Identifying and Solving Network Performance Problems on zEnterprise

SHARE Boston
August 4, 2010

Mac Holloway (mhollowa@us.ibm.com)
Dean Butler (butlerde@us.ibm.com)
Agenda

- Intro
- Problem spaces
- Configuration issues
- Operational issues
- Modern Applications
- Scenarios
- Summary
Intro

- **Host network management definition**
  - Enabling effective use of z/OS Communications Server, VTAM and OSA communications as part of End-To-End solutions in IT shops.

- **Presenters**
  - **Dean Butler** – 15+ years of network and network management development both inside and outside IBM. Currently a System z software architect in Tivoli.
  - **Mac Holloway** – 20 years of IBM networking in NSD, NHD and Tivoli including work on NWAYs, zNV, Mainframe Networks and z/VM and Linux
Mainframe Networking Performance Problem spaces
Problem spaces

- What we hear
  - “A critical application is “broken”. We all get on a bridge call/line. Everyone says “Mine stuff is okay. It must be the network. I need to be able to say it is not the network or at least not my part of the network.”

- Configuration issues
  - This is the most common type of problem we see. This is the most common type of problem zCS sees
  - Examples of VTAM, zCS, SNMP, IPSec configuration issues
  - An approach

- Operational issues
  - Problem sources
    - Protocol – malformed packets, unusable ports, …
    - Resources – packets dropped, buffer overflows
    - Indirect – packet reassembly, response time, rate changes
  - Location indicators
    - Outboard – from the MAC out into the network and beyond
    - Stack – from the MAC to the buffer interfaces above TCP and UDP
    - App – above the stack including FTP, TN3270
Mainframe Networking Performance Configuration Issues
Configuration Issues - examples

- **OSA**
  - OSA-Express Direct SNMP subagent (IOBSNMP) or OSA/SF application (IOAOSASF) and the OSA/SF sockets application (IOASNMP) – running?
  - RACF for OSA/SF SNMP sub-agent (IOASNMP) – security messages

- **IPSec**
  - IKE Daemon started?
  - PAGENT Daemon started?
  - IPSec NMI access authorized?

- **TN3270 & FTP**
  - z/OS Communications Server real-time SMF data network NMI enabled?
  - Monitoring app authorized?
  - Sliding window or bucket count data – configured in Telnet profile?

- **SNA NMI**
  - zCS SNA NMI not enabled,?
  - OMVS segment created for VTAM?
  - Monitoring app authorized
Configuration Issues - examples

- **VTAM**
  - Is it running? Have you done a vary to activate it?
  - Is SNA data collection configured?
  - Is your monitoring app in the VTAMLST?
  - Is the PMI exit available to VTAM? Add DD card, quiesce VTAM, restart VTAM

- **SNMP**
  - Is it configured? SNMPD.CONF
  - Is it running? OSNMP
  - Are you accessing the right address? loopback address (127.0.0.1)
  - Do you have the right community name? Check SNMPD.CONF
  - Are you using the right port? 161
  - SNMP requests are timing out? Your application
Configuration Issues – an approach

Status on:
• SNA NMI
• VTAM
• PMI
• PAGENTD
• IKED
• OSNMP
• …

Situations/actions based on status
Mainframe Networking Performance Operational Issues
Operational Issues Indicators

Types of indicators
- Protocol errors – malformed packet, unmonitored port
- Resource constraints – dropped packets, buffer overflows
- Indirect – packet reassembly, response time, rates

Direction of indicator
- Outboard – from the MAC out into the network and beyond
- Stack – from the MAC to the buffer interfaces above TCP and UDP
- App – above the stack including FTP, TN3270
The layers – a quick look
Operational issues – more detail

Each of these parts has data that can indicate
Protocol indicators

TCP/IP
- Input Discards
- Output Discards
- UDP Discard
- UDP Input Errors
- UDP No Port

OSA
- Fragments, jabber, length error, CRC, alignment
- Unknown IP Frames

Interfaces
- Inbound Packet – discarded, in error
- Outbound Packet – discarded, in error
- Utilization
- Transmission Rates
- Unknown IP Frames
Resource constraint indicators

TCP/IP
- CPU Percentage
- CSA – Allocated, In Use
- Authorized Private Storage – Allocated, In Use
- ECSA storage – max, allocated, in use, pools
- Datagrams Discarded
- Backlog Connections Rejected

UDP
- Datagrams Discarded

HPR
- Throughput rate – allowed
- Unacknowledged Buffers – high water mark

Interfaces
- Receive/Transmit Bandwidth utilization

OSA
- PCI Utilization
- Processor Utilization
- Missed Packets

VTAM
- CPU Percentage
- CSA – Allocated, Allowed, In Use
Indirect indicators

TCP
- Percent Segments Retransmitted
- Response Time
- Segments Retransmitted
- Fragmentation
- Reassembly
- TCP Retransmit
- Out of Order Segments
- Segments Retransmitted
- Remote Window Size Frequency
- Response Time Variance
- TCP Keep-Alive Drops

HPR
- Out of Sequence Buffers
- Packet Retransmission Rate
- Path Switches
- Response Time Variance
- Smoothed Round Trip Time

TN3270
- Average IP Response Time and variance
- Average SNA Response Time and variance
Modern Applications … Integrated Management
Workflow for Resolving Composite Application Problems

## Sense
Detect that a threshold has been breached and that a problem occurred, or is about to happen

## Isolate
Pinpoint the problem to a specific part of the environment and hand-off to the appropriate specialist

## Diagnose
Drill down into the details and get to the root cause of the problem

## Repair
Fix the faulty component, validate the fix and roll back into production
zEnterprise with zBX (z Blade Extension)

- Workload 1
- Workload 2
- Workload 3

Key:
- Workload 1
- Workload 2
- Workload 3

Key:
- Black = Intra-Ensemble Data Network (IEDN)
- Pink = Customer managed data network
- Purple = Customer managed management network
- Yellow = Intra-Node Management Network (INMN)
- Customer managed management network
- Intra-Node Management Network (INMN)
- Intra-Ensemble Data Network (IEDN)
- Customer managed data network
Multiple Virtual Networks – Isolation

1. Define Multiple Virtual Networks

   - “Production Network” VLAN ID 300
   - “Development Network” VLAN ID 500

   …each having unique VLAN IDs and IP subnets

2. Then add virtual servers to each virtual network as needed...

   …which isolates “Production Servers” from “Development Servers”
Which zEnterprise network components are of interest from a monitoring perspective?

- OSAs – Support is already available to monitor OSAs today.
  - New OSA types (OSX and OSM)

- A private and physically isolated management network (the intranode management network - INMN), connects all zEnterprise System resources (CPCs, BladeCenters, etc.) for management purposes. This INMN is pre-wired, internally switched, configured, and managed with full redundancy for high availability.
  - Throughput, dropped packets, fragmentation, etc are NOT of interest

- A private and secure OSA-Express Ethernet intraensemble data network (IEDN) that connects all elements of a zEnterprise System ensemble. The IEDN is access-controlled using integrated virtual local area network (VLAN) provisioning. IEDN management provides enforcement of strict access control across heterogeneous environments, further augmenting security and simplicity.
  - Throughput, dropped packets, fragmentation, etc are of interest

- VLANs – Defined in zHMC. Includes one or more virtual servers per VLAN ID. Key performance metrics (throughput, dropped packets, fragmentation, etc) would be of interest per interface.
Scenarios
Scenario A: Its not the Network!

The setting:
A company relies on batch FTP to copy files between a mainframe at headquarters and each of its retail stores every night (local store time). Sales and inventory data is uploaded and product and pricing changes are downloaded to the stores. One morning, a systems administrator notices that some of the FTP jobs have not finished. He reports the problem to the IT help desk. The problem is routed to the mainframe networks systems programmer.
User reports batch FTP failures

- Start with checking current activity: FTP transfers & FTP sessions

**FTP Session Summary Table**

<table>
<thead>
<tr>
<th>Collection Time</th>
<th>Application Name</th>
<th>FTP Type</th>
<th>Remote IP Address</th>
<th>Remote Port</th>
<th>Local IP Address</th>
<th>Local Port</th>
<th>User ID on Server</th>
<th>Client User ID</th>
<th>Session Start</th>
<th>Session End</th>
<th>Session Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>08/1/08 23:52:22</td>
<td>FTPD1</td>
<td>Server</td>
<td>9.65.126.164</td>
<td>3000</td>
<td>9.42.45.179</td>
<td>21</td>
<td>USER2</td>
<td></td>
<td>08/1/08 23:52:18</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

**FTP Transfer Summary Table**

<table>
<thead>
<tr>
<th>Collection Time</th>
<th>Remote IP Address</th>
<th>Remote Port</th>
<th>Local IP Address</th>
<th>Local Port</th>
<th>User ID on Server</th>
<th>Client User ID</th>
<th>Role</th>
<th>Transmission Start</th>
<th>Transmission End</th>
<th>Transmission Duration</th>
<th>Bytes Transmitted (in GB)</th>
<th>Bytes Transmitted</th>
</tr>
</thead>
<tbody>
<tr>
<td>08/1/08 23:53:39</td>
<td>9.65.126.164</td>
<td>3010</td>
<td>9.42.45.179</td>
<td>20</td>
<td>USER2</td>
<td>Server</td>
<td>08/1/08 23:52:49</td>
<td>08/1/08 23:53:39</td>
<td>11420</td>
<td>0</td>
<td>1440054</td>
<td></td>
</tr>
<tr>
<td>08/1/08 23:53:39</td>
<td>9.65.126.164</td>
<td>3010</td>
<td>9.42.45.179</td>
<td>20</td>
<td>USER2</td>
<td>Server</td>
<td>08/1/08 23:53:39</td>
<td>08/1/08 23:53:39</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
Check Applications and Connections

**Applications:**
- Accepting connections?
  - Rate, Backlog, Rejections
- Last activity time
- Response Times
- Retransmissions
- Transmit / Receive Rates
- Out of order segments
- CICS, IMS, WAS, z/OS

**Connections:**
- Start time/duration
- Response Time
- Response Time Variance
- Retransmissions
- Transmit / Receive Rates
- Out of order segments
Check OSA and Interfaces

**OSA**
- Online Status
- Configuration
- Microcode Level
- Utilization
- Transmission Rates
- Unknown IP Frames
- By LPARS
- By Ports

**Interfaces**
- Packet Errors
- Bandwidth Utilization
- MTU Size
Check TCP/IP Stack

- **Retransmits** - Network congestion
- **Out of Order** - Routing or network congestion
- **Fragmentation** - MTU size
- **Discards** - Resource constraints
- **Timeouts** - Connectivity
- **UDP input Errors** - Attack
- **UDP Discards** - Wrong Sockets
- **High storage utilization** - Could indicate Network congestion
  - This can result in requests backing up in storage
Network is OK, then where is the problem?

Check job logs, SYSLOG, and syslogd for clues. Dataset access issue? FTP session login failure?
Scenario B: Slow response time in web service

The setting:
A company recently deployed a set of web services that replaced a very high profile application. The operations team monitors the performance closely. When performance degrades, its time to investigate…
1. An alert identifies a response time problem. Annette, an operator, determines that slow response times are being recorded for the new web services.

2. Annette checks the number of requests and the message size and determines this is a normal volume of traffic. Annette passes the issue to Johann, a SME.
Slow response time in web service ...

3. Johann begins by looking closer at the web services. Identifies flows and response time for each step.

4. Problem appears to be with the network between the CICS and DB2 servers. These two LPARs are connected by a data center network.
5. Johann views metrics for connections between CICS and DB2 on the two LPARs.

6. Johann notices there have been retransmits and out-of-order segments between CICS and DB2 servers. But what is the root cause?
7. Johann checks the OSA cards and discovers the OSA on the DB2 server has high PCI and processor utilization.

8. Further checks reveal contention on OSA with other LPARs in the CEC is causing the performance issues.

Each OSA is dedicated to an LPAR, but also serves as backup OSA for a 2nd LPAR. Switch other LPAR to its primary OSA.
Scenario C: DB2 is working, it must be the network

The setting:
A multi-tier application framework is being used by a team of programmers to develop a Java application. The application is stored as large binary objects (BLOBs) in a DB2 on z/OS database. Each programmer retrieves, changes, and then saves a BLOB. Long delays that occur sporadically during the save are frustrating the application team.
DB2 is working, it must be the network …
DB2 is working, it must be the network ...

1. Facing revolt from his team, the team leader asks the DB2 systems programmer to check for performance problems.

2. The DB2 systems programmer checks thread CPU time, lock contention, and query plan, among other things. He determines that DB2 is not the cause of the slowdown.
3. Expecting that the problem may be due to an underlying network problem, the team leader turns to Johann for help.

4. Johann views the DB2 application and associated connections. Large amounts of data is being transferred over the DB2 connections with no retransmits or out of order segments.
DB2 is working, it must be the network …

5. Interesting… Response time and response time variance are higher than expected (0.5+ sec, 0.5+). Also, much more data is being sent from DB2 than received from the remote system.

Why is ACK from remote system taking so long?

6. Working with distributed network and other SMEs to identify and resolve.
Scenario D: Erratic response times for TN3270 application

The setting:
Users are becoming frustrated at response times with an SNA application. All access is through TN3270. The response times are on average very fast, but vary widely over the course of a day.
Erratic TN3270 response times …

1. A user opens a trouble ticket. Annette contacts the user who identifies a TN3270 session (TCP00072) that exhibits the erratic behavior.

2. The average response time and average SNA response time are fairly high. In contrast, average IP response time is good, so does not appear to be a network problem.
Erratic TN3270 response times …

3. Looking further, the bucket counts show that there have been a number of transactions with poor response time and a number with good response times but not much in between.

4. Annette passes the problem to the SNA application support team, which identifies and resolves the issue.

High average SNA response time? Investigate:
- High application workload spike
- z/OS system resource constraints.
Scenario E: Application Performance Problem

The setting:
A company is starting to protect more and more of its IP traffic using encryption. The deployment has gone well and the IT operations staff is trained and ready. A user calls the help desk because a file transfer is taking a long time.
z/OS IP Security Support

Common Problems and Symptoms:

- Filter added in wrong order
  - Loss of connectivity to applications
- Security policies at endpoints are incompatible
  - Loss of connectivity to applications
  - Tunnel activation failures
- Loss of network connectivity between security endpoints
  - Loss of connectivity to applications
  - Tunnel activation failures
- Cryptographic services unavailable, misconfigured, or insufficient
  - Application performance is slow
  - Loss of connectivity to applications
  - Tunnel activation failures
Application performance problem ...

- End user calls help desk because the transfer of large files is taking a long time.
- The operator looks at the Applications workspace and sees that the user’s FTP client ID is experiencing retransmissions.
Application performance problem ...

- The systems programmer finds the IP filters for src/dst IP address, then finds the associated dynamic tunnels.
- There are a high number of expired tunnels.
- The tunnel associated with the user's transfer has data rates of 0 and there are many tunnels with the same tunnel ID indicating it has been refreshed many times.
Application performance problem …

- The systems programmer examines the tunnel refresh and expiration information.
- The tunnel is being refreshed every 2 to 10 seconds.
- The systems programmer corrects the refresh time for the tunnel, which fixes the performance problem.
Questions?

Identifying and Solving Network Performance Problems on zEnterprise

Mac Holloway (mhollowa@us.ibm.com)
Dean Butler (butlerde@us.ibm.com)
Backup