Summer 2010 – Session 6886
A journey through the layers of EE
How to translate VTAM messages into IP talk

Matthias Burkhard
mburkhar@de.ibm.com
IBM Germany

Wednesday, August 4, 2010: 4:30PM- 6:00 PM
Hynes Convention Center, Room 109

Session Content

When it is not enough to present a set of VTAM messages to your network provider to solve a HPR problem in the IP network, the time has come to learn a second language.
Join this session to learn how

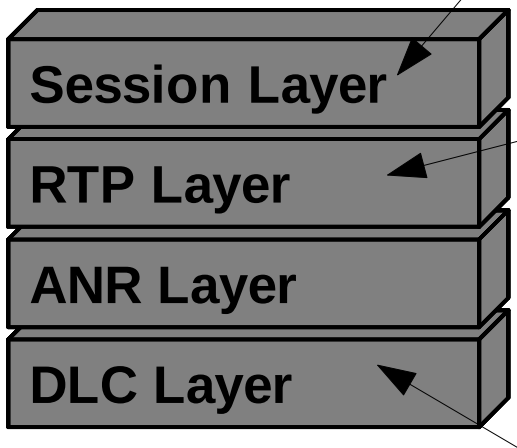
```
IST1494I PATH SWITCH STARTED FOR RTP CNR00062 TO netid.cpname
IST1818I PATH SWITCH REASON: SHORT REQUEST RETRY LIMIT EXHAUSTED
```

translates into the '4 and a half UDP Firewall Filter Rule' problem.
Come and understand why a High Performance Routing (HPR) pipe is sometimes performing Low and is still not using LPR protocol.
Get to know the underlying architecure of HPR: RTP, ANR, ARB
Walk with us through the layers of an EE packet and say Hello to all the bits and bytes that you better call a friend from now on. It will speed up problem resolution in heterogenous networks as most problems that result in nasty VTAM messages are not to be solved within z/OS!

EE problem categories
The external symptoms of EE problems

- Connectivity issues
  - Links don't set up / Links INOP
- Session hang problems
  - Sessions don't setup (PSEST, PBIPLUBF)
  - Sessions hang X-Clock
  - Sessions don't terminate (PSESEND)
- HPR PATHSWITCH
  - SRQ Retry Limit Exhausted
- Performance Problems
  - Slowdown and Retransmissions
  - CPU utilisation

VTAM messages
The layers that trigger them

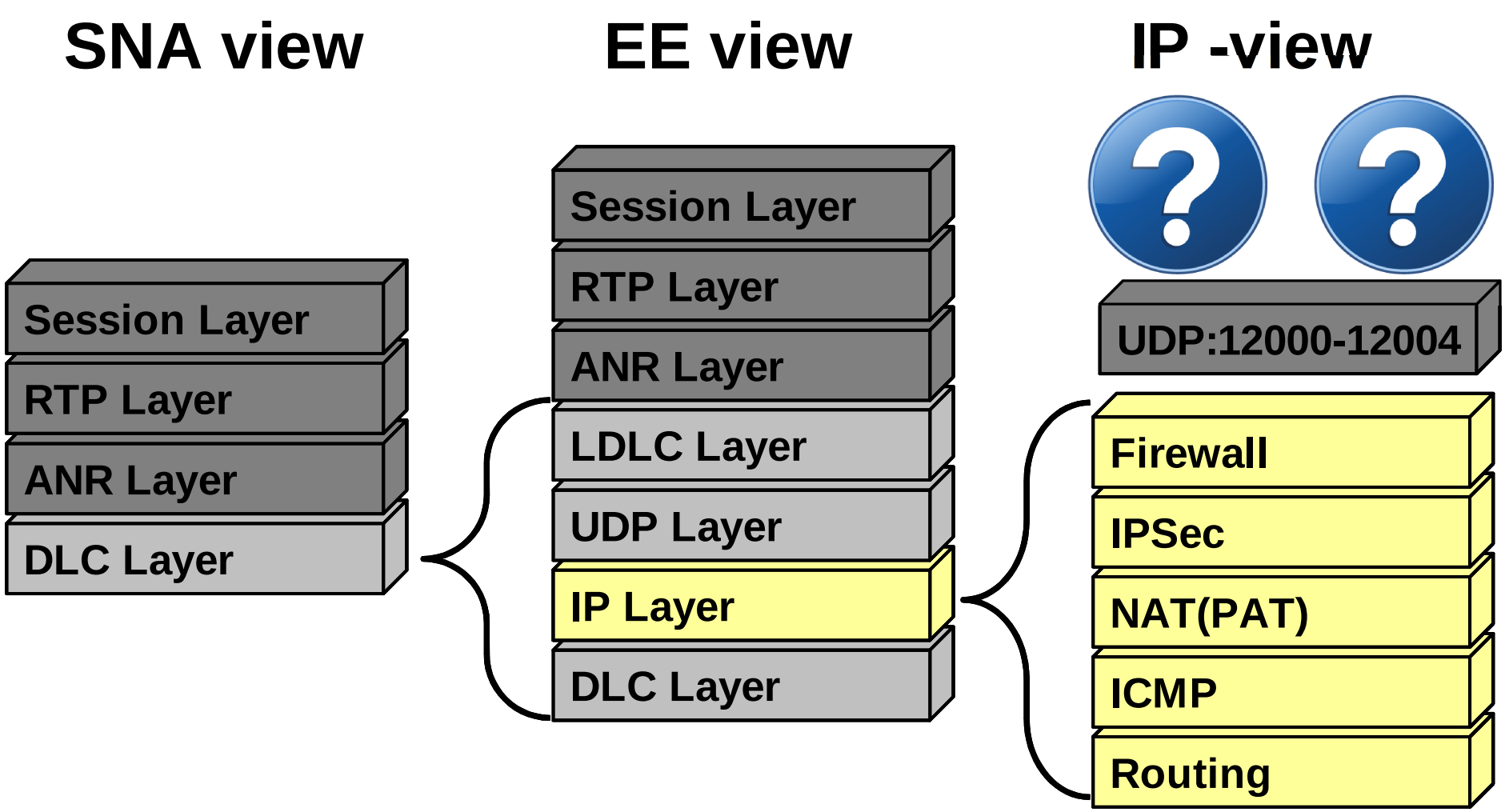



```
IST1097I  CP-CP SESSION WITH APPN.AS400NN     TERMINATED
IST1280I  SESSION TYPE = CONLOSER  - SENSE = 80020000
IST314I  END
```

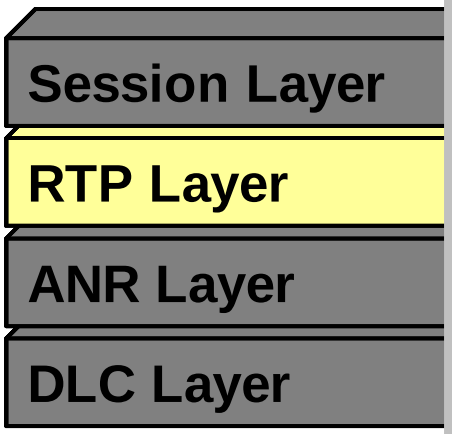
```
IST1494I PATH SWITCH STARTED FOR RTP CNR0F621 TO APPN.AS40
IST1818I PATH SWITCH REASON: SHORT REQUEST RETRY LIMIT EXH

IST1494I PATH SWITCH FAILED FOR RTP CNR0F621 TO APPN.AS400
IST1495I NO ALTERNATE ROUTE AVAILABLE
```

```
IST1411I INOP GENERATED FOR E000000B
IST1430I REASON FOR INOP IS XID OR LDLC COMMAND TIMEOUT
IST314I END
IST1141I SOFT INOP FOR LCBCP1 OVERRIDDEN BY SOFT INOP
```

Enterprise Extender
Just another HPR-only APPN DLC type?



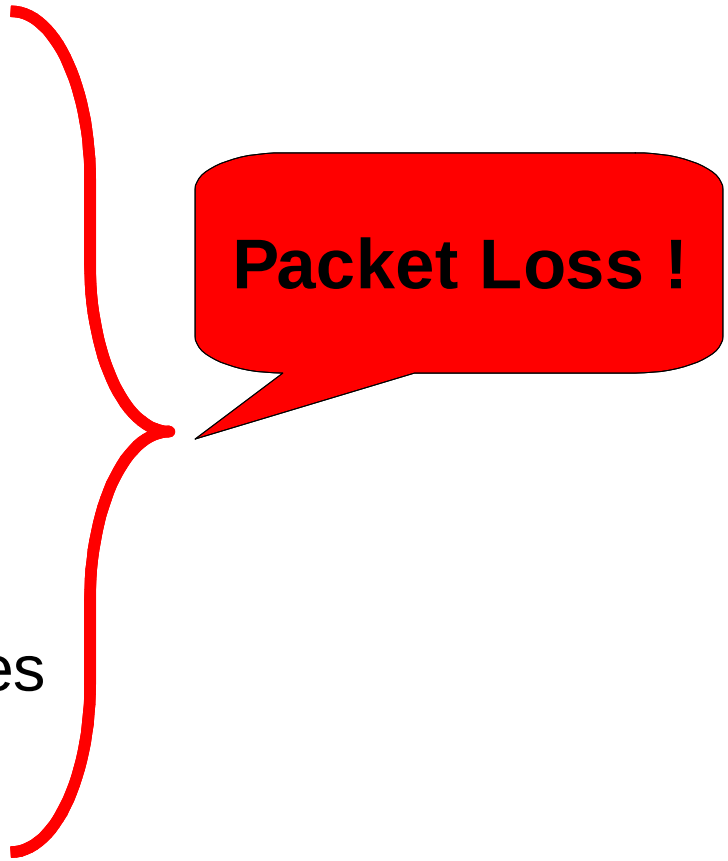
HPR Pipe Display – Lots of Information




```
IST1968I  ARB INFORMATION:
IST1844I  ARB MODE = RED
IST1477I  ALLOWED DATA FLOW RATE =     217 KBITS/SEC
IST1516I  INITIAL DATA FLOW RATE =    3200 KBITS/SEC
IST1841I  ACTUAL DATA FLOW RATE =        0  BITS/SEC
IST1969I  MAXIMUM ACTUAL DATA FLOW RATE =        0  BITS/SEC
IST1862I  ARB MAXIMUM SEND RATE =      32 MBITS/SEC
...
IST1970I  RATE REDUCTIONS DUE TO RETRANSMISSIONS =     0
IST924I   -------------------------------------------------
IST1971I  TIMER INFORMATION:
IST1852I  LIVENESS TIMER =         180 SECONDS
IST1851I  SMOOTHED ROUND TRIP TIME =         400 MILLISECONDS
IST1972I  SHORT REQUEST TIMER =        3200 MILLISECONDS
IST924I   -------------------------------------------------
IST1973I  OUTBOUND TRANSMISSION INFORMATION:
IST1974I  NUMBER OF NLPS SENT =                1421323 (   1M )
IST1975I  TOTAL BYTES SENT =             1153218385 (   1G )
IST1849I  LARGEST NLP SENT =        1039 BYTES
IST1980I  SEQUENCE NUMBER = 1092124629 (X'41187FD5')
IST1842I  NUMBER OF NLPS RETRANSMITTED =       8983
IST1976I  BYTES RETRANSMITTED =              10156534 (  10M )
IST1478I  NUMBER OF UNACKNOWLEDGED BUFFERS =      7
IST1958I  NUMBER OF ORPHANED BUFFERS =            0
IST1843I  NUMBER OF NLPS ON WAITING-TO-SEND QUEUE =     16
```

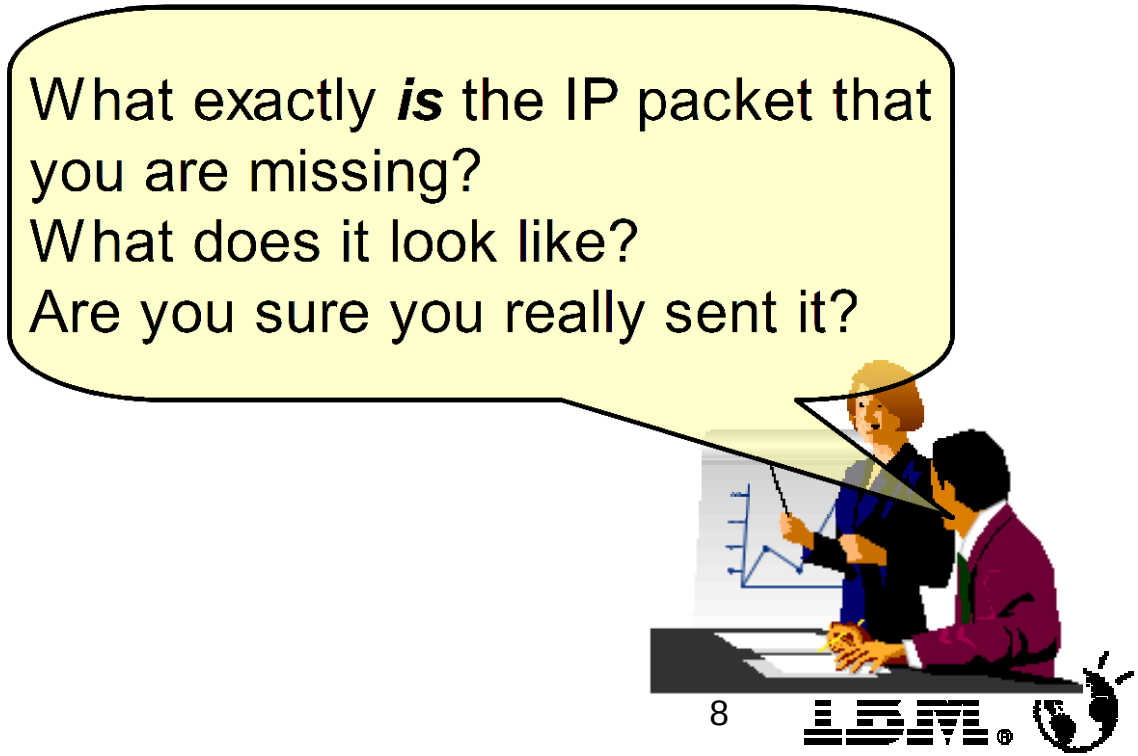
EE problem categories
The #1 root cause of EE problems

- Connectivity issues
  - IP routing issues
  - FW filter rules
- Session hang problems
  - IP Fragmentation
  - IPSec tunnels
- HPR PATHSWITCH
  - FW connection tables / FW filter rules
- Performance
  - Retransmissions
  - Slowdowns , delays, congestion
  - Insufficient resources (storage/CPU)



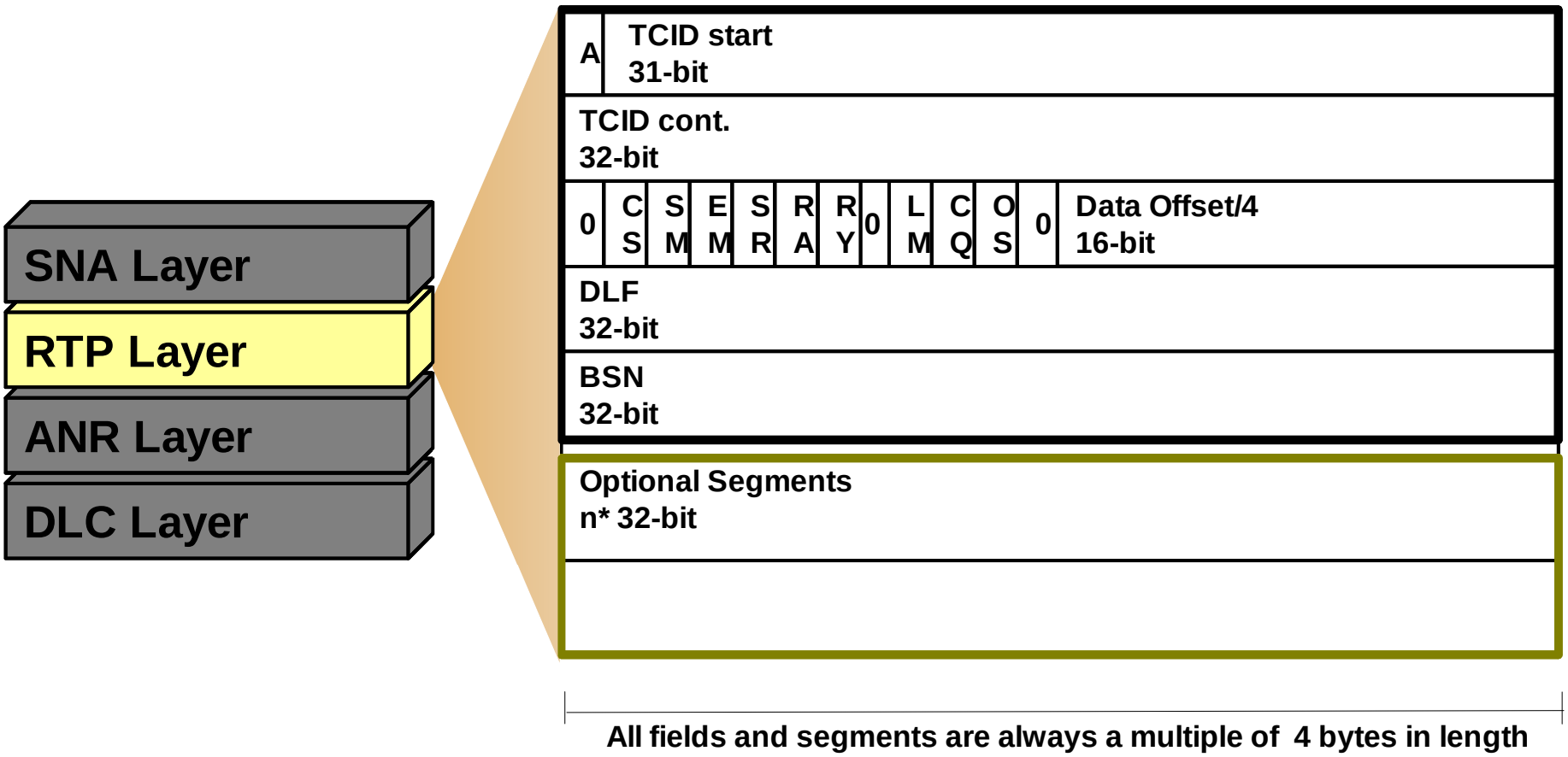
EE problems
The top 3 questions

- Where did my packet get lost ?
  - Locally (close to the source)
  - Remotely (close to the destination)
  - Somewhere in between (network, firewalls ...)
- Why was it discarded?
  - Incorrect routing ?
  - Congestion ?
  - Security Policies ?
- How can I prove it?
  - VTAM messages?
  - VTAM Traces?



RTP layer
Transport Header – Optional Segments

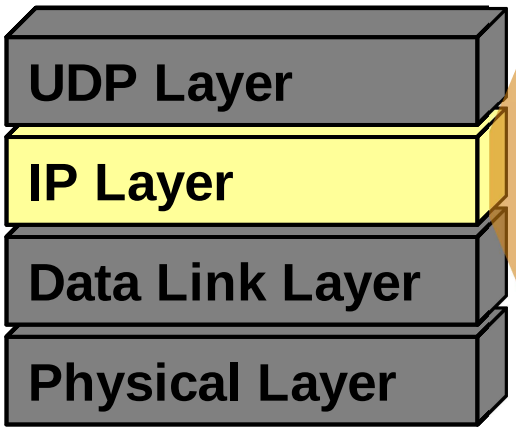
- Where HPR vocabulary lives



SETUP: 0D	CIDEX: 10	SWINFO: 14	ARB: 22	STATUS: 0E	COOB: 0F	CFAULT: 12

The IP header
IP Header and additional protocols

- Where IP vocabulary lives



version 4-bit	header length	3-bit TOS field 5-bit reserved	total length 16-bit			
identification 16-bit			0	D F	M F	fragment offset 13-bit
time to live TTL 8-bit	protocol 8-bit		header checksum 16-bit			
source IP address 32-bit						
destination IP address 32-bit						
options e.g. record route, timestamp, padding, etc. maximum 40 bytes						

- and

DNS	OSPF	ICMP	IPSec / VPN	MTU size	Firewalls	NAT

A typical EE packet
What this session will focus on

- **IP Header**
  - Length
  - **Identifier**
  - **Flags/FragOffs**
  - **TTL**
  - **IP addresses**
- **ICMP**
- UDP Header
  - Ports
  - Length
  - Checksum
- LDLC
  - SAPs
  - Control
    - XID/TEST/UI/DISC
- HPR
  - NHDR
  - **THDR**
  - **Optional Segments**
- SNA PIU (TH,RH,RU)
  - Sense Codes
  - FMH7s

IP Header
IP_ID and addresses

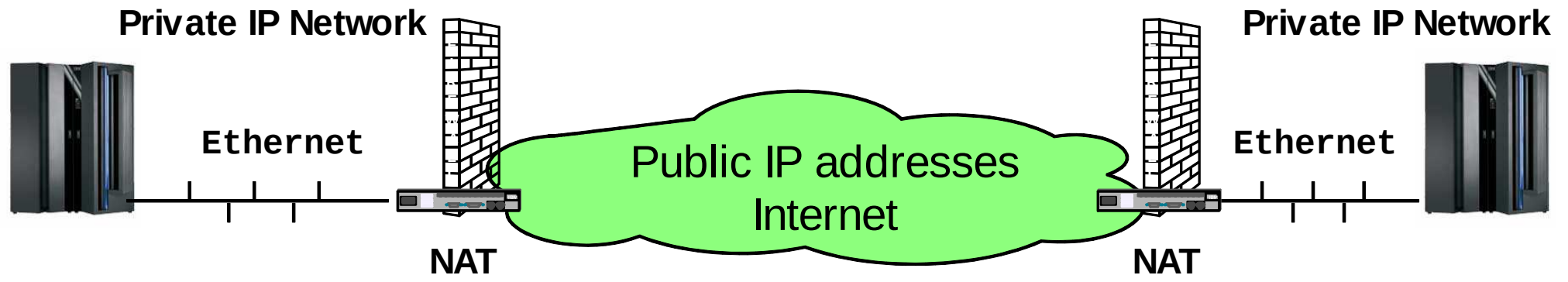
- **IP Header**
  - TOS precedence
  - Length
  - **Identifier**
  - Flags/FragOffs
  - TTL
  - **IP addresses**

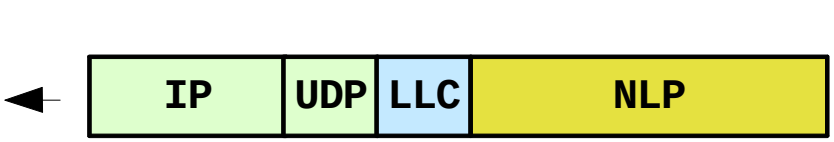
4	5	3-bit TOS field 5-bit reserved	total length 16-bit			
identification 16-bit			0	D F	M F	fragment offset 13-bit
time to live TTL 8-bit		protocol 8-bit	header checksum 16-bit			
source IP address 32-bit						
destination IP address 32-bit						

- **IP_ID is a 2 byte field incremented with every new packet generated at the sending stack (unique until it wraps)**
- **IP addresses: in IPV4 (header 45) 4 byte long**
  - **Can be changed by NAT devices in flight**

IP Identifier
Following a packet through the network

- NAT will change the addresses – the IP_ID will stay the same



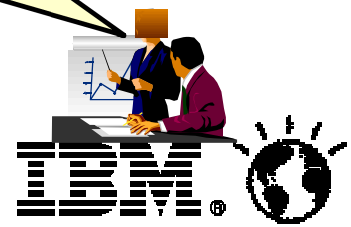


ip.src:192.168.105.166,ip.id:7DB0

ip.src:194.60.161.49,ip.id:7DB0

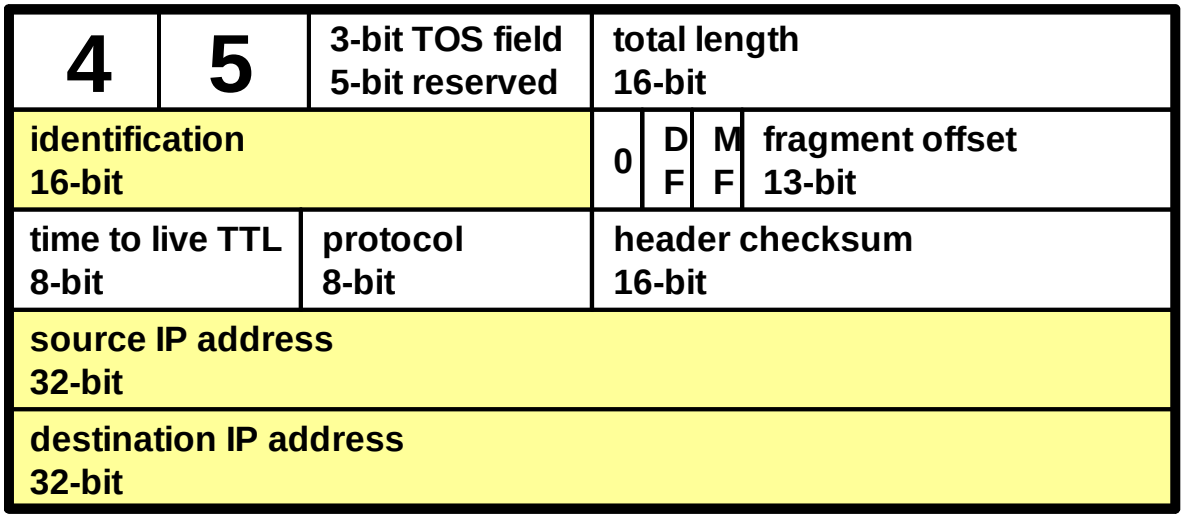
Here it is:
IP.ID 7DB0!

Where is my
IP.ID 7DB0!



IP Header
IP Identifier

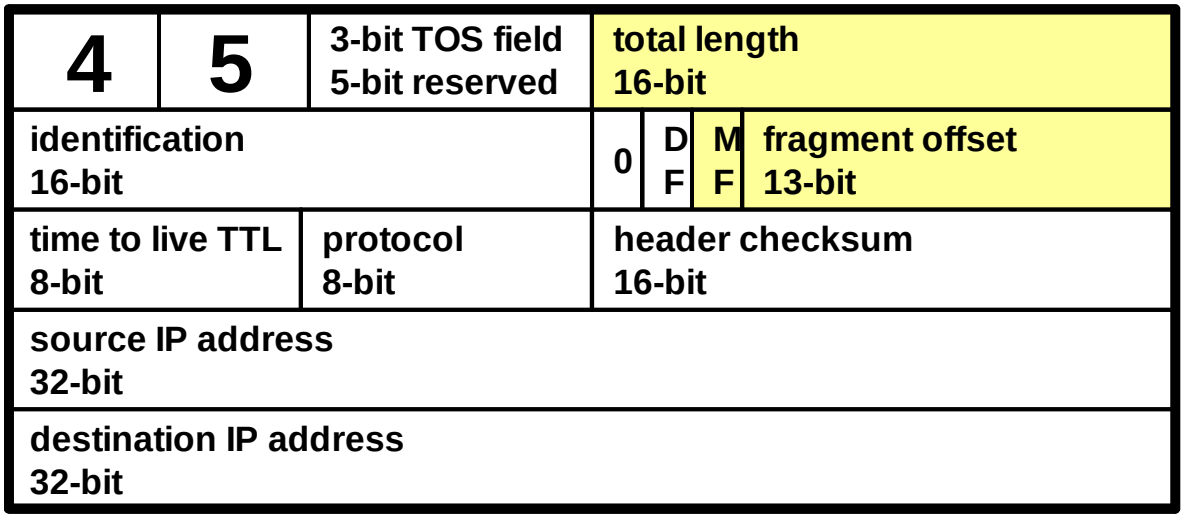
- Required to track down a packet
- Will not change in flight




```
                  IPID          10.186.86.243 --> 10.217.18.47
IP    45C00058 9C730000 3C1162AD 0ABA56F3 0AD9122F
UDP   2EE12EE1 0044D9DF
LLC   040403
NLP   C608 80 FF00
RTP   00000000100257B0004000D000000000000E2CD
      032259040000000000000000                              ARB REPLY
      050E0000004000200 00CF2F0000000000000000              STATUS
```

IP Header
IP length and Fragmentation

- **IP Header**
  - TOS precedence
  - **Length**
  - Identifier
  - **Flags/Frag_Offs**
  - TTL
  - IP addresses



- **The length indicates how large the IP datagram is**
- **If the datagram exceeds the MTU size of the weakest link an intermediate router will fragment the packet .**
- **Reassembly is then required at the destination IP stack.**

IP Fragmentation
Two fragments arriving

- IPID is the same
- MF flag is set in 1st fragment
- Fragment Offset is > 0
- No protocol header in 2nd Fragment

4	5	3-bit TOS field 5-bit reserved	total length 16-bit			
identification 16-bit			0	D F	M F	fragment offset 13-bit
time to live TTL 8-bit		protocol 8-bit	header checksum 16-bit			
source IP address 32-bit						
destination IP address 32-bit						

```
 IP_V4 Header
     lgth ID   ,-MF
45000304 5A482000 2F11BA0B 0AC7E821
0ABA56F3
 UDP Header
2EE22EE2 059C8EF5
 Logical Link Control Header
080403
```

```
 IP_V4 Header
     lgth ID    ,-- fgmt_offs
450002C0 5A48005E 2F11D9F1 0AC7E821
0ABA56F3
 IP FRAGMENT
40404040 40404040 40404040 40404040
40404040 40404040 40404040 40404040
40404040 40404040 40404040 40404040
```

IP Fragmentation
Issues and Path MTU discovery

- Fragmentation increases CPU at receiver
- Fragmentation causes an additional delay
- Fragmentation across Firewall infrastructure often not allowed
  - FW filter rules check on IP@ ,protocol and port numbers
    - 2nd fragment does not have port numbers
- Path MTU Discovery (PMTUD) available in VTAM V1R10
  - DF bit is set causing ICMP message from router if fragmentation is required
  - Contains MTU size of next hop

```
 IP_V4 Header
                 ,-DF
45C00058 275D4000 40110000 0A03000C
0A00021C
 UDP Header
2EE12EE1 00441680
 Logical Link Control Header
C80403
```

IP Fragmentation
Out of order arrival

- 2nd Fragment of 5001 arriving after 5002
- IP layer will reassemble the 2 fragments (if both arrive...)

```
Description       TTL IP Address   <> IP Address (+ PortN Iden Lengt
-------------------------------------------------------------------
EE_HIG FID5        48  10.186.86.24 <- 10.199.232.33(12002 4FFE   300
EE_HIG First Frag  47  10.186.86.24 <- 10.199.232.33(12002 5001   836
EE_HIG continued   48  10.186.86.24 <- 10.199.232.33(12002 5002   845
IP/FRAGMENT Last   47  10.186.86.24 <- 10.199.232.33        5001   768
EE_HIG continued   48  10.186.86.24 <- 10.199.232.33(12002 5003 1,488
EE_HIG continued   48  10.186.86.24 <- 10.199.232.33(12002 5004   830
EE_HIG FID5        48  10.186.86.24 <- 10.199.232.33(12002 5005 1,488
```

- The upper layer protocol needs to reorder the data
  - Reordering is done in TCP protocol, not in UDP
  - For Enterprise Extender , HPR RTP will perform this function

IP Header
TTL - Time To Live

- **IP Header**
  - TOS precedence
  - Length
  - Identifier
  - Flags/Frag_Offs
  - **TTL**
  - IP addresses

4	5	3-bit TOS field 5-bit reserved	total length 16-bit				
identification 16-bit			0	D F	M F	fragment offset 13-bit	
time to live TTL 8-bit	protocol 8-bit	header checksum 16-bit					
source IP address 32-bit							
destination IP address 32-bit							

- 1 byte field that controls how far an IP packet can travel
- It gets decremented by every router on the path
- When it reaches 1, the router will discard the packet

IP TTL
Initial TTL values

- The initial TTL value is configurable in every IP stack
  - Most IP stacks use the default though
    - Why not?

OS / Protocol	AIX	z/OS	i5OS	Win	Routers
TCP	x3C	x40	x40	x80	xFF
UDP	x1E	x40	x40	x80	xFF

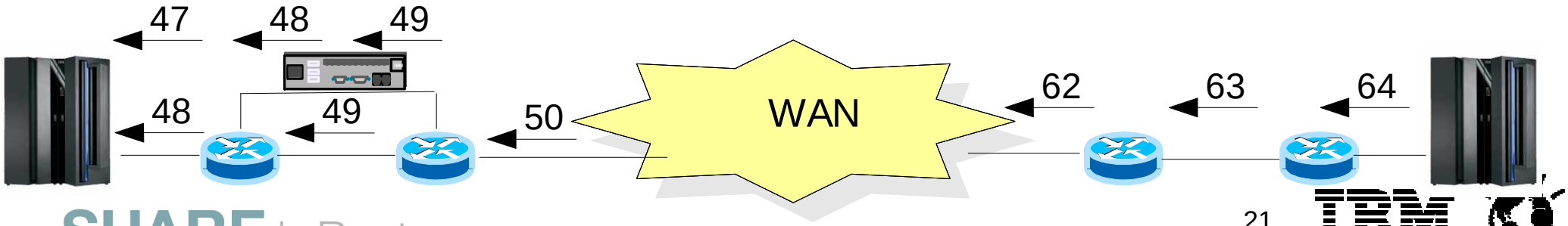
- Knowledge of the initial TTL at the source can be used to determine (guess) the operating system and the distance of a remote host (# of hops)
- An ICMP error message will be sent when a router discards a packet because of an inbound TTL of 1

IP TTL
Guessing the topology

- Packets arriving with a TTL of 48
  - The sending IP stack is 16 hops away (Initial TTL=64)

```
Description       TTL IP Address   <> IP Address (+ PortN Iden Lengt
-------------------------------------------------------------------
EE_HIG FID5        48  10.186.86.24 <- 10.199.232.33(12002 4FFE   300
EE_HIG First Frag  47  10.186.86.24 <- 10.199.232.33(12002 5001   836
EE_HIG continued   48  10.186.86.24 <- 10.199.232.33(12002 5002   845
IP/FRAGMENT Last   47  10.186.86.24 <- 10.199.232.33       5001   768
EE_HIG continued   48  10.186.86.24 <- 10.199.232.33(12002 5003 1,488
EE_HIG continued   48  10.186.86.24 <- 10.199.232.33(12002 5004   830
EE_HIG FID5        48  10.186.86.24 <- 10.199.232.33(12002 5005 1,488
```

- However, fragmented packets arrive with a TTL of 47



IP TTL
D NET,EEDIAG,TEST=YES

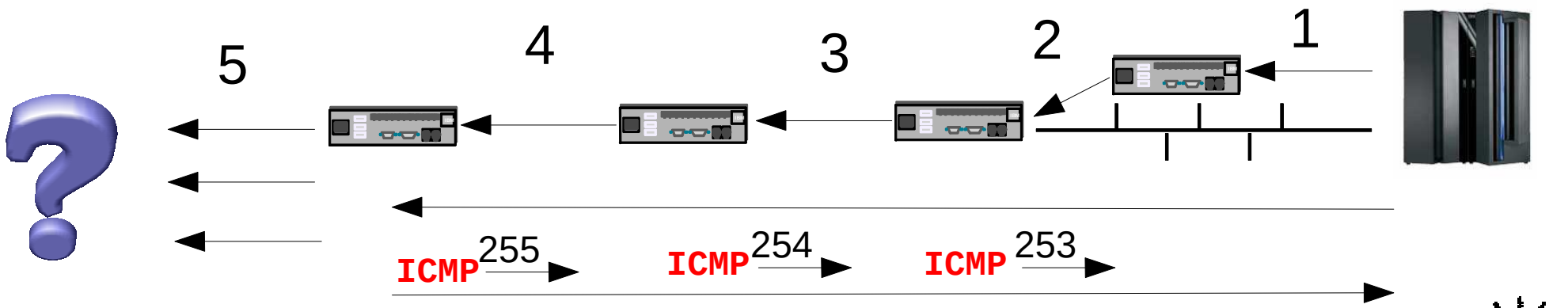
- Sets TTL purposely too short to learn the IP route's RTT

```
D NET,EEDIAG,TEST=YES,LIST=ALL,IPADDR=(198.239.65.137,198.238.232.3)
IST350I DISPLAY TYPE = EEDIAG
IST2130I ENTERPRISE EXTENDER CONNECTIVITY TEST INFORMATION
IST2119I ENTERPRISE EXTENDER DISPLAY CORRELATOR: EE00001A
IST2131I EEDIAG DISPLAY COMPLETED ON 06/26/09 AT 08:02:26
IST2132I LDLC PROBE VERSIONS: VTAM = V1          PARTNER = UNKNOWN
IST1680I LOCAL IP ADDRESS 198.239.65.137
IST1680I REMOTE IP ADDRESS 198.238.232.3
IST924I -------------------------------------------------------
IST2133I INTFNAME: L113420                    INTFTYPE: IPAQENET
IST2135I   CONNECTIVITY UNSUCCESSFUL    SENSE: ***NA***   PORT: 12000
IST2136I   CONNECTIVITY TEST ENDED - MAXIMUM TIME LIMIT EXCEEDED
IST2137I     1  *                   (3)      RTT: *N/A*
IST2137I     7  *                   (3)      RTT: *N/A*
IST2135I   CONNECTIVITY UNSUCCESSFUL    SENSE: ***NA***   PORT: 12001
IST2136I   CONNECTIVITY TEST ENDED - MAXIMUM TIME LIMIT EXCEEDED
IST2137I     1  *                   (3)      RTT: *N/A*
```

IP TTL
D NET,EEDIAG,TEST=YES Trace

- ICMP TTL_Timeout messages required to come back

```
 Delta   Description (short)  TTL IP Address      Iden <> IP Address (+ PortNu
4.668.333 EE_CMD               1  198.238.232.3        <- 198.239.65.137(12000
0.000.169 ICMP/TIMEOUT (TTL)  255 147.55.217.3   B7C0 -> 198.239.65.137
0.000.328 EE_CMD               2  198.238.232.3        <- 198.239.65.137(12000
0.000.212 ICMP/TIMEOUT (TTL)  255 147.55.215.2   5906 -> 198.239.65.137
0.000.897 EE_CMD               3  198.238.232.3        <- 198.239.65.137(12000
0.000.712 ICMP/TIMEOUT (TTL)  254 147.55.215.6   430C -> 198.239.65.137
0.000.017 EE_CMD               4  198.238.232.3        <- 198.239.65.137(12000
0.000.878 ICMP/TIMEOUT (TTL)  253 10.2.3.3       3B31 -> 198.239.65.137
0.000.010 EE_CMD               5  198.238.232.3        <- 198.239.65.137(12000
3.301.891 EE_CMD               5  198.238.232.3        <- 198.239.65.137(12000
```



ICMP
Listen to the network music

- ICMP protocol is used in IP to
  - Test connectivity (PING)
  - Report errors
    - Destination unreachable
    - Time-out conditions
  - Propagate information
    - MTU size of next hop with PMTU Discovery
- ICMP packets are very important in diagnosis
  - Often not allowed through secured infrastructure
    - Firewall rules block ICMP in general
  - Often not traced because of trace filters
    - Source IP address of ICMP packets not predictable
- PMTUD (V1R10) depends on receipt of ICMP messages

ICMP
Error message

- ICMP protocol type 1
- Contains Error information in the ICMP header
  - 0301 – cannot route any further, host unreachable
- Contains the original IP header that caused this error
  - As seen at the ICMP sender (IPID:B31A, TTL=3A, UDP,12000)

```
IN  149.83.5.17 <- 172.17.60.1 ICMP/DESTUNR(03):Host Unreachable(01)
 IP_V4                 IP_V4 Header
                0000 45000038 B5210000 F9018A2C AC113C01
                0010 95530511
 ICMP                  ICMP Internet Control and Messaging
                0000 030156FF 00000000
 IP_V4                 IP_V4 Header
                0000 45C000A0 B31A0000 3A111952 95530511
                0010 AC1B6CA1
 ICMP_IPFRAG_DATA      ICMP_IPFRAG
                0000 2EE02EE0 008C47B3
```

A typical EE packet
What's left ...

- **IP Header**
  - Length
  - **Identifier**
  - **Flags/FragOffs**
  - **TTL**
  - **IP addresses**
- **ICMP**
- UDP Header
- Ports
- Length
- Checksum

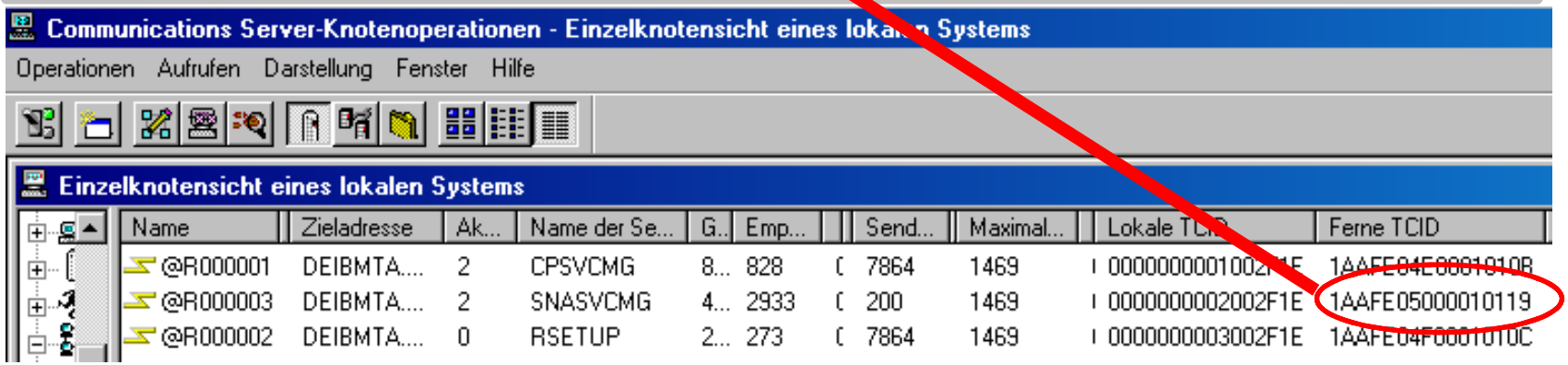
- LDLC
  - SAPs
  - Control
    - XID/TEST/UI/DISC
- HPR
  - NHDR
  - **THDR**
  - **Optional Segments**
- SNA PIU (TH,RH,RU)
  - Sense Codes
  - FMH7s

VTAM
HPR Pipes – RTP PUs - TCIDs

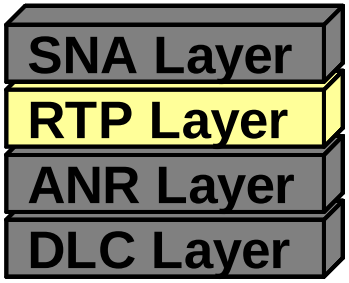
- VTAM knows a pipe by a PU name in ISTRTPMN
  - Typically starts with CNRxxxxx
- The names are different, depending on the RTP node
  - If the remote RTP is also a VTAM, it will have another CNRxxxxx name
  - If it is a distributed SNA stack, the name will be
    - @Rnnnnnn
- The TCIDs must be used to correlate the display output
  - The 'local TCID' here is the 'remote TCID' at the other end
- Both TCIDs must be used to follow a pipe's traffic in a trace
  - All NLPs carry the receiver's local TCID

RTP - TCID
Different Names, same TCIDs

```
IST097I DISPLAY ACCEPTED
IST350I DISPLAY TYPE = RTPS
IST1695I PU NAME   CP NAME      COSNAME SWITCH CONGEST STALL SESS
IST1960I CNR000C4 BROWN.KEN    SNASVCMG  NO      NO      NO      2
IST1960I CNR000C3 BROWN.KEN    RSETUP    NO      NO      NO      0
IST1960I CNR000C2 BROWN.KEN    CPSVCMG   NO      NO      NO      2

D NET,ID=CNR000C4,E
IST097I DISPLAY ACCEPTED
IST075I NAME = CNR000C4, TYPE = PU_T2.1 667
IST1043I CP NAME = KEN - CP NETID = BROWN - DYNAMIC LU = YES
IST1962I APPNCOS = SNASVCMG - PRIORITY = NETWORK
IST1476I TCID X'1AAFE05000010119' - REMOTE TCID X'0000000002002F1E'
IST1481I DESTINATION CP BROWN.KEN - NCE X'80'
```



RTP THDR TCID

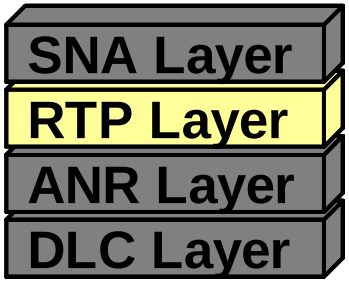


A	TCID start 31-bit													**29B2A158**
TCID cont. 32-bit														**000100D6**
0	CS	SM	EM	SR	RA	RY	0	LM	CQ	OS	0	Data Offset/4 16-bit		
DLF 32-bit														
BSN 32-bit														

- The TCID is the first 8 bytes in the THDR
  - It uniquely identifies the pipe at the receiving RTP node
  - To follow traffic in both directions, both TCIDs must be used when filtering a trace.

```
IP      45000063 41B10000 72118824 0AD9122F 0ABA56F3
UDP     2EE12EE1 004F3669
LLC     040403
ANR     C600 80 D900140100000 0 8002000C01000000 D000000000000000 FF00
RTP     29B2A158000100D60004000A000000000000001E1
STATUS  050E000000000020000010600000000000000000
```

RTP Transport Header THDR BSN



A	TCID start 31-bit											
TCID cont. 32-bit												
0	CS	SM	EM	SR	RA	RY	0	LM	CQ	OS	0	Data Offset/4 16-bit
DLF 32-bit												00000000
BSN 32-bit												000001E1

- The Byte Sequence Number keeps track of the data sent
  - Increments with every byte of payload (DLF field)
  - Also increments if End_of_Message bit is set

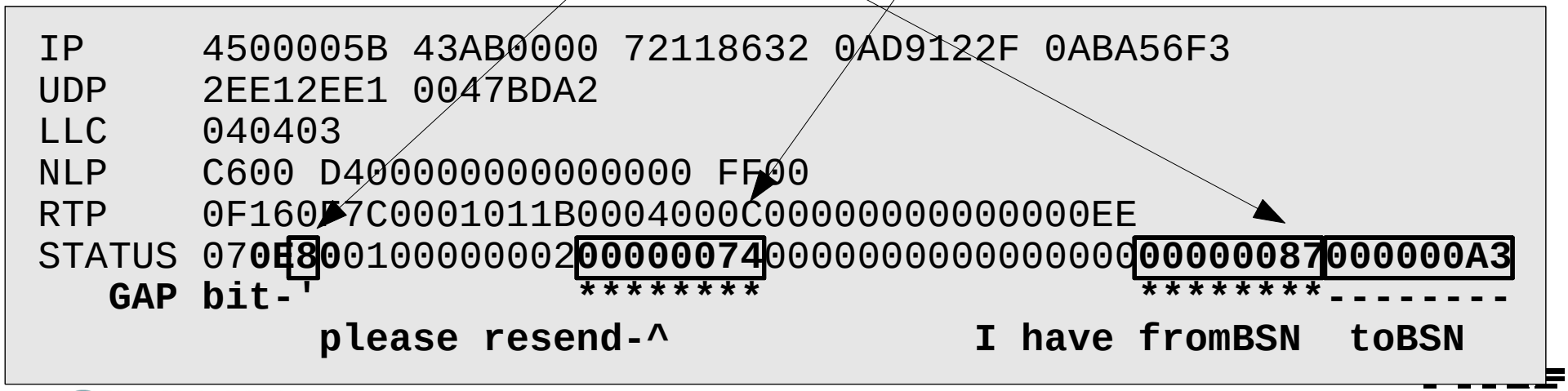
```
IP      45000063 41B10000 72118824 0AD9122F 0ABA56F3
UDP     2EE12EE1 004F3669
LLC     040403
ANR     C600 80 D900140100000 0 8002000C01000000 D000000000000000 FF00
RTP     29B2A158000100D600004000A00000000000001E1
STATUS  050E000000000020000010600000000000000000
```

Translating from VTAM to IP
Where is my lost packet?

- How can I identify an HPR packet in an IP Packet Trace?
- At the sender
  - Note the unique BSN/DLF on a given pipe (TCID)
  - Remember the IPID in the IP header
- In the network/ at the receiver
  - Look for the IP identifier of the sent IP packet
    - Remember: the IP addresses may be NAT'ed
    - The IP datagram may have been fragmented
    - The trace may not show the full packet
  - Verify the BSN and TCID
- Network Support people will be happy to track down a lost IPID

RTP Retransmission
STATUS Segment reporting a GAP

- When a packet is lost, RTP will detect a gap and initiate selective retransmission of the lost NLP
- A STATUS segment with the GAP bit will report
  - The next expected BSN (this is the BSN of the lost NLP)
  - One ore more Byte_Span_Pair
    - The BSNs that were received successfully but out of order)



RTP Retransmission
STATUS Segment reporting a GAP

```
----------  BSN 00000074 DLF 00000012 sent out -------------
IP      45C0004A 43500000 3C11BBDE 0ABA56F3 0AD9122F
UDP     2EE12EE1 0036B528
NLP     C600 80 FF00
RTP     00000000010053033000000500000012 00000074     SOM EOM
FID5TH 5D000000000020000000000
----------  GAP report coming in asking for 00000074
IP      4500005B 43AB0000 72118632 0AD9122F 0ABA56F3
UDP     2EE12EE1 0047BDA2
LLC     040403
NLP     C600 D40000000000000 FF00
RTP     0F160F7C0001011B0004000C0000000000000EE
STATUS 070E800100000002 00000074 00000000000000000000087000000A3
----------  BSN 00000074 DLF 00000012 sent out again ------
IP      45C0005E 43620000 3C11BBB8 0ABA56F3 0AD9122F
UDP     2EE12EE1 004A9AF8
LLC     040403
NLP     C608 80 FF00
RTP     00000000010053033C04000A00000012 00000074     SOM EO
STATUS 05E000000300000000000EE000000000000000
FID5TH 5D000000000020000000000
```

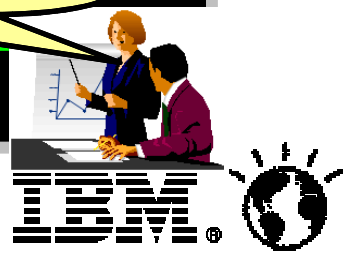
- A lost packet is retransmitted

D NET,EEDIAG,REXMIT=
Finding retransmitting connections

```
D NET,EEDIAG,REXMIT=1
IST097I  DISPLAY  ACCEPTED
IST350I  DISPLAY TYPE = EEDIAG
IST2065I  ENTERPRISE EXTENDER CONNECTION REXMIT INFORMATION
IST2067I  EEDIAG DISPLAY ISSUED ON 08/25/09 AT 16:08:59
IST924I  -------------------------------------------------
IST1680I  LOCAL  IP ADDRESS 10.232.72.11
IST1680I  REMOTE IP ADDRESS 129.35.231.237
IST2024I  CONNECTED TO SWITCHED PU PUSA01
IST924I  -------------------------------------------------
IST2033I  PORT PRIORITY = MEDIUM
IST2036I    NLPS SENT           =                  95  ( 000K )
IST2038I    NLPS RETRANSMITTED  =                   7  ( 000K )
IST2068I    NLP RETRANSMIT RATE =                  7%
IST924I  -------------------------------------------------
IST2035I  TOTALS FOR ALL PORT PRIORITIES
IST2036I    NLPS SENT           =                      )
IST2038I    NLPS RETRANSMITTED  =                      )
IST2068I    NLP RETRANSMIT RATE =
IST2069I  REXMIT COUNTERS LAST CLEARED ON 08/25/09 AT 13:2
IST314I  EN
```

We're losing way too many packets!

Wireshark Filter to find GAP reports:
sna.nlp.thdr.optional.0e.gap == 1



RTP - Data Flow Control
ARB algorithm

- Adaptive **Rate Based** algorithm
  - Operates on Send Rates (bits/s)
  - Initial Sendrate (5% of TG's capacity in APPN Topology)
  - Allowed Sendrate
    - Controlled by the receiving RTP
- The goal is to avoid congestion *before* packets get lost
  - BASE ARB was too polite to compete with greedy TCP/IP
  - Responsive Mode ARB (ARB2) is standard these days
    - I5 OS still uses BASE ARB per default!
  - Progressive Mode ARB available in V1R11 and CS V6R4
- ARB Segments (22) are used to adjust the allowed sendrate
  - Based on network delay changes

RTP Data Flow Control
ARB in action

```
-------------------------------------------------------------------
07:20:00.56  ARB RPNCB:1157C800 SR(kb/s):468 SZ:1401
07:20:00.64  ODPK EE_LOW 25BBF0C300010296 BSN:0025D375 continued me OUT
07:20:00.64  ODPK EE_LOW 25BBF0C300010296 BSN:0025D8BE continued me OUT
07:20:00.64  ODPK EE_LOW 25BBF0C300010296 BSN:0025DE07 continued me OUT
07:20:00.64  ODPK EE_LOW 25BBF0C300010296 BSN:0025E350 continued me OUT
07:20:00.65  ODPK EE_LOW 25BBF0C300010296 BSN:0025E899 continued me OUT
07:20:00.65  ODPK EE_LOW 25BBF0C300010296 BSN:0025EDE2 continued me OUT
07:20:00.66  ODPK EE_LOW 25BBF0C300010296 BSN:0025F32B continued me OUT
07:20:00.66  ODPK EE_LOW 25BBF0C300010296 BSN:0025F874 continued me OUT
07:20:00.67  ODPK EE_LOW 253ABFFE000100BF BSN:00000467 ARB REP -25  IN
07:20:00.68  ARB RPNCB:1157C800 SR(kb/s):387 SZ:1401
07:20:00.75  ODPK EE_LOW 25BBF0C300010296 BSN:0025FDBD ARB REQ cont OUT
07:20:00.76  ODPK EE_LOW 25BBF0C300010296 BSN:00260306 continued me OUT
07:20:00.76  ODPK EE_LOW 25BBF0C300010296 BSN:0026084F continued me OUT
07:20:00.76  ODPK EE_LOW 25BBF0C300010296                          OUT
07:20:00.77  ODPK EE_LOW 25BBF0C300010296                    me OUT
07:20:00.77  ODPK EE_LOW 25BBF0C300010296                    me OUT
07:20:00.77  ODPK EE_LOW 25BBF0C300010296 BSN:00261962       OUT
07:20:00.78  ODPK EE_LOW 253ABFFE000100BF BSN:00000467 ARB R  IN
```

- RTP sending at 468 kb/s
- Slowdown2 coming in
- Sendrate reduced to 387 kb/s

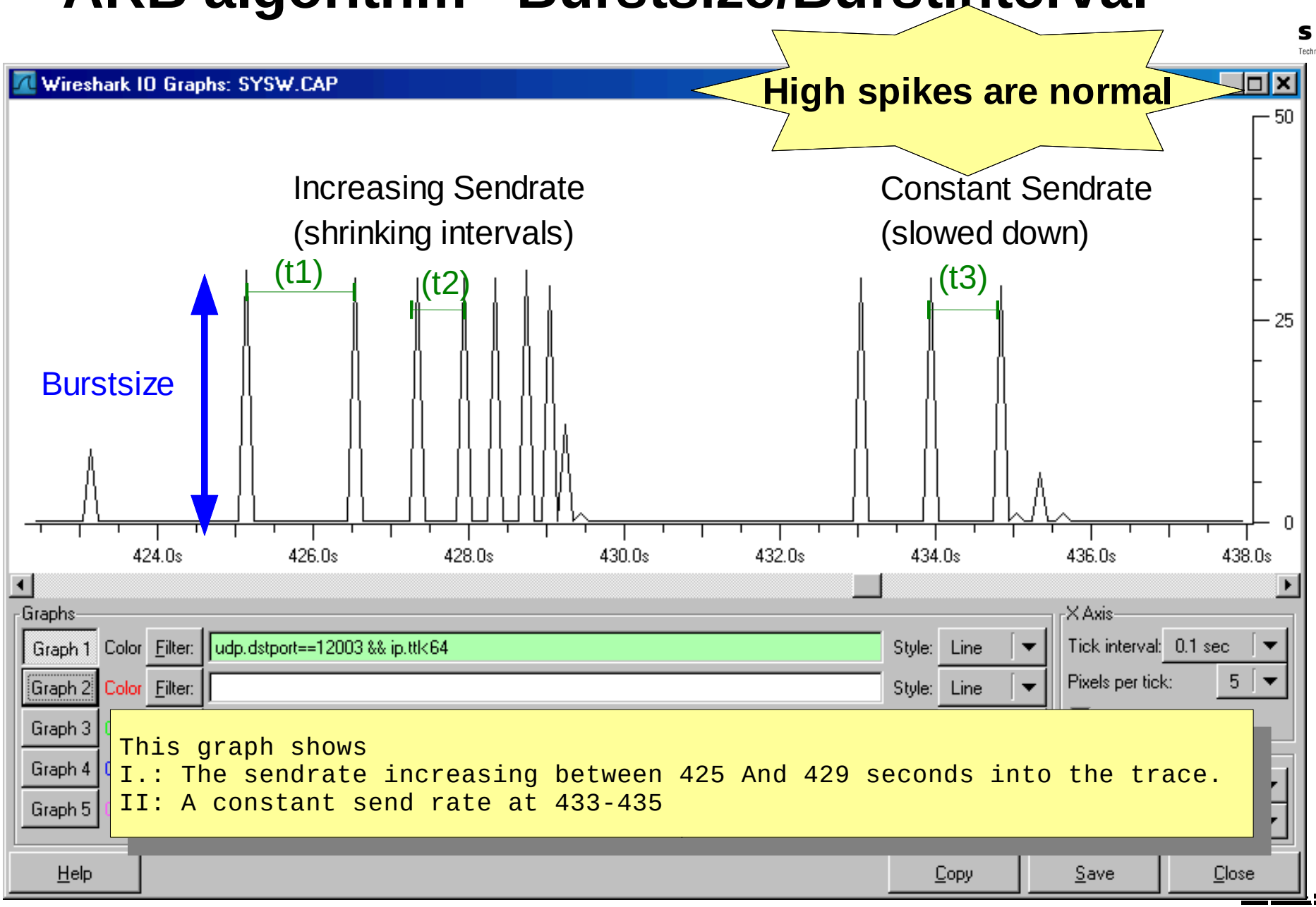
Here are the slowdown segments:
Network delays building up!

Wireshark Filter to find ARB SLOWDOWN:
sna.nlp.thdr.optional.22.raa >1

RTP - Data Flow Control
ARB algorithm – VERY time sensitive

- The receiving RTP is measuring the network delays
- Increasing network delays are treated as an indication of congestion building up in the network
  - Queues in bottleneck routers are building up
- To avoid a queue overflow in the network, ARB reduces the sendrate before packets are dropped allowing the network to recover sooner.
- Sometimes the delays are caused *within* the RTP nodes
  - CPU constrained systems (zSeries with few real CPUs)
  - z/OS and Linux under z/VM
  - Windows/Linux under VMWARE/Citrix environments
- Result is unnecessary slowdowns and poor performance

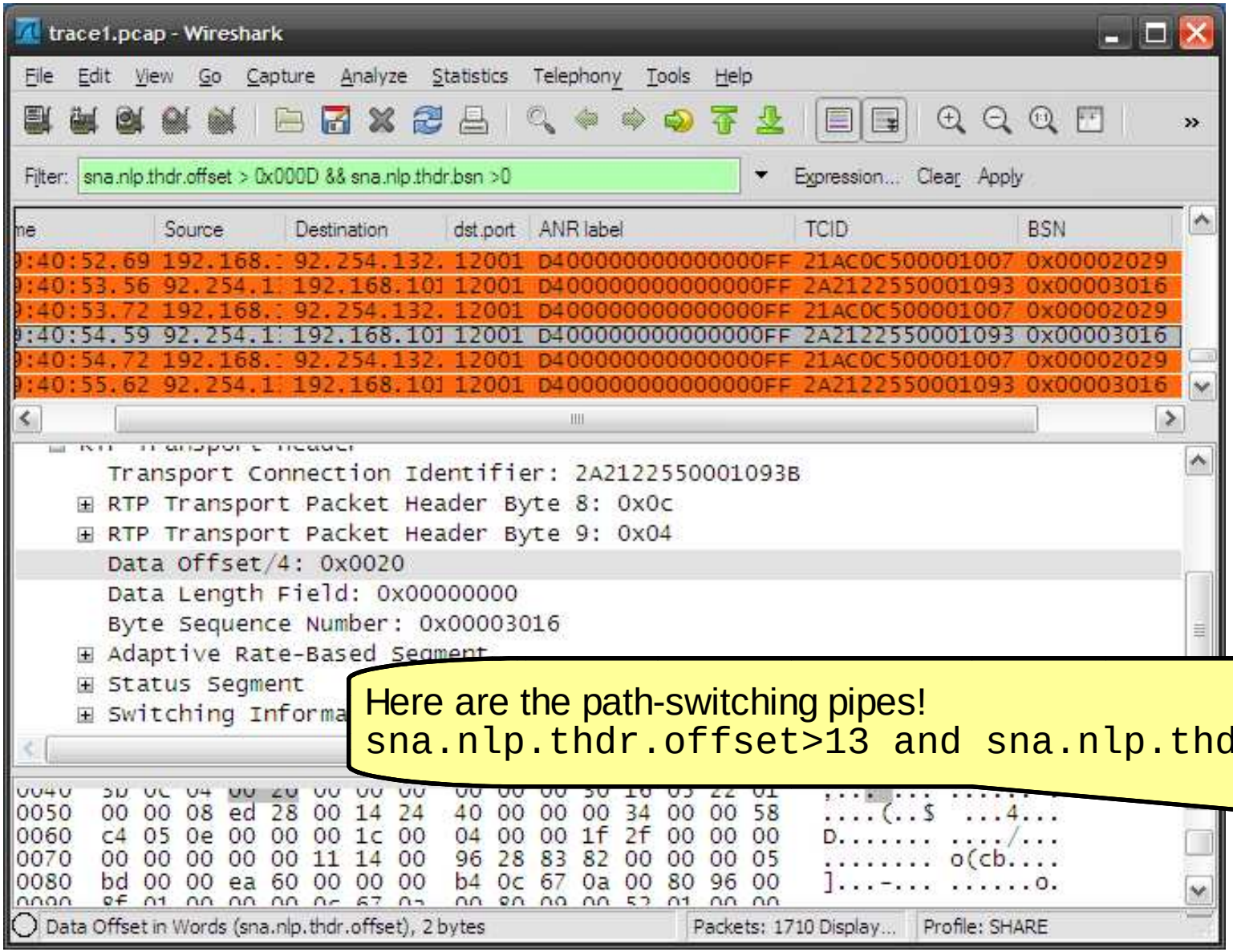
ARB algorithm - Burstsize/Burstinterval



RTP – PATHSWITCH
Identifying switching pipes in a trace

- Most HPR problems show up as path-switching pipes
- NLPs of an existing path-switching pipe will contain following information
  - The BSN will higher than zero
  - ARB SETUP is present to re-initialize the ARB algorithm
    - Length is 5 words
  - SWINFO is present to describe the characteristics of the new path
    - Length is variable but typically larger than 8 words
- In abbreviated traces the segments might not be traced

Wireshark filter - PATHSWITCH



Here are the path-switching pipes!
`sna.nlp.thdr.offset>13 and sna.nlp.thdr.bsn>0`

IPCS Formatter
Export SYSTCPDA to sniffer

- Convert your SYSTCPDA/SYSTCPOT traces in sniffer format

```
//SYSTSIN  DD *
 PROFILE MSGID
  IPCS NOPARM
  SETD PRINT NOTERM LENGTH(160000) NOCONFIRM FILE(INDMP)
  DROPD
 CTRACE COMP(SYSTCPDA)  -
        OPTIONS((SNIFFER(9000,TCPDUMP))) SUB((TCPIP))
 END
```

- Download the resulting file in binary
- Run a trace tool against the trace to filter the down to the problematic packets
- Talk to your network personnel using ***their language***!

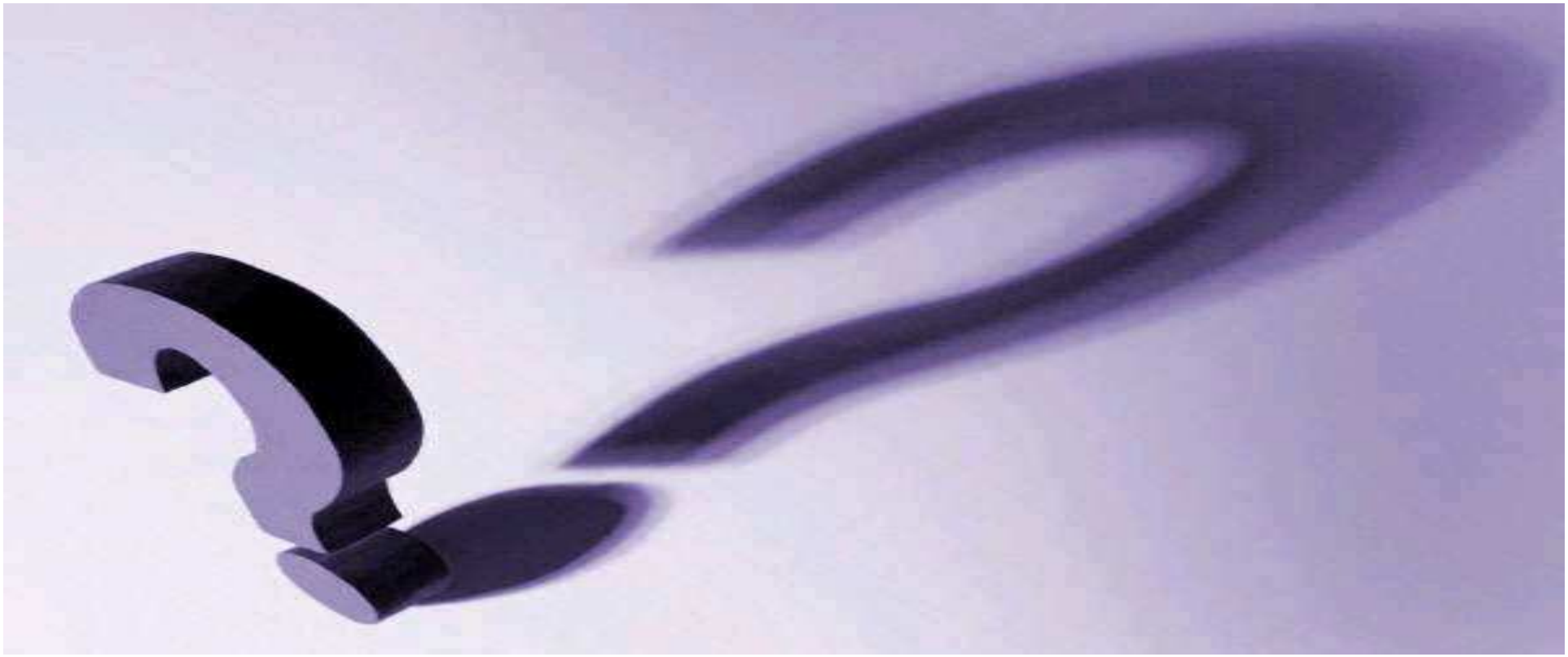
wireshark
Display Filters

- Stored as dfilters in
  - C:\Documents and Settings\Administrator\Application Data\Wireshark

```
"IP Fragmentation" ip.frag_offset>0 or ip.flags.mf == 1  or icmp.code == 4
"Low TTL - Routing Loop" ip.ttl < 8 and !ip.ttl==1 and !udp.srcport==33434
"ICMP Errors" icmp and !icmp.type==8 and !icmp.type==0
"PING" icmp.type==8 or icmp.type==0
"Time Delta" frame.time_delta > 0.2
"Warning" expert.severity>=1024
"EEonly " (udp.dstport > 11999 && udp.dstport<12005) or ICMP
"SNA BIND/UNBIND" (data.data[0] == 31 or data.data[0] == 32) and sna.rh.ru_category == 0x03
"UNBIND Cleanup" data.data[0] == 32 and data.data[1] >02 and sna.rh.ru_category == 0x03
"CPSVCMG Pipe " sna.nlp.nhdr.anr == d4:00:00:00:00:00:00:00:ff
"Pipe SETUP  " sna.nlp.thdr.bsn == 0 or sna.gds.type == 0x12ce
"PATHSWITCH" sna.nlp.thdr.offset > 0x000D && sna.nlp.thdr.bsn >0 or sna.gds.type == 0x12ce
"Pipe Termination" sna.nlp.thdr.optional.type == 0x12
"GAP " sna.nlp.thdr.optional.0e.gap == 1
"SLOWDOWN" sna.nlp.thdr.optional.22.raa >1
```

Questions?

- Come to Thursday's session on wireshark and EE problems



- Come to NYC Nov. 1st to learn all about EE -> WRB055US

WRB055cc – ITSO Workshop Agenda
EE Implementationand Problem Determination

1. Introduction to APPN
   - APPN Node Types
   - Topology and Routing Services
   - Directory Services
   - High Performance Routing
2. Implementation
   - VTAM definitions
   - Personal Communications
   - iSeries
3. Diagnosis
   - IP Header
   - HPR Architecture
   - Firewalls and IPSec
4. A Journey through the layers
   - VTAM messages
   - IP traces
5. Wireshark Trace Tool
   - EE Profile
   - Filters / Colors
   - Sample Traces