IBM

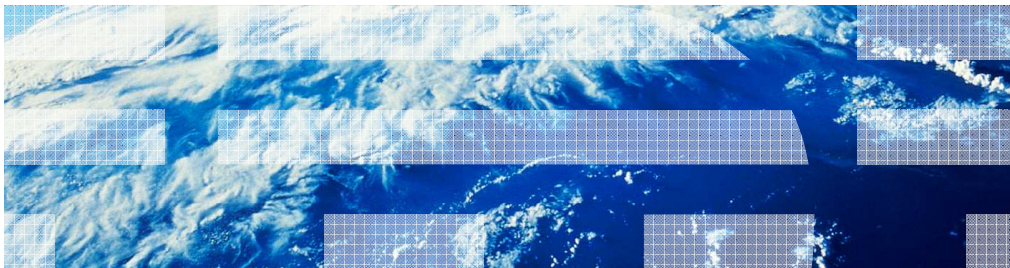# z/OS Communications Server
SNA and Enterprise Extender

This presentation describes the SNA and Enterprise Extender enhancements in z/OS V1R12 Communications Server.

SNA and Enterprise Extender

✓Enterprise Extender connection health verification

✓Multi-path control for Enterprise Extender

✓Improved recovery from RTP pipe stalls

✓Enhancements to topology database diagnostics

SNA and Enterprise Extender

The enhancements to SNA and Enterprise Extender (EE) are:

(1) EE now verifies the health of an EE connection during activation and after connection establishment. It can verify that traffic over all five UDP ports used for EE traffic flow are operational for an EE connection.

(2) EE connections can now be disabled for multi-path support (by default multi-path is disabled). EE connections in the past have suffered performance issues related to out-of-order packets which can often be attributed to the use of multi-path support.

(3) Additional enhancements allow RTP pipes to learn of MTU size changes when traversing an EE connection and one of its endpoints is not the same as the EE connection's endpoints.

(4) Topology display enhancements build on the TDUDIAG function introduced in V1R11. These display enhancements help determine the instigators of a "TDU war". A topology database update (TDU) war occurs when two or more APPN network nodes erroneously think they are responsible for reporting topology information about a node or transmission group (TG).

Enterprise Extender connection health verification before V1R12

IBM

12000
12001
12002
12003
12004

Firewall
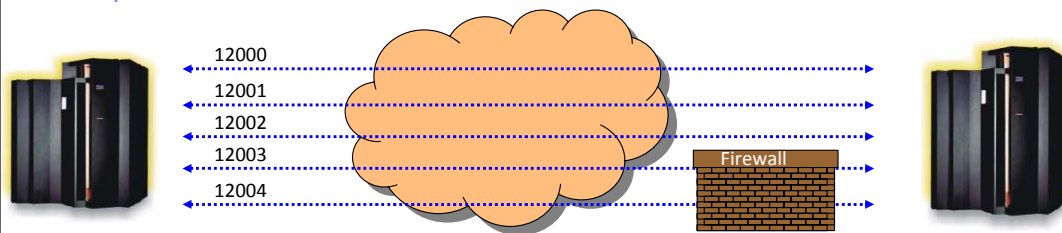
- z/OS V1R8 introduced the D NET,EEDIAG,TEST=YES command to detect possible EE connectivity issues
  - This command is a manual method to determine potential connectivity issues between two EE endpoints
- But EE connectivity issues can be intermittent; problems can exist during the activation of the EE connection or develop after connection activation
- One or more UDP ports used by EE can be blocked due to firewall or some other condition
  - There is no notification of this unhealthy condition

3   SNA and Enterprise Extender   © 2010 IBM Corporation

This slide gives some background about the difficulty of solving an EE connectivity issue prior to V1R12.

When the operator issues the command D NET,EEDIAG,TEST=YES, VTAM sends a test probe using all five ports to detect any connectivity issues.

This is a manual method to find any connectivity problems for each EE connection. When you have many EE connections, it is not easy to detect EE connection problems manually.

VTAM uses ports 12000 through 12004 for EE. One or more ports can be blocked when the EE connection is activated and later after the EE connection is active.
Ports can be blocked due to a firewall or other conditions in the network.

VTAM can activate an EE connection without being able to communicate over all five EE UDP ports. Even if connectivity exists for all five ports during EE session initiation, one or more of these ports can experience problems later during the life of the EE connection.

Enterprise Extender connection health verification in V1R12
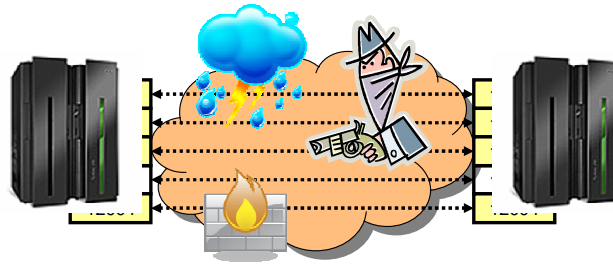
- Are all five EE ports reachable at EE connection initialization point in time? Do all five EE ports remain reachable?
- Apart from something not working correctly, you really do not know!
- z/OS V1R12 adds optional probing logic during EE connection initialization and during its lifetime to verify the health of the EE connection.
  - EEVERIFY=NEVER
    - Do not send any probes
  - EEVERIFY=ACTIVATE
    - Probe during connection initialization
  - EEVERIFY=timer-interval
    - Probe during initialization and periodically at the specified timer-interval

4   SNA and Enterprise Extender   © 2010 IBM Corporation

In z/OS V1R12, VTAM verifies the health of the EE connection during activation. Additionally connection health can be verified for the duration of the EE connection. VTAM sends a probe to the remote partner on all five ports during the activation to verify the health of the EE connection. If it finds one or more ports blocked or unreachable, VTAM issues an error message and does not bring up the EE connection. If VTAM finds that the remote partner does not support the probe, it issues a message and does bring up the EE connection.

VTAM also sends a 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 the console until the operator clears the message or all active EE connections receive a successful EE health verification.

Enterprise Extender connection health verification – all failed connections

- 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
```

SNA and Enterprise Extender

If 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 EE health verification failures, the display output does not include message IST2325I.

Enterprise Extender connection health verification – specific failed connection

- To see a specific failed connection, issue this command:

```
d net,ee,id=SWIP2A1
 IST097I DISPLAY ACCEPTED
 IST350I DISPLAY TYPE = EE
 IST2001I ENTERPRISE EXTENDER CONNECTION INFORMATION
 IST075I NAME = SWIP2A1, TYPE = PU_T2.1
 IST1680I LOCAL IP ADDRESS 9.67.1.1
 IST1680I REMOTE IP ADDRESS 9.67.1.2
 IST2022I EE CONNECTION ACTIVATED ON 12/21/09 AT 16:21:57
 IST2114I LIVTIME:   INITIAL =   10   MAXIMUM =    0   CURRENT =   10
 IST2023I CONNECTED TO LINE LNIP1
 IST2327I EE HEALTH VERIFICATION OPTION - EEVERIFY = 2 MINUTES
 IST2328I EE HEALTH VERIFICATION FAILED ON 12/21/09 AT 20:55:57
 IST2339I EE HEALTH VERIFICATION LAST SUCCESS ON 12/21/09 AT 20:52:57
 IST2025I LDLC SIGNALS RETRANSMITTED AT LEAST ONE TIME   =        0
 IST2026I LDLC SIGNALS RETRANSMITTED SRQRETRY TIMES      =        0

 IST314I END
```
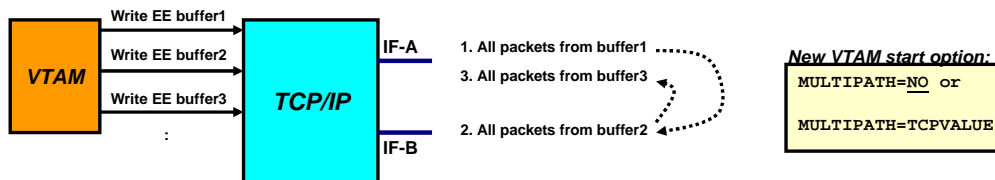
6　　　　　SNA and Enterprise Extender　　　　　© 2010 IBM Corporation

If the D NET,EE,ID=puname command is issued, the display shows a message group with the failed EE health verification.

The example shows the message group consisting of IST2327I, IST2328I and IST2339I because the most recent EE health verification was a failure.

Multi-path control for Enterprise Extender

- For EE multi-path, all packets in one EE write buffer are sent over an 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 multi-path is generally beneficial for other TCP/IP traffic

- New support allows TCP/IP to use multi-path, but disable it by default for EE traffic

```
            Write EE buffer1
  VTAM      Write EE buffer2      IF-A     1. All packets from buffer1 .......
            Write EE buffer3     TCP/IP    3. All packets from buffer3
                  :              IF-B      2. All packets from buffer2
```

New VTAM start option:
```
MULTIPATH=NO or

MULTIPATH=TCPVALUE
```

The z/OS 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 used. The fast UDP component 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 path 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 will still get a 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 V1R12 Communications Server is enhanced to allow multi-path routing to be enabled in the stack but disabled for Enterprise Extender. This allows TCP applications to benefit from using multiple paths and prevents 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.

SNA_EE.ppt

Improved recovery from RTP pipe stalls

- z/OS V1R10 provided a version of Path MTU Discovery (PMTU) for Enterprise Extender.
  - However, MTU size changes are only communicated to the two endpoints of that EE link (NN2 and NN3 below)
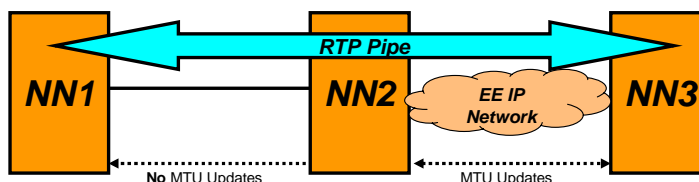
- If an existing RTP pipe begins on a node other than the EE link endpoint, it will not learn the PMTU-discovered MTU size
- Therefore the RTP pipe will continue to send packets at a non-optimal size, potentially resulting in packet loss and transmission stalls.
- z/OS V1R12 adds logic for VTAM to drive the path switch logic if multiple retransmissions occur (stall detection)
  - Thereby letting NN1 here 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 EE connection when changes are detected in the IP network. If the HPR endpoint is not the same as the EE endpoint, the HPR endpoint does not benefit from the MTU updates.

Additional function was added to detect a transmission stall. When a stall is detected, VTAM issues message IST2245I informing the operator of the stalled 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 will continue 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.

V1R12 enhances the RTP transmission stall logic to force a path switch. The path switch sends a route setup on the connection and calculates 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 that the new NLP size is sent to the RTP end-point which will alleviate the transmission stall.
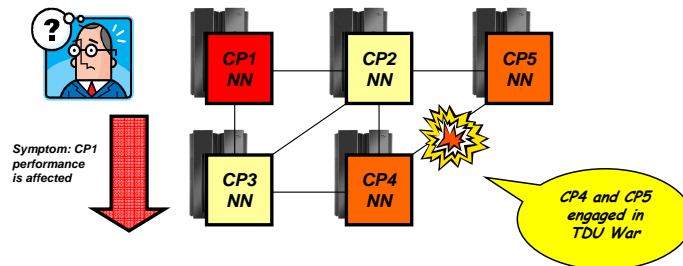
The term "TDU war" defines a situation where two or more APPN network nodes erroneously think they are responsible for reporting topology information about a node or a TG. This leads to an endless flood of TDU flows, where each node attempts to correct the topology information provided by the other nodes. 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.

Enhancements to topology database diagnostics

- 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

SNA and Enterprise Extender   © 2010 IBM Corporation

The topology display enhancements in z/OS V1R12 Communications Server build on the TDUDIAG function introduced in V1R11.

The DISPLAY TOPO,LIST=TDUINFO command 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 (RSNs) are updated most frequently by the node where the display is entered.

New DISPLAY TOPO,LIST=TDUDIAG options 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.

New TDUDIAG display for a TG

▪ Displaying topology database update (TDU) diagnostic information for a TG:

```
>>___DISPLAY NET,TOPO__,LIST=TDUDIAG__,ORIG=cp_name___>

 >__,DEST=cp_name__,TGN=tg_number_____>

       __,NUM=10_____
 >__|_____|_____>
       |_,NUM=_number_of_TDU_updates_|

       _,CLEAR=NO_____
 >__|_____|_____><
       |_,CLEAR=_ _YES_ __|
            |_NO__|
```

SNA and Enterprise Extender   © 2010 IBM Corporation

Now let's look at the DISPLAY TOPO,LIST=TDUDIAG command for a TG or node. To display TDU diagnostic information for a TG, you specify the ORIG, DEST, and TGN operands to identify the TG you want to display. The display output provides the information saved from the x'4E' control vectors received and sent for this TG. The DISPLAY TOPO,LIST=TDUDIAG command for a node provides the same information by replacing the ORIG, DEST, and TGN operands with the ID operand.

If you include the CLEAR=YES operand on the LIST=TDUDIAG command for a TG or node, only the TDU counters and RSN update information for that topology resource are cleared.

You have already identified the three resources that are in contention in the TDU war, so you can continue diagnosis by displaying the TDUDIAG information for each of the TGs in contention.

TDUDIAG display command example for a TG

- DISPLAY NET,TOPO,LIST=TDUDIAG,ORIG= SSCP1A,DEST=SSCPBA, TGN=21,NUM=1 to see the TDUDIAG information for a TG:

```
IST350I DISPLAY TYPE = TDU DIAGNOSTICS
IST2311I TDU DIAGNOSTIC INFORMATION FOR TG: TGN = 21
IST2256I ORIG = NETA.SSCP1A - DEST = NETA.SSCPBA
IST2290I TDUDIAG START OPTION = 1000
IST2312I CURRENT RSN = 249644 - HEX RSN = 0003CF2C
IST924I  ----------------------------------------------------------
IST2275I TDU INFORMATION SINCE LAST RESET ON 03/23/10 AT 10:44:39
IST1769I LAST TDU RECEIVED - 03/23/10 16:46:57 FROM NETA.SSCP2A
IST2281I LAST TDU SENT - 03/23/10 16:46:57
IST2282I TDU COUNTS:
IST2352I   SENT     = 261842      RECEIVED = 259158
IST2353I   ACCEPTED = 46668       REJECTED = 206289
IST2354I   IGNORED  = 6201
IST2313I TOTAL RSN UPDATES BY LOCAL HOST NODE = 0
IST924I  ----------------------------------------------------------
IST2294I TDUDIAG RSN UPDATES:
IST2295I                    TIME      HEX RSN    HEX RSN
IST2296I CP NAME           UPDATED    BEFORE     AFTER      REASON
IST2297I NETA.SSCP1A        16:46:57  0003CF28   0003CF2C   TDU GREATER
IST2300I   RECEIVED FROM: NETA.SSCP1A
IST2314I 1 OF 50 RSN UPDATES DISPLAYED
```

This slide shows the output from the DISPLAY NET,TOPO,LIST=TDUDIAG command for the TG whose origin node is SSCP1A, destination node is SSCPBA, and TG number is 21.

There are three sections of display output for this command. The first section contains general TDUDIAG information. Notice that the current RSN in message IST2312I is displayed in both decimal format and hexadecimal format. The decimal format allows you to compare the RSNs in different DISPLAY TOPO command output. The hexadecimal format allows you to compare the RSN in the display to the TDU RSNs in a VTAM internal trace.

The second section contains information about TDUs associated with the node, including the TDU counters.

The third section of the DISPLAY TOPO,LIST=TDUDIAG output lists RSN update information from the x'4E' control vectors that were created when the RSN was updated. The RSN update at the top of the list is the most recent RSN update in the list.

Notice that all of the TDUDIAG RSN updates in this section include two lines of output. Because message IST2300I follows message IST2297I for every RSN update, SSCPBA is recording only updates from inbound TDUs. The IST2297I identifies the node that actually updated the RSN. In this case, SSCP1A is listed in all of the RSN updates. The reason of "TDU GREATER" indicates that SSCP1A received a TDU, but the RSN received in the inbound TDU was greater than the RSN in SSCP1A's topology database (HEX RSN BEFORE). In that case, SSCP1A increments the RSN from the inbound TDU and sends out a new TDU with the updated RSN (HEX RSN AFTER).

Also notice that the RSN in the RSN updates are not in numeric order in the display on SSCPBA. This is because there are thousands of TDUs flowing between multiple networks nodes, and SSCPBA is receiving TDUs from multiple adjacent nodes. However, all of the RSNs are even numbers. Typically the node that owns the TG is the node that updates the RSN with even numbers. This is consistent as the origin node of the TG displayed is SSCP1A.

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_SNA_EE.ppt

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

SNA and Enterprise Extender   © 2010 IBM Corporation

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