



Cisco[®] MDS DIRL SAN Congestion Software
Performance Validation



October 2021

DR210721G

Contents

- 1.0 Executive Summary 3
- 2.0 Introduction..... 5
- 3.0 How We Did It..... 7
- 4.0 DIRM Software Validation..... 8
 - 4.1 Slow Drain Congestion 8
 - 4.2 Over-utilization Congestion 11
- About Cisco 14
- About Miercom Performance Verified 14
- About Miercom..... 14
- Use of This Report..... 15

1.0 Executive Summary

Storage Area Networks (SAN) are dedicated high-speed networks that allow devices to access large amounts of data. To build SANs, most enterprises use Fibre Channel technology to achieve higher levels of scalability, reliability, and performance. As high-performance all-flash and NVMe storage arrays are connected to Fibre Channel SANs, the issue of congestion and slow drain is becoming a bigger problem, resulting in the need for enhanced techniques to identify, correct and prevent SAN congestion.

The Cisco® MDS 9000 Series directors and switches provide a solution to correct this problem with its Dynamic Ingress Rate Limiting (DIRL) software, a new innovation that identifies and prevents SAN congestion. SAN congestion caused by slow drain and over-utilization, is prevented by the MDS DIRL software with no impact on the target device and other connected devices. DIRL resolves the root cause of congestion by dynamically adapting traffic which is being created or caused by the culprit SAN device. In simple terms, DIRL alleviates egress congestion on a switch port by limiting ingress data.

The MDS DIRL congestion prevention software does not drop frames. Instead, it uses the B2B (Buffer to Buffer) credit pacing mechanism to control frames to allow both slow and fast network devices to coexist in the same SAN. DIRL does not require a software license or dependency on endpoint devices, and it is a standard feature of the Cisco MDS 9000 directors and switches.

Key Findings

- **No need to upgrade end devices.** MDS DIRL software is fully integrated with Cisco MDS 9000 directors and switches. It is not dependent on the end devices to identify and prevent congestion.
- **Zeros in on SAN congestion issues.** MDS DIRL software efficiently alleviates SAN congestion caused by Slow Drain and Over-Utilization.
- **Isolates congestion issues and affected devices.** MDS DIRL software targets the affected device without disrupting performance of other devices in the fabric – automatically pacing ingress data to prevent the spread of congestion.
- **Proves minimal impact on affected device.** MDS DIRL software quickly adapts to traffic and minimally impacts the congested device for a smooth end user experience.
- **New innovation for the Cisco MDS 9000 Series.** MDS DIRL software is a standard feature on new Cisco MDS 9000 switches, and it is available as a free upgrade for all existing Cisco MDS 9000 Series switches
- **Topology independent.** MDS DIRL software works in all environment types: single-switch fabric, edge-core fabrics, or edge-core-edge fabrics.

Based on our findings, the unique Dynamic Ingress Rate Limiting (DIRL) software of the Cisco MDS 9000 Series demonstrates a superior ability to identify and resolve SAN congestion without impacting network devices or performance. We proudly award the Cisco MDS 9000 Series DIRL software the **Miercom Performance Verified** certification.

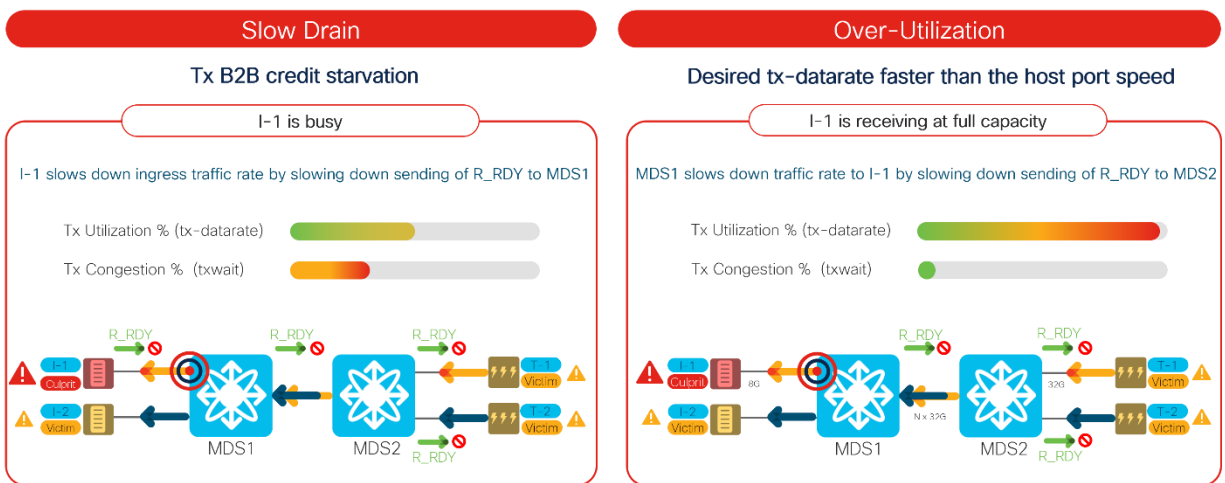
Robert Smithers
CEO, Miercom



2.0 Introduction

Cisco is a leader in SAN technology – offering solutions for slow drain and over-utilization, that cause network congestion. In block-storage networks, storage arrays do not automatically send data. A server must request it by initiating I/O commands to the storage arrays. SAN congestion is created when the server asks for more data than it can handle. Cisco's MDS DIRL software identifies the congestion, where egress congestion originates on the switch ports, and then limits ingress data from that device. DIRL dynamically adjusts ingress frame rate until egress congestion disappears.

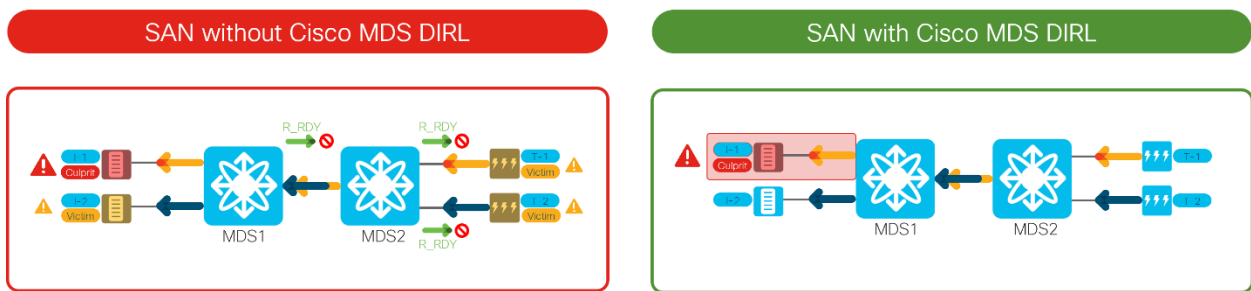
Common Causes of SAN Congestion



Frames are not dropped in FC fabric. Frames consume switch buffers causing a fabric-wide congestion

Source: Cisco Systems 2021

SAN Congestion vs No SAN Congestion



Source: Cisco Systems 2021

The Cisco MDS 9000 Series provides protection against slow-drain and over-utilization scenarios by using the following:

Slow Drain Device Detection

- Uses high-fidelity metrics that are available on Cisco Fibre Channel port-ASICs
 - ✓ *TxWait*: duration a switch port cannot transmit a frame due to lack of transmit B2B credits; collected every 2.5 microseconds with 1 second automatic alert
 - ✓ *Slowport monitor*: duration a switch port cannot transmit a frame due to lack of transmit B2B credits; collected every 1 millisecond with 1 second automatic alert only on the continuous duration of transmit B2B credit unavailability
- Sends automatic notifications via the Cisco Port-Monitor (PMON) feature, providing policy-based configuration to detect, notify, and take automatic actions to prevent congestion and slow drain
- Cisco Nexus Dashboard SAN Controller provides long-term trending and correlation using the slow-drain analysis feature
- Cisco SAN Analytics technology provides I/O flow metrics, such as Exchange Completion Time (ECT), Data Access Latency (DAL), I/O sizes and more

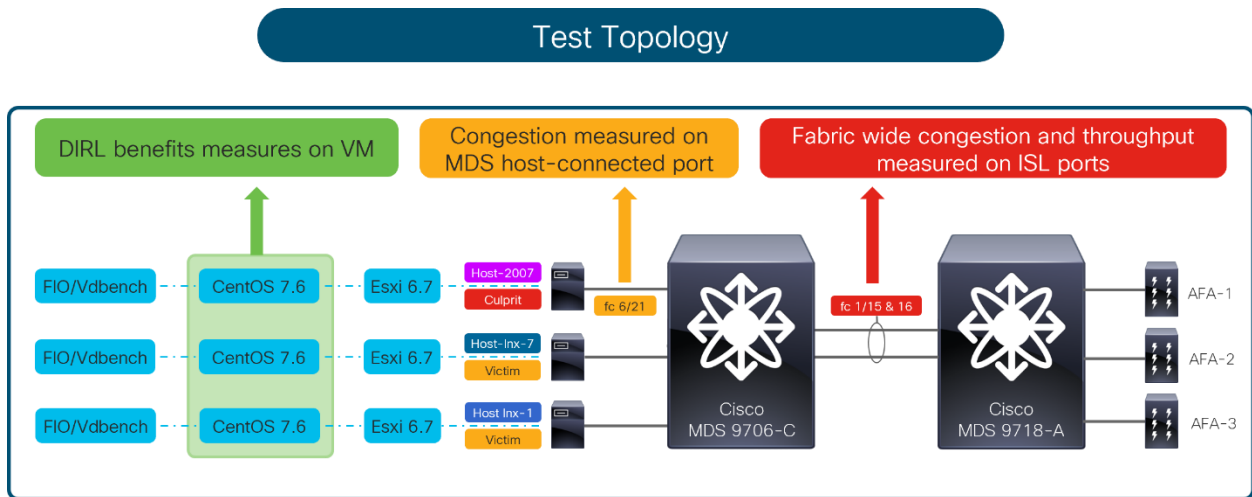
Congestion Prevention

Cisco MDS DURL software prevents SAN congestion and slow drain:

- Limits ingress traffic by automatically controlling frame flow using B2B credit pacing mechanism of the Cisco Fibre Channel ASICs
- Does not drop any frames
- End device upgrades are not needed to use DURL
- Dynamically adjusts to traffic profile of the host
- Rate-limiting is only applied to the congested host, as to not affect other devices
- Works in edge-core, edge-core-edge, or collapsed core (single switch fabric) topologies

3.0 How We Did It


Using a realistic network environment, we tested the capabilities of the Cisco MDS 9000 Series and the ability of the DIRL software to relieve various SAN congestion scenarios.




The Cisco MDS 9700 Series Multilayer Directors were connected to the servers. The demonstration focused on Host 2007 presenting congestion simulations for Over-utilization and Slow Drain use cases. Impacts were viewed on Host Inx-7 and Inx-1.

Test Tools

The following tools are a representative list of software tools we used to carry out our analysis.

- 

Grafana 7.5.7
Grafana is an open-source multi-platform tool that offers analytics and visualizations using charts, graphs, and alerts to observe and monitor data.
- 

CentOS 7.6.1810
An open-source Linux distribution operating system, commonly used for server administration.

4.0 MDS DURL Software Validation

4.1 Slow Drain Congestion

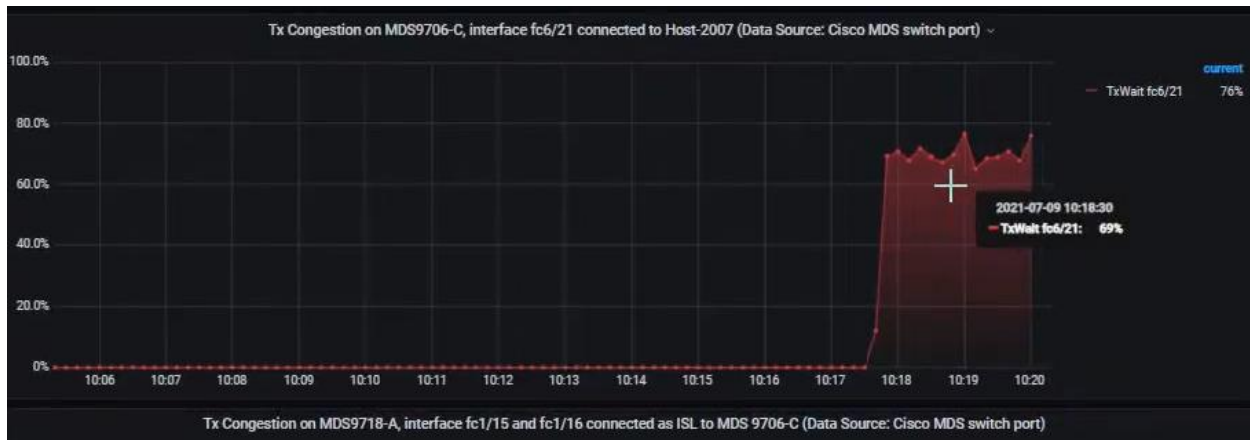
We observed the creation of “read” commands at a size of 256K on Host2007, making it the culprit device causing the slow drain congestion. The 256K I/O size was chosen for ease of congestion simulation. The behavior of DURL on the victims Host Inx-1 and Host Inx-7 was observed.



Throughput for the Inx-1 and Inx-7 without any SAN congestion.



Traffic increased for the Host2007 and declined for Host Inx-7 and Host Inx-1, causing a slow-drain congestion event.



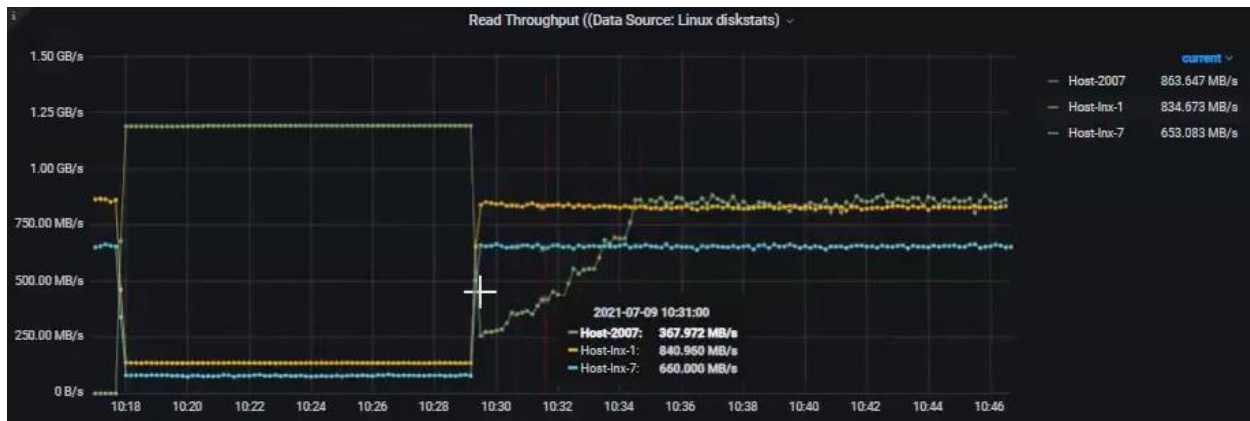
In the slow drain scenario, the TxWait fc6/21, connected to the Host-2007, increased when traffic started. TxWait represents the amount of time that the port will not transmit because of congestion. In the screenshot shown, the port fc6/21 could not transmit 69 percent of the time because of congestion.

The port fc6/21, connected to the Cisco MDS 9706 Multilayer Director, could detect any congestion conditions, and react accordingly when DIRL was enabled.

All Cisco MDS 9000 Series have a feature called the Port Monitor that must be enabled. This feature enables automatic monitoring and alerting on the switch. Port Monitor detects data plane events at a low granularity. DIRL was one of the actions available within the Port Monitor, which can be configured and enabled as needed.

```
port-monitor name fabricmon_edge_policy
 logical-type edge
 counter link-loss poll-interval 30 delta rising-threshold 5 event 4 falling-threshold 1 event 4 alerts syslog rmon portguard FPIN
 counter sync-loss poll-interval 30 delta rising-threshold 5 event 4 falling-threshold 1 event 4 alerts syslog rmon portguard FPIN
 counter signal-loss poll-interval 30 delta rising-threshold 5 event 4 falling-threshold 1 event 4 alerts syslog rmon portguard FPIN
 counter invalid-words poll-interval 30 delta rising-threshold 1 event 4 falling-threshold 0 event 4 alerts syslog rmon portguard FPIN
 counter invalid-crc poll-interval 30 delta rising-threshold 5 event 4 falling-threshold 1 event 4 alerts syslog rmon portguard FPIN
 counter state-change poll-interval 60 delta rising-threshold 5 event 4 falling-threshold 0 event 4 alerts syslog rmon
 counter tx-discards poll-interval 60 delta rising-threshold 200 event 4 falling-threshold 10 event 4 alerts syslog rmon
 counter lr-rx poll-interval 60 delta rising-threshold 5 event 4 falling-threshold 1 event 4 alerts syslog rmon
 counter lr-tx poll-interval 60 delta rising-threshold 5 event 4 falling-threshold 1 event 4 alerts syslog rmon
 counter timeout-discards poll-interval 60 delta rising-threshold 200 event 4 falling-threshold 10 event 4 alerts syslog rmon
 counter credit-loss-reco poll-interval 1 delta rising-threshold 1 event 4 falling-threshold 0 event 4 alerts syslog rmon
 counter tx-credit-not-available poll-interval 1 delta rising-threshold 10 event 4 falling-threshold 0 event 4 alerts syslog rmon
 counter rx-datarate poll-interval 10 delta rising-threshold 80 event 4 falling-threshold 70 event 4 alerts syslog rmon obfl
 counter tx-datarate poll-interval 10 delta rising-threshold 80 event 4 falling-threshold 70 event 4 alerts syslog rmon obfl portguard DIRL
 no monitor counter err-pkt-from-port
 no monitor counter err-pkt-to-xbar
 no monitor counter err-pkt-from-xbar
 counter tx-slowport-oper-delay poll-interval 1 absolute rising-threshold 50 event 4 falling-threshold 0 event 4 alerts syslog rmon
 counter txwait poll-interval 1 delta rising-threshold 30 event 4 falling-threshold 10 event 4 alerts syslog rmon portguard DIRL
 counter txwait warning-signal-threshold 40 alarm-signal-threshold 60 portguard congestion-signals
 counter rx-datarate-burst poll-interval 10 delta rising-threshold 5 event 4 falling-threshold 1 event 4 alerts syslog rmon obfl datarate 90
 no monitor counter tx-datarate-burst
 counter input-errors poll-interval 60 delta rising-threshold 5 event 4 falling-threshold 1 event 4 alerts syslog rmon
```

Above is an example of a Port Monitor configuration, with the highlighted area displaying the specific DIRL configuration for the demonstration.



Activating the DIRL software provided instantaneous benefit to Host Inx-1 and Host Inx-7 which were affected by congestion due to slow-drain. The traffic to the slow-drain device, Host 2007, was initially rate-limited. Then, we viewed gradual traffic recovery by seeing traffic congestion decreasing, and the gradual stability in traffic. It is important to note that the DIRL limited the rate of the ingress traffic on the culprit device rather than cutting it off.

```
MDS9706-C# show fpm ingress-rate-limit events interface fc6/21
```

Interface	Counter	Event	Action	Operating port-speed Mbps	Input rate Mbps	Output rate Mbps	Current RL%	Applied RL%	Time
fc6/21	txwait	recovery	rate-recovery	8000	128.52	1149.37	3.3816	4.2270	Fri Jul 9 10:34:15 2021
fc6/21	txwait	recovery	rate-recovery	8000	121.26	936.89	2.7053	3.3816	Fri Jul 9 10:33:15 2021
fc6/21	txwait	recovery	rate-recovery	8000	91.14	751.74	2.1642	2.7053	Fri Jul 9 10:32:15 2021
fc6/21	txwait	recovery	rate-recovery	8000	73.27	624.10	1.7314	2.1642	Fri Jul 9 10:31:14 2021
fc6/21	txwait	recovery	rate-recovery	8000	58.14	530.74	1.3851	1.7314	Fri Jul 9 10:30:13 2021
fc6/21	txwait	rising	rate-reduction	8000	235.47	2054.48	100.0000	1.3851	Fri Jul 9 10:29:12 2021
fc6/21	txdatarate	recovery	rate-recovery	8000	0.40	0.41	94.9202	100.0000	Wed Jun 16 19:07:31 2021
fc6/21	txdatarate	recovery	rate-recovery	8000	0.35	0.34	75.9362	94.9202	Wed Jun 16 19:06:30 2021
fc6/21	txdatarate	recovery	rate-recovery	8000	0.39	0.41	60.7489	75.9362	Wed Jun 16 19:05:30 2021
fc6/21	txdatarate	recovery	rate-recovery	8000	0.18	0.22	48.5991	60.7489	Wed Jun 16 19:04:30 2021
fc6/21	txdatarate	recovery	rate-recovery	8000	0.27	0.29	38.8793	48.5991	Wed Jun 16 19:03:29 2021
fc6/21	txdatarate	recovery	rate-recovery	8000	0.28	0.26	31.1035	38.8793	Wed Jun 16 19:02:29 2021
fc6/21	txdatarate	recovery	rate-recovery	8000	0.26	0.29	24.8828	31.1035	Wed Jun 16 19:01:29 2021
fc6/21	txdatarate	recovery	rate-recovery	8000	0.21	0.29	19.9062	24.8828	Wed Jun 16 19:00:28 2021

The internal CLI (command line interface) is another way clients can view the Cisco MDS functioning via the log history information. As shown, we were viewing ingress rate-limited events located on the Cisco MDS 9000. At the timestamp "10:29:12", the switch accurately detected that the txwait counter was 'rising' and took the appropriate action as viewed in the Action column; rate-reduction was followed by rate-recovery.

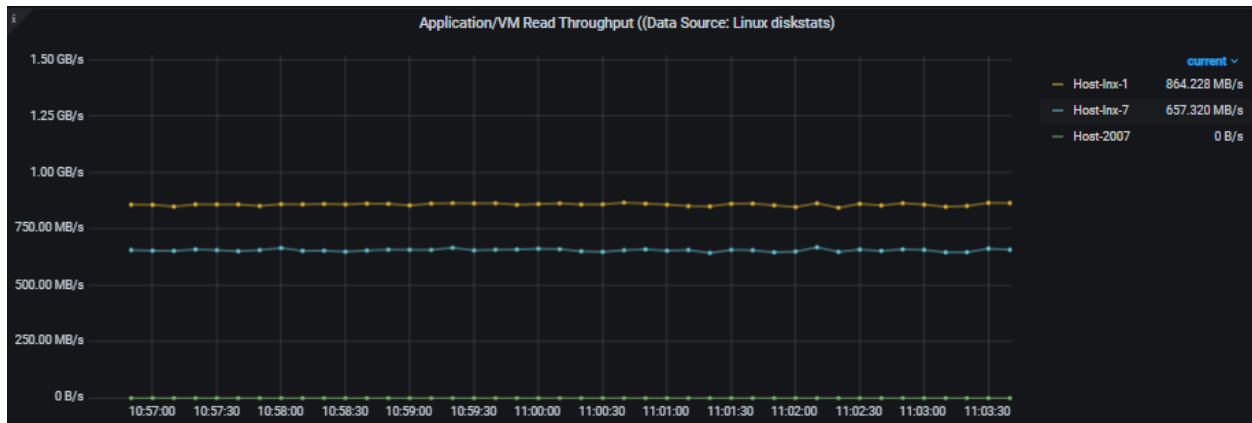
Slow drain is one of the various scenarios seen in SAN congestion. When enabling the DIRL software, we observed instantaneous effects on the congested device only, without affecting traffic for other network devices. The result was a gentle and gradual recovery towards traffic equilibrium. Analytic information was viewable on the switch CLI or other analytic platforms. DIRL was shown to be effective in this use case scenario.

The Cisco Advantage

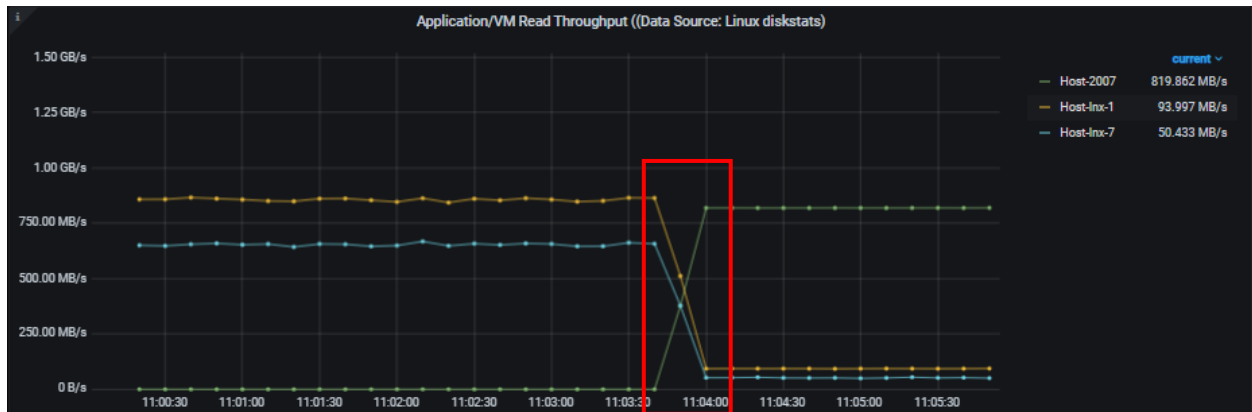
Cisco MDS DIRL software identifies and prevents SAN congestion, without traffic loss or affecting other devices. Cisco MDS DIRL SAN Congestion software is a standard feature of the Cisco MDS 9000 Series directors and switches. Simply connect the Cisco MDS 9000 Series to your existing SAN, implement the DIRL software, and eliminate network congestion. Existing Cisco MDS 9000 Series customers can update to the latest software version at no cost and implement DIRL today.

4.2 Over-utilization Congestion

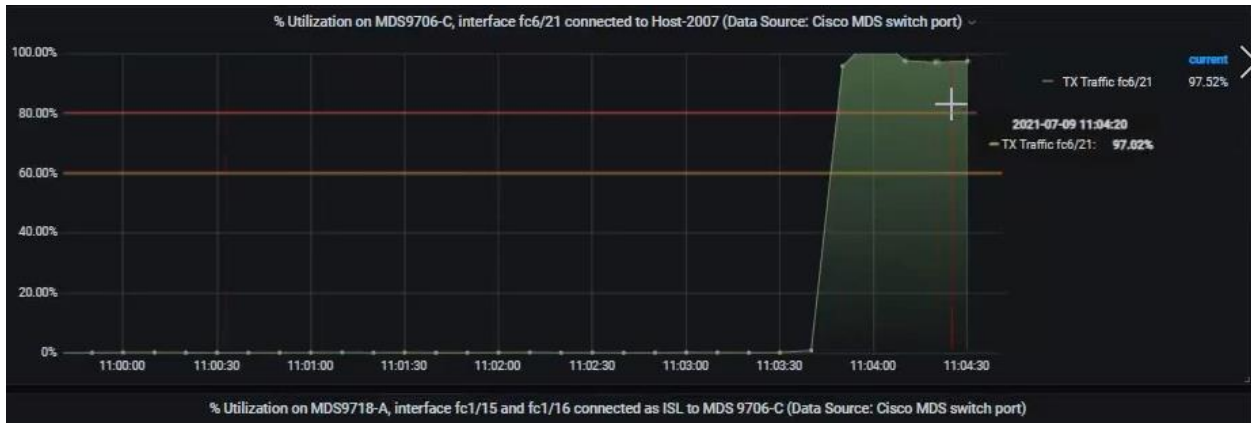
We viewed a demonstration of over-utilization congestion with the expectation of a similar result as slow-drain congestion. This was done by reducing the speed of the culprit device, Host 2007 and viewing the impact on the Host Inx-7 and Host Inx-1.



Shown above is the original throughput for the Inx-1 and Inx-7, without any SAN congestion and Host 2007 pushing zero traffic.



We observed the dip in Host Inx-7 and Host Inx-1 traffic, as well as the sharp increase in the Host 2007 culprit device. Traffic was limited to approximately 800 MBps because the link is running at 8 Gbps, and the maximum traffic was approximately 800 MBps. Therefore, the link had a physical limitation. These events occurred because there was a speed mismatch between the servers, which was connected at 8G, and the all-flash storage array, which was connected at 32G FC.



In the over-utilization scenario, we saw an increase in TxTraffic to approximately 97 percent, which represents the port utilization rather than txwait. Since the link was fully utilized, there was a backpressure.

```
port-monitor name fabricmon_edge_policy
logical-type edge
counter link-loss poll-interval 30 delta rising-threshold 5 event 4 falling-threshold 1 event 4 alerts syslog rmon portguard FPIN
counter sync-loss poll-interval 30 delta rising-threshold 5 event 4 falling-threshold 1 event 4 alerts syslog rmon portguard FPIN
counter signal-loss poll-interval 30 delta rising-threshold 5 event 4 falling-threshold 1 event 4 alerts syslog rmon portguard FPIN
counter invalid-words poll-interval 30 delta rising-threshold 1 event 4 falling-threshold 0 event 4 alerts syslog rmon portguard FPIN
counter invalid-crc poll-interval 30 delta rising-threshold 5 event 4 falling-threshold 1 event 4 alerts syslog rmon portguard FPIN
counter state-change poll-interval 60 delta rising-threshold 5 event 4 falling-threshold 0 event 4 alerts syslog rmon
counter tx-discards poll-interval 60 delta rising-threshold 200 event 4 falling-threshold 10 event 4 alerts syslog rmon
counter lr-rx poll-interval 60 delta rising-threshold 5 event 4 falling-threshold 1 event 4 alerts syslog rmon
counter timeout-discards poll-interval 60 delta rising-threshold 200 event 4 falling-threshold 10 event 4 alerts syslog rmon
counter credit-loss-reco poll-interval 1 delta rising-threshold 1 event 4 falling-threshold 0 event 4 alerts syslog rmon
counter tx-credit-not-available poll-interval 1 delta rising-threshold 10 event 4 falling-threshold 0 event 4 alerts syslog rmon
counter rx-datarate poll-interval 10 delta rising-threshold 80 event 4 falling-threshold 70 event 4 alerts syslog rmon obfl
counter tx-datarate poll-interval 10 delta rising-threshold 80 event 4 falling-threshold 70 event 4 alerts syslog rmon obfl portguard DIRL
no monitor counter err-pkt-from-port
no monitor counter err-pkt-to-xbar
no monitor counter err-pkt-from-xbar
counter tx-slowport-oper-delay poll-interval 1 absolute rising-threshold 50 event 4 falling-threshold 0 event 4 alerts syslog rmon
counter txwait poll-interval 1 delta rising-threshold 30 event 4 falling-threshold 10 event 4 alerts syslog rmon portguard DIRL
counter txwait warning-signal-threshold 40 alarm-signal-threshold 60 portguard congestion-signals
counter rx-datarate-burst poll-interval 10 delta rising-threshold 5 event 4 falling-threshold 5 event 4 alerts syslog rmon obfl datarate 90
no monitor counter tx-datarate-burst
counter input-errors poll-interval 60 delta rising-threshold 5 event 4 falling-threshold 1 event 4 alerts syslog rmon
```

Above is an example of a Port Monitor configuration with the highlighted area displaying the specific DIRL configuration for the demonstration. There were multiple policies for different possible workplace scenarios.



After enabling, we saw the instantaneous effect of DIRL by limiting the data rate followed by a gradual return of traffic stability.

fc6/21	txdatarate	rising	rate-reduction	8000	952.68	8366.04	100.0000	5.6040	Fri Jul 9 11:06:05 2021
fc6/21	-	cli recovery	full-recovery	8000	128.52	1149.37	4.2270	100.0000	Fri Jul 9 10:56:24 2021
fc6/21	txwait	recovery	rate-recovery	8000	128.52	1149.37	3.3816	4.2270	Fri Jul 9 10:34:15 2021
fc6/21	txwait	recovery	rate-recovery	8000	121.26	936.89	2.7053	3.3816	Fri Jul 9 10:33:15 2021
fc6/21	txwait	recovery	rate-recovery	8000	91.14	751.74	2.1642	2.7053	Fri Jul 9 10:32:15 2021
fc6/21	txwait	recovery	rate-recovery	8000	73.27	624.10	1.7314	2.1642	Fri Jul 9 10:31:14 2021
fc6/21	txwait	recovery	rate-recovery	8000	58.14	530.74	1.3851	1.7314	Fri Jul 9 10:30:13 2021
fc6/21	txwait	rising	rate-reduction	8000	235.47	2054.48	100.0000	1.3851	Fri Jul 9 10:29:12 2021
fc6/21	txdatarate	recovery	rate-recovery	8000	0.40	0.41	94.9202	100.0000	Wed Jun 16 19:07:31 2021
fc6/21	txdatarate	recovery	rate-recovery	8000	0.35	0.34	75.9362	94.9202	Wed Jun 16 19:06:30 2021

The internal CLI was another way that clients could view the Cisco MDS 9000 functioning via the log history information. As shown, we viewed the event when txdatarate counter was detected on the switch port and rate reduction was applied at the timestamp '11:06:05'.

Just like the slow drain scenario, we observed similar instantaneous effects for the over-utilization use case. Cisco MDS DIRL software was shown to be effective; the result was a gentle and gradual recovery towards traffic equilibrium. Analytic information was viewable on switch CLI or other analytic platforms.

The Cisco Advantage

Cisco MDS DIRL software identifies and prevents SAN congestion without traffic loss or affecting other devices. Cisco MDS DIRL San Congestion software is a standard feature of Cisco MDS 9000 Series directors and switches. Simply connect the Cisco MDS 9000 Series to your existing SAN, implement the DIRL software and eliminate network congestion. Existing Cisco MDS 9000 Series customers can update to the latest software version at no cost and implement DIRL today.

About Cisco

Cisco® (NASDAQ: CSCO) is the worldwide leader in technology that powers the Internet. Cisco inspires new possibilities by reimagining your applications, securing your data, transforming your infrastructure, and empowering your teams for a global and inclusive future. Discover more on [The Network](#) and follow us on [Twitter](#).

Cisco Systems, Inc.
300 East Tasman Dr.
San Jose, CA 95134
USA

About Miercom Performance Verified

This report was sponsored by Cisco Systems, Inc. The data was obtained completely and independently by Miercom engineers and lab-test staff as part of our Performance Verified assessment. Testing such as this is based on a methodology that is jointly co-developed with the sponsoring vendor. The test cases are designed to focus on specific claims of the sponsoring vendor, and either validate or repudiate those claims. The results are presented in a report such as this one, independently published by Miercom.

About Miercom

Miercom has published hundreds of network product analyses in leading trade periodicals and other publications. Miercom's reputation as the leading, independent product test center is undisputed.

Private test services available from Miercom include competitive product analyses, as well as individual product evaluations. Miercom features comprehensive certification and test programs, including Certified Interoperable™, Certified Reliable™, Certified Secure™ and Certified Green™.

Products may also be evaluated under the Performance Verified™ program, the industry's most thorough and trusted assessment for product usability and performance.

Use of This Report

Every effort was made to ensure the accuracy of the data contained in this report, but errors and/or oversights can occur. The information documented in this report may also rely on various test tools, the accuracy of which is beyond our control. Furthermore, the document relies on certain representations by the vendors that were reasonably verified by Miercom but beyond our control to verify to 100 percent certainty.

This document is provided “as is,” by Miercom and gives no warranty, representation or undertaking, whether express or implied, and accepts no legal responsibility, whether direct or indirect, for the accuracy, completeness, usefulness, or suitability of any information contained in this report.

All trademarks used in the document are owned by their respective owners. You agree not to use any trademark in or as the whole or part of your own trademarks in connection with any activities, products or services which are not ours, or in a manner which may be confusing, misleading, or deceptive or in a manner that disparages us or our information, projects, or developments.

By downloading, circulating, or using this report in any way you agree to Miercom’s Terms of Use. For full disclosure of Miercom’s terms, visit: <https://miercom.com/tou>.

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: www.cisco.com/go/trademarks. Third-party trademarks mentioned are the property of their respective owners.

© 2021 Miercom. All Rights reserved. No part of this publication may be reproduced, photocopied, stored on a retrieval system, or transmitted without the express written consent of the authors. Please email reviews@miercom.com for additional information.