HPC and the Linux Kernel

Roland Dreier, Cisco

THE
LINUX
FOUNDATION

Linux is the OS for 439 of the Top 500 supercomputers (9 of top 10, including #1, Roadrunner)

(source: October 2008 Top 500 list, top500.org)

THE
LINUX
FOUNDATION

#1: Roadrunner

6,480 Opteron processors, 12,960 Cell processors

> 100 TB RAM

296 racks

> 2 MW power

1 petaflops

...running Linux.

InfiniBand is the interconnect for 141 of the Top 500

(4 of top 10, incl. #1)

(source: October 2008 Top 500 list, top500.org)

THE **LINUX** FOUNDATION

What the heck is InfiniBand?

InfiniBand is an RDMA interconnect.

RDMA:

Remote Direct Memory Access
Asynchronous work and completion queues
Kernel bypass

Linux has great RDMA support

7 low-level hardware drivers

Networking: IPoIB, RDS

Storage: SRP, iSER

Filesystems: NFS/RDMA, 9p/RDMA

Linux supports RDMA on Ethernet

3 of 7 RDMA HW drivers are for iWARP adapters

(what the heck is iWARP?)

InfiniBand-over-Ethernet maybe in the future

Future of Ethernet

10 Gb/sec now, 40 & 100 Gb/sec someday

plus

Data Center Bridging (DCB)
aka
Data Center Ethernet (DCE)
aka
Converged Enhanced Ethernet (CEE)

DCB standards from IEEE

Priority-Based Flow Control (802.1Qbb)

Enhanced Transmission Selection (802.1Qaz)

Congestion Notification(802.1Qau)

"No drop" Ethernet becomes possible

DCB support in Linux

Initial support is there (DCBXP, multiqueue, FCoE)
Full implications still to be explored