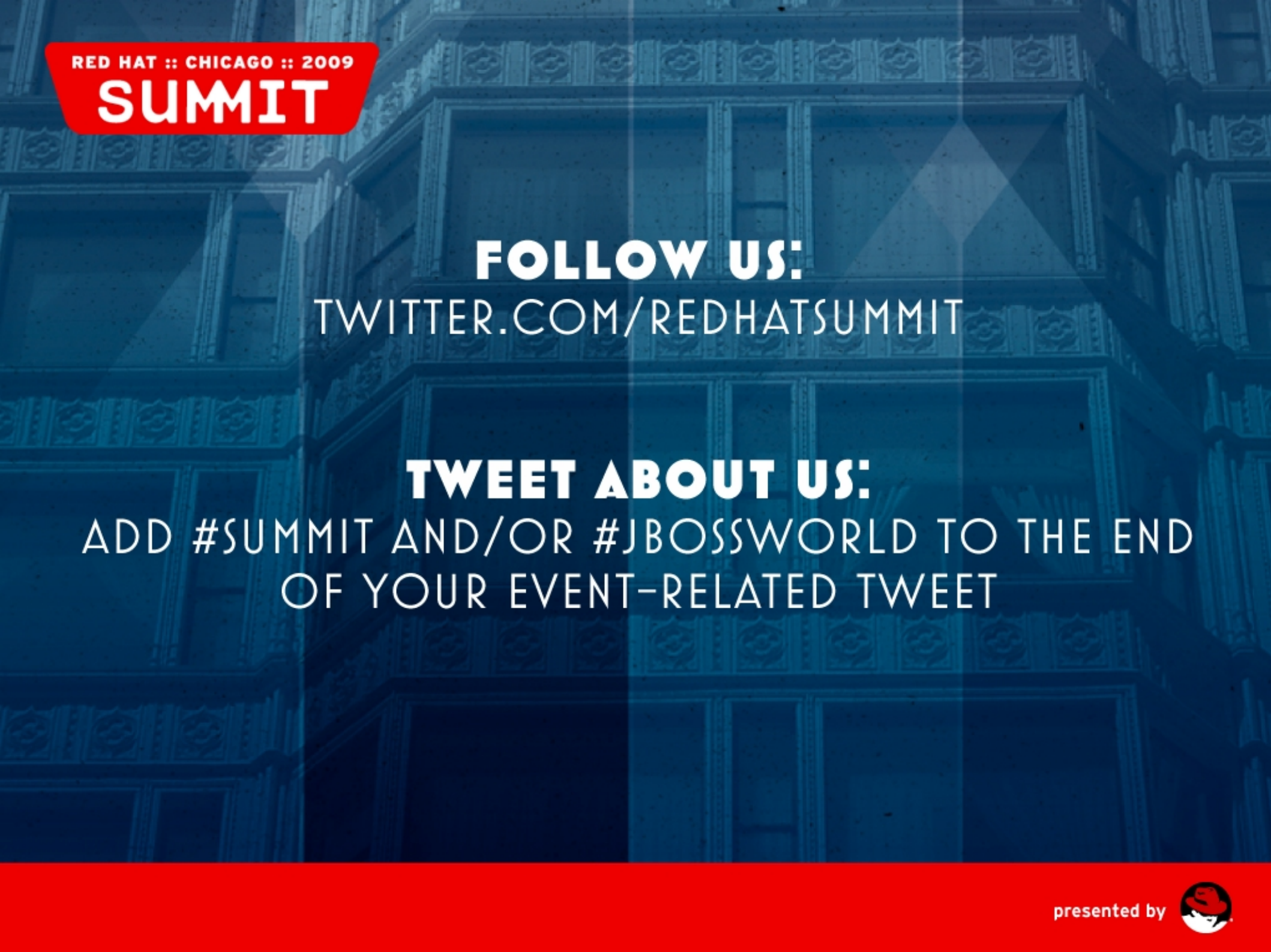

RED HAT :: CHICAGO :: 2009
SUMMIT
FOLLOW US:
TWITTER.COM/REDHATSUMMIT
TWEET ABOUT US:
ADD #SUMMIT AND/OR #JBOSSWORLD TO THE END
OF YOUR EVENT-RELATED TWEET
presented by

RED HAT :: CHICAGO :: 2009
SUMMIT

# Red Hat Enterprise MRG Update

Bryan Che
Product Manager, Red Hat
September 2, 2009

presented by

# About Red Hat Enterprise MRG

Integrated platform for high performance distributed computing

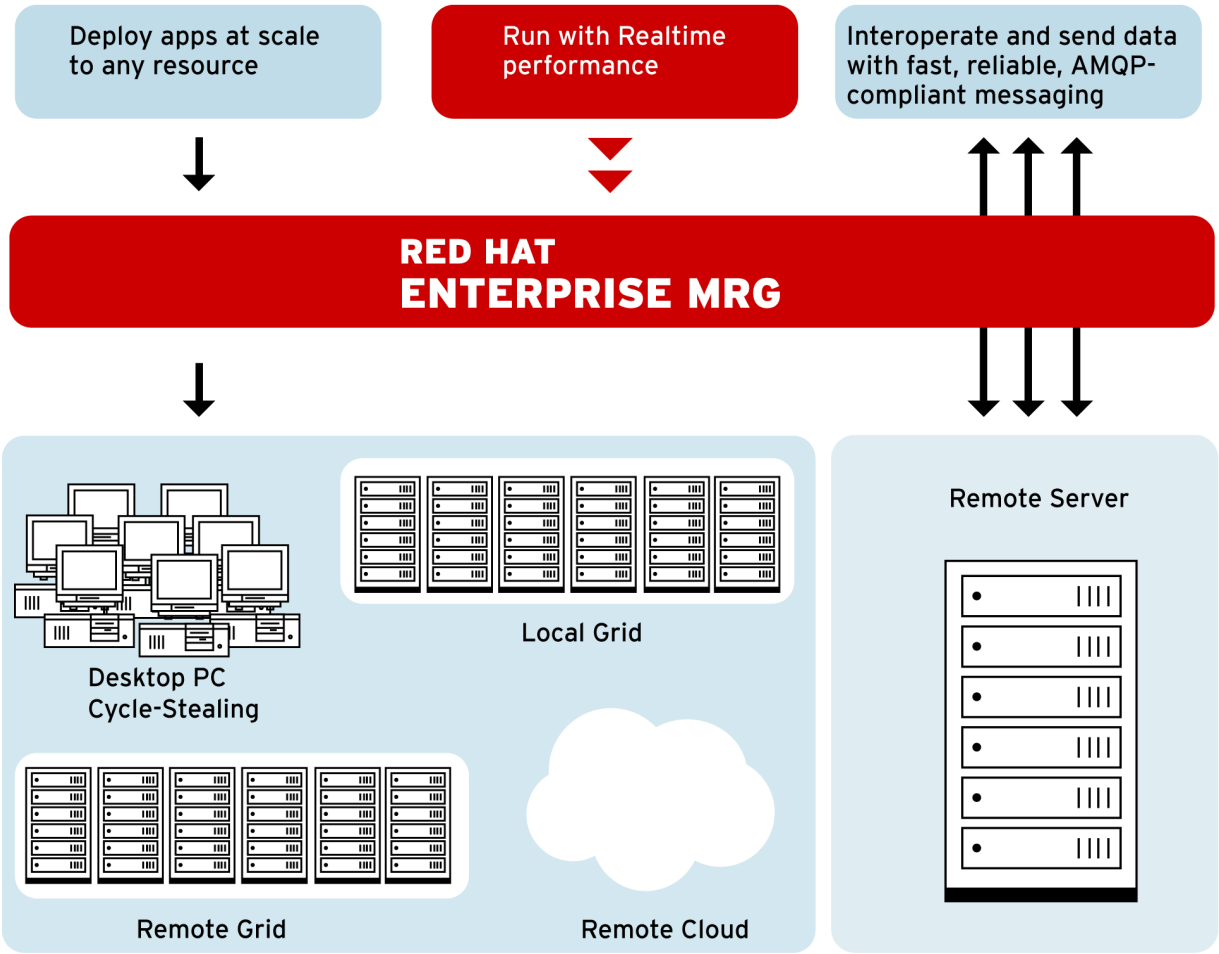
High speed, interoperable, open standard AMQP **Messaging**

Deterministic, low-latency **Realtime** kernel

High performance & throughput computing **Grid** scheduler for distributed workloads and Utility/Cloud computing

MRG 1.0 launched at Red Hat Summit 2008 in Boston

MRG 1.1 released February 2009



# MRG Messaging

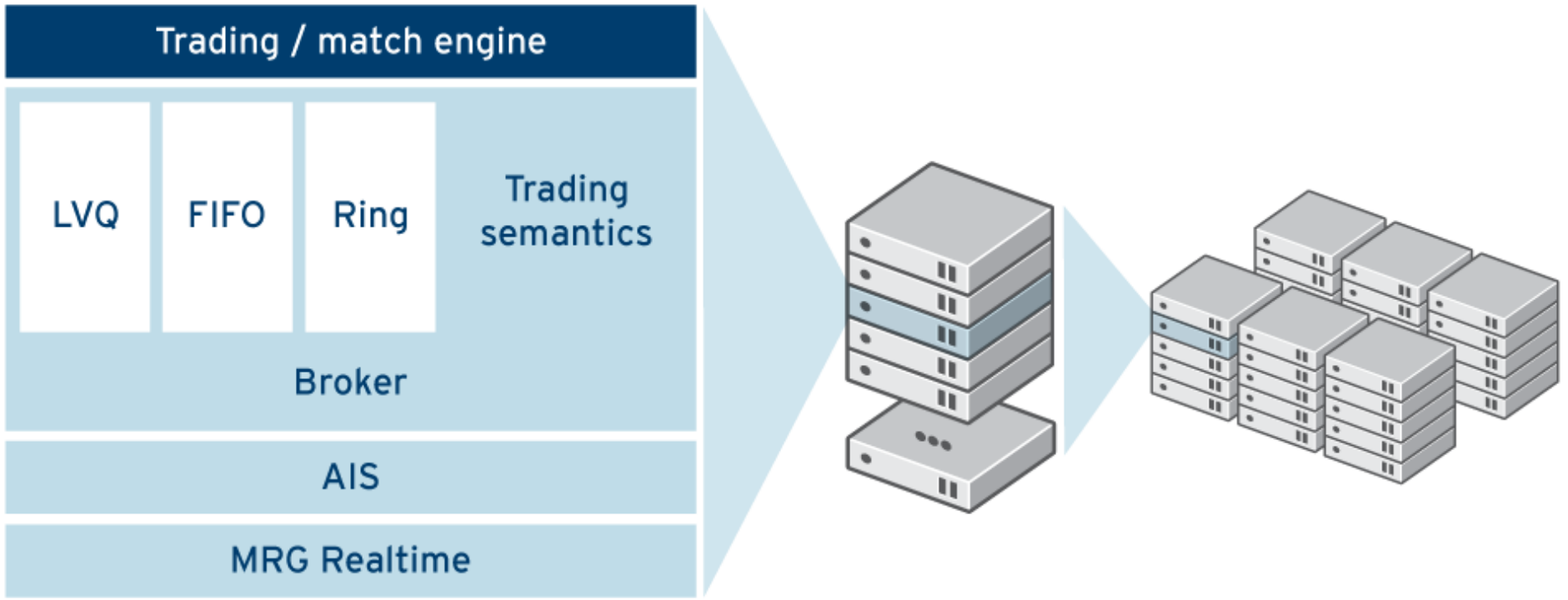
## Enterprise Messaging System that

- **Implements AMQP** (Advanced Message Queuing Protocol), the first open messaging standard
- **Spans all use cases in one implementation** to consolidate architectural silos (fast messaging, reliable messaging, large file transfer, publish/subscribe, eventing, etc)
- Uses Linux-specific optimizations to achieve **breakthrough performance** on Red Hat Enterprise Linux and MRG Realtime
- **Runs on non-Linux platforms** without the full performance and quality of service benefits that Red Hat Enterprise Linux provides

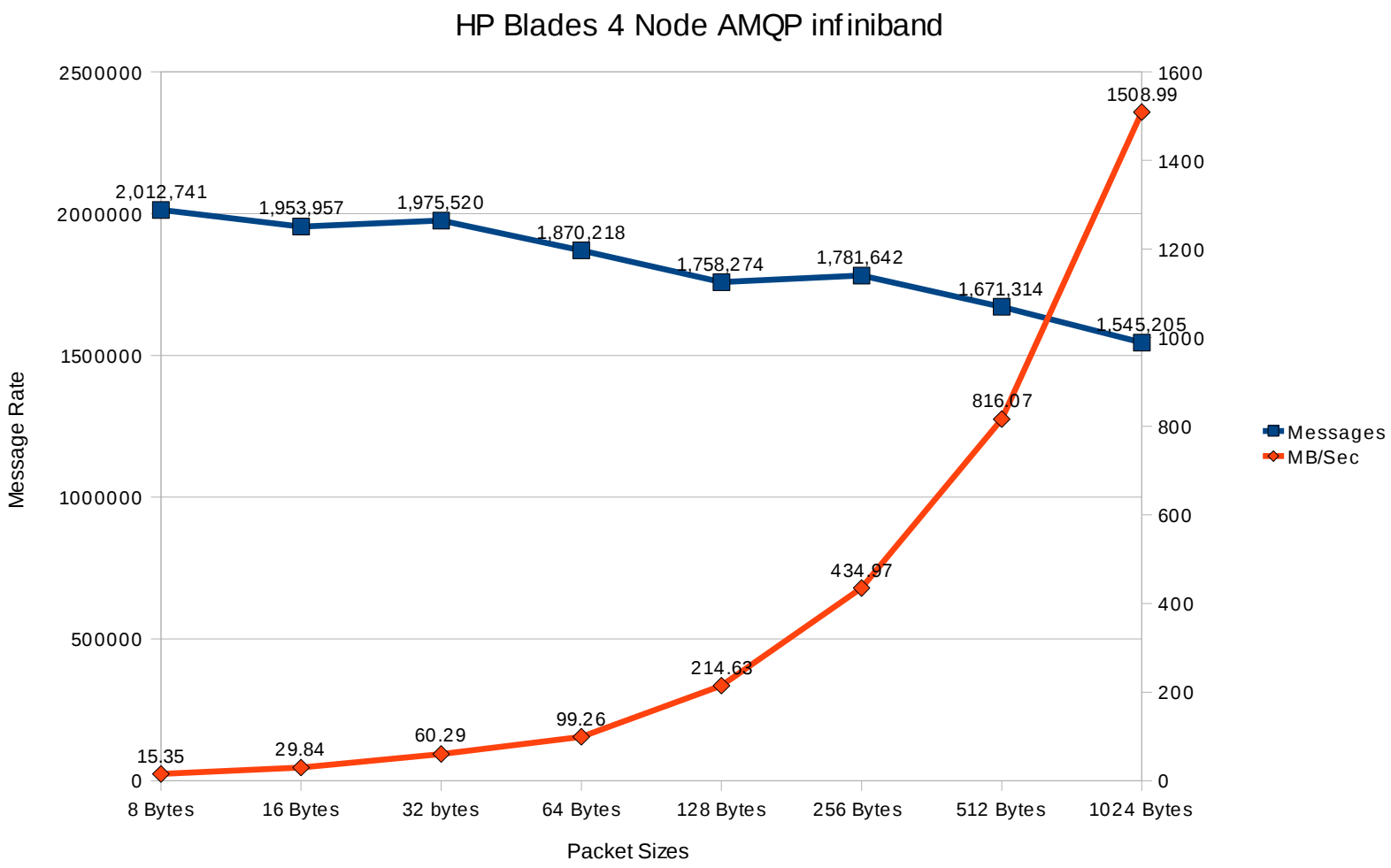
## Features introduced in MRG 1.1

- Infiniband RDMA support for ultra low latency messaging
- Active-Standby/Active-Active Broker Clustering
- Security: SASL auth, SSL encryption, role-based access control
- Native .NET Messaging Client, including WCF
- Performance enhancements
- Queue Semantics: Ring Queue, Last Value Queue, TTL
- Federation dynamic routes
- JMSXUserId support

# MRG Messaging Trading System Deployment

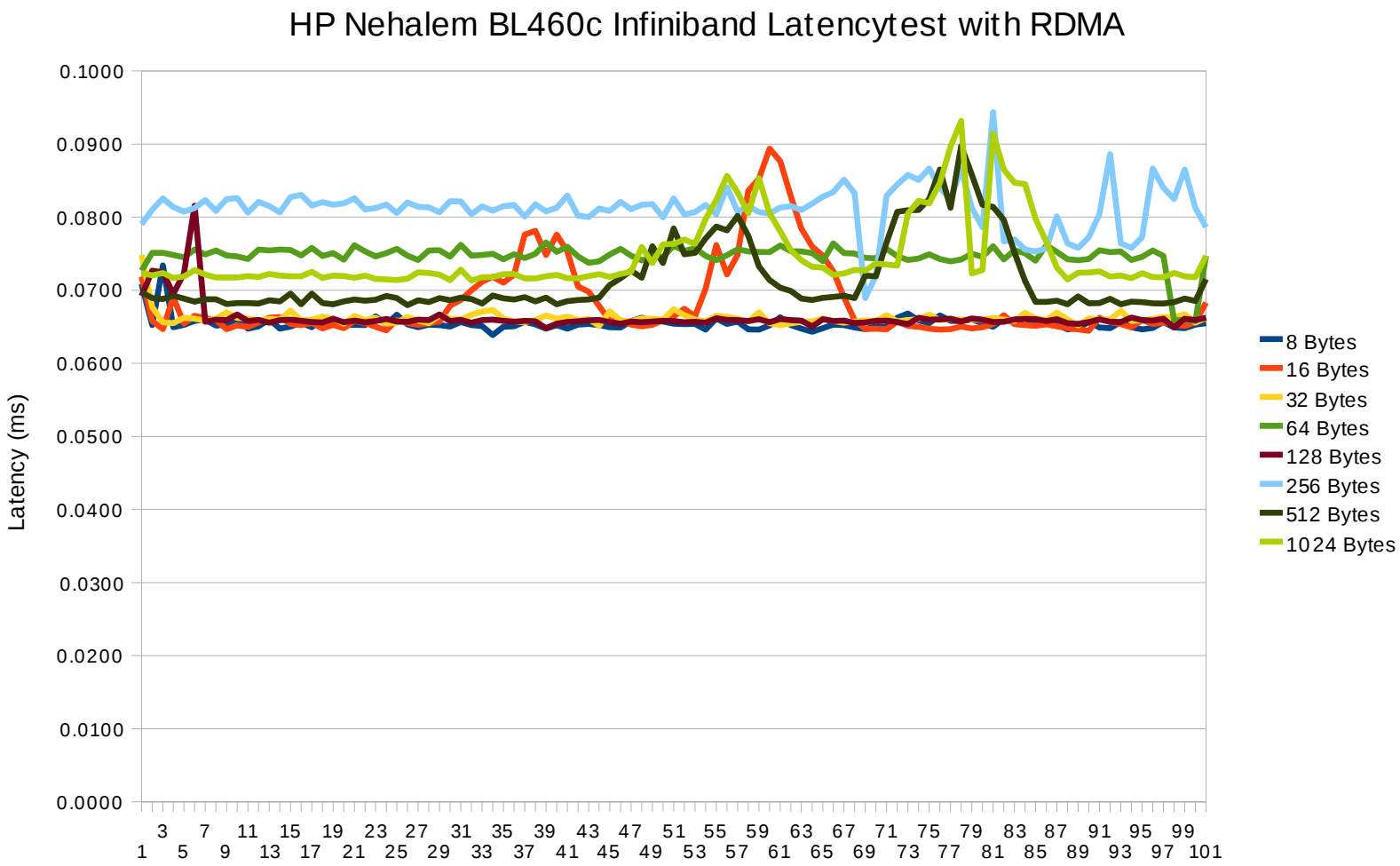


# MRG Messaging Infiniband Throughput: Over 2 Million Reliable Messages/second




- 8 systems: 4 load drivers, 4 compute nodes
- HP ProLiant BL460c G6
- Dual Intel(R) Xeon(R) CPU X5570 @ 2.93GHz
- 24GB RAM
- Infiniband 4X QDR IB Dual-port Mezzanine HCAs(1 port connected)
- BLc 4X QDR IB Switch

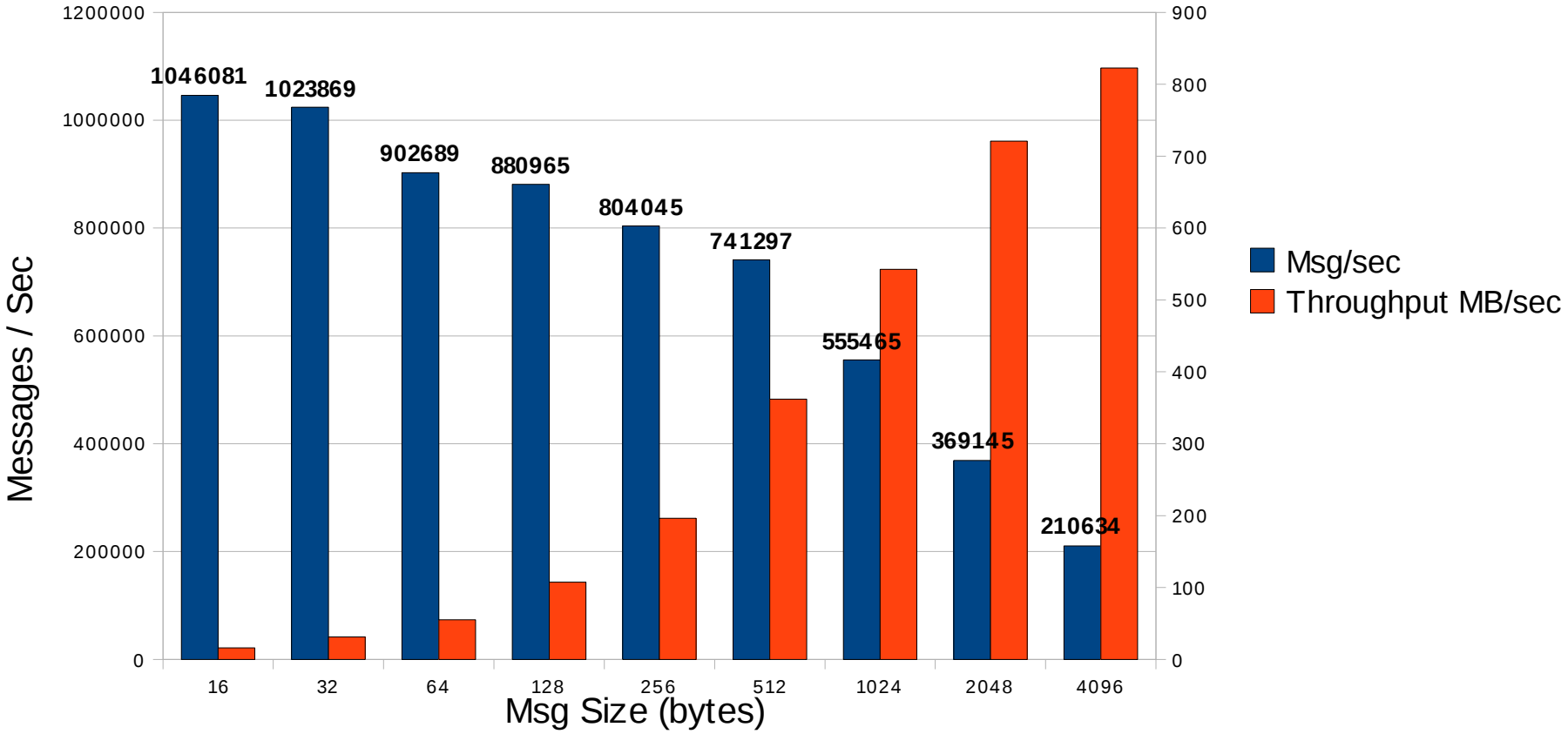
# MRG Messaging Infiniband RDMA Latency: Under 65 Microseconds Reliably Acknowledged



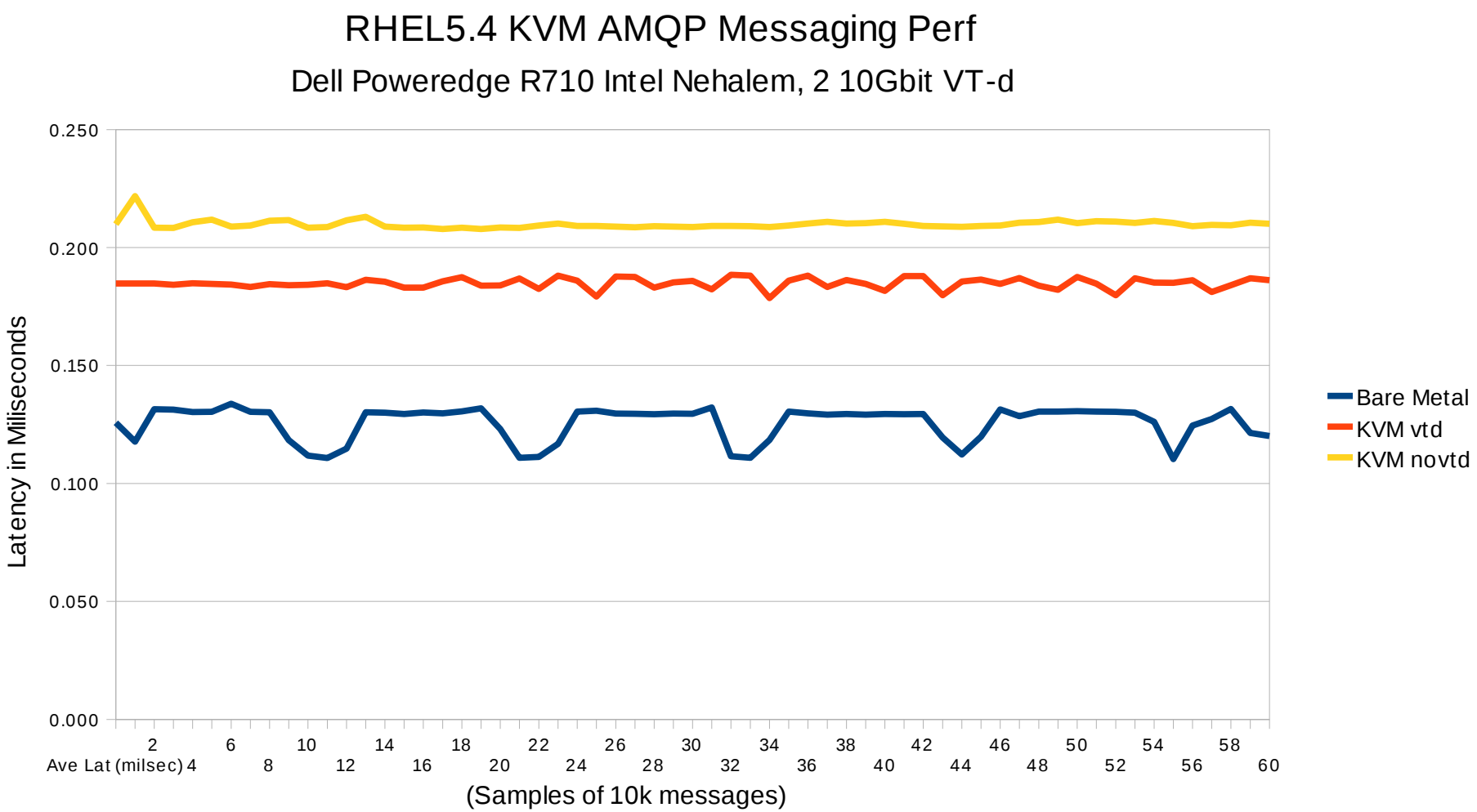
# MRG Messaging on KVM Virtualized Performance: Over 1 million Messages/Second Throughput

RHEL 5.4 KVM AMQP 2-Guest

Dell Poweredge R710 Intel Nehalem, 2 10Gbit VT-d



# MRG Messaging on KVM Virtualized Performance: <200 Microsecond Latency



# MRG Realtime

Enables applications and transactions to run predictably, with guaranteed response times

Provides microsecond accuracy

Provides competitive advantage & meets SLA's

Travel web site: missed booking

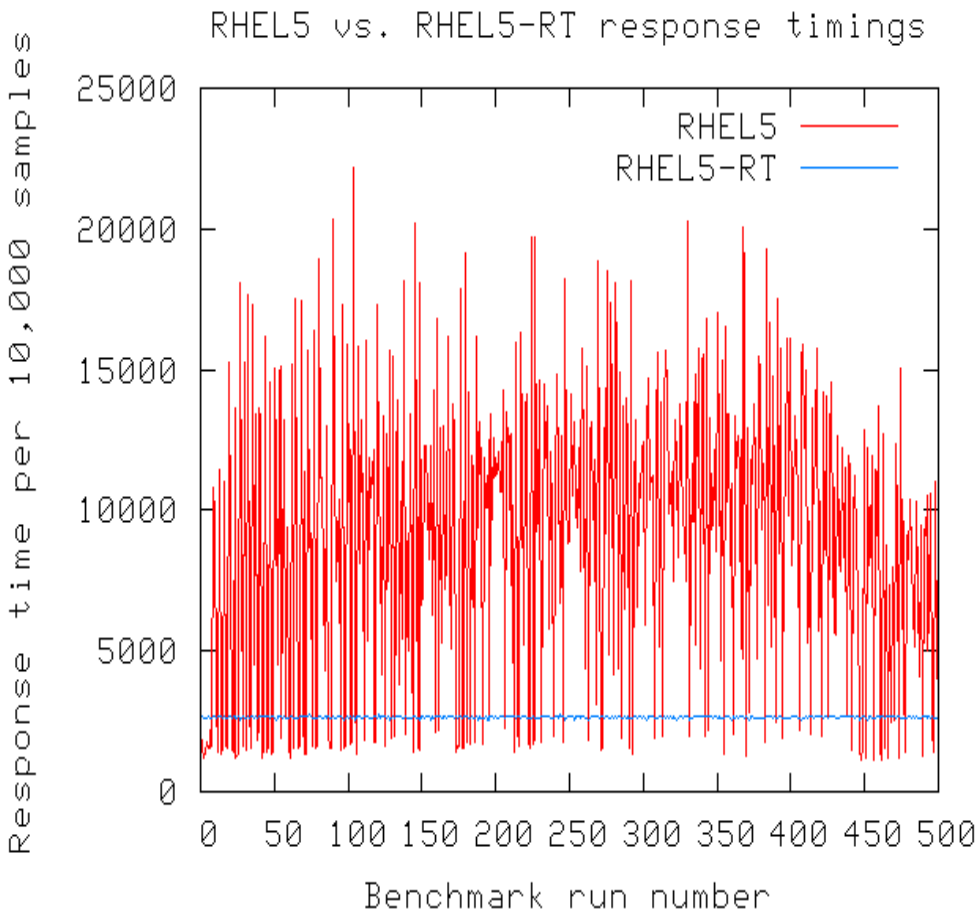
Program trading: missed trades

Command & Control: life & death

Upgrades RHEL5 to Realtime OS

- Provides replacement kernel for RHEL 5; x86/x86_64
- Preserves RHEL Application Compatibility

Red Hat Leads Upstream Linux Realtime Development



# MRG Realtime Tools

MRG includes a new tuning tool, TUNA

Dynamically control tuning parameters like process affinity, parent & threads, scheduling policy, device IRQ priorities, etc.

MRG includes a new MRG Realtime Latency Tracer

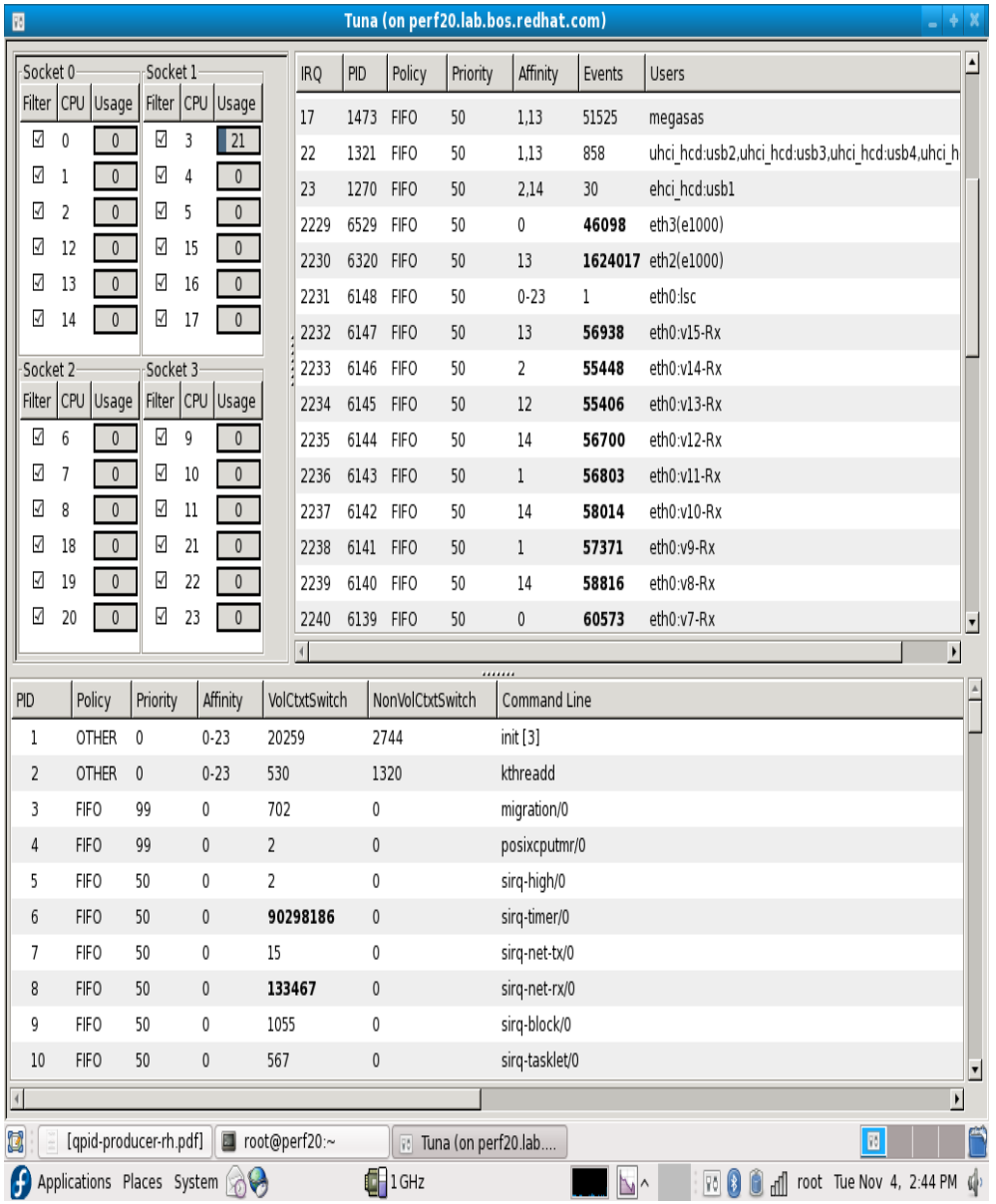
Runtime trace capture of longest latency codepaths – both kernel and application. Peak detector

Selectable triggers for threshold tracing

Detailed kernel profiles based on latency triggers

Existing standard RHEL5 based performance monitoring tools remain relevant

Gdb, OProfile, SystemTap, kprobe, kexec/kdump, etc.



# Hardware Matters

Hardware can have a big effect on realtime performance

Hardware drivers may need to be updated to handle threaded interrupts

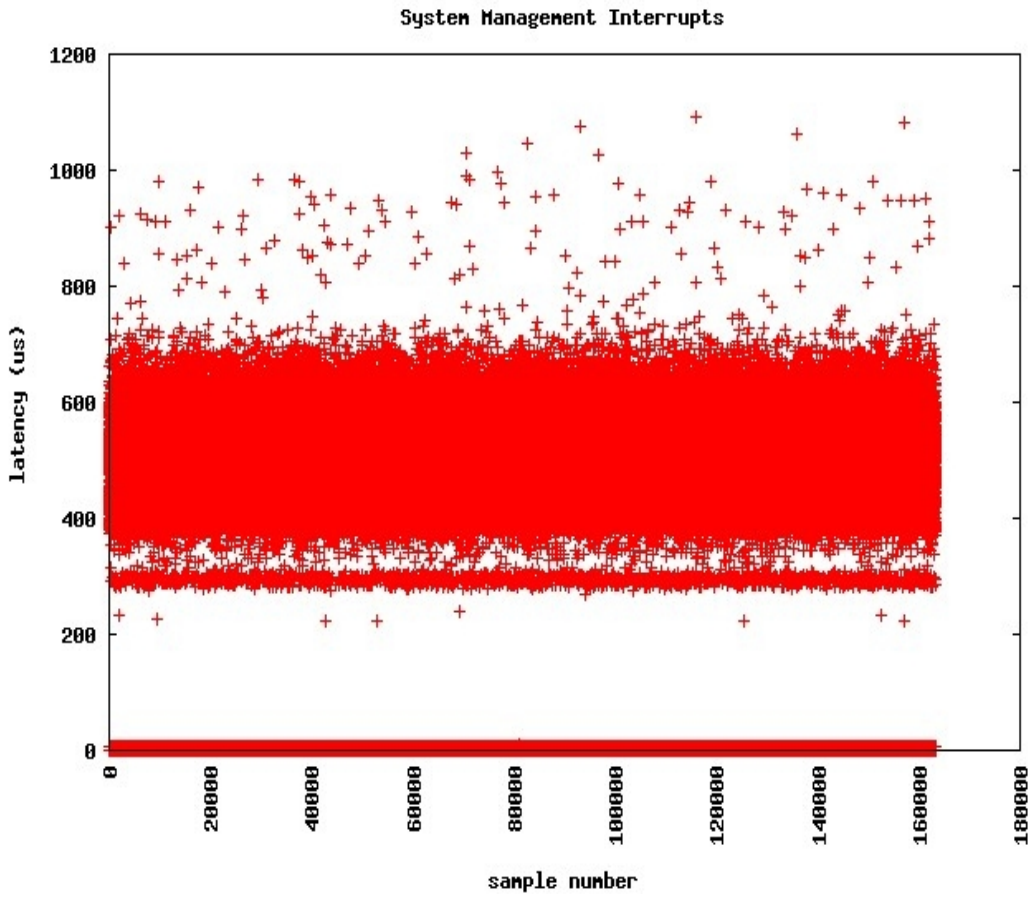
Many system BIOS's include System Management Interrupts (SMIs)

Cause non-deterministic latency *beneath* the operation system by taking CPU cycles for things like power management, administration

SMI latencies *cannot* be resolved by realtime linux—they require the hardware OEM to remove SMIs or make them configurable

Red Hat has worked with OEMs to certify systems for MRG Realtime

http://www.redhat.com/mrg/hardware/



# MRG Grid

## Provides both HPC/HTC and Utility/Cloud Computing

Brings advantages of scale-out and flexible deployment to any application or workload

Delivers better asset utilization, allowing applications to take advantage of all available computing resources

Handles “Holiday Rushes”

Executes across multiple platforms and in virtual machines

## Provides seamless and flexible computing across:

Local grids

Remote grids

Remote clouds (Amazon EC2)

Cycle-harvesting from desktop PCs

## Fully supported in MRG 1.1

- Web-Based Management Console
  - Unified management across all of MRG for job, system, license management, and workload management/monitoring
- Low Latency Scheduling
  - Sub-second job scheduling via AMQP Messaging
- Virtualization Support
- Cloud Integration with Amazon Ec2
- Concurrency Limits
  - limits how much of a certain resource can be used at once
- Dynamic Slots
  - Mark slots as partitionable

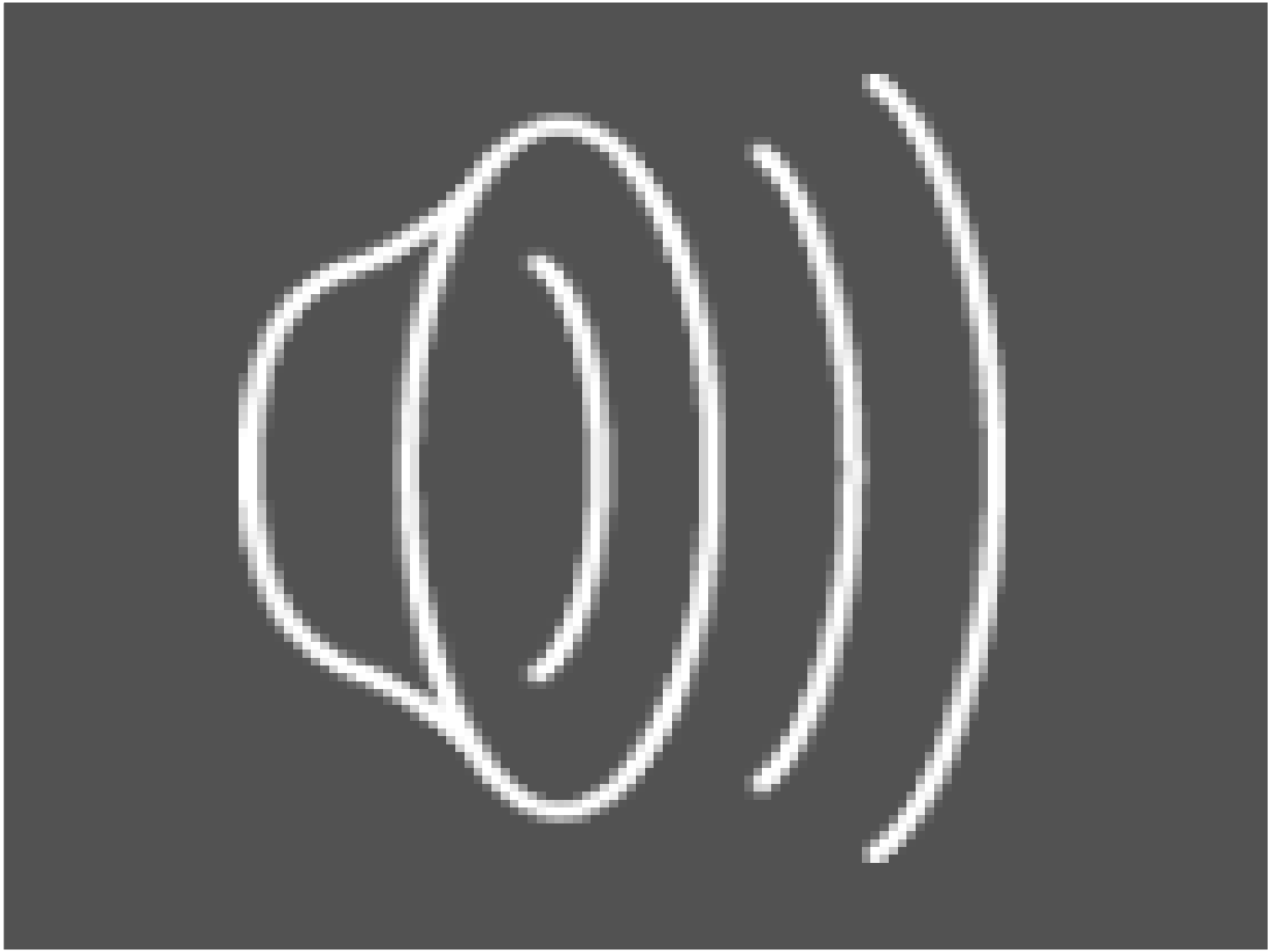
# Cloud Computing with MRG Grid

MRG integrates internal clouds, hybrid clouds, and public clouds from one interface

- Internal Cloud: Build your own cloud with Red Hat Enterprise Virtualization and Red Hat Enterprise MRG
- Hybrid Cloud: Augment your internal cloud with resources from remote grids/clouds, idle resources, virtual machines, bare metal, etc
- Public Cloud: Schedule jobs to Amazon EC2 from the same interface as for local clouds

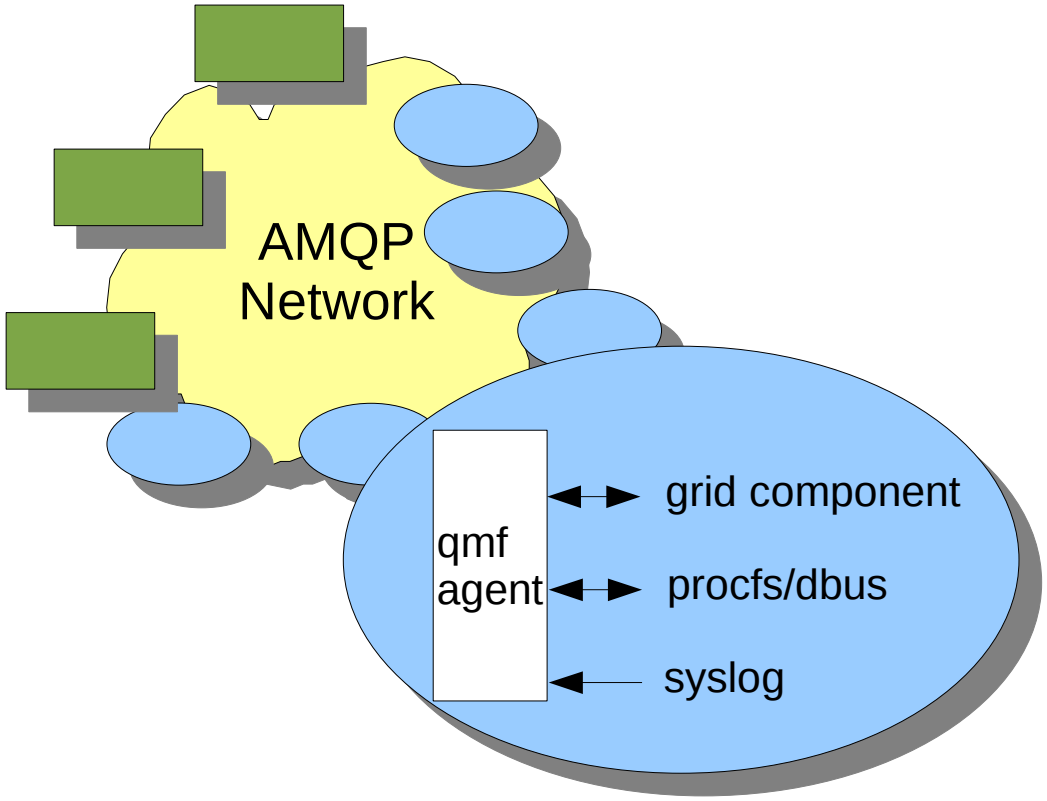
Next Session at 4:30: Building and Leveraging Compute Clouds with Red Hat Enterprise MRG

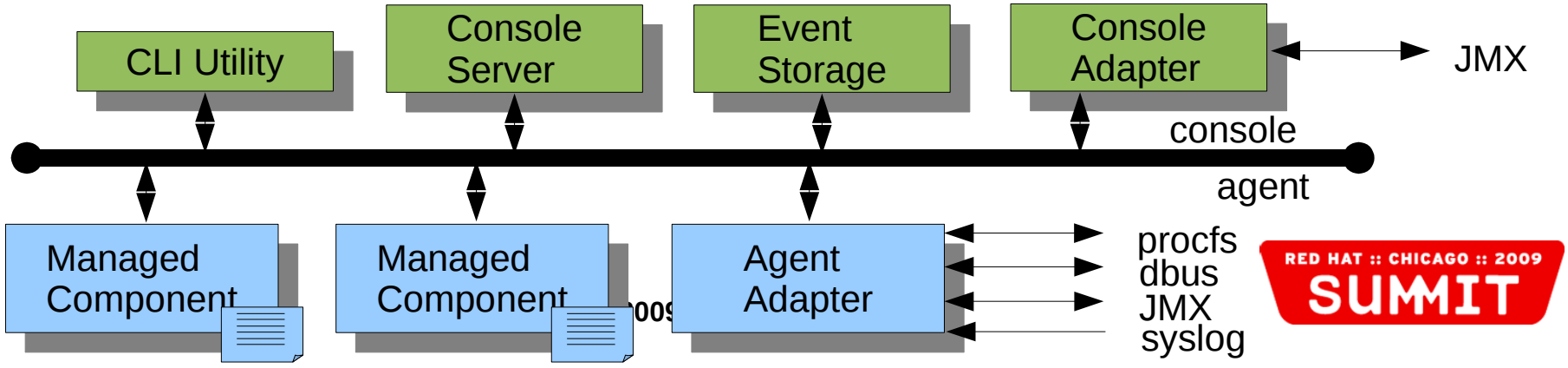
# Cloud Computing with MRG Grid Demo Video



# QMF: AMQP Messaging-Based Management

- Red Hat Enterprise MRG's entire management/monitoring system is AMQP messaging-based
  - Asymmetric, Efficient, Scalable, and Secure
  - Any messaging client can manage
- QMF: AMQP Messaging-Based Management Framework
  - Agent-defined management model (self-describing)
  - Objects (properties, statistics, and methods/controls), Events
  - Ease of development and extensibility



# MRG Messaging Roadmap

## MRG 1.2 Highlights

- Protocol independent API for c++, python and ruby clients
- iwarp support for RDMA
- Durable stores will function in clustered environments
- Performance enhancements

## Beyond MRG 1.2

- AMQP 1.0
- Unix Domain Sockets
- Enhanced performance
- Enhanced tools

**Roadmap Features and Schedules Subject to Change*

# MRG Realtime Roadmap

## MRG 1.2 Highlights

- 2.6.31 kernel
- Improved Throughput Performance
- Rteval realtime evaluation toolkit
- New perf tools
  - New Performance Counter System interface to underlying hardware registers that track events that affect performance. e.g. cache misses, number of context switches, etc.
    - Profile at hardware speeds
  - New perf command for interfacing with the Performance Counter System. e.g. perf top, perf record, perf report

## Beyond MRG 1.2

- Enhanced performance
- Continue to track upstream kernel and merge realtime into the mainline trunk
- Enhanced tools
  - Ftrace user space latency tracing
  - Enhance rteval to provide qualitative assessments
  - Integrate TUNA with Web management console
- Additional hardware support and certification

**Roadmap Features and Schedules Subject to Change*

# MRG Grid Roadmap

## MRG 1.2 Highlights

- New User Tools, including GUI Submit
- Updated Admin Tools
- Windows Execute Node Support
- Resource restrictions
- Updated Amazon EC2 Enhanced Support
- Enhanced scalability

## Beyond MRG 1.2

- Fine-grained SELinux integration & policies (SELinux Sandbox)
- Configuration for diskless Condor
- Dynamic provisioning preemption support
- Pre-built configurations and job templates
- Integration with RHEV

**Roadmap Features and Schedules Subject to Change*

# Additional Information

http://www.redhat.com/mrg


Next Session at 4:30:

Building and Leveraging Compute Clouds with Red Hat Enterprise MRG

# QUESTIONS?

## TELL US WHAT YOU THINK:
## REDHAT.COM/SUMMIT-SURVEY