Power Systems

SMC-R on AIX for SAP

Experience report

Jutta Land
IBM Deutschland Research & Development GmbH
Created on January 14th, 2019 – Version 1.0

Note: Before using this information and the product it supports, be sure to read the general information under ”Copyrights and Trademarks” on page 18 as well as “Disclaimer and Special Notices” on page 19.

Edition Notice and Version Information

IBM Corporation
Systems and Technology Group
Route 100
Somers, NY 10589
Edition Notice: This is version 1.0 of this document.
Focus: SMC-R on AIX for SAP
Target: Experience report

Preface

Abstract

Running your SAP business requires optimal performance.
There are multiple areas of an SAP landscape, like CPU, storage, and networking which have significant impact on your business performance.

With IBM® AIX® 7.2 TL 2, the AIX® operating system supports the new communication protocol solution Shared Memory Communications over Remote Direct Memory Access (SMC-R).

If, with SMC-R your network area would improve then this improves your overall business performance and results in a better revenue.

This white paper summarizes the results of SAP test scenarios that demonstrate the benefits that SMC-R has for SAP workloads in a distributed SAP environment. The usage of SMC-R improves network latency, network throughput, and leads to a reduction of CPU usage.
Based on this test results IBM recommends exploiting SMC-R to optimize network performance in an SAP environment. See: **SAP Note 2699513** - Recommendation of new AIX Feature SMC-R for SAP on AIX.

Scope of this document

This white paper is intended for customers running SAP on AIX® POWER®. It describes the performed verification and the experiences with SMC-R in an SAP on AIX® POWER® development environment.

The most recent document version can be downloaded from IBM TechDocs:
http://www-03.ibm.com/support/techdocs/atsmastr.nsf/Web/WhitePapers

Note: This document does not explain operations and technical concepts of AIX, SMC-R, SAP or network. Please refer to the AIX or SAP product documentation for more information.

Author

Jutta Land is working with IBM Germany, Research and Development Lab, as a member of the SAP on IBM Power (AIX, Linux) development team.

Document change history

Version	Date	Changes
1.0	14 Jan 2019	Initial Publication

Table of Contents

Table of figures

Introduction

Good performance in an SAP application requires optimizations in different areas, one part of an SAP 3-tiers (Database, Application Server and Client) infrastructures is the network.

With IBM® AIX® 7.2 TL 2, the AIX® operating system supports the new communication protocol solution Shared Memory Communications over Remote Direct Memory Access (SMC-R).

In the AIX documentation you can find the concept, benefits, and more on the SMC-R topic:
https://www.ibm.com/support/knowledgecenter/en/ssw_aix_72/com.ibm.aix.rdma/smc_r.htm

Technology

SMC-R is a new communication protocol based on sockets over RDMA (Remote Direct Memory Access) communications. RDMA enables a host to register memory and allow a remote host to directly read / write into this memory. It is confined to applications using Transmission Control Protocol (TCP) sockets over IPv4 or IPv6.

Benefit

With SMC-R, an improvement of the performance of transactional workloads via network like SAP workloads is expected by saving CPU resources, increasing throughput, and reducing latency.

Verification of SMC-R on AIX for SAP

The following chapters describe the verification of the SMC-R feature in a 3-tier SAP environment (Database, Application Server and Client). They summarize the verification test results, interpret them and leads to a recommendation.

Possible tools for the verification

You can use typical SAP workloads like the SAP SD benchmark, niping, enqt, or SGEN to verify the improvements of SMC-R in network latency, throughput and CPU utilization:

- The SAP Standard Application Benchmarks and in particular the SD – SAP Sales and Distribution Benchmark can be used to test the hardware and database performance of SAP applications and components.

- SAP niping – network interface ping - is a tool delivered by SAP for generating network traffic to measure network throughput and latency under real SAP conditions using SAP's communication layer.

- SAP enqt is a tool which can measure performance of the SAP lock mechanism that is implemented by the SAP Enqueue Server.

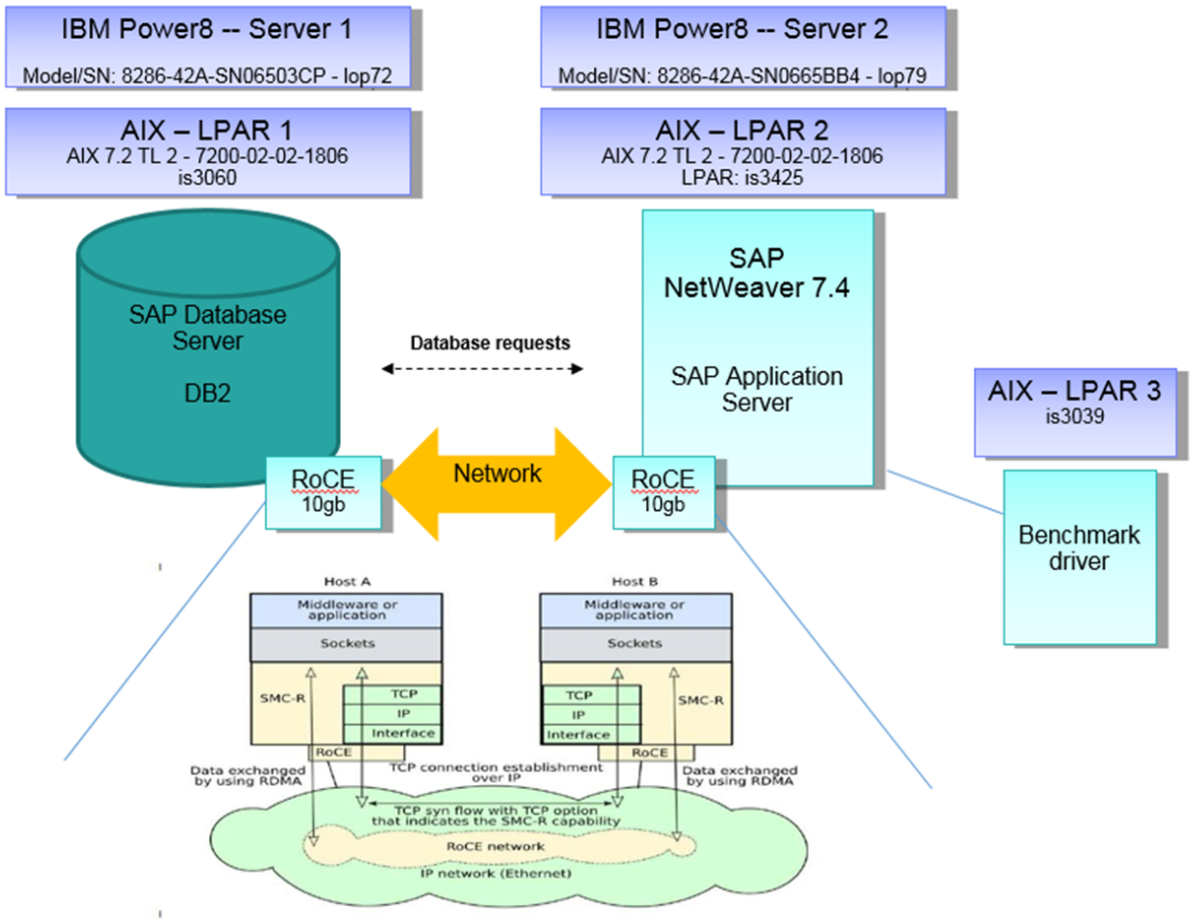
- SGEN is an SAP transaction used to compile ABAP repository objects. You can use SGEN to demonstrate SMC-R benefits like saving CPU.

Verification environment

Environment setup

- Two POWER 8 systems.
- Two dedicated LPARs, each with one CPU (smt=4), 50 GB RAM, and the required AIX level (7.2 TL2 7200-02-1806).
- Two RoCE cards with 10 Gigabit Ethernet (EC38 PCIe3 2-port 10 GbE NIC & RoCE SFP+ Copper Adapter).
- A distributed (3-tier) SAP NetWeaver 7.4 installation (including SAP SD Benchmark).
    - LPAR 1: IBM® DB2® Version 10.1 for Linux, UNIX, and Windows database
    - LPAR 2: SAP ASCS and DI instances
    - LPAR 3: SAP SD benchmark driver.

Figure 1 - Verification environment



SMC-R Configuration

In the AIX SMC-R documentation, the following 4 attributes are described, which need to be specified if you want to exploit SMC-R:

- enabled Specify if SMC-R is enabled or disabled
- ip_addr_list Specify allowed IP Address List
- port_range Specify allowed TCP Port Range
- max_memory Max Memory in MB

The ‘enabled’ attribute provides an easy way to enable (1) or disable (0) the SMC-R feature.
The ip_addr_list and port_range attributes are required. Only if they are specified the SMC-R protocol solution is used. You can find the relevant SAP ports in: https://help.sap.com/viewer/ports .
To keep it simple, you can run the complete communication traffic over SMC-R by specifying 'any' for ip_addr_list and port range. You can find documentation about SMC-R parameters and their values in the official IBM documentation and in the following blog:

https://www.ibm.com/developerworks/community/wikis/home?lang=en#!/wiki/Power%20Systems/page/Shared%20Memory%20Communications%20over%20RDMA%20(SMC-R)

SAP niping tool results

You can use the SAP niping tool to measure latency and throughput of the network between any two machines running SAP software. You can find details in the following SAP note:
https://launchpad.support.sap.com/#/notes/500235

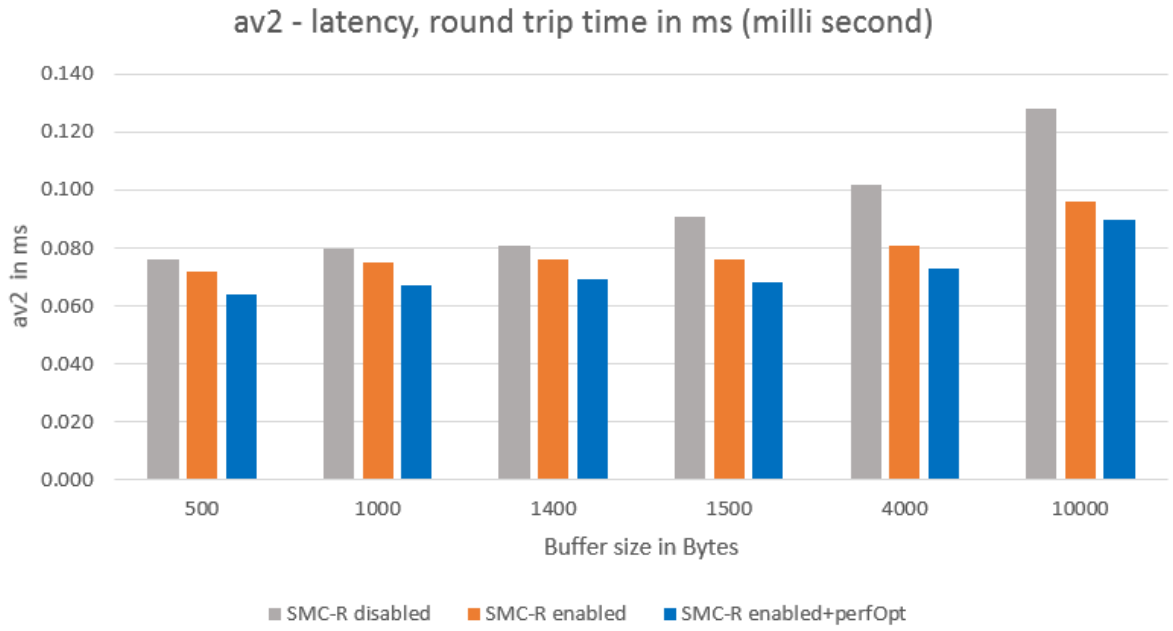
SAP niping has a server and client part. On one LPAR, the niping server is started and on the other LPAR, the niping client sends a specified number (loop) of data packets with a specified buffer size. When the client terminates, statistical data about latency and throughput are provided. The niping command of the client-side syntax is:

niping -c -H <server hostname> -S <port> -B <buffersize> -L <loop>

Network latency

The following figure shows the differences in network latency.
The latency is shown in milliseconds for different buffer sizes which are given in bytes. The grey bars show the latency when the SMC-R feature is disabled, the orange ones when it is enabled and the blue ones after additional performance optimizations have been applied (enabled and performance optimized). Details about the used attributes values are described for each scenario.

Figure 2 - SAP niping network latency



Depending on the buffer size we observe a total improvement from **15-29%** of the latency between the grey bars when the SMC-R is disabled, and the blue bars when SMC-R is enabled with additional performance optimizations. Results are shown in table 1.

Table 1 - SAP niping network latency total improvements

av2 - latency, round trip time in ms (milli second)				
Buffersize	SMC-R disabled	SMC-R enabled	SMC-R enabled+perfOpt	Total improvements in %
500	0.076	0.072	0.064	15.79
1000	0.080	0.075	0.067	16.25
1400	0.081	0.076	0.069	14.81
1500	0.091	0.076	0.068	25.27
4000	0.102	0.081	0.073	28.43
10000	0.128	0.096	0.090	29.69

These results were achieved in the mentioned environment setup with the following SMC-R attributes (SMC-R enabled – orange bars):

conns_per_lg	16	Number of connections per link group
enabled	1	SMCR Enabled
init_snd_pools	2	Number of send buffer pools to allocate quickly
ip_addr_list	any	IP Address List
max_memory	512	Max Memory in MB
port_range	any	TCP Port Range
rx_intr_packets	128	Number of packets to process in interrupt context
tx_intr_cnt	128	Tx Interrupt event coalesce counter
tx_intr_time	10000	Tx Interrupt event coalesce timer (microseconds)

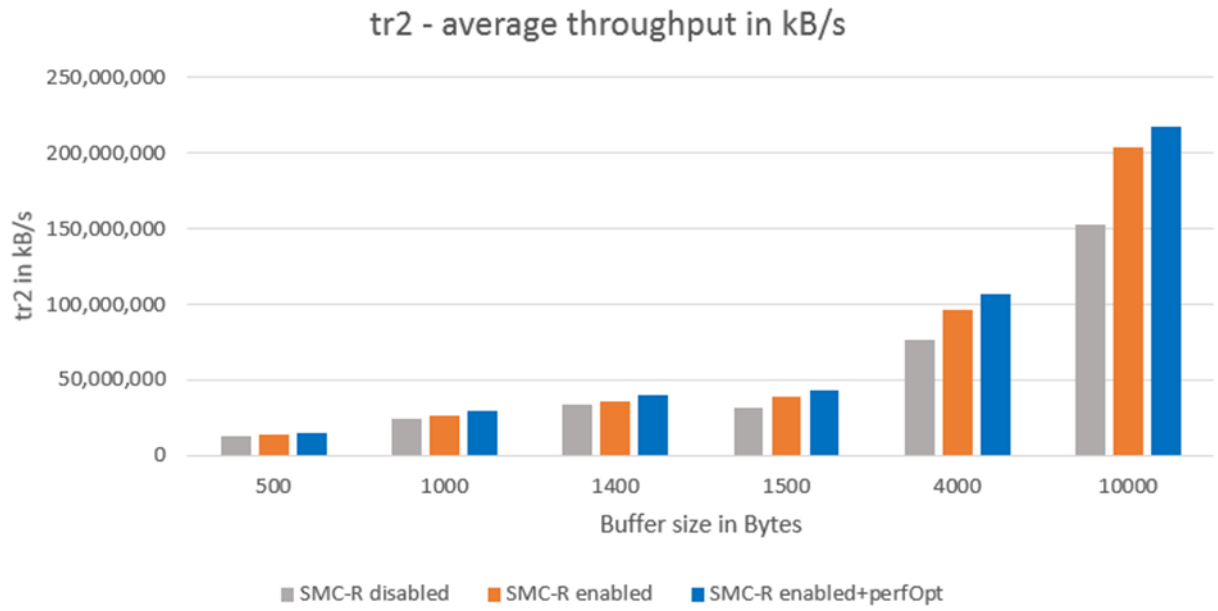
Further improvements were achieved in this environment by specifying the following additional attributes (SMC-R enabled, and performance optimized – blue bars):

conns_per_lg	64	Number of connections per link group
init_snd_pools	16	Number of send buffer pools to allocate quickly
rx_intr_packets	256	Number of packets to process in interrupt context

Network throughput

The following figure demonstrates the differences in network throughput.
The throughput is shown in kB per second for different buffer sizes which are given in bytes. The grey bars show the throughput when the SMC-R feature is disabled, the orange bars when it is enabled and the blue bars with the additional performance optimizations specified.

Figure 3 - SAP niping network throughput



Depending on the buffer size, we saw a total improvement from **15-29%** of the throughput between SMC-R being disabled (grey), and SMC-R being enabled with the additional performance optimizations (enabled and performance optimized, blue). Results are shown in the table 2.

Table 2 - SAP niping network throughput total improvements

tr2 - average throughput in kB/s				
Buffersize	SMC-R disabled	SMC-R enabled	SMC-R enabled+perfOpt	Total improvements in %
500	12,828,407	13,563,368	15,229,045	15.76
1000	24,567,610	26,172,529	29,205,607	15.88
1400	33,602,151	35,978,618	39,700,544	15.36
1500	32,150,206	38,675,743	43,004,587	25.24
4000	76,781,327	96,302,003	107,020,548	28.26
10000	152,737,048	203,450,521	217,618,384	29.81

CPU load

Improvements concerning the CPU load are shown with the help of the nmon tool, a computer system monitoring tool.

Information about the nmon tool is available in the AIX documentation:
https://www.ibm.com/support/knowledgecenter/en/ssw_aix_71/com.ibm.aix.cmds4/nmon.htm
There is also a nmon analyzer available:
https://www.ibm.com/developerworks/community/wikis/home?lang=en#!/wiki/Power+Systems/page/nmon_analyser

In the following nmon snippets, the CPU utilization is shown with SMC-R being disabled or enabled.

Figure 4 shows the total CPU usage when sending 3 million packets of the buffer size 4k bytes via SAP niping.

Figure 4 - SAP niping cpu load 1

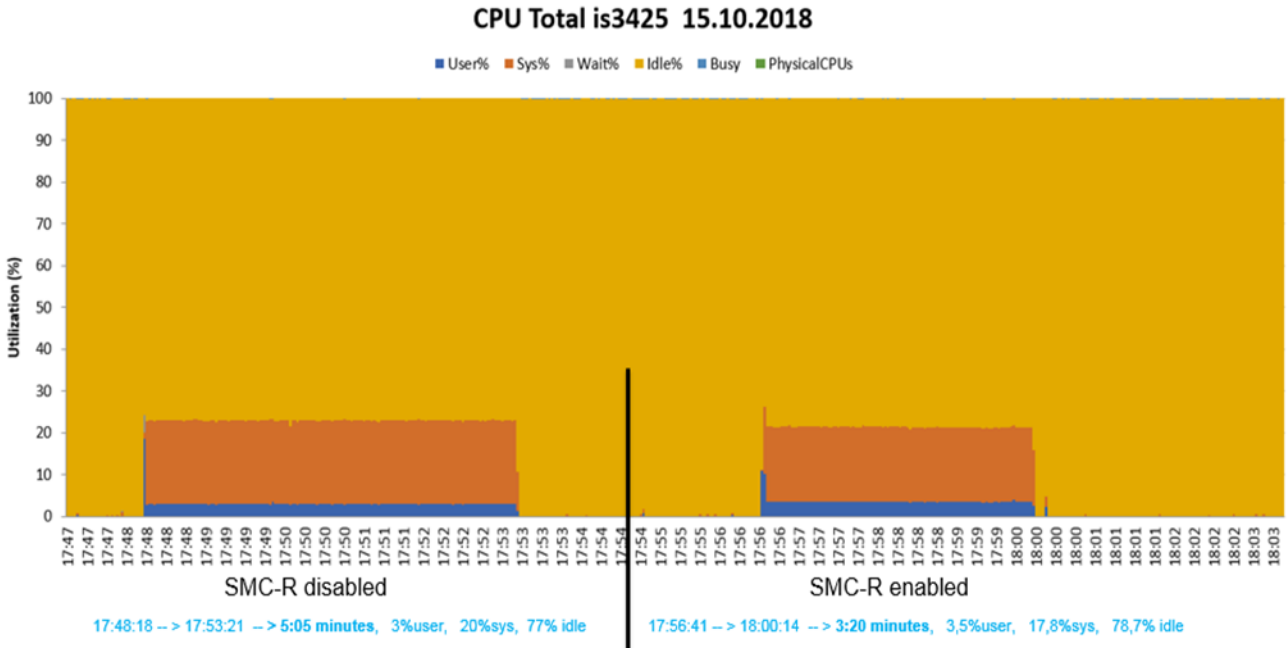


Figure 4 clearly shows the advantage of SMC-R. The workload generated by SAP niping takes longer (x-axis) to complete, 5 minutes (= 300s) with SMC-R disabled and 3 minutes and 20 seconds (= 200s) with SMC-R enabled. This is a performance improvement of 33%. As an additional benefit, the usage of SMC-R frees up CPU cycles. The CPU idle time increases from 77% to 78,7%.

A more detailed analysis of the CPU values documented by nmon reveals less user time, less system time and more free cycles for other workload:

- %sys (AIX kernel code): 17.8% over 200 sec with SMC-R vs 20% over 300 sec without SMCR
- %user (application code): 3.5% over 200 sec with SMC-R vs 3% over 300 sec without SMCR
- %idle (idle time): 78.7% over 200 sec with SMC-R vs 77% over 300 sec without SMCR

Figure 5 shows a nmon snippet of the CPU utilization when sending 3 million packets of buffer size 1KB once (bar number 1 and 5) and 5 times parallel (bar 2 and 6) and then 4KB packets (bar 3 and 7) and 5 times parallel (bar 4 and 8) via SAP niping. The parallel scenario is typical for an SAP environment, which has parallel work processes (WP), where each WP has a connection to the SAP Database. A significant performance improvement concerning the used time (x-axis) can be seen.

Figure 5 - SAP niping cpu load 2

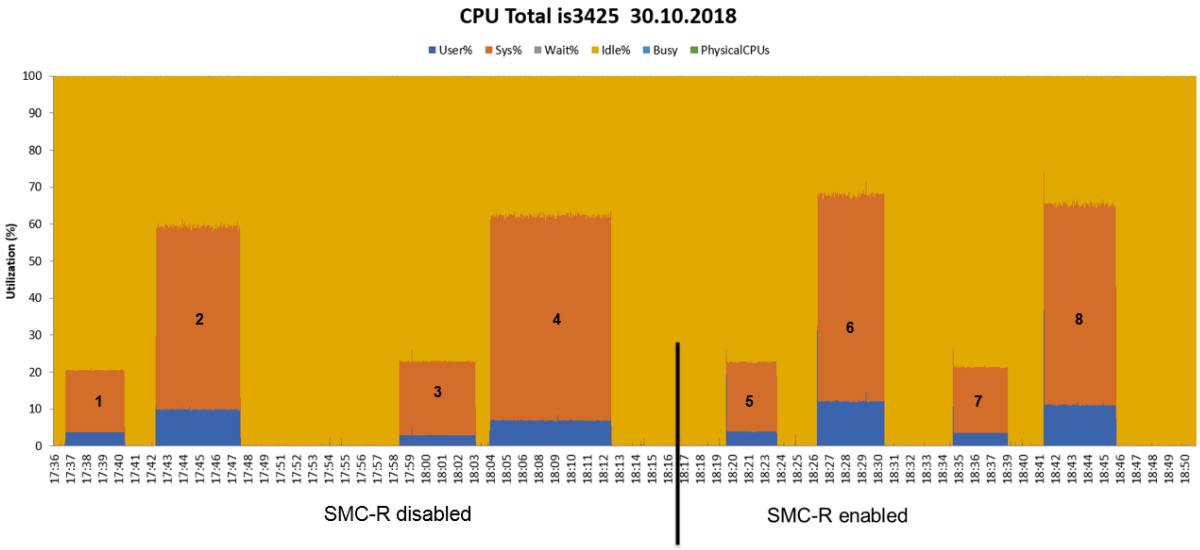


Table 3 shows in detail the improvement, when using SMC-R for different buffer sizes and having parallel send/receive streams. It is also visible, that with SMC-R enabled the CPU utilization is slightly higher, but the overall CPU utilization is better.

For example, bar 1 and 5:
1 disabled: 240s * 20.6 %CPU = 49.44 s of 100% CPU
5 enabled: 200s * 22.9 %CPU = 45.80 s of 100% CPU
(7,4%)
and bar 2 and 6:
2 disabled: 360s * 58.8 %CPU = 211.68 s of 100% CPU
6 enabled: 300s * 67.9 %CPU = 203.70 s of 100% CPU
(3.8%)

Table 3 - SAP niping CPU load 2

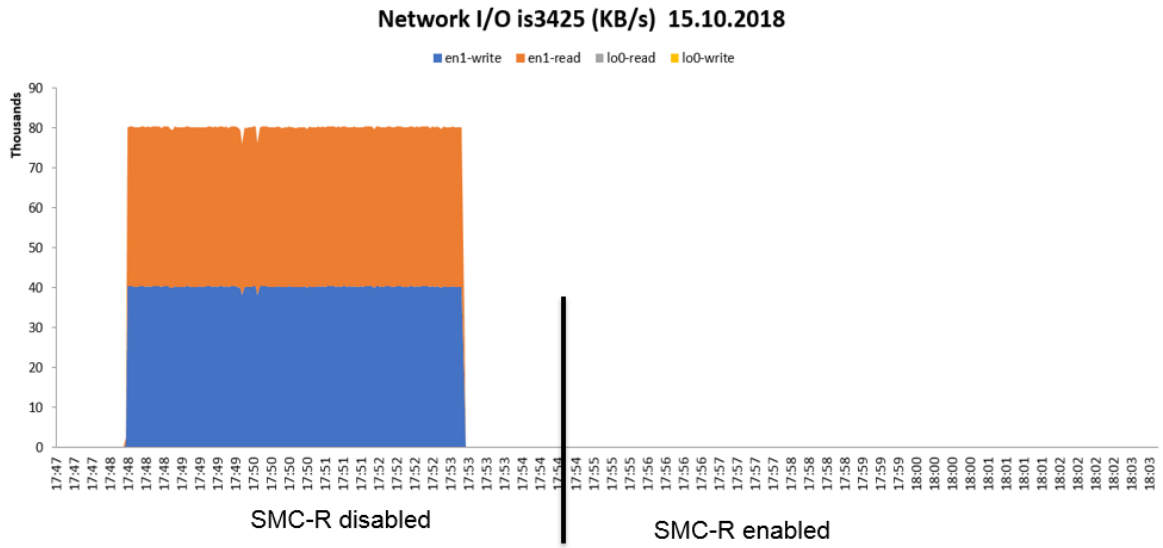
bar	enabled	Buffer size	parallel	User %	Sys %	Idle%	CPU %	Used time second
1	n	1000	1x	3.7	16.9	79.4	20.6	240
2	n	1000	5x	9.9	48.9	41.2	58.8	360
3	n	4000	1x	3	20	77	23	300
4	n	4000	5x	7	55.4	37.6	62.4	480
5	y	1000	1x	3.9	19	72.1	22.9	200
6	y	1000	5x	12.1	55.8	32.1	67.9	300
7	y	4000	1x	3.6	17.9	78.4	21.6	210
8	y	4000	5x	11.3	37.6	34.1	65.9	310

Additional experiences

nmon is a useful tool to verify CPU utilization, but it does not provide network information when SMC-R is enabled. SMC-R uses RDMA. Since the data is sent directly by the adapter, and it does not flow through the usual network stack, nmon is not able to capture these packets. Therefore, the

network I/Os for "SMC-R enabled" are not shown in **Figure 6**. Figure 6 covers the same test intervals as the previous scenario and shows the network activity instead of the CPU load.

Figure 6 - SAP niping nmon network information



Enqueue performance analysis, SAP enqt tool results

The **SAP enqueue server** (also known as the lock server) is the SAP system component that manages the lock table. There is only one SAP enqueue server in a distributed SAP system. The SAP enqueue server receives a lock request and checks the lock table to determine whether the lock request collides with an existing lock. If it does, the SAP enqueue server rejects it. If it does not collide, the SAP enqueue server sets the lock and makes the entry in the lock table. The purpose of the lock mechanism is to prevent two SAP transactions from changing the same data in the database simultaneously.

The SAP enqueue server must run with high performance. The benefits of SMC-R between SAP application server and a standalone SAP enqueue server can help as shown below.

The purpose of the standalone **enqueue test** is to measure enqueue rates that are not influenced by any SAP system activity. The tests use the standard SAP standalone enqueue test tool called enqt that is delivered with the SAP kernel. The program enqt initiates client connections to the SAP enqueue server and triggers SAP enqueue lock operations. The SAP enqt tool is used to analyze the number of enqueue operations in a time interval. In a standard SAP installation, the enqt program is in the /usr/sap/<SID>/SCS<instance number>/exe directory located.

Remark: enqt is the enqueue test tool for the old enqueue server. Starting with SAP kernel 773 a new SAP standalone enqueue server (ENSA2) is the default. The new tool to measure enqueue performance in an ENSA2 environment is enq_admin. Because most customers have the old environment active, the focus here is on the old SAP standalone enqueue server.

For more information about enqt refer to the command line help by typing enqt –help. For this test, the following enqt OpCodes were used:

OpCode	Description
8	Live Monitoring of enqueue Statistics
10	Emulate SD Benchmark Load
11	Create enqueue locks
12	Destroy enqueue locks
20	Verify entries

Emulate SD Benchmark Load with enqt

The enqueue rates are measured by running the enqt tool on the LPAR where the SAP enqueue server is running (is3425). There the enqt tool uses OpCode 8, which prints enqueue statistics every 5 seconds and shows the accumulated enqueue Ops/min rate for different time intervals:

```
enqt pf=<enq_server_profile> 8 5
```

On the remote LPAR (is3060), the enqt tool generates an enqueue workload like an SAP SD benchmark workload, with OpCode 10 during a time interval of 900 seconds:

```
enqt pf=<central services profile> 10 900
```

Table 4 shows the results when sending the SAP SD benchmark workload one time:

Table 4 - SAP enqt SD benchmark

SMC-R	Operation	actual	10 sec	1 min	5 min
Enabled with 1 process	Enqueue	412272	412944	414001	385301
Disabled with 1 process	Enqueue	381000	382080	381159	333055

--- > Operations per minute: 381159 : 414001 * 100 = 92,07% shows an improvement of 7,93%

Table 5 shows the results when sending the SAP SD benchmark workload five times parallel:

Table 5 - SAP enqt SD benchmark

SMC-R	Operation	actual	10 sec	1 min	5 min
Enabled with 5 processes parallel	Enqueue	1178400	1180374	1178502	1180909
Disabled with 5 processes parallel	Enqueue	1122900	1123536	1125291	1125226

--- > Operations per minute: 1125291 : 1178502 * 100 = 95,48% shows an improvement of 4,52%

Create enqueue locks

Another scenario is to generate a fixed amount of enqueue locks and to measure the needed time. Therefore, the enqueue table size was increased to 640 000 KB, via the following profile parameter of the SAP enqueue server:

B13_ASCS01_is3425:enque/table_size = 640000

On the remote LPAR (is3060), the enqt tool generates for each of 10 user 55000 locks:

enqt pf=<central services profile> 11 10 55000

The used time and required CPU can be captured via nmon:
Figure 7 and 8 show that the needed time (x-axis) is smaller, to see the benefit more granular the detailed number must be extracted.

Figure 7 - Create enqueue locks – LPAR simulates load

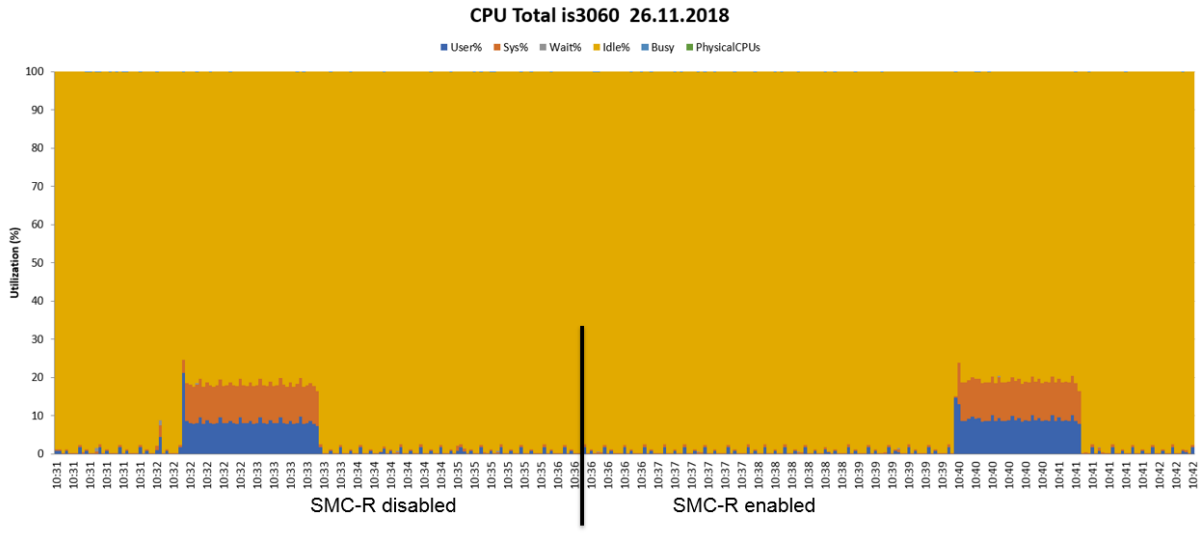
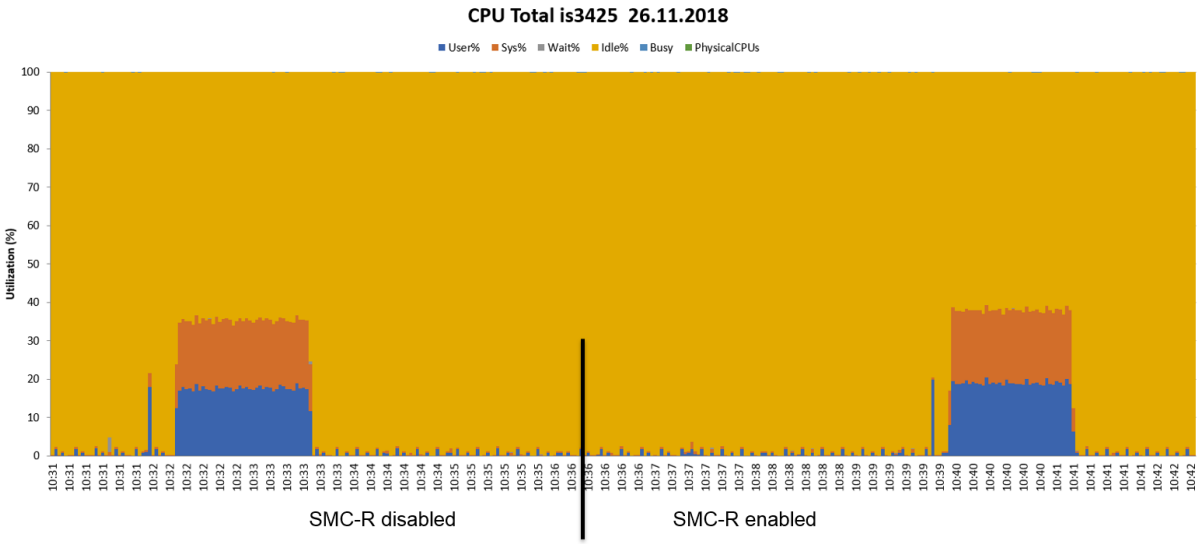


Figure 8 - Create enqueue locks – LPAR with enqueue server



The extracted numbers from nmon with SMC-R disabled/enabled are:

Table 6 – SAP enqt create locks

	Time	user	sys	idle	Busy = CPU	
SMC-R disabled	82 sec	17.3%	17.7%	65%	35%	
SMC-R enabled	77 sec	18.8%	19.2%	62%	38%	

The creation of 55000 enqueue locks results in a performance improvement of 6% when SMC-R is enabled, the CPU usages is higher in a shorter timeframe (35 % in 82s / 38% in 77s).

Summary

All tests show the benefit of SMC-R, lower latency, higher throughput and lower CPU consumption. Therefore, IBM recommends using SMC-R in a distributed SAP environment.

Referenced documents

In the AIX documentation you can find the concept, benefits, and more to the topic SMC-R:
https://www.ibm.com/support/knowledgecenter/en/ssw_aix_72/com.ibm.aix.rdma/smc_r.htm

More information is available concerning configuration and performance tuning for SMC-R on the AIX wiki blog.

https://www.ibm.com/developerworks/community/wikis/home?lang=en#!/wiki/Power%20Systems/page/Shared%20Memory%20Communications%20over%20RDMA%20(SMC-R)

Copyrights and Trademarks

IBM Corporation
Systems and Technology Group
Route 100
Somers, NY 10589

Edition Notice: 2018
This is version 1.0 of this document.

Disclaimer and Special Notices

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems.

Furthermore, some measurement may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products.

All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs

ANY INFORMATION HEREIN IS PROVIDED “AS IS” WITHOUT WARRANTY OR INDEMNIFICATION OF ANY KIND BY IBM AND DO NOT ANY EXPRESS OR IMPLIED, WARRANTIES OR CONDITIONS OF MERCHANTABILITY, FITNESS OR USAGE FOR PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT.