



IBM Software Group Enterprise Networking Solutions
z/OS® V1R11 Communications Server

z/OS V1R11 Communications Server – hardware

z/OS Communications Server Development, Raleigh, North Carolina



© Copyright International Business Machines Corporation 2009. All rights reserved.

This presentation will give you an overview of the enhancements to the Communications Server in z/OS V1R11 for virtualization (hardware). The virtualization (hardware) theme in this release of Communications Server is related to OSA.

Virtualization

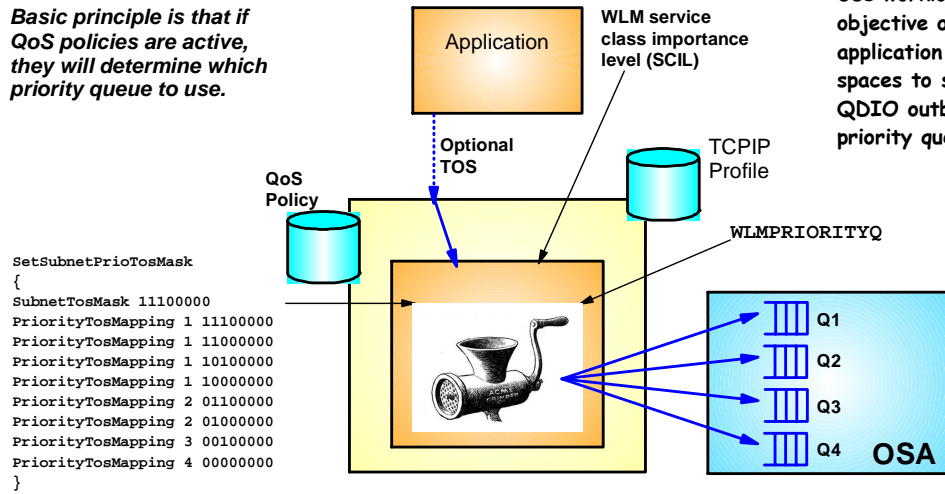
- 🔑 QDIO enhancements for WLM IO priority
- 🔑 QDIO support for OSA interface isolation
- 🔑 OSA-Express3 optimized latency mode

z/OS V1R11 Communications Server enhancements provide a simple mechanism to select QDIO outbound priority queue. The new support allows installations who do not use QoS networking policies to set QDIO outbound queue priority based on WLM input. This support can be used to extend the WLM importance level of the application that sends data to the selection of QDIO outbound queue.

QDIO enhancements for WLM IO priority

Basic principle is that if QoS policies are active, they will determine which priority queue to use.

Use workload objective of application address spaces to select QDIO outbound priority queue



```

SetSubnetPrioTosMask
{
SubnetTosMask 11100000
PriorityTosMapping 1 11100000
PriorityTosMapping 1 11000000
PriorityTosMapping 1 10100000
PriorityTosMapping 1 10000000
PriorityTosMapping 2 01100000
PriorityTosMapping 2 01000000
PriorityTosMapping 3 00100000
PriorityTosMapping 4 00000000
}
    
```

With z/OS V1R11 Communications Server, a new GLOBALCONFIG profile statement establishes a mapping of Work Load Manager (WLM) service class importance levels to outbound QDIO priorities.

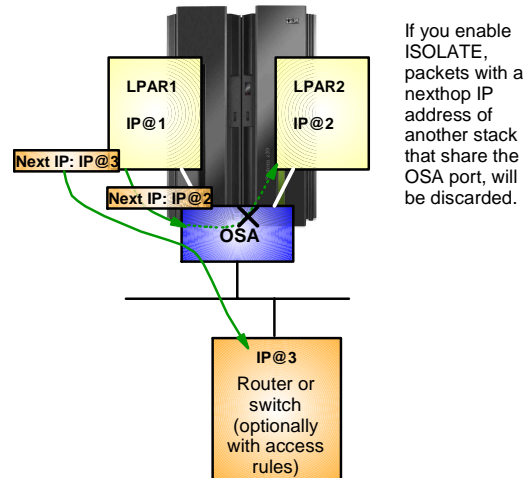
Using the new WLM PRIORITYQ parameter allows you to map outbound OSA-Express data with an IPv4 Type of Service (ToS) byte or IPv6 Traffic Class of zeros.

Taking advantage of already established importance levels eases the extension of this prioritization through z/OS and z/OS Communications Server, using OSA-Express and onto the LAN. In addition, WLM PRIORITYQ allows the outbound priority to be applied to forwarded packets containing a ToS or Traffic Class of zeros.

OSA interface isolation

- New function added to the OSA adapter
 - z/OS Communications Server adds support for this new function in z/OS V1R11
- Allow customers to disable shared OSA local routing functions
 - ISOLATE/NOISOLATE option on QDIO network interface definition
- OSA local routing can in some scenarios be seen as a security exposure
- Depends on OSA MCL update

Be careful using ISOLATE if you use OSPF and share a subnet between stacks that share an OSA port.



If you enable ISOLATE, packets with a nexthop IP address of another stack that share the OSA port, will be discarded.

In some environments where strict control over routing between IP nodes must be enforced, the loop-back feature of a shared OSA port can prevent such rules from being enforced. Assume sending an IP packet from LPAR1 to a home IP address of LPAR2 (without VLAN tagging or attached to the same VLAN). In that case OSA will send that packet up to LPAR2 directly without sending it out to the switch. If the switch is there to enforce access rules, that behavior is an issue.

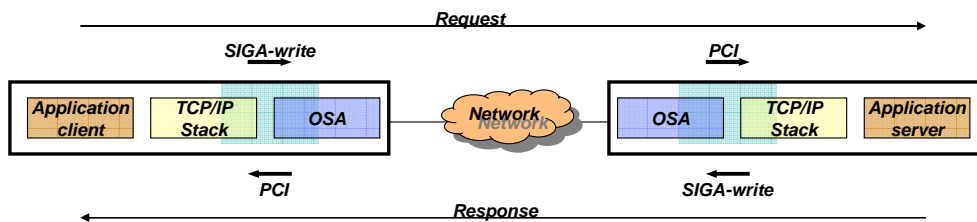
OSA-Express connection isolation provides a way for a stack using an OSA-Express to prevent packets from flowing directly between two stacks sharing the OSA. When connection isolation is in effect, OSA-Express will discard any packets when the next hop address was registered by a sharing stack. OSA-Express requires that both stacks sharing the port be non-isolated for direct routing to occur.

OSA-Express connection isolation is only supported for OSA-Express features in QDIO mode.

OSA-Express connection isolation is not supported when the OSA-Express is defined using a DEVICE and LINK statement.

OSA-Express3 optimized latency mode (OLM)

- OSA-Express3 has significantly better latency characteristics than OSA-Express2
- The z/OS software and OSA microcode can further reduce latency:
 - If z/OS Communications Server knows that latency is the most critical factor
 - If z/OS Communications Server knows that the traffic pattern is not streaming bulk data
- Inbound
 - OSA-Express signals host if data is “on its way” (“Early Interrupt”)
 - Host looks more frequently for data from OSA-Express
- Outbound
 - OSA-Express does not wait for SIGA to look for outbound data (“SIGA reduction”)



OSA-Express3 and z/OS V1R11 Communications Server are allowing configuration of a new mode of operation for an OSA-Express3 in QDIO mode – optimized latency mode (OLM). When in this mode, several processing improvements are being implemented.

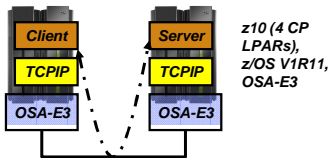
For inbound processing, the OSA signals the host when data is “on its way”. On inbound processing, z/OS V1R11 Communications Server looks more frequently for available data to process, ensuring any new data is read from the OSA-Express3 without requiring another PCI.

For outbound processing, the OSA-Express3 also looks more frequently for available data to process, thus not requiring a Signal Adapter (SIGA) instruction to always know more data is available. OSA-Express3, as with previous generations of OSA-Express, supports four write priority queues. This additional scan for available data is only be done for write priority queue 1.

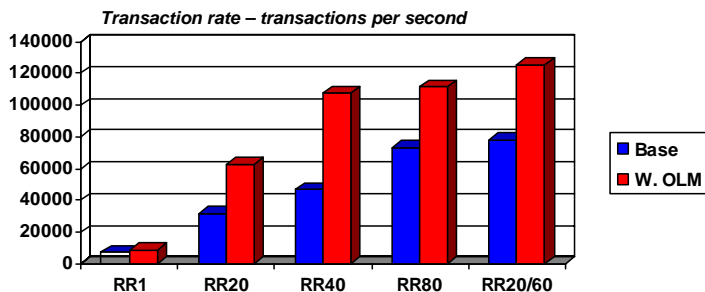
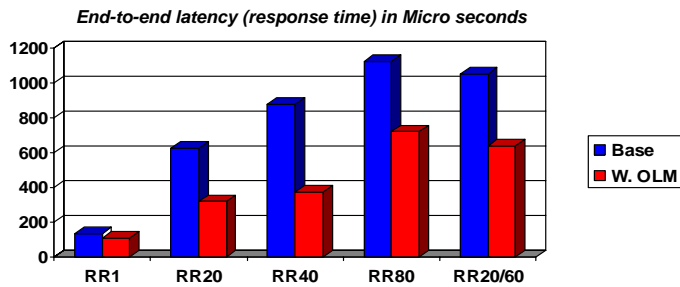
You must define an OSA-Express3 for QDIO mode, using the INTERFACE statement, to allow the OSA-Express3 to operate in optimized latency mode. You also must install the correct OSA-Express Licensed Internal Code (LIC) level. See the appropriate PSP bucket for your level of z10 for the level that supports optimized latency mode.

You should only configure an OSA-Express3 to operate in OLM mode when the traffic over this OSA-Express3 demands the best latency possible. This is typically a high volume, interactive workload. In this type of workload, OLM can significantly improve both latency time per transaction and throughput - the number of transactions processed. Preliminary performance runs, though not officially verified, show total end to end time for one transaction reduced by 17%. More significantly, in a typical customer environment with 20 simultaneous interactive workloads being processed at any given time, average latency was almost cut in half, and throughput improved 95%.

Preliminary performance indications of OLM for interactive workload



- Client and Server have almost no application logic
- RR1 with one session
 - One byte in, one byte out
- RR20 with 20 sessions
 - 128 bytes in, 1024 bytes out
- RR40 with 40 sessions
 - 128 bytes in, 1024 bytes out
- RR80 with 80 sessions
 - 128 bytes in, 1024 bytes out
- RR20/60 with 80 sessions
 - Mix of 100/128 bytes in and 800/1024 out



Note: The performance measurements discussed in this presentation are preliminary z/OS V1R11 Communications Server numbers and were collected using a dedicated system environment. The results obtained in other configurations or operating system environments can vary.

Optimized latency mode is meant for interactive workloads. The maximum number of TCP/IP stacks (LPARs) that can share an OSA adapter operating in OLM mode is limited to four. The general assumption is that customers will use dedicated OSA ports for workloads where they plan to use the OLM capability.

OLM requires a new level of OSA microcode that has not yet (March 2009) been announced.



Trademarks, copyrights, and disclaimers

IBM, the IBM logo, ibm.com, and the following terms are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both:
z/OS

If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of other IBM trademarks is available on the Web at "Copyright and trademark information" at <http://www.ibm.com/legal/copytrade.shtml>

Other company, product, or service names may be trademarks or service marks of others.

Product data has been reviewed for accuracy as of the date of initial publication. Product data is subject to change without notice. This document could include technical inaccuracies or typographical errors. IBM may make improvements or changes in the products or programs described herein at any time without notice. Any statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business. Any reference to an IBM Program Product in this document is not intended to state or imply that only that program product may be used. Any functionally equivalent program, that does not infringe IBM's intellectual property rights, may be used instead.

THE INFORMATION PROVIDED IN THIS DOCUMENT IS DISTRIBUTED "AS IS" WITHOUT ANY WARRANTY, EITHER EXPRESS OR IMPLIED. IBM EXPRESSLY DISCLAIMS ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NONINFRINGEMENT. IBM shall have no responsibility to update this information. IBM products are warranted, if at all, according to the terms and conditions of the agreements (for example, IBM Customer Agreement, Statement of Limited Warranty, International Program License Agreement, etc.) under which they are provided. 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 in connection with this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products.

IBM makes no representations or warranties, express or implied, regarding non-IBM products and services.

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents or copyrights. Inquiries regarding patent or copyright licenses should be made, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. All customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput or performance improvements equivalent to the ratios stated here.

© Copyright International Business Machines Corporation 2009. All rights reserved.

Note to U.S. Government Users - Documentation related to restricted rights-Use, duplication or disclosure is subject to restrictions set forth in GSA ADP Schedule Contract and IBM Corp.