

Understanding Enterprise Extender

Part I. Concepts and Considerations

Sam Reynolds

samr@us.ibm.com

IBM z/OS Communications
Server Design

Agenda

- **Understanding EE: Concepts and Considerations**

- ▶ **SNA: Dead or Alive?**
- ▶ **Enterprise Extender - What is it?**
- ▶ **High Performance Routing (HPR) Concepts**
- ▶ **Link Characteristics and TGPs**
- ▶ **Connection Network**
- ▶ **Planning for Enterprise Extender**

- **Understanding EE: Nuts and Bolts**

- ▶ **HPR and EE: Layers and Timers**
- ▶ **EE Enhancements in V1R5**
- ▶ **Coding and Operating EE on z/OS**
- ▶ **Testing and Migration for EE**
- ▶ **EE Performance**

SNA: Dead or Alive?

Actual quotes passed on to me in the past year:

- "SNA is going away..." Not likely in foreseeable future.
- "IBM has announced our intent to discontinue service of SNA..." False.
- "EE is an interesting technology, but nobody has really deployed it yet..." No longer true.

SNA: Dead or Alive?...

"This report of my death was an exaggeration."

– Mark Twain, 1897

– SNA, 2005

- SNA Applications: Over a trillion lines of customer written application code based on CICS, IMS, and DB2
- 70% of all business data still accessed via SNA applications
- Numerous market factors including the continued convergence of enterprise networks onto IP technologies, and the withdrawal of the venerable 3745 from marketing, have led to a very rapid adoption of Enterprise Extender as a key component of SNA application access strategy amongst the IBM customer set.
- Recent SHARE and Networking Solutions Technical Conferences have included EE user experience presentations from a number of companies:
 - iT-Austria
 - Bank of America
 - MCI/Worldcom
 - Caterpillar
 - Mitsubishi Motors
 - GAD

IBM SNA Statement of Direction

It is IBM's intent to support VTAM in z/OS Communications Server for the foreseeable future. Customers have a substantial investment in 3270 and SNA applications. We continue to support and enhance VTAM's capabilities while integrating it with new technologies. IBM has no plans at this time to discontinue SNA support in z/OS Communications Server.

From: <http://www.ibm.com/software/network/commserver/os390> (2002)

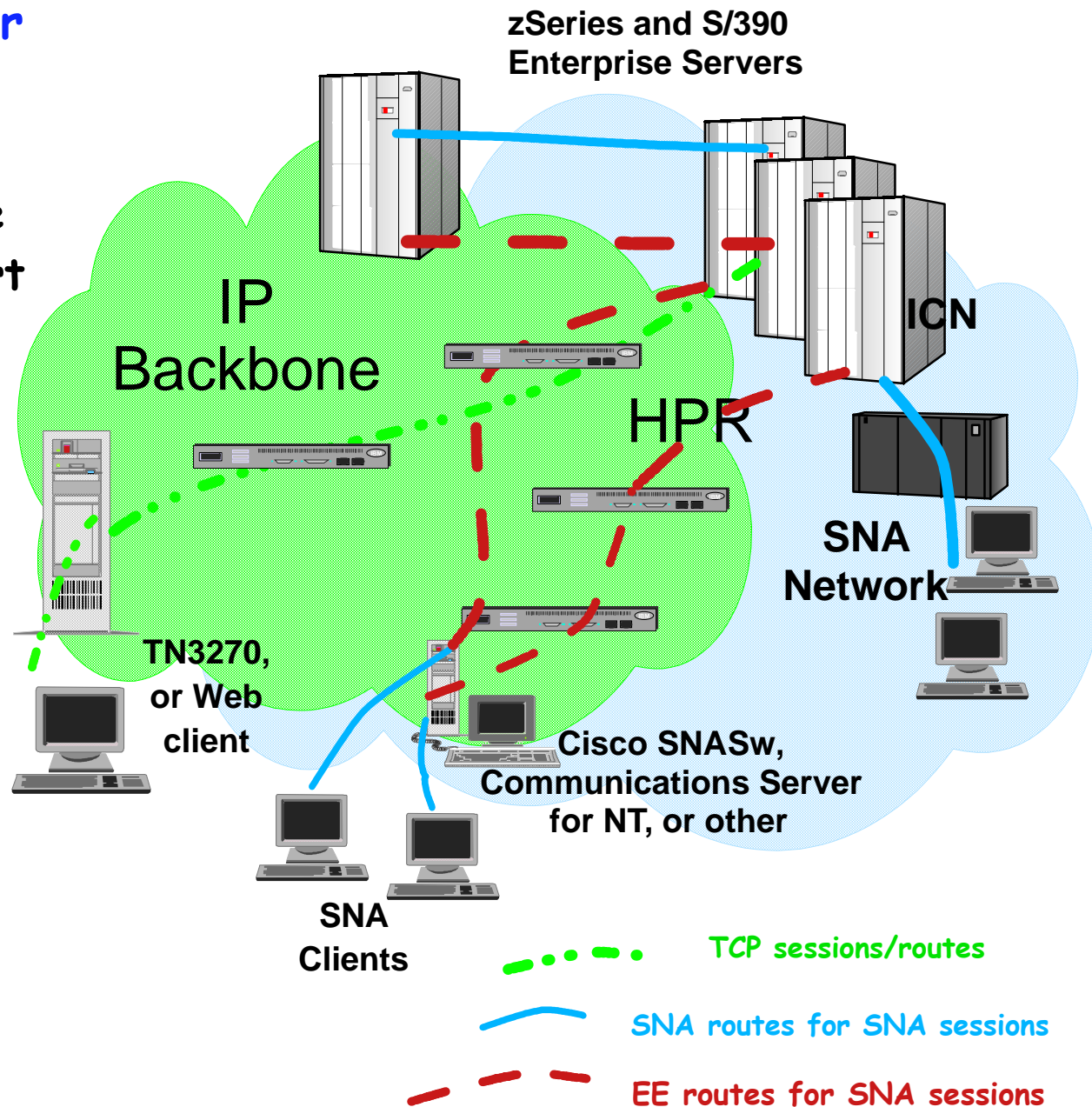
IBM's plans to support SNA workloads have not changed since the Statement of Direction made in 2002. As of June 2004, customers can -- for selected SNA workloads -- use Communications Server products for Linux®, Linux on IBM® zSeries®, Microsoft® Windows®, and AIX® to replace some of the old SNA infrastructure components, such as the IBM 3745/46 or other channel-attached SNA controllers. z/OS® Communications Server can replace some (SNA Network Interconnect) SNI workloads using Enterprise Extender and Extended Border Node functions.

It is IBM's intent to introduce an additional solution in 2005 that uses NCP software running within Linux on zSeries. The intent is to provide a migration path for customers who use traditional SNA (including SNA Network Interconnect (SNI)) to communicate with their Business Partners. This solution can allow them to continue using traditional SNA without a dependency on IBM 3745 and 3746 Communications Controller hardware.

What is Enterprise Extender ?

What is Enterprise Extender?

- Allows use of IP network for SNA sessions
 - EE allows enablement of IP applications and convergence on a single network transport while preserving SNA application and endpoint investment.
- Conceptually, IP network looks like APPN/HPR TG in session route
 - An EE link represents IP connectivity from this host to the specified IP address or host name.
- Typically isolates SNA footprints to the "outside" of the network.

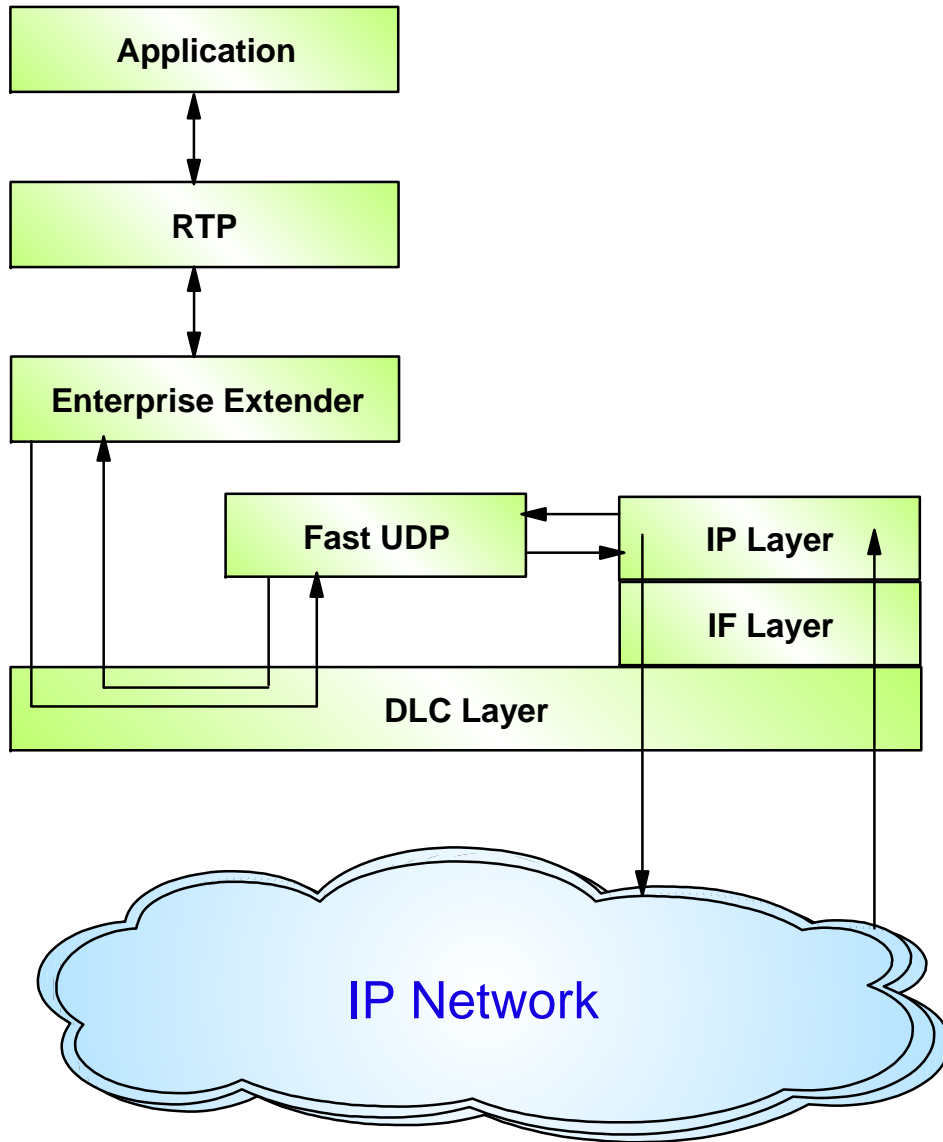


Advantages of Enterprise Extender

- **SNA transport over native IP network**
 - Native IP routing within network maximizes router efficiency
 - Enables SNA applications to take advantage of advances in IP routing
 - SNA traffic can exploit OSA Gigabit Ethernet & HiperSockets
 - EE can use any zSeries or S/390 IP network connection -- channel attached router, OSA, etc.
- **No changes to SNA applications**
- **End-to-End failure protection and data prioritization**
 - SNA priority mapped to IP Type of Service (TOS)
- **Can reduce APPN network complexity while exploiting IP alternate routing/redundancy technologies**
- **EE works with IPSEC and SNA Session Level Encryption**
 - If using NAT, address administration is required for EE
 - EE Connection Network does not work with NAT prior to z/OS

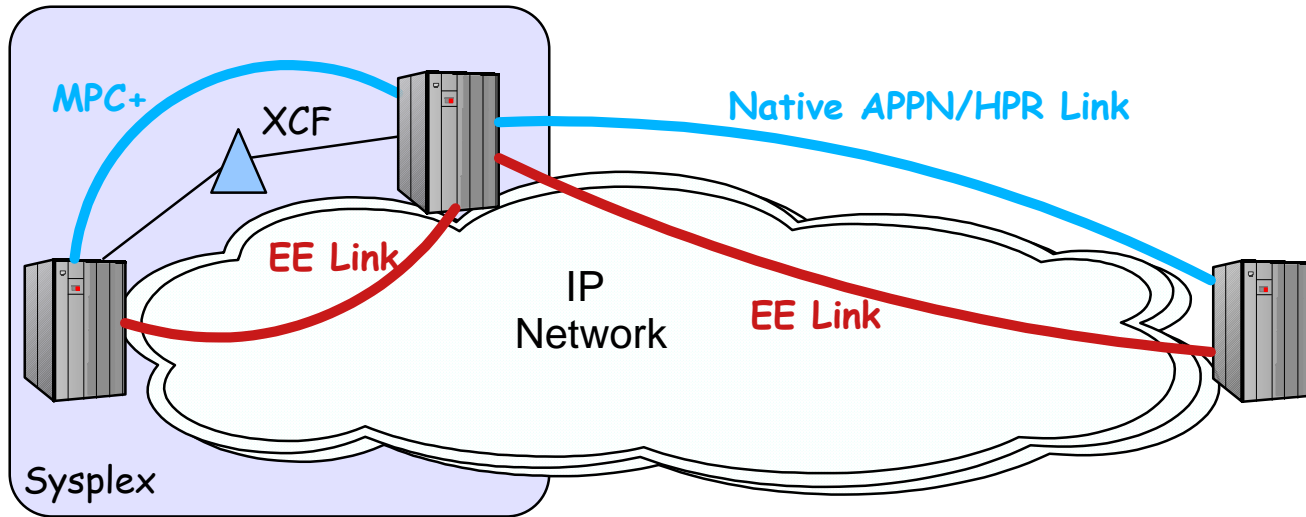
V1R5 CS

Enterprise Extender on z/OS



- For Enterprise Extender, z/OS CS implemented a separate UDP layer (Fast UDP) optimized for EE communications
- Fast UDP communicates with EE (the APPN over UDP component in VTAM) via the IUTSAMEH device

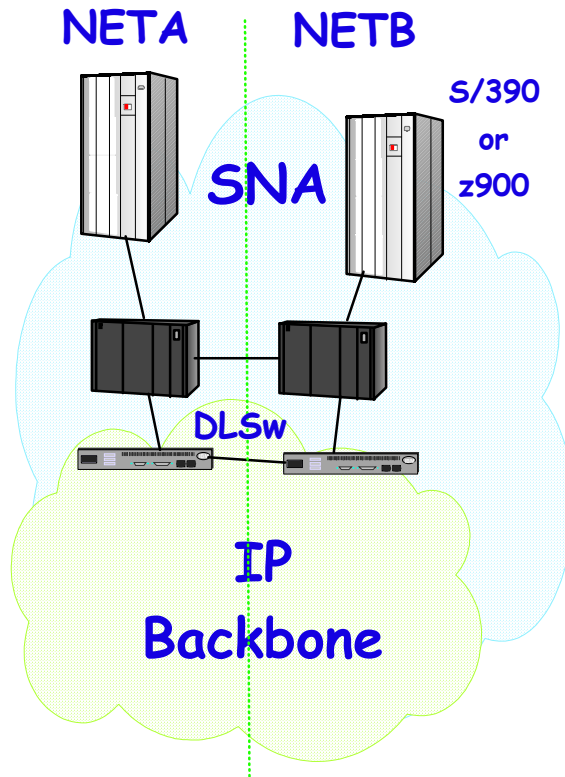
EE Is Just Another Link



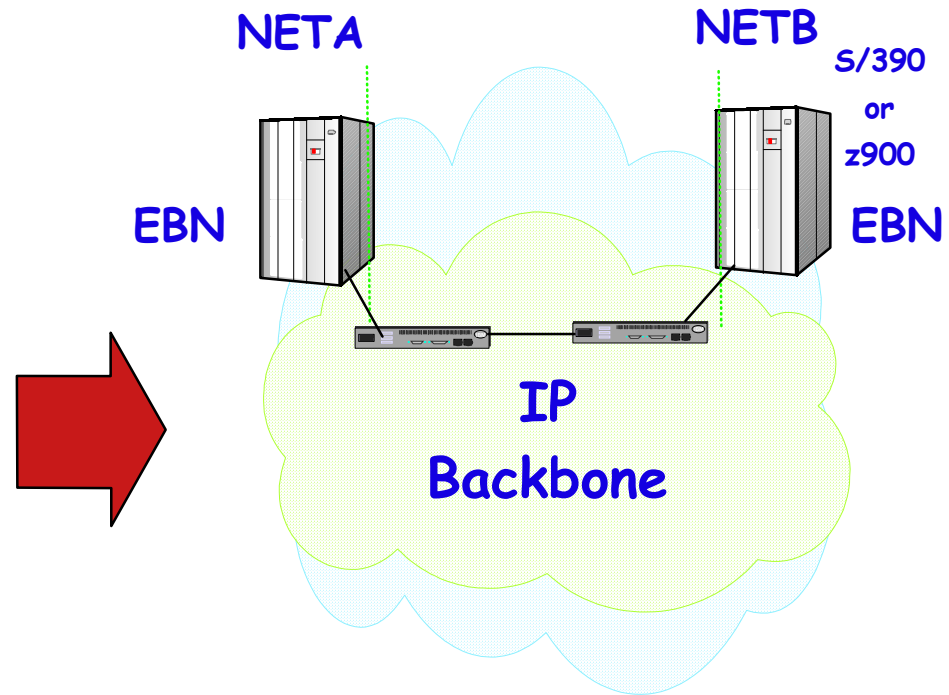
- SNA's view of EE: EE is just another DLC
- IP's view of EE: EE is just another (UDP) application

- Since EE is just another APPN/HPR link type, it can easily coexist with other types of APPN/HPR connectivity
- Selection between EE and other links subject to usual APPN route selection rules
- EE can be used in parallel with MPC+ and XCF in the sysplex
- During EE migration, EE link can be preferred or not preferred based on characteristics/COS, and can be dynamically introduced/removed via F TOPO, FUNCTION=QUIESCE (or NORMAL) commands

EE/EBN As An SNI Alternative



- Traditional SNI
 - ▶ Requires 37xx for interconnectivity
 - ▶ Complex to define and reconfigure
 - ▶ HPR (e.g. non-disruptive session switch) not available



- Enterprise Extender with Extended Border Node (EBN)
 - ▶ Configure border node in z/OS CS only
 - ▶ Single hop SNA connection
 - ▶ SNA apps unchanged
 - ▶ Eliminates DLSw

High Performance Routing (HPR) Concepts

High Performance Routing

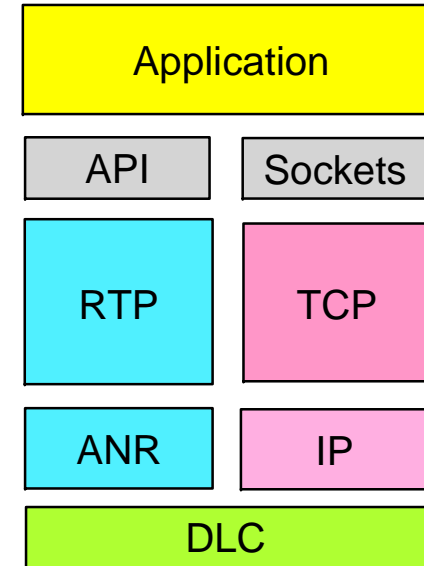
- HPR has 2 major components

- ▶ Rapid Transport Protocol (RTP)

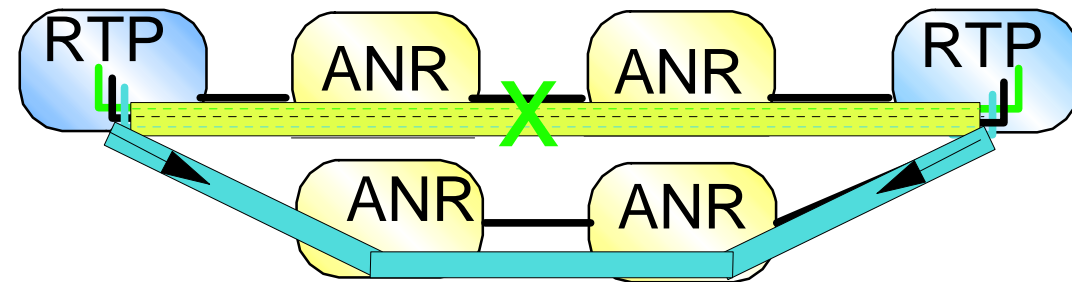
- Error detection and retransmission
- Congestion control
- Prioritization

- ▶ Automatic Network Routing (ANR)

- Label contains all routing information
- SNA router strips label and forwards packet



- HPR preserves sessions across intermediate node/link failure



- ▶ RTP reroutes sessions in event of a planned or unplanned node or link failure
- ▶ Sessions rerouted to new route determined by Class of Service
- ▶ Sessions may be maintained across link "hit" without any switch occurring
- ▶ Discarded data retransmitted using HPR selective retransmission

HPR Path Switch

- There are several triggers that will cause an RTP pipe to enter path switch state:
 - ▶ **General ARB-detected failure**
 - × Failure-driven
 - ▶ **Local link failure**
 - × Failure-driven
 - ▶ **Modify RTP command**
 - ✓ Operator-initiated
 - ▶ **PSRETRY start option**
 - ✓ Timer-initiated

Detecting the Need for Path Switch

- The RTP endpoint frequently asks its partner for status
- When asking for status, RTP starts a Short Request Timer (SRQ) that is based on the Smoothed Round Trip Time (SRTT) and the Smoothing Deviation (SDEV)
- If the timer pops, RTP resends the status request (once again starting the SRQ timer with the same SRTT-based value)
- RTP will retry the status request six times
- If no status reply received, RTP enters path switch mode
- Example:
 - ▶ Assumes responsive mode ARB and "re-fifo" media type (such as EE)

error = measured_RTT - SRTT

SDEV(smooth deviation) = SDEV + 1/4 * (|error| - SDEV)

SRQ = 8*SRTT + 16*SDEV

Detection_Time = SRQ + SRQ + 5*SRQ = 7 * SRQ = 56 * SRTT + 112 * SDEV

Original
Status
Request

Retry 1

Retries 2 - 6

Detecting the Need for Path Switch...

- If the local link (the first hop of the RTP connection) inops, RTP detects this and immediately enters path switch without waiting on SRQ expiration and retry attempts
- The previous discussion assumes data is being sent (Status requests are typically conveyed with data packets)
- If no activity on the pipe, the ALIVE timer must trip to force the first status request
- The ALIVE timer defaults to three minutes, but may be less if the RTP's path contains a delayed disconnect link. It may be disabled entirely if APAR OA04393 is applied.

Path Switch Steps

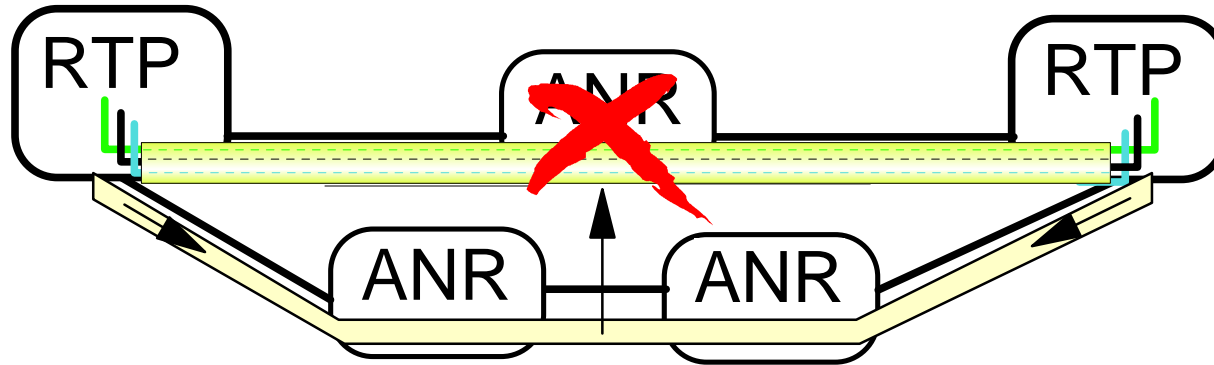
- **When Path Switch Starts:**

1. A path switch timer is started. If that timer pops, the path switch attempt will be terminated, and the pipe will be inactivated. (VTAM control: HPRPST start option)
2. RTP asks for an HPR-capable route to its partner. In a network node, this is an intra-node request between RTP and Topology and Routing Services. In an end node, a locate must be sent to the network node server to ask for the route.
3. When RTP receives the route, a route setup is sent along the path to gather the ANR routing information, etc.
4. If successful so far, RTP sends the new routing information to the partner and the path switch is complete.
5. If not, start a "New Route Timer" (see next foil)
6. When the "New Route Timer" expires, go back to step 2

Path Switch Steps: New Route Timer

- To prevent excessive CPU use by continuous retries for a new route, VTAM will wait "a little while" between path switch route requests.
- The "New Route Timer" (NRT) is an internal, non-configurable timer that VTAM sets according to the following rules:
 1. After the first attempt, the NRT is set to 1 second.
 2. After the second attempt, the NRT is set to 2 seconds.
 3. After each subsequent failed attempt, the value of the NRT is doubled, up to a maximum value of 10% of the path switch timer value (as set by HPRPST)
- Example: By default, the HPRPST value for medium transmission priority pipes is 240 seconds. So, for an RTP pipe with medium transmission priority, we could expect to see the following sequence:
 1. RTP asks for a new route when initially entering path switch
 2. If that fails, RTP waits 1 second, then requests another route
 3. If that fails, RTP waits 2 seconds, then requests another route
 4. If necessary, RTP will retry again after 4 seconds, then 8, then 16, and then 24 seconds (10% of the 240 second HPRPST value)
 5. RTP will continue to retry at 24 second intervals until the 240-second path switch timer pops, indicating that it is time to abandon the path switch attempt and bring down the pipe.
 - That works out to be about 14 attempts, including the initial attempt

PSRETRY & Path Switch

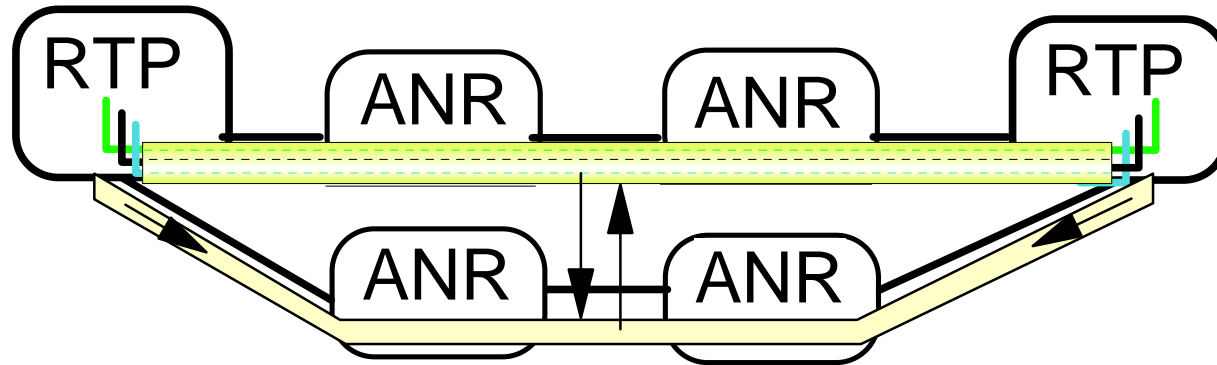


1. Failure occurs on preferred path
2. HPR pipe switches to backup path automatically
3. Preferred path recovers
4. Will HPR pipe return to preferred path?

Solution: PSRETRY

- Start option which determines how often the VTAM RTP endpoint will look for an alternate route
- Specified on a transmission-priority basis
- Example: `PSRETRY=(10M,5M,2M,1M)`
- Default is `(0,0,0,0)` which disables the function

PSWEIGHT & Path Switch



- What if multiple "equal-weight" paths are available?

Alternative: PSWEIGHT

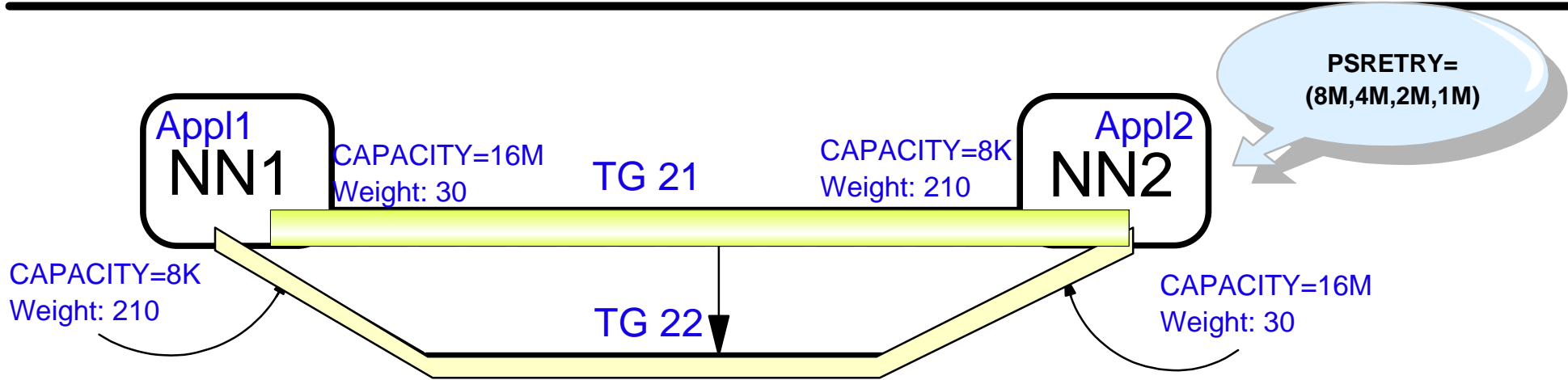
- Indicate to VTAM preference to switch only to "better route"
- Reduces (but doesn't eliminate) switching between equal-weight routes
- PSWEIGHT = LESSTHAN | EQUAL | SAMEROUT

Link Characteristics and Transmission Group Profiles (TPGs)

Importance of TG Characteristics

- When link characteristics are not defined (CAPACITY, PDELAY, TGP, etc.), z/OS CS must assume default values:
 - Enterprise Extender TGs: Could represent any type of IP connectivity between the EE endpoints.
 - Defaults: CAPACITY=4M, PDELAY=TERRESTR
 - Virtual Route Based TGs (VR-TGs): Could represent any mix of subarea connectivity between the two endpoints of the VR-TG.
 - Defaults: CAPACITY=8K, PDELAY=TERRESTR
 - Non-EE Switched PUs: 9600, 56K Frame, 16M Token Ring?
 - Defaults: CAPACITY=8K, PDELAY=TERRESTR
- Not having reasonably accurate and consistent values for the link characteristics may result in a number of problems:
 - Unexpected and possibly non-optimal routes.
 - Reduced initial throughput when using HPR since the defined (or defaulted) speeds are used as input to setting HPR's initial send rate.
 - Excessive path switches and even redundant RTP pipes.

Importance of TG Characteristics...



1. App1 initiates a session with App2. NN1 calculates the route, and chooses TG 21 as the optimal route.
2. Since no RTP pipe exists between NN1 and NN2 over TG 21, one is created.
3. The PSRETRY timer in NN2 pops, so NN2 looks for better routes.
4. NN2 sees TG 22 as the optimal route, and switches the pipe.
5. App1 initiates another session with App2. NN1 chooses TG 21 for the route.
6. Since no pipe exists between NN1 and NN2 over TG21, a pipe is created.
7. PSRETRY runs in NN2, and ...

Importance of TG Characteristics...

- **Recommendation:** Assign Transmission Group Profiles (TGPs) that reflect the media type being used.
- **Especially Important:** Coding TGPs for EE TGs and VR-TGs.
- The set of TGPs shipped with VTAM (in the IBMTGPS member) may be used as examples.
 - It is recommended that you customize the *CAPACITY* operand on the TGP to reflect the media speed of your network's underlying connectivity.
 - **OA12948** will provide five new TGPs for IBMTGPs:
 - FASTENET
 - GIGENET
 - HIPERSOC
 - FICON
 - FICONEXP
 - **OA12948** is a V1R7 APAR, but the TGPs can be copied into IBMTGPS on any level of z/OS CS
- **Recommendation:** Use D TOPO,ORIG=,DEST= commands to verify APPN connectivity, capacity values, and weights
- **Example:** D NET,TOPO,ORIG=CP1,DEST=CP2,APPNCOS=#CONNECT

Coding TG Characteristics with EE

- IBM provides several TGPs in a dataset called IBMTGPS
 - ▶ TGPs are a set of link characteristics like CAPACITY, SECURITY, COSTBYTE, etc. which may be associated with an APPN link
 - ▶ Like other Switched Definitions, the TGP associated with an EE link is coded on the PU in the Switched Major Node

```
CSS1SWEE VBUILD TYPE=SWNET
CSS1PUE  PU      ADDR=22, TGP=TRING16M, DISCNT=NO, *
                CPNAME=CSS1, NETID=CSSNET, PUTYPE=2
CSS1EEPT PATH   IPADDR=9.82.5.120, GRPNM=GRPEEP
```

- ▶ Alternatively, link characteristic parms may be coded on the link definition

```
CSS1SWEE VBUILD TYPE=SWNET
CSS1PUE  PU      ADDR=22, CAPACITY=16M, PDELAY=NEGLIGIB, DISCNT=NO, *
                CPNAME=CSS1, NETID=CSSNET, PUTYPE=2
CSS1EEPT PATH   IPADDR=9.82.5.120, GRPNM=GRPEEP
```

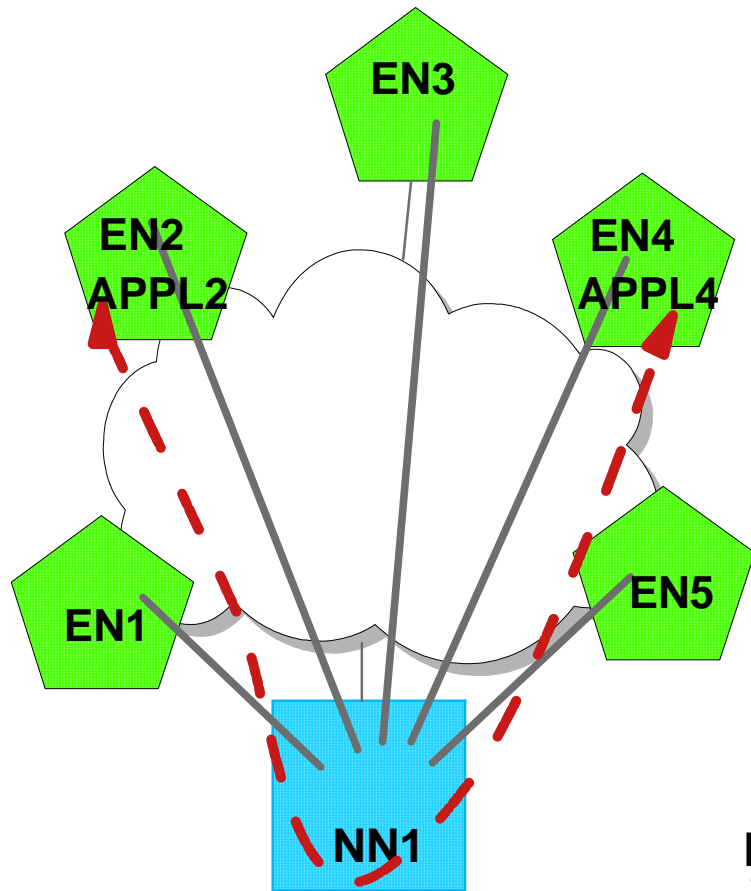
- EE XCA specifies Connection Network Link TGPs (or link parms):

```
XCAEE      VBUILD TYPE=XCA
PORTEE     PORT   MEDIUM=HPRIP
GRPEE      GROUP  DIAL=YES, AUTOGEN=(10,E,X), DYNPU=YES, *
                CALL=INOUT, ISTATUS=ACTIVE, *
                VNNAME=CSSNET.HPRIP, TGP=TRING16M
```

Connection Network

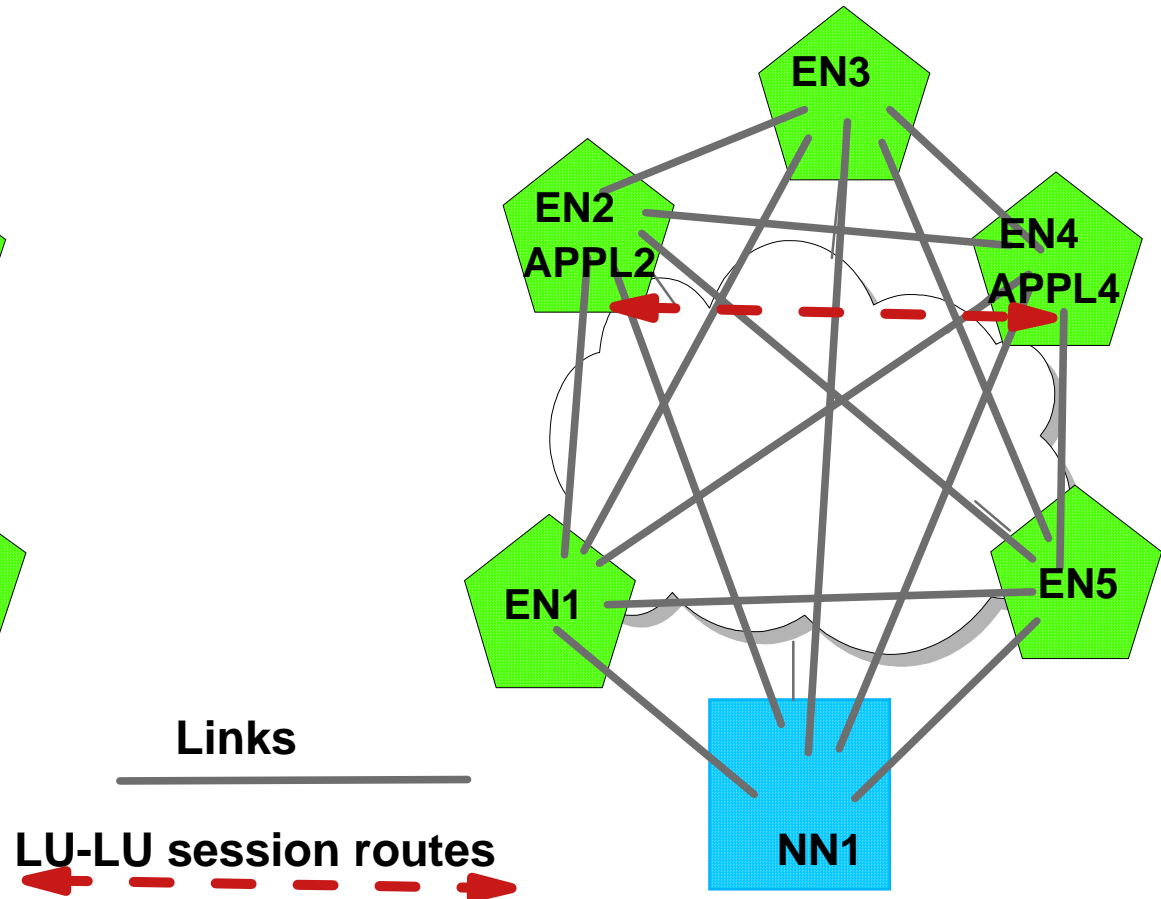
APPN TGs without Connection Network

Few Link Definitions,
But Routing not Optimal



Sessions take 2 Hops

Routing Optimal, But
Too Many Link Definitions

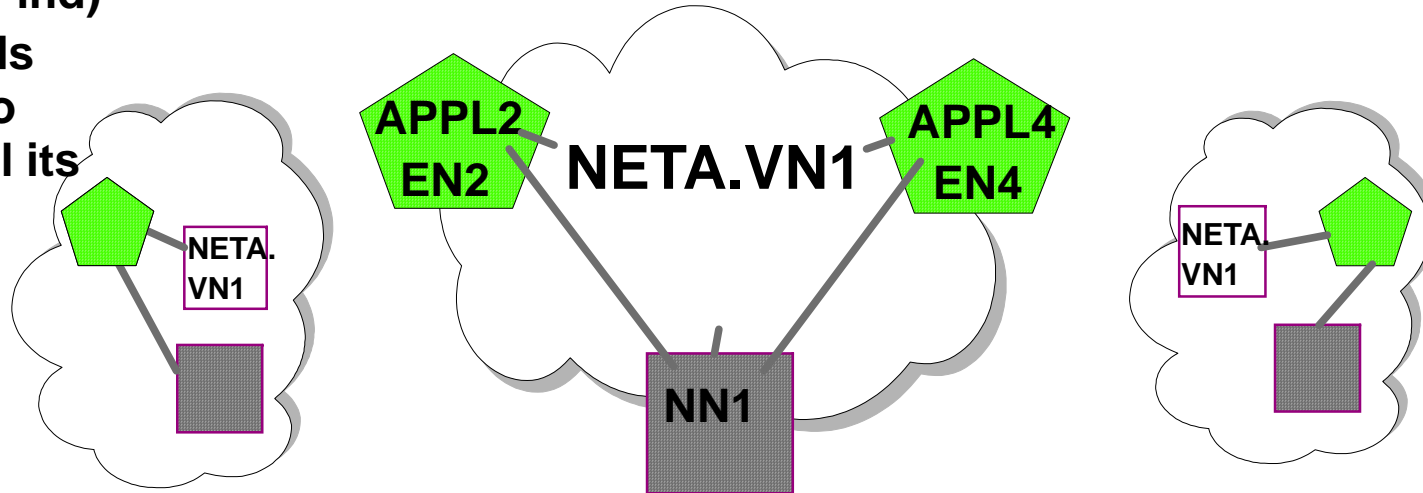


Sessions take 1 Hop

- These nodes are on a common network such as an IP network or a LAN
- Common network also referred to as an SATF - Shared Access Transport Facility

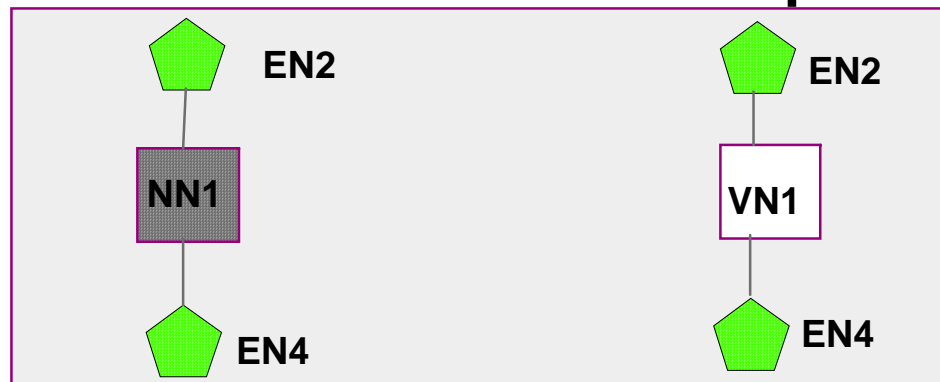
Connection Network Description

Locate(Find)
EN sends
DLC info
about all its
links



Locate(Found)
EN replies
with DLC info
about all its
links

NN understands there are two possible routes:



- NN1 picks "least weight route"

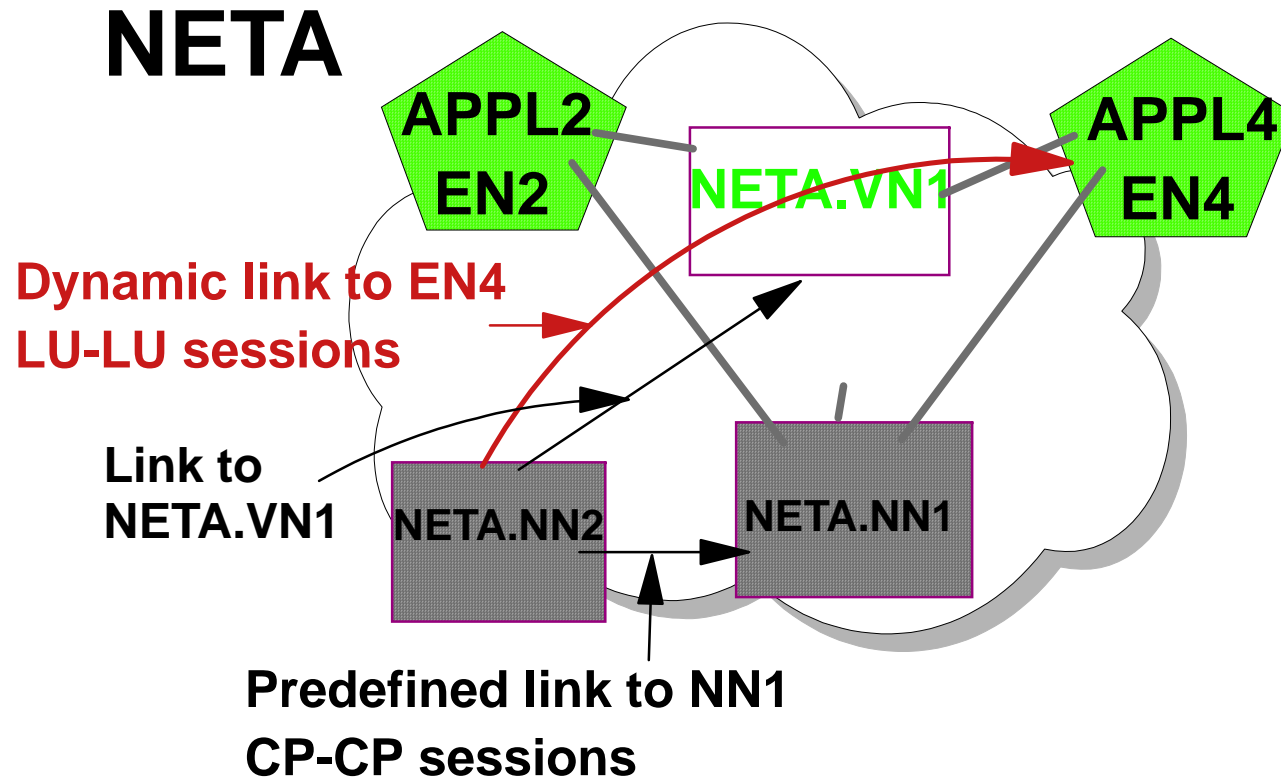
- ▶ If connection network is chosen, NN passes ENs DLC info (for EE it is partner's IP address)
- ▶ EN2 and EN4 bring up "dynamic link" for life of session(s)
- ▶ With NNs, DLC info about connection network links is in topology database

- Connection Network is also referred to as VRN - Virtual Routing Node

Connection Network Details

- Connection Network links are NOT used for CP-CP sessions
 - ▶ ENs predefine a link to their network node server, and usually another link to a backup network node server
 - This predefined link is used for CP-CP sessions
 - Dynamic connection network links are used for LU-LU sessions and the links normally come down when all LU-LU sessions using them have ended
 - ▶ NNs predefine links to adjacent NNs
 - These predefined links are used for CP-CP sessions with other NNs
 - Dynamic connection network links can be used for LU-LU sessions between NNs and other nodes (ENs or NNs) which define the SAME connection network name

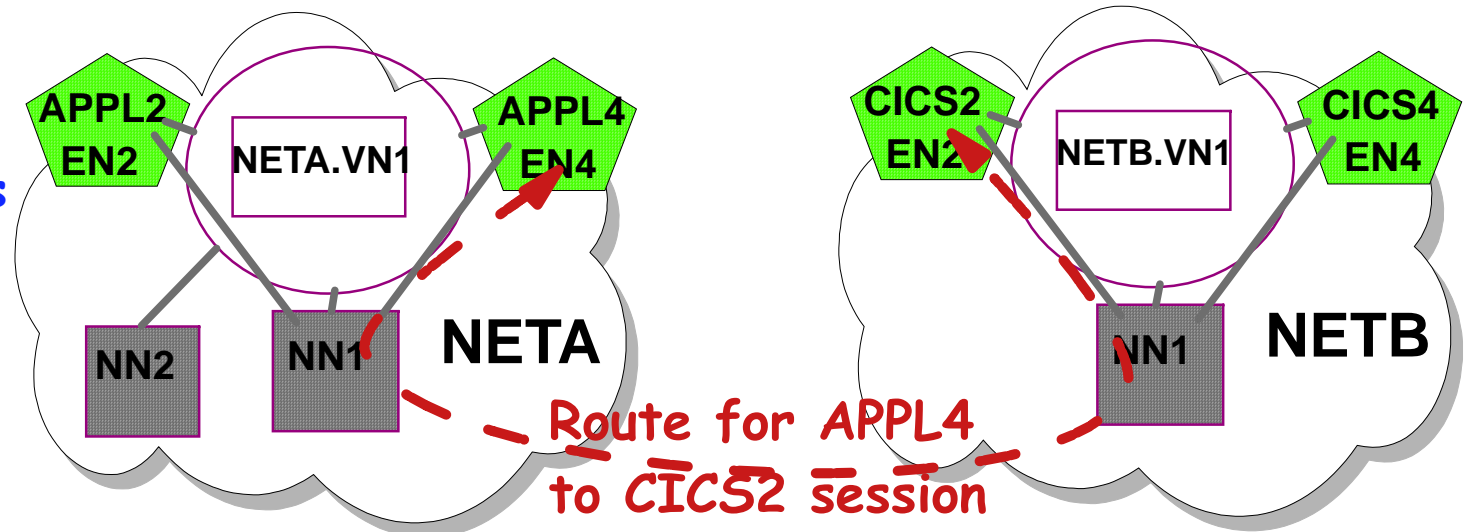
Connection Network-Adding a new node



- NN2 can communicate with all nodes on same SATF by simply defining link to other NN and the same Connection Network name
- NN2 predefines link address of NN1
 - ▶ Link used for CP-CP sessions, topology updates, locate requests, etc. and can also be used for LU-LU sessions
- NN2 defines link to virtual node NETA.VN1
 - ▶ Dynamic links can be used for LU-LU sessions with any other node on the same connection network, NETA.VN1

Connection Network - NETIDs

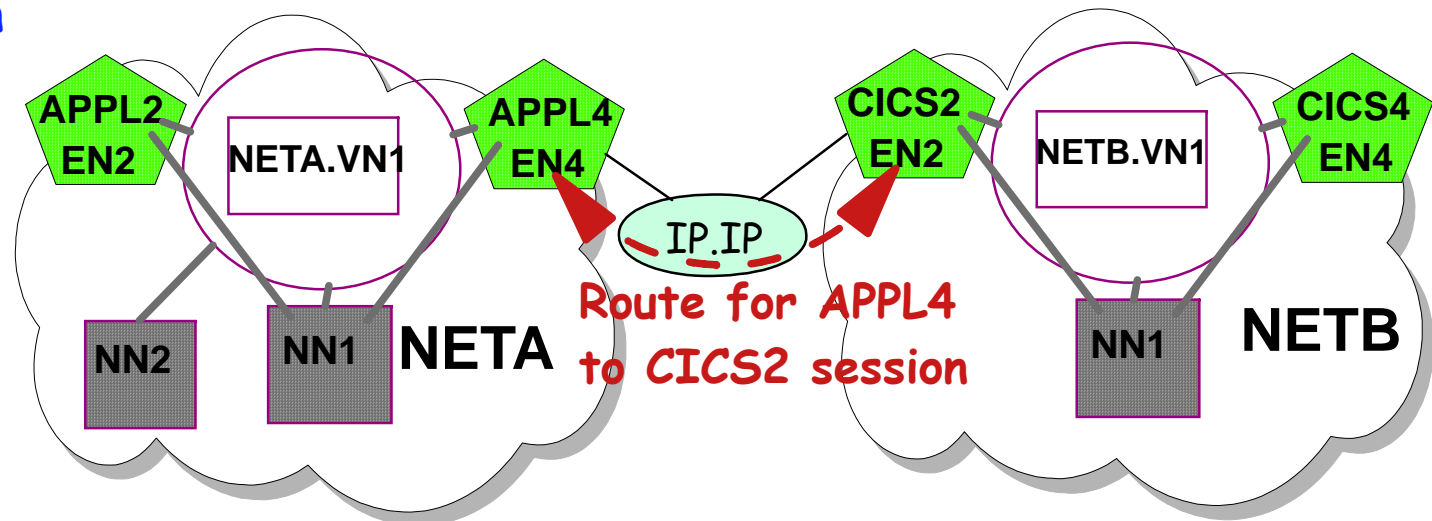
- When NNs have different NETIDs, border node functions are required to allow sessions to traverse different networks



- Prior to EE Global Connection Network, although Border Node allows sessions between resources in separate APPN networks, Connection Network links cannot be used between separate networks

- With EE Global Connection Network (GVRN), Connection Network links may be used between nodes in different APPN networks

- ▶ GVRN Applies only to Enterprise Extender SATFs, not other types of SATFs



Planning for Enterprise Extender

Planning for Enterprise Extender

- Products which have Enterprise Extender:
 - ▶ OS/390 V2R6 and higher
 - ▶ Communications Server for AIX, Linux, Windows, and OS/2
 - ▶ PCOMM
 - ▶ Cisco SNA Switch
 - ▶ Microsoft HIS
- Some products have EE and Branch Extender which can be used together
 - ▶ Can simplify APPN network topology
- Some products have EE Global Connection Network
 - ▶ Communications Server for z/OS V1R2
 - ▶ Enabled in Cisco SNASw with ddtS CSCdu65439

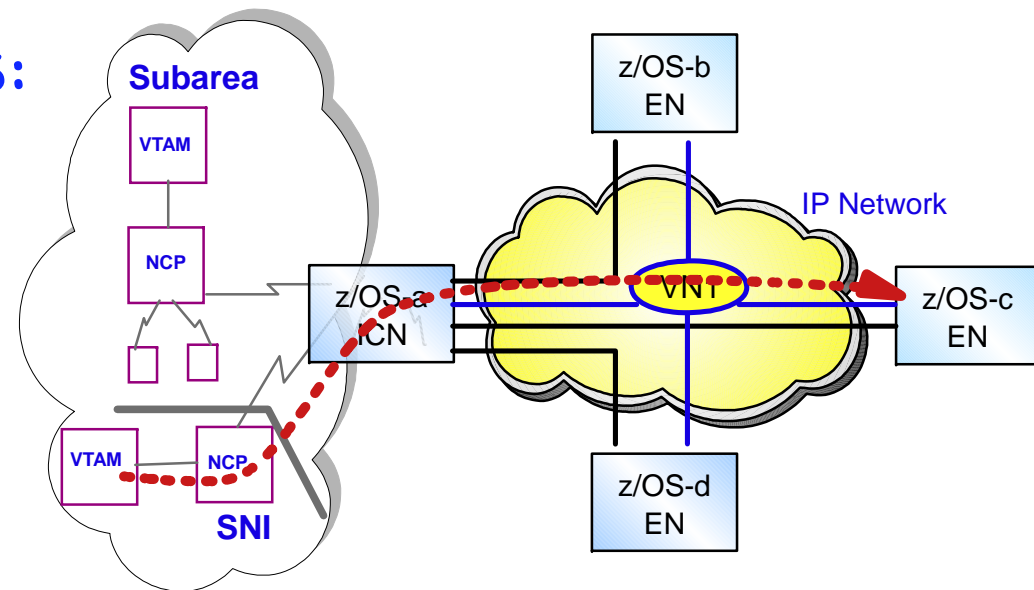
Enterprise Extender Planning

- **Planning issues**

- ▶ **Must implement VTAM APPN first and understand subarea/APPN interoperability**
 - This may be an educational/skill issue
- ▶ **IP Routing and Addressing**
 - Virtual IP address is required
 - Dynamic Routing should be used to allow redundancy
- ▶ **APPN Link Weights**
 - TGP's for EE provided with VTAM, but capacity values of these TGP's are low for most of today's environments
- ▶ **Router setup if prioritization in network is desired**

EE Evolution on OS/390 & z/OS

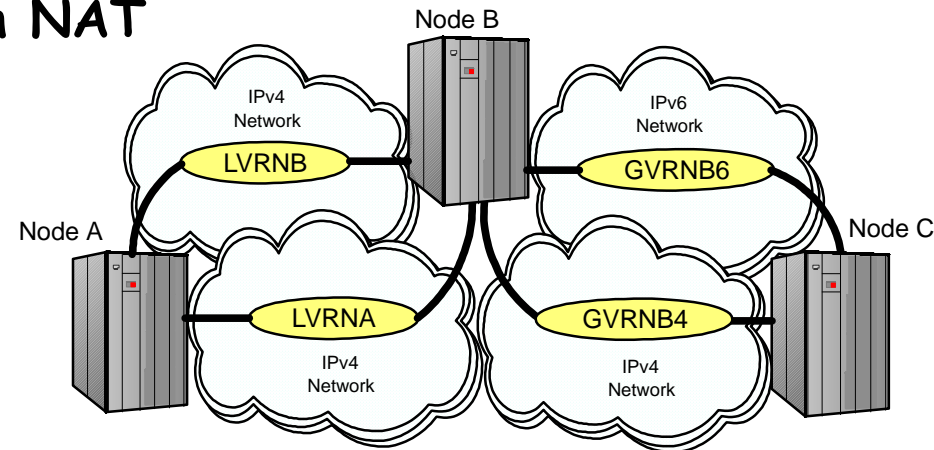
- Enterprise Extender was first made available on the host in CS for OS/390 V2R7 (or via PTF to V2R6)
- CS for OS/390 V2R10 addressed a significant limitation by allowing the use of EE links for interchange sessions.
 - Also added DYNTYPE=VN and DYNTYPE=RTP model PU support
- z/OS CS V1R2 added support for Global Connection Network and HPR Route Test
- EE enhancements in z/OS V1R4 CS:
 - Dial usability improvements
 - Enhanced diagnostics on RTP PU display
 - Lifts restriction on defining EE connection networks at ICNs
 - LINES and PUs in EE XCA major node are allocated from the high-order element address space (up to 16M line)
 - Maximum value on AUTOGEN parameter in EE XCA major node is raised from 4K to 64K



EE Evolution on z/OS...

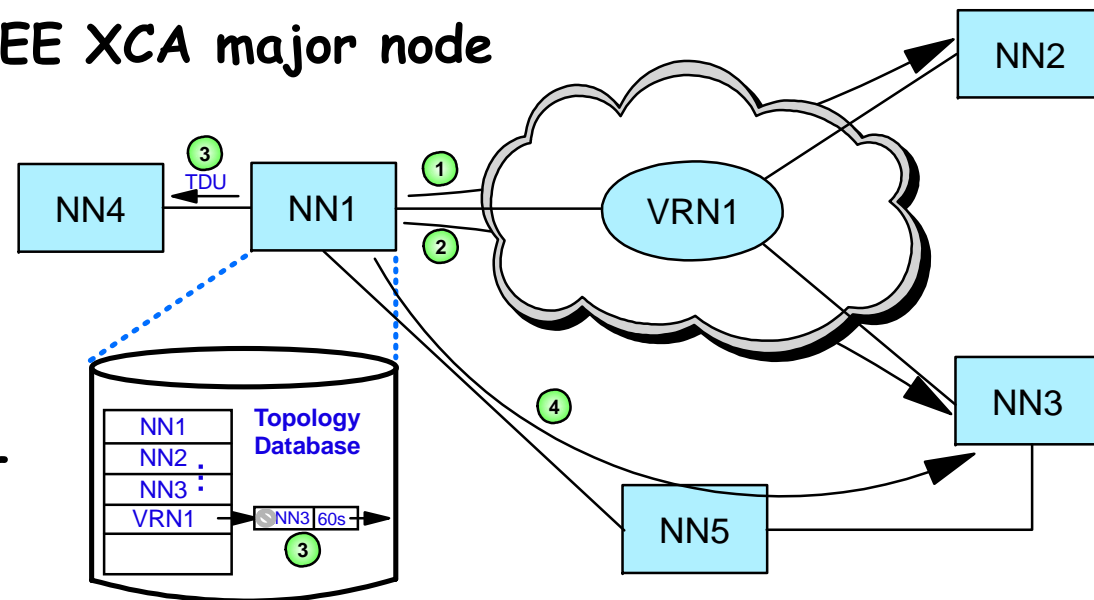
• EE enhancements in z/OS V1R5 CS:

- ▶ EE connection network compatibility with NAT
- ▶ IPv6 support for EE
- ▶ Support for multiple EE virtual routing nodes
- ▶ EE model PU support (DYNTYPE=EE)



• EE enhancements in z/OS V1R6 CS:

- ▶ VARY ACT,UPDATE support for EE XCA major node
- ▶ Display EE command
- ▶ D RTPS enhancement
 - TCID operand
- ▶ EE Connection Network Reachability Awareness
- ▶ Enhanced D ID=<RTP PU> output
- ▶ RTP PU element addresses are allocated from high-order space



• EE enhancements in z/OS V1R7 CS:

- ▶ Display EEDIAG command



SNI to EE/EBN Considerations

- Partner network also needs to define APPN, BN, EE
 - ▶ VSE and VM VTAM do not provide EE
 - ▶ IP addressing coordination
- With SNI, sessions between nodes in different networks normally route through 3745s directly to partners
 - ▶ With Border Node - sessions may route through VTAMs acting as BNs (HPR routing)
 - Global Connection Network should be considered
 - ▶ Security considerations include:
 - Use of NAT (incompatible with EE VRN prior to z/OS V1R5 CS)
 - SME changes or use of DSME instead
 - Firewalls must allow UDP packets on ports 12000-12004

Summary

- **Understanding EE: Concepts and Considerations**

- ▶ **SNA: Dead or Alive?**
- ▶ **Enterprise Extender - What is it?**
- ▶ **High Performance Routing (HPR) Concepts**
- ▶ **Link Characteristics and TGPs**
- ▶ **Connection Network**
- ▶ **Planning for Enterprise Extender**

- **Understanding EE: Nuts and Bolts**

- ▶ **HPR and EE: Layers and Timers**
- ▶ **EE Enhancements in V1R5**
- ▶ **Coding and Operating EE on z/OS**
- ▶ **Testing and Migration for EE**
- ▶ **EE Performance**