



Shared Memory Communications over RDMA (SMC-R) using RoCE

Optimize Communications with Memory to Memory Communications

Today's digital transformation is driving extreme growth in data. There is a massive demand for access to the data, creation of new services using the data and the ability to get actionable insights from the data. This in turn is creating the need for high performance server to server communications. Shared Memory Communications over RDMA (SMC-R) is the perfect solution.

New SMC-R revolutionizes data transfer

SMC-R is an open protocol that can be exploited by any platform or OS that implements the protocol. Today, SMC-R is supported by z/OS®, AIX® (P) and Linux® on IBM Z®.

On the IBM Z having SMC-R, along with a RoCE¹ Express adapter, is a solution designed for both internal and cross server communications that takes advantage of high-speed protocols and *direct memory placement* of data for faster communications. Use of SMC-R allows organizations to transfer huge amounts of data quickly, with low CPU overhead, and with low latency. Best of all, there are no application changes needed. The use of SMC-R is *totally transparent* to applications enabling rapid time to value.

Using memory to a new way

SMC-R uses a sockets over RDMA communications protocol. RDMA is a standard that enables network adapters to transfer data using direct memory to

memory communications, eliminating buffer write overhead. Using RDMA, SMC-R provides host-to-host direct memory communications without incurring significant TCP/IP processing costs. SMC-R allows existing sockets-based applications to transparently exchange data over a Converged Ethernet fabric without any code changes. In addition, SMC-R protocol provides extra investment protection as it allows customers to leverage existing Ethernet infrastructure using a well accepted communications standard.

SMC-R uses existing TCP connections

SMC-R uses existing TCP connections to establish or terminate a connection and monitor heartbeat functions. Once the initial handshake is complete, communications then uses SMC-R sockets-based communications. Socket application data is then exchanged via RDMA while the TCP connection remains active. This model preserves critical operational and network management features of TCP/IP.

Qualities of Service Preserved

The qualities of service known to IBM Z are preserved with SMC-R. RoCE and SMC-R offer high availability, traffic load balancing and failover, minimal operational changes, dynamic discovery of partner RDMA capabilities and dynamic connection setup.

RoCE also preserves most network administrative and operational models and supports security configurations to help organizations reduce risk.

The combination of a RoCE Express adapter and the SMC-R protocol can deliver improvements in network performance for processor to processor communications with reduced latency, improved throughput, and application transparency.

Value

IBM SMC-R delivers faster communications with:

- Transparent application use
- Low CPU utilization and latency
- Ability to leverage existing infrastructure
- Preserves TCP/IP security, management
- Standards-based interface

SMC-R Can Benefit Many Workloads

Use cases include:

- Network intensive and transaction oriented workloads that require many round trip communications to complete a single transaction
- Transactional workloads exchanging large messages (e.g. WAS Type 4 connections to Db2® or CICS®)
- Applications that use TCP-based communications for large file transfer programs or any large streaming workloads (bulk data)
- z/OS Sysplex Distributor (SD) clients with a RoCE connection to SD using VIPARoute
- Streaming (or bulk) file transfer workloads (e.g. FTP) can experience improvements in CPU and throughput
- Applications where data must be transferred between two locations in the data center (300 meter distance limitation per link)
- Any IT organizations requiring fast file transfers between processors running z/OS

SMC-R Requirements

The minimum HW requirements are:

- IBM z15 (z15), IBM z14 (z14), IBM z13s® (z13s), IBM z13™ (z13), IBM zEnterprise® EC12, or zEnterprise BC12
- RoCE Express adapters
 - 25 GbE RoCE Express2.1 (IBM z15 only)
 - 25 GbE RoCE Express2 (IBM z15 or IBM z14)
 - 10 GbE RoCE Express2
 - 10 GbE RoCE Express
- A minimum of two adapters is recommended
- Layer 2 OSA port connection is required

Planning

A z/OS V2.2 Communications Server planning tool, SMC-AT, is available to help clients gain insight into the applicability of this technology for specific z/OS application workloads. SMC-AT is not available for Linux on Z.

Summary

SMC-R over RoCE represents a low latency, low CPU overhead alternative for data transfers to help organizations prepare for the huge volumes of enterprise data exchange from increased traffic. It provides the strengths of SMC-D with the benefits of application transparency plus outstanding performance.

¹ RDMA over Converged Ethernet