Scheduling the Datacenter Network

Jonathan Perry
Joint work with Hari Balakrishnan, Devavrat Shah, Amy Ousterhout, Hans Fugal
Datacenters have changed computing

Large web/mobile properties, cloud providers, enterprises, universities

Photo: Connie Zhou, Google Mayes County, OK
Ideal datacenter properties

No current design satisfies all of these simultaneously!

Not enough information?

Alizadeh et al, “DCTCP”, SIGCOMM’10

Burst Control
Predictable
Multiple Objectives
Load Balancing
Centralized arbiter schedules and assigns paths to all packets.

Concerns with centralization:

- Latency
- Scaling
- Fault tolerance

Chuck Norris doesn't wait in queues. He schedules every packet in the datacenter!
Example: packet from A to B

5 µs  \( A \rightarrow \text{Arbiter} \)  
1-20 µs  \( \text{Arbiter} \)  
15 µs  \( \text{Arbiter} \rightarrow A \)  

no queuing  \( A \rightarrow B \)  

“\( A \) has packet for \( B \)”  

timeslot allocation & path selection  

“@t=107: \( A \rightarrow B \) through R1”  

sends data  

\( A \leftarrow \text{B} \)

\( B \)

\( A \)

\( R1 \)

\( R2 \)

\( \text{A} \)

\( \text{Arbiter} \)
Timeslot allocation
Experiments@Facebook: Queues & delays

- Software-only
- 1.6% overhead
MIT invention to speed up data centers should cheer developers

Fastpass uses parallel processing to eliminate the need for complicated network queues, researchers say

Stephen Lawson (IDG News Service) on 19 July, 2014 06:23

Data Center ▶ Data Networking

Kill queues for fast data centres: MIT boffins

Arbitration for in-DC network traffic

By Yawn Himmelsbach / August 27, 2014 / Comment

MIT, Facebook Create Low-Latency Fastpass TCP Replacement
Schedule the Datacenter Network

Benefits
- Predictable Burst Control
- Multiple Objectives
- Load Balancing

A: "Now I can see pictures of other people's food and children so much more quickly... can't wait..."

B: "You forgot about [...] cats. I will say, faster pics of cats is probably worth some merit."

Lower User Service Time

Better Developer Productivity

Lower Infrastructure Cost

Burst Control
- Predictable
- Multiple Objectives
- Load Balancing
Background

• Networks & Mobile Systems group at CSAIL (nms.csail.mit.edu)
• Datacenter/cloud systems, mobile/sensor computing, Internet architecture
• Jonathan: PhD student near completion, previous industry experience
• Hari: multiple startups from projects – Cambridge Mobile Telematics, Meraki (Cisco), StreamBase (TIBCO), Sandburst (Broadcom)
Fastpass: Treat network as a BIG switch!

- Core
- ToR
- Nodes

**Diagram:**
- Endpoint
  - Host networking stack
  - FCP client
  - NIC
- Arbiter
  - Timeslot allocation
  - FCP server
  - Path Selection

- Fastpass Arbiter

**Text:**
- NIC
- Path
- Core
- ToR
- Nodes
- Fastpass Arbiter
- Host networking stack
- FCP client
- Timeslot allocation
- Path Selection
- FCP server