UCS Network Utilization Monitoring: Configuration and Best Practice

Steve McQuerry
Technical Marketing Engineer
Unified Computing Systems
Cisco Systems, Inc.
Document Version 1.0

Table of Contents

What You Will Learn

One of the many goals for any operation team is to be a proactive with their network components monitoring resources for use and errors in order to quickly identify upward trends of use and excessive errors. The goal of this paper is to establish a baseline practice of statistics that can and should be monitored in a UCSM environment.

Within UCSM there are many statistics available. While these statistics can be viewed through SNMP tools they are also available within the UCSM GUI. In addition to being visible for some statistics it is possible to configure UCSM to monitor statistics through a mechanism called a threshold policy. When a statistic crosses a user defined threshold UCSM will generate a system alert and a SNMP trap notifying administrators.

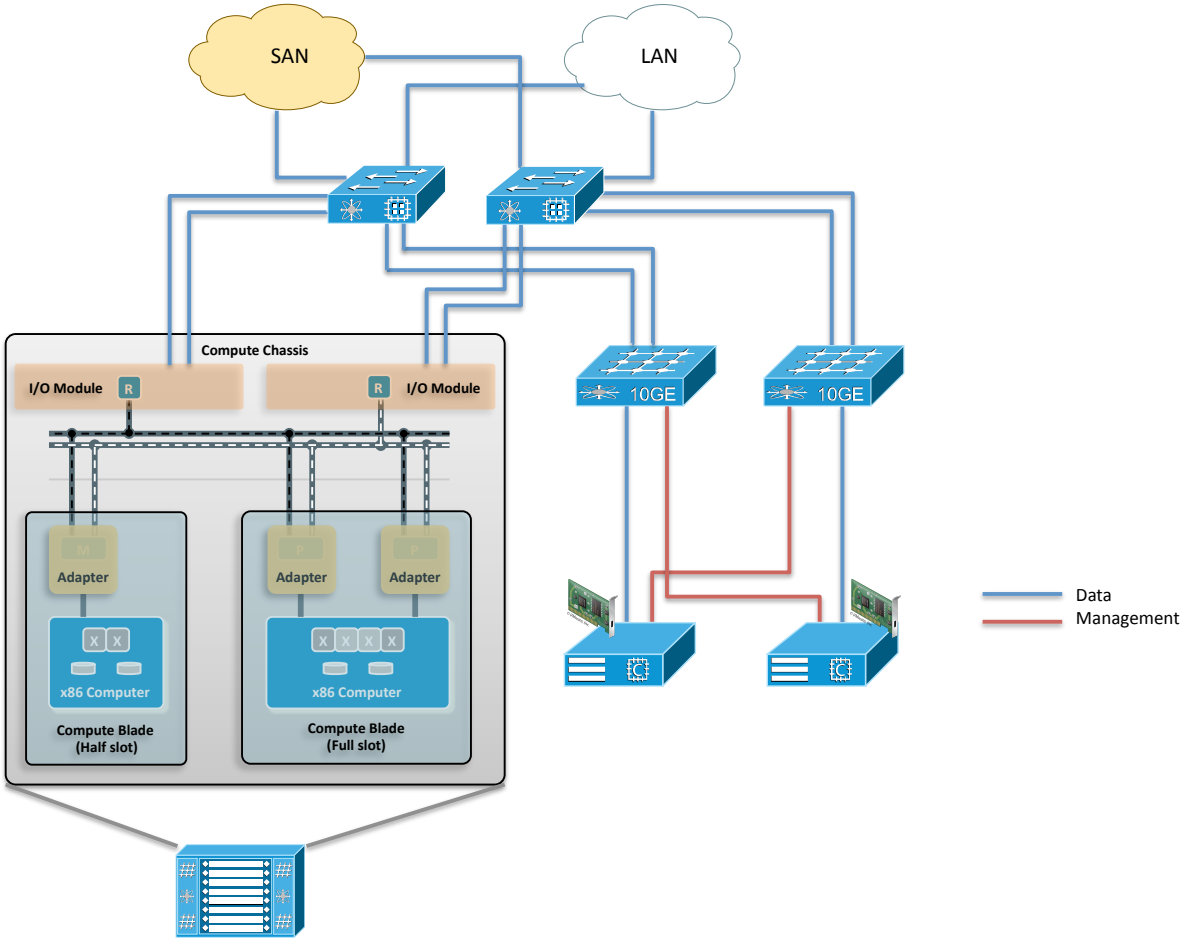
While there are many components of the network system that are visible within UCSM only a few can be monitored today with the threshold policy. The goal of this paper is to:

- Describe the operation of the statistics gathering and threshold policy within UCMS GUI
- Describe the configuration of a threshold policy within the UCSM GUI
- Recommended practices or guidelines for monitoring network traffic and errors

Unified Computing System (UCS) Overview

Cisco UCS tightly integrates unified fabric and compute. The makeup of the system is comprise of 4 different elements:

- Compute system (blade chassis or rack mount) – There are 2 different compute form factors. The first one is the blade chassis that houses multiple server blades, for both half width and full width blades. The other form factor is the rack server, which comes in 1, 2, or 4 RU.
- Port Adapters – Compute's network interface. The available supported port adapter for the Cisco UCS compute can either be from a qualified 3rd party network interface card (NIC) or Cisco's port adapter, virtual interface card (VIC).
- Fabric Interconnect (FI) – The fabric interconnect is a low latency unified fabric which provides 10G, fibre channel (FC), and fibre channel over Ethernet (FCoE) connectivity. It provides a passage for the compute system to the ethernet and storage network.
- Fabric Extender (FEX) – The fabric extender is an extension of the fabric interconnects, acting as a linecard that the compute element directly attaches to. The FEX module for the blade chassis is I/O module that is housed within the blade chassis. In the case of the rack server, it connects into the Nexus 2232PP.



	Blade Chassis (B-series)	Rack Servers (C-series)	
Fabric Interconnect (FI)	UCS 6100/6200		
Fabric Extender (FEX)	I/O Module UCS 2100/2200	10GE Nexus 2232PP	
Port Adapters	Adapter UCS 81KR UCS 82598KR UCS VIC 1240/1280 M71KR-E/Q		1GbE BCM95709A0906G Intel ET2 Quad Port Intel Quad Port Gb ET
			10GbE UCS P81E Intel X520-DA2 Emulex OCe10102-FX-C

Figure 1: Cisco UCS system

UCS Manager (UCSM)

Cisco UCS Manager is a unified management that is embedded within the fabric interconnects. It gives the end user a single point of management for both hardware and software across multiple chassis, rack-mount servers, virtual machines, and fabric interconnects. Cisco UCS Manager exposes all its functions through native XML API, which can be access through Cisco UCS Manager's GUI or CLI.

Monitoring Overview

UCSM has access to the detailed statistics and sensors of the various components. In the case of networking these statistics are collected from the NXOS running on the Fabric Interconnect and populated in the Data Management Engine (DME) within UCSM.

Collection Policy

Stats Management in the system consists of collection policies. Collection policies exist for Adapters, Chassis, FEX (IOM), Hosts, Ports, and Servers. Currently the Hosts collection policy is inactive and does not collect specific statistics from any component within the system.

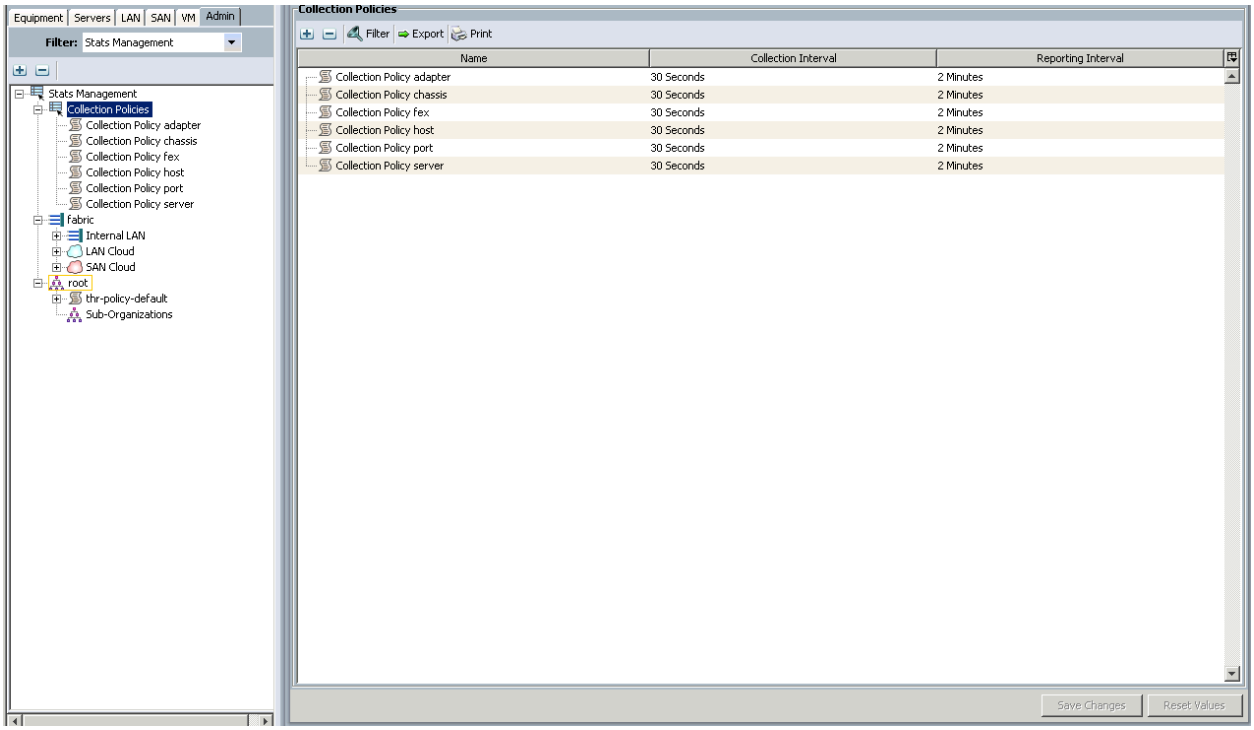


Figure 2: Collection Policies

A collection policy consists of a collection interval and a reporting interval, which are configurable parameters. The collection interval is how frequently the statistics are pulled from the system, where as the reporting interval is how frequently statistics are stored in the DME. The system can store values

from the last 5 reporting intervals and these values are used for computing averages as shown in the statistic output within the UCSM GUI in Figure 3.

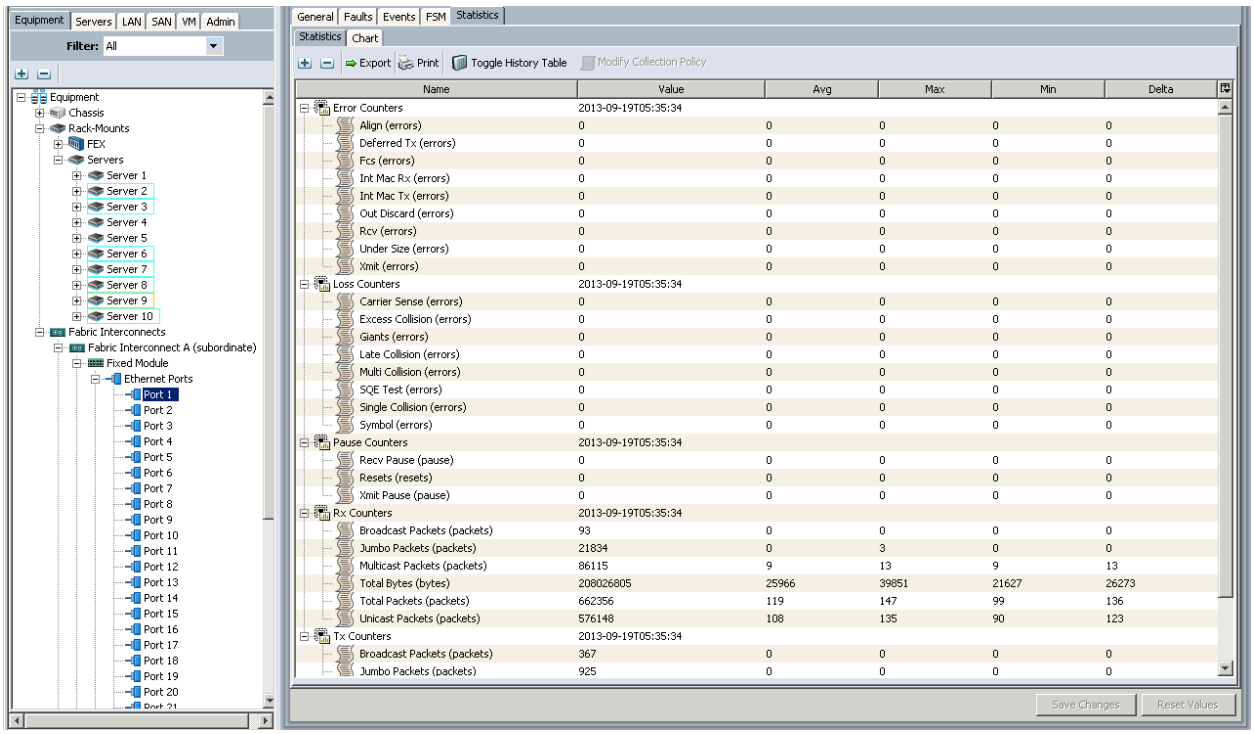


Figure 3: Fabric Interconnect Port Statistics

The default setting for the collection interval is 1 minute while the default setting for the reporting interval is 15 minutes. To provide a more granular look into the system it is recommended to change the polling interval to 30 seconds. This has no noticeable (< 0.1%) affect on the CPU load of the FI. The value of the reporting interval becomes important when defining threshold policies.

Threshold Policy

The threshold policy is another important component of monitoring the system. A threshold policy is a user definable policy within UCSM that provides a mechanism for alerts to be raised when a statistic crosses a given threshold. Network threshold policies can be defined within the LAN tab. In general there are 3 separate policies: LAN Cloud, LAN Internal, & Server. LAN Cloud policies relate to uplink ports for traffic leaving the Fabric Interconnect, LAN internal are the links between the FI and the IOM, and Server policies deal with the Server Port adapter and any vNICs on those adapters. These policies can all be configured from the LAN tab as illustrated in Figure 4.

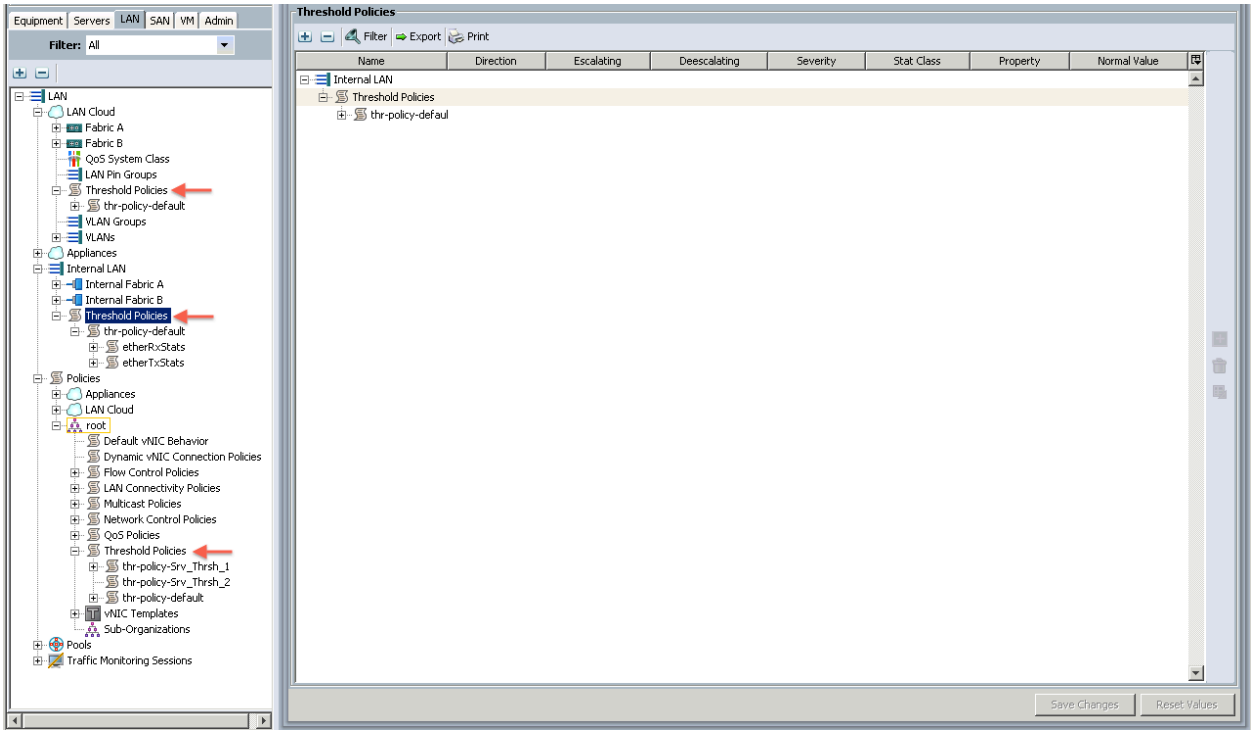


Figure 4: LAN Threshold Policies

Threshold policies can also be defined in the SAN tab for FC ports. This is only definable for the uplink ports (SAN Cloud) as shown in figure 5, this is because internal system traffic is FCoE which is Ethernet by nature. However the FC interface on the server can be monitored independently from the Server Ethernet traffic. This is definable within the server threshold policies, which show up under root in the LAN, SAN, and Server tab.

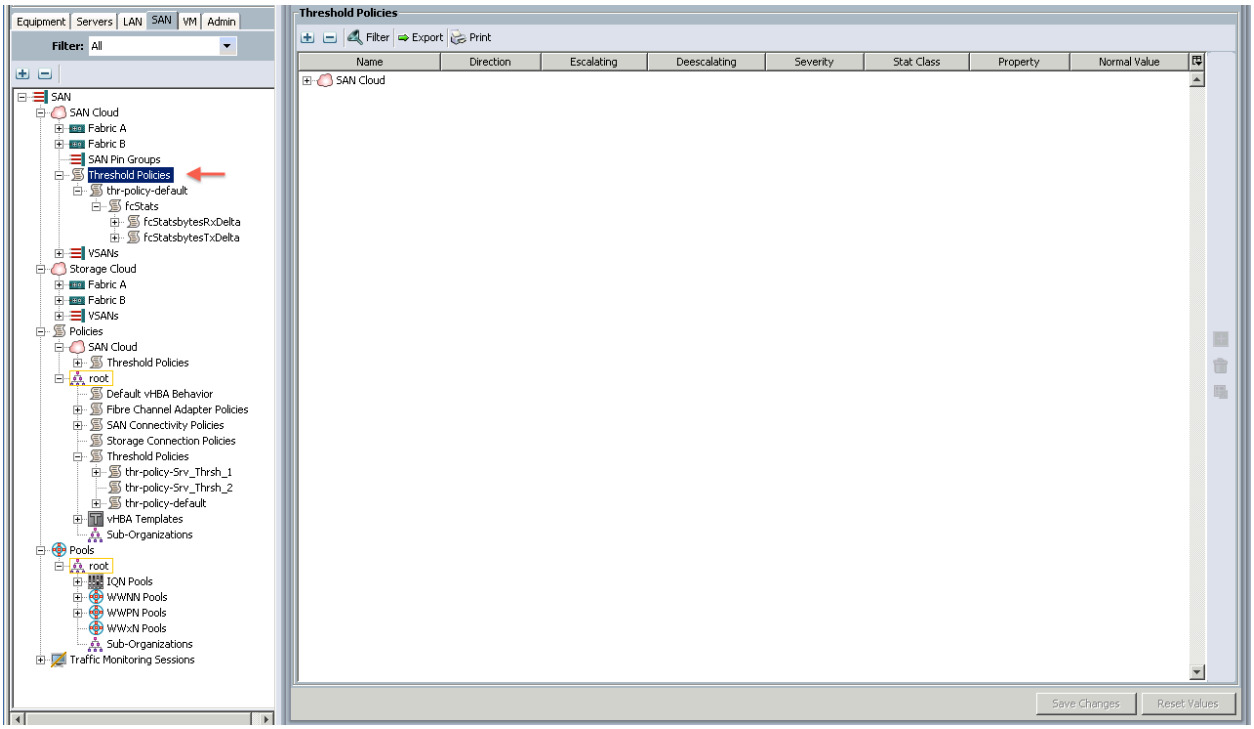


Figure 5: SAN Threshold Policies

Monitoring Best Practices

When monitoring UCSM there are 2 items that should be monitored. The first item is throughput for hotspot detection. The key areas to monitor are at the aggregation points, such as the links between FEX and fabric interconnect, fabric interconnect and upstream devices (SAN and/LAN switches), and between FEX and port adapters as show in Figure 6. These areas provide a good birds eye view of the traffic load on the system. If any of these areas are saturated, then the system can raise a flag so the administrators can monitoring it closer and take appropriate actions before it becomes a problem. Administrators can focus and drill down to the specific problem area to determine the cause and then come up with solution to alleviate the load either by redirecting the traffic or moving/reducing the workload.

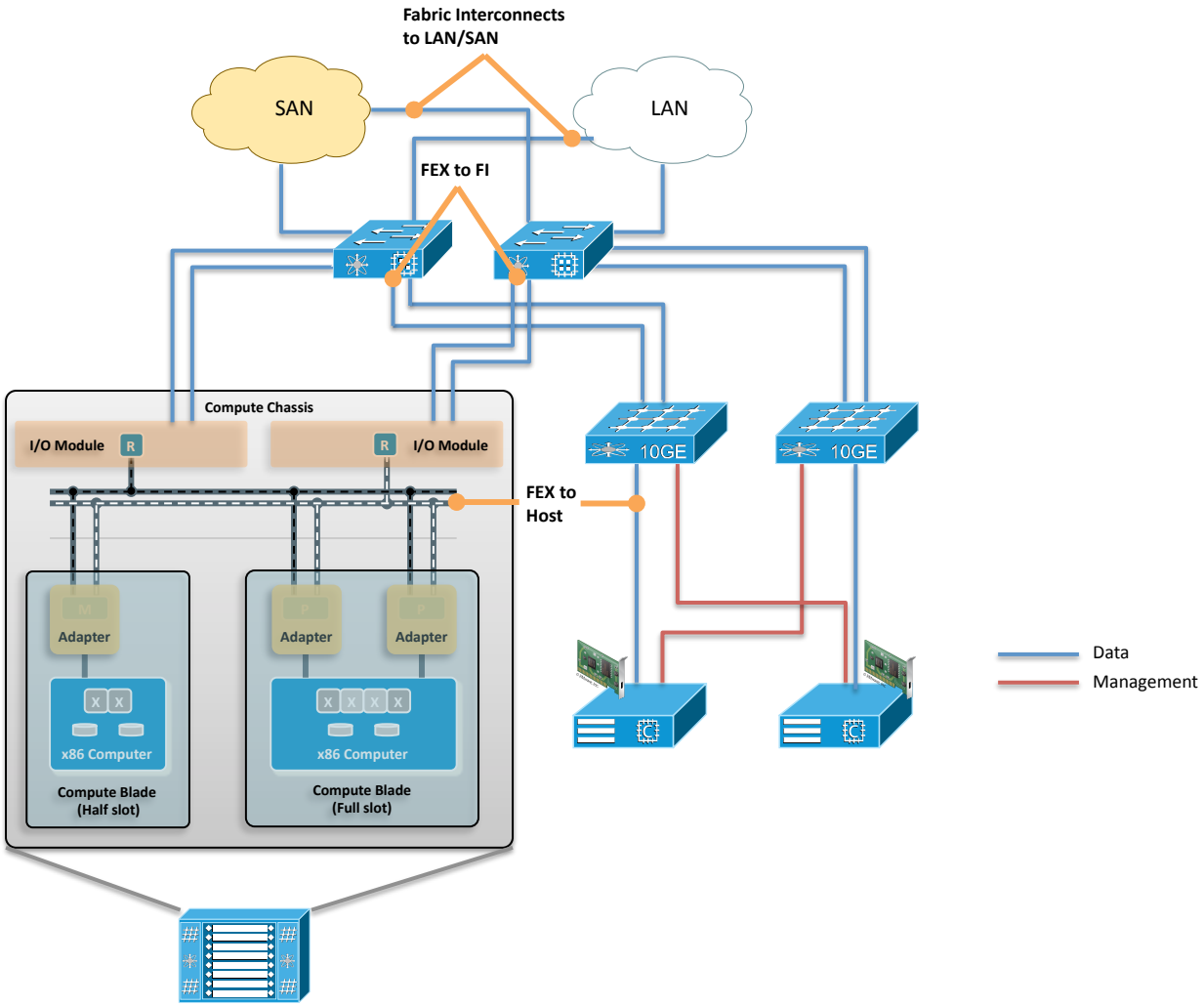


Figure 6: Potential hot spots (congestion) within Cisco UCS.

To monitor for hotspots a threshold policy will be defined in the LAN tab for the cloud (uplinks) and the Internal LAN (IOM connections). In a threshold policy the user creates a threshold class. The class defines which statistics will be monitored. For LAN bandwidth monitoring the user will create two threshold classes: EtherRxStats & EtherTxStats. Within the threshold class they will create a threshold definition. Within the threshold definition, different severity variances (ranging from low to critical) can be defined. This will determine how the alert is seen within UCSM. In this definition, the user defines the normal (expected) value and the up and down values for a respective severity level. The up and down values are compared against the normal value. The value defined under the "up" field is the trigger point, which UCS Manager will raise the flag when it is crossed. When the condition drops down below the "down" value, then UCS Manager will remove the flag.

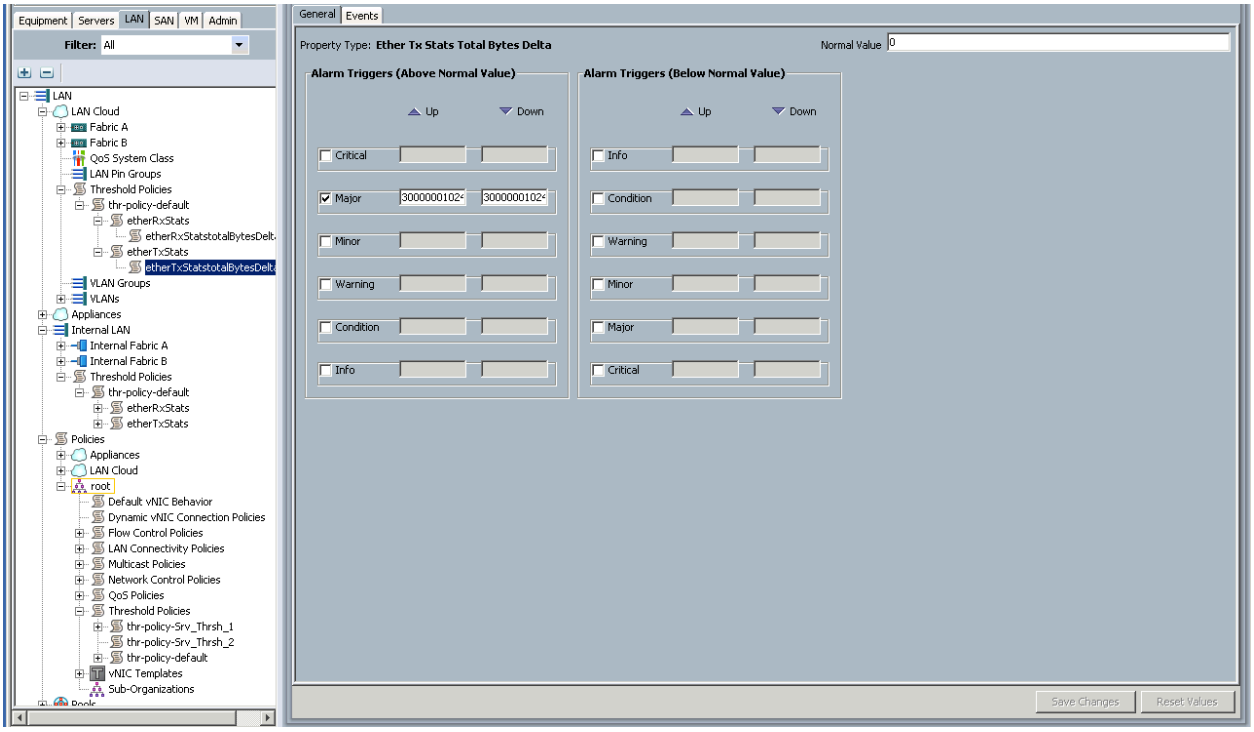


Figure 7: Threshold definition.

For each class created Tx and Rx the administrator needs to choose a value on which to alert. The values are compared against the Normal value and the statistics which are gathered by the system. Many of the statistics gathered are counters that increase every time they are read. These counters are not useful for alerting since once the threshold has been crossed the alert will stay on until the counters are cleared. The best values to be used with alerts are Avg. or Delta counters. Since Avg counters are the period of the last 5 historical data points, they are not as granular as Delta counters. A Delta counter is the change since the last collection interval. Since we set the collection interval at 30 seconds this gives us the most granular look at data within the system. We need to calculate the change over this interval and compare it against the normal value. For throughput monitoring the best practice is to set the normal value to 0 so that we can compare the traffic on a scale of 0 to 100%.

Network throughput is measured in bits per second (bps) but there is no specific statistic category within UCSM that provides that detail. However UCSM does provide total delta bytes sent or received on a link. Given this information and the collection interval of 30 seconds we can calculate the number of bytes sent in that time. We use this to extrapolate our threshold values based on the rate that we do not want our traffic to exceed. To calculate we need to do the following:

Maximum Traffic Rate (Threshold) in bits per second.

(Threshold) bits/s / 8 bits/byte = (Threshold) bytes/s
(Threshold) bytes/s * 30 seconds = (Threshold) bytes

Every 30 seconds, the expected change in bytes should be the value. So if the threshold is 75% of 10G (7.5Gbps), then the total bytes received in a 30 second time span should be roughly 28 gigabytes (GB) – as shown below.

[7,500,000,000 / 8] x 30 = 28,125,000,000

So the “up” value should be 28,125,000,000 and the “down” value should be any value less than 28,125,000,000. The tool will compare these against the delta value. If the delta value is >28GB then the throughput was more than 7.5Gbps.

Table 1: The below table provides the expected bytes over a period of 30 second polling.

Percentage of BW	Rate (Gbps)	Bytes	Total of Bytes over 30 seconds
100	10	1,250,000,000	37,500,000,000
90	9	1,125,000,000	33,750,000,000
85	8.5	1,062,500,000	31,875,000,000
80	8	1,000,000,000	30,000,000,000
75	7.5	937,500,000	28,125,000,000
70	7	875,000,000	26,250,000,000

While the threshold policy allows user to define multiple trigger events, it isn't necessary to configure all of them. Configuring just the major and critical level should be sufficient as shown in Figure 7. Using the values that has been provided in Table 1, the up and down can be derived from it.

Recommendation: Traffic should be monitored in both directions by creating a threshold definition for both the RX and the TX class. Also a threshold policy should be configured for the LAN Cloud and Internal LAN. If there are any port channels, do not configure the threshold based on the aggregate speed of the link. This could lead to an inability to detect the saturation of a single link. In general it is recommended that a warning should be set at 80% utilization and critical at 90% utilization. The early warning signals the operation team should keep a closer watch, and if it reaches critical level then action should be taken before services are impacted.

Another Key area to measure within UCSM would be the errors on the links. Errors can also be measured within the same threshold policies defined in the LAN tab for monitoring throughput. To monitor errors users would need to add an additional set of classes for error counters and loss counters. The key Error counters to monitor are: FCS, Int Max Rx, Int Max Tx, and Out Discard. These errors are typically key indicators of hardware or cabling issues. Another important item to monitor is the carrier sense in the loss counters. If the carrier sense for a port is transitioning frequently (flapping) it can have undesirable results, especially if that port is an uplink. These values are shown in Figure 8.

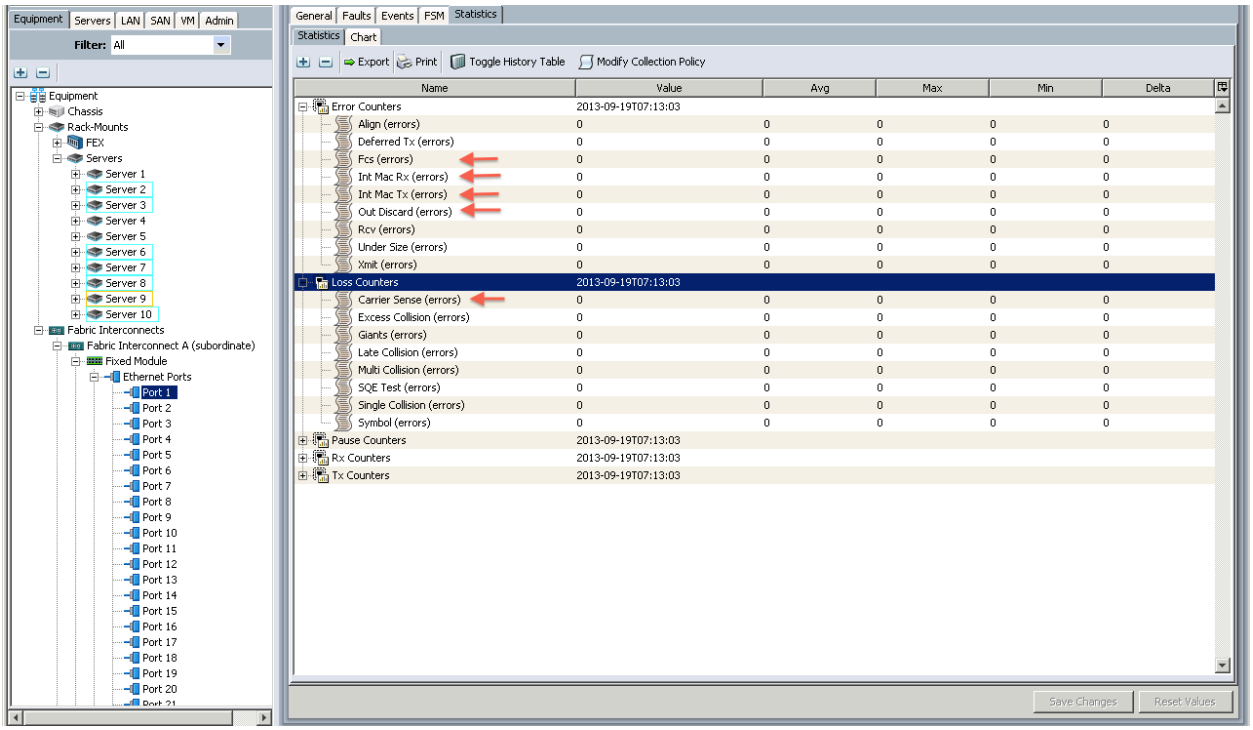


Figure 8: Error counters.

To configure threshold policies a user would need to add two additional classes: EtherErrStats, and EtherLossStats. Under each class you would need to add thresholds for the items identified as shown in figure 9.

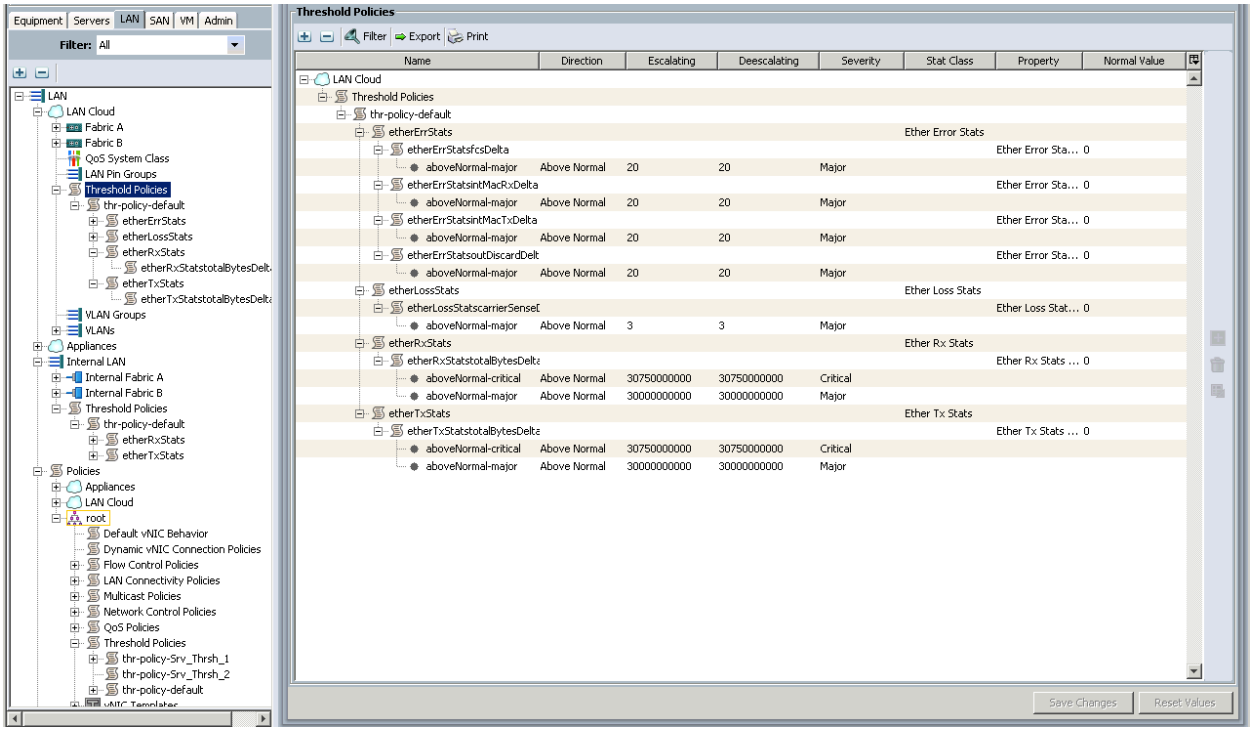


Figure 9: Threshold definitions.

The recommendations for error counters are a delta of 20 over the collection interval. This would mean that 20 errors were encountered in a 30 second period. In general we expect the error rates to be near 0 for a system, therefore 20 in a short time would be significant. For carrier loss the threshold should be set to 3. This would be 3 link transitions in a 30 second period which would be abnormal even if someone were moving cables within the infrastructure.

These general recommendations for traffic monitoring and error monitoring are specific to the LAN Cloud and Internal LAN. These can also be used as a guideline for the SAN cloud settings as well. You could choose to do these for the individual server connections using a threshold policy for a service profile, but it may make more sense to monitor BW statistics based on OS type. For example bare metal OS 50-60%, Hypervisor 70-80%. There are no specific recommendations from Cisco for server monitoring as this is highly dependent on the user application.

Conclusion

Cisco Unified Computing System was architected from the ground up with management and operation being top of mind. The ability to manage the unified fabric and compute as a single entity and with the added monitoring feature, it helps lower the overall OPEX at the same time allows operation team to be more proactive instead of reactive organization.

Fore more information

Contact your local Cisco representative or visit:
Cisco Unified Computing System
http://www.cisco.com/go/unifiedcomputing
Cisco UCS Manager Architecture White Paper
http://www.cisco.com/en/US/prod/collateral/ps10265/ps10281/white_paper_c11-525344.html

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883

Asia Pacific Headquarters
Cisco Systems, Inc.
168 Robinson Road
#28-01 Capital Tower
Singapore 068912
www.cisco.com
Tel: +65 6317 7777
Fax: +65 6317 7799

Europe Headquarters
Cisco Systems International BV
Haarlerbergpark
Haarlerbergweg 13-19
1101 CH Amsterdam
The Netherlands
www-europe.cisco.com
Tel: +31 0 800 020 0791
Fax: +31 0 20 357 1100

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**