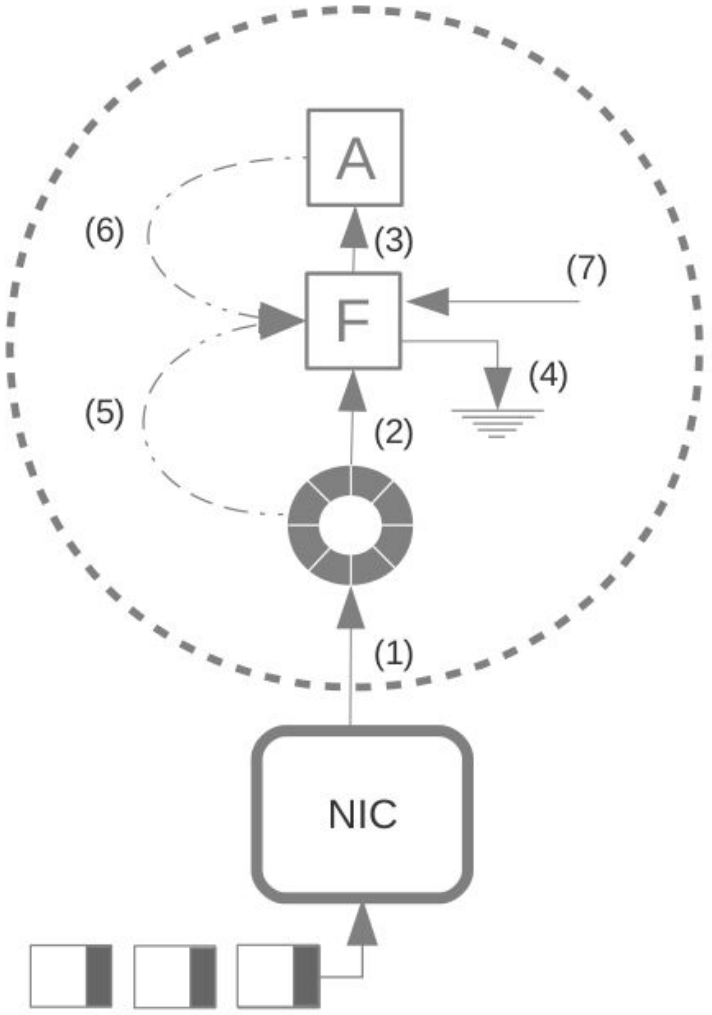
DNAC: Dynamic Network Acceleration for multi-Core

DNAC: Feedback Scenarios

Feedback type	Description	Controller type
Packet congestion	Packets are dropped by the application as a consequence of the application being congested.	closed-loop control
Packet relevance	Packets are dropped by the application as a consequence of the packet not being relevant to the application.	open- or closed-loop control

DNAC: Data Flow

(1) Packet DMA to memory
(2) Forwarder processes packet
(3) Forwarder passes packet to application
(4) Forwarder drops packet
(5) Closed loop feedback from ring
(6) Closed loop feedback from application
(7) Open loop configuration



DNAC: Feedback Classes

Feedback type: packet congestion | controller type: closed-loop

Algorithm: *TED Queuing (Tail early dropping)*. A queuing policy that, upon congestion, drops packets from the tails of each connection, preserving the heads, by dynamically computing a connection cut threshold.

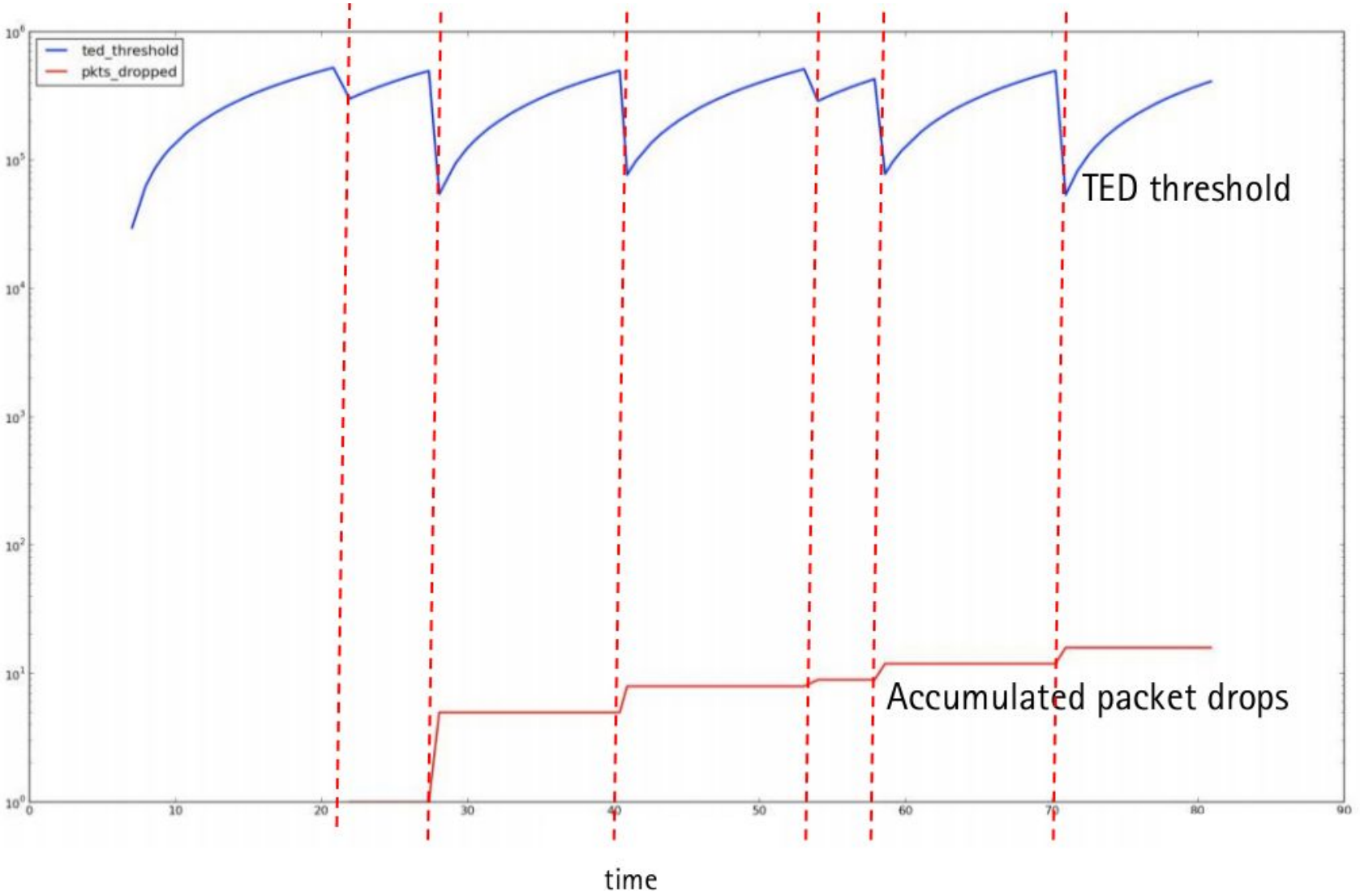
Feedback type: packet relevancy | controller type: closed-loop

Algorithm: *Packet shunting*. Bro workers communicate packet shunting decisions to the forwarder.

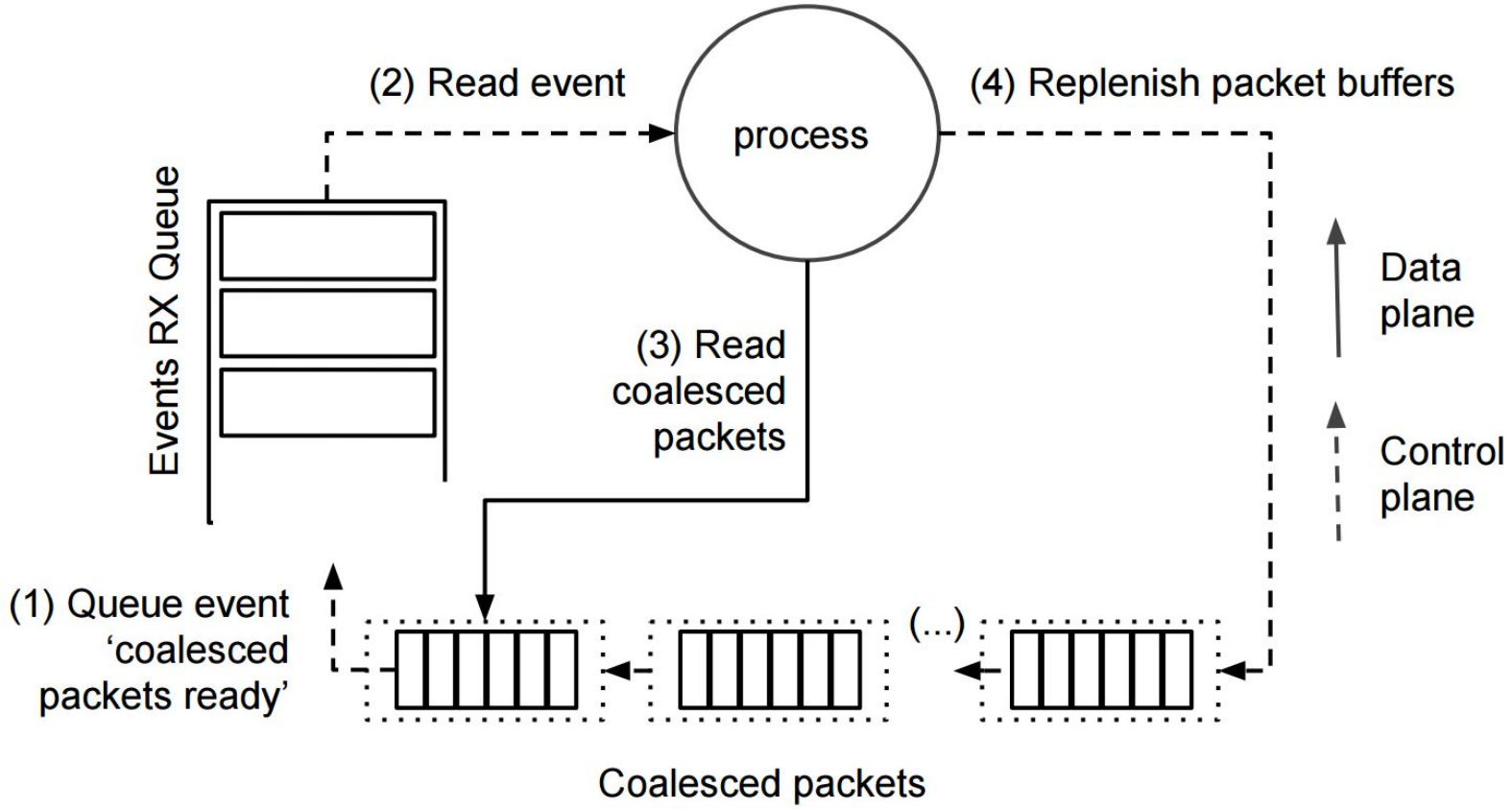
Feedback type: packet relevancy | controller type: open-loop

Algorithm: *Packet prioritization.* Certain packets carry higher degrees of information. Example: dropping a FIN packet has both semantic implications at the protocol level and impacts performance as the upper layer needs to rely on expensive timeouts to free connection context.

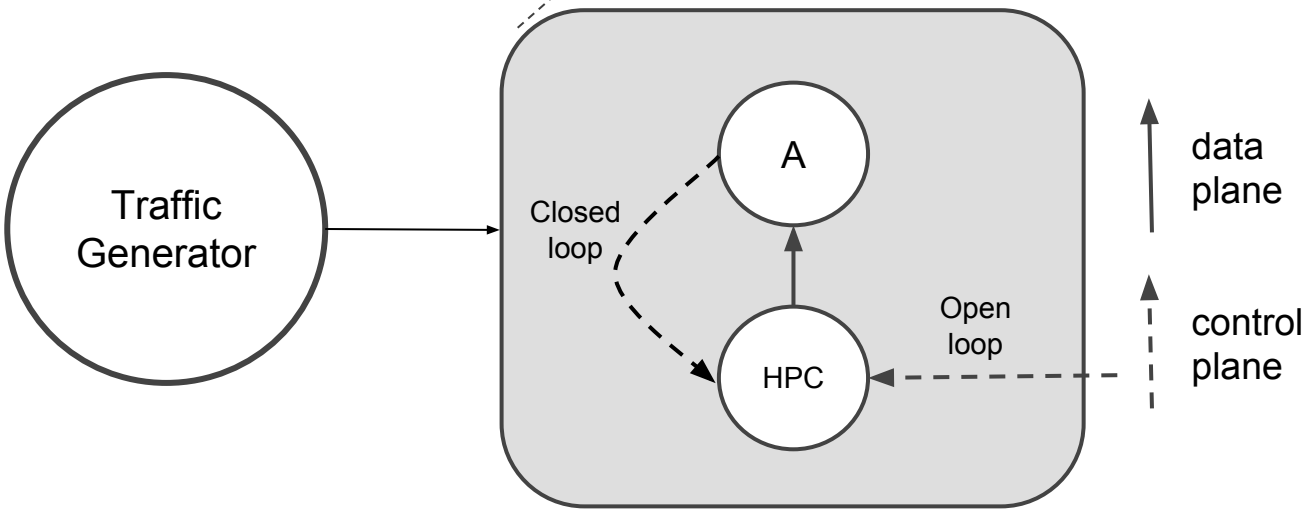
TED Queuing


ted_threshold
pkts_dropped
TED threshold
Accumulated packet drops
time

DNAC: Packet Processing Engine



DNAC: Performance Evaluation



R-Scope Appliance

- High-performance network analyzer running Bro*

R-Scope Manageability

Bro

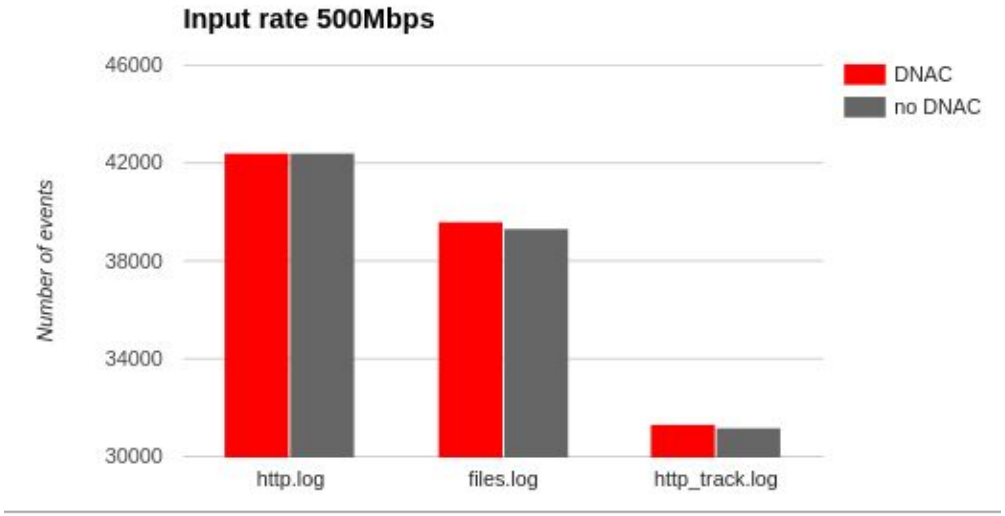
R-Scope Acceleration

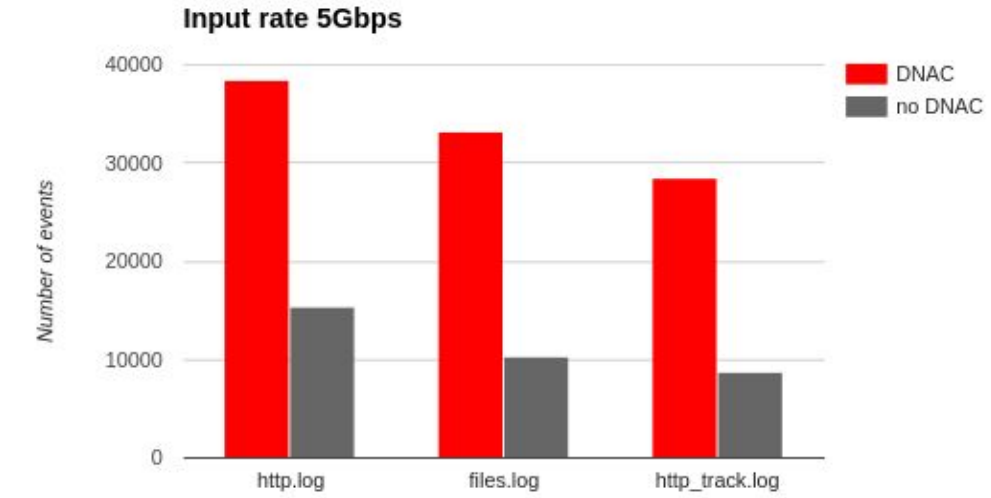
R-Scope Appliance Specifications

Model	R-Scope PACE
Operational Mode	Network Sensor
Maximum Throughput	20 Gbps (and stackable)
Traffic Interfaces	4 x 10 Gbps SFP+
Dimensions (H x W x D)	1U rack-mount (1.7" x 17.1" x 27.5")
Weight	39 lbs.
Power	120V/240 VAC, 50/60 Hz, 750 W Redundant PSUs
Support	On-site, email, and phone

* Bro project: www.bro.org

DNAC: Performance Evaluation



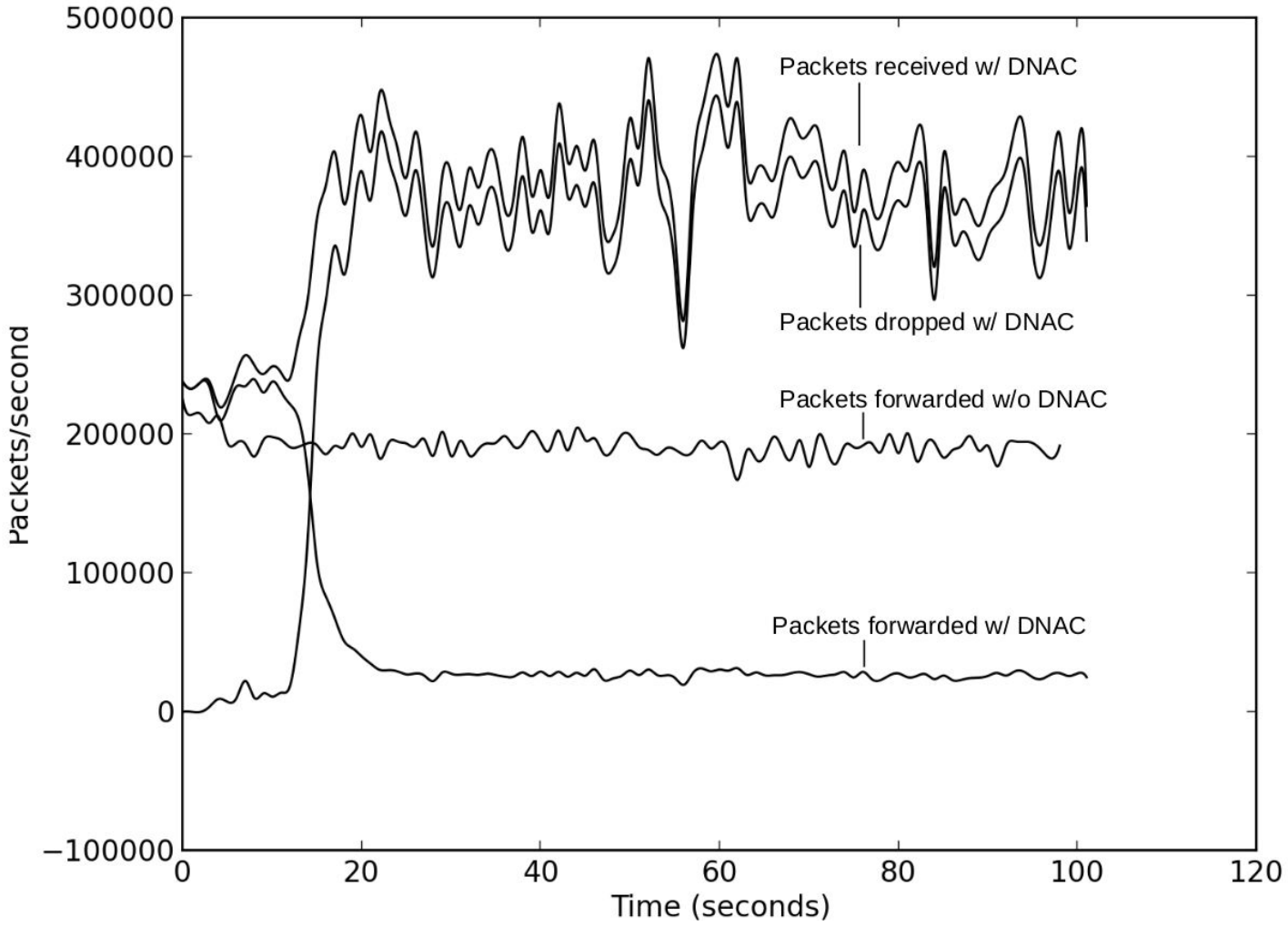


DNAC: Performance Evaluation

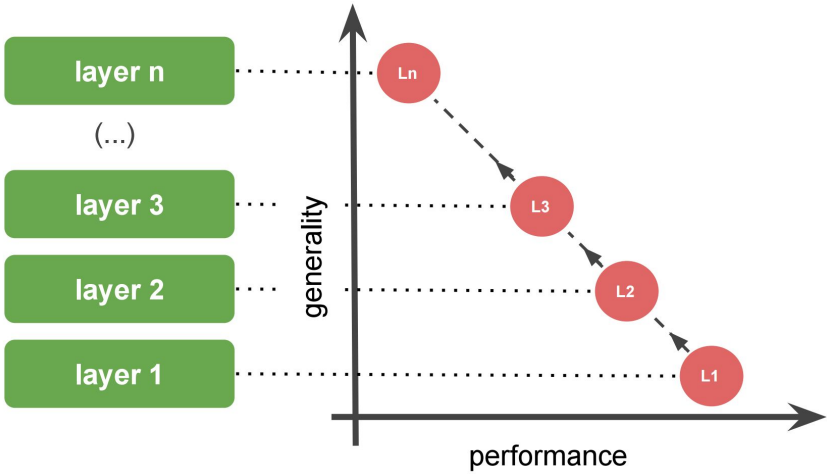
Table 3. Number of events detected

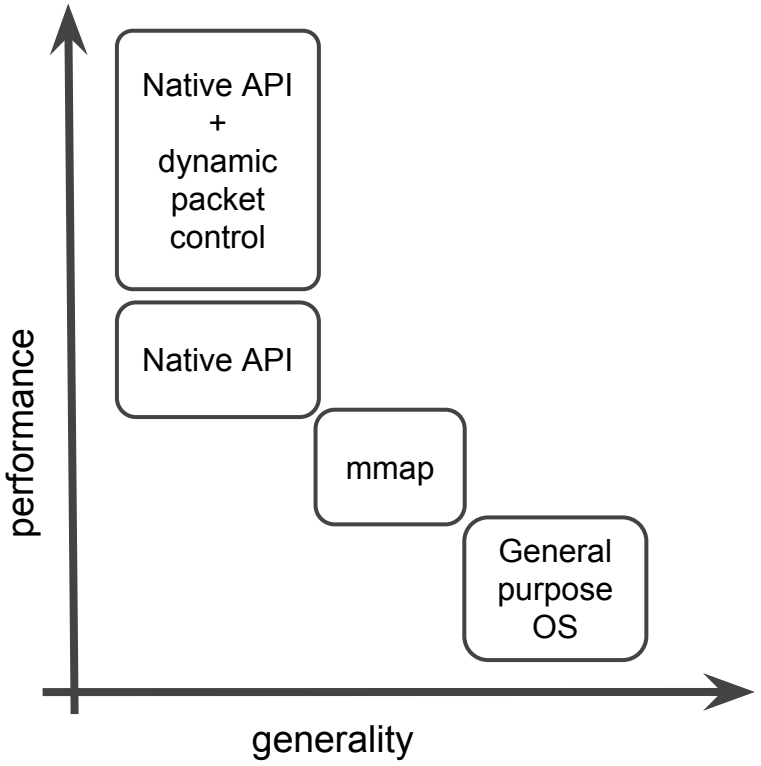
	500Mbps input rate			5Gbps input rate		
	http	files	http_track	http	files	http_track
w/ DNAC	42449	39594	31300	38425	33137	28400
w/o DNAC	42434	39314	31200	15314	10376	8700
gain	1	1	1	2.5	3.2	3.2

DNAC: Performance Evaluation



Trade-Off Generality Versus Performance





Providing Network Visibility at SC2015 (SCinet)

6 x R-Scope systems (6 x 10 Gbps) providing real time Bro* analytics

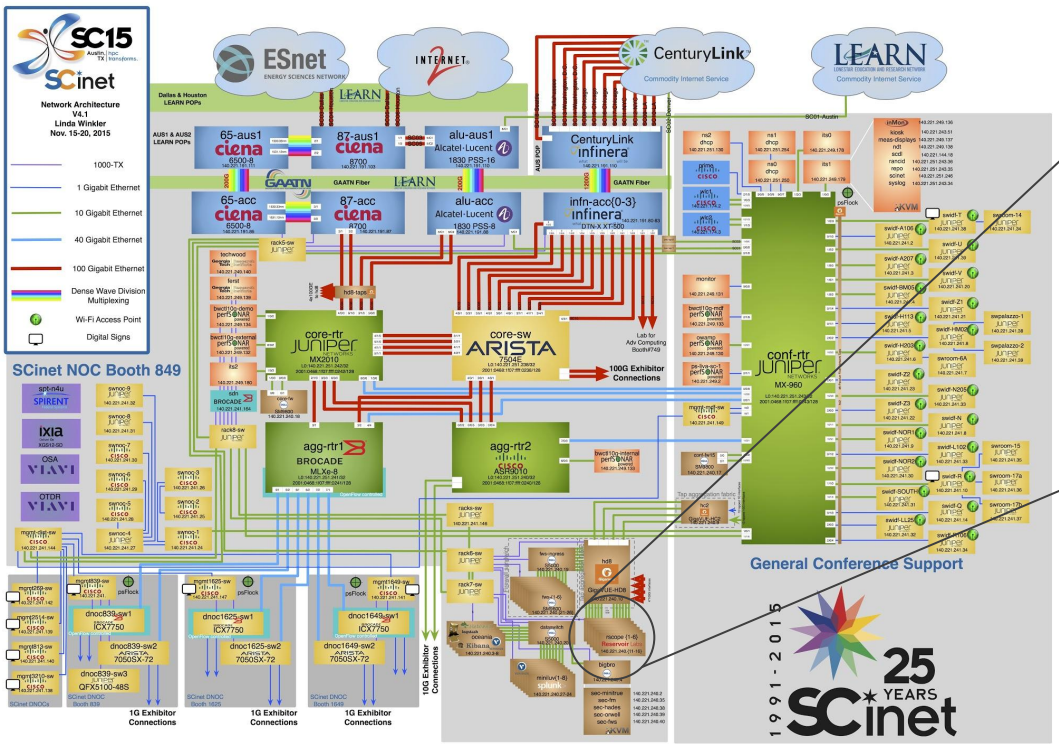
* Bro project: www.bro.org

Providing Network Visibility at SC2015 (SCinet)

Sunday, November 15, 2015

World's Most Powerful Computer Network is Live in Austin

The SCinet network, SC's Supercomputing Internet, is now live! On November 14, the Austin Convention Center became home to the fastest and most innovative computer network in the world, delivering more than 1.6 terabits per second of network bandwidth to the SC conference (SC15).



R-Scope Appliance running DNAC

Thank You

Reservoir Labs

632 Broadway
Suite 803
New York, NY 10012

812 SW Washington St.
Suite 1200
Portland, OR 97205