Thomson Reuters Enterprise Platform

ADS Standard Mode

Performance Test Results on IBM BladeCenter HS22V

DISCLAIMER: The test results and recommendations contained in this report are made available for informational purposes only. By issuing this report, IBM does not guarantee similar performance results. All information contained herein is provided on an “AS-IS” BASIS WITHOUT WARRANTY OF ANY KIND. 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. IBM is pleased to provide this report for our customers. We appreciate that each organization has unique requirements, and therefore may find this information insufficient for its needs. Customers wishing to obtain custom analysis for their systems are encouraged to contact their local IBM representative.

Issue 1.1

Date of Issue: Aug 9th 2011

Contents

1 General Information ... 3

1.1 Objective ... 3

1.2 Testing Methodology ... 3

1.3 Software Versions ... 3

1.4 Hardware ... 4

2 Preparation for Performance Test ... 5

2.1 Network ... 5

2.2 Hardware ... 5

2.3 Operating System Configuration ... 5

2.4 TREP Configuration ... 5

2.5 Miscellaneous Notes ... 5

3 Detailed Results ... 7

3.1 RSSL/RWF Update Throughput ... 7

3.2 End-to-End RSSL/RWF Latency ... 7

4 HS22V uEFI Parameter Settings ... 8

© IBM Corporation 2011. All Rights Reserved.

IBM, by publishing this document, does not guarantee that any information contained herein is and will remain accurate or that use of the information will ensure correct and faultless operation of the relevant service or equipment. IBM, its agents and employees, shall not be held liable to or through any user for any loss or damage whatsoever resulting from reliance on the information contained herein.

This document contains information proprietary to IBM and may not be reproduced, disclosed, or used in whole or part without the express written permission of IBM.

Any Software, including but not limited to, the code, screen, structure, sequence, and organization thereof, and Documentation are protected by national copyright laws and international treaty provisions. This manual is subject to U.S. and other national export regulations.

1 General Information

1.1 Objective

The objective of this document is to report the performance test results for standard mode ADS 2.1.0.L4, for a particular hardware and software platform. The test procedures were conducted as described in *Reuters Advanced Data Hub and Advanced Distribution Server Performance Test Procedures and Results* document.

The goal of these tests is to measure throughput and latency through Thomson Reuters Enterprise Platform (TREP) infrastructure components, specifically the ADS. The tests are grouped into two categories:

- Update throughput using RSSL/RWF data (see 3.1)
- End-to-end RSSL/RWF latency using embedded timestamp (see 3.2)

1.1.1 Results Summary

- **End-to-end latency test:** 500 thousand updates per second with a mean latency less than 180 microseconds and 1.6 million updates per second with a mean latency of less than 1 millisecond
- **ADS (no fan-out) throughput:** 1.80 million updates per second
- **ADS (producer 50/50) throughput:** 4.7 million updates per second

1.2 Testing Methodology

For throughput testing, the ***sink_driven_src*** utility was used to generate update traffic, and the ***rmdstestclient*** utility was used to consume the updates. Level 1 data was used, with a Marketfeed (MF) update size of 140 bytes, and an equivalent Reuters Wire Format (RWF) update size of 74 bytes. Tests with no fan-out of updates used a 100,000 item watchlist. The infrastructure is tuned for maximum throughput, and the update rate was increased until the CPU limit was reached with no errors reported.

The embedded timestamp approach was used to calculate end-to-end latency for Level 1 (Quotes and Trades) data. ADH/ADS end-to-end update latency is measured by using ***sink_driven_src*** as the publisher and ***rmdstestclient*** as the subscriber.

In the embedded timestamp approach, the publisher embeds timestamps into selected updates which the subscriber uses for latency calculations. In this scenario, the publisher and subscriber must be running on the same node for accurate timestamps.

1.3 Software Versions

1.3.1 TREP Components

ADH ver. 2.1..0.L4
ADS ver. 2.1.0.L4
rrcp as included in ADH ver 2.1.0.L4

1.3.2 TREP Test Tool

sink_driven_src (from ADH load above)
rmdstestclient (from ADS load above)

1.3.3 Operating Systems

- Red Hat Enterprise Linux 5.5 64-bit, Linux kernel 2.6.18-194.el5

1.4 Hardware

The performance tests were performed on a single IBM BladeCenter HS22V with the following components:

1.4.1 Compute nodes

Three IBM Blade Server HS22V (7870) blades.
Each has:
Two QC Intel Xeon X5680 processors (3.3 GHz);
Six 4GB 1333 MHz DDR3 Registered DIMMs;
One 50GB SSD;
Two integrated Broadcom 1GbE controllers;
One 2-port Mellanox CNX2 10GbE Mezzanine Expansion Card (Cffh);

1.4.2 Blade chassis

One BC-H (8852) which contains
One Advanced Management Module;
Two Power Modules;
One Cisco 3012 1GbE Switch Module (I/O Bay 1);
Two BNT Virtual Fabric 10GbE Switch Modules (I/O Bays 7 & 9);
Three HS22V blades (as mentioned above) in Blade Bays 1 to 3.

1.4.3 Network

Each blade is on three networks:
One 1-Gigabit Ethernet (GbE) network solely for management purposes and
Two 10-Gigabit Ethernet (10GbE) networks for low latency and high throughput TREP communications.

Port 1 of the built-in dual-port1GbE is connected to Cisco 1GbE switch module.
Port 2 of the built-in dual-port1GbE is not used.
Port 1 of the dual-port Mellanox 10GbE card is connected to BNT 10GbE switch module.
Port 2 of the dual-port Mellanox 10GbE card is connected to another BNT 10GbE switch module.

2 Preparation for Performance Test

2.1 Network

All the performance tests were run where the machines were connected to a private network via 10 Gbps switches. All the network cards and switch ports were set to Auto Negotiate.

2.2 Hardware

All ADS components were run on the same machine class.

2.3 Operating System Configuration

Earlier tests have shown that the value chosen for ticks per second (tps) on the test application machine has a significant impact on latency measurement. Accordingly, a tps value of 1000 was used in these tests.

2.3.1 TCP and UDP Buffers

Any settings changed from the defaults are noted below:

Step	Procedure		
1	**OS**	**Enter the following lines in system file noted**	**System File**
	Linux	net.core.wmem_max = 1677216	***/etc/sysctl.conf***
		net.core.wmem_default = 1677216	
		net.core.optmem_max = 16777216	
		net.ipv4.tcp_timestamps = 0 net.ipv4.tcp_sack = 0 net.ipv4.tcp_low_latency = 1 (0 for throughput tests)	
		net.core.rmem_max = 8388608	
		net.ipv4.tcp_dma_copybreak = 262144	
		net.core.rmem_default = 8388608	
		net.ipv4.tcp_rmem = 8192 8388608 16777216	
		net.ipv4.tcp_wmem = 8192 8388608 16777216	
		net.ipv4.tcp_mem = 16777216 16777216 16777216	
		net.ipv4.udp_rmem_min = 16384 net.ipv4.udp_wmem_min = 16384 net.ipv4.udp_mem = 8388608 12582912 16777216	

2.4 TREP Configuration

The configuration template ***rmds.cnf.template*** was customized for the tests.

Config File	Description	Path
rmds.cnf.template	Configuration file	***./config***

2.5 Miscellaneous Notes

Any other significant deviations from the standard test procedures, or clarifications, are noted below (such as number/type of machines used, CPU binding policy, etc.):

Test	Deviation	Comments
All	CPU Binding	NIC interrupts were handled by core 1 The *rrcpd* daemon were bound to cores 2 and 3. ADH threads were bound to cores 4 and 5. ADS processes (running on a separate blade) were bound to cores 4 and 5. The s*ink_driven_src* process was bound to core 4 or core 6. *rmdstestclient* (on a separate blade) was bound to core 5.

3 Detailed Results

3.1 RSSL/RWF Update Throughput

- All the throughput numbers quoted here are for Level 1 data.
- The data file used in these tests has 1 update, with an update (data, not including header) size of 74 bytes in RWF.
- All of the tests with no fan-out used 100,000 item watchlist.
- For all tests the individual processes were bound to particular cores.
- ***sink_driven_src*** and ***rmdstestclient*** were used as the publisher and consumer of data
- ***sink_driven_src*** and ***ADH*** were started on blade 1. ***ADS*** was on blade 2. and ***rmdstestclient*** was on blade 3.

3.1.1 ADS/LAN

Configuration Option	Mounts : Commonality	Transport	Max Throughput	Comments
Cache Disabled	No fan-out	RRCP	1.80 Million updates per second	A single instance ADS and ADH each were used
Cache Disabled	100 mounts Producer 50/50	RRCP	4.7 Million updates per second	A single instance of ADS and ADH were used

3.2 End-to-End RSSL/RWF Latency

Latency is defined as the time for a data item to propagate through one or more TREP components. "End to end" latency is defined as the delta between the time an update is posted by the publisher application to its API and the time the same update is received by the consuming application from its API, i.e. it includes both the latency contribution from the API and the core infrastructure components.

NOTES:

- Caching was disabled in both the ADH and the ADS during these tests.
- Optimized binaries of the TREP infrastructure components were used.
- NTP was disabled on the tools node, as any drifts in time will affect the reported latency.
- Tests were run with 100,000 item watch list and RWF data update size of 74 bytes [Data file (***sample.xml***) was used]. The update size is equivalent to a 140-byte IDN update.
- Latency tests were run at each update rate for at least 5 minutes, up to the maximum sustainable update rate for a given setup.
- Decode of data was turned on in these tests.

3.2.1 RRCP Backbone Results

Update Rate [74-byte RWF messages]	Mean Latency (microsecs)	Std Deviation (microsecs)	Maximum Latency (microsecs)	Minimum Latency (microsecs)	Number of Latency Points
50000	105.37	10.80	258	86	3070
100000	113.75	22.37	482	82	3070
200000	131.77	27.81	580	85	3070

300000	147.25	36.64	559	85	3070
400000	163.43	46.65	700	84	3070
500000	178.70	52.65	846	76	3070
600000	208.01	103.45	1470	90	3070
700000	226.28	94.12	1121	88	3070
800000	249.72	129.20	1508	92	3060
900000	257.03	129.39	1575	69	3060
1000000	299.89	189.34	1933	78	3060
1100000	340.58	265.53	2534	90	3070
1200000	367.41	256.27	2337	94	3070
1300000	480.94	741.66	8638	91	3070
1400000	541.45	501.67	3989	92	3070
1500000	598.61	524.34	4864	91	3075
1600000	724.12	587.74	3712	93	3065
1650000	1063.23	1005.02	5033	96	3040

4 HS22V uEFI Parameter Settings

uEFI.OperatingMode	"Performance Mode”
uEFI.ProcessorEistEnable	Disable
uEFI.ProcessorCcxEnable	Disable
uEFI.PackageCState	"ACPI C2”
uEFI.ProcessorC1eEnable	Disable
uEFI.ProcessorHyperThreading	Disable
uEFI.ExecuteDisableBit	Disable
uEFI.AesEnable	Enable
uEFI.ProcessorVmxEnable	Disable
uEFI.IPStreamerPrefetcher	Enable
uEFI.HardwarePrefetcher	Enable
uEFI.AdjacentCacheLinePrefetch	Enable
uEFI.DCUStreamerPrefetcher	Enable
uEFI.DataReuseDisable	Enable
uEFI.QPISpeed	"Max Performance"
IMM.ThermalModePolicy	Performance
uEFI.DdrSpeed	"Max Performance"
uEFI.SocketInterleave	Numa
IMM.LanOverUSB	Disable
uEFI.PatrolScrub	Enable