IBM z/OS Shared Memory Communications: Security Considerations

Version 2 - May, 2016

Contents

Abstract

This paper is provided for both IBM and IBM customers who have an interest in the security characteristics of the IBM z/OS® Shared Memory Communications (SMC) functions. This includes SMC over RDMA (SMC-R), introduced in z/OS V2R1 and SMC - Direct Memory Access (SMC-D), introduced in z/OS V2R2. It is assumed that readers already have a basic background in TCP/IP protocols and the related z/OS implementation of those protocols.

SMC-R is an open protocol defined in the informational RFC entitled *IBM's Shared Memory Communications over RDMA* (https://tools.ietf.org/html/rfc7609). The SMC-R portions of this paper, however, focus exclusively on the IBM z/OS implementation of the SMC-R protocol.

SMC-D is closely related to SMC-R, but is a proprietary mechanism within the IBM z Systems™ family of products that is based on the Internal Shared Memory (ISM) capabilities introduced with the IBM z13™ (z13) hardware model.

The primary purpose of this paper is to describe how SMC communications are secured as well as the way existing z/OS Communications Server network security features apply to SMC. The paper is organized into the following topics:

1. **Shared Memory Communications Overview**
   is an introduction and overview of Shared Memory Communications (SMC-R and SMC-D) concepts, architecture and z/OS implementation. This portion of the paper provides sufficient background to understand the subsequent topics. If you are looking for a detailed discussion of SMC itself, please refer to the SMC reference materials at http://www.ibm.com/software/network/commserver/SMC/index.html
2. **Security characteristics of SMC connections**
   examines the SMC connection from a security perspective and compares it to IP packets and TCP connections.
3. **z/OS network security features and SMC**
   is an overview of the z/OS Communications Server network security features and how each of them applies to SMC communications.

The reader is assumed to have a basic understanding of network communication protocols including Ethernet, IP and TCP as well as a basic understanding of network security technologies like firewalls, packet filtering, IPsec, TLS/SSL and intrusion detection.

z/OS Shared Memory Communications Overview

Shared Memory Communications (SMC) allows two SMC capable peers to communicate using memory buffers that each peer allocates for the partner's use. There are two forms of Shared Memory Communications:

- SMC over Remote Direct Memory Access (SMC-R)
- SMC - Direct Memory Access (SMC-D)

Both forms allow your TCP sockets applications to benefit from direct, high-speed, low-latency, memory-to-memory (peer-to-peer) communications transparently – no changes are required in your application programs.

SMC provides services that are designed for enterprise class data center networks. Communicating peers (the z/OS TCP/IP stacks) dynamically learn about the shared memory capability by using traditional TCP/IP connection establishment flows. With this awareness, the TCP/IP stacks can switch from TCP network flows to more efficient direct memory access or RDMA flows, as appropriate. The application programs are unaware of the switch to shared memory communications.

Let's take a closer look at each form of SMC.

z/OS Shared Memory Communications over RDMA

Remote Direct Memory Access (RDMA) is a communications technology that enables a host to make a subset of its memory directly available to a remote host. By doing so, data can be transferred between hosts very efficiently and without any help from the CPU on the source or target host. Historically, RDMA has been confined to high-performance computing environments where the cost of maintaining RDMA-capable network fabrics such as InfiniBand® was justified given the emphasis of performance over cost. However, RDMA is now available on standard Ethernet-based networks by using the industry (InfiniBand Trade Association) standard referred to as RDMA over Converged Ethernet (RoCE). With RoCE, the cost of adopting RDMA is lower because it can flow over the Ethernet fabrics that are already in place to carry IP network communications. Both standard TCP/IP and RDMA traffic can flow over the same physical LAN fabric at the same time, but RDMA network interface cards (RNICs, also referred to as RoCE host channel adapters (HCAs)), are required to do so. On z Systems, the 10Gb RoCE Express adapter serves as the RNIC.

z/OS Communications Server V2R1 introduced a new capability that combines the performance benefits of RDMA with the widely-used TCP/IP sockets

programming interface. This function, called *Shared Memory Communication – RDMA (SMC-R)* allows your TCP sockets applications to benefit from direct, high-speed, low-latency, memory-to-memory (peer-to-peer) communications over RDMA transparently – no changes are required in your application programs.

SMC-R provides a set of RDMA services that are designed for enterprise class data center networks. Communicating peers (the z/OS TCP/IP stacks) dynamically learn about the shared memory capability by using traditional TCP/IP connection establishment flows. With this awareness, the TCP/IP stacks can switch from TCP network flows to more efficient direct memory access flows that use RDMA. The application programs are unaware of the switch to shared memory communications. Once a TCP connection switches to the SMC protocol, all data from that point forward is exchanged using RDMA. The TCP connection remains active over the IP network. If the TCP connection were to terminate for any reason, then the SMC connection would also be immediately terminated.

The remainder of this section will describe relevant characteristics of SMC-R communications in only enough detail to provide a basis for the later security discussion. For a more complete description of the z/OS SMC-R implementation, refer to the *z/OS Communications Server IP Configuration Guide Version 2 Release 2 (SC27-3650)*, Chapter 10.

SMC-R: A Hybrid Protocol

Shared Memory Communications over RDMA is a *hybrid protocol* that uses RDMA technology within an existing IP network topology. SMC-R connections are established and operate transparently to applications within the context of their TCP socket connections. IP communications occur over OSA adapters (as they have in the past) and the associated SMC-R connections are established over the RNICs. As such, the RNICs must be attached to the same network infrastructure as the OSAs.

SMC-R's reliance on existing IP network topology and TCP connection setup preserves critical TCP/IP operational and network management features, including compatibility with transport layer load balancers (e.g., Sysplex Distributor), preserving the IP security model (IP filters, VLANs, TLS/SSL, and so forth), and minimal (or zero) topology changes to accommodate the use of RDMA. This reliance on IP topology is a foundational element of the SMC-R security landscape.

SMC-R Eligibility

In order for two nodes to be eligible to communicate with SMC-R, several criteria must be met:

- Both must be enabled for SMC-R
- Both must have direct access to the same physical LAN fabric
- Both must have direct access to the same IP subnet and VLAN (if VLANs are defined)
- The enterprise does not require IPsec protection of network and transport layer headers on the LAN segment (note, however, that TLS/SSL can be used to protect application data across SMC-R enabled connections). We'll discuss the role of IPsec in more detail later in a few pages.

The "direct access" requirements are based on the fact that the underlying RDMA connections are non-routable. This means that *SMC-R connections are not routable* as well. The direct access requirements ensure that a direct communication path exists at layer 2 between the SMC-R capable nodes, with no intervening IP router. The additional VLAN requirement further confines the traffic within the physical LAN fabric in cases where VLANs are in use.

The topology requirements are illustrated in Figure 1 below. As you can see, the SMC-R enabled TCP connections between HOST A and HOST B are allowed since both hosts have OSA and RoCE adapters connected to the same physical LAN fabric, VLAN and IP subnet. On the other hand, even though HOST C is attached to the same physical LAN fabric, it cannot establish any IP connections to HOST A or HOSTB without an intervening IP router since it is connected to a different VLAN and IP subnet.

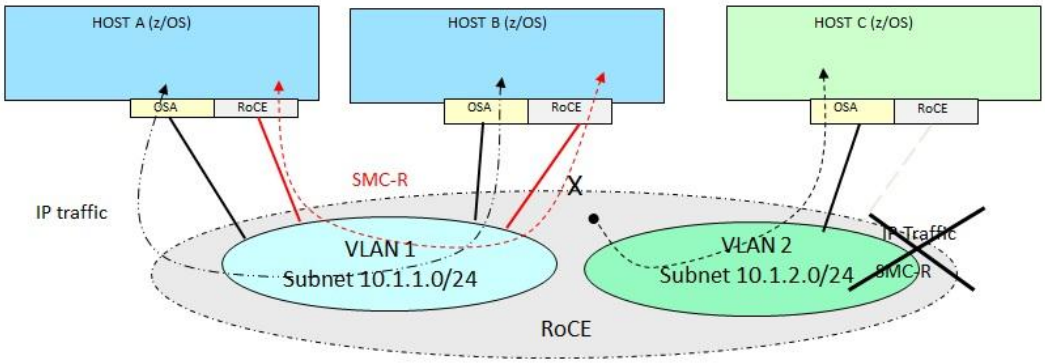

Figure 1 Network topology and SMC-R eligibility

Since SMC-R connection processing leverages existing IP topology (TCP/IP connection setup) SMC-R connections transparently “inherit” the same VLAN and IP subnet connection eligibility attributes of the associated TCP connection. When VLANs are in use, SMC-R connections then become VLAN qualified.

Since SMC-R’s topology and eligibility requirements mimic those of IP, the level of trust that an enterprise has in its IP network infrastructure should mirror its trust in that infrastructure for SMC-R purposes.

Enabling SMC-R and Connection Setup

SMC-R is enabled by specifying the SMCR parameter of the GLOBALCONFIG statement in the TCPIP profile data set and including one or more Peripheral Component Interconnect Express (PCIe) function ID (PFID) values. Each PFID value represents an RNIC adapter that is configured by using the traditional hardware configuration definition (HCD) tools. TCP/IP activates the RNICs when the first SMC-R capable IP interface is started. IPAQENET or IPAQENET6 interfaces with the OSD channel path ID type can be configured for SMC-R capability.

Any TCP connections that traverse SMC-R capable IP interfaces are eligible for SMC-R communications. The decision about whether an eligible connection uses SMC-R communications is reached during traditional TCP connection establishment. The sequence of flows that determine whether or not to use SMC-R on a given TCP connection is called *Rendezvous processing*, which is illustrated in Figure 2 below

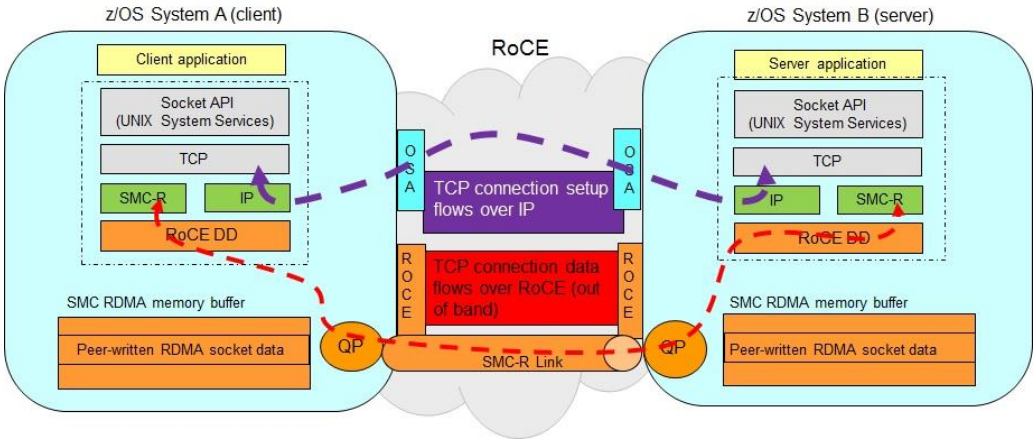

Figure 2: Rendezvous processing

The Rendezvous exchange of information occurs in three stages:

1. TCP connection establishment flows:
   TCP connections are still established using the standard three-way handshake mechanism. When SMC-R communications are enabled, the client adds TCP options settings in the SYN request to indicate that it supports SMC-R protocols. The server, when SMC-R communications are enabled, likewise responds with TCP options settings for SMC-R in the SYN-ACK response. No additional exchange of information is required in this stage of the rendezvous processing.

2. In-band SMC-R Connection Layer Control (CLC) messages:
   After the TCP three-way handshake succeeds, the client and server negotiate the use of SMC-R for this TCP connection by using SMC-R CLC messages that flow as in-band data over the TCP connection. Conceptually, these flows are similar to the TLS/SSL handshake processing that occurs after the TCP connection is established, but they occur before any data is allowed to flow over the TCP connection (including the TLS/SSL handshake). The CLC messages exchange the following information:

   - Layer 2 addressing information (MACs and GIDs)
   - RoCE credentials, consisting of
     – Remote memory buffer access information
     – Queue Pair and related information

3. SMC-R Link Layer Control (LLC) messages:
   Using the RoCE credentials exchanged in phase 2, an RDMA connection called an *SMC-R Link* is established between the two peers across *Reliable Connected Queue Pairs* (RC QPs). SMC-R LLC messages are then exchanged across the SMC-R Link to confirm that the RoCE information is correct and that the RC QPs that comprise the SMC-R link have connectivity. This stage is skipped if an existing RoCE connection is used for this TCP connection.

Since the z/OS TCP/IP stack does not allow the client and server applications to exchange application data before or during rendezvous processing, the TCP connection can revert to IP protocols if there is a failure during the setup of the SMC-R communications. However, once the RoCE connection is confirmed by using the LLC messages, the TCP connection is committed to using SMC-R protocols and cannot fall back to using IP protocols if SMC-R communications encounter an error.

Even though application data is sent out of band from the TCP connection with SMC-R communications, the TCP connection remains active in order to

preserve the connection state for monitoring and management functions, load balancers, etc., and to support various stack functions, including connection termination processing.

z/OS Shared Memory Communications - Direct Memory Access

This section will describe relevant characteristics of SMC-D communications in only enough detail to provide a basis for the later security discussion. For a more complete description of the z/OS SMC-D implementation, refer to the *z/OS Communications Server IP Configuration Guide Version 2 Release 2 (SC27-3650-05 or later)*, Chapter 10.

Figure 3 provides an overview of SMC-D connectivity.

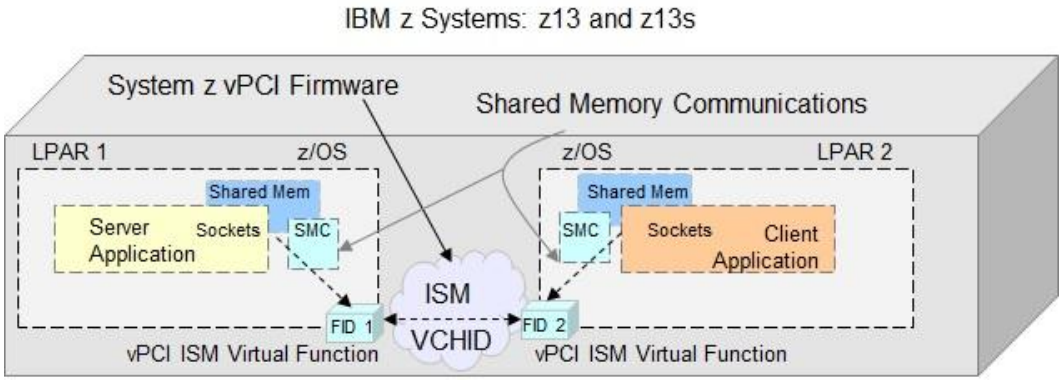

Figure 3 SMC-D overview

Shared Memory Communications - Direct Memory Access (SMC-D) uses the z System Internal Shared Memory (ISM) virtual PCI function for communications between two SMC capable peers that are located on the same central processor complex (CPC). ISM is a z System firmware solution available on IBM z13/IBM z13s™ (z13s) and later models that leverages existing virtual memory capabilities.

Although ISM is not a physical resource, the ISM technology is represented by a z System logical (virtual) channel ID (CHID). ISM CHIDs are configured in HCD. z13 Systems support up to 32 unique ISM CHIDs. Each ISM CHID is analogous to a unique physical network (LAN), each having a unique Physical Network ID (PNet ID). In some cases z System users deploy multiple unique business lines (for example, customer A and customer B) within the same CPC. Often, each unique business line is provisioned a separate set of LPARs, and in some cases the workloads for the unique business lines must be completely isolated. This isolation can be accomplished by provisioning a

unique ISM CHID for each set of LPARs. Alternatively, or in addition to unique ISM CHIDs, workloads can be isolated on the same ISM CHID using unique VLANs. Figure 4 illustrates both approaches.

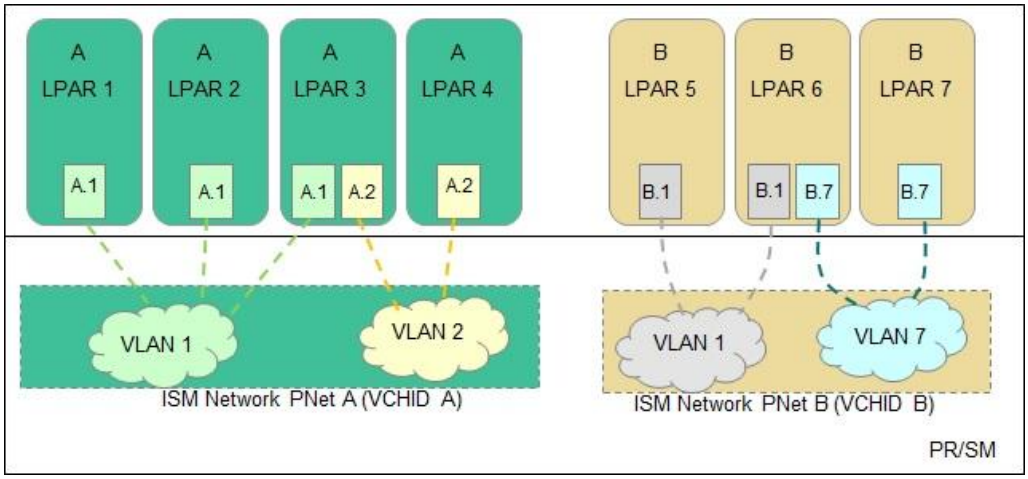

Figure 4 ISM isolation approaches

SMC-D is very similar to SMC-R, extending the benefits of SMC-R to same CPC operating system instances without requiring physical resources (RoCE adapters, PCI bandwidth, NIC ports, I/O slots, network resources, 10GbE switches etc.). In a manner very similar to SMC-R, the communicating peers, such as TCP/IP stacks, dynamically detect the shared memory capability by using traditional TCP/IP connection establishment flows. The shared memory enables the TCP/IP stacks to switch from TCP network flows to more optimized direct memory access flows that use ISM.

SMC-D is enabled by specifying the SMCD parameter on the GLOBALCONFIG statement of the TCPIP profile data set.

All of major points regarding SMC-R made in the previous section (general architecture, hybrid protocol, eligibility conditions (except those related to physical LAN connectivity), and SMC negotiation via rendezvous processing) apply to SMC-D as well. Generally speaking, the key differences between SMC-R and SMC-D stem from the interface differences between using an RNIC versus using ISM, and all of these differences are transparent to application programs.

SMC-R and SMC-D coexistence

As Figure 5 illustrates, both forms of SMC can be used concurrently to provide a highly optimized connectivity solution.

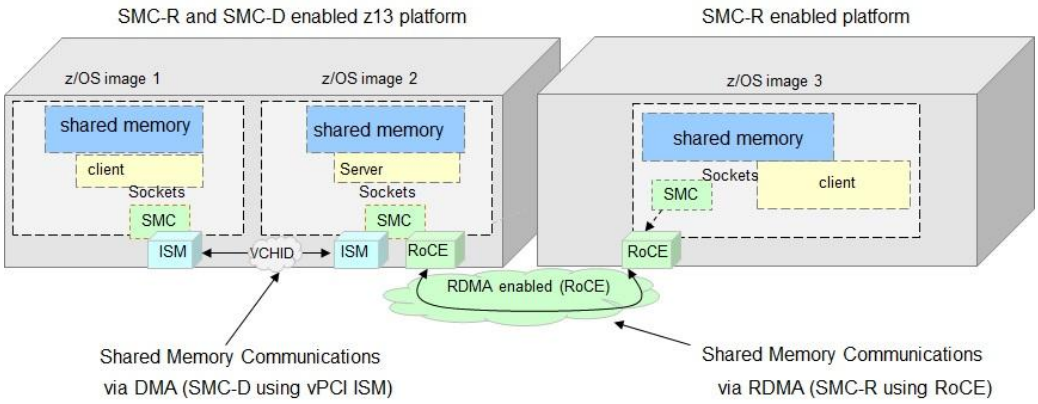

Figure 5 SMC-R and SMC-D coexistence

The SMC protocol supports concurrent exploitation of both SMC-R and SMC-D. The SMC CLC protocol allows the server to dynamically detect when both peers are enabled and eligible for SMC-D (i.e. both peers are on the same CPC) and will select SMC-D when possible. When SMC-D is not possible, then SMC-R will be used if possible. When both SMC variations are not possible, then the TCP connection will be established over traditional IP-based communications.

Security characteristics of SMC connections

This section examines SMC connections from a security perspective and compares them to IP packets and TCP connections within a single LAN segment, since that is the scope of an SMC connection.

Protecting application data

First, let's consider the protection of application-level data as it traverses the LAN. Many enterprises need to protect application-level data (transactions, database query results, transferred files, etc.) as it crosses any network between two nodes. When those applications use TCP-based protocols (FTP, HTTP, CICS® transactions, etc.), Transport Layer Security (TLS), which is the modern day version of Secure Sockets Layer (SSL) protocols can be

used to provide endpoint authentication, data authentication, data integrity and data privacy (encryption) protections for the application data. Since this protection is done above the transport layer, it can be used across SMC (both SMC-R and SMC-D) enabled connections just as effectively as it can for TCP connections over a regular IP network.

Protecting network protocol headers

This topic only applies to SMC-R. Since SMC-D involves only direct memory transfers within the same CPC, it has no concept of network frames or network transmission.

We've seen that application data can be protected over SMC-R enabled links using the TLS just like it can be over regular TCP/IP. However, what about protecting the TCP and IP headers of the packets that contain the application data? Let's look a little more closely at the types of attacks that can be directed at the lower-level networking headers and then discuss protection scenarios as they apply to regular IP networks and those carrying RDMA traffic.

TCP/IP, RDMA and the underlying Ethernet-based link layer protocols must all carry enough information in each transmission unit to ensure the data is properly handled as it traverses the network and after it arrives at its destination. Specifically, such a protocol provides information to:

- identify where it came from (source identity/location)
- identify who its intended target is (destination identity/location),
- identify where the data is supposed to go within that target (upper-layer tagging) and,
- convey other types of control information to ensure proper handling of the transmitted data.

Ethernet provides the source and destination identity/location information in the form of MAC addresses. Along with the MAC addresses, Ethernet uses an associated upper-layer addressing value (an IP address in the case of IP or a Global ID (GID) in the case of RDMA) to assist in the routing of a given frame through the LAN and to the proper device driver once it arrives at its destination. Other control information is included in each Ethernet frame header.

IP provides the source and destination identity/location information through IP addresses and the control information through various flags or headers. In the IP cases relevant to our discussion, TCP provides the upper-layer tagging in the form of TCP port information and also provides its own control information. Other control information is included in each IP packet header.

When an IP packet traverses an Ethernet-based LAN, it is susceptible to different kinds of attack techniques, including the following:

- *MAC and IP address spoofing*, where a LAN-attached device generates or modifies an Ethernet frame containing an IP packet using another device's MAC address and/or IP address as the source address information in order to impersonate that device. This technique is usually used in combination with packet injection or man-in-the-middle (MITM) attacks as described below.

- *Packet injection*, where a LAN-attached device assembles a new Ethernet frame containing an assembled IP packet using spoofed source MAC and IP addresses and sends it to a destination with the intent that it will be consumed as if it originated at the spoofed address. Note that this requires pre-knowledge of other pieces of control information as well (sequence numbers, etc.).

- *Man-in-the-middle (MITM)* attacks, wherein a LAN-attached device captures an Ethernet frame carrying an IP packet that was sent by a legitimate IP node on the LAN, modifies some of the contents, and then forwards it on to the destination node with the intent that the destination node will consume the modified data as if it came from the original sender.

- *Denial of Service (DoS)* attacks, wherein a LAN-attached device takes some sort of action to disrupt a connection that already exists between two legitimate IP nodes on that LAN segment or else to consume the resources of one of the nodes such that other nodes cannot use its services. An example of a DoS attack on a TCP connection is that of injecting a packet into an existing TCP connection with incorrect sequence information in order to force the receiver of the injected packet to close the TCP connection in response to an error condition.

There are a few of ways to mitigate the risks involved in these types of IP-based attacks.

- For LANs that are easily accessed (for example, wireless LANs, those with easily accessible wired ports, those accessible from less secure portions of the network through layer 3 routing, etc.), IPsec can be used to protect the IP packets, including the IP headers.

- For more secure LANs such as those connecting an enterprise's high performance (multi-tiered) clusters, closely controlled physical access along with firewall control for external access may be enough to ensure that only authorized IP nodes have access to the LAN.

- Regardless of LAN accessibility, z/OS Communications Server's TCP Traffic Regulation support (part of integrated Intrusion Detection Services) can be used to limit the number of TCP connections that any single client can establish against a given TCP port. This can be useful in mitigating DoS risks.

SMC-R connections are expected to be established on the second type of LAN, where the physical access is controlled tightly enough to obviate the need for cryptographic protection of network and transport layer headers.

Since SMC-R is based on RoCE which in turn is based on RDMA, it relies on the RoCE and RDMA protocols which use the Ethernet MAC and the RDMA GID for their source and destination identity/location. The RC QPs that comprise the SMC Link provide the "session" level information (analogous to a TCP connection). The RC QPs are identified by source and destination queue pairs. An SMC Link is further bound to a specific area of memory in each partner called a remote memory buffer which is identified by a special token and key, called RToken and RKey, respectively. All of this information is carried in the various session establishment flows over IP and RoCE as described above, as well as over the RNICs during the underlying RDMA connection establishment flows.

While the SMC-R sessions can be attacked using the very similar techniques as those described for IP, these connections are no more or less vulnerable to such attacks than regular TCP/IP connections. However, since there is no RDMA analog to IPsec, SMC-R connections should only be enabled on LAN segments that are physically secure.

If an enterprise does not trust in the physical security of a LAN segment enough to flow data without IPsec protection, then that segment should not be enabled for SMC-R.

Though the preceding point should be used as a general guideline regarding the appropriateness of SMC-R traffic over any given LAN segment, it is worth noting that the use of IPsec to protect a TCP session actually disables the use of SMC-R on z/OS, as described below.

Firewalls and Deep Packet Inspection (DPI) devices

One other important consideration in the use of SMC-R is its effect on firewalls and deep packet inspection (DPI) devices like "bump-in-the-wire" intrusion prevention devices. These devices are designed to process Ethernet frames, IP packets, and other well-known protocols built upon IP and they are usually deployed at network zone boundaries along with a layer 3 (IP) router. Since SMC-R traffic is not routable (due to its RDMA roots), it is not expected to reach traditional layer 3 firewalls. Likewise, it would be unusual for SMC-R

traffic to reach a non-traditional firewall or a DPI device. You should still be aware, however, that such devices may be incompatible with RDMA traffic and that product documentation for the specific device would need to be consulted to determine its behavior should it encounter SMC-R (RDMA) traffic.

z/OS network security features and SMC

This section provides an overview of the key security features of z/OS Communications Server and how each relates to SMC. Note that this section only discusses the subset of security features that are specifically relevant to the SMC functions in the TCP/IP stack – it is not a comprehensive survey of all the Communications Server security functions. Note that unless otherwise specified, this discussion applies equally to SMC-R and SMC-D.

Interface-based SMC enablement

The TCPIP profile data set's INTERFACE statement allows you to specify whether or not a given IPAQENET or IPAQENET6 interface (with the OSD channel path ID type) is enabled for SMC-R by specifying the SMCR or NOSMCR parameter. Likewise, the same interfaces can be enabled or disabled for SMC-D by specifying the SMCD or NOSMCD parameters, respectively. If not specified, the default is to enable SMC-R and SMC-D for these types of interfaces. SMC-D is also configurable via the SMCD and NOSMCD parameters for IPAQIDIO and IPAQIDIO6 (HiperSockets™) interfaces, and again, SMC-D is enabled by default. Finally, you can also enable or disable SMC-D for Dynamic XCF Links that use HiperSockets using the SMCD and NOSMCD parameters of the IPCONFIG and IPCONFIG6 statements and on the DYNAMICXCF parameter. In this case, however, SMC-D is enabled by default.

Port-based SMC exclusion

The TCPIP profile data set's PORT and PORTRANGE statements allow you to exclude a port or set of ports from SMC eligibility by specifying the NOSMC parameter. This can be a useful feature for disabling SMC for a specific z/OS TCP server application. You can use the SMC parameter of these statements to override the TCP/IP stack's autonomic application of SMC to the specified server ports to ensure that SMC is always used when it is supported by both endpoints of a given connection. For more information on SMC autonomics, refer to the SMCGLOBAL AUTOSMC parameter of the GLOBALCONFIG statement in the TCPIP profile data set.

SAF-based network access controls

z/OS Communications Server provides two types of SAF resources that allow system administrators to control access of z/OS userids to network resources. One controls access to different network zones (hosts, subnets and networks) and the other controls access to local TCP ports.

EZB.NETACCESS.*sysname.tcpname.zonename* resources control which z/OS userids are allowed to access different network zones. Userids with permission to a NETACCESS resource are allowed to send and receive data to and from the network zone represented by the *zonename*. Network access control provides an additional layer of security on top of any authentication and authorization mechanisms that are used in the network or at the peer system by preventing unauthorized users from communicating with the peer network resource.

EZB.PORTACCESS.*sysname.tcpname.portname* resources determine which z/OS userids are allowed to bind to specific TCP ports. Userids with permission to a PORTACCESS resource are allowed to bind to the TCP port represented by the resource.

Since both NETACCESS and PORTACCES controls are enforced at the socket layer, they apply fully to SMC enabled connections.

For more information on NETACCESS and PORTACCESS resources, refer to the *z/OS Communications Server IP Configuration Guide Version 2 Release 2 (SC27-3650)*, Chapter 3.

IP filter rules

z/OS Communications Server's IP security function includes IP packet filtering that controls the flow of IP packets into and out of the TCP/IP stack. An administrator can define IP security policy to permit or deny any type of IP traffic from entering/exiting the TCP/IP stack. Both local and routed traffic are supported and filter rules can be defined using a wide variety of criteria.

Basic permit/deny filters for local traffic are honored for TCP connections that use SMC. A permit filter allows TCP connections to be established while a deny filter prevents them. This is true regardless of the use of SMC. Once a TCP connection is successfully established, any change in IP filtering configuration that installs a deny filter for the connection will immediately cause the connection to be dropped for both IP and SMC traffic.

Note that since RoCE nor ISM (and hence, SMC-R and SMC-D) traffic is not routable, IP filters that apply to routed traffic do not apply to SMC traffic.

For more information on IP filtering, refer to the *z/OS Communications Server IP Configuration Guide Version 2 Release 2 (SC27-3650)*, Chapter 18.

IPsec

In addition to IP filtering, Communications Server's IP security function provides a complete IPsec implementation. IPsec is a cryptography-based set of technologies that allow participating peers to establish a wide variety of VPN tunnels between themselves. As described earlier, IPsec protects entire IP packets (not just the application data within those packets) and is completely transparent to application programs. z/OS applies IPsec protection under the direction of IP filter rules that specify an action of "protect with IPsec."

Since IPsec protocols are bound tightly to IP packet formats, they are simply not compatible with SMC. As such, when any TCP connection is protected by an IPsec filter rule, z/OS makes that connection ineligible for SMC. Furthermore, if an SMC enabled TCP connection is already established when such an IPsec filter rule is installed, then that connection will be terminated.

For more information on IP filtering, refer to the *z/OS Communications Server IP Configuration Guide Version 2 Release 2 (SC27-3650)*, Chapter 18.

SSL/TLS, including Application Transparent TLS (AT-TLS)

Transport Layer Security (TLS) and its predecessor, Secure Sockets Layer (SSL) are cryptography-based technologies that are applied just above the transport layer, typically under direction of the application program that is generating the traffic to be protected. z/OS offers two RFC-compliant TLS/SSL implementations:

- System SSL, a component of the z/OS Cryptographic Services element. System SSL can either be invoked directly through its own set of APIs or it can be invoked transparently on the application's behalf based on Communications Server Application Transparent TLS (AT-TLS) policies. As its name implies, the latter approach protects application traffic without requiring application code changes.

- Java® Secure Socket Extension (JSSE), a 100% Java implementation for WebSphere®-based and other pure Java applications. JSSE2 (the most current version of JSSE) is provided as part of the z/OS Java Runtime Environment.

Since TLS (and SSL) protection is applied to application data just above the transport layer, it is completely compatible with SMC. The TCP/IP stack, including the SMC function sees it all simply as application data.

For more information on System SSL, refer to *z/OS Cryptographic Services System Secure Sockets Layer Programming Version 2 Release 2 (SC14-7495).* For more information on AT-TLS, refer to *z/OS Communications Server IP Configuration Guide Version 2 Release 2 (SC27-3650)*, Chapter 21.

To download a copy of the documentation for the IBMJSSE2 provider, see the Security documentation on developerWorks®. In addition to this cross platform information, z/OS specific information for JSSE2 can be found in the z/OS JSSE Reference Guide.

SSH

The Secure Shell (SSH) protocol is widely-used in to cryptographically protect traffic between UNIX environments (including z/OS UNIX®). SSH operates at above the transport layer, so it applies to SMC enabled TCP connections in the same manner as TLS/SSL.

For more information on the z/OS OpenSSH implementation, refer to *z/OS OpenSSH User's Guide Version 2 Release 2 (SC27-6806).*

Application layer security protocols and features

Many security protocols exist at the application layer and some applications have their own security features. For example, Web Services Security (WS-Security), Domain Name System Security (DNS-SEC), security features within SNMP and so forth. As with TLS/SSL and SSH, the fact that these are applied above the transport layer means that it all just looks like application data to the TCP/IP stack and to SMC. Because of this, application layer security functions will continue to work over SMC enabled connection.

Integrated Intrusion Detection Services (IDS)

z/OS Communications Server Intrusion Detection Services (IDS) provide reporting and, in some cases, defensive protections, for a wide variety of intrusion-related events that fall into the following categories:

- Scan detection and reporting
- Traffic regulation
- Attack detection, reporting and prevention

Scan detection and traffic regulation are both enforced during TCP connection setup and therefore complete before SMC enablement is ever considered for a given connection.

Attack detection includes a wide range of checks. Since most of the TCP-related checks apply to inbound TCP packets, they are not relevant to the direct memory transfers employed by SMC communications. However, two

attack types do apply to SMC enabled connections. Let's examine each of these in more detail.

TCP queue size events

You can configure IDS policy to detect when a TCP connection's send or receive queues become constrained due to the amount or age of the data on the queue. When a queue becomes constrained, the IDS rule can either reset the TCP connection or continue to monitor the condition until the queue is no longer constrained. Send or receive queues for SMC enabled TCP connections are considered to be constrained in the following circumstances:

- The send queue is considered to be constrained when available outbound data cannot be sent because the peer's RMB element (SMC-R) or DMB element (SMC-D) has no available space for a prolonged period of time (30 seconds). This condition typically identifies a problem with the peer application not consuming the available data in a timely manner.

- The receive queue is considered to be constrained when inbound data that has been stored into the local memory buffer is not received by the application for more than 30 seconds.

- When either queue becomes constrained, the TCP connections are monitored or stopped according to the relevant IDS policy rule.

Global TCP stall events

You can configure IDS policy to detect attacks that are designed to consume system resources by creating many TCP connections and causing them to stall, making them unable to send data. A global stall condition is in effect when at least 50% of the active TCP connections are stalled and at least 1000 TCP connections are active. A global TCP stall IDS rule can either reset stalled connections or continue to monitor the condition. TCP connections that traverse SMC links are considered during global TCP stall detection. Such a connection is considered to be stalled when the TCB is write-blocked.

For more information on Intrusion Detection Services, refer to *z/OS Communications Server IP Configuration Guide Version 2 Release 2 (SC27-3650)*, Chapter 17.

Multilevel Security (MLS)

Multilevel security is an enhanced security environment in which the z/OS security server and trusted resource managers enforce mandatory access control policies in addition to the usual discretionary access control policies. To

participate in an MLS environment, the user IDs associated with tasks trying to access z/OS resources and those resource profiles in the SERVAUTH class need to have security labels defined.

MLS is supported for SMC enabled TCP connections with one exception. If a TCP connection requires MLS packet tagging (that is, if the source and destination zones are both SYSMULTI), then SMC will be disabled for that connection. If an SMC enabled TCP connection is already established and a dynamic configuration update introduces a requirement for MLS packet tagging on that connection, it will be dropped.

For more information on Multilevel Security, refer to *z/OS Communications Server IP Configuration Guide Version 2 Release 1 (SC27-3650)*, Chapter 4.

Summary

SMC-R and SMC-D are powerful features of z/OS that provide significant improvements in terms of CPU utilization and latency. z/OS Communications Server also provides a rich set of network security functions that function at varying levels with SMC connections. For application data protection, TLS/SSL and SSH remain strong choices even with SMC enabled. IPsec, on the other hand does not work with SMC. If a connection is configured to use IPsec, z/OS makes it ineligible for SMC. In general, SMC-R connections should only be enabled on LAN segments whose physical security measures provide enough trust to negate the need for IPsec protection. Finally, the applicability of other z/OS Communications server security features depends on the layer at which the features are enforced for network traffic. Those that are closely tied to the network layer will generally not apply to SMC, but those that are enforced by the transport layer or above often will apply.

It is important to understand how each of these security functions affects the underlying use of SMC in your TCP/IP applications. We hope that this paper is helpful in establishing that understanding.

Acknowledgments and Contributions

This paper was a collaborative effort. Thanks to the following individuals for their contributions to this paper.

- Jerry Stevens, IBM Raleigh NC, USA
- Gus Kassimis, IBM Raleigh NC, USA
- Alexandra Winter, IBM Boeblingen, Germany
- Mike Fox, IBM Raleigh NC, USA
- Linwood Overby, IBM Raleigh NC, USA

About the Author:

Chris Meyer, CISSP is a Senior Software Engineer in IBM Software Group's Enterprise Networking Solutions design team, focusing on communications security. He has over 30 years of experience developing IBM operating systems and security-related software products. Chris can be reached at meyerchr@us.ibm.com.

IBM Systems
Route 100
Somers, New York 10589
U.S.A.
Produced in the United States of America,
04/2016

ZSW03255USEN-02