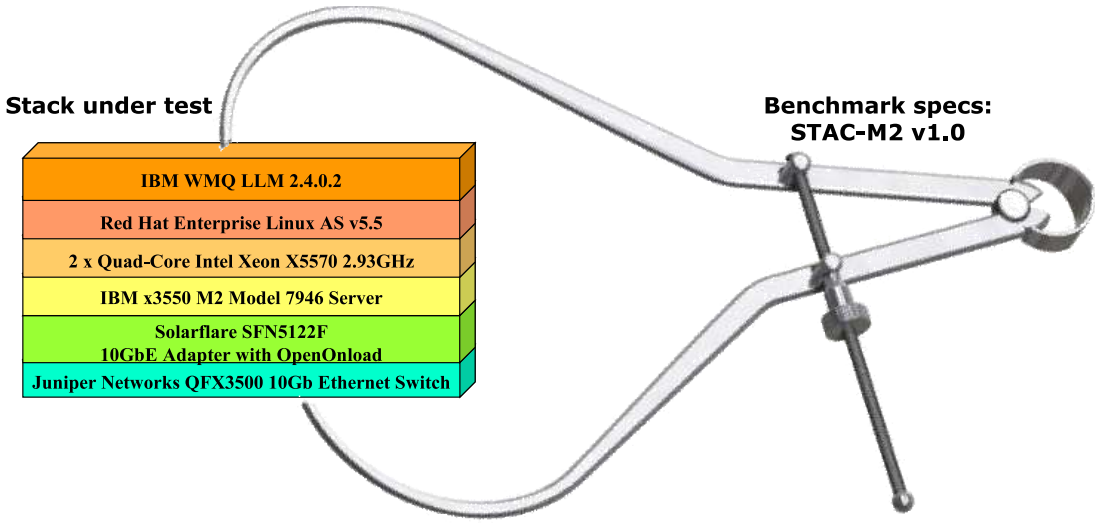
STAC Report™ Highlights

IBM WMQ LLM with IBM x3550 servers, Solarflare SFN5122F/Open Onload and Juniper Networks QFX3500 10GbE Switch
(SUT ID: LLM110203)

STAC-M2 Benchmarks™ v1.0

Tested by: STAC
Test date: 3 February 2011
Report 1.0, 23 February 2011



STAC BENCHMARK COUNCIL

THESE TESTS FOLLOWED STAC BENCHMARK SPECIFICATIONS APPROVED BY THE STAC BENCHMARK COUNCIL (SEE WWW.STACRESEARCH.COM). BE SURE TO CHECK THE VERSION OF ANY SPECIFICATION USED IN A REPORT. DIFFERENT VERSIONS MAY NOT YIELD RESULTS THAT CAN BE COMPARED TO ONE ANOTHER.

STAC® SECURITIES TECHNOLOGY ANALYSIS CENTER

Disclaimer

The Securities Technology Analysis Center, LLC (STAC®) prepared this report at the request of Juniper Networks. It is provided for your internal use only and may not be redistributed, retransmitted, or published in any form without the prior written consent of STAC. All trademarks in this document belong to their respective owners.

The test results contained in this report are made available for informational purposes only. STAC does not guarantee similar performance results. All information contained herein is provided on an "AS–IS" BASIS WITHOUT WARRANTY OF ANY KIND. STAC has made commercially reasonable efforts to adhere to published test procedures and otherwise ensure the accuracy of the contents of this document, but the document may contain errors. STAC explicitly disclaims any liability whatsoever for any errors or otherwise.

The evaluations described in this document were conducted under controlled laboratory conditions. Obtaining repeatable, measurable performance results requires a controlled environment with specific hardware, software, network, and configuration in an isolated system. Adjusting any single element may yield different results. Additionally, test results at the component level may not be indicative of system level performance, or vice versa. Each organization has unique requirements and therefore may find this information insufficient for its needs.

Customers interested in analyzing their own environment using the same methodology used in this report are encouraged to contact STAC or visit www.STACresearch.com/harnesses.

Contents

Summary

STAC ran STAC-M2 Benchmarks on a stack consisting of IBM's WebSphere MQ Low Latency Messaging (LLM) running on IBM xSeries® servers with Solarflare's 10GbE Ethernet adapter and Juniper Networks' QFX3500 switch (see Section 7 for system details). This report documents the results.

LLM is a messaging transport that IBM targets at the very high-volume, low-latency requirements of financial markets firms. The product is daemonless and provides peer-to-peer transport for one-to-many and many-to-many multicast messaging, as well as point-to-point unicast messaging. More product information is available in Section 2 of this report and at www.ibm.com/software/integration/wmq/llm.

The Solarflare SFN5122F is a low-latency, low-power 10GbE server adapter. Solarflare Server Adapters are designed to provide high performance in the most demanding application environments. Solarflare's OpenOnload application acceleration middleware was used in combination with the SFN5122F to enable full operating system bypass, which Solarflare reports dramatically reduces host processing overheads and enables high transaction rates while substantially reducing application latency with very low jitter. OpenOnload performs network processing at user-level and is binary compatible with existing applications that use TCP/UDP with BSD sockets.

Juniper Networks' Quantum Fabric 10GbE switch features 48 dual-mode small form-factor pluggable transceiver (SFP+/SFP) ports and four quad small form-factor pluggable plus (QSFP+) ports in a 1 U form factor, delivering feature rich Layer 2 and Layer 3 connectivity to networked devices such as rack servers, blade servers, storage systems, and other switches in highly demanding, high-performance data center environments. For converged server edge access environments, the QFX3500 is also a standards-based Fibre Channel over Ethernet (FCoE) Transit Switch and FCoE to Fibre Channel (FCoE-FC) Gateway, enabling customers to protect their investments in existing data center aggregation and Fibre Channel storage area network (SAN) infrastructures.

STAC-M2 tests the ability of messaging middleware to handle real-time market data under a variety of conditions. The tests provide key performance metrics such as latency, throughput, power efficiency, and CPU/memory consumption under several scenarios, including both undisturbed flow and exception conditions like slow consumers. Version 1.0 of STAC-M2 was approved by the STAC Benchmark Council on 11 September 2009. Section 1 includes both an overview of STAC-M2 and diagrams that summarize each test case used to produce the STAC Report Card. The diagrams have labels that correspond to the spec IDs in the Report Card. Understanding the tests in more detail requires access to the detailed STAC-M2 specifications, which are available to Contributor and Recipient members of the Council (membership is open to trading organizations and vendors; see www.STACresearch.com/council).

STAC tested LLM against all of the required STAC-M2 Benchmark specifications, except for those that required the stack under test to provide a last-known-value cache. IBM attested that LLM does not support this functionality, and this exception is allowed under STAC-M2. IBM says that its customers are able to use LLM as the base for custom-built market data systems that include caching functionality and that LLM is also integrated with the IBM WebSphere Front Office for Financial Markets product (www.ibm.com/software/integration/wfo), which supports caching, among other features. For this report, IBM chose not to involve products beyond LLM.

In all, the version 1.0 STAC-M2 specifications deliver hundreds of test results, which are presented through a variety of tables and visualizations in the full STAC Report™. Of these, Juniper, IBM and Solarflare wish to highlight the following:

Using standard Ethernet and UDP protocols, at the base message rates set by the specs, the mean latency of the solution did not exceed 10 microseconds, while standard deviation of latency was measured at less than 1 microsecond. At the highest tested rate of 2.3 million messages/second, the mean latency of the solution was just 14 microseconds.

Get the rest of the story!

These STAC Report Highlights provide some insight into the performance of the stack that was tested, but there's more to the story. The full STAC Report contains additional information, including:

1) Results from additional test cases:
   - Tests with three distinct types of slow consumers
   - Tests where one or a group of clients restarts during message flow
   - Tests with two Producers, not just one

2) Latency versus throughput for a range of message rates, including the highest rate tested.

3) Deeper latency analysis:
   - Full latency statistics for the highest message rate tested, not just the base rate
   - Latency percentiles all the way to six nines (99.9999th), not just 99th
   - Visualizations such as histograms and time plots of latency over test runs

4) Additional metrics (recovery times in restart cases, an additional latency metric that shows underlying transport speed, factoring out any latency due to queuing prior to send)

The full STAC Report is in the STAC Vault™. The Vault is only accessible to those firms with an appropriate membership in the STAC Benchmark Council. For more information, see www.STACresearch.com/vault.

The STAC Report Card summarizes results from several of the STAC-M2 test cases. Each benchmark has a unique identifier. If you are comparing these results to other STAC-M2 Benchmark results, make sure the identifiers match exactly. If they do not, they cannot be fairly compared. To obtain a copy of the STAC-M2 Test Harness to run the same STAC Benchmarks™on a system in your own lab, contact info@STACresearch.com.

STAC and the STAC-M2 working group have taken great pains to identify and describe the limitations inherent in the STAC-M2 methodology (see Section 6.3). Some tradeoffs were made in version 1.0 in the interest of getting a set of specifications to market that could begin the process of standards-based benchmarking of high-performance messaging systems. As the industry gains experience with STAC-M2 v1.0 and provides feedback, the STAC-M2 working group will consider additional specifications and perhaps modifications to the existing specifications that will improve a future version of STAC-M2. To participate in this process, please contact info@STACresearch.com.

STAC-M2 REPORT CARD

IBM WMQ Low Latency Messaging v2.4.0.2 (LLM110203)

Type	Spec ID	Description	VALUE	MEAN	MED	99P	MAX	STDV
Latency	STAC.M2.v1.0.BASELINE.LAT1	SupplyToReceive Latency (Hybrid) at base rate in the 1:5 setup with no watchlist overlap (µsec)		9	9	11	23	0
	STAC.M2.v1.0.OVERLAP.LAT1	SupplyToReceive Latency (Hybrid) at base rate in the 1:5 setup with some watchlist overlap (µsec)		9	9	12	23	0
	STAC.M2.v1.0.FLEXIBLE.LAT1	SupplyToReceive Latency (Hybrid) at base rate in the setup with flexible Consumer resources (µsec)		9	9	11	20	0
	STAC.M2.v1.0.CACHING.LAT1	SupplyToReceive Latency (Hybrid) at base rate in the 1:5 setup with caching enabled (µsec)		NS	NS	NS	NS	NS
Throughput	STAC.M2.v1.0.BASELINE.TPUT1	Highest supply rate tested in the 1:5 setup with no watchlist overlap (msg/sec)	1,500,000					
	STAC.M2.v1.0.OVERLAP.TPUT1	Highest supply rate tested in the 1:5 setup with some watchlist overlap (msg/sec)	1,600,000					
	STAC.M2.v1.0.FLEXIBLE.TPUT1	Highest supply rate tested in the setup with flexible Consumer resources (msg/sec)	1,500,000					
	STAC.M2.v1.0.CACHING.TPUT1	Highest supply rate tested in the 1:5 overlap setup with caching enabled (msg/sec)	NS					
CPU	STAC.M2.v1.0.SINGLE.CPU1	Consumer CPU utilization at base rate in the 1:1 setup (core equivalents)	Active Cores = 2	1.99			2.00	
	STAC.M2.v1.0.SINGLE.CPU2	Producer CPU utilization at base rate in the 1:1 setup (core equivalents)	Active Cores = 3	2.21			2.24	
	STAC.M2.v1.0.FIELDS1.CPU1	Consumer CPU utilization at base rate in the 1:1 setup with complete field consumption (core equivalents)	Active Cores = 2	1.99			2.00	
	STAC.M2.v1.0.TINYLISTS.CPU1	Consumer CPU utilization in the 1:5 setup with four tiny-watchlist Consumers at base rate (core equivalents)	Active Cores = 2	1.99			2.00	
Memory	STAC.M2.v1.0.SINGLE.MEM1	Max memory utilization of the Consumer machine at base rate in the 1:1 setup (MB)	331					
	STAC.M2.v1.0.SINGLE.MEM2	Max memory utilization of the Producer machine at base rate in the 1:1 setup (MB)	345					
	STAC.M2.v1.0.FIELDS1.MEM1	Max memory utilization of the Consumer machine at base rate in the 1:1 setup with complete field consumption (MB)	328					
Data-center Efficiency	STAC.M2.v1.0.FLEXIBLE.MPW	FLEXIBLE.TPUT1, divided by adjusted watts in the setup with flexible Consumer resources (msg/Watt)	505					
	STAC.M2.v1.0.FLEXIBLE.MPU1	FLEXIBLE.TPUT1, divided by rack units in the setup with flexible Consumer resources (msg/U)	75,000					
	STAC.M2.v1.0.FLEXIBLE.MPO	FLEXIBLE.TPUT1, divided by OS impressions in the setup with flexible Consumer resources (msg/impression)	93,750					

NOTE: "NS" means not supported. The vendor attests that the product configuration tested does not support the functionality necessary to produce this benchmark.

1. Background on the STAC-M2 Benchmark specifications

1.1 Overview

The STAC-M2 Benchmark specifications test the ability of a solution such as messaging middleware to handle real-time market data in a variety of configurations. The specs reflect the input of seven leading trading firms and six vendors of high-performance messaging. They provide key performance metrics such as throughput, latency, power efficiency, and CPU/memory consumption under several scenarios.

As shown in Figure 1, the software binaries required for a STAC-M2 test are Producer and Consumer applications written to interface to the "stack under test" (SUT). These incorporate the STAC-M2 Library to control behavior. The test specs make no assumption about the architecture of the SUT (brokerless vs broker-based, appliance vs software, Ethernet vs InfiniBand, etc.). The fundamental hardware requirement is a small number of machines to host the Producers and Consumers. Some specs are designed to enable arbitrarily large test harnesses for scale testing, but such testing is not required. STAC-M2 does not require any proprietary hardware for time sync or wire capture.

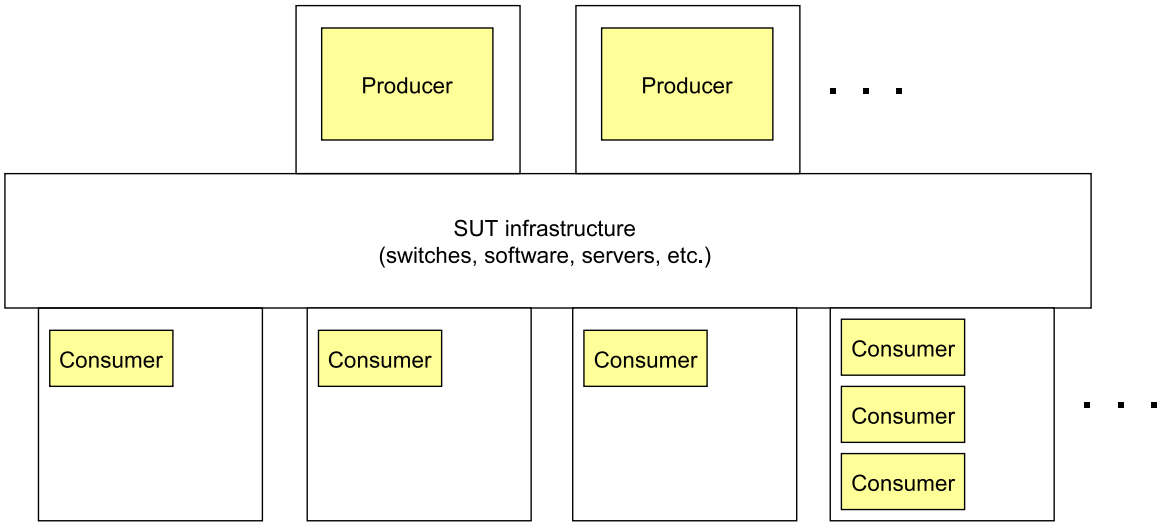


Figure 1 - High-level view of STAC-M2 test harness

The harness uses a "reflection" methodology for round-trip latency measurement, illustrated in Figure 2. A Producer transmits a "primary" message; a Consumer consumes the message and selectively republishes it as a "reflected" message; and the Producer consumes the reflected message. Multiple Consumers can receive primary messages from a given Producer, which in turn can receive reflected messages from multiple Consumers. Latency is measured from the earliest moment a primary message is available for sending in the Producer to the moment the reflected message payload is available for consumption in the proper format in the Producer (ReceiveTime minus SupplyTime in Figure 2). This result is divided by two to provide a "Hybrid Latency." Hybrid Latency is indicative of one-way latency but may not be exact, because the reflected message rate is always a fraction of the primary message rate.

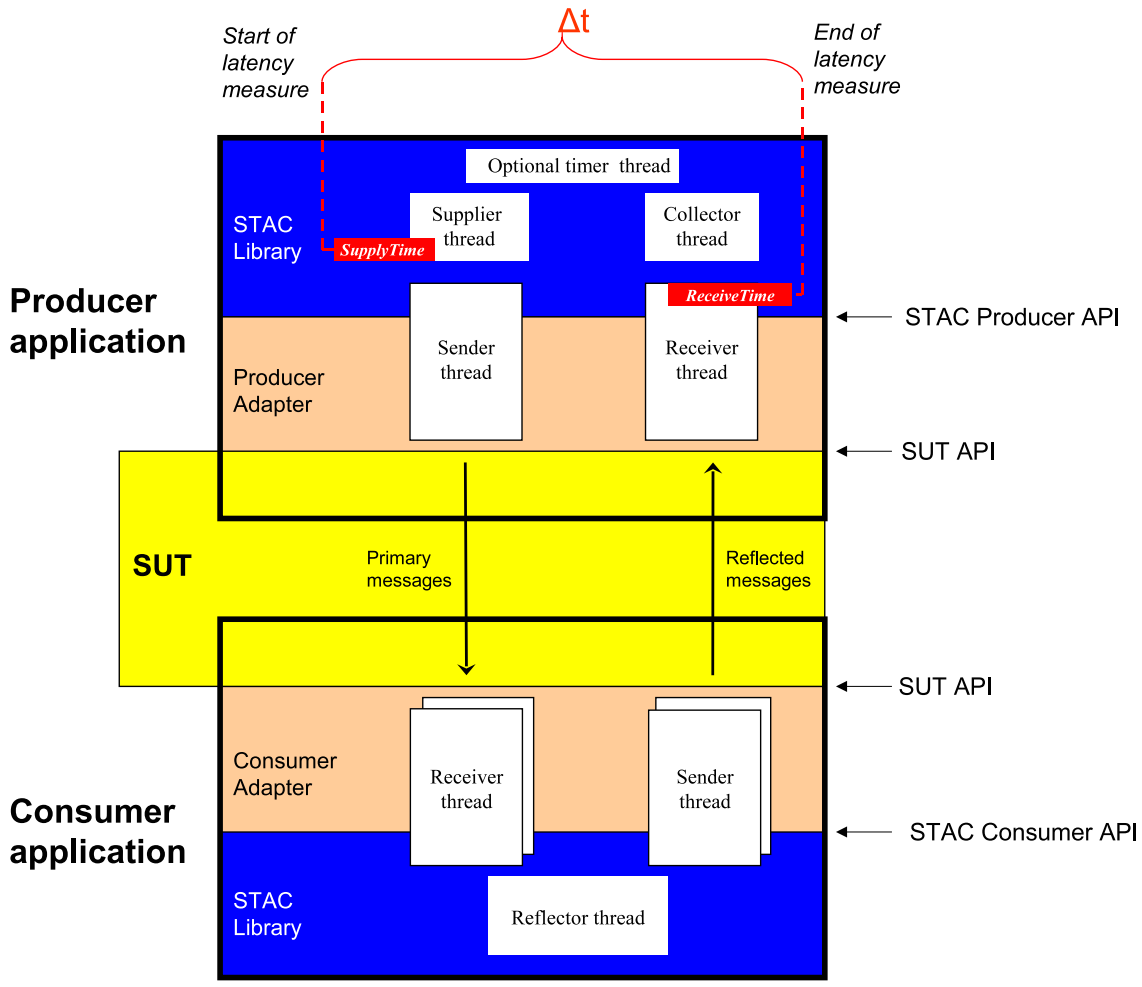


Figure 2 - Component-level view of STAC-M2 Producer/Consumer methodology

Version 1.0 of STAC-M2 draws from equities and options use cases: smart order routing, pairs trading, market-making, and black-box trading. Some specs emulate "latency minimizing" deployments, where application owners tend to over-provision resources to avoid contention. Other specs emulate "cost-minimizing" deployments, where application owners care more about total cost, meaning that they may load multiple apps onto a given machine.

The specs vary the number of Producers and Consumers, the watchlist sizes and commonality among Consumers, whether the SUT must cache, and whether it must deliver parsed or opaque messages. They test common exception conditions like slow Consumers and application restarts.

Version 1.0 supplies messages modeled on observed output of US equities order-book feed handlers deployed in the field. The STAC Library supplies messages to the Producer Adapter as pointers to C structs containing a 24- to 32-byte header and a 232-byte payload containing 136 bytes of data. The Producer Adapter can send the payload as an opaque buffer or use a SUT-specific encoding algorithm. Note that these sizes are purely the size in memory of data going into and out of the adapter. They say nothing about the size of a message as a SUT represents it internally or across a wire. Version 1.0 of STAC-M2 supplies messages at a steady-state rate. Future versions can include more message types, sizes, and timing patterns.

For more detail on STAC-M2, see www.STACresearch.com/m2. Most of the materials are restricted to Council members. To gain access or to provide input on future versions of STAC-M2, please contact council@STACresearch.

1.2 Test cases used to produce the STAC-M2 Report Card

Below are diagrams that summarize the test sequences (setups and scenarios) used to generate the results in the STAC Report Card (additional sequences were used to generate the full test results contained in the complete STAC Report). The specification IDs in the STAC Report Card embed the name of the relevant sequence below. The sequences are presented in the order in which they are referenced in the STAC Report Card.

Sequence BASELINE:

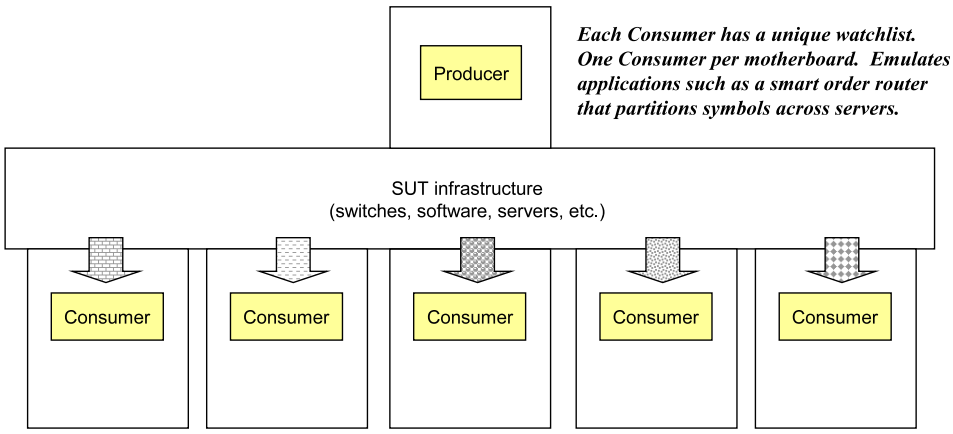


Figure 3

Sequence OVERLAP:

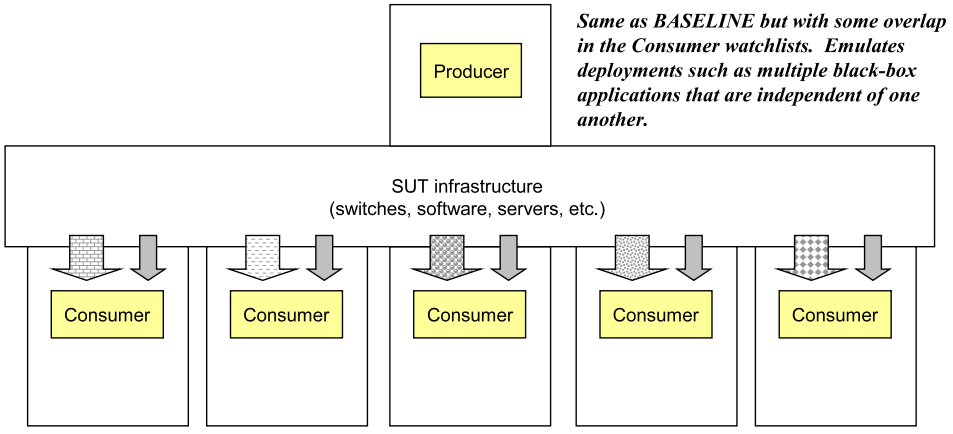


Figure 4

Sequence FLEXIBLE:

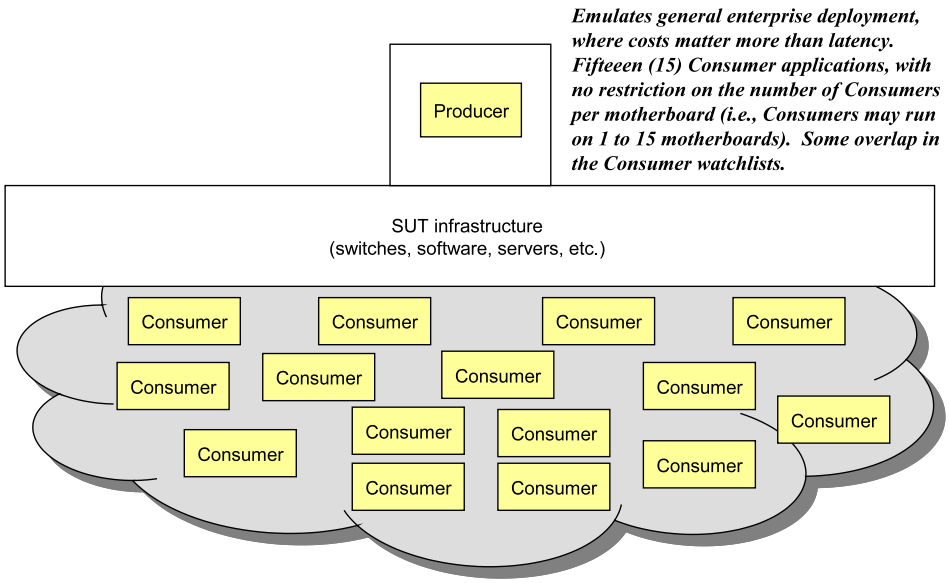


Figure 5

Sequence CACHING:

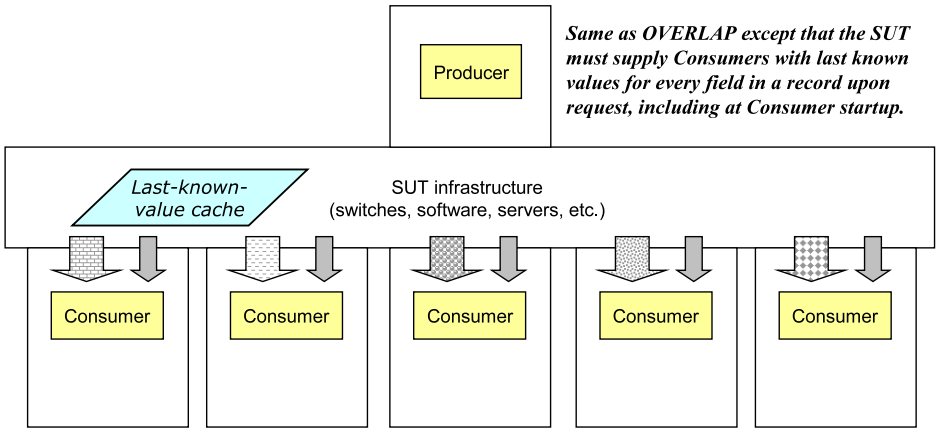


Figure 6

Sequence SINGLE:

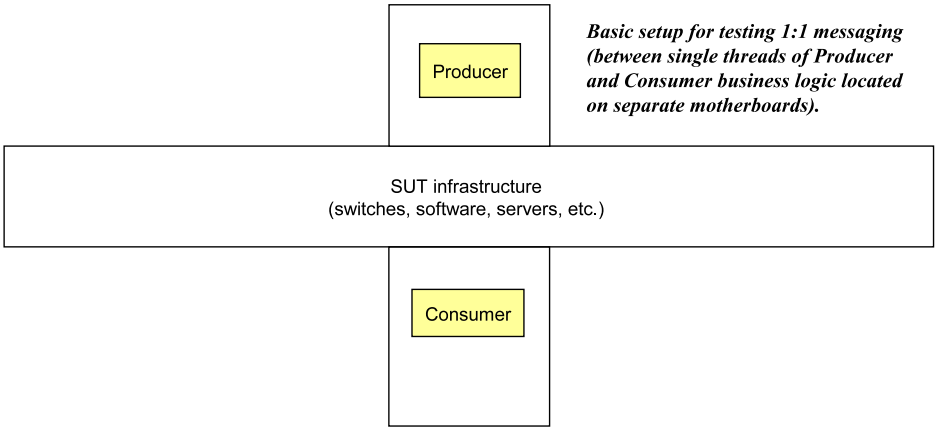


Figure 7

Sequence FIELDS1:

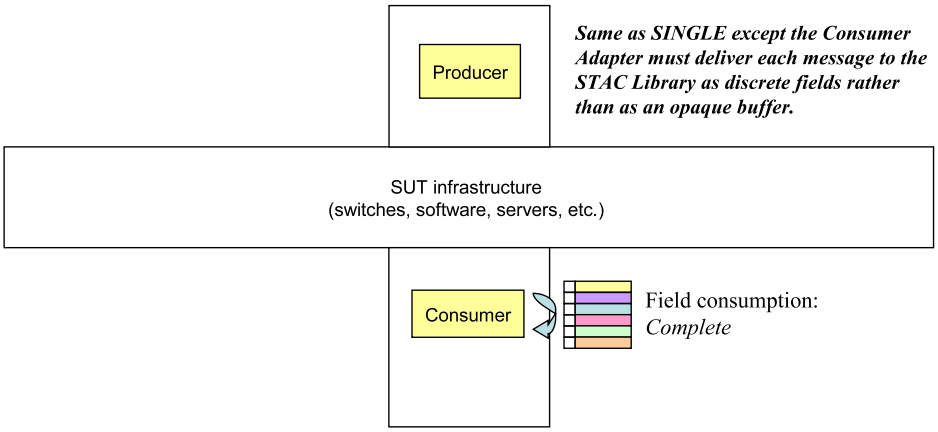


Figure 8

Sequence TINYLISTS:

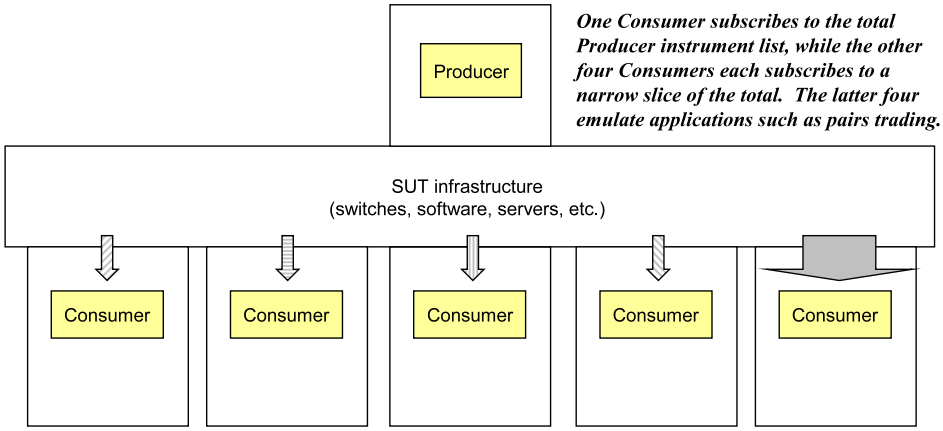


Figure 9

2. Product background

Juniper submitted the following information and claims about its products:

The Juniper Networks QFX3500 delivers feature rich Layer 2 and Layer 3 connectivity to networked devices such as rack servers, blade servers, storage systems, and other switches in highly demanding, high-performance data center environments.

Featuring 48 pluggable dual-mode SFP+/SFP ports and four QSFP+ ports in a 1U form factor, the QFX3500 switch delivers feature-rich Layer 2 and Layer 3 connectivity to networked devices such as rack servers, blade servers, storage systems and other switches in highly demanding, high-performance data center environments. For converged server-edge access environments, the QFX3500 is also a standards-based Fibre Channel over Ethernet (FCoE) Transit Switch and FCoE to Fibre Channel (FCoE-FC) Gateway, enabling customers to protect their investments in existing data center aggregation and Fibre Channel SAN infrastructures.

In this test scenario, a Juniper EX4500 switch was used to perform the function of IGMP Querier, simulating an upstream peering router. Complete Data Sheets and other technical information about the Juniper products are available at: www.juniper.net/us/en/products-services/

IBM submitted the following information and claims about its products:

IBM WebSphere MQ Low Latency Messaging (LLM) is a messaging transport that is highly optimized for the very high-volume, low-latency requirements of financial markets firms.

Applications for LLM include the high-speed delivery of market data, transactional data, reference data, and event data in or between front-, middle-, and back-office. This includes:

- *Market data from exchanges to data consumers.*
- *Market and reference data within the enterprise to analytic or trading applications.*
- *Trade data such as positions or orders to direct-market access and other trading applications.*
- *Event notifications for systems-monitoring, risk-analytics, and compliance applications.*
- *Gateway to matching engine high-availability connection with total ordering for stock exchange scenarios.*

The product is fully daemonless and provides peer-to-peer transport for one-to-one, one-to-many and many-to-many data exchange. In addition to this message-delivery flexibility and the performance demonstrated in this report, LLM's capabilities include:

- *Reliable message delivery with fine-grained control of message delivery assurance.*
- *High availability to maintain system service levels and to protect the integrity of the data stream when components fail.*
- *Message persistence at wire speeds for message recovery and auditing.*
- *Monitoring and congestion control to automatically detect bottlenecks and streamline data flow.*
- *High-speed message filtering that supports fine-grained data multiplexing and efficient data segmentation.*

NOTE: Not all of these features were necessarily enabled during the STAC-M2 tests.

To learn more about LLM's abilities or to evaluate LLM for a specific use case, IBM asks customers to visit www.ibm.com/software/integration/wmq/llm or to contact IBM at IBM_WFO_Solutions@us.ibm.com.

The tests were conducted on IBM's xSeries® , with the x3550 M2 servers, equipped with the Intel® Xeon® Processor X5570 series.

The xSeries Servers and WebSphere MQ Low Latency Messaging (LLM) offers compelling benefits for customers looking to increase performance of market data delivery, while reducing power consumption and management costs. For more information, please visit http://www-03.ibm.com/systems/x/.

Solarflare submitted the following information and claims about its products:

The Solarflare SFN5122F is a low-latency, low-power 10GbE server adapter. In combination with Solarflare OpenOnload application acceleration middleware, the Solarflare 10GbE adapter can achieve very low TCP/UDP application latency and dramatic increases in messages per second. By improving the host CPU efficiency, OpenOnload enables applications to leverage more server resources, resulting in dramatically accelerated application performance without any need to rewrite applications or change the existing Ethernet and TCP/IP infrastructure. OpenOnload performs network processing at user-level and is binary-compatible with existing applications that use TCP/UDP with BSD sockets. It comprises a user-level shared library that implements the protocol stack, and a supporting kernel module. For more information, see www.openonload.org and www.solarflare.com. To request Solarflare 10GbE server adapters for evaluation, contact sales@solarflare.com.

3. Project participants and responsibilities

The following firms participated in the project and should be considered "Project Participants" for the purpose of the non-disclosure agreement for information from the STAC Vault:

- IBM
- Juniper Networks
- Solarflare Communications
- STAC

The Project Participants had the following responsibilities:

- IBM and Solarflare implemented the STAC-M2 adapter using the STAC Library.
- Juniper, IBM and Solarflare configured and optimized the stack under test (SUT).
- IBM provided the servers and messaging middleware
- Juniper Networks provided the 10GbE switch
- Solarflare provided the 10GbE network adapters with OpenOnload

- Juniper sponsored the Audit.
- STAC conducted the STAC-M2 Benchmark Audit, which included inspecting the source code to the STAC-M2 Producer and Consumer, inspecting the STAC Test Harness configuration, validating required SUT functionality, and executing the tests.

4. Contacts

For more information on Juniper products, please visit www.juniper.net/us/en/contact-us/.

For questions about this system configuration or the test results, contact IBM_WFO_Solutions@us.ibm.com or visit www.ibm.com/software/integration/wmq/llm.

For more information on Solarflare products, contact btolley@solarflare.com.

For questions about STAC-M2, contact info@STACresearch.com.

5. Benchmark Status

- These test results were audited by STAC or a STAC-certified third party, as indicated in the Responsibilities section above. As such, they are considered official results. For details, please see www.STACresearch.com/reporting.
- The Vendor attests that the Producer and Consumer binaries provided for the audit were compiled from the same source code approved by STAC or a STAC-certified third party as part of this audit engagement.
- The Vendor attests that it did not modify the SUT during the audit.

6. Methodology

6.1 Specifications

This project followed the 1.0 version of the STAC-M2 Benchmark specifications. Full members of the STAC Benchmark Council can access these specifications at www.STACresearch.com/m2.

6.2 STAC Test Harness

STAC-M2 Library Version	v1.0.3
STAC Lib Timing Config	"system" (gettimeofday)

6.3 Limitations

- The latency metrics in version 1.0 STAC-M2 specifications are not true one-way latencies. They are "hybrid" values (round trip divided by 2), which should be interpreted as imperfect proxies for one-way latency. Several factors potentially limit the accuracy of these proxies: 1) since the reflected message rate is a small fraction of the primary message rate, the workload is not symmetric, which means latency might not be symmetric (for example, latency from Producer to Consumer might be higher than from Consumer to Producer); 2) while the specs call for the primary and reflection paths to have symmetric configurations, such symmetry may not be possible for some SUTs, which could cause asymmetries in latency.
- The STAC-M2 Test Harness relies on software instrumentation for measurement, which has an impact on performance. These limitations are consistent for all systems tested with this harness but are worth noting. In particular: 1) Because it is a reflection methodology, the latency measurements include some latency for the STAC-M2 Library to reflect messages (which is divided by 2 along with the rest of the round-trip latency); 2) reflection in the Consumer also takes CPU time, which increases the observed CPU utilization of the Consumer and may decrease throughput; 3) the Producer is required to call the STAC-M2 Library to record a "Send" timestamp, which consumes some CPU and may reduce the max sustainable rate.

- When deployed in a Consumer, the threading model of the STAC Library used for this version of the specifications emulates an application that requires business logic to execute in a different thread from message-consumption threads and requires serialized input to the business-logic thread from those other threads. This imposes a single-threaded bottleneck on performance of a single Consumer. Another common application pattern in the industry is for an application to scale across multiple cores by partitioning business logic by symbol and running multiple partition threads, each performing its own in-line message consumption. The latter pattern would potentially allow a single client to achieve higher throughput but is not supported in this version of the specs. Similarly, each Producer has a single Supplier thread, which limits the total possible message supply rate for a given Producer. Specifications that allow an arbitrary number of Suppliers on a given machine may be added in the future but are not part of the 1.0 specs.
- Power measurements in this version of STAC-M2 are allowed to be rudimentary. It is acceptable to take periodic amp readings and multiply by voltage supplied to the power distribution units (PDU). This method is known to have two sources of error: 1) accuracy of the meters in the PDUs, 2) lack of accounting for the power factor. More accurate power readings may be taken if available.
- The timing of Consumer coordination is only accurate to +/- one second. This could lead to some small variation in results from test run to test run.

7. Stack under test

7.1 Overview

The stack under test (SUT) consisted of 16 IBM x3550 M2 Model 7946 rack-mounted servers. Each server was interconnected using one port of a Solarflare 10 Gbps Ethernet Dual Port Adapter networked through a Juniper QFX3500 10GbE Switch. STAC-M2 Producer and Consumer applications used IBM's WebSphere MQ Low Latency Messaging (LLM) API for all communications and were configured to use different multicast channels for primary messages from Producer(s) to Consumer(s) than for reflected messages from Consumer(s) to Producer.

The STAC-M2 Adapter was dynamically linked with LLM, while LLM was in turn dynamically linked with Solarflare OpenOnload using the LD_PRELOAD mechanism. OpenOnload provided a standard POSIX socket API, allowing LLM and the STAC-M2 Adapter to work with a user-space kernel bypass protocol stack instead of the usual Linux kernel TCP/UDP stack. The LD_PRELOAD mechanism means that the standard shipping LLM binaries could be used with OpenOnload without any need for modification. The Juniper QFX3500 was configured with default Layer 2 configurations to support low-latency, cut-through switching.

A Juniper EX4500 Ethernet switch, connected via 10Gb Ethernet was used to provide the IGMP Querier functionality required for the Multicast messages sent during the STAC test.

For more detailed configuration information, contact IBM_WFO_Solutions@us.ibm.com.

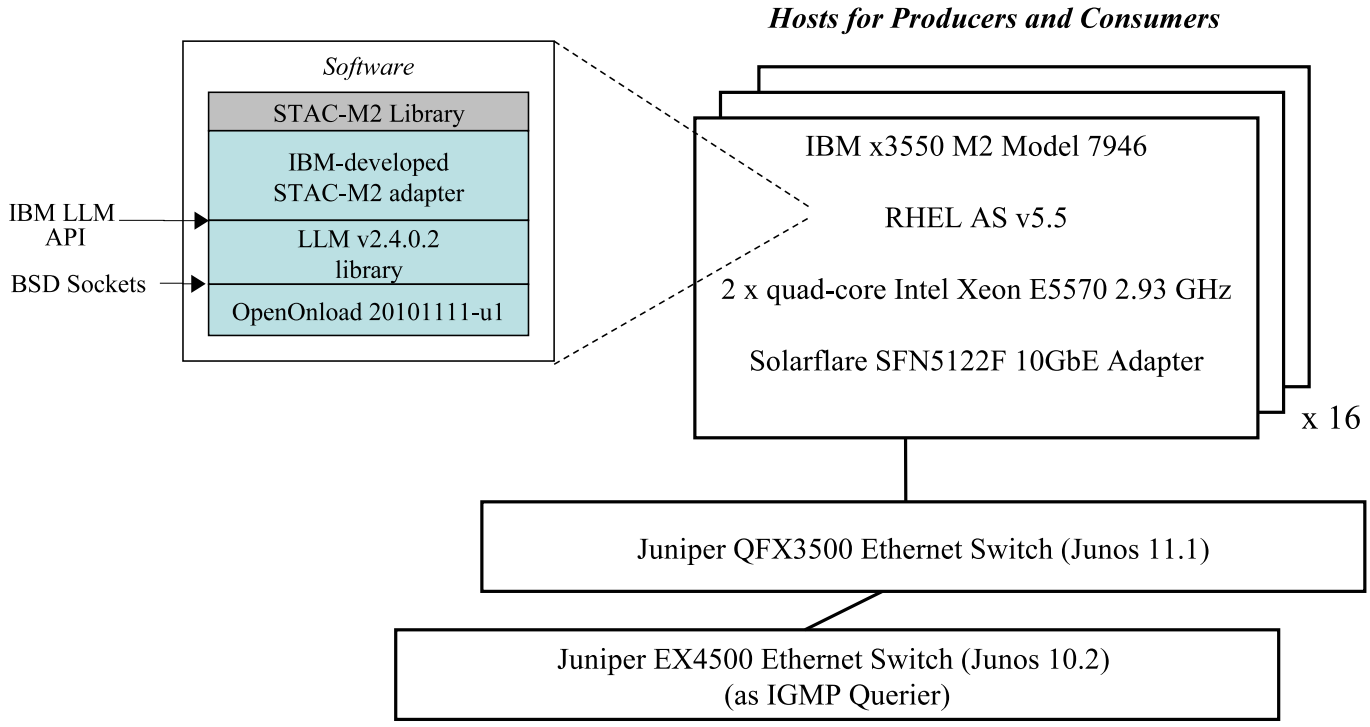


Figure 10 - Test configuration

7.2 Producer

7.2.1 Servers

Vendor Model	IBM System x3550 M2 Model 7946
Rack Units	1
Power Supply	675 W
Processor type	2 x Intel Xeon CPU X5570 @ 2.93 GHz
Cache	L1 I cache: 32K, L1 D cache: 32K, L2 cache: 256K, L3 cache: 8192K
Bus Speed	1333 MHz
Memory	48GB (6 x 8GB) DIMM
Disk	146GB
BIOS Vendor	IBM
BIOS Version	D6E150AUS-1.10
BIOS Release Date	Build id: D6E150AUS-1.10 (12/15/10)

7.2.2 Network interfaces

NIC 2 – port 2	Solarflare SFN5122F with OpenOnload
Used for	Message Traffic
NIC Driver Version	3.0.8.2221
NIC Firmware Version	3.0.8.2216

NIC 1 – port 1	Broadcom Corporation NetXtreme II BCM5709 Gigabit Ethernet (rev 20)
Used for	Test Harness Admin
NIC Driver Version	2.0.2
NIC Firmware Version	5.2.3 NCSI 2.0.10

Network Stack Configuration	Contact Vendor for details

7.2.3 Operating System

Version	Red Hat Enterprise Linux Server AS 5.5
Kernel Version	2.6.18-194.el5 (x86_64)

7.3 Consumer

7.3.1 Servers

Vendor Model	IBM System x3550 M2 Model 7946
Rack Units	1
Power Supply	675 W
Processor type	2 x Intel Xeon CPU X5570 @ 2.93 GHz
Cache	L1 I cache: 32K, L1 D cache: 32K, L2 cache: 256K, L3 cache: 8192K
Bus Speed	1333 MHz
Memory	48GB (6 x 8GB) DIMM
Disk	146GB
BIOS Vendor	IBM
BIOS Version	D6E150AUS-1.10
BIOS Release Date	Build id: D6E150AUS-1.10 (12/15/10)
Hyper-Threading	Enabled

7.3.2 Network interfaces

NIC 2 – port 2	Solarflare SFN5122F with OpenOnload
Used for	Message Traffic
NIC Driver Version	3.0.8.2221
NIC Firmware Version	3.0.8.2216

NIC 1 – port 1	Broadcom Corporation NetXtreme II BCM5709 Gigabit Ethernet (rev 20)
Used for	Test Harness Admin
NIC Driver Version	2.0.2
NIC Firmware Version	5.2.3 NCSI 2.0.10

Network Stack Configuration	Contact Vendor for details

7.3.3 Operating System

Version	Red Hat Enterprise Linux Server AS 5.5
Kernel Version	2.6.18-194.el5 (x86_64)

7.4 Interconnects

Switch	Juniper Networks QFX3500 10GbE Converged Ethernet Switch, Junos 11.1
Switch (for IGMP)	Juniper Networks EX4500 10GbE Converged Ethernet Switch, Junos 10.2

7.5 Application Software

Product/component	WebSphere MQ Low Latency Messaging
Version	2.4.0.2
Build ID	20101014.1

Product/component	STAC Pack
Author	IBM
Interface	IBM MQ LLM API v2.4.0.2
Source version	v2.4.0.2
Producer build ID	v 1.4 :: Feb 2 2011 - 18:33:19
Consumer build ID	v 1.4 :: Feb 2 2011 - 18:33:16

7.6 Configuration for Slow Consumers

STAC-M2 requires all Consumers to be configured the same. In an Audit, the vendor does not know which Consumers STAC will configure to exhibit slow consumption in a given test run. LLM allows for the specification of slow-consumer policies that can be used to automatically suspend or expel receivers that are not performing well. According to IBM, the LLM policies in these tests applied thresholds to the number of NAKs (negative acknowledgements) generated by a Consumer and expelled slow Consumers from the system.

8. Vendor Commentary

IBM, Juniper, and Solarflare provided the following comments as a joint statement:

Readers of the report should note the very, very low standard deviation of latency (jitter) that was observed. Various elements of the solution stack and its configuration contribute to the low jitter.

The server systems were tuned such that all interrupts and background processing tasks were run on a single dedicated CPU core.

The Solarflare OpenOnload library was configured to enable the application threads to run without contention with each other, and without requiring any interrupt processing or other interaction with the kernel active on the dedicated core.

These techniques implemented in Solarflare's OpenOnload enabled the system to be entirely vertically separated on CPU cores, thus avoiding any application-level crosstalk. As a result, extremely low jitter was observed in the results. The very low latency of the Solarflare SFN5112F Server adapter when used in combination with OpenOnload results in a low degree of buffering in the adapter and protocol stack, which also makes a significant net reduction in overall application-level latency.

The Juniper Networks QFX3500 switch used in this test has shown extremely low latency and jitter when configured in optional cut-though mode. The principal reason for the consistency is that the packet-processing path has been reduced to a minimal number of ASICs, which also contributes to the very low power draw observed in the test. We would also like to highlight that the QFX3500 runs the Junos Operating System (11.1 and higher), easing integration for the full portfolio of Junos based switching, routing and security solutions.

IBM invites anyone interested in more information about the products or specific configurations of this solution to contact them at: IBM_WFO_SOLUTIONS@us.ibm.com.

9. STAC Notes

None

About STAC

STAC is a technology-research firm that facilitates the STAC Benchmark Council™ (www.STACresearch.com/council), an organization of leading trading organizations and vendors that specifies standard ways to measure the performance of trading solutions. STAC Benchmarks cover workloads in market data, analytics, and trade execution.

STAC helps end-user firms relate the performance of new technologies to that of their existing systems by supplying them with STAC Benchmark reports as well as standards-based STAC Test Harnesses™ for rapid execution of STAC Benchmarks in their own labs. End users do not disclose their results. Some STAC Benchmark results from vendor-driven projects are made available to the public, while those in the STAC Vault™ are reserved for members of the Council (see www.STACresearch.com/vault).

To be notified when new STAC Benchmark results become available, please sign up for free at www.STACresearch.com.