Finding and Remediating Slow Draining Devices on Fibre Channel SANs

Slow Draining Devices

Virtual Instruments' VirtualWisdom® is the leading product in the industry that helps SAN administrators to quickly find the root cause of SAN performance problems, including slow draining devices. VirtualWisdom's comprehensive instrumentation and Storage Area Network (SAN) I/O measurement capabilities reduce application response time, increase availability, and improve resource utilization.

Problem

Fibre Channel network becomes unresponsive; I/O becomes bottlenecked; no apparent cause

Impact

Application response time slows to unacceptable levels. Latency can cause time-outs.

Finding the problem

Without VirtualWisdom, troubleshooting is often a "trial and error" process and a guessing game, though switch vendors are adding features to their switches to help with diagnostics by doing a better job on monitoring buffer-to-buffer credits. However, such additions are incomplete and still require the engagement of the vendor's "expert" to interpret the proprietary data and, in most cases, you are still unable to find the root cause – the slow draining device. To reduce time to resolution, the correlation of a unique set of metrics, end to end visibility of the environment, and easy to use applied analytics are required, to not only successfully find the slow draining device, but to do it quickly.

> "The most common performance problem that gets reported to us on our Dell/EMC flash arrays isn't media access time, it's slow draining devices".
>
> Adnan Sahin
> Senior Director,
> Distinguished Engineer
> Dell/EMC

VirtualWisdom offers a unique set of metrics to help determine if a slowdown is caused by a server side (slow draining) device, on the storage side, or other issue. For instance, VirtualWisdom's SAN Availability Probe solution can pick up on one of the hearts of the issue, lack of credit, which eventually leads to cause class 3 discards. While credit issues can be detected or observed with the VirtualWisdom SAN Availability Probe, the offending link can be hard to diagnose, as the link that reports the discard usually isn't the one that is causing the credit issue. Though there are other causes of class 3 discards, the pattern of their occurrence is quite different than the pattern of occurrence of class 3 discards due to credit issues.

Another metric that the switches offer is often referred to as "time with zero credit" or something similar; it varies by vendor and switch. But the meaning is basically the same. Unfortunately, this counter has been historically inaccurate and has been more of a red herring than useful for analyzing credit issues. But as mentioned above, some of the switch vendors have fixed this counter for their newer switches, and we are testing this to see if it is indeed valuable.

With our SAN Performance Probe hardware, Read/Write/Other Exchange Completion Times can be easily set to trigger alerts. Using VirtualWisdom is the best way to detect and diagnose credit issues. This is because it sees the
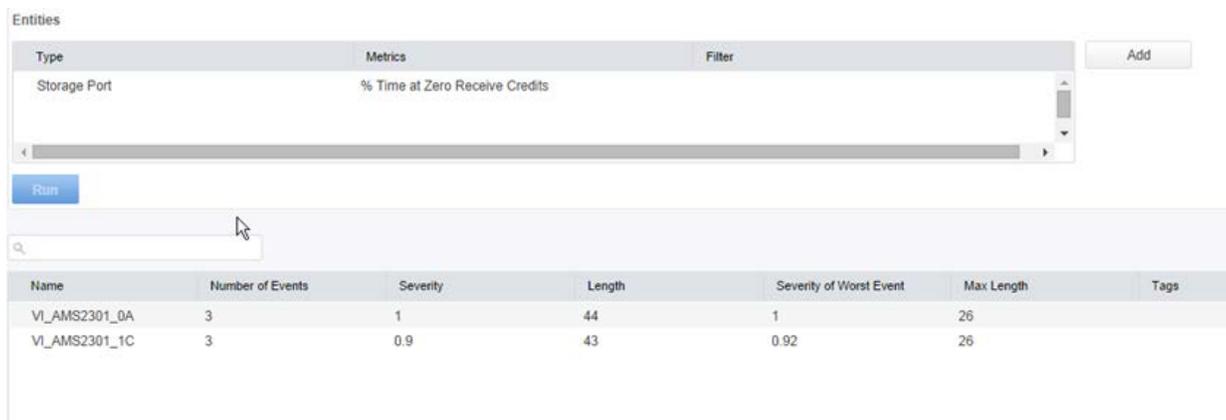
full symptoms of credit issues … which are high latency and large numbers of exchanges that remain pending. This does not usually require that the VirtualWisdom probes be placed on the link having the credit issue, and while there are a few cases where this is true, by and large, the device or devices causing credit starvation can be detected by monitoring the storage ports they are accessing. This is because switches have only two choices when they receive a frame. They have to either route it or drop it within a very short timeframe. If the switch drops the frame, then the class 3 discard counter is incremented. In order to avoid having to drop frames, lack of credit on a port quickly flows back through an environment with a balance between delaying credit and dropping frames. As a result, exchange completion times increase. The number of pending and minimum pending exchanges rapidly increases for the devices that are having the biggest impact to credit in the environment. There is also a much larger difference between the "time to first data" measurement (which is how long the array takes to provide the first frame of data for a read) and the overall transaction time because the arrays are ready with the data but waiting for credit to send it onto the link.

The following example below shows potential steps for locating a slow draining device.

The VirtualWisdom Analytics Event Advisor is configured to search for events where % of Time at Zero Received Credits by a Storage ports is abnormal.

Name	Number of Events	Severity	Length	Severity of Worst Event	Max Length	Tags
VI_AMS2301_0A	3	1	44	1	26	
VI_AMS2301_1C	3	0.9	43	0.92	26	

The Event Advisor returns two possibilities: The screen above shows one with Severity 1, which occurred three times.

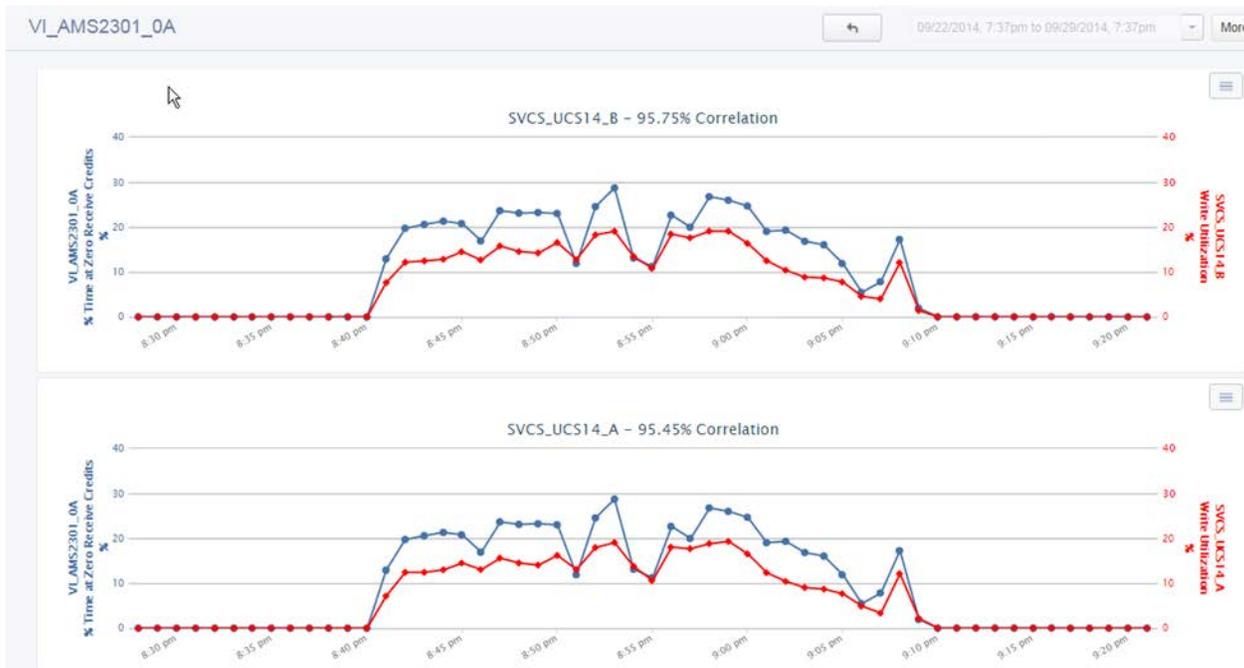Next, the Trend Match capability for the top chart is used, and finds 5 matches, below:

The Trend Match capability will search for an HBA Port where the Write Utilization matches the "Buffer at Zero" profile.

SVCS_UCS14_B – 95.75% Correlation

SVCS_UCS14_A – 95.45% Correlation

The Trend Match returns a greater than 95% match with Host SVCS_UCS14 HBAs A and B; so the Event Advisor and Trend Matching capabilities helped locate the offending device.

Mitigating the Problem

To resolve latency issues, it is important to consider how many commands are outstanding, how quickly the storage array is responding to each request, how the exchange times relate to the demand in the environment, and whether recent change requests could be contributing to the problem. Remediation may include re-configuration, re-routing or replacing equipment, depending upon the underlying cause.

Generally speaking, you must identify the demand. The quickest way to identify and understand the performance is to look at the response times. Live Reports showing "Performance and Demand" can be used to find any latency in the environment, as well as to determine the relationship between the demand for the specific devices and the performance.

If high latency is detected but there appears to be no correlation to demand, a Live Report showing "Performance and Pending Exchanges" can be run to determine if there is a correlation between the response times and the queue depth settings or pending exchanges. queue depth settings often can help to mitigate slow draining devices, which are the most common cause of credit issues. Queue depth settings impact the number of transactions that can be open by a device at any one time and credit issues deal with individual frames rather than transactions. By limiting the

number of transactions, you can throttle slow devices down to levels of data that they can consume and prevent them from impacting other devices that share some of the same resources such as ISLs and storage ports. Typically, queue depth settings are set too high for optimum performance. The VirtualWisdom SAN Performance Probe allows users to really see the true impact of the queue depth settings and overall latency in the environment. Adding a VirtualWisdom Server Virtualization Probe enables users to see not only how the queue depth setting impacts the storage environment, but also to see how it affects overall server response times.

Performance should also be configured and understood at the Server and Storage level. This is typically the closest match to the values that Server teams are seeing and using to evaluate SAN health. It is also typically the best level to use when setting alert thresholds.

It is often important to separate and categorize the performance by the size of the transactions, the types of devices in the environment, and any tiering or other big differences in expected performance. Reporting and understanding the performance at these different levels can be critical to the successful remediation of latency problems.

VirtualWisdom has the capability to profile IO Performance for an Application, Server, Storage or other devices.

Armed with all the Performance information provided by VirtualWisdom, proactive measures, such as Alarms, can be implemented to detect deviations from SLAs or Baseline values, and avoid slowdowns.

VIRTUAL INSTRUMENTS®