z/OS Communications Server

Communications Server support for OSA-Express7S 25 GbE features

Version 2 Release 3

Note:
Links to related publications are from original documents and might not work. The links to publications are included for reference purposes only.

Contents

Tables

Conventions and terminology that are used in this information

Commands in this information that can be used in both TSO and z/OS® UNIX environments use the following conventions:

- When describing how to use the command in a TSO environment, the command is presented in uppercase (for example, NETSTAT).
- When describing how to use the command in a z/OS UNIX environment, the command is presented in bold lowercase (for example, **netstat**).
- When referring to the command in a general way in text, the command is presented with an initial capital letter (for example, Netstat).

All the exit routines described in this information are *installation-wide exit routines*. The installation-wide exit routines also called installation-wide exits, exit routines, and exits throughout this information.

The TPF logon manager, although included with VTAM®, is an application program; therefore, the logon manager is documented separately from VTAM.

Samples used in this information might not be updated for each release. Evaluate a sample carefully before applying it to your system.

Note: In this information, you might see the following Shared Memory Communications over Remote Direct Memory Access (SMC-R) terminology:

- RoCE Express®, which is a generic term representing IBM® 10 GbE RoCE Express, IBM 10 GbE RoCE Express2, and IBM 25 GbE RoCE Express2 feature capabilities. When this term is used in this information, the processing being described applies to both features. If processing is applicable to only one feature, the full terminology, for instance, IBM 10 GbE RoCE Express will be used.
- RoCE Express2, which is a generic term representing an IBM RoCE Express2® feature that might operate in either 10 GbE or 25 GbE link speed. When this term is used in this information, the processing being described applies to either link speed. If processing is applicable to only one link speed, the full terminology, for instance, IBM 25 GbE RoCE Express2 will be used.
- RDMA network interface card (RNIC), which is used to refer to the IBM® 10 GbE RoCE Express, IBM® 10 GbE RoCE Express2, or IBM 25 GbE RoCE Express2 feature.
- Shared RoCE environment, which means that the "RoCE Express" feature can be used concurrently, or shared, by multiple operating system instances. The feature is considered to operate in a shared RoCE environment even if you use it with a single operating system instance.

Clarification of notes

Information traditionally qualified as Notes is further qualified as follows:

Attention
: Indicate the possibility of damage

Guideline
: Customary way to perform a procedure

Note Supplemental detail

Rule Something you must do; limitations on your actions

Restriction
Indicates certain conditions are not supported; limitations on a product or facility

Requirement
Dependencies, prerequisites

Result Indicates the outcome

Tip Offers shortcuts or alternative ways of performing an action; a hint

Chapter 1. New Function Summary

Communications Server support for OSA-Express7S 25 GbE features

z/OS V2R3 Communications Server, with TCP/IP APAR PI95703 and SNA APAR OA55256, is enhanced to support the OSA-Express7S feature with 25 GbE bandwidth.

To enable Communications Server support for OSA-Express7S 25 GbE features, complete the appropriate tasks in Table 1.

Table 1. Task topics to enable Communications Server support for OSA-Express7S 25 GbE features

Task	Reference
Display the generation level and speed for an active OSA-Express7S QDIO interface by issuing the **DISPLAY TCPIP,,OSAINFO** command.	DISPLAY TCPIP,,OSAINFO in z/OS Communications Server: IP System Administrator's Commands
Display the interface speed value for an active OSA-Express7S QDIO interface by issuing the **Netstat DEvlinks/-d** command.	Netstat DEvlinks/-d report in z/OS Communications Server: IP System Administrator's Commands
Display the read storage value for an active OSA-Express7S QDIO interface by issuing the **Netstat DEvlinks/-d** command.	Netstat DEvlinks/-d report in z/OS Communications Server: IP System Administrator's Commands
Display the read storage value for an active OSA-Express7S QDIO data device by issuing the **D TRL, TRLE=***trle* command.	DISPLAY TRL command in z/OS Communications Server: SNA Operation
Determine the amount of fixed storage that will be allocated for each OSA-Express QDIO interface.	Fixed storage considerations for OSA-Express interfaces in QDIO mode in z/OS Communications Server: IP Configuration Guide
Consider whether to increase the FIXED MAX setting in your IVTPRM00 parmlib member.	Fixed maximum storage for CSM buffers in z/OS Communications Server: IP Configuration Guide

To find all related topics about Communications Server support for OSA-Express7S 25 GbE features, see Table 2.

Table 2. All related topics about Communications Server support for OSA-Express7S 25 GbE features

Book name	Topics
IP Configuration Guide	• Fixed storage considerations for OSA-Express interfaces in QDIO mode • Fixed maximum storage for CSM buffers • Additional fixed storage for OSA interfaces using 8 MB of read storage
IP System Administrator's Commands	• DISPLAY TCPIP,,OSAINFO • Netstat DEvlinks/-d report • Reply field descriptions
SNA Operations	• DISPLAY TRL command

Chapter 2. IP Configuration Guide

Fixed storage considerations for OSA-Express interfaces in QDIO mode

OSA-Express7S with 25 GbE bandwidth

The OSA-Express7S feature has 25 GbE bandwidth compared to a maximum of 10 GbE for earlier generation OSAs. The z/OS Communications Server default value for read storage is insufficient to accommodate a high-volume workload over a 25 GbE adapter and could result in packet loss and sub-optimal performance.

Obtaining 8 MB of fixed read storage

The largest amount of read storage that can be allocated for an OSA-Express QDIO interface is 8 MB.

Guideline: If you expect a high-volume workload over an interface with at least 25 GbE of bandwidth, you should run with 8 MB of read storage to achieve optimal performance.

The only way to obtain 8 MB of fixed CSM 4K HVCOMMON read storage for an OSA-Express QDIO interface which has at least 25 GbE of bandwidth is to specify a VTAM start option of QDIOSTG=126 and to specify (or default) the READSTORAGE value on the OSA INTERFACE statement to GLOBAL. This will make 8 MB the default for all OSAs. To get a lower value for a specific OSA interface, you will need to configure a value other than GLOBAL on the READSTORAGE parameter on the INTERFACE statement for that OSA.

Tip: To see the amount of read storage in MB being used for a specific QDIO interface, use the **`Netstat DEvlinks/-d`** command.

Additional fixed storage for OSA interfaces using 8 MB of read storage

For OSA interfaces which have a bandwidth of at least 10 GbE and which are using 8 MB of read storage, z/OS Communications Server allocates additional fixed CSM 4K HVCOMMON storage for work element processing. If the interface is not using IWQ, the additional storage is 96 KB (plus another 32 KB if QDIO Accelerator is enabled). If the interface is using IWQ, the additional storage is 224 KB (plus another 96 KB if QDIO Accelerator is enabled).

Fixed maximum storage for CSM buffers

The IVTPRM00 parmlib member defines parameters for CSM. The default value of the FIXED MAX setting is 200 MB. This setting defines the maximum amount of storage that is dedicated to fixed CSM buffers.

Tip: If you use OSA-Express interfaces with at least 25 GbE of bandwidth, you should consider increasing the FIXED MAX setting in your IVTPRM00 parmlib member to 512 MB. If you increase this setting, you should also review your

HVCOMMON setting in IEASYSxx and determine whether you need to increase this value to account for the fact that the system reserves a larger portion of this storage for CSM buffers.

OSA configuration recommendations

If your configuration includes OSA-Express QDIO interfaces with different speeds (for example 10 GbE and 25 GbE), you should consider how you plan to use these interfaces in your network configuration.

Guideline: In general, you should avoid mixing interfaces with different speeds (especially when you expect a heavy workload over these interfaces) in the following configurations:

- In the same multipath group
- In the same LAN group
- In a sysplex distributor environment where traffic reaches the distributor stack over a higher bandwidth OSA and is distributed to the target stack over a lower bandwidth OSA

Tip: You can display the speed of an active OSA-Express QDIO interface by using either **`Netstat DEvlinks/-d`** or the **`DISPLAY TCPIP,,OSAINFO`** command.

Chapter 3. IP System Administrator's Commands

DISPLAY TCPIP,,OSAINFO

Use the DISPLAY TCPIP,,OSAINFO command to retrieve information for active IPAQENET and IPAQENET6 interfaces. An interface represents a single datapath device of an OSA-Express feature. The information is retrieved directly from the OSA-Express feature.

Tips:

- If you have an INTERFACE and a DEVICE or LINK definition with the same port name and both are active, specifying either the INTERFACE or link name on the command will generate a report with both IPv4 and IPv6 information.
- You can use the DISPLAY TCPIP,,OSAINFO command to retrieve information for an active OSAENTA interface although only the base portion of the report is pertinent.

The command output provides the following sections of information:

Base Contains physical characteristics and attributes for the interface and OSA-Express feature.

Registered addresses
Contains the Layer 2 MAC addresses or Layer 3 unicast and multicast IPv4 and IPv6 addresses registered to the OSA-Express feature.

QDIO inbound workload queueing routing variables
If QDIO inbound workload queueing is in effect for the interface, this section contains the routing variables for the ancillary input queues. Routing variables identify which inbound packets are to be presented on an ancillary input queue. For more information about ancillary input queues, see QDIO inbound workload queueing in z/OS Communications Server: IP Configuration Guide.

Both the modifiers and the MAX parameters can be used to limit the number of output lines.

Reply field descriptions

Interface
Interface name from the display command.

Base section:
The Base section of the report is displayed if the BASE modifier is specified or none of the modifiers are specified.

PortName
Portname specified on the INTERFACE definition, specified as the device name, or both when the datapath device is shared by both definitions. This name also matches *portname* on the VTAM TRLE definition.

PortNum
Physical port on the OSA-Express that is used for the interface.

Datapath
Hexadecimal datapath device address on the OSA-Express that is used for the interface.

RealAddr
Hexadecimal logical address and unit address of the interface.

PCHID
Physically installed channel path that is used by this QDIO datapath device.

CHPID
Channel path identifier that is used by this QDIO datapath device.

CHPID Type
The CHPID type of the interface, which can have the following values:

OSD External network

OSM Intra node management network

OSX Intra ensemble data network

OSA code level
OSA-Express processor code level of the QDIO datapath device.

Gen Generation of the OSA-Express feature. The following values are supported:

OSA-E3
OSA-Express3

OSA-E4S
OSA-Express4S

OSA-E5S
OSA-Express5S

OSA-E6S
OSA-Express6S

OSA-E7S
OSA-Express7S

Active speed/mode
Switch speed and duplex mode of the interface. The following values are supported:

- 10 mb/sec half duplex
- 10 mb/sec full duplex
- 100 mb/sec half duplex
- 100 mb/sec full duplex
- 1000 mb/sec half duplex
- 1000 mb/sec full duplex
- 10 gigabit/sec full duplex
- 25 gigabit/sec full duplex
- Unknown

Media Transmission media (copper or fiber). If fiber is the transmission media, it can be single-mode fiber (LR/LX) or multimode (SR/SX). The following values are supported:

- Copper
- Multimode Fiber
- Single-mode Fiber

Jumbo frames
Indicates whether jumbo frames are supported.

Isolate
Indicates whether this TCP/IP stack is prohibited from communicating directly through the interface with other TCP/IP stacks that are sharing the OSA-Express feature.

PhysicalMACAddr
Physical Medium Access Control (MAC) LAN address for the interface.

LocallyCfgMACAddr
Local Medium Access Control (MAC) LAN address for the interface.

Queues defined

Out Number of output priority queues that are defined for this interface.

In Number of input queues that are defined for this interface. A value of 1 indicates only the primary queue is defined. A value larger than 1 indicates that QDIO inbound workload queueing ancillary queues are defined and the number of ancillary queues is 1 less than the value reported.

Ancillary queues in use
Number of QDIO inbound workload queueing ancillary input queues (SYSDIST, BULKDATA, and so on) in use by this interface.

Connection Mode
Connection mode of the interface. The following values are supported:

- Layer 2
- Layer 3

IPv4 Indicates whether an IPv4 interface is active for the datapath device.

IPv6 Indicates whether an IPv6 interface is active for the datapath device.

SAPSup
Information used for problem analysis by IBM support.

SAPEna
Information used for problem analysis by IBM support.

IPv4, IPv6, or Layer 2 attributes
This section displays the attributes for the interface.

VLAN ID
Decimal virtual LAN identification number that is defined on this interface.

VMAC Active
Indicates whether the interface is using a virtual MAC address.

Defined Router
The defined router attribute. This field is displayed for Layer 3 only when VMAC is not active. The following values are supported:

Pri The interface is a primary router.

Sec The interface is a secondary router.

Non The interface is not a router.

Active Router
Indicates whether this interface is the active router for the OSA-Express feature. This field is displayed for Layer 3 only when VMAC is not active and is applicable only for PRIROUTER and SECROUTER interface configurations.

VMAC Addr
Displays the virtual MAC address in use for this interface. This field is displayed only when VMAC is active.

VMAC Origin
Indicates the origin of the virtual MAC address. This field is displayed only when VMAC is active. The following values are supported:

Cfg The virtual MAC address was configured in the TCP/IP stack PROFILE

OSA The virtual MAC address was assigned by the OSA-Express

VMAC Router
This field is displayed for Layer 3 only when VMAC is active. The following values are supported:

All Indicates that the OSA-Express is routing everything that was received on the virtual MAC address to the interface without regard for registered addresses.

Local Indicates that the OSA-Express is routing everything that is received on the virtual MAC address, and to a registered IP address, to the interface.

AsstParms
This field is displayed only for Layer 3. It is information used for problem analysis by IBM support.

OutCkSumEna
This field is displayed only for Layer 3. It is information used for problem analysis by IBM support.

InCkSumEna
This field is displayed only for Layer 3. It is information used for problem analysis by IBM support.

Registered Addresses
This is the registered addresses section of the report and is displayed if the REGADDRS modifier is specified or none of the modifiers are specified and only if there are registered addresses.

For Layer 3, there are four subsections that are included only if there are addresses to report:

- IPv4 Unicast Addresses
- IPv4 Multicast Addresses
- IPv6 Unicast Addresses
- IPv6 Multicast Addresses

For Layer 2, there is only one subsection which is included only if there are addresses to report:

- Layer 2 Multicast MAC Addresses

ARP
Indicates whether the OSA-Express is providing address resolution for the corresponding registered IPv4 address.

Addr
IPv4 or IPv6 address.

Total number of IPv4 addresses or Total number of IPv6 addresses
Shows the cumulative number of IPv4 or IPv6 addresses immediately preceding this message.

MAC
The Medium Access Control (MAC) LAN address corresponding to the Layer 2 or registered multicast IP address.

Total number of Layer 2 MAC addresses
Shows the cumulative number of MAC addresses immediately preceding this message.

Ancillary Input Queue Routing Variables

The Ancillary Input Queue Routing Variables section of the report is displayed if any of the following modifiers were specified, or none were specified:

- BULKDATA
- SYSDIST

Queue Type
Displays the workload name for an ancillary queue. The following values are supported:

BULKDATA
Specifies that the input queue is used for streaming workloads.

EE Specifies that the input queue is used for enterprise extender workloads.

IPSEC
Specifies that the input queue is used for IP security workloads.

SYSDIST
Specifies that the input queue is used for sysplex distributor workloads.

Queue ID
Ancillary queue number.

Protocol
TCP

Src
Source address and port. This information is displayed only for the BULKDATA queue.

Dst
Destination address and port. This information is displayed only for the BULKDATA queue.

Total number of IPv4 connections or Total number of IPv6 connections
Displays the cumulative number of BULKDATA IPv4 or IPv6 Src/Dst combinations immediately preceding this message.

Chapter 4. z/OS Migration

Determine the storage impact if QDIOSTG=126 is in effect

Description

If you specify QDIOSTG=126 in your VTAM start options, each OSA-Express QDIO interface which uses the default READSTORAGE setting of GLOBAL on the INTERFACE or LINK statement gets 8 MB of fixed CSM for read storage. For any OSA-Express QDIO interfaces which have a bandwidth of at least 10 GbE and which are using 8 MB of read storage, z/OS allocates additional fixed CSM 4K HVCOMMON storage for work element processing.

Table 3 provides more details about this migration action. Use this information to plan your changes to the system.

Table 3. Information about this migration action

Element or feature:	Communications Server.
When change was introduced:	z/OS V2R3 with TCP/IP APAR PI95703 and SNA APAR OA55256.
Applies to migration from:	z/OS V2R3 without TCP/IP APAR PI95703 and SNA APAR OA55256.
Timing:	Before IPLing of z/OS V2R3 with TCP/IP APAR PI95703 and SNA APAR OA55256.
Is the migration action required?	Yes.
Target system hardware requirements:	None.
Target system software requirements:	None.
Other system (coexistence or fallback) requirements:	None.
Restrictions:	None.
System impacts:	The system allocates additional fixed CSM HVCOMMON storage for work element processing for each OSA-Express QDIO interface which is using 8 MB of read storage.
Related IBM Health Checker for z/OS check:	None.

Steps to take

See *Additional fixed storage for OSA interfaces using 8 MB of read storage* in *z/OS Communications Server: IP Configuration Guide* to understand how much additional storage z/OS will allocate for work element processing. If you do not want the system to allocate this extra storage for a specific interface, you can update your INTERFACE or LINK statement to specify a READSTORAGE value other than GLOBAL.

Reference information

- *Additional fixed storage for OSA interfaces using 8 MB of read storage* in *z/OS Communications Server: IP Configuration Guide*.

- *INTERFACE - IPAQENET OSA-Express QDIO interfaces statement* and *INTERFACE - IPAQENET6 OSA-Express QDIO interfaces statement* in *z/OS Communications Server: IP Configuration Reference*.

Printed in USA