



Detecting, Troubleshooting, and Preventing Congestion in Storage Networks

Detecting, Troubleshooting, and Preventing Congestion in Storage Networks

Paresh Gupta
Edward Mazurek

Cisco Press

Hoboken, New Jersey

Detecting, Troubleshooting, and Preventing Congestion in Storage Networks

Table of Contents

Cover

Title Page

Copyright Page

Contents at a Glance

Contents

Introduction

Chapter 1 Introduction to Congestion in Storage Networks

Types of Storage in a Data Center

Storage TypeBy Location

Local Storage

Remote Storage

Storage TypeBy Access Level

Block Storage

File Storage

Object Storage

Storage for Clustered and Distributed File Systems

SDS, HCI, and Everything Else

Storage Protocols, Transports, and Networks

Network TypeBy Framing and Encoding

Ethernet

Fibre Channel (FC)

InfiniBand (IB)

Table of Contents

Network TypeBy Use of Flow Control

Lossy Networks

Lossless Networks

Converged Ethernet Networks

Crossing the Boundaries of Network Types

Fibre Channel over Ethernet (FCoE)

RDMA over Converged Ethernet (RoCE)

Climbing Up the Networking Layers

Internet Protocol

Transmission Control Protocol (TCP)

User Datagram Protocol (UDP)

iSCSI

NVMe/TCP

NFS

SMB

HTTP

Crossing the Boundaries of Network TypesAgain

Fibre Channel over IP (FCIP)

RDMA-Capable Protocols

Storage Protocols That Use RDMA

Storage Networks

Storage Network Designs

Single-Switch Design

Edge-Core Design

Edge-Core-Edge Design

Mesh Design

Spine-Leaf Design

Terminology

Fibre Channel and FCoE Terminology

Choice of Storage

Table of Contents

Choice of Storage Network

Dedicated Versus Shared Networks for Storage Traffic

Common Questions on Storage Networks

Q: What is the difference between a network and a fabric?

Q: Whats the difference between a storage area network (SAN) and a storage network?

Q: Do storage networks have a role in the cloud?

Q: Do storage networks have a role in container storage?

Congestion in Storage Networks: An Overview

Congestion Spreading

Causes of Congestion in Storage Networks

Congestion Due to Slow End Devices

Congestion Due to Overutilization of a Link

Bit Errors on a Link

Lack of Buffers for the Distance, Frame Size, and Speed of a Link

Source of Congestion in Storage Networks

Congestion from End Devices

Congestion on ISLs

Congestion Within Switches

Common Questions About Congestion in Storage Networks

Q: What is backpressure?

Q: What are traffic burst and microburst?

Q: Isnt increasing network capacity the ultimate solution to network congestion?

Q: I was told that unlike Fibre Channel, RoCEv2 does not suffer from slow drain. Is this correct?

Q: Is slow drain the same as PFC storm?

Q: Would moving to the cloud eliminate congestion in storage networks?

Q: Would moving to HCI or SDS eliminate congestion in storage networks?

NVMe over Fabrics

Common Questions on NVMe over Fabrics

Table of Contents

Q: I have heard that NVMe supports 64K queues, each with 64K commands. How can I be ready for it?

Q: Doesnt NVMe have mechanisms to control network congestion?

Q: I built a new environment with NVMe over Fabrics, but the network throughput did not increase. Why?

Q: What effects does NVMe over Fabrics have on network congestion?

Q: Someone told me that congestion in their networks vanished after they upgraded to NVMe over Fabrics. Is that possible?

Q: Is building a dedicated network for NVMe over Fabrics best for congestion management?

Quality of Service (QoS)

Sources of Delay in a Network

Forwarding Delay

Propagation Delay

Serialization Delay

Queuing Delay

Common Questions on QoS in Storage Networks

Q: Why do network devices need buffers?

Q: What is the difference between buffers and queues?

Q: What is the difference between buffers, pause buffers, and B2B credits?

Q: Why is queue a common term in IP/Ethernet networks but not in Fibre Channel fabrics?

Q: What are some common misconceptions about using QoS in storage networks?

Q: Why is QoS not commonly used in Fibre Channel fabrics?

Q: Which is better for storage traffic in Ethernet networks: policing or shaping?

Q: What is the difference between priority and bandwidth in the context of QoS?

Summary

References

Chapter 2 Understanding Congestion in Fibre Channel Fabrics

Table of Contents

Fibre Channel Flow Control

- Initial Communication of B2B Credits
- Return of B2B Credits During Frame Flow
- B2B credit counters
- Important Details About R_RDYs and B2B Credits
- B2B Flow Control in a Multi-Hop Fabric
- B2B Flow Control in a Multi-hop Fabric Without Congestion
- B2B Flow Control in a Multi-hop Fabric with Congestion
- Buffer Overrun Situation
- Frame Rate Equalization Using B2B Flow Control

Congestion Spreading in Fibre Channel Fabrics

- Congestion Due to Slow-Drain Devices
- Congestion Due to Overutilization
- Congestion Due to Overutilization on HostEdge Links
- The Culprit Host
- Comparing Congestion Due to Slow Drain and Overutilization
- Effect on the Culprit Host
- Effect on the Culprits Port and Its Connected Switchport
- Effect on the Fabric
- Congestion in Single-Switch Fabrics
- Congestion in an ISL
- Congestion Spreading Due to Edge Devices
- Overutilization of an ISL
- Lack of B2B Credits for the Distance, Speed, and Frame Size of an ISL
- Buffering and the Ability to Absorb Congestion
- Dependency on Traffic Patterns
- Effects on Latency
- The Number of Buffers
- User Action

Frame Flow Within a Fibre Channel Switch

Table of Contents

- Frame Switching Within a Cisco MDS Switch
- Frame Switching Architecture of a Fibre Channel Switch
- Location of Buffers: Ingress, Egress, or Both
- Number of Buffers
- Preventing Head-of-Line Blocking
- Store-and-Forward Versus Cut-Through Switching
- The Ability to Detect and Drop CRC-Corrupted Frames
- Load-Balancing Schemes on ISLs
- Congestion Management Features

The Effects of Bit Errors on Congestion

- Fibre Channel Frame Format
- Fibre Channel Levels
- Data Transmission on Fibre Channel Media
- Transforming an I/O Operation to FC Frames
- Encoding the Frames and Special Functions
- Special Functions: Delimiters, Primitive Signals, and Primitive Sequences
- Transmitting Bits on the Media
- Fibre Channel Baud Rate
- Fibre Channel Bit Rate
- Fibre Channel Data Rate
- Difference Between Fibre Channel Speed and Bit Rate
- The Effects of Primitive Signals on Data Rate
- Counters on Fibre Channel Ports
- Link Initialization Counters
- Invalid Transmission Words
- CRC
- Forward Error Correction (FEC)
- Case Study: An Online Retailer
- Observations
- Conclusions

Table of Contents

Lessons Learned

Effect of Bit Errors on Congestion: Summary

B2B Credit Loss and Recovery

Loss of Tx B2B Credits Due to Bit Errors

Zero Tx B2B Credits for an Extended Duration

Credit Loss Recovery Using the B2B State Change Mechanism

Negotiation at Link Initialization

Periodic Detection and Recovery of Credit Loss

Important Details About the B2B State Change Mechanism

Credit Loss Recovery Using Link Reset Protocol

Comparison of the B2B State Change Mechanism and Link Reset Protocol

Fibre Channel Counters Summary

Summary

References

Chapter 3 Detecting Congestion in Fibre Channel Fabrics

Congestion Detection Workflow

Effects of Congestion (Congestion Severity)

Cause of Congestion

Source of Congestion (Culprits)

Spread of Congestion (Victims)

Time of Congestion Events

How to Detect Congestion

Reactive Approaches

Proactive Approaches

Predictive Approaches

Reactive, Proactive, Predictive, or All?

Where to Detect Congestion

Detecting Congestion on Network Devices

Detecting Congestion on Remote Monitoring Platforms

Table of Contents

Congestion Direction: Ingress or Egress

Egress Congestion

Ingress Congestion

Congestion Detection Metrics

Congestion Detection Metrics on Cisco MDS Switches

Tx Credit Unavailability in Microseconds: TxWait

Raw TxWait

Percentage TxWait

TxWait History Graphs

TxWait History in the OBFL Buffer

Rx Credit Unavailability in Microseconds: RxWait

Continuous Tx Credit Unavailability in Milliseconds: Slowport-monitor

Slowport-monitor Events in Real Time

Slowport-monitor History in OBFL

Continuous Tx Credit Unavailability for 100 ms: Tx-credit-not- available

Tx-credit-not-available in Real Time

Tx-credit-not-available History in the OBFL Buffer

Differences Between TxWait, Slowport-monitor, and Tx-credit-not-available

When to Enable Slowport-monitor?

Continuous Rx Credit Unavailability for 100 ms: Rx-credit-not-available

Timeout Discards and Timeout-Drops

Tx Credit Loss Recovery

Link Failure: Link Reset Failed Nonempty Recv Queue (LR Rcvd B2B)

Credits and Remaining Credits

Credit Transition to Zero

Link Utilization

Tx-datarate

Tx-datarate-burst

Rx-datarate

Table of Contents

Rx-datarate-burst

Bit Errors

Automatic Alerting

Port-Monitor on Cisco MDS Switches

Port-Monitor Policy Types

Port-Monitor Policy Parameters

Port-Monitor Counters

Detecting Congestion Using Remote Monitoring Platforms

NDFC Congestion/Slow-Drain Analysis

The MDS Traffic Monitoring (MTM) App

MTM Architecture

MTM Use Cases

Metric Export Mechanisms

Parsing the Command-Line Output over SSH

Simple Network Management Protocol (SNMP)

Application Programming Interfaces (APIs)

Streaming Telemetry

Recommendations

The Pitfalls of Monitoring Network Traffic

Percentage Utilization of Fibre Channel Ports

Average and Peak Utilization

Detecting Congestion Due to Slow Drain and Overutilization

Slow Drain and Overutilization at the Same Time

Detecting Congestion on long-distance links

Summary

References

Chapter 4 Troubleshooting Congestion in Fibre Channel Fabrics

Table of Contents

Troubleshooting Methodology and Workflow

Congestion Severities and Levels

Mild Congestion (Level 1 and Level 1.5)

Moderate Congestion (Level 2)

Severe Congestion (Level 3)

Goals of Troubleshooting

Identifying the Source (Culprits) and Cause of Congestion

Identifying the Affected Devices (Victims)

Methodology

Step 1: Troubleshooting Congestion in Decreasing Severity Levels

Step 2: Chasing the Source of Congestion (Culprit)

Hints and Tips for Troubleshooting Congestion

Investigating Higher Congestion Levels First

Finding Level 3 Congestion: Credit Loss

Finding Level 2 Congestion: Frame Drops

Finding Level 1/1.5 Congestion: TxWait and Overutilization

Using the show tech-support slowdrain Command

Synchronizing Clocks and Considering Timing

Timeout-Drop Anomaly

Enabling and Using Automatic Alerting

Using a Remote Monitoring Platform (NDFC/DCNM)

Cisco MDS NX-OS Commands for Troubleshooting Congestion

The show interface Command

The show interface counters [detailed] Command

The show interface txwait-history and rxwait-history Commands

The OBFL Commands: show logging onboard

TxWait

RxWait

Error Statistics

Flow Congestion Drops

Table of Contents

Generic Troubleshooting Commands

The show topology Command

The show flogi database Command

The show fcns database Command

The show zone member Command

The show zone name Command

The show zoneset active Command

The show fcs le Command

The show fcdomain Command

The show fspf database Command

The show rdp Command

The show fdmi database Command

System Messages: show logging log

Link failure Link Reset failed nonempty recv queue System Message

Link failure Link reset failed due to timeout System Message

TCP conn. closed - retransmit failure System Message

Case Study 1: Finding Congestion Culprits and Victims in a Single-Switch Fabric

Fabric A Analysis

Loss of Information Due to Clearing the OBFL Counters

TxWait Analysis

Traffic Utilization (Tx-datarate) Analysis

Graphical Correlation of Congestion Symptoms

Fabric B Analysis

Culprit Analysis

Victim Analysis

Direct Victims

Same-Path Victims

Indirect Victims

Case Study 1 Summary

Table of Contents

Case Study 2: Credit Loss Recovery Causing Frame Drops

Initial Investigation

Fabric A Analysis

Edge Switch Fab_A_MDS_9396T_14

Core Switch Fab_A_MDS_9718_01

Core Switch Fab_A_MDS_9718_02

Fabric A Conclusion

Fabric B Analysis

Edge Switch Fab_B MDS_9396T_14

Core Switch Fab_B MDS_9718_01

Core Switch Fab_B MDS_9718_02

Fabric B Conclusion

Culprit Analysis

Victim Analysis

Direct Victims

Same-Path Victims

Indirect Victims

Case Study 2 Summary

Case Study 3: Overutilization on a Single Device Causing Massive Congestion Problems

Level 3

Level 2

MDS_9513_03

MDS_9710_03

MDS_9710_01

MDS_9513_01

Culprit Analysis

Victim Analysis

Direct Victims

Table of Contents

Same-Path Victims

Indirect Victims

Case Study 3 Summary

Case Study 4: Long-Distance ISLs Causing Congestion

Level 3

Level 2

Level 1.5

MDS_9148S_01

MDS_9148S_02

MDS_9148S_03

Culprit Analysis

Victim Analysis

Case Study 4 Summary

Summary

References

Chapter 5 Solving Congestion with Storage I/O Performance Monitoring

Why Monitor Storage I/O Performance?

How and Where to Monitor Storage I/O Performance

Storage I/O Performance Monitoring in the Host

Storage I/O Performance Monitoring in a Storage Array

Storage I/O Performance Monitoring in a Network

Cisco SAN Analytics Architecture

Traffic Inspection

Metric Calculation

Metric Export

Understanding I/O Flows in a Storage Network

I/O Flows in Fibre Channel Fabrics

Table of Contents

I/O Flows Versus I/O Operations

I/O Flow Metrics

Latency Metrics

Exchange Completion Time

Data Access Latency

Host Response Latency

Using Latency Metrics

The Location for Measuring Latency Metrics

Performance Metrics

I/O Operations per Second (IOPS)

I/O Size

Throughput

Outstanding I/O

I/O Operations and Network Traffic Patterns

Read I/O Operation in a Fibre Channel Fabric

Write I/O Operation in a Fibre Channel Fabric

Network Traffic Direction

Network Traffic Throughput

Correlating I/O Operations, Traffic Patterns, and Network Congestion

Case Study 1: A Trading Company That Predicted Congestion Issues Using SAN Analytics

Background

Initial Investigation: Finding the Cause and Source of Congestion

A Better Host Upgrade Plan

Case Study 1 Summary

Case Study 2: A University That Avoided Congestion Issues by Correcting Multipathing Misconfiguration

Background

Investigation

Case Study 2 Summary

Table of Contents

Case Study 3: An Energy Company That Eliminated Congestion Issues

Background

Investigation

Case Study 3 Summary

Case Study 4: A Bank That Eliminated Congestion Through Infrastructure
Optimization

Background

Investigation

Case Study 4 Summary

Summary

References

Chapter 6 Preventing Congestion in Fibre Channel Fabrics

An Overview of Eliminating or Reducing Congestion

Defining the Outcome of an Approach

Manual Versus Automatic Approaches

Link Capacity

Congestion Recovery by Disconnecting the Culprit Device

Considerations for Disconnecting a Culprit

How to Disconnect?

Congestion Recovery by Dropping Frames

Dropping Frames Based on Their Age in the Switch

Configuring Congestion-Drop Timeout on Cisco MDS Switches

Details on Congestion-Drop Timeout

Dropping Frames Based on Slow Drain on an Edge Port

Enabling No-Credit-Drop Timeout on Cisco MDS Switches

Details on No-Credit-Drop Timeout

No-Credit-Drop Timeout in Action

Finding the Optimum No-Credit-Drop Timeout Value

Traffic Segregation

Table of Contents

Categorizing Traffic for Segregation

Traffic Segregation to Dedicated ISLs

Using VSANs for Traffic Segregation on Dedicated ISLs

Considerations for Traffic Segregation to Dedicated ISLs Using Multiple VSANs

Case Study 1: A Bank That Avoided Congestion with Traffic Segregation

Background and Investigation

Solution: Traffic Segregation to Dedicated ISLs

Case Study 1 Summary

Traffic Segregation Using Virtual Links

Understanding Virtual Links

Flow Control in a Virtual Link

Congestion Segregation Using Virtual Links

Scope of Congestion Segregation Using Virtual Links

Extending Virtual Links to the End Devices

Enabling Virtual Links on ISLs on Cisco MDS Switches

Traffic Assignment to Virtual Links

Automatic Assignment of Traffic to Virtual Links: Congestion Isolation

Manual Assignment of Traffic to Virtual Links

Comparing No-Credit-Drop Timeout with Congestion Isolation

No-Credit-Drop Timeout and Congestion Isolation in Action

Too Many VLs: The Hidden Side Effects

Traffic Segregation Considerations

Comparing Traffic Segregation Using VSANs and Virtual Links

Congestion Segregation Using Virtual Links: Caution

Congestion Prevention Using Rate Limiters on Storage Arrays

Congestion Prevention Using Dynamic Ingress Rate Limiting on Switches

How DURL Prevents Congestion

How DURL Prevents Congestion Due to Overutilization

Table of Contents

How DIRL Prevents Congestion Due to Slow Drain

Details of DIRL

Benefits of DIRL

Enabling and Using DIRL on Cisco MDS Switches

Enable FPM

Configure Port-Monitor

DIRL in Action

Test Setup

Scenario 1: Congestion Due to Slow Drain Without Spreading

Scenario 2: Congestion Due to Slow Drain with Spreading

Scenario 3: Preventing Congestion Due to Slow Drain Using DIRL

Scenario 4: Preventing Congestion Due to Overutilization Using DIRL

Comparing DIRL with Other Approaches

DIRL Versus No-Credit-Drop Timeout

DIRL Versus Traffic Segregation Using Virtual Links

Preventing Congestion by Notifying the End Devices

Readiness of Notifications and Signals in Fibre Channel

Notifications and Signals in Fibre Channel Fabrics

Register Diagnostic Functions

Exchange Diagnostic Capabilities

Fabric Performance Impact Notification (FPIN)

Congestion Signals

Examples of RDF, EDC, FPIN, and Congestion Signals

Comparing FPIN Frames and Congestion Signals

The Possible Results of FPIN Frames and Signals

Configuring Sending of FPIN Frames and Congestion Signals on Cisco MDS
Switches

Using DIRL Versus Notifying the End Devices for Congestion Prevention

Network Design Considerations

Table of Contents

- Lowering the Link Speed of Storage Ports
- Edge-Core-Edge or Edge-Core or Collapsed-Core Design
- Increased Traffic Localization to a Single Switch
- Splitting Large Fabrics into Smaller Islands

Summary

References

Chapter 7 Congestion Management in Ethernet Storage Networks

Ethernet Flow Control

- How Ethernet Flow Control Works
- Pause Time
- When Are Pause Frames Sent?
- Ingress and Egress Queues
- Location of Ingress No-Drop Queues
- Number of Ingress No-Drop Queues per Port
- Implementation Differences and the Scope of This Book
- Pause Threshold and Resume Threshold
- Ethernet Pause Frames Compared with Fibre Channel B2B Credits
- Priority Flow Control
- Mapping Traffic Classes to the Pause Frame Class Enable Vector Field
- Layer 2 Priority Flow Control
- Layer 3 Priority Flow Control
- Converged Ethernet Networks
- Configuring Lossless Ethernet
- Dedicated and Converged Ethernet Network

Understanding Congestion in Lossless Ethernet Networks

- Slow Drain in Lossless Ethernet Networks
- Overutilization of a Link in Lossless Ethernet Networks
- Bit Errors

Table of Contents

Congestion Spreading in a Single-Switch Lossless Ethernet Network

Congestion Spreading in an EdgeCore Lossless Ethernet Network

Congestion Spreading in a Lossless SpineLeaf Network

Slow Drain in a Lossless Ethernet SpineLeaf Network

Overutilization of a HostEdge Link in a Lossless Ethernet SpineLeaf Network

Comparing Congestion Due to Slow Drain and Overutilization in a Lossless Ethernet SpineLeaf Network

Detecting Congestion in Lossless Ethernet Networks

Congestion Direction: Ingress or Egress

Congestion Detection Metrics

Duration of Traffic Pause: TxWait and RxWait

The Number of Pause Frames

Frame Drops or Discards

Bit Errors

Link Utilization

PFC Storms

Storage I/O Performance Monitoring

UDP Flow Monitoring Versus I/O Flow Monitoring

Unavailability of I/O Flow Monitoring in Lossless Ethernet Networks

Alternative Approaches

FCoE I/O Operations

RoCE I/O Operations

Correlating I/O Operations, Traffic Patterns, and Network Congestion

Detecting Congestion on a Remote Monitoring Platform

Congestion Detection Using Cisco Nexus Dashboard Insights

Metric Export Mechanisms

Troubleshooting Congestion in Lossless Ethernet Networks

Goals

Congestion Severities and Levels

Methodology

Table of Contents

Troubleshooting Congestion in SpineLeaf Topology

Reality Check

Troubleshooting Congestion by Using a Remote Monitoring Platform

Comparative Analysis

Trends and Seasonality

Monitoring a Slow-Drain Suspect

Monitoring an Overutilization Suspect

FC and FCoE in the Same Network

Congestion Spreading Due to Slow Drain

Congestion Spreading Due to Overutilization

Bit Rate Differences Between FC and FCoE

Multiple No-Drop Classes on the Same Link

Bandwidth Allocation Between Lossless and Lossy Traffic

The Effect of Lossy Traffic on the No-Drop Class

Case Study 1: An Online Gaming Company

Case Study 2: Converged Versus Dedicated Storage Network

Preventing Congestion in Lossless Ethernet Networks

Eliminating or Reducing Congestion: An Overview

Congestion Recovery by Dropping Frames

Dropping Frames Based on Their Age in the Switch

Dropping Frames Based on Slow Drain on an Edge Port

Congestion Notification in Routed Lossless Ethernet Networks

Solution Components

RoCEv2 Transport Overview

RoCEv2 Congestion Management

RoCEv2 Congestion Management Considerations

PFC and ECN

Lossless Traffic with VXLAN

VXLAN Overview

Table of Contents

VXLAN Transport

Physical Topology

MAC Address Learning

Lossless Traffic over VXLAN

VXLAN Encapsulation

VXLAN Decapsulation

Congestion Notification over VXLAN

Flow Control and Congestion Notification with VXLAN

Congestion Management in VXLAN

Summary

References

Chapter 8 Congestion Management in TCP Storage Networks

Understanding Congestion in TCP Storage Networks

Comparison with Lossless Networks

How iSCSI and NVMe/TCP Exchange Data

Bit Errors in Lossy Ethernet Networks with TCP Transport

How TCP Provides Reliable Data Transfer

TCP Flow Control

TCP Congestion Control

Congestion in TCP Storage Networks

Congestion Due to Overutilization of the Host Link

Congestion Within the Host

Storage I/O Performance Monitoring

TCP Flow Monitoring Versus I/O Flow Monitoring

Unavailability of I/O Flow Monitoring in TCP Storage Networks

Alternative Approaches

iSCSI I/O Operations

NVMe/TCP I/O Operations

Correlating I/O Operations, Traffic Patterns, and Network Congestion

Table of Contents

Comparison with Lossless Networks

Estimating I/O Flow Performance from TCP Flow Performance

IP MTU and TCP MSS Considerations

The Number of Packets for an I/O Operation

Packet Fragmentation

Comparison with Lossless Networks

Preventing Congestion in TCP Storage Networks

Eliminating or Reducing Congestion: An Overview

Congestion Notification in TCP Storage Networks

Solution Components

Explicit Congestion Notification in TCP/IP Networks

Comparison with RoCEv2 Networks

Comparison with Fibre Channel Fabrics

ECN Considerations for Block-Storage Traffic

Switch Buffer Management

Queue Utilization

Queue Utilization Considerations

User Actions

Comparison with Lossless Ethernet

Comparison with Fibre Channel Fabrics

Active Queue Management

Tail Drop

Random Early Detect (RED)

Weighted Random Early Detection (WRED)

Approximate Fair Dropping (AFD)

Dynamic Packet Prioritization (DPP)

Detecting Congestion in TCP Storage Networks

Source of Congestion Within the End Devices

Congestion Detection Notes

Table of Contents

Comparison with Lossless Networks

The Source of Congestion Within the Network

Packet Drops or Discards

ECN Counters

Link Utilization

Queue Depth Monitoring and Microburst Detection

Bit Errors

Detecting Congestion Using a Remote Monitoring Platform

Comparative Analysis

Trends and Seasonality

Congestion Detection Using Cisco Nexus Dashboard Insights

Metric Export Mechanisms

Troubleshooting Congestion in TCP Storage Networks

Goals

Congestion Severities and Levels

Methodology

Load Balancing in TCP Storage Networks

QoS Considerations for Dedicated and Shared Storage Networks

The Effect of Other Traffic Classes on Storage Traffic Class

Configuring Versus Operating a Shared Storage Network

QoS Expertise

FCoE, RoCE, iSCSI, and NVMe/TCP in the Same Network

iSCSI and NVMe/TCP in a Lossless Network

iSCSI and NVMe/TCP with VXLAN

Fibre Channel over TCP/IP (FCIP)

TCP Optimizations for Storage Traffic on Cisco FCIP Switches

Detecting Congestion on FCIP Links

Modified TCP Implementations

Summary

Table of Contents

References

Chapter 9 Congestion Management in Cisco UCS Servers

Cisco UCS Architecture

UCS Domain

Traffic Flow in a UCS Domain

Flow Control in a UCS Domain

Understanding Congestion in a UCS Domain

Detecting Congestion in a UCS Domain

Ingress Congestion

Egress Congestion

Congestion Between FI Server Ports and IOM/FEX Fabric Ports

UCS Congestion Detection Notes

The UCS Traffic Monitoring (UTM) App

The Journey of UTM

Getting Started with UTM

UTM Architecture

An Overview of Using UTM

Troubleshooting Congestion Using UTM

Congestion Troubleshooting Workflow in UTM

Proactively Detecting Congestion Due to Slow Drain

Proactively Detecting Congestion Due to Overutilization

Case Study 1: Finding the Cause and Source of Congestion in a UCS Domain

Background

Investigation

Conclusion

Solution

Case Study 1 Summary

Case Study 2: Congestion Due to Slow Drain on the Backplane Port

Investigation

Table of Contents

Conclusions

Case Study 2 Summary

Case Study 3: Non-Uniform Utilization of FI Uplink Ports

Investigation

Conclusion

Solution

Case Study 3 Summary

Case Study 4: Congestion Due to Multipathing I/O Imbalance

Investigation

Conclusion

Solution

Case Study 4 Summary

Summary

References

Index