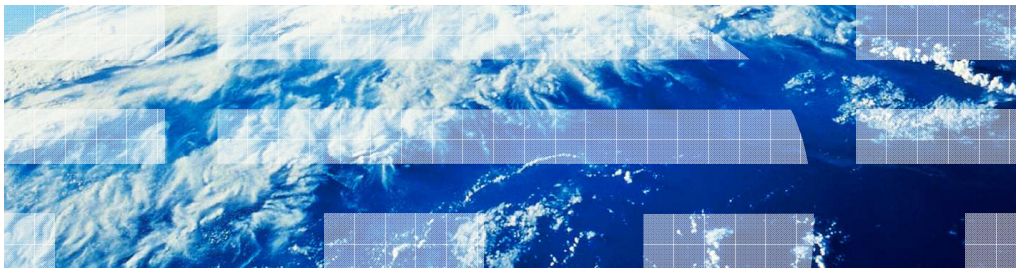
z/OS Communications Server

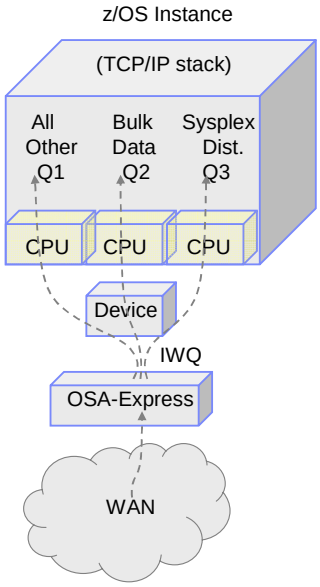
Performance improvements



This presentation describes the new functions in z/OS® V1R13 Communications Server for performance improvements, called the economics and platform efficiency theme. These enhancements include performance improvements for Enterprise Extender, IPv6 checksum and segmentation offload, and support for additional VLANs. In addition to these changes, z/OS Communications Server has an ongoing mandate to reduce path length and increase throughput each release, and V1R13 includes a variety of improvements to reach this goal.

Performance improvements for Enterprise Extender traffic: Background

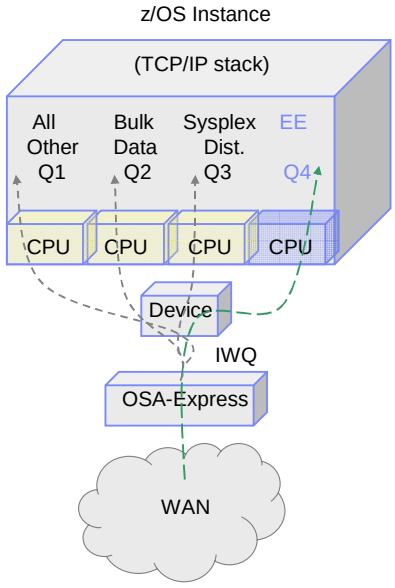
- OSA separates inbound packets and routes them over three different read queues on the same interface
  - Bulk Data traffic (for example FTP)
  - Sysplex Distributor traffic
  - All other traffic
- Each inbound queue can be serviced concurrently by a separate processor



Beginning in V1R12, QDIO inbound workload queueing (IWQ) allows OSA to presort sysplex distributor traffic, bulk data traffic and all other traffic and route it separately to z/OS Communications Server. The traffic is then placed on unique ancillary input queues (AIQs). With IWQ enabled in z/OS Communications Server, these queues can be processed concurrently and independently.

Performance improvements for Enterprise Extender traffic

- New ancillary input queue for EE traffic
- EE traffic serviced on its own processor
- Processing of the EE queue is optimized since the only traffic on the queue is EE
- Both IPv4 and IPv6 are supported



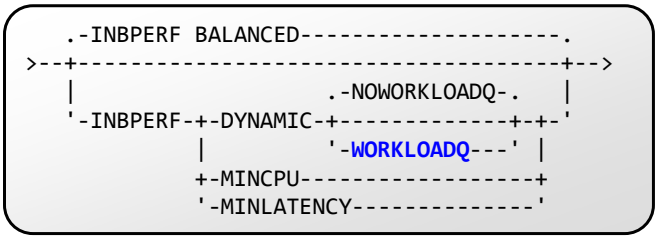
3 Performance improvements

With z/OS V1R13 Communications Server, inbound traffic separation for EE traffic is now supported using a new ancillary input queue. TCP/IP will register with OSA routing variables (RVs) designating the EE traffic to be received on the new AIQ. The OSA-Express Data Router function routes traffic to the correct queue. Since each ancillary input queue can be serviced by a separate process, EE traffic is serviced by its own process. Also, by separating the EE traffic, the non-EE traffic on the primary queue can be serviced sooner.

Both IPv4 and IPv6 are supported for EE traffic.

Performance improvements for EE traffic: Configuration

- Existing INBPERF DYNAMIC WORKLOADQ enables QDIO inbound workload queueing for all supported queues
- Will now also enable inbound workload queueing for EE if HPR=RTP and OSA support



- IPAQENET and IPAQENET6 INTERFACE statements only, no support for DEVICE/LINK definitions!
- INBPERF DYNAMIC WORKLOADQ requires VMAC

The QDIO inbound workload queueing function is enabled with the existing INBPERF DYNAMIC WORKLOADQ setting on IPAQENET and IPAQENET6 INTERFACE statements. That keyword will enable workload queueing support for all function types supported by the OSA-Express and requested by z/OS Communications Server. If the OSA-Express supports inbound workload queueing for only Sysplex Distributor and streaming workloads (BULKDATA), z/OS Communications Server will support just those two. If the OSA-Express also supports inbound workload queueing for EE, and VTAM® is capable of EE support (start option HPR=RTP is specified), z/OS Comm Server will support all three types.

WORKLOADQ is not supported for INBPERF DYNAMIC on IPAQENET LINK statements. For steps to convert from IPv4 IPAQENET DEVICE, LINK, and HOME definitions to the IPv4 IPAQENET INTERFACE statement see the *z/OS Communications Server: IP Configuration Guide*.

The VMAC parameter is also required and can be specified with or without *macaddr*.

For more information, see the IPAQENET INTERFACE and IPAQENET6 INTERFACE statements in *z/OS Communications Server: IP Configuration Reference*.

Performance improvements for EE traffic: Netstat DEvlinks/-d

```
D TCPIP,TCPCS1,NETSTAT,DEVLINKS,INTFNAME=QDIO4101L
EZD0101I NETSTAT CS V1R13 TCPCS1
INTFNAME: QDIO4101L          INTFTYPE: IPAQENET   INTFSTATUS: READY
    PORTNAME: QDIO4101  DATAPATH: 0E2A     DATAPATHSTATUS: READY
    CHPIDTYPE: OSD
    SPEED: 0000001000
...
    READSTORAGE: GLOBAL (4096K)
    INBPERF: DYNAMIC
      WORKLOADQUEUEING: YES
    CHECKSUMOFFLOAD: YES
    SECCLASS: 255                     MONSYSPLEX: NO
    ISOLATE: NO                       OPTLATENCYMODE: NO
...
1 OF 1 RECORDS DISPLAYED
END OF THE REPORT
```


The Netstat DEvlinks/-d report is unchanged; it already indicates whether IPAQENET and IPAQENET6 interfaces have enabled WorkloadQueueing. The possible values are Yes, indicating the function is enabled, No, indicating the function is not enabled, and Unsupported, indicating that the OSA-Express does not support the function. This field is not displayed when InbPerf is not Dynamic or for IPAQENET DEVICE/LINK definitions.

For more information, see the Netstat DEvlinks/-d report in *z/OS Communications Server: IP System Administrator's Commands*.

This existing information can also be retrieved with a network management application that has been updated to use information returned by the GetIfs callable NMI.

Performance improvements for EE traffic: Display OSAINFO

```
D TCPIP,TCPCS2,OSA,INTFN=QDIO4102I
EZZ0053I COMMAND DISPLAY TCPIP,,OSAINFO COMPLETED SUCCESSFULLY
EZD0031I TCP/IP CS V1R13  TCPIP Name: TCPCS2    13:05:20 881
Display OSAINFO results for IntfName: QDIO4102I
PortName: QDIO4102  PortNum: 00  Datapath: 0E32   RealAddr: 0E32
.
.
Ancillary Input Queue Routing Variables:
  Queue Type: SYSDIST   Queue ID:  3  Protocol: TCP
    Addr: 10.91.2.2
    Total number of IPv4 addresses:      1
  Queue Type: EE        Queue ID:  4  Protocol: UDP
    Dst: 10.81.2.2..12000
    Dst: 10.81.2.2..12001
    Dst: 10.81.2.2..12002
    Dst: 10.81.2.2..12003
    Dst: 10.81.2.2..12004
    Total number of IPv4 addresses and ports:      5
38 of 38 lines displayed
End of report
```


The DISPLAY TCPIP,,OSAINFO command is used to retrieve information for active IPAQENET and IPAQENET6 interfaces. The Ancillary Input Queue Routing Variables section of the OSAINFO now shows the routing variables that are registered for the EE queue when QDIO inbound workload queueing is enabled.

Modifiers are supported for this command to specify which types of information about the OSA should be returned. A new modifier has been added to display EE RVs. The default is to display all information pertaining to this OSA.

For more information, see the DISPLAY TCPIP,,OSAINFO report in *z/OS Communications Server: IP System Administrator's Commands*.

Performance improvements for EE traffic: Display TRLE

```
D NET,TRL,TRLE=QDIO101
IST097I DISPLAY ACCEPTED
...
IST2263I PORTNAME = QDIO4101   PORTNUM =   0   OSA CODE LEVEL = ABCD
...
IST1221I DATA  DEV = 0E2A STATUS = ACTIVE     STATE = N/A
IST1724I I/O TRACE = OFF  TRACE LENGTH = *NA*
IST1717I ULPID = TCPCS1
IST2310I ACCELERATED ROUTING DISABLED
IST2331I QUEUE   QUEUE     READ              QUEUE
IST2332I ID      TYPE      STORAGE           STATUS
IST2205I ------  --------  ---------------   ----------------------
IST2333I RD/1    PRIMARY   4.0M(64 SBALS)    ACTIVE
IST2333I RD/2    BULKDATA  4.0M(64 SBALS)    ACTIVE
IST2333I RD/3    SYSDIST   4.0M(64 SBALS)    ACTIVE
IST2333I RD/4    EE        4.0M(64 SBALS)    ACTIVE
...
IST924I -------------------------------------------------------------
IST314I END
```


The output for the Display ID=trlename and Display TRL,TRLE=trlename commands indicates whether the QDIO inbound workload queueing function is in use for the QDIO interface. For each input queue, it includes the queue ID and queue type in addition to the read storage. The queue type is PRIMARY for the primary input queue, BULKDATA for the bulk data AIQ, SYSDIST for the sysplex distributor connection routing AIQ, and EE for the Enterprise Extender AIQ. The read storage value indicates how much storage is reserved for this queue.

The queue status value indicates whether the queue is active, not currently in use by the TCP/IP stack, or not supported by the OSA.

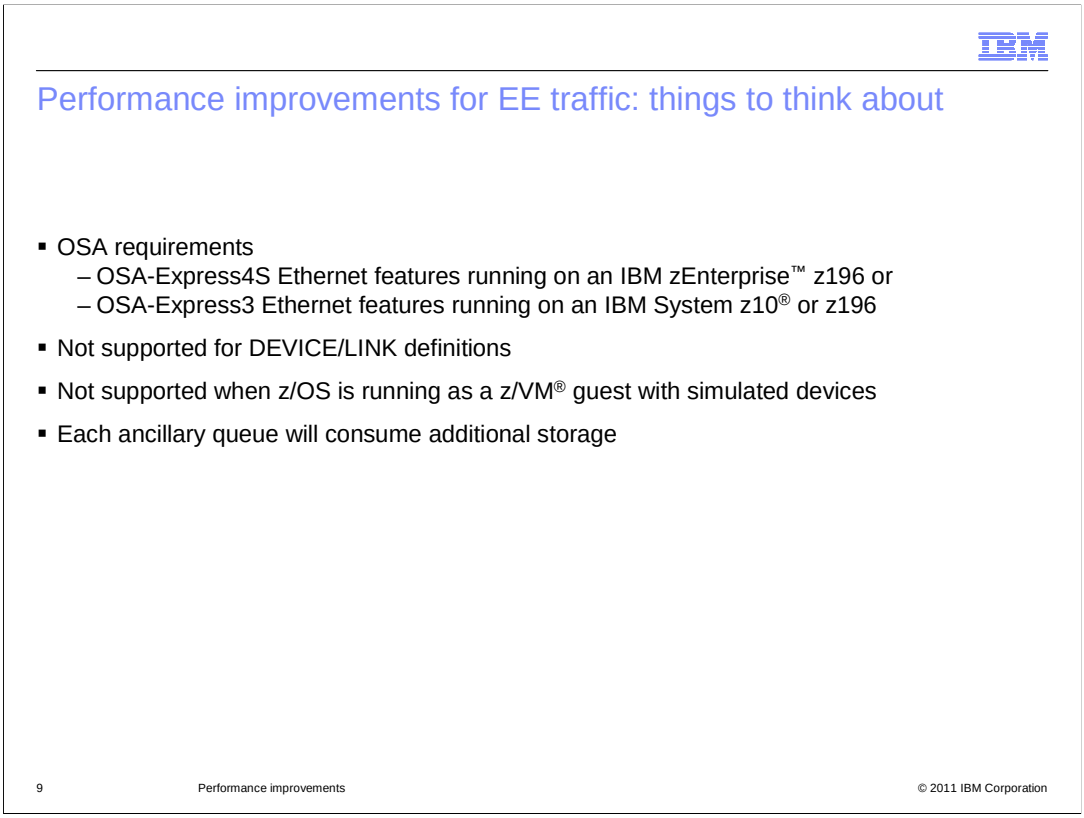
The queue ID and queue type can be used to correlate with VTAM tuning statistics, packet trace, and OSA-Express Network Traffic Analyzer (OSAENTA) trace output for the QDIO interface. For more information, see the DISPLAY ID command and DISPLAY TRL command in *z/OS Communications Server: SNA Operation*.

Performance improvements for EE traffic: Traces

```
7 MVS279   PACKET   00000004 13:32:15.193763 Packet Trace
From Interface    : OSDC1INT4        Device: QDIO Ethernet    Full=95
 Tod Clock        : 2010/08/23 13:32:15.193763                Intfx: 13
 Segment #        : 0                Flags: Adj Tunnel In QID
 Source           : 10.81.5.5
 Destination      : 10.81.1.1
 Source Port      : 12001            Dest Port: 12001 Asid: 0022 TCB: 00000000
 QID              : 4 (EE)
IpHeader: Version : 4                Header Length: 20
 Tos              : C0               QOS: Internetwork Normal Service
 Packet Length    : 95               ID Number: 008F
 Fragment         :                  Offset: 0
 TTL              : 64               Protocol: UDP            CheckSum: 5E98 FFFF
 Source           : 10.81.5.5
 Destination      : 10.81.1.1
UDP
 Source Port      : 12001 (EE-Network) Destination Port: 12001 (EE-Network)
 Datagram Length  : 75               CheckSum: FCD4 FFFF
```


Packet trace records and OSA-Express Network Traffic Analyzer (OSAENTA) trace records now show the ancillary input queue ID (QID) and queue type (EE in this example). The QID is 1 for the primary input queue and 2 or greater for the AIQs. The QID can be used to correlate with VTAM display and tuning statistics output. In addition, the flags field includes a QID flag when an AIQ is used. For more information, see the *z/OS Communications Server: IP Diagnosis Guide.*

Note also that VTAM VIT traces can now include a QID for Enterprise Extender traffic.

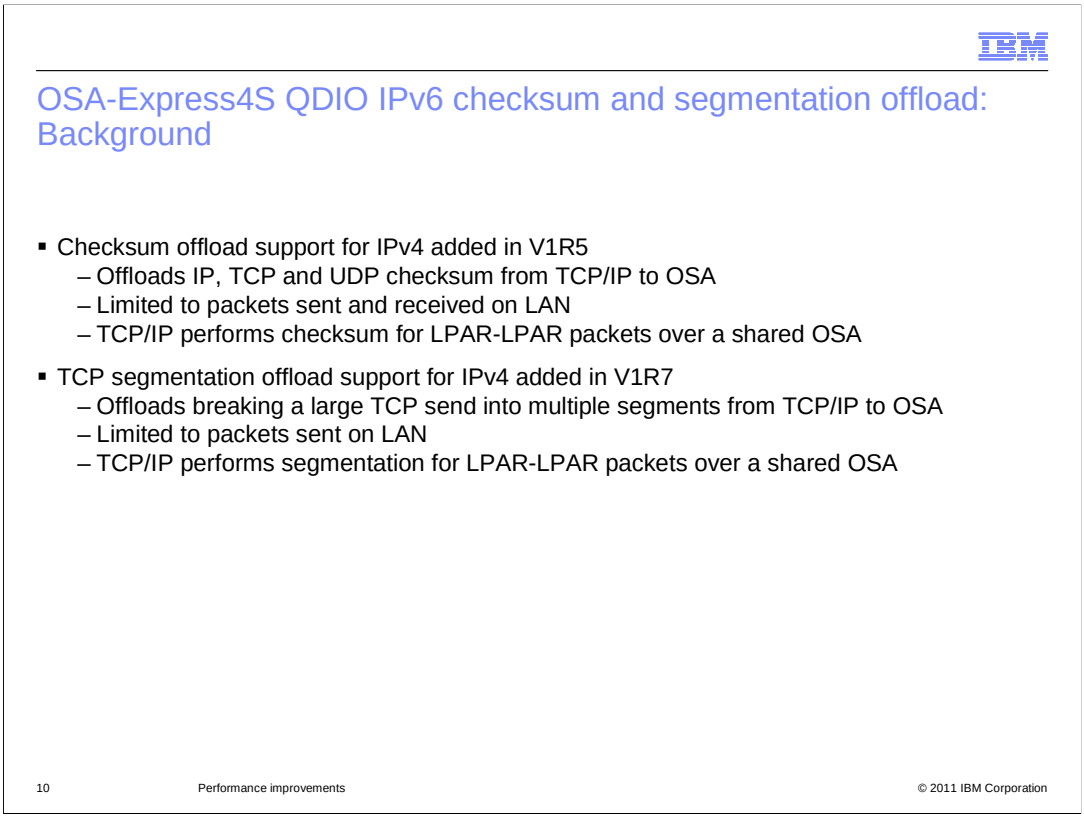


QDIO inbound workload queueing requires an OSA-Express3 (or later) Ethernet feature running in QDIO Mode on an IBM System z10 or later. For more information, see the 2097DEVICE and 2098DEVICE Preventive Service Planning (PSP) buckets.

The QDIO inbound workload queueing function is not supported for IPAQENET interfaces that are defined by using the DEVICE, LINK, and HOME definitions. You must convert your IPAQENET definitions to use the INTERFACE statement to enable this support.

The QDIO inbound workload queueing function is not supported for a z/OS guest on z/VM using simulated (virtual) devices, such as virtual switch (VSWITCH) or guest LAN.

Each ancillary queue increases storage utilization by an amount equal to 36K of fixed ECSA (approximately nine additional 4k pages of ECSA). An additional but tunable amount of fixed CSM data space as specified by the READSTORAGE parameter (64K times the number of SBALs) of fixed CSM 4K data space storage can potentially be needed.



In V1R5, z/OS Communications Server added checksum offload support for OSA-Express in QDIO mode. This function applies to IPv4 TCP and UDP packets both outbound and inbound. It provides a performance benefit and reduces processor utilization by offloading checksum cycles to the OSA.

In V1R7, z/OS Communications Server added segmentation offload support for OSA-Express2 in QDIO mode. This function applies to IPv4 TCP outbound transmissions which are larger than the MSS. It provides a performance benefit and reduces processor utilization by offloading segmentation cycles to the OSA.

Both of these functions are implemented in the NIC portion of the OSA, so they are limited to packets sent and received over the LAN. For packets sent and received to another stack sharing the same OSA, the stack performs the checksum and segmentation functions.

OSA-Express4S QDIO IPv6 checksum and segmentation offload

- OSA-Express4S on the IBM System z196 adds support for:
  - IPv6 checksum offload
  - IPv6 segmentation offload
  - LPAR-LPAR *checksum* offload over a shared OSA (both IPv4 and IPv6)
- These offloads apply to QDIO mode for the OSD and OSX CHPID types - No offloads for OSM (which is Layer 2)
- No OSA-Express4S support for LPAR-LPAR *segmentation* offload over a shared OSA

In V1R13, z/OS Communications Server takes advantage of the offload enhancements in the OSA-Express4S on the IBM System z196 for the OSD and OSX CHPID types when running in QDIO mode. This OSA function supports IPv6 checksum offload and IPv6 segmentation offload. It also supports checksum offload for packets which flow directly between two stacks sharing an OSA for both IPv4 and IPv6 (but not segmentation offload for these packets). On OSA-Express4S, the checksum offload function is performed in the application-specific integrated circuit (ASIC) portion of the OSA rather than the NIC portion. However, when segmentation is also offloaded, the outbound checksum offload is performed in the NIC.

Note that IPv6 checksum offload is not supported for packets with IPv6 extension headers. Checksum for these packets is performed by TCP/IP.

OSA-Express4S QDIO checksum offload: configuration

```
IPCONFIG6
. . .

    .-CHECKSUMOFFLoad----.
  +-+--------------------+--------------+
    '-NOCHECKSUMOFFLoad--'

. . .
```

```
IPCONFIG
. . .

  .-CHECKSUMOFFLoad----.
+-+--------------------+--------------+
  '-NOCHECKSUMOFFLoad--'

. . .
```


A new parameter is added to the IPCONFIG6 statement to control IPv6 checksum offload. This parameter defaults to enabling IPv6 checksum offload.

In earlier releases, IPv4 checksum offload cannot be disabled. A new parameter is now added to the IPCONFIG statement to control IPv4 checksum offload for all OSA devices that support it. This parameter defaults to enabling IPv4 checksum offload.

OSA-Express4S QDIO segmentation offload: Configuration

```
IPCONFIG6
. . .

  .-NOSEGMENTATIONOFFLoad--.
+-+------------------------+--------------+
  '-SEGMENTATIONOFFLoad----'

. . .
```

```
IPCONFIG
. . .

  .-NOSEGMENTATIONOFFLoad--.
+-+------------------------+--------------+
  '-SEGMENTATIONOFFLoad----'

. . .
```

```
GLOBALCONFIG (existing option – deprecated and only applies to IPv4)
. . .

  .-NOSEGMENTATIONOFFLoad--.
+-+------------------------+--------------+
  '-SEGMENTATIONOFFLoad----'

. . .
```


A new parameter is added to the IPCONFIG6 statement to control IPv6 TCP segmentation offload. This parameter defaults to disabling IPv6 TCP segmentation offload.

Previously, IPv4 TCP segmentation offload was controlled using the GLOBALCONFIG statement. The use of the GLOBALCONFIG segmentation offload parameter is deprecated, and a new parameter is added to the IPCONFIG statement to control IPv4 TCP segmentation offload for all OSA devices that support it. This parameter defaults to disabling IPv4 TCP segmentation offload.

OSA-Express4S QDIO IPv6 checksum and segmentation offload: Netstat COnfig/-f

```
IPv6 Configuration Table:
Forwarding: Yes    HopLimit: 00255      IgRedirect: No
...
ChecksumOffload: Yes            SegOffload: Yes
.....................................................................

IP Configuration Table:
Forwarding: Yes    TimeToLive: 00064    RsmTimeOut: 00060
...
ChecksumOffload: Yes            SegOffload: Yes
.....................................................................

Global Configuration Information:
TcpIpStats: Yes    ECSALimit: 2096128K  PoolLimit: 2096128K
...
SegOffload: Yes    SysplexWLMPoll: 060  MaxRecs: 100
```


This is an example report from a Netstat CONFIG/-f command. It shows whether checksum offload and segmentation offload are enabled or disabled for the stack. The segmentation offload field is no longer in the GLOBALCONFIG section because that parameter has been deprecated.

OSA-Express4S QDIO IPv6 checksum and segmentation offload: Netstat DEvlinks/-d

```
EZD0101I NETSTAT CS V1R13 TCPCS1
INTFNAME: QDIO6101L         INTFTYPE: IPAQENET6  INTFSTATUS: READY
    PORTNAME: QDIO6101  DATAPATH: 0E2B     DATAPATHSTATUS: READY
    CHPIDTYPE: OSD
    SPEED: 0000001000
...
    READSTORAGE: GLOBAL (4096K)
    INBPERF: DYNAMIC
      WORKLOADQUEUEING: YES
    CHECKSUMOFFLOAD: YES              SEGMENTATIONOFFLOAD: YES
    SECCLASS: 255                     MONSYSPLEX: NO
    ISOLATE: NO                       OPTLATENCYMODE: NO
...
```


This is an example report from a Netstat DEVLINKS/-d command. It shows whether checksum offload and segmentation offload are enabled or disabled for an *active* OSA-Express QDIO interface. If an offload function is enabled but OSA does not support that function, then the field reflects the value Unsupported.

OSA-Express4S QDIO IPv6 checksum and segmentation offload: Packet trace

- Shows info about IPv6 packets for which segmentation is offloaded
- Also in session report

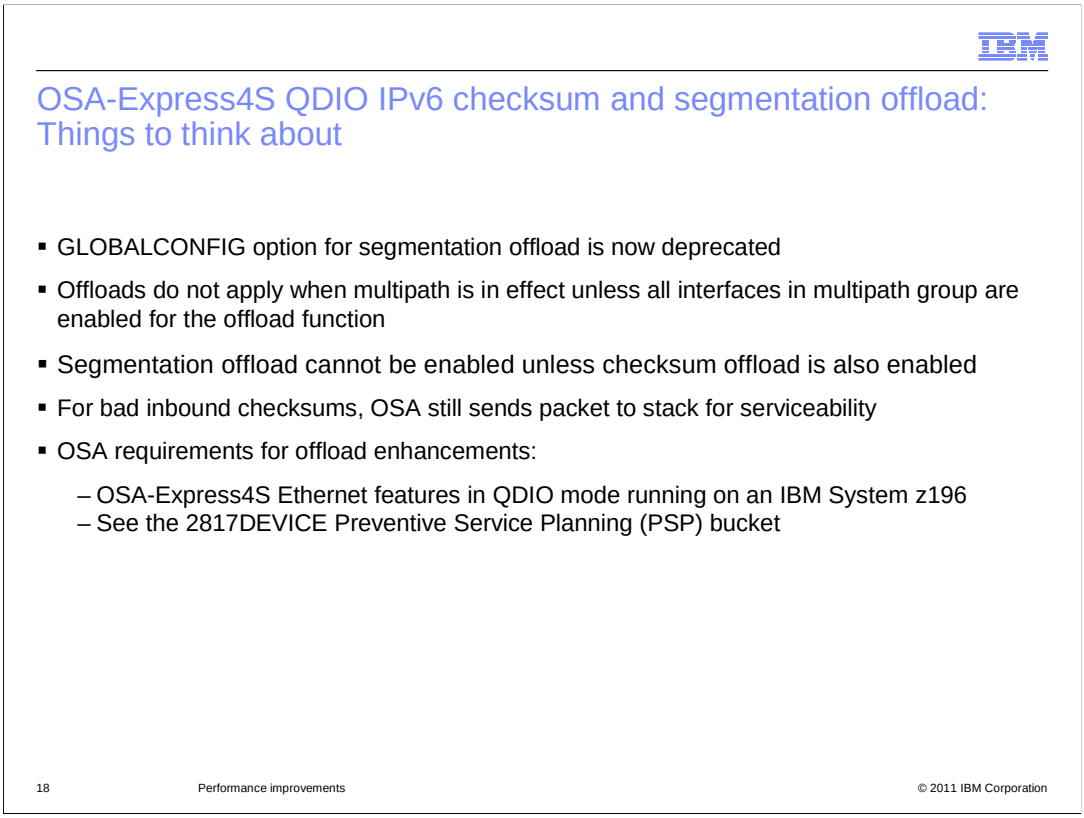
```
IpHeader: Version : 6             Header Length: 40
 Class:           : 00            Flow: 000000
 Offload Length   : 5400
 Hops             : 255           Protocol: TCP
 Source           : 2001::1
...
TCP
 Source Port : 1026 ()           Destination Port: 1026 ()
...
 Window Size : 32768             CheckSum: 120B 0000 Urgent Data Pointer: 0000
 Offload Segments : 4            Length: 1448            Last: 456
.....................................................................

Data Segment Stats:              Inbound,      Outbound
 ...
 Offload Sends:                        3         ( 50%)
 Offload Segments:                     6
 Offload Bytes:                    43616        (72.69%)
 Total Packets(normal + offload):     18        (83.33%)
```


This is an example report from the packet trace formatter which shows an IPv6 send for which segmentation was offloaded. This is similar to the existing report for IPv4 and shows the number of segments to be offloaded and the size of each segment including the last one.

OSA-Express4S QDIO IPv6 checksum and segmentation offload: NMI, SNMP and SMF

- Configured IPCONFIG/IPCONFIG6 settings
  - GetProfile callable NMI
  - SMF 119 subtype 4 record
- Checksum and segmentation offload indicators for an interface
  - GetIfs callable NMI
  - SNMP ibmMvsIfflagType MIB object

This chart summarizes the NMI, SNMP, and SMF externals for the checksum offload and segmentation offload functions. You can use the NMI or SMF to see the configured profile settings. You can use the NMI or SNMP to see whether these offloads are enabled for a given interface.



The GLOBALCONFIG option for enabling IPv4 segmentation offload is now deprecated. You should use IPCONFIG instead.

IPv6 checksum offload and segmentation offload only apply for multipath if all interfaces in the multipath group support the offload function. This is consistent with the existing IPv4 support.

You cannot enable segmentation offload unless you also enable checksum offload.

If OSA detects a bad checksum for an inbound packet, OSA still sends the packet to the stack for serviceability reasons. This way, the stack verifies the checksum and discards the packet with an exception CTRACE. This is also consistent with the existing IPv4 support.

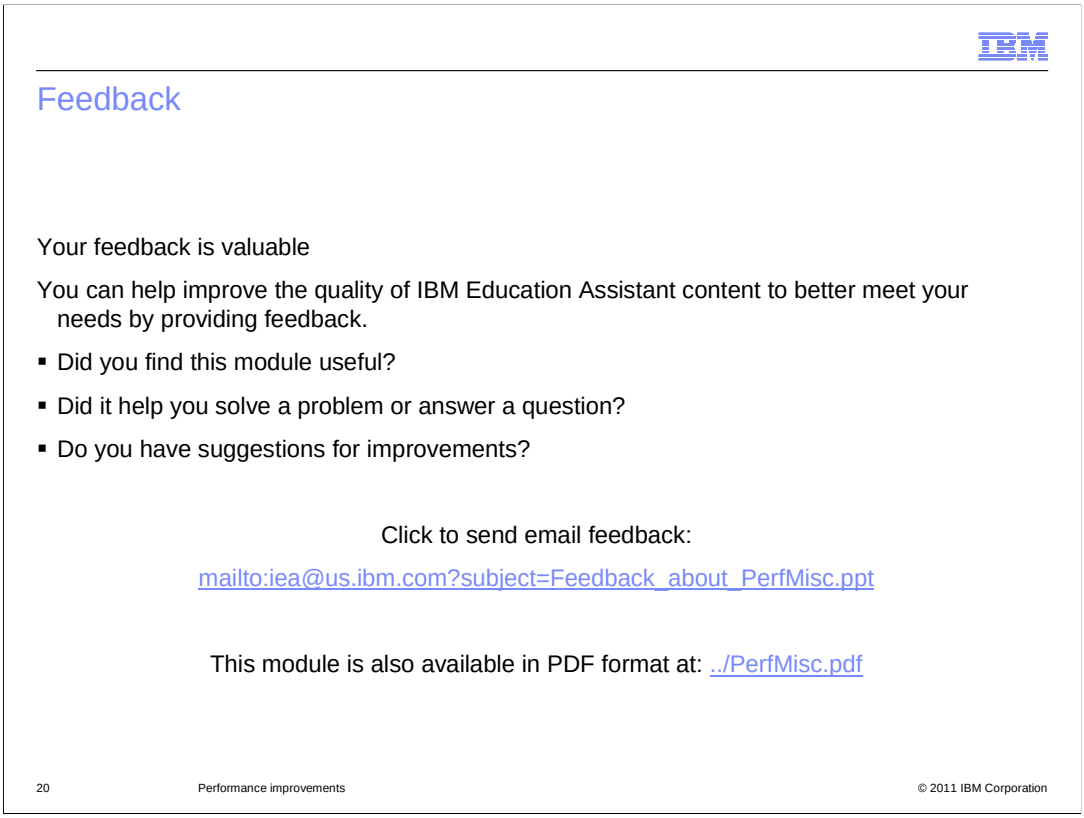
The checksum offload and segmentation offload enhancements are exclusive to the OSA-Express4S in QDIO mode which requires a z196. For more details, see the PSP bucket listed on the chart.

Support for additional VLANs for an OSA-Express QDIO port

- Current limit is up to eight IPv4 and eight IPv6 VLANs per OSA port per stack
  - Requires separate INTERFACE statement and data path device per VLAN
- Raised limit from eight to 32
  - Software changes only; no OSA dependencies required for this function
  - VTAM-generated TRLEs for OSX contain up to 17 datapath devices
    - You can configure a TRLE for OSX if needed

Before V1R13, you can configure only up to eight separate IPv4 and eight separate IPv6 INTERFACE statements for distinct VLANs to the same OSA. Beginning in V1R13, you can configure up to 32 VLANs for IPv4 and 32 VLANs for IPv6 for the same OSA port. This function works for any OSA.

If you configure OSX using the CHPID parameter, then VTAM creates a TRLE with up to 17 datapath devices. If you need more than 17, use the existing option to configure the OSX using the PORTNAME parameter and configure a TRLE with sufficient datapath devices.



You can help improve the quality of IBM Education Assistant content by providing feedback.

Trademarks, disclaimer, and copyright information