RoCE
RDMA over Converged Ethernet

July 16, 2010
Bill Lee
bill@mellanox.com
What is RoCE?

• Light-weight RDMA transport over Ethernet
  – Data movement between application memory without CPU involvement
  – RDMA read/write, send-receive and kernel bypass

• Proven, widely deployed technology
  – Server efficiency and scaling to 1000s of nodes
  – Scales to 40GigE support and beyond
  – Supports existing low latency (RDMA) applications

• IBTA standard
RoCE Standardization Timeline

2009
June  IBTA RoCE working group formed
August RoCE in a HOTI paper and panel
November OpenFabrics Board of Directors votes to adopt RoCE and include support in the OFA driver stack

2010
February RoCE spec review completed
February RoCE presented at EA TEF
March RoCE spec approved by IBTA Steering Committee
March OFED 1.5.1 with RoCE support is released by OFA
IO Stack

- Applications / OS Services
  - Sockets
  - SDP
  - OFA RDMA Verbs
  - iWARP
  - iSCSI
  - FCP/FC
  - TCP
  - IP
  - FCoE
  - RoCE

- DCB Ethernet
Target Applications & Benefits

Financial Applications

Sample Apps: Wombat/NYSE, IBM WebSphere MQ, Red Hat MRG, 29West etc.

Benefits: High performance, scalable market data processing, faster analytics and algorithmic trading

Data Warehousing

Sample Apps: Oracle RAC, IBM DB2 PureScale, Microsoft SQL etc.

Benefits: Significantly higher job operations per second, linear scaling with cluster size, maintain table scan time in the face of exponential growth in DB table sizes

Clustered Cloud Computing

Sample Apps: VMware, Citrix, Microsoft, Amazon EC2, Google App Engine etc.

Benefits: Improved SLAs through deterministic performance, efficient clustering allowing for elastic/scale out computing and storage with higher ROI

Delivers Compelling Benefits to High Growth Markets
Determinism in Performance and Profitability

MRG 1.3 Red Hat Enterprise 6.0 over RoCE*

- Consistent latency across message rate
- 1.2M messages per second
- In-box support in RHEL 6.0

*Presented by Red Hat at Red Hat Summit’10
Industry Support

- Cisco
- IBM
- Red Hat
- NYSE Euronext
- Arista
- OpenFabrics Alliance
- InfiniBand® (TRADE ASSOCIATION MEMBER)
- Ethernet Alliance
Thank You