Network Strategies and Design Workshop
April 4, 2012

Agenda

- Status of 10G, 40G, 100G Technologies
- Troubleshooting the LAN

Mark Williamson
mawillia@cisco.com
Status of 10G, 40G, and 100G Technology

Mark Williamson
mawillia@cisco.com
Physical Layer Standards vs Form Factor

- IEEE defines physical network interfaces (Physical Medium Dependent specs)
  - 10GBASE-LR, 40GBASE-SR4, 100GBASE-SR10 etc.
- MSA (multi-source agreements) define transceiver form factors
  - X2, SFP+, CFP, QSFP, etc.
  - Usually defined by a group of transceiver vendors based on technology availability and roadmap
- PMDs of the same speed/reach/protocol must interoperate, irrespective of the form factor
  - X2-10GB-SR interoperates with SFP-10G-SR, both being defined by the 10GBASE-SR physical interface
- Multiple 40G and 100G form factors are being released simultaneously at Cisco and other system vendors
40G/100G Industry Standards

- **IEEE 802.3ba**: 40Gb/s and 100Gb/s Ethernet Task Force
  40G and 100G Ethernet (Parallel transmission)
  Physical interfaces for Backplane, Copper, Fiber PMDs

- **IEEE 802.3bg**: 40Gb/s SMF Ethernet Task Force
  40G **Serial** PMD optimized for carrier applications

- **ITU Study Group 15**: Optical and Transport Networks
  OTU4 frame format
  Single mapping for 40GE/100GE into OTU3/OTU4
  OTL protocol enabling OTU3/4 over multi-lane (low cost) optics

- **OIF**: 100G Long-distance DWDM Transmission
  Industry consolidation around a single 100G DWDM solution
For 40G/100G, Ethernet Standard defines ‘parallel’ lanes to handle the flow of data.
IEEE 802.3: Parallel Data Streams

• Electrical interfaces in “host” currently cannot handle single 40Gb/s and 100Gb/s data streams
• IEEE has defined ‘parallel lanes’ to handle flow of data
• 40GE/100GE interfaces all based on multiple lanes (coax cables, fibers, wavelengths, etc).

Media Access Control (MAC), Physical Coding Sublayer (PCS)
CFP and QSFP+ Transceivers – 40G

Ribbon fiber cables terminated with MPO 12-fiber connector

One pair – 4 wavelengths Singlemode

4 pair multimode

4 separate 10G

40G CFP - Catalyst 6500 (Nexus 7K future)

4X CFP - Catalyst 6500

QSFP+ for ASR9000, Nexus 3000 and Nexus 7000
CFP and CXP Transceivers – 100G

2 pair Singlemode

Dual SC Connectors
2 Lanes per connection
100GBASE-LR4
CRS-3, ASR 9000 and Nexus 7000

CXP – later this year
Summary

• Standards for 40G and 100G have been Ratified

• Doesn’t matter what the Form Factor is – The IEEE Standard defines interoperability:
  
  40GBASE-LR4, 40GBASE-SR4, 100GBASE-LR4
  
  CFP, QSFP+, CXP

• Cisco offers 5 Platforms that support 40 or 100 Gig today

• Limitations on the electrical interface result in the need for multiple traffic lanes

• In many cases, multiple traffic lanes means multiple fiber optic pairs (2, 4, 10 pairs). Requires OM3 / OM4 grade multimode fiber

• Development continues to achieve:
  
  Smaller form factors
  
  Moving components from transceiver to Line Card interface
  
  Fewer fiber pairs
  
  Parallel to serial, more WDM
Troubleshooting the LAN

Mark Williamson
mawillia@cisco.com
Agenda

- Troubleshooting Tools
- Basic Troubleshooting Methodology
- Troubleshooting Scenarios
  - Layer 1 & 2 connectivity problems
  - High CPU utilization
- Summary
Troubleshooting Tools

- **Command Line Interface (CLI)**
  - Manual line-by-line configuration
  - Most informative, detailed expert interface

- **Built-in Web Interface**
  - SmartPorts configuration
  - Limited troubleshooting compared to CLI

- **Cisco Network Assistant (CNA)**
  - Free – supports networks up to ~ 40 networking devices
  - SmartPorts configuration
  - Automated Topology Services

- **Prime LMS and NCS**
  - Licensed by number of networking devices
  - Automated tools (Troubleshooting tips, dialogs)

- **Syslog Server**
  - Event correlation and logging
  - Kiwi (Solarwinds), Splunk, Prime LMS
Caution!

- **debug** and **show platform** commands must be used with caution.

- Excessive debug output can disable the Switch. “Debug all” will take even and 8-port Compact Switch down immediately. Do not ignore the CLI warnings!

- **show platform** commands are intended for in-depth troubleshooting by Cisco engineers.

- Use **debug** and **show platform** commands only when advised by a knowledgeable Network Support Engineer or TAC.
Shortcuts And Keystrokes

When you type enough of the characters that uniquely identify the command in a given context you can use 'tab' key and the CLI will complete the command on the screen for you.

Keystrokes for editing:

CTRL-z - takes the cursor back to the 'enabled exec' mode from any other context (also the ESC key)
CTRL-a - takes the cursor to the beginning of the line
CTRL-e - takes the cursor to the end of the line (Left and Right arrow keys will move the cursor one character @ a time)
CTRL-k - erases everything to the right of the cursor
CTRL-u - erases everything to the left of the cursor

Upper Arrow or CTRL-P - displays previous command from the history buffer
Lower Arrow or CTRL-N - displays next command from the history buffer

System keeps the record of up to 10 or 20 (default) commands you typed in. You can increase/decrease the history buffer up to 256 commands.

If you want to check the history buffer size, type in:

Router# show terminal | include history
Changing the buffer size can be accomplished with the following command in the 'enabled mode'

Router# terminal history size 256
The above command will increase the size of the history buffer to 256 commands.

If you want to see the buffer (which commands were typed) use this command:

Router# show history

Question Mark provides context sensitive help – Magic!

Other Tips: [http://packetlife.net/wiki/ios-tips](http://packetlife.net/wiki/ios-tips)
Basic Troubleshooting Method

1. Define Problem
2. Gather Facts
3. Consider Possibilities
4. Create Action Plan
5. Execute Action Plan
6. Observe Results

• Need to have a good understanding about how the system looks like when it is healthy

• Documentation:
  • Baseline: Response Time
    Link Utilization
  • Archive: Configuration files
    Trouble tickets
  • Topology Diagrams are **very** helpful
Basic Troubleshooting Method

1. Define Problem
2. Gather Facts
3. Consider Possibilities
4. Create Action Plan
5. Execute Action Plan
6. Observe Results

- Symptoms?
- System Messages?
- User Input?
- When?
- Frequency?
- Impact?
- Scope?

- Any recent network changes?
- Any previous trouble tickets with similar issue?
Basic Troubleshooting Method

1. Define Problem
2. Gather Facts
3. Consider Possibilities
4. Create Action Plan
5. Execute Action Plan
6. Observe Results

Brainstorm potential root causes

<table>
<thead>
<tr>
<th>Category</th>
<th>Possible Cause</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hardware Issue</td>
<td>Bad Hardware?</td>
</tr>
<tr>
<td>Transport Issue</td>
<td>Link down?</td>
</tr>
<tr>
<td>Saturated?</td>
<td></td>
</tr>
<tr>
<td>Software Issue</td>
<td>Application problem?</td>
</tr>
<tr>
<td>Bug, Virus?</td>
<td></td>
</tr>
<tr>
<td>Config/Design</td>
<td>Mis-configuration?</td>
</tr>
<tr>
<td>Reaching Capacity?</td>
<td></td>
</tr>
<tr>
<td>Traffic</td>
<td>DOS Attack?</td>
</tr>
<tr>
<td>Traffic Pattern Change?</td>
<td></td>
</tr>
<tr>
<td>Utilization on server?</td>
<td></td>
</tr>
</tbody>
</table>
Basic Troubleshooting Method

1. Define Problem
2. Gather Facts
3. Consider Possibilities
4. Create Action Plan
5. Execute Action Plan
6. Observe Results

Documentation

What needs to be done to isolate each potential root cause?
Make a change, measure results, rollback change if problem persists

Problem solved? If not, continue action plan
Users Complain about Slow App Performance

All users complaining
App moved to new server installed on Monday – slow since then

Let’s have a look at Interface Fa0/2………..
Users Complain about Slow App Performance

Switchport is Auto
Server NIC is set for Auto!

After 512 bits have already been sent. Indication the other end may be in Full Duplex mode.
Users Complain about Slow App Performance

- Try to upgrade the NIC driver using the server vendor's web site.

- Hard code speed and duplex. You have to do this on both ends of the connections. This disables the AUTO NEGOTIATION feature (cannot do this on just one end of the connection).

- Could be a bad cable, but we would expect see some other layer 1 errors (CRC errors, lost carrier).

- Replace the NIC on the server

SW2#configure terminal
SW2(config)#interface fastethernet0/2
SW2(config-if)#speed 100
SW2(config-if)#duplex full
Some Users Cannot Access Server (intermittent)

Only some users affected
Sometimes they can access it, most of the time they cannot
Affects users on the same subnet as server
Problem started this morning
You cannot access the server from your desktop
OSI Reference Model

- **Application**
  - Interacts with user applications

- **Presentation**
  - Ensures that information sent from one system will be readable by another system

- **Session**
  - Manages communication sessions between presentation layer entities

- **Transport**
  - Implements data transport services that are transparent to upper layers

- **Network**
  - Allows multiple data links to be combined into an internetwork. Uses logical addresses

- **Data Link**
  - Provides transport of data across a physical network link. Uses physical (MAC) addresses

- **Physical**
  - Specifications for the physical link between systems

Open Systems Interconnection model (OSI model)
International Organization for Standardization
Some Users Cannot Access Server (intermittent)

- PING 192.168.1.1 (192.168.1.1) 56(84) bytes of data.
  64 bytes from 192.168.1.1: icmp_seq=1 ttl=64 time=0.372 ms
  64 bytes from 192.168.1.1: icmp_seq=2 ttl=64 time=0.349 ms
  64 bytes from 192.168.1.1: icmp_seq=3 ttl=64 time=0.371 ms
  64 bytes from 192.168.1.1: icmp_seq=4 ttl=64 time=0.374 ms

Layer 3 appears to be working............

- What if another device has that IP address? (not the server)
- Disconnecting the server would be disruptive
- Let’s just confirm with arp –a

  C:\>arp –a
  Interface: 192.168.20.106 --- 0xf
  Internet Address   Physical Address   Type
  192.168.1.1        00-18-39-b4-b8-ee  dynamic

Compare to “ipconfig /all” on Server or “sh mac address-table” on Switch
Some Users Cannot Access Server (intermittent)

- MAC address is different!
- The device that answers first is the one that gets into the ARP table
- Where is the “rogue device”?
- Show MAC address-table on switches

```
2960G#sh mac address-table address f0de.f14c.3d8d

Mac Address Table

-------------------------------------------
Vlan  Mac Address       Type      Ports
----  -----------       ------     ----
    f0de.f14c.3d8d DYNAMIC   Gi0/7

Total Mac Addresses for this criterion: 1
```

2960G#
Troubleshooting High CPU Utilization on Switches

- A non-stackable switch, the CPU is busy at least 5 percent of the time. If the switch is stacked, the CPU is busy at a minimum of 7 or 8 percent utilization.

- In a switch stack, CPU utilization is measured only on the master switch. The number of members in the stack affects the overall CPU utilization.

- 3750 w/IP Base = 475 active system processes
- 2960G w/LAN Base = 257 active processes

- Utilization of 50% or less is normal - CPU % does not affect switching
- Spikes will occur (write mem, show tech, topology changes, etc.)

<table>
<thead>
<tr>
<th>Example CPU Processes:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>system health checks</td>
<td>LED indicators</td>
</tr>
<tr>
<td>routing information updates</td>
<td>processing broadcasts</td>
</tr>
<tr>
<td>spanning-tree processing</td>
<td>processing BPDUs</td>
</tr>
<tr>
<td>port status</td>
<td>address learning</td>
</tr>
</tbody>
</table>
# Troubleshooting High CPU Utilization on Switches

## 2960G# show processes cpu sorted 5min

CPU utilization for five seconds: 5%/0%; one minute: 5%; five minutes: 5%

<table>
<thead>
<tr>
<th>PID</th>
<th>Runtime(ms)</th>
<th>Invoked</th>
<th>uSecs</th>
<th>5Sec</th>
<th>1Min</th>
<th>5Min</th>
<th>TTY</th>
<th>Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>255998</td>
<td>12870</td>
<td>19891</td>
<td>0.31</td>
<td>0.55</td>
<td>0.67</td>
<td></td>
<td>Check heaps</td>
</tr>
<tr>
<td>170</td>
<td>42</td>
<td>128</td>
<td>328</td>
<td>0.31</td>
<td>0.05</td>
<td>0.01</td>
<td></td>
<td>Virtual Exec</td>
</tr>
<tr>
<td>174</td>
<td>803</td>
<td>4788</td>
<td>167</td>
<td>0.00</td>
<td>0.02</td>
<td>0.00</td>
<td></td>
<td>IP Input</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>SpanTree Helper</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>6740</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>Load Meter</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>2</td>
<td>4000</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>Pool Manager</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>Transport Port A</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>Timers</td>
</tr>
<tr>
<td>9</td>
<td>115</td>
<td>1117</td>
<td>102</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>ARP Input</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>AAA_SERVER_DEADT</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>Crash writer</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>Policy Manager</td>
</tr>
<tr>
<td>13</td>
<td>17</td>
<td>4</td>
<td>4250</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>Entity MIB API</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>IFS Agent Manage</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>564</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>IPC Dynamic Cach</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>AAA high-capacity</td>
</tr>
<tr>
<td>17</td>
<td>0</td>
<td>33589</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>IPC Periodic Tim</td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>IPC Managed Time</td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>33589</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>IPC Deferred Por</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>IPC Seat Manager</td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>IPC Session Serv</td>
</tr>
<tr>
<td>16</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>IPC Zone Manager</td>
</tr>
<tr>
<td>23</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>XML Proxy Client</td>
</tr>
<tr>
<td>24</td>
<td>0</td>
<td>33589</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>Dynamic ARP Insp</td>
</tr>
<tr>
<td>25</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>ARP Snoop</td>
</tr>
<tr>
<td>26</td>
<td>8</td>
<td>33524</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>GraphIt</td>
</tr>
<tr>
<td>27</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>Critical Bkgnd</td>
</tr>
<tr>
<td>28</td>
<td>177</td>
<td>9145</td>
<td>19</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>Net Background</td>
</tr>
<tr>
<td>29</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>IDB Work</td>
</tr>
<tr>
<td>30</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>Logger</td>
</tr>
<tr>
<td>31</td>
<td>25</td>
<td>33520</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>TTY Background</td>
</tr>
<tr>
<td>1</td>
<td>9</td>
<td>7</td>
<td>1285</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>Chunk Manager</td>
</tr>
<tr>
<td>33</td>
<td>9554</td>
<td>565</td>
<td>16909</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td></td>
<td>Per-minute Jobs</td>
</tr>
<tr>
<td>34</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td>IF-MGR control p</td>
</tr>
</tbody>
</table>
Troubleshooting High CPU Utilization on Switches

- Definitions and Explanations are in the document “The show processes Command”

- [1] Check Heaps: process is run every 60 seconds to verify the consistency of the memory heaps. Pool of unused memory area called the heap. Basically, it collects and verifies all the memory structures are valid and intact and that there is no memory corruption. Depending on how much memory you have, check heaps may have to work more to verify the consistency of the memory.

- [2] Virtual Exec: handles vty lines (Telnet, SSH sessions to the router). Will spike during “write mem”, “show tech-support”, “show run”, etc.

- [3] IP Input: IP packets that must be handled directly by the CPU (Telnet sessions, ICMP, etc.). If High CPU is due to IP Input, don’t rely on Telnet and SSH sessions to the device (could be unresponsive) – use the Console Port!

- Let’s Consider some common causes............
Common Causes of High CPU Utilization

- High **SNMP Engine** Process (over 40%):
  - SNMP polling rate to high
  - Multiple SNMP Servers polling the device
  - Queries to the Flash File System of the Switch

- High **Hulc LED Process** (over 8%)
  - Link(s) flapping
  - Check Syslog for event messages
  - If lots of unused ports, shut them down – will reduce this process

- High **HLFM address learning** (over 10%)
  - Local Forwarding Manager manages MAC to IP mappings and address learning
  - Unstable Layer 2 environment
  - MAC flooding attack

- High **IP Input** Process (accompanied by high interrupt – over 10%)
  - Excessive IP packets getting to the CPU
  - See next slide
Common Causes of High CPU Utilization

- High Interrupt (over 10%):

```
Switch# show processes cpu sorted 5sec
CPU utilization for five seconds: 64%/19%; one minute: 65%; five minutes: 70%
PID Runtime (ms) Invoked uSecs 5Sec 1Min 5Min TTY Process
186 19472027 64796535 300 35.14% 37.50% 36.05% 0 IP Input
192 24538871 82738840 296 1.11% 0.71% 0.82% 0 Spanning Tree
458 5514 492 11207 0.63% 0.15% 0.63% 2 Virtual Exec
61 3872439 169098902 22 0.63% 0.63% 0.41% 0 RedEarth Tx Mana
```

- Too many packets getting to the CPU:

  "show controllers cpu-interface"

```
cpu-queue-frames retrieved dropped invalid hol-block stray
----------------- --------- ----------- -------- ----------- -------
rpc 2811788 0 0 0 0
stp 944641 0 0 0 0
ipc 280645 0 0 0 0
routing protocol 813536 0 0 0 0
L2 protocol 8787 0 0 0 0
remote console 2808 0 0 0 0
```

"debug platform cpu-queues"

```
L3If:Vlan200 L2If:GigabitEthernet1/0/3 DI:0xB4, LT:7, Vlan:200 SrcGID:3, SrcGIP:3,
ACLDLogIdx:0x0, MacDA:001d.46be.7541, MacSA: 0000.0300.0101 IP_SA:10.10.200.1
IP_DA:10.10.200.5 IP_Protol:1
```
Troubleshooting Summary

• Multiple Troubleshooting Tools are available – CLI sometimes fastest way to pinpoint a problem

• Documentation is invaluable
  
  Provides a baseline

  Topology diagrams save troubleshooting time

• Fully define the problem

• Gather as much detail as possible

• Work through the possible root causes

• High Switch CPU may indicate traffic anomalies
  
  Does not affect packet forwarding

• Several Troubleshooting Guides available

  Cisco.com, Cisco Press, CCNP, other online resources
Thank you.