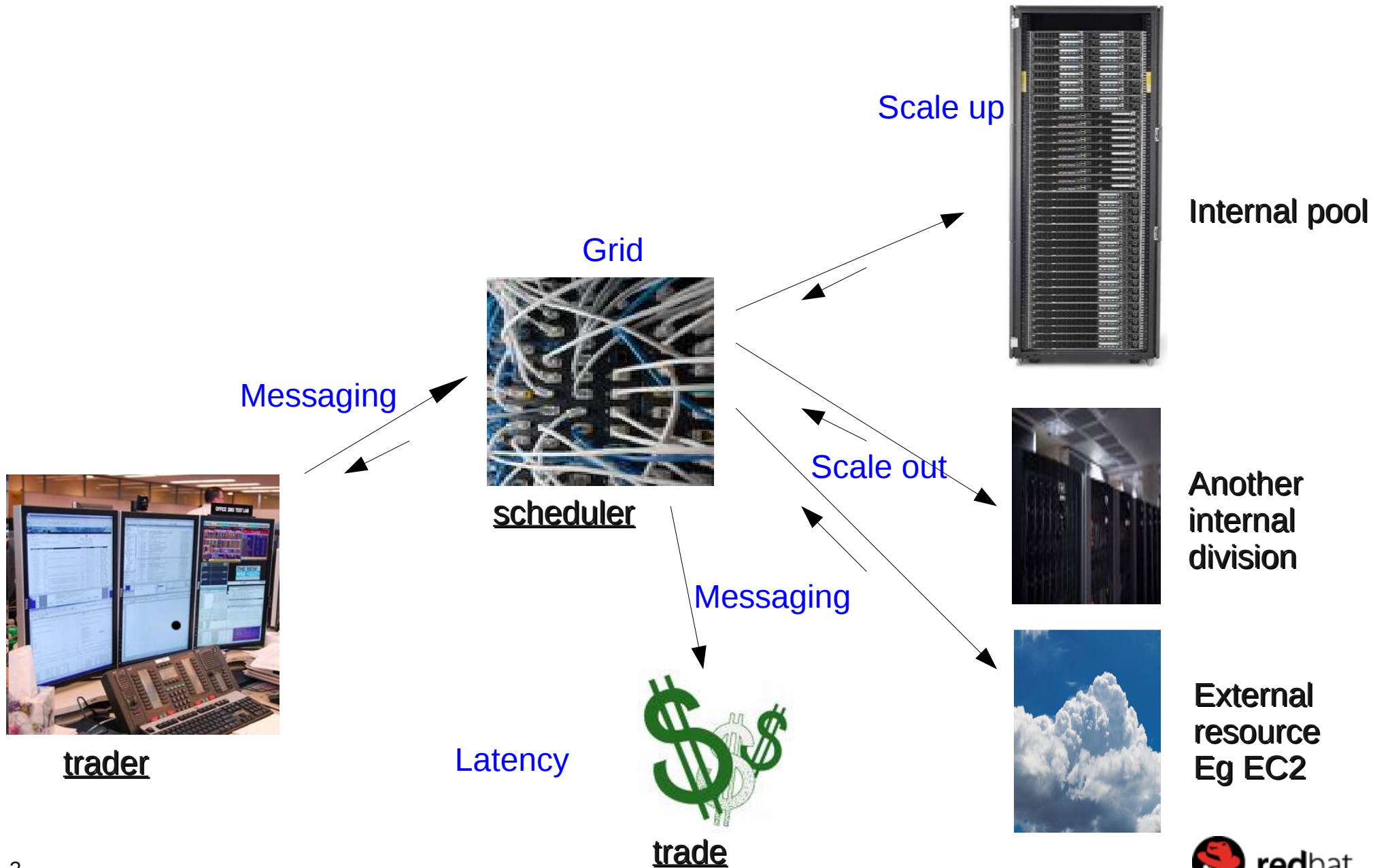




Cloud, MRG, and more...

Carl Trieloff, Technical Director
cctrieloff@redhat.com

From simulation to trade



What is Cloud

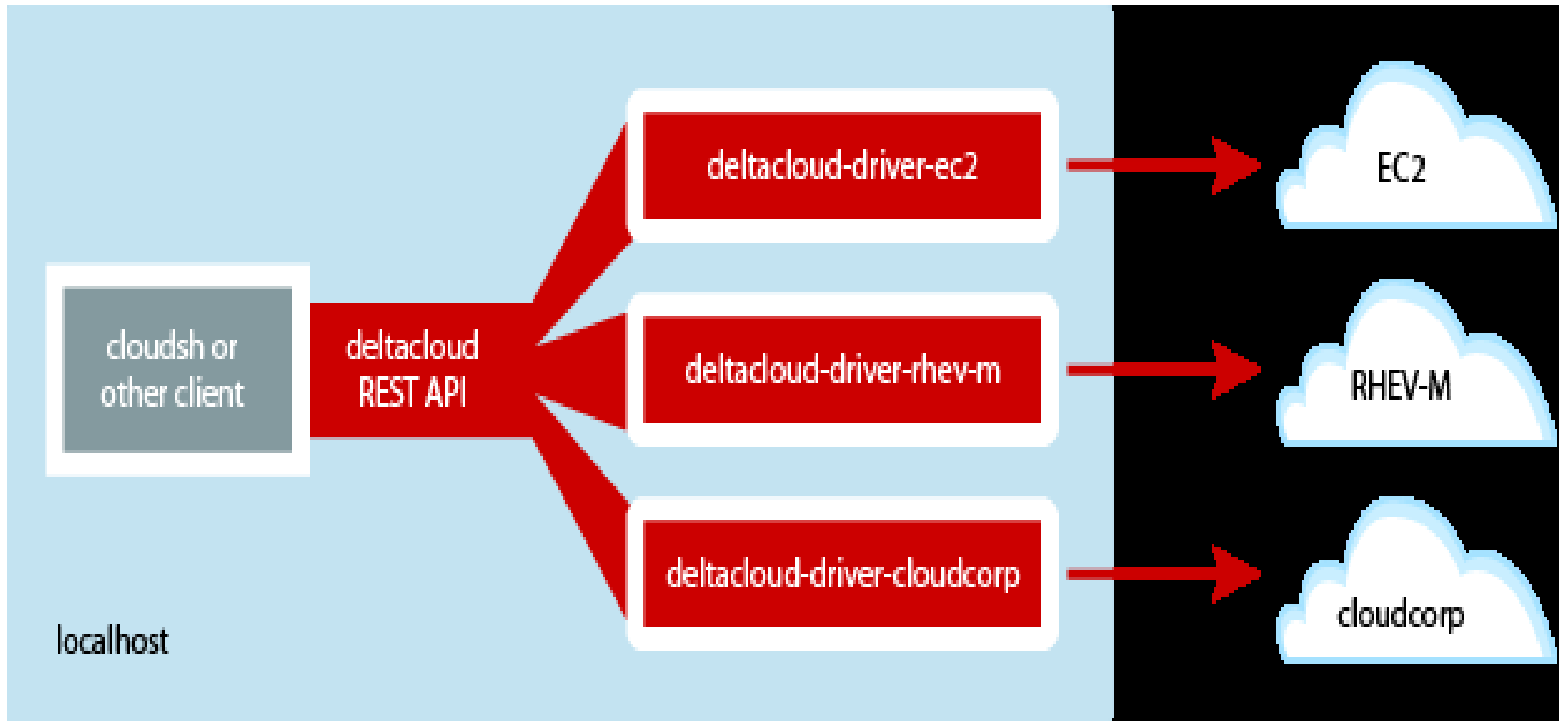
- Other than Hype? What is it?

- Cloud creates another layer of **abstraction**, and with this abstraction there are properties that we find valuable, these include
 - Self service
 - Location transparency
 - Abstraction from the resource
 - High grained accounting
 - UGAPy (User, Group, Accounts, and Permissions)
 - Location transparent storage
 - Location transparent services
 - Higher level API's
 -

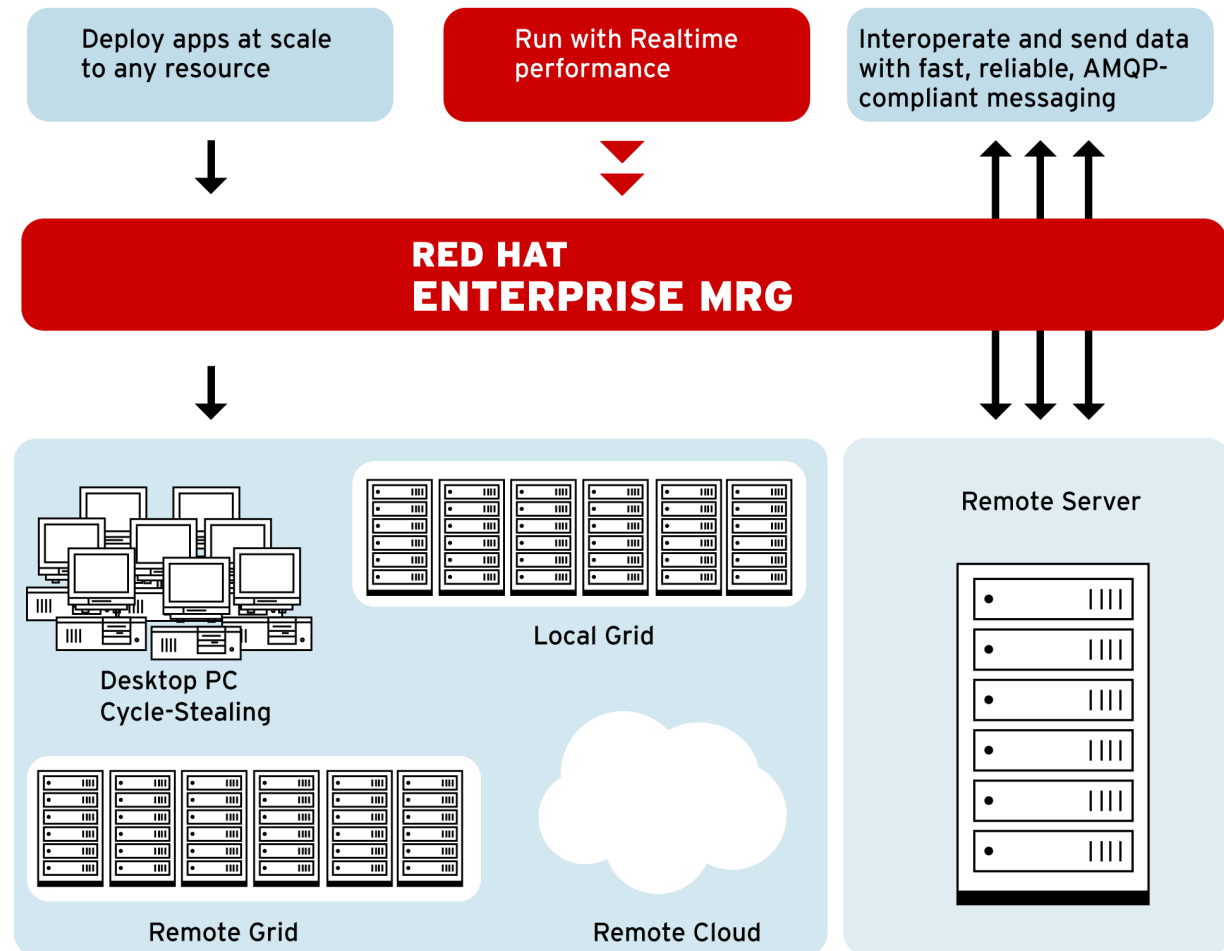
What is Cloud

- Cloud: Abstracts the Resource
- Visualization: Abstracts the Hardware
- Base Metal: Full access

- The Cloud Engine '*Cloud provider engine*'
 - UGAPY, Resources, Accounting, Image Management, Scheduling
- Guest/Image creation tool chain
 - Everything required to build RHEL/ Red Hat/ Jboss/ your guests
- Services with node API's
 - Inter-node Storage & Messaging, Jboss, other PaaS services
- Self Service API's (Delta-Cloud Framework) --- www.deltacloud.org
 - API and public cloud drivers, EC2/other tool drivers

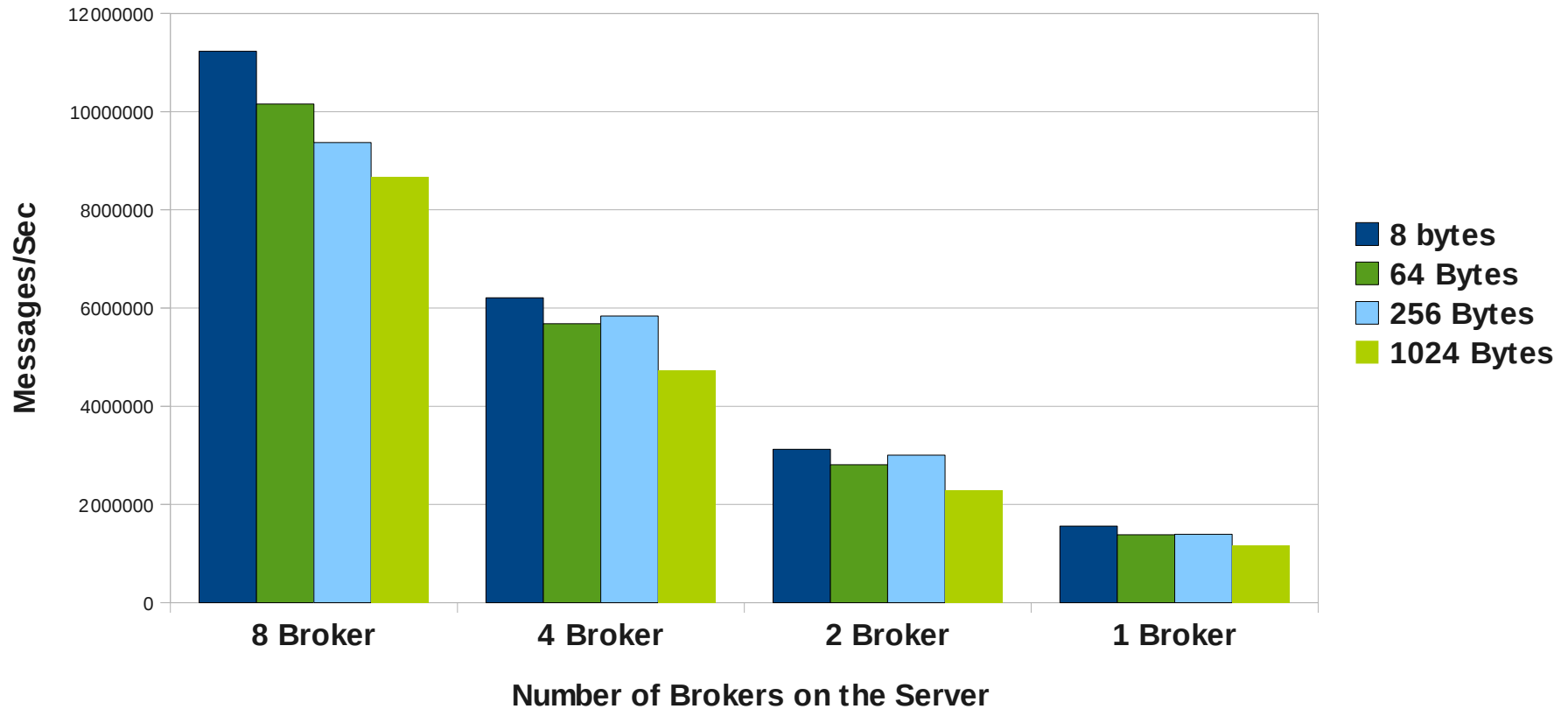


- Integrated platform for high performance distributed computing
 - High speed, interoperable, open standard **Messaging**
 - Deterministic, low-latency **Realtime** kernel
 - High performance & throughput computing **Grid** scheduler for distributed workloads and Cloud computing



AMQP, HP Performance, scale up.

Single HP Nehalem BL460c 40G Infiniband AMQP Perf test



two Intel(R) Xeon(R) CPU X5570 @ 2.93GHz per blade (Nehalem) (2.93 GHz, 8MB L3 cache, 95W,

Memory 24GB(6x4GB) , Memory Type DDR3-1333, HT, Turbo 2/2/3/3)

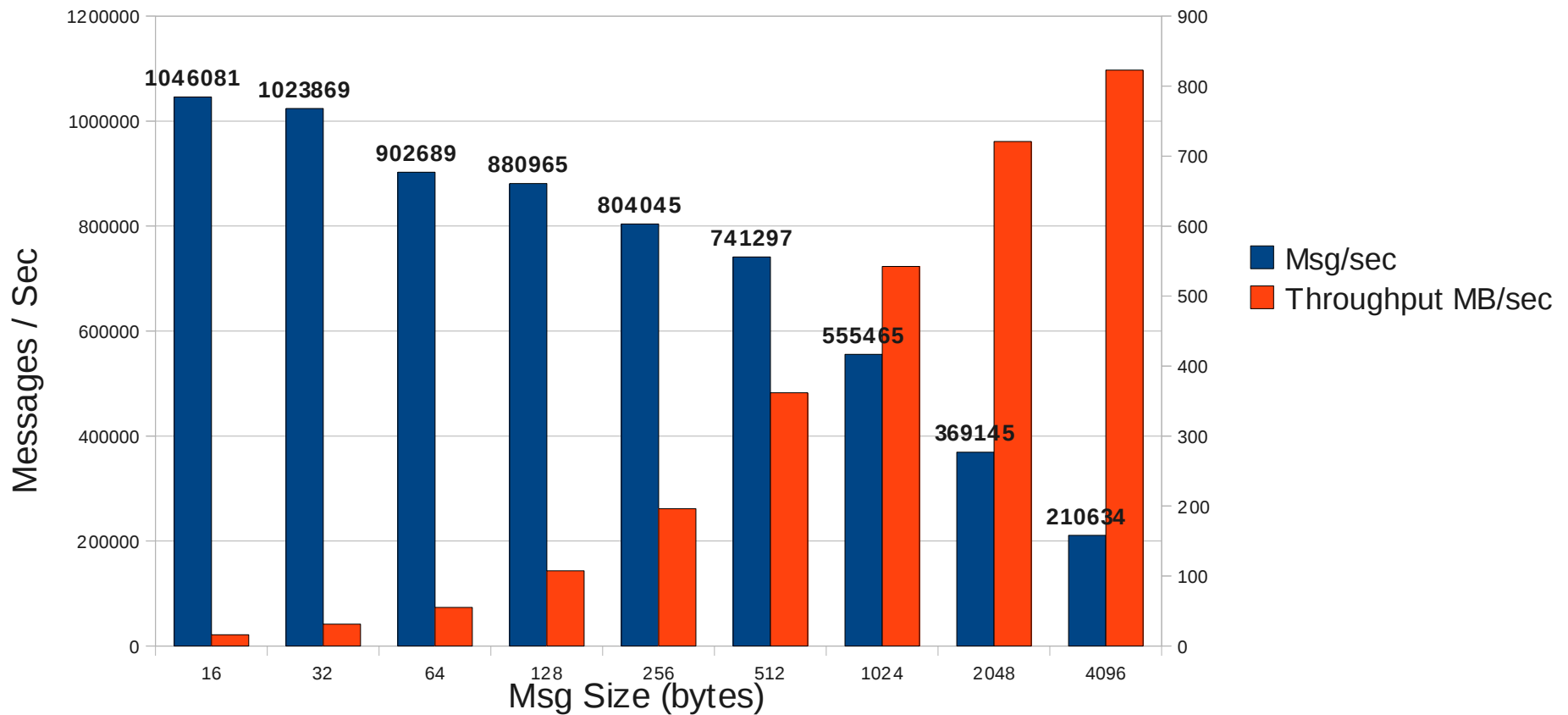
Infiniband 4X QDR IB Dual-port Mezzanine HCAs(1 port connected)

Infiniband Switch BLc 4X QDR IB Switch

KVM Performance - AMQP Messaging

Note, < 5% performance loss from Bare Meta

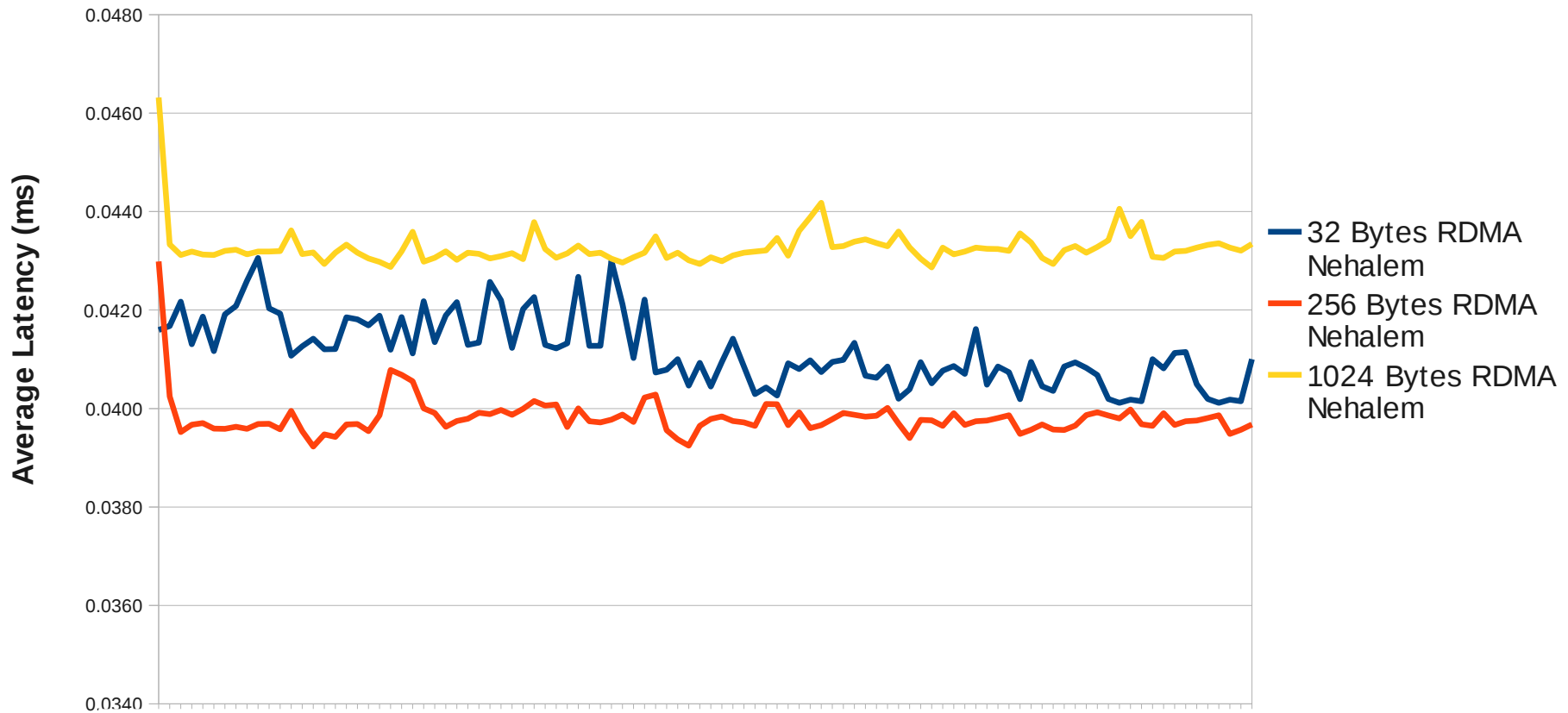
RHEL 5.4 KVM AMQP 2-Guest



MRG Messaging RDMA Latency: OFED

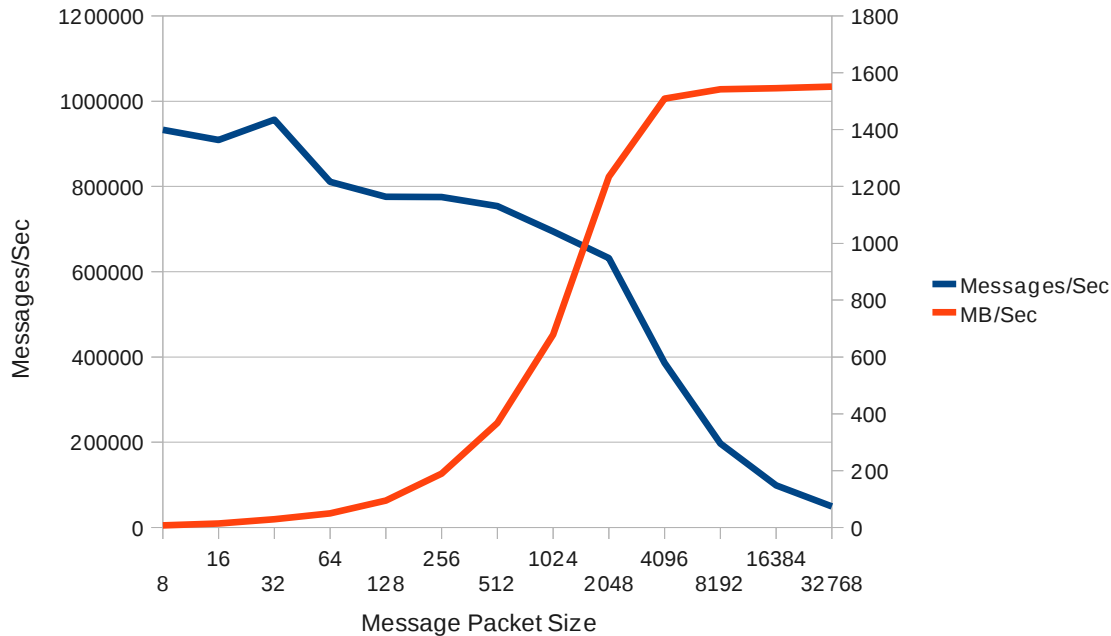
Under 40 Microseconds Reliably Acknowledged

MRG Messaging Latency Test on HP BL460c G6 Infiniband
100K Message Rate



10G has two options iWARP & RoCEE

RHEL55 Mellanox 10Gb AMQP Perfetest



OFED verbs latency for RoCEE:

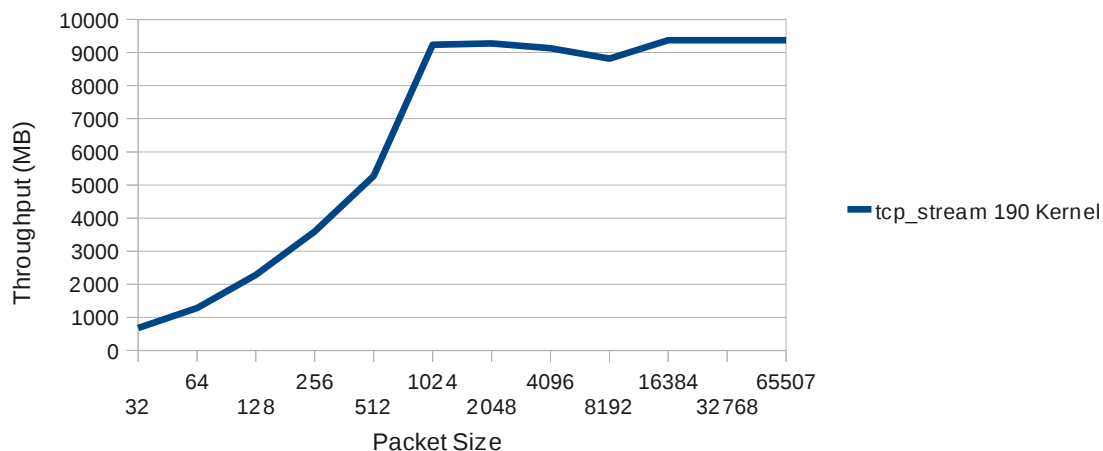
Not published yet - dependencies from OFED 1.5.1 are not yet in the upstream kernel.

Holding our breath for RoCEE latency numbers....

Want it in RHEL?

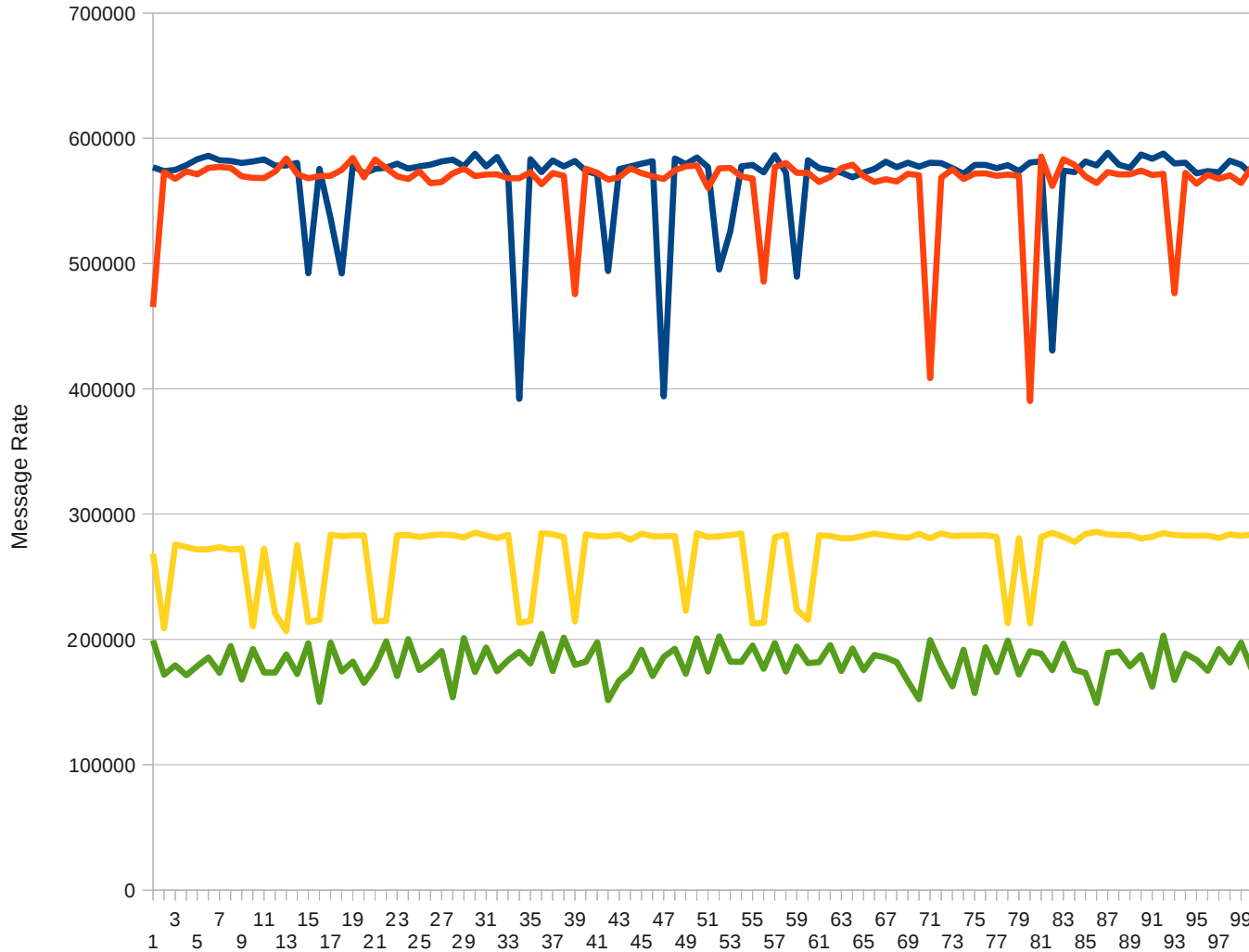
- get it upstream
- work with us to get it upstream
- Anyhow, just get it upstream!

RHEL 5.5 Mellanox 10Gb Netperf



MRG Messaging Durable Messaging Throughput

MRG Durable Messaging Throughput Across Different Storage Types

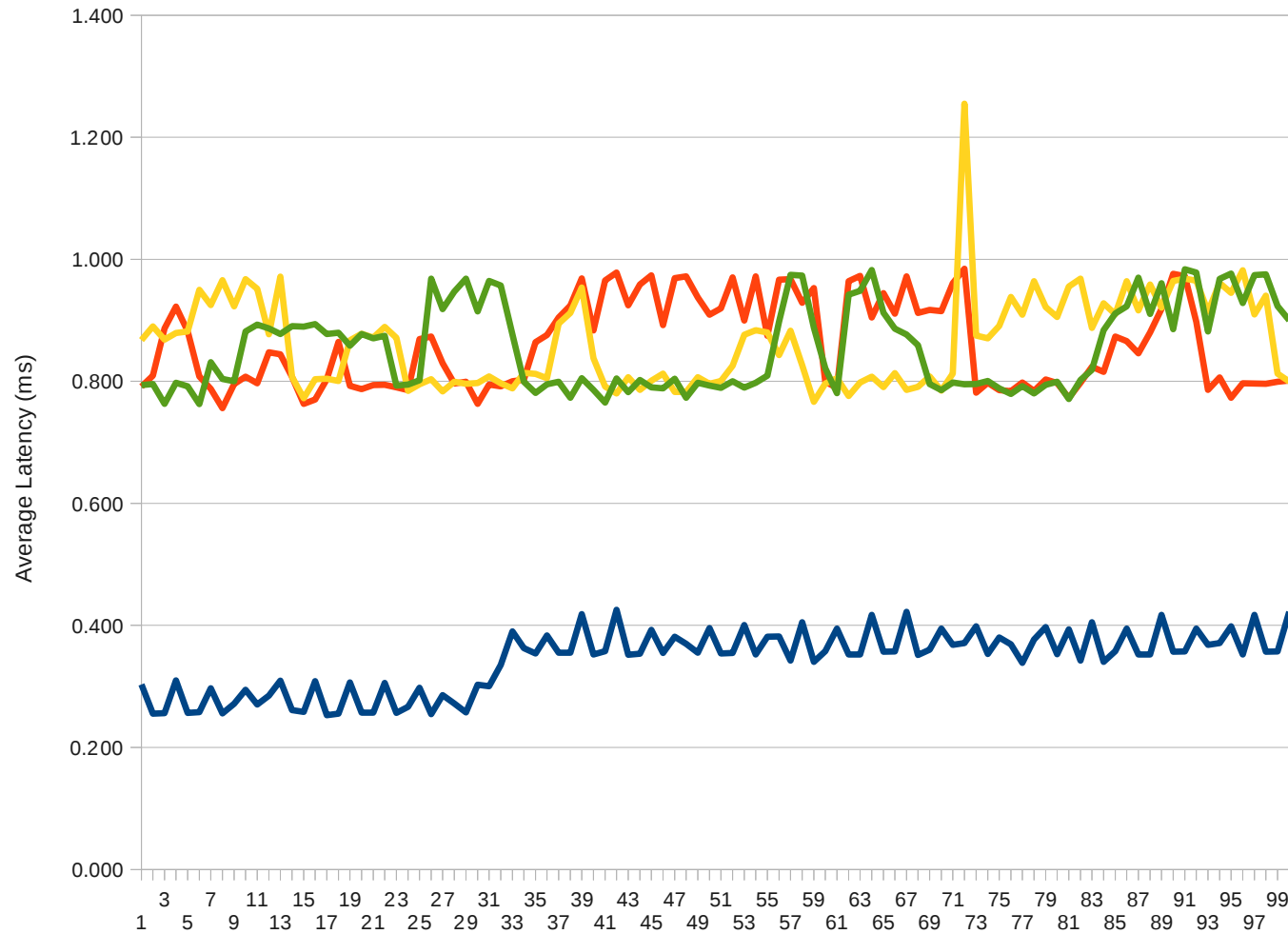


- Intel 16 CPU Hapertown
- 12GB memory 667 Memory speed
- HP IO Fusion
- 32-byte messages

- 1 NIC
- 1 NIC Durable IO Fusion Card
- 1 NIC Durable Fiber Disk
- 1 NIC Durable Internal SCSI drive

MRG Messaging Durable Messaging Latency

Latencytest with Durable Store Different Storage Types



- Intel 16 CPU Hapertown
- 12GB memory 667 Memory speed
- HP IO Fusion
- 32-byte messages

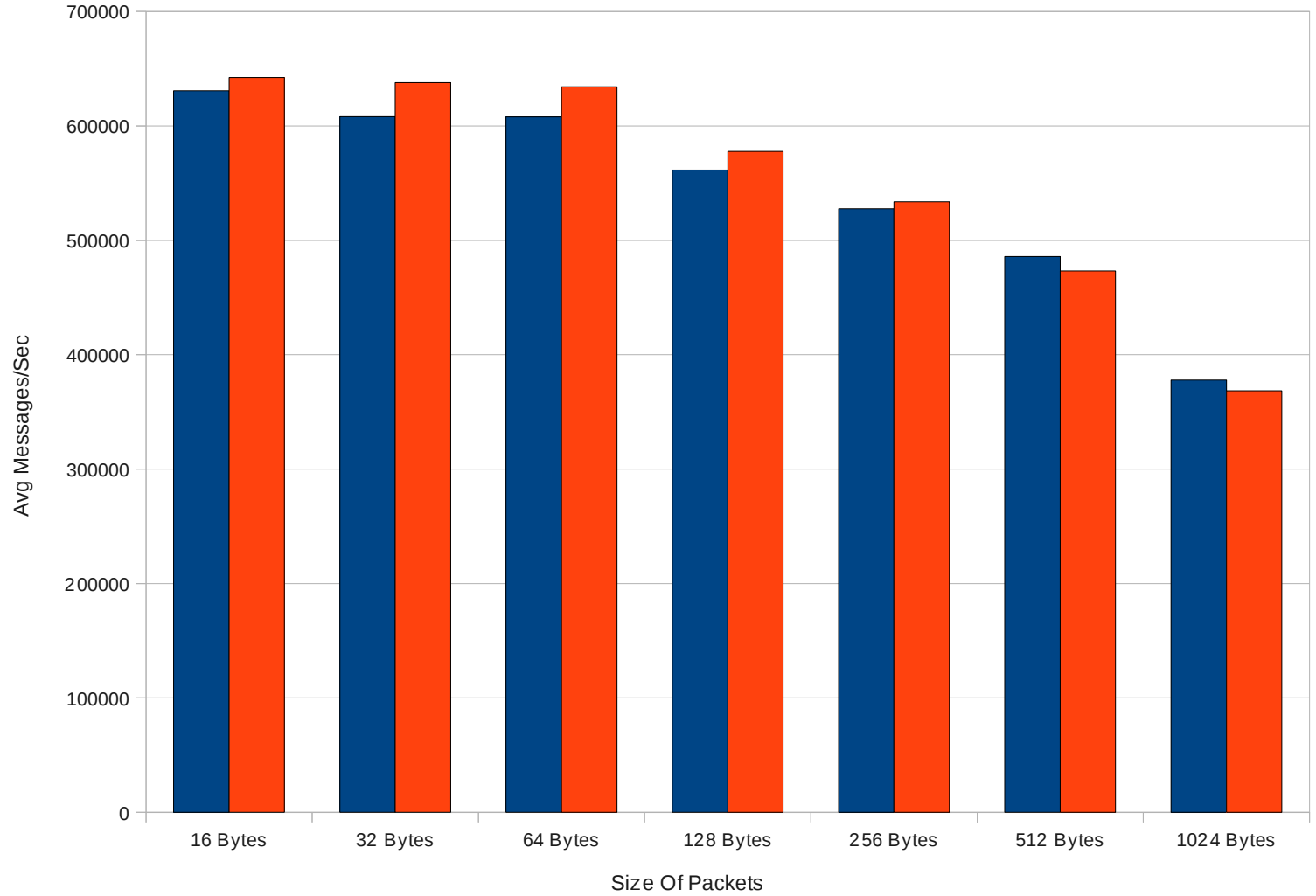
- 1 NIC No Durable
- 1 NIC Iofusion Durable
- 1 NIC Fiber on durable
- 1 NIC Sata Durable

MRG Messaging Clustered Throughput with RDMA



Note: RDMA allows us to go full HA with < 5% performance impact!

AMQP Perfetest tesing Corosync



- 3-Node cluster
- IBM x3550
- 2 x 4-core Xeon E5420 2.5MHz (Harpertown)
- 16GB ram 266 MHz
- Mellanox MT25204 [InfiniHost III Lx HCA]

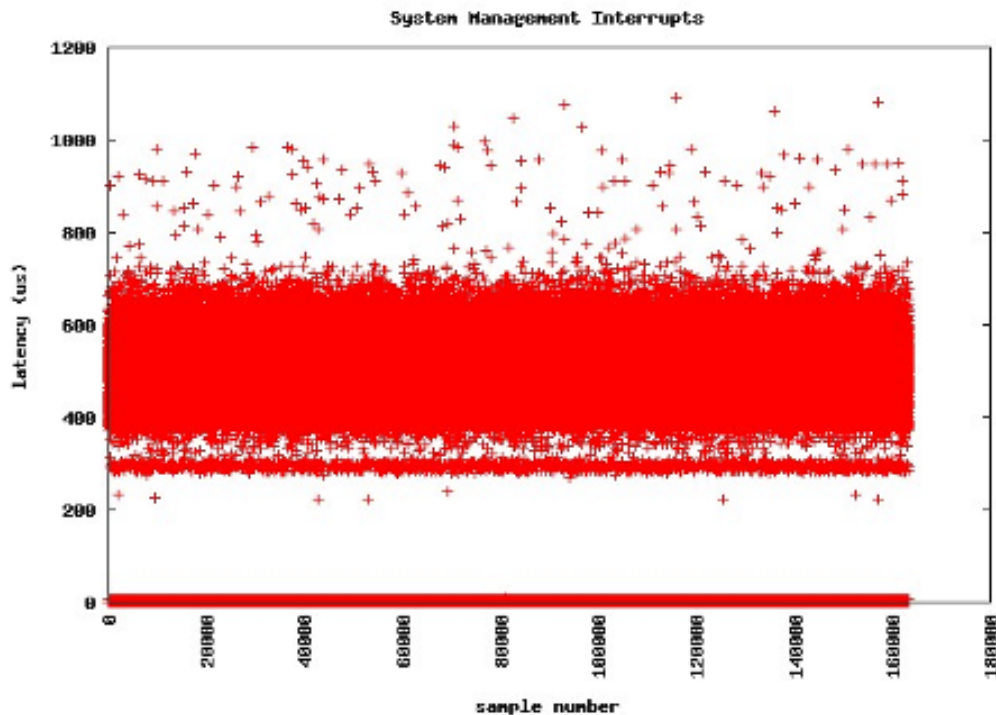
■ No CoroSyc
■ CoroSyc



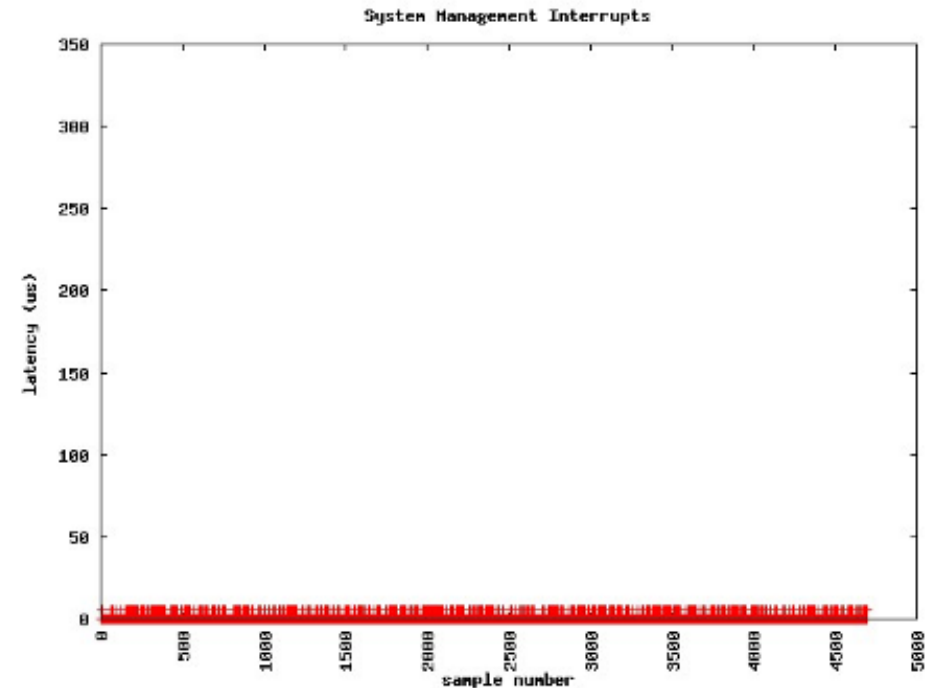
BIOS Option for Low Latency Apps

- Disable frequent SMIs used for Dynamic Power Savings Mode, CPU Utilization monitoring, P-state monitoring and ECC monitoring
- Benefits both RHEL & MRG operating environments.

Latency spikes with standard BIOS settings



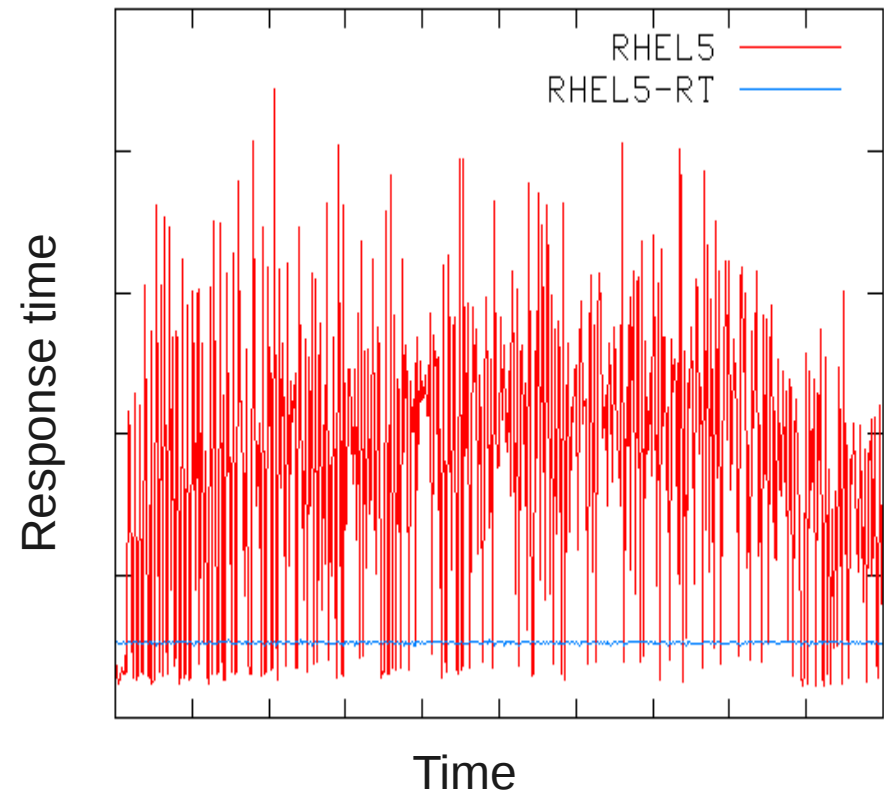
Latencies when SMIs disabled in BIOS



MRG – Realtime RHEL & RDMA

- Enables applications and transactions to run predictably, with guaranteed response times
- Upgrades RHEL 5 to realtime OS
 - Provides replacement kernel for RHEL 5; x86/x86_64
 - Preserves RHEL Application Compatibility
- Note that Tuning has less impact on RDMA, as IRQ affinity is less of an issue for verbs. However SMIs is no discriminator of RDMA.

RHEL5 vs. RHEL5-RT response timings



https://hardware.redhat.com/list.cgi?version=5&field0-0-0=cf_fixed_in&type0-0-0=substring&value0-0-0=MRG

<http://www.redhat.com/mrg>
<http://www.deltacloud.org>