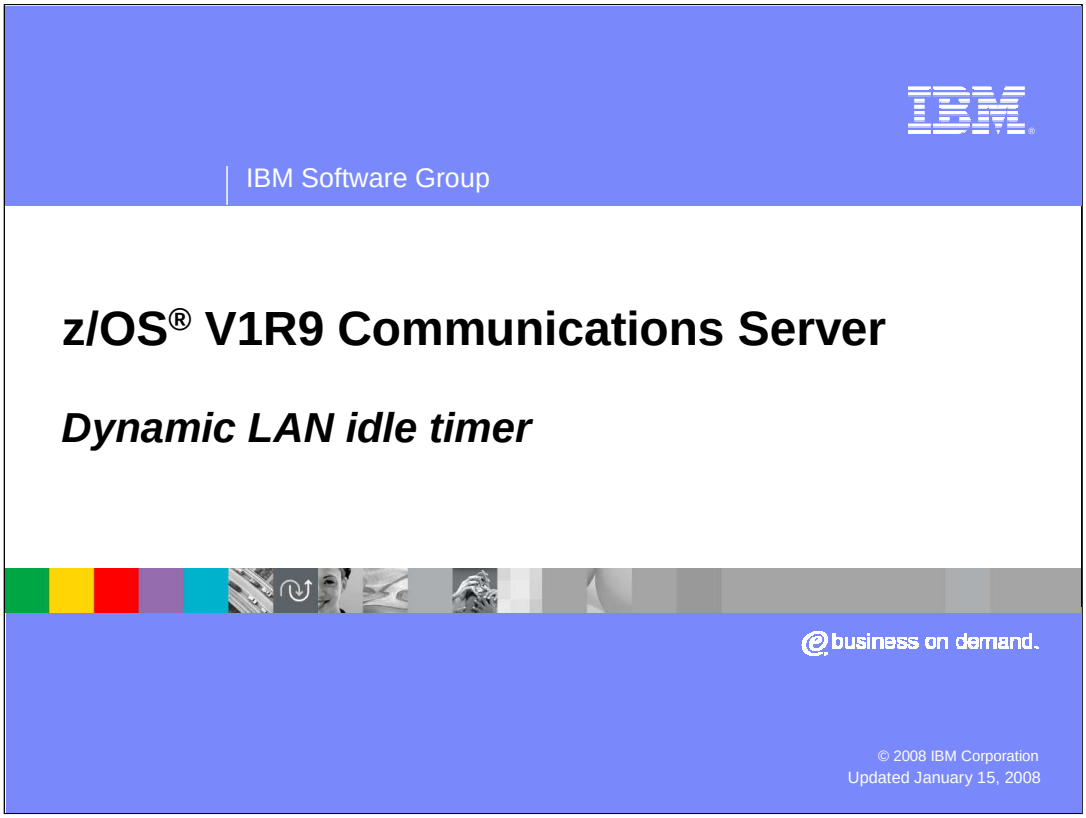
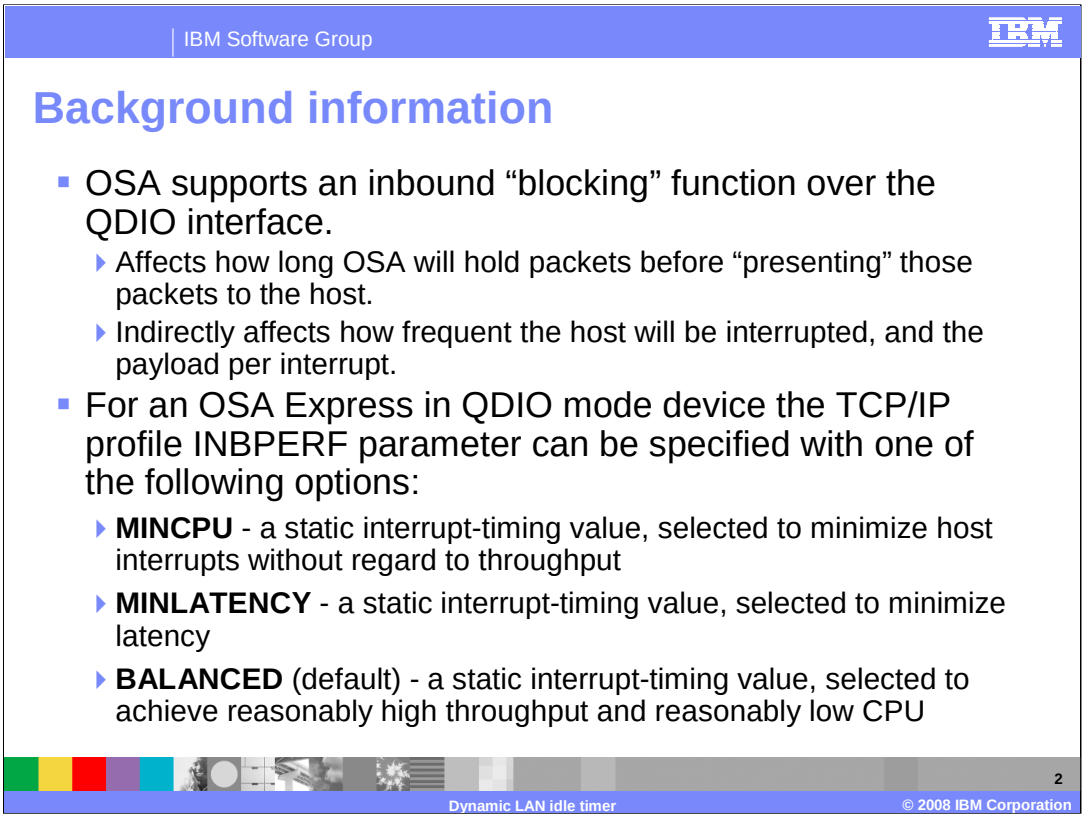


This presentation discusses the Dynamic LAN idle timer the z/OS V1R9 Communications Server.



OSA supports an inbound “blocking” (or packing) function over the QDIO interface. This function affects how long OSA will hold packets before “presenting” those packets to the host. Here “presenting” means assigning the read buffer to the host, which is a matter of updating the state of the host buffer to host owned. In most cases this same action will result in an interrupt to the host for this QDIO data device. Therefore, this function indirectly affects the QDIO interrupt processing (that is, how frequent the host will be interrupted, and the payload per interrupt).

This function is referred to as the OSA “LAN Idle timer”. Today the host can pass various time intervals to OSA when the QDIO data device is activated. In the z/OS case, the system administrator can adjust this setting. However, the setting is static and can not be changed unless the connection to OSA is terminated (device is stopped) and reestablished (restart the device).

Currently you can not directly configure explicit LAN Idle settings. Instead the user provides a constant value to OSA which represented the best “compromise” setting. In the TCP/IP profile the user can define a LAN Idle setting for an OSA Express in QDIO mode device. This is performed by specifying the TCP/IP profile INBPERF parameter with one of the following options:

MINCPU setting

OSA should increase packet hold time. Holding packets longer minimizes CPU utilization by decreasing interrupt frequency and providing a better payload per interrupt)

MINLATENCY setting

OSA should decrease packet hold time. Presenting packets to host sooner reduce network latency (but drive up CPU utilization by causing more frequent interrupts and a smaller payload per interrupt)

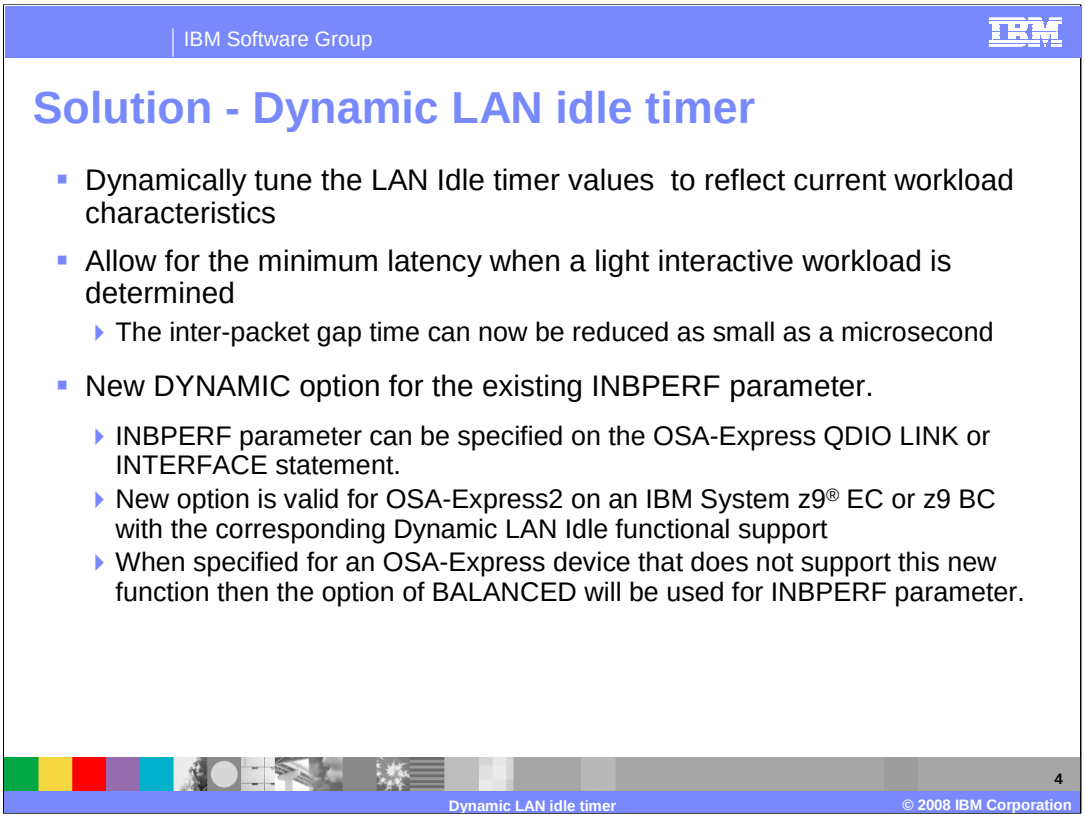
BALANCED setting (default)

CS compromises and sets timer interval values that split the difference in the above two settings

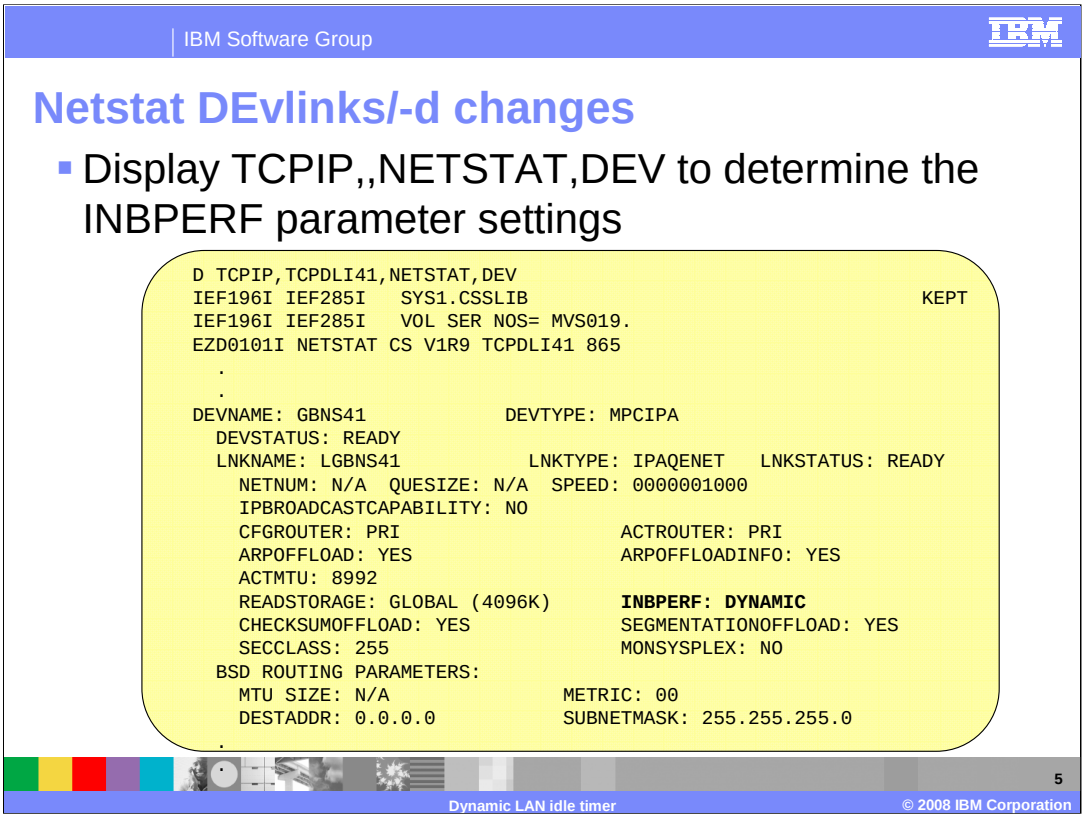
Problem - Network latency on zSeries®

- LAN idle timer settings have contributed to network latency on zSeries
  - Even when the INBPERF parameter is specified with a value of MINLATENCY the permitted inter-packet gap is set to 20 microseconds
- LAN idle timer settings are static and can not be changed unless the connection to OSA connection is terminated and reestablished.

Our current LAN idle timers are having OSA hold the packets to save CPU when sometimes CPU is not an issue. Note that at higher utilizations, dispatch delay becomes a determining factor in network turnaround time rather than LAN idle delay.



Performance studies have shown network latency improvements in environments where the CEC is under low utilization of up to 35% by tuning the Lan Idle timer within the OSA Express2 using a dynamic algorithm that takes workload characteristics. This dynamic algorithm involves taking the current default inter-packet gap of 40 microseconds to as low as 1 microsecond.
A new INBPERF parameter option of DYNAMIC will now be permitted. This new configurable setting allows the TCP/IP stack to dynamically calculate the best values for the LAN idle timer settings. These settings will indirectly determine how frequently the OSA adapter will interrupt the host for inbound traffic.



The **Netstat DEvlinks/-d** command displays information about devices and defined interfaces or links defined to the TCP/IP stack.

The **INBPERF** field is significant only for active IPAQENET links, IPAQTR links, and IPAQENET6 interfaces. This field indicates how frequently the adapter should interrupt the host. The possible values are:

MinCPU

Indicates that the adapter is using a static interrupt-timing value that minimizes host interrupts, and therefore minimizes host CPU consumption.

MinLatency

Indicates that the adapter is using a static interrupt-timing value that minimizes latency delay by more aggressively presenting received packets to the host.

Balanced

Indicates that the adapter is using a static interrupt-timing value that strikes a balance between MinCPU and MinLatency.

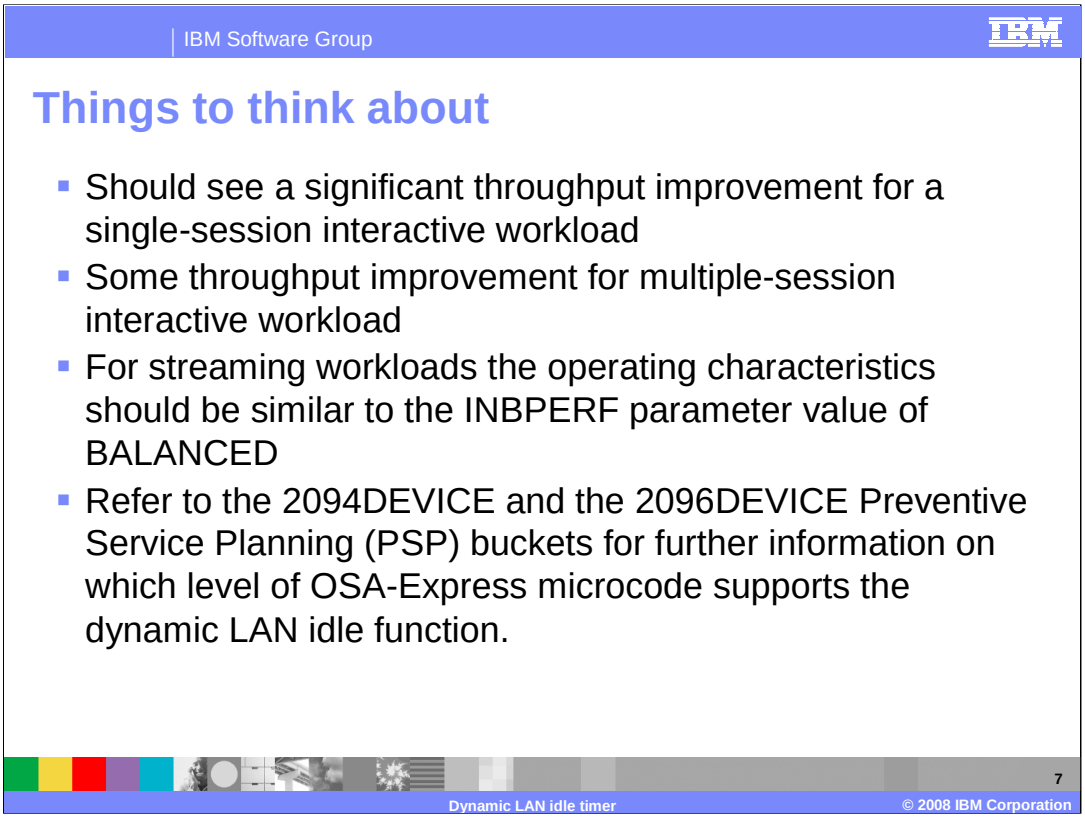
Dynamic

Indicates that the stack and the adapter are dynamically updating the frequently with which the adapter should interrupt the host for inbound traffic.

Diagnosis

- When an INBPERF parameter is specified and the OSA-Express device does not support dynamic LAN idle support, then:
  - EZD0036I error message will be received when the INBPERF value was specified on a LINK statement.
  - EZD0037I error message will be received when the INBPERF value was specified on a INTERFACE statement.
  - For either of the above errors the INBPERF parameter will default to a setting of BALANCED
- For dynamic LAN idle settings update failures a EZD0038I message will be issued.
  - OSA will continue running with whatever values were last successfully set.

For EZD0036I and EZD0037I, install a level of OSA-Express microcode that supports the dynamic LAN idle function. Use the VTAM DISPLAY TRL command to determine your current OSA-Express microcode level. If you choose to not use the dynamic LAN idle function, remove the DYNAMIC value from the LINK statement or INTERFACE statement.



A new dynamic LAN idle algorithm will be used to compute the optimal OSA inter-packet gap timers to be utilized. These LAN Idle Timers will be updated by this algorithm to attempt to optimize throughput. The new algorithm should be effective for all protocols. The new DYNAMIC option for the existing INBPERF parameter is only valid for OSA-Express2 on an IBM System z9 EC or z9 BC with the corresponding Dynamic LAN Idle functional support. You should see a significant throughput improvement for a single-session interactive workload. A Latency improvement of 30% or more with a reduction in the CPU cost per transaction.
The dynamic LAN idle timer algorithm will adjust the LAN idle timer settings to best fit the traffic characteristics.

Feedback

Your feedback is valuable

You can help improve the quality of IBM Education Assistant content to better meet your needs by providing feedback.

- Did you find this module useful?
- Did it help you solve a problem or answer a question?
- Do you have suggestions for improvements?

Click to send e-mail feedback:

mailto:iea@us.ibm.com?subject=Feedback_about_DynLAN.ppt

This module is also available in PDF format at: ../DynLAN.pdf

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

Trademarks, copyrights, and disclaimers

The following terms are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both:

IBM System z9 z/OS zSeries

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.

Information is provided "AS IS" without warranty of any kind. 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.

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.