IBM Software Group   Enterprise Networking Solutions
z/OS® V1R12 Communications Server

z/OS Communications Server – Overview
SNA and Enterprise Extender

The charts in this presentation cover those new functions in the SNA and EE theme in z/OS® V1R12 Communications Server.
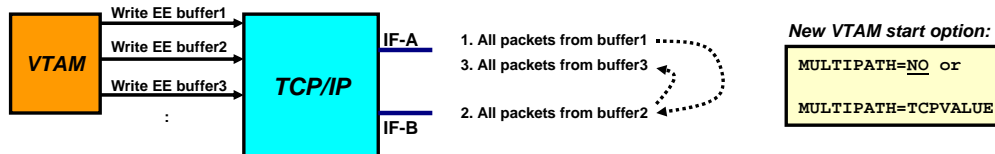
SNA and EE

- Multipath control for Enterprise Extender
  - Separate configuration of multipath for EE and the TCP/IP stack
  - Now possible to enable multipath for TCP/IP while disabling it for EE
- Improved recovery from RTP pipe stalls
  - RTP pipe endpoints which are not the EE endpoint do not learn the TCP/IP path MTU
    - Endpoint continues to send packets at a non-optimal size, which can lead to packet loss and transmission stalls
  - EE will trigger a path switch attempt when entering stall state rather than just slamming the MTU to the minimum size
- Enterprise Extender connection health verification
  - Verify health of EE connections during activation and on active connections
- Enhancements to topology database diagnostics

Several enhancements are made to the SNA and Enterprise Extender support in z/OS V1R12 Communications Server. A new VTAM® option can disable multipath support for Enterprise Extender while allowing multipath to be enabled for TCP/IP applications. RTP pipe stall recovery is improved in cases where the path MTU changes on an existing RTP pipe. VTAM can now verify that all five UDP ports can be used to communicate with an Enterprise Extender partner, both at start and on a periodic basis. And topology database diagnostics are improved for cases where a TDU war can be occurring.

Multipath control for Enterprise Extender

- With multipath enabled in TCP/IP
  - All packets in one EE write buffer are sent over one interface
  - All packets in the next EE write buffer is sent over another interface
  - A modified per-packet algorithm – really a per-EE-buffer algorithm
- Same behavior independent of PERCONNECTION / PERPACKET setting in TCP/IP
- EE traffic can incur performance issues if the different paths are not truly equal in terms of bandwidth and delay
- Per-connection multipath is generally beneficial for other TCP/IP traffic
- New support to allow TCP/IP to specify use of Multipath, but disable it by default for EE traffic

```
                  Write EE buffer1                                                    New VTAM start option:
 VTAM             Write EE buffer2          IF-A    1. All packets from buffer1
                                                    3. All packets from buffer3        MULTIPATH=NO or
                  Write EE buffer3  TCP/IP
                        :                    IF-B    2. All packets from buffer2        MULTIPATH=TCPVALUE
```

The z/OS Communications Server TCP/IP stack allows two values to be coded for multi-path routing. The stack can perform multi-path routing on a per-packet or a per-connection basis. When multi-path routing was enabled for EE it was designed to use a hybrid of the two functions. The fast UDP component has no knowledge of SNA sessions therefore per connection multi-path routing cannot be done. The fast UDP layer will send all the data over the same path for each dispatch which can contain multiple EE packets. The next dispatch of fast UDP will use the next route in a round-robin fashion.

You should enable multi-path routing using per-connection instead of per-packet. The EE implementation uses an algorithm that is similar to per-packet multi-path routing. If you want to use per-connection multi-path for other applications, you need to specify per-packet multi-path routing behavior for Enterprise Extender. Analysis has shown that HPR performance can decrease when packets are consistently delivered out of order. Multi-path routing increases the chances of packets arriving out of order and can therefore negatively affect HPR performance.

z/OS Communications Server has a new function to allow multi-path routing to be enabled in the stack but disabled for Enterprise Extender. This will allow TCP applications to benefit from utilizing multiple paths and prevent HPR performance problems on EE connections. The start option will allow EE connections to use the MULTIPATH setting in the TCP/IP profile or to disable multi-path routing for EE connections only. Because of the known concerns with EE multi-path routing, the start option will default to off which disables multi-path routing for EE.

Improved recovery from RTP pipe stalls

- z/OS V1R10 provided a version of Path MTU Discovery (PMTU) for Enterprise Extender.
  - However, MTU size changes over an active EE link are only communicated to the endpoints of that EE link (NN2 and NN3)

RTP Pipe

NN1   **NN2**   *EE IP Network*   **NN3**

No MTU Updates   MTU Updates
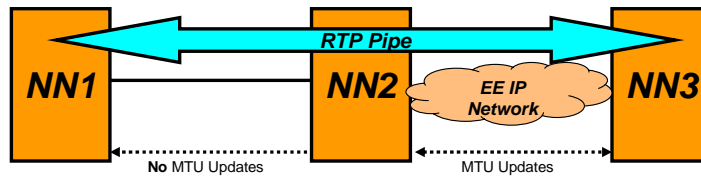
- If an existing RTP pipe begins on a node other than the EE link endpoint
  - It will not learn the PMTU-discovered MTU size
  - It will continue to send packets at a non-optimal size, potentially resulting in packet loss and transmission stalls.
- z/OS V1R12 adds logic for VTAM to drive the path switch logic if multiple retransmissions occur (stall detection)
  - Thereby letting NN1 learn the new current MTU size and adapt
  - `IST2335I PATH SWITCH REASON: XMIT STALL RECOVERY`

z/OS V1R10 Communications Server implemented the path MTU discovery function for Enterprise Extender. This function dynamically updates the EE MTU size for the Enterprise Extender connection when changes are detected in the IP network. If the HPR endpoint is not the same as the EE endpoint then the HPR endpoint does not benefit from the MTU updates.

Also, VTAM can detect a transmission stall. When a stall is detected, VTAM issues message IST2245I informing the operator of the stall condition. VTAM will issue message IST2246I every thirty seconds while in the stall state. Finally message IST2247I is issued when the stall is alleviated. The transmission stall is detected by one of the RTP endpoints of the pipe. Therefore the node detecting the problem cannot have the EE connection as the first hop and it will not receive MTU updates.

The HPR connection can start with an NLP size of 4K. A change in the IP network can cause the data size of the EE TG to decrease from 4K to 1500. The RTP endpoint at NN1 still has 4K as the NLP size and continues sending 4K data. Data is discarded in the IP network and the HPR connection will detect a transmission stall.

The current HPR architecture does not contain a method for one node to pass an updated NLP size to another. This information is calculated in the route setup RU when the connection is started or during a path switch.

This solution will enhance the RTP transmission stall logic to force a path switch. The path switch will send a route setup on the connection and calculate the updated MTU/NLP information from the EE ANR hop of the connection. It is anticipated that the path switch will choose the same route as before. The benefit is the new NLP size is sent to the RTP end-point that will alleviate the transmission stall.

Enterprise Extender connection health verification
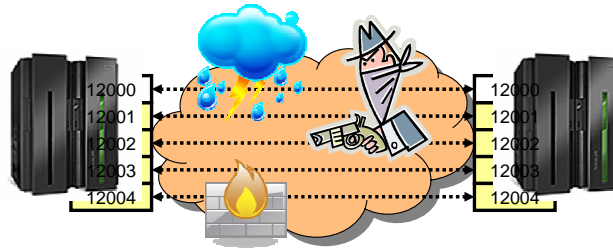
- Questions:
  - Are all five EE ports reachable at EE connection initialization point in time?
  - Do all five EE ports remain reachable?
- Apart from something not working correctly, you really do not know!
- z/OS V1R12 adds optional probing logic
  - Probe during EE connection initialization
  - Probe during the lifetime of the EE connection to verify the health of the EE connection.

VTAM uses ports 12000 through 12004 for Enterprise Extender connections. One or more ports can be blocked at the time of the activation of the EE connection and when the EE connection is active. A port can be blocked due to a firewall or some other condition. VTAM can still activate an EE connection, even if one or more ports is blocked.

When the operator issues the command D NET,EEDIAG,TEST=YES, VTAM sends the test probe using all five ports to detect the connectivity problem. It is a manual method to determine the connectivity problem of all EE connections. When a customer has many EE connections, it is not easy to detect EE connection problems manually.

In z/OS V1R12, VTAM is updated to verify the health of the EE connection during the activation of the EE connections and to verify the health of the active EE connections. VTAM sends a probe to the remote partner on all five ports during the activation of the EE connection to verify the health of the EE connection. If it finds one or more ports blocked or unreachable, it issues an error message and it does not bring up the EE connection. If VTAM finds that the remote partner does not support the probe, it issues a message and it brings up the EE connection.

VTAM also sends the probe to the remote partner at a user specified interval on an active EE connection. If the remote partner is not reachable using any port for any reason, VTAM issues a highlighted error message if it is not already present. The highlighted error message stays on console until the operator clears the message or all active EE connections receive a successful EE health verification.

There are three user specified interval options. The first option is EEVERIFY=NEVER which means do not send any probes. The second option is EEVERIFY=ACTIVATE which means probe during connection initialization. The third option is EEVERIFY=timer-interval which means probe during initialization and periodically at the specified timer-interval.

wnsna.ppt

### Enterprise Extender connection health verification - example

- To see all failed connections, issue this command:

```
d net,ee,list=eeverify
 IST097I DISPLAY ACCEPTED
 IST350I DISPLAY TYPE = EE
 IST2000I ENTERPRISE EXTENDER GENERAL INFORMATION
 IST1685I TCP/IP JOB NAME = TCPCS
 IST2003I ENTERPRISE EXTENDER XCA MAJOR NODE NAME = XCAIP
 IST2004I LIVTIME = (10,0)    SRQTIME =    15  SRQRETRY =      3
 IST2005I IPRESOLV =    0
 IST2231I CURRENT HPR CLOCK RATE = STANDARD
 IST924I -------------------------------------------------------------
 IST2006I PORT PRIORITY =  SIGNAL     NETWORK     HIGH   MEDIUM     LOW
 IST2007I IPPORT NUMBER =   12000       12001    12002    12003   12004
 IST2008I IPTOS VALUE   =      C0          C0       80       40      20
 IST924I -------------------------------------------------------------
 IST2324I EE HEALTH VERIFICATION: FAILED CONNECTION INFORMATION
 IST2325I LINE LNIP1 PU SWIP2A1 ON 12/21/09 AT 15:56:39
 IST2326I EE HEALTH VERIFICATION TOTAL CONNECTION FAILURES = 1
 IST2017I TOTAL RTP PIPES =          1      LU-LU SESSIONS =          2
 IST2018I TOTAL ACTIVE PREDEFINED EE CONNECTIONS          =          1
 IST2019I TOTAL ACTIVE LOCAL  VRN EE CONNECTIONS          =          0
 IST2020I TOTAL ACTIVE GLOBAL VRN EE CONNECTIONS          =          0
 IST2021I TOTAL ACTIVE EE CONNECTIONS                     =          1
 IST314I END
```

When the D NET,EE,LIST=EEVERIFY command is issued, the display output includes the header message IST2324I, the health verification failed message IST2325I (one per connection), and the total number of failed health verification connections message IST2326I. If there were no failed EE health verification connections, the display output does not include message IST2325I.

Enhancements to topology database diagnostics

- Enhancements in V1R11 defined a new control vector for TDU flows
  - Topology Resource Sequence Number Update (x'4E') control vector to identify node that set the RSN

- TDUDIAG start option available to control frequency of when new control vector is included

- Still required dumps and traces to diagnose many TDU-related problems

- z/OS V1R12 enhances various commands to improve the ability to better diagnose issues related to TDU flows:
  - Enhance existing DISPLAY TOPO,LIST=TDUINFO output
  - New DISPLAY TOPO,LIST=TDUDIAG summary command
  - Diagnostic information from the Topology RSN Update control vector added in V1R11 is saved
  - New displays of diagnostic information from the x'4E' control vector
    - DISPLAY TOPO,LIST=TDUDIAG command for a TG
    - DISPLAY TOPO,LIST=TDUDIAG command for a node

Symptom: CP1 performance is affected

CP4 and CP5 engaged in TDU War
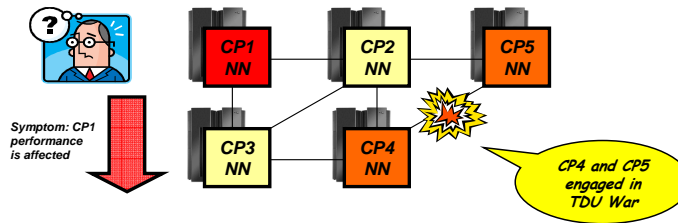
The term "TDU war" defines a situation where two, or more, APPN network nodes erroneously think they are responsible for reporting topology information about a node or a TG. This leads to an endless flood of TDU flows, where each node attempts to correct the topology information provided by the other node or nodes. Topology information is distributed across the network, impacting not just the two network nodes involved in the TDU war, but every node in the network. The volume of TDU flows can significantly degrade bandwidth throughout the network. A major concern with TDU wars, even beyond the performance impact, is that it is very challenging to determine the root cause of the TDU war. The biggest challenge is that the node experiencing the performance issues is often not one of the nodes that is generating all the topology updates flooding the network. In order to stop the TDU war, you need to identify which nodes are generating all the updates.

The topology display enhancements in z/OS V1R12 Communications Server build on the TDUDIAG function introduced in V1R11. The DISPLAY TOPO,LIST=TDUINFO command output originally provided information about topology resource control vectors included in inbound TDUs. The command output is enhanced to also display information about topology resource control vectors included in outbound TDUs. In addition, information is displayed about the topology resources whose resource sequence numbers (RSN) are updated most frequently by the node where the display is entered. New DISPLAY TOPO,LIST=TDUDIAG commands are now available to display RSN update information received and sent in TDUs. A LIST=TDUDIAG summary command identifies topology resources with the most TDU activity that have associated RSN update information saved in the topology database. Once the resources are identified, the LIST=TDUDIAG command for a node or TG displays the RSN update detailed information.

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 email feedback:

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

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

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

Trademarks, copyrights, and disclaimers

IBM, the IBM logo, ibm.com, VTAM, and z/OS are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. 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

THE INFORMATION CONTAINED IN THIS PRESENTATION IS PROVIDED FOR INFORMATIONAL PURPOSES ONLY. WHILE EFFORTS WERE MADE TO VERIFY THE COMPLETENESS AND ACCURACY OF THE INFORMATION CONTAINED IN THIS PRESENTATION, IT IS PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. IN ADDITION, THIS INFORMATION IS BASED ON IBM'S CURRENT PRODUCT PLANS AND STRATEGY, WHICH ARE SUBJECT TO CHANGE BY IBM WITHOUT NOTICE. IBM SHALL NOT BE RESPONSIBLE FOR ANY DAMAGES ARISING OUT OF THE USE OF, OR OTHERWISE RELATED TO, THIS PRESENTATION OR ANY OTHER DOCUMENTATION. NOTHING CONTAINED IN THIS PRESENTATION IS INTENDED TO, NOR SHALL HAVE THE EFFECT OF, CREATING ANY WARRANTIES OR REPRESENTATIONS FROM IBM (OR ITS SUPPLIERS OR LICENSORS), OR ALTERING THE TERMS AND CONDITIONS OF ANY AGREEMENT OR LICENSE GOVERNING THE USE OF IBM PRODUCTS OR SOFTWARE.