LS-DYNA Performance
With MPI Collectives Acceleration

April 2011

Note

- **The following research was performed under the HPC Advisory Council HPC|works working group activities**
  - Participating vendors: HP, Intel, Mellanox
  - Compute resource - HPC Advisory Council Cluster Center

- **We would like to thank Mellanox for providing early access to its MPI Collectives Acceleration solution (FCA version 2.1)**

- **For more info please refer to**
  - http://www.hp.com/go/hpc
  - www.intel.com
  - www.mellanox.com
  - www.lstc.com

Objectives

- **The presented research was done to provide best practices**

  – LS-DYNA interconnect performance benchmarking

  – Application profiling

  – Understanding LS-DYNA communication patterns

- **Preview on available MPI collectives accelerations**

- **First performance results with LS-DYNA**

  – Utilizing MPI collectives accelerations

- **LS-DYNA**
  - A general purpose structural and fluid analysis simulation software package capable of simulating complex real world problems
  - Developed by the Livermore Software Technology Corporation (LSTC)
- **LS-DYNA used by**
  - Automobile
  - Aerospace
  - Construction
  - Military
 - Manufacturing
  - Bioengineering
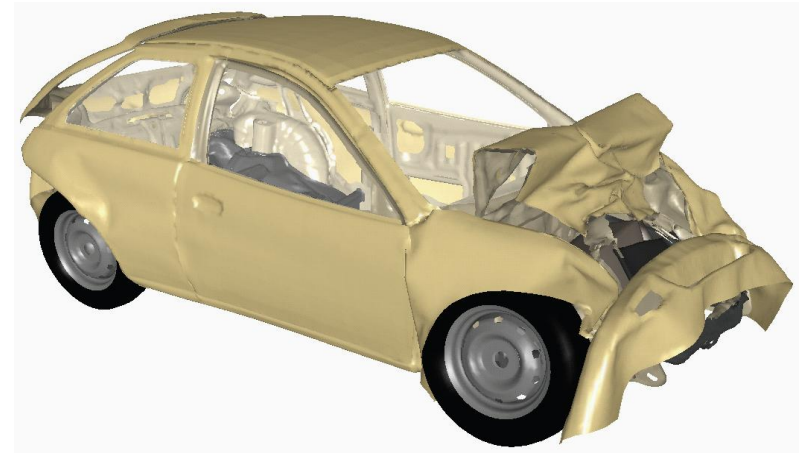
- **LS-DYNA SMP (Shared Memory Processing)**
  - Optimize the power of multiple CPUs within single machine
- **LS-DYNA MPP (Massively Parallel Processing)**
  - The MPP version of LS-DYNA allows to run LS-DYNA solver over High-performance computing cluster
  - Uses message passing (MPI) to obtain parallelism
- **Many companies are switching from SMP to MPP**
  - For cost-effective scaling and performance

Test Cluster Configuration
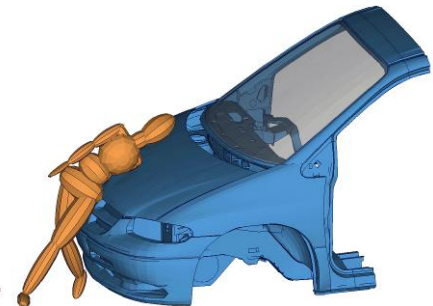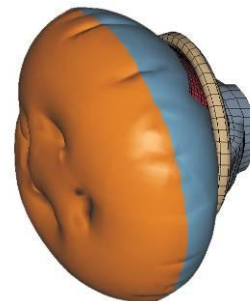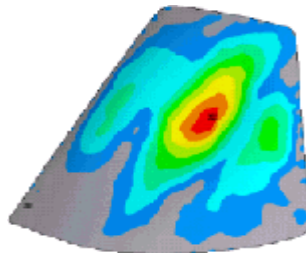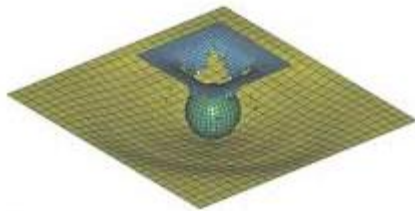
- **HP ProLiant SL2x170z G6 16-node cluster**

  – Six-Core Intel X5670 @ 2.93 GHz CPUs

  – Memory: 24GB per node

  – OS: CentOS5U4, OFED 1.5.2 InfiniBand SW stack

- **Mellanox ConnectX-2 InfiniBand QDR adapters and switches**

- **MPI: Open MPI 1.4.3**

- **Mellanox Fabric Collective Accelerator™ (FCA™) version 2.1**

- **Application: LS-DYNA mpp971_s_R5.0**

- **Benchmark workload**

  – 2001 Ford Taurus, Detailed model (1,057,113 elements)

- **Solution-optimized for extreme scale out**

ProLiant SL160z G6

ProLiant SL165z G7

Large memory
-memory-cache apps
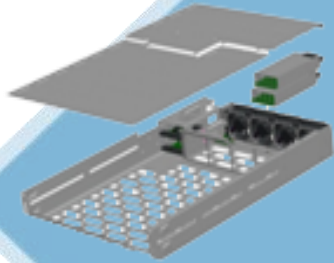
ProLiant z6000 chassis

Shared infrastructure
– fans, chassis, power
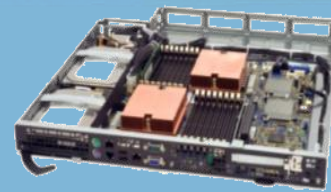
ProLiant SL170z G6

Large storage
-Web search and database apps
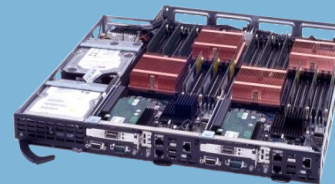
ProLiant SL2x170z G6

Highly dense
- HPC compute and
web front-end apps

Save on cost and energy -- per node, rack and data center
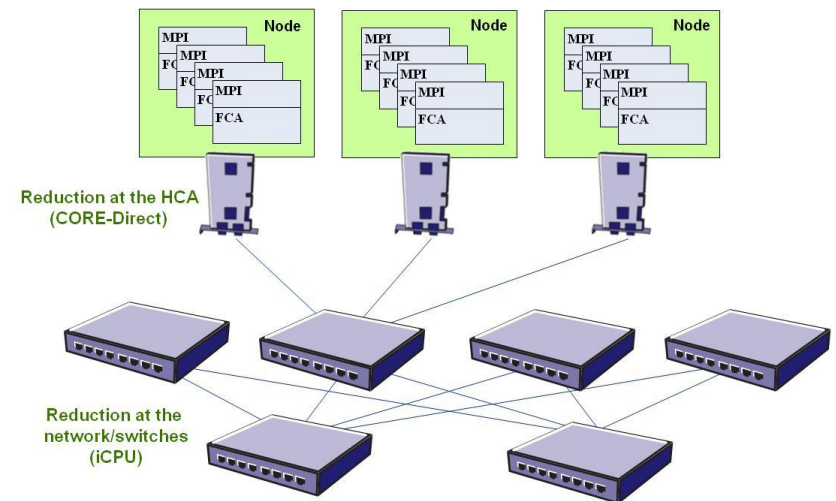
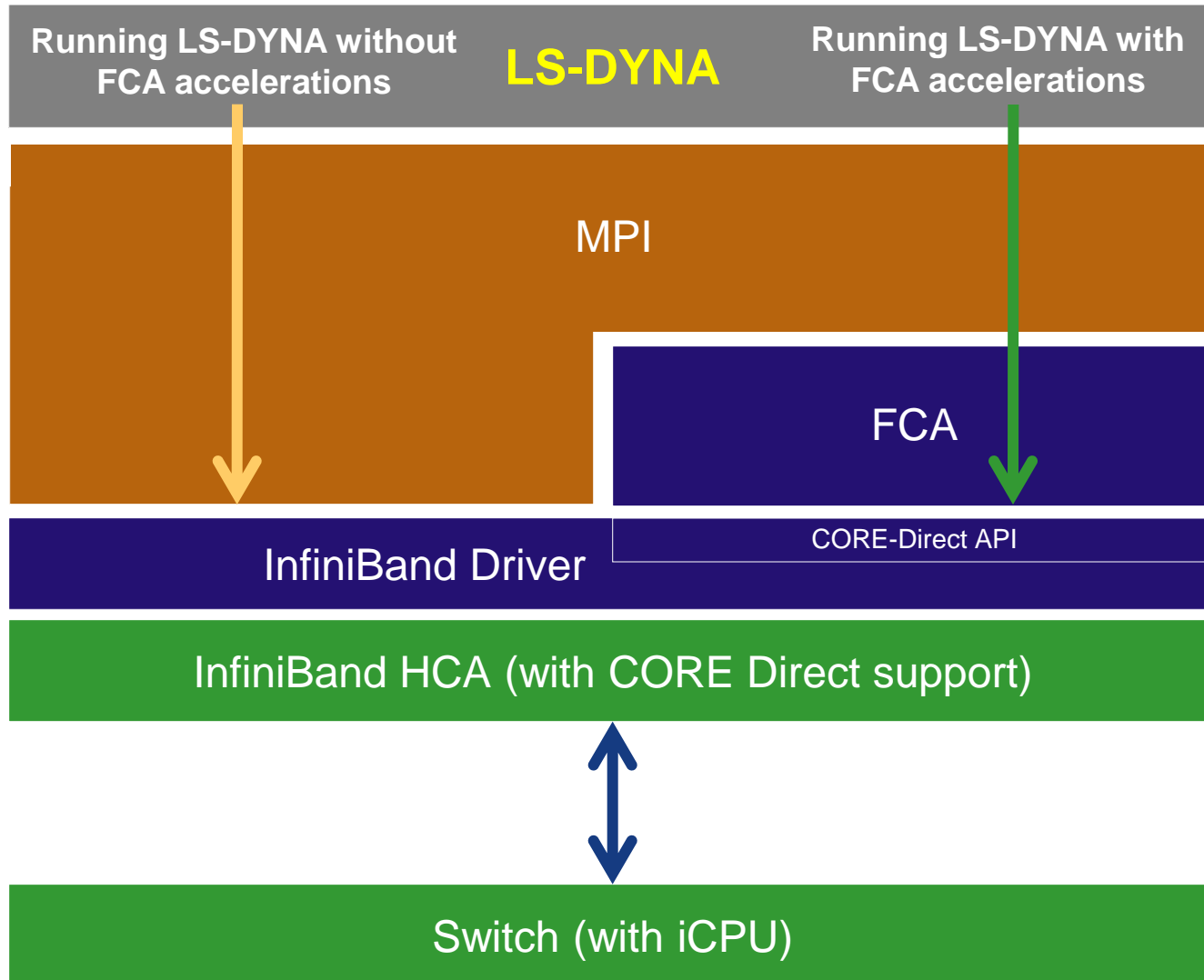Mix and match configurations

Deploy with confidence

#1
Power
Efficiency*

* SPECpower_ssj2008
www.spec.org
17 June 2010, 13:28

About Mellanox FCA

- **Mellanox Fabric Collectives Accelerator (FCA)**
 - Utilized hardware accelerations on the adapter (CORE-Direct)
  - Utilized managed switches capabilities (iCPU)
  - Accelerating MPI collectives operations
  - The world first complete solution for MPI collectives offloads

- **FCA 2.1 supports accelerations/offloading for**
  - MPI Barrier
  - MPI Broadcast
  - MPI AllReduce and Reduce
  - MPI AllGather and AllGatherV

Software Layers Overview

Running LS-DYNA without FCA accelerations   **LS-DYNA**   **Running LS-DYNA with FCA accelerations**

MPI

FCA

InfiniBand Driver   CORE-Direct API

InfiniBand HCA (with CORE Direct support)

Switch (with iCPU)

LS-DYNA Benchmark Results - FCA

- **Input Dataset**
  - Ford Taurus (1M elements)
- **FCA enables nearly 20% performance gain at 16 nodes / 192 cores**
  - Bigger advantage expected at higher node count / core count

LS-DYNA (Ford Taurus Benchmark)

■ Without FCA ■ FCA
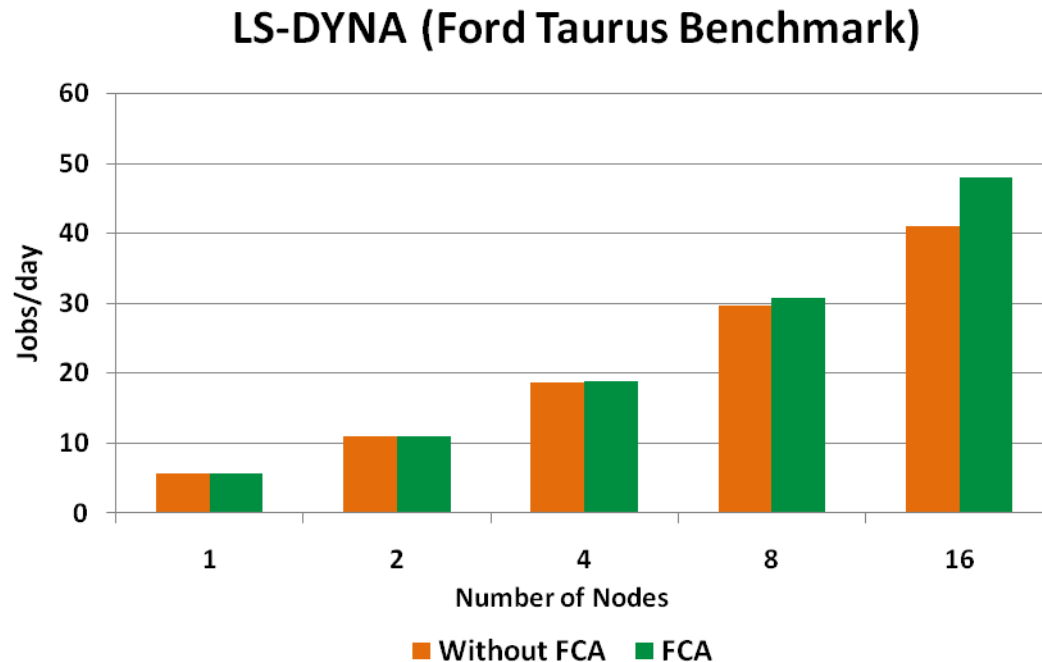
Higher is better

12-cores per node

LS-DYNA MPI Profiling – MPI Functions

- **MPI communication time decreases by more than 10%**
  - With 16 nodes / 192 cores
  - More communication time should be decreased as cluster size increases

LS-DYNA

(Ford Taurus)

InfiniBand QDR

NETWORK OF EXPERTISE

LS-DYNA Benchmark Summary
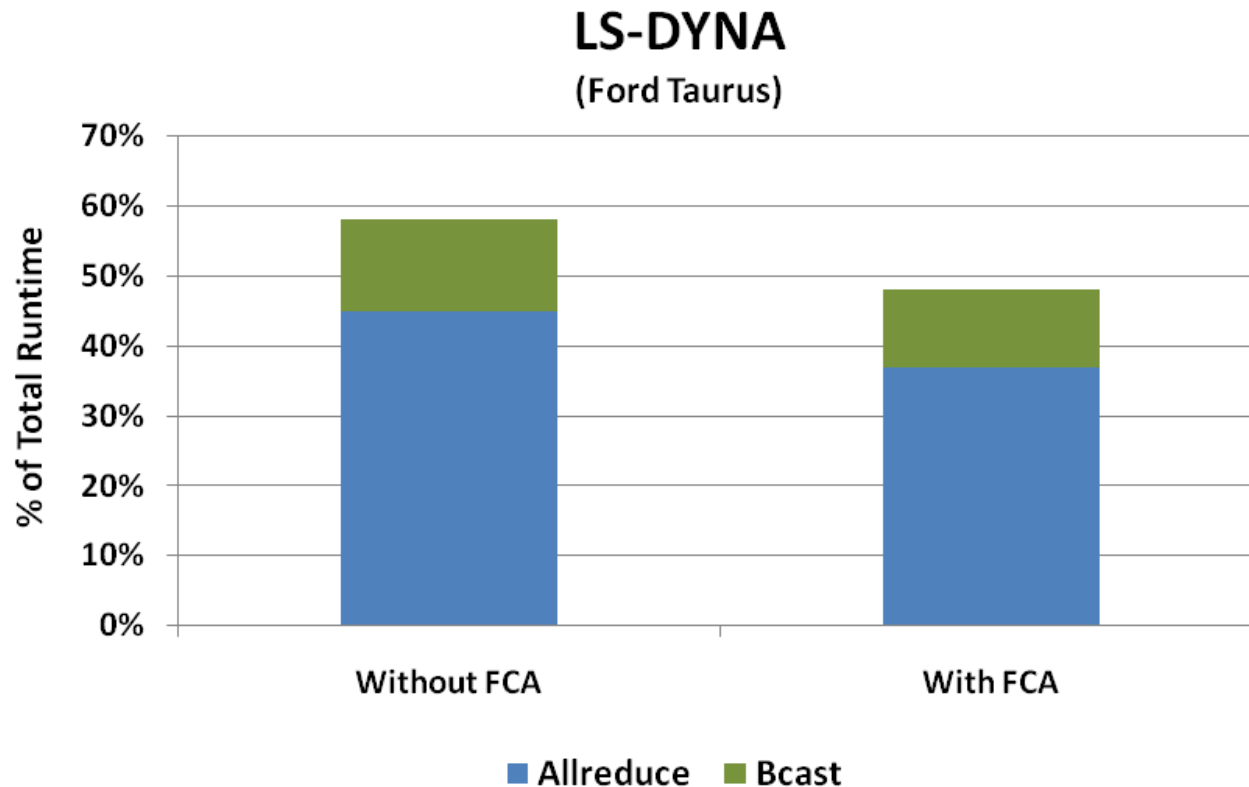
- **MPI Collectives accelerations can dramatically accelerate HPC applications performance**
  - The data presented here reviewed acceleration at small scale
  - Large scale systems will get bigger benefit from such acceleration

- **LS-DYNA MPI profiling**
  - MPI_Bcast and Allreduce are the main MPI routine impacts LS-DYNA performance

- **FCA package has proven to accelerate application even at small scale**
  - Nearly 20% at 16 nodes / 192 cores for LS-DYNA
  - Higher performance boost expected at larger scale

Thank You
HPC Advisory Council

NETWORK OF EXPERTISE