

White paper

Come evitare la diffusione della congestione

In questo documento vengono descritti il possibile impatto della diffusione della congestione (nota anche come rallentamento del flusso) sulla SAN (Storage Area Network), le metriche utilizzate per descrivere la severity di ciascun tipo di congestione per Connectrix B-Series e per la serie MDS nonché le misure preventive da mettere in atto per evitare gli effetti della diffusione della congestione

Maggio 2019

Come evitare la diffusione della congestione | H17762.2 |

Revisioni

Data	Descrizione
Maggio 2019	Release iniziale

Ringraziamenti

Questo documento è stato prodotto dai seguenti membri del team Storage Engineering di Dell EMC: Autori:

Alan Rajapa Erik Smith

Le informazioni contenute nella presente documentazione sono fornite "come sono". Dell Inc. non fornisce alcuna dichiarazione o garanzia in relazione alle informazioni contenute nel presente documento, in modo specifico per quanto attiene alle garanzie di commerciabilità o idoneità per uno scopo specifico.

L'utilizzo, la copia e la distribuzione dei prodotti software descritti in questo documento richiedono una licenza d'uso valida per ciascun software.

©Pubblicato a maggio 2019: Dell Inc. o sue società controllate. Tutti i diritti riservati. Dell, EMC, Dell EMC e gli altri marchi sono marchi di Dell Inc. O di sue società controllate. Altri marchi possono essere marchi dei rispettivi proprietari.

Dell ritiene che le informazioni contenute nel presente documento siano esatte alla data di pubblicazione. Le informazioni sono soggette a modifiche senza preavviso.

Sommario

Prefazi	one	.4
Panora	mica	. 6
PREREQ	UISITI	. 6
Che co	s'è la diffusione della congestione?	. 8
Diffusio	ne della congestione dovuta a sottoscrizione in eccesso	11
Baselin	e dell'applicazione	12
Genera	zione di grafici dei profili di base dell'applicazione	12
4.1.1	Brocade	18
4.1.2	Cisco	19
AVVISI D	I DIFFUSIONE DELLA CONGESTIONE DI UNISPHERE	20
CONCLU	SIONI	25
Risoluz	ione	26
PREVEN	ZIONE	26
Rappor	to larghezza di banda	26
Implem	entazione dei limiti di larghezza di banda	27
Append	lice	29
BILITAZ	IONE DEL MONITORAGGIO DELLE PRESTAZIONI	29
ΙΟΝΙΤΟΙ	RAGGIO DELLA DIFFUSIONE DELLA CONGESTIONE IN CONNECTRIX	31
6.1.1	Brocade	31
6.1.2	Cisco	36
6.1.3	Dell EMC	44
6.1.4	Brocade	44
6.1.5	Cisco	45
6.1.6	Sottoscrizione in eccesso	46
6.1.7	Brocade	47
6.1.8	Cisco	48
	Prefazio Panora PREREQ Che con Diffusio Baselin Genera 4.1.1 4.1.2 AVVISI D CONCLU Risoluz PREVEN Rappor Implem Append ABILITAZ MONITOF 6.1.1 6.1.2 6.1.3 6.1.4 6.1.5 6.1.6 6.1.7 6.1.8	Prefazione Panoramica PREREQUISITI Che cos'è la diffusione della congestione? Diffusione della congestione dovuta a sottoscrizione in eccesso Baseline dell'applicazione Baseline dell'applicazione Generazione di grafici dei profili di base dell'applicazione 4.1.1 Brocade 4.1.2 A.1.1 Brocade 4.1.2 A.1.2 Cisco AVVISI DI DIFFUSIONE DELLA CONGESTIONE DI UNISPHERE CONCLUSIONI Risoluzione PREVENZIONE PREVENZIONE Rapporto larghezza di banda Appendice Appendice ABILITAZIONE DEL MONITORAGGIO DELLE PRESTAZIONI. MONITORAGGIO DELLA DIFFUSIONE DELLA CONGESTIONE IN CONNECTRIX. 6.1.1 Brocade 6.1.2 Cisco 6.1.2 Cisco 6.1.4 Brocade 6.1.4 Brocade 6.1.5 Cisco 6.1.5 Cisco 6.1.6 Sottoscrizione in eccesso. 6.1.7 Brocade 6.1.8 Cisco

1	Prefazione	
		In questo documento vengono descritti il possibile impatto della diffusione della congestione (nota anche come rallentamento del flusso) sulla SAN (Storage Area Network), le metriche utilizzate per descrivere la severity di ciascun tipo di congestione per Connectrix B-Series e per la serie MDS nonché le misure preventive da mettere in atto per evitare gli effetti della diffusione della congestione.
		Nell'ambito di un impegno volto a migliorare e ottimizzare le prestazioni e le funzionalità della propria linea di prodotti, Dell EMC pubblica periodicamente revisioni del proprio hardware e software. Pertanto, alcune delle funzioni descritte in questo documento potrebbero non essere supportate da tutte le versioni del software o dell'hardware attualmente in uso. Per informazioni più aggiornate sulle funzionalità del prodotto, fare riferimento alle note di rilascio del prodotto.
		Nel caso in cui un prodotto non funzioni in modo corretto o secondo quanto descritto nella documentazione, rivolgersi al proprio referente Dell EMC.
Audience		Questo TechBook è destinato al field Dell EMC, inclusi i consulenti tecnologici e gli Storage Architect, gli amministratori e gli operatori che si occupano dell'acquisto, della gestione, del funzionamento o della progettazione di un ambiente di network storage contenente dispositivi EMC e host.
Documenta correlata	azione	Tutta la documentazione correlata e le note di rilascio sono disponibili all'indirizzo <u>https://dell.com/support</u> . Fare clic su Support by Product , immettere il nome del prodotto e fare clic su Documentation .
Dell EMC S Matrix ed E Navigaton Interoperat	upport -Lab bility	Per le informazioni più aggiornate, consultare sempre la <i>Dell EMC Support Matrix</i> , disponibile tramite E-Lab Interoperability Navigator (ELN) all'indirizzo <u>https://www.dellemc.com/en-us/products/interoperability/elab.htm#tab0=2</u>
Dove otten assistenza	ere	Informazioni sul supporto, sui prodotti e sulle licenze Dell EMC sono disponibili sul sito Web di supporto online Dell EMC come indicato di seguito.
		Nota: Per aprire una Service Request nel sito web di supporto online Dell EMC, è necessario disporre di un contratto di supporto valido. Per informazioni dettagliate su come ottenere un contratto di supporto valido o per altre domande sul proprio account, contattare il proprio responsabile vendite Dell EMC.
Informazion sui prodott	ni i	Per ottenere la documentazione, le note di rilascio, gli aggiornamenti software o le informazioni sui prodotti, le licenze e i servizi Dell EMC, accedere al sito web di supporto online Dell EMC (registrazione obbligatoria) all'indirizzo: <u>https://www.dell.com/support</u>
Supporto to	ecnico	Dell EMC offre una vasta gamma di opzioni di supporto.
Supporto per prodott	:o —	Dell EMC offre informazioni consolidate e specifiche per prodotto sul sito web all'indirizzo: https://support.dell.com/products Le pagine web di supporto per prodotto offrono link rapidi a documentazione, white paper, strumenti di consultazione (quali gli articoli più consultati della knowledgebase) e download, oltre ad altri contenuti dinamici quali presentazioni, discussioni, argomenti del forum del Customer Service e un link a Dell EMC Live Chat.

Dell EMC Live Chat e supporto eLicensing Permette di avviare una chat o una sessione di messaggistica istantanea con un tecnico del supporto Dell EMC. Per attivare le proprie autorizzazioni e ottenere i file di licenza, visitare il centro di assistenza

Per attivare le proprie autorizzazioni e ottenere i file di licenza, visitare il centro di assistenza all'indirizzo <u>https://dell.com/support</u>, come indicato nella lettera License Authorization Code (LAC) inviata tramite e-mail.

2 Panoramica

L'obiettivo di questo white paper è:

- 1. Descrivere in che modo la diffusione della congestione (nota anche come rallentamento del flusso) può influire sulla SAN (Storage Area Network);
- 2. Definire le metriche utilizzate per descrivere ciascuna severity e il tipo di congestione per i prodotti Connectrix B-Series e della serie MDS;
- 3. Descrivere le misure preventive che possono essere messe in atto per evitare gli effetti della diffusione della congestione; e
- 4. Dimostrare come utilizzare le informazioni di cui sopra per rilevare, prevenire e correggere la diffusione della congestione a causa della sottoscrizione in eccesso.

PREREQUISITI

Tenere presente quanto segue:

Ai fini del presente documento si presuppone che siano in uso le seguenti versioni del software. I passaggi possono differire nelle versioni precedenti. Fare riferimento all'Appendice per i dettagli che descrivono come abilitare le funzionalità richieste.

 Dell EMC Unisphere for PowerMax e VMAX è installato e in esecuzione e l'array è stato registrato per raccogliere i dati sulle prestazioni. https://www.dell.com/support/products/27045 Unisphere-for-

https://www.dell.com/support/products/27045 Unisphere-for /Documentation/?source=promotion

- 2. Le GUI di SAN Management sono installate.
 - a. Per Brocade Fabrics: Connectrix Manager Data Center Edition (CMCNE) 14.x o versione successiva Download: <u>https://www.dell.com/search/?text=CMCNE%2014&searchLang=en_US&facetResource=DOWN</u> Guida per l'amministratore: <u>https://www.dell.com/search/?text=CMCNE%2014% 20admin%20guide&searchLang=en_US</u>
 - b. Per Cisco Fabrics: Cisco Data Center Network Manager(DCNM) 10.x o versione successiva Download: https://www.dell.com/support/search/?text=DCNM%2010&facetResource=DOWN

Guida per l'amministratore:

https://www.cisco.com/c/en/us/support/cloud-systems-management/prime-data-centernetwork-manager/products-installation-guides-list.html

- 3. Il firmware switch SAN deve essere il seguente:
 - a. Brocade: OS Fabric 7.4.1d o versione successiva **Download:** <u>https://www.dell.com/support/search/?text=Brocade%20FOS%20download&searchLang=en_US&facetResource=DOWN</u>
 - b. Cisco: NX-OS 6.2(13) o versione successiva
 Download: https://www.dell.com/support/search/?text=NX-OS%20download

- 4. Vengono installate tutte le licenze di monitoraggio delle prestazioni necessarie.
 - a. Brocade richiede una licenza MAPS: https://docs.broadcom.com/docs/53-1005239-04
 - b. Cisco richiede la licenza del pacchetto DCNM-SAN Server: <u>https://www.cisco.com/c/en/us/support/cloud-systems-management/prime-data-center-network-manager/products-installation-guides-list.html</u>
 - c. PowerMAX e VMAX richiedono una eLicense Dell EMC Unisphere. Fare riferimento alla pagina 21 del seguente PDF per maggiori informazioni: <u>https://www.dell.com/collateral/TechnicalDocument/docu88904.pdf</u>

3 Che cos'è la diffusione della congestione?

Il trasporto dei dati da e verso un array di storage richiede che tutti i dati giungano a destinazione in modo tempestivo. In particolare, si tratta di un requisito essenziale per i protocolli di storage basati su block che utilizzano SCSI (ad esempio Fibre Channel - FCP). Anche se i motivi esatti non rientrano nell'ambito di questo white paper, è possibile trovare maggiori informazioni nella sezione "Congestion and Backpressure" del TechBook *Networked Storage Concepts and Protocols*:

(https://www.dellemc.com/en-us/products/interoperability/elab.htm#tab0=1hardware/technical-documentation/h4331networked-storage-cncpts-prtcls-sol-gde.pdf).

Come qualsiasi altro protocollo di rete, Fibre Channel (FC) deve garantire questa consegna tempestiva dei dati in un'ampia gamma di situazioni comuni di congestione della rete. Il meccanismo utilizzato da FC si concentra sulla prevenzione della perdita di frame utilizzando il controllo del flusso buffer-to-buffer. Per questo motivo, FC è considerato un "protocollo lossless".

Sebbene i meccanismi di controllo del flusso utilizzati da ciascun protocollo siano leggermente diversi, FC e altri protocolli lossless (ad es., DCB Ethernet e Infiniband) impediscono l'overflow del buffer a entrambe le estremità di un link, consentendo al trasmettitore di determinare quando il ricevitore all'altra estremità del link sta per esaurire la capacità. Una volta effettuata questa determinazione, una porta interromperà la trasmissione dei dati fino a quando l'altra estremità del link non indicherà di essere pronta a ricevere ulteriori dati. Mentre un trasmettitore è in questo stato, non è in grado di trasmettere frame e diciamo che è congestionato. Se un trasmettitore rimane in stato di congestione per un periodo di tempo sufficientemente lungo, questa congestione può propagarsi all'indietro verso la fonte. Questo fenomeno è noto come diffusione della congestione e un esempio è mostrato nella seguente sequenza di diagrammi.

La Figura 1 è un esempio di rete SAN non soggetta a congestione. Sia Host 1 che Host 2 stanno eseguendo i comandi READ sull'array.

Poiché sia l'array che l'host sono collegati a 16 Gbps e la larghezza di banda ISL è sufficiente (ad es., 32G), non vi è alcuna congestione nella rete SAN.



Figura 1 Nessuna congestione

La Figura 2 mostra un esempio di rete SAN con diffusione della congestione dovuta a sottoscrizione in eccesso. Si noti che l'unica differenza tra le due figure consiste nel fatto che nella Figura 3 l'interfaccia su Host 1 è stata impostata per l'esecuzione a 4 Gbps anziché a 16 Gbps. Non appena effettuata questa operazione, se l'interfaccia dell'array trasmette i dati a una velocità superiore alla velocità dell'HBA collegato (ad es., 4G), Host 1 non sarà in grado di ricevere i dati alla velocità di trasmissione, e la conseguenza immediata sarà l'accodamento dei frame. Man mano che Queue 1 si riempie, la congestione si diffonde all'origine dei dati. Poiché sia Host 1 che Host 2 condividono lo stesso Inter Switch Link (ISL), questa congestione influisce sul "flusso innocente" tra Host 2 e Storage 2, riducendo il throughput effettivo da 16 Gbps a 4 Gbps.



Figura 2 Congestione

Ulteriori informazioni sulla congestione e sulla diffusione della congestione sono disponibili nella sezione Congestion and Backpressure del TechBook *Networked Storage Concepts and Protocols*: (https://www.dell.com/collateral/hardware/technical-documentation/h4331-networked-storage-cncpts-prtcls-sol-gde.pdf). È importante notare che la sottoscrizione in eccesso è solo una delle potenziali cause di diffusione della congestione. Altre cause verranno illustrate nelle sezioni seguenti.

• Rapporto di congestione (Rapporto C)

Il rapporto di congestione, o rapporto C, è un valore calcolato che può aiutare a rilevare il verificarsi della diffusione della congestione. Ad esempio, la Figura 3 mostra un host (nel caso specifico, Host 1) in grado di ricevere i dati a una velocità di 4 Gbps ma che sta ricevendo i dati da un'interfaccia di storage in grado di trasmettere i dati a 16 Gbps.



Figura 3 Rapporto di congestione

Che cos'è la diffusione della congestione?

Una delle cose che rende particolarmente difficile rilevare e risolvere questi problemi è il fatto che, dal punto di vista dell'interfaccia 4G sullo switch 1, non vi è alcun problema. L'interfaccia dello switch trasmette i frame alla massima velocità consentita dal link. Tuttavia, poiché lo storage sta trasmettendo i dati alla velocità consentita dal suo link (ovvero 16 Gbps), ci saranno 12 Gbps (16 Gbps - 4 Gbps) di larghezza di banda che verranno trasmessi dall'array e che dovranno essere accodati da qualche parte. Questo accodamento si verifica in genere nel fabric ed è la causa della diffusione della congestione. Come accennato in precedenza, un metodo che può essere utilizzato per rilevare la presenza della diffusione della congestione è quello di calcolare il rapporto di congestione. A tale scopo, è sufficiente dividere il valore del contatore "Time Spent at zero transmit credit" e dividerlo per il contatore Frames Transmitted per ottenere un numero generalmente compreso tra 0 e 1. Se questo numero è maggiore di 0,2, si ha una congestione. A proposito, questo numero deve essere calcolato per ogni singola interfaccia, quindi è probabilmente meglio creare uno script di processo per il controllo di questo valore.

4 Diffusione della congestione dovuta a sottoscrizione in eccesso

Il seguente caso di studio si basa sulla diffusione della congestione dovuta a sottoscrizione in eccesso. La topologia per questo caso di studio è mostrata nella Figura 4 di seguito. In questo caso di studio verranno illustrati gli strumenti e le tecniche attualmente disponibili per rilevare e prevenire il verificarsi di questo problema.

Nota: la diffusione della congestione è un problema estremamente difficile da rilevare e risolvere. Ciò è dovuto principalmente all'incapacità dell'attuale generazione di strumenti di gestione di offrire una chiara indicazione del verificarsi del problema, e ancor meno di fornire eventuali indicazioni effettive su come risolvere il problema stesso. Di conseguenza, per risolvere un problema di questo tipo l'utente finale deve essere in grado di comprenderne i termini e sapere come utilizzare gli strumenti attualmente disponibili per trarre conclusioni dai dati limitati disponibili.



Figura 4 Caso di studio della topologia di sottoscrizione in eccesso

• Scenario:

L'utente 1 ha un'applicazione esistente in esecuzione su Host 2 (HBA 16G) che esegue I/O a varie dimensioni di block, profondità della coda e pattern di I/O. Questa applicazione è in esecuzione da molto tempo in questo ambiente e fino a poco tempo fa non ha dato luogo ad alcun problema. All'inizio di questo mese, l'utente 2 ha deciso di caricare un'applicazione su Host 1 (HBA 4G) a scopo di test. Inizialmente non è stato rilevato alcun problema nell'ambiente relativamente a prestazioni e latenza. Tuttavia, l'utente 1 ha recentemente iniziato a notare problemi di prestazioni con l'applicazione.

• Panoramica della risoluzione dei problemi:

Per risolvere qualsiasi problema in generale, è necessario innanzitutto comprendere le prestazioni e la configurazione dei singoli componenti quando si opera nelle condizioni ideali. Come noto, una rete SAN dispone di molte parti mobili che costituiscono l'ecosistema, pertanto è molto importante creare un profilo di ambiente costituito da profili per i tre componenti principali che costituiscono una rete SAN: applicazioni, fabric SAN e storage.

La creazione di questi profili baseline in vari componenti dell'ambiente permetterà di individuare facilmente i problemi quando si presentano. Occorre tenere presente che questi profili non sono estemporanei. È necessario raccogliere costantemente i dati della linea baseline per tutta la durata dell'ambiente, in modo non solo da risolvere i problemi, ma anche di pianificare un'ulteriore crescita e un'espansione.

Nelle prossime sezioni, mostreremo che è necessario raccogliere queste statistiche baseline dall'array di storage, in modo da avere la possibilità di individuare rapidamente la root cause quando si verifica un problema come quello indicato nello scenario precedente.

Baseline dell'applicazione

Con Dell EMC PowerMax e VMAX, quando si abilita il monitoraggio delle prestazioni è possibile tornare alla cronologia (fino a un anno dall'abilitazione della funzionalità), in modo da comprendere il profilo dell'applicazione in termini di IOPS e tempi di risposta prima che venisse apportata qualsiasi modifica. Disporre di questo profilo base dell'applicazione consente di utilizzare i grafici generati e di determinare facilmente l'origine dei problemi.

Generazione di grafici dei profili di base dell'applicazione

In Dell EMC Unisphere, fare clic su Performance > Charts

900197600237 👻	🗧 Storage 🛄 Hosts 🕻	Data Protecti	on als P	entormance	Databoses	System	Support
Dashboards Analyze Ch	arts						
Tame Frame	Symmetria		Sillect Char	1 0237 : Host	Host MDs/wet		
E Last Hour	000197600237	4	- Tappe	1			
Category			2				
Aerays 👒			atl.	1400 -			
Avg 🔾 Max			F				
ID	Host IOs/sec	*	Add Te	1200 -			
200102002223	A144		Chieft				

Selezionare un **intervallo di tempo**. Dovrà essere un momento qualsiasi PRIMA che si notasse il problema di prestazioni. Nel menu a discesa Categoria, selezionare **Hosts > Hosts**.

↑ 000197600237 ~ €	Storage 📮 Hosts 📑 Data Protec	tion II.
Dashboards Analyze Charts		
Time Frame	Symmetrix	Select C
Last Week v	000197600237 🗸	(Abr.
Category	New York Contraction Contracti	~
Hosts v		1.4
🗑 Avg 🕠 Max		111
ID	Host 10s/sec.	Add To
host_118	12950.8	Chart
NATE_HOST183_FC_ig	12659.4	1.4%
SITEA_IG	3084.3	C.T.
l0p061048	2409.4	
iop061046	1063.9	
yuhao_sgh_ucs	1	
Norm_UCS153_blade1_2_XLA0_1D9	0.7	
uci153_1	0.2	
epun_host_233_v1	0.1	
Chassis1_server3_ig	0.1	1
epun_host_136	0	
amin hort 138	0.	
✓ KPI Metrics		
Search KPI Metrics		
Host IOs/sec		
Host MBs Read/sec		
Host MBs Written/sec		
Host Reads/sec		
Host Writes/sec		
Read RT (ms)		
Response Time (ms)		
Syscall Count/sec		

Selezionare l'host in questione. Per **KPI Metrics** verranno generati sette grafici diversi. Ripetere questa sezione per ogni metrica KPI. Se si fa clic su tutte le metriche contemporaneamente, queste verranno tutte inserite in un unico grafico.

- a. I/O host/sec
- b. MB host/sec
- c. Letture host/sec
- d. Scritture host/sec
- e. Lettura RT (ms)
- f. Tempo di risposta (ms)
- g. Scrittura RT (ms)

Nella Figura 5, "I/O host e MB/sec", stiamo osservando gli I/O dell'host e i MB/sec. Da questi grafici possiamo vedere quando e per quanto tempo l'applicazione ha eseguito il maggior numero di I/O e ha utilizzato la maggiore quantità di larghezza di banda del link così come i momenti di minore attività.

Nota: nella legenda si noterà che ci sono due host, ma attualmente stiamo mostrando solo l'I/O per uno di essi perché l'altro host non è impegnato in attività di I/O.



01:25 01:35 01:45 01:55 02:05 02:15 02:25 02:35 02:45 02:55 03:05 03:15 03:25 03:35 03:45 03:55 04:05 04:15 04:25 04:35 04:45 04:55 05:05 Eiop061046

Figura 5 I/O host e MB/sec

0

I grafici mostrati nella Figura 6, "Operazioni di lettura e scrittura/sec", forniscono un quadro dettagliato del tipo di I/O generati dall'applicazione. Sulla base di questi grafici, possiamo determinare quale percentuale di I/O delle applicazioni sono di tipo READ o di tipo WRITE. In questo caso possiamo confermare che l'applicazione opera con un rapporto all'incirca 70/30 tra operazioni di lettura (READ) e scrittura (WRITE).



Figura 6 Operazioni di lettura e scrittura/sec

La Figura 7 e la Figura 8 sono probabilmente i grafici più utili per la risoluzione dei problemi. Forniscono un quadro dettagliato dei tempi di risposta tra operazioni READ e WRITE che ci permette di comprendere la latenza dell'applicazione. Ciò è estremamente utile quando occorre risolvere i problemi di prestazioni, poiché è possibile correlare eventuali picchi a eventi specifici mediante i grafici precedenti.



Figura 7 RT lettura e scrittura (ms)



Figura 8 Tempi di risposta (ms)

Ora abbiamo un profilo dell'applicazione per l'app dell'utente 1, sappiamo che opera con un rapporto 70/30 tra operazioni di lettura e scrittura con tempi di risposta medi di ~0.7ms e un tempo di risposta massimo di 2,3 ms.

Come indicato nella <u>sezione dedicata allo scenario</u>, abbiamo recentemente aggiunto un nuovo host che ha creato dei problemi di prestazioni nell'ambiente, quindi dobbiamo vedere come risolvere questo problema.

Poiché abbiamo un problema di prestazioni nel nostro ambiente, dovremo implementare le funzionalità disponibili nella SAN per determinare la tipologia dei problemi sorti.

Dal momento che conosciamo i tempi medi di risposta, in base al profilo dell'applicazione, sappiamo che questo problema è identificabile con tempi di risposta superiori al previsto.

Avvisi di diffusione della congestione di SAN Connectrix

In questa sezione verrà esaminato il tipo di eventi di congestione segnalati dal lato dello switch SAN. Assicurarsi di aver completato l'abilitazione di queste funzionalità nell'ambiente in base a quanto indicato nei prerequisiti.

- 4.1.1 Brocade
- 1. Assicurarsi di selezionare almeno **Top Port Traffic** e **BB Credit Zero** sul dashboard. In caso contrario, è possibile fare clic sulla chiave inglese nell'angolo in alto a sinistra per aggiungerli.

200	
Water part Chenry	
Trieffer A4 SAN TE Port Collector, A5 SAN FCP Tunnel Collector, A9 SAN FC Port Collector, A4 SAN G4	246
BB Credit Zwo Al Sala FC Part Calenter	
Nervery Utilization Percentege All SAN Product collector	
CPU Lititization Percentage A4 SAN Product collector	
Temperature All SAIs Product collector	
Response Tare All SAU Product submitter	
Port UNicides Percentage Ad SAN TE Port Collector, At SAN FOR Torreal Collector, Ad SAN FC Port Collector, Ad SAN Gu	
CRC Drom All SAIR TE Part Collector, All SAIR FC. Part Collector	
Link Resets AR SAIL PC Port Celeziter	
Line Fedures All Sold FC Part Collector	
13 Danuette Al SAS ET Dur Columbra	
De la la cita a citul	Const Zero C

Figura 9 Dashboard CMCNE

- 2. Quando si verifica una diffusione della congestione a causa di sottoscrizione in eccesso, come nell'esempio della Figura 9, "Dashboard CMCNE", in genere vengono visualizzati i seguenti avvisi nel dashboard CMCNE:
 - a. Porta F con utilizzo intensivo
 - b. Credito BB zero

Adand - Des EMC Wen	M Dense SAM									
419 527 - Tap Per	n Traffic		4.2	B X	San200621 -	tt Credit Zer			410×	
Part	Min TX Traffi.	TX Traff. 7	New TA Traff.		Port	Connected	88 Credit Zer	1: 98 Credit Zero	Nec Produc	
2 skt3 pert47	#18.20	STAIL NOT	421.111		alute portf.	skt4 portf.	The local division of	332868.345	awas.	
and sector pointed	400.015	491 907	402.071	8.00	alet4 portfi	STOR HOLE	Alter Said	332509.345	awach.	
alat4 port4	203.864	306.121	209.096		attit port?	Stop Hotel	Concession of the	332569 345	810/822	
B sur4 port5	202.238	296.23	294.395	20	B skt4 port4	skol posti	1000	332760.905	awach:	
stot4 port7	201,845	204.529	207,714	80	Attrop Class M	12.05.05.9	37470	20.817	84/825	
B stel4 port6	201.751	204743	267.513	44						
🔳 slot4 porté	0.103	0.104	0.100	844						
stot3 pert46	0.053	0.054	0.054	500						
alut2 port47	0.051	0.051	0.051	**						
B slot4 port5	0	0	0							
and the second second second						4			1.0	
Refreshed-1:22 PM					Refreshed-126	PM				
		-								
I slat4 port5 Retreated: 1.22 PM	4		0	÷	Retreated 1.26	PM				

Figura 10 Dashboard CMCNE con visualizzazione degli avvisi

- 3. La combinazione di questi due eventi, ovvero la porta F con utilizzo intensivo e i crediti buffer-to-buffer a zero sugli ISL possono indicare la presenza di un potenziale problema di prestazioni da risolvere. Fare riferimento al capitolo Risoluzione per indicazioni sugli elementi da esaminare.
 - 4.1.2 Cisco
- Nel dashboard DCNM dovrebbe essere presente il dashlet Top SAN End Ports. In caso contrario, è possibile aggiungerlo dal menu a discesa. In Top SAN End Ports sono visualizzati uno o più dispositivi con un utilizzo superiore al 90%. DCNM avrà soglie predefinite che fanno sì che una porta venga contrassegnata in giallo o in rosso quando supera l'utilizzo predefinito. Questo avviso di per sé non significa necessariamente che vi sia un problema di prestazioni nella rete SAN. Dovremo cercare anche altri avvisi nei fabric.

		entre Colla Caritter Pathae	u warda					1.1.1.4.+	Descine .
O Datitions	•	Deshboard / Summary							Casterna
	-	Martris		Top SA	End Ports		0.6×	Link Traffic	TT 14 TANK
X. mercenter		0 Critical 1		i Serapi	rese.				Date Certer
	10	Same Management, Party, Spirit Address	-+.0		days. But	and fa	Amond N.		Autto
Source and the		A firm 1		100	-	11 10 10		+	Name Status
d Manner	•	Select numbers 1900 AMER MIDDATE -	ALL 0 10	10.00	1010				22 No. Westweet
		Tarty Raman Tarty MOTOCANOIN	0 111	10.00		*	- m =		The Arts Every Horse (Bills for
Contigues		Index Surfaces Value AMER ARTISTS -		44	0.000	*==	10. CC		14-CPU (LILO)
		Sectored Ages, AND ACCRED-1		10.00		- m ()	19.72	-	Burnstone (Parties)
Administration		Index Indexes View, MAR ADDRESS		0000	*=	1 M	- MR		manney (Madaine)
		8 Marrier 144		A 1878	10.000	99.7002	10.5	the she was use when	inventory (Third)
		1 m 1	tas homosticipal like	-			100		monthly (sophali
		Top ISLs/Tranks							maning (Sublime) maning (Part Carachy)
		Station Report	Avg. No.						- Complex
		C Cumparia April - 10 Jack	en						
		2 Budring Admillionatives.	10. Total	U.					
		∠ the state + 2 0 the eff.	FR	1					
		2 Ottom to at a 177	10 Fairs						
		2 20040000000000000000	(m						
		2 Pasteria 1109-autobal (c).	PR	1					
		≤ @using h12-solute(h1)	10. T						
		C BORNARD and March	-						
		2 Durking LOA - orbital htt	Ph 11111	3					

2. Se viene segnalata una porta F con un uso intensivo, eseguire lo strumento di analisi del flusso lento. Fare clic su **Monitor—>SAN > Slow Drain Analysis.**



3. Eseguire lo strumento di analisi del flusso lento per 10 minuti. Al completamento del report si noterà che durante l'esecuzione dello stesso si è verificato un notevole incremento del contatore TxWait. La combinazione di questi avvisi e l'uso intensivo della porta F indicano la presenza di una congestione SAN dovuta a sottoscrizione in eccesso.

Diffusione della congestione dovuta a sottoscrizione in eccesso

tom NaN-NaN Stow Drain	-NeN NeNe Details fo	Nak to: 2017-07-19 16:03:09 In undefined									Z Totel 13	com 10 minutes M4
										Show	Non Zero data	rawk with 👻 🌱
-			180.00 M		Lawait 3		Level	2		Laiv	el 1	
ministration .	sheen	Connectio	Type	TxCreditLoss	TxLinkReset	RxLinkRe	TxTimeoutD	IxDiscard	TaWiAvg10	Red32Bite0	Tx83281to0	TxWait(2.5] *
≓ fc2/36	160b	IDP054151 fc2/36 (port channelli)	Switch	D	0	0	0	0	0	0	41429192	37.3350
≥ 1c2/37	1606	IDP054151 (c2/37 (port-channell))	Sultch	0	0	0	0	0	0	0	30150028	28.8393
€ fc2/38	10Gb	I0P054151 (c2/38 (port-channel6))	Switch	0	0	0	0	0	0	0	26941217	25,6649
	124.44	10 at 1 (1 47 1) 444444)	-				14					

La combinazione di questi due eventi, ovvero la porta F con utilizzo intensivo e i crediti buffer-to-buffer a zero sugli ISL, possono indicare la presenza di un potenziale problema di prestazioni da risolvere. Fare riferimento alla sezione <u>Risoluzione</u> per indicazioni sugli elementi da esaminare.

AVVISI DI DIFFUSIONE DELLA CONGESTIONE DI UNISPHERE

In questa sezione verrà illustrato come utilizzare Unisphere per PowerMAX e VMAX per correlare gli eventi dello switch SAN all'array di storage.

Assicurarsi di aver completato l'abilitazione di queste funzionalità nell'ambiente in base a quanto indicato nella sezione Prerequisiti.

 Utilizzare la procedura illustrata in <u>Generazione di grafici dei profili di base dell'applicazione</u>. Generare gli stessi sette grafici, aggiungendo l'host dell'utente 2 nel mix (perché era una delle modifiche recenti nell'ambiente prima che si verificasse il problema di prestazioni). Esaminare i dati.

Tenere presente che il nostro profilo base dell'applicazione = 70/30 lettura/scrittura con tempi di risposta medi di ~0,7 ms e un tempo di risposta massimo di 2,3 ms.

Nella Figura 11 di seguito, il raffronto tra gli I/O e i MB/sec non sembra indicare un possibile problema. Infatti, se confrontato con il grafico dell'applicazione baseline originale, il numero di IOPS è addirittura aumentato.

Inoltre, in alcuni punti siamo vicini alla velocità di linea (evidenziati di seguito). Questi punti entreranno in gioco in un successivo momento.







Figura 11 IO host e MB/sec

La *Figura 12* mostra il confronto in termini di operazioni di lettura/scrittura tra i due server, e non c'è molta differenza tra il profilo di I/O dei due server a questo punto.



Figura 12 Operazioni di lettura/scrittura dell'host/sec

La Figura 13 fornisce le informazioni più utili. Tenendo presente il nostro profilo dell'applicazione, stiamo registrando tempi medi di risposta di 0,7 ms con un massimo di 2,3 ms. Dal grafico qui sotto, possiamo vedere che c'è un enorme picco nei tempi di risposta in cui raggiungiamo l'intervallo di 8 ms e anche i nostri tempi di risposta medi complessivi sono aumentati.

Se riprendiamo la Figura 11, possiamo vedere che questi tempi di risposta elevati sono correlati a quando entrambi i server sono prossimi alla velocità di linea.

In genere, in Fibre Channel, sono necessari I/O con block di grandi dimensioni (maggiori di 128 k) per saturare un link.



Figura 13 Tempi di risposta delle operazioni di lettura e scrittura dell'host (ms)

CONCLUSIONI

Per ricapitolare tutte le informazioni acquisite finora in questo caso di studio:

• SAN Connectrix:

La nostra SAN sta segnalando un numero elevato di crediti buffer-to-buffer prossimi allo zero.
 Stiamo riscontrando un elevato utilizzo del traffico sulle nostre porte F.

• Dell EMC PowerMAX e VMAX:

1. Tempi di risposta elevati durante l'utilizzo completo del link

Come detto in precedenza, la congestione dovuta alla mancata corrispondenza della larghezza di banda è estremamente difficile da rilevare e confermare con il set di strumenti oggi disponibili. Tuttavia, in base agli avvisi di cui sopra, possiamo dedurre che il problema è dovuto alla mancata corrispondenza della larghezza di banda e a operazioni di lettura/scrittura di block di grandi dimensioni. Ciò si evince dai tempi di risposta elevati durante l'utilizzo completo del link

Un altro modo per rilevare questo problema è mediante il rapporto di congestione. Oggi dobbiamo calcolarlo manualmente nell'ambiente (oppure è possibile provare a creare uno script), ma sappiamo che se il rapporto C è superiore a 0,2 si verificherà una congestione poiché verrà esercitata una pressione a monte nell'ambiente SAN. Il rapporto C è la prima indicazione di rallentamento del flusso.

5 Risoluzione

PREVENZIONE

Per questo caso di studio specifico (diffusione della congestione a causa di sottoscrizione in eccesso), sono disponibili alcune opzioni che è possibile distribuire nell'ambiente per evitare che si verifichi questo problema.

Rapporto larghezza di banda

- Quando si esamina la rete SAN, si desidera identificare i dispositivi in esecuzione a velocità inferiori e comprenderne il tipo di profili di traffico dell'applicazione. La mancata corrispondenza della larghezza di banda NON implica necessariamente un problema.
- Esaminare il fabric end-to-end per assicurarsi che tutti i dispositivi finali siano in esecuzione alla stessa velocità di link.
- Garantire un'ampia quantità di larghezza di banda sugli ISL. Una buona regola generale è che la larghezza di banda ISL totale deve essere uguale o maggiore della quantità totale di larghezza di banda di storage nell'infrastruttura, ove possibile.
- È possibile modernizzare l'intera rete SAN aggiornando tutti i componenti end-to-end, come mostrato nella Figura 14 di seguito. La controindicazione di questo approccio è che una sottoscrizione in eccesso pari a zero dall'end-to-end in ambienti di grandi dimensioni non è sostenibile, oltre a essere estremamente costosa. Pertanto, la soluzione migliore è concentrarsi sull'aggiornamento dell'host specifico, dello switch e della connettività di storage.

Figura 14 Modernizzazione

Un altro modo potrebbe essere quello di procedere a una nuova suddivisione in zone, come mostrato nella Figura 15.

Figura 15 Prima e dopo la nuova suddivisione in zone

Implementazione dei limiti di larghezza di banda

Le piattaforme Dell EMC VMAX e Dell EMC Unity creano limiti di larghezza di banda per i gruppi di storage (VMAX) o i LUN (Unity). Nel caso di studio precedente avevamo una diffusione della congestione dovuta a sottoscrizione in eccesso, e una volta implementati i limiti di larghezza di banda le prestazioni erano tornate alla normalità come indicato nella Figura 16 di seguito. Questa operazione può essere effettuata direttamente tramite Unisphere nel gruppo Storage.

Con i limiti di I/O, è importante notare che questo approccio non sarà risolutivo con i cluster. Prendiamo la Figura 17, di seguito, come esempio. Quando il limite dell'host viene applicato a un host 4 Gb che causa la pressione a monte, l'array inizia a limitare la quantità di dati inviati all'host 4 Gb (in base ai limiti di I/O impostati); così facendo si elimina completamente la pressione a monte e gli altri flussi possono operare alla massima velocità di linea.

Figura 17 Limiti di I/O con i cluster

In questo esempio, abbiamo due host in esecuzione a 4 Gb/s che si trovano in un cluster. Poiché si trovano in un cluster, entrambi gli host avranno accesso al volume tramite ciascun fabric, e ciò significa che dobbiamo impostare un limite I/O di larghezza di banda di 1.600 MB/s (800 MB/s per ciascun FA). Tuttavia, con questo approccio non c'è nulla che impedisca a un singolo HBA di consumare tutti gli 800 MB/s.

• Isolamento

Un altro modo per prevenire questo problema consiste nell'isolare il traffico più lento dal traffico ad alta velocità utilizzando ISL dedicati. Questo risultato può essere ottenuto creando fabric virtuali (Brocade) o VSAN (Cisco) come indicato nella Figura 18 di seguito. Lo svantaggio di questo approccio è rappresentato dalla necessità di utilizzare porte dedicate, ma in questo modo si impedisce che il traffico più lento abbia un impatto sul traffico a velocità superiore. Per abilitare i fabric virtuali su Brocade, sarà necessario prevedere un downtime, poiché l'intero switch deve essere riavviato. Quando si sposta una porta in una VSAN diversa su Cisco, l'operazione incide solo sui dispositivi finali spostati.

Figura 18 Isolamento

ABILITAZIONE DEL MONITORAGGIO DELLE PRESTAZIONI

In questa sezione vengono descritte le operazioni necessarie per abilitare ed esaminare li dati di monitoraggio delle prestazioni in Unisphere per VMAX.

1. Eseguire il login alla GUI di Unisphere.

FMC Unisphere for VMAX	
ente onisphere for vinitit	
User	
Password	
Login Help	
	-
	EMC

2. Assicurarsi che l'array sia registrato per raccogliere i dati sulle prestazioni. In caso contrario, registrare l'array.

	000196/	01146					A (144	ч	
M	Wodel WALLION	Compare Lange	etion.	Ø	HYPESHAX O'S SATE HIDS 1125				
CAPACITY			103	FTEN UT	NUCCESS.				
Value +0	58 GB (# 1171) GB	Rept	er this syst	ante la co	And performance	tata.			
Overal Efficie	66 O.B. 14 22737 O.B. May 5.1	Repo	er the syst	am la co	Red performant e	lan.			
Urbail 40 Overal Efficie	ut an lé 22729 (an 1829 – S. t 2. Dectembres 10	1	er the syst	nor la co	Ani, petivenni e	uis.			
Votual 40	at an of 22727 an aty 5.1 • Performence =	1	er this syst	arria o	And performance	tata Bar	of Cause Andres	s Internal	
Virtual 40 Overat Efficie annecent par	of an of service an may 5.1 Dechanges a Bool Trace of	1 Frequencies	er the syst	an la co rectore Gal	Reci performance i Incluent Racio every (Maarsc)	tata Rec	of Cassie Analysi	s Enternal (Minutur)	Barrar ()

Selected Symmetrix(s)	000196701146	

MONITORAGGIO DELLA DIFFUSIONE DELLA CONGESTIONE IN CONNECTRIX

- 6.1.1 Brocade
- Discovery del fabric
- 1. Eseguire il login al server CMCNE e fare clic su Discover > Fabrics (SANnav?)

53											Vi	ew All	- CMCNE 14.0.0
Server £	(dt Ye	w Discov	er <u>C</u> onfigure	Honitor	Bep	orts	Icols	Help					
2		Host A	s dantare	a	1	0	-	Decimal 🔻	Name	•	Name	•	Legend
Dashboy	and - Proc	Inc An Ha	inagers										
Vew.	A8 +	VCEN	Servera			8	4 B	a (

2 Nella nuova finestra, fare clic su Add.

3		Discover Fabri	CS	
lee the Add button to discover a	Fabric.			
Discovered Fabrics				
Name	P Address	WWN	Discovery Status	1
				Add
				Cor.
				Delitte
				-
				- Reincover
				Seed Switz
				uns-unitio
				Referen
				1
	1.2.1			6354.

3 Inserire le informazioni necessarie per uno degli switch del fabric. CMCNE eseguirà automaticamente la discovery di tutti gli switch in tale fabric, presumendo che il nome utente e la password siano gli stessi per tutti gli switch nel fabric.

1
1
••••

- 4 Ripetere questa sezione per tutti gli altri fabric.
 - Abilitare MAPS e FPI
 - 1. Fare clic su Monitor > Fabric Vision > MAPS > Configure

Server	<u>E</u> dit ⊻iew	Discover Co	nfigure	Monitor Reports Tools	He	p		8	
3	2 🗟 🔅 📽 🖷		100	Eabric Vision	•	MAPS >	Violations	-	Na
-	<u> </u>			Fabric Watch	÷	Elow .	Configure	H	
Dasht	board - Dell El	MC World Demo	SAN	Configuration Policy Manage	r		Enable		
EMO	C World DEMO	0-10:00:c4:f5:7c	2d:b0:36	Performance	•				

2. Evidenziare il fabric e abilitare FPI.

Nota: FPI è abilitato per impostazione predefinita sugli switch che eseguono FOS 8.0 e versioni successive.

3. Da questo menu è possibile configurare ciascuno switch nell'infrastruttura e impostare la policy MAPS desiderata.

Nota: CMCNE fornisce policy predefinite che è possibile clonare e quindi modificare. Non è possibile modificare le policy predefinite. Fare riferimento alla guida dell'amministratore di MAPS sulle policy e sulle impostazioni specifiche.

In questo caso attiveremo la policy aggressiva predefinita. A tale scopo, evidenziare "dflt_aggressive_policy" e fare clic su activate. Questa operazione deve essere ripetuta su TUTTI gli switch nel fabric per i quali si desidera abilitare la policy, poiché al momento non è possibile abilitarla per l'intero fabric.

Stiamo attivando la policy aggressiva prima di avere un'idea immediata dei problemi nel fabric. Una volta completato questo passaggio, è possibile modificare e utilizzare le altre policy se riceviamo troppi avvisi. Facendo clic su **View**, è possibile esaminare le soglie per ciascun evento.

oducts	Policies	License	RAS Log Event	SNMP Trap	Port Decommission	Fence	E-ma	
All Fabrics						l en lannonnan		Violations
E Switch 5	offit aggressive policy	1	1	1			1	Enable FP
Active Policy	dft_aggressive_policy							Actions
- Policy	dflt_base_policy							E mail Set
Policy	dflt_moderate_policy							E-mail Set
Policy	dflt_conservative_policy							Policies
switch_5	dflt_aggressive_policy	1	1	1			1	Activate
Active Policy	offt_aggressive_policy							View
Policy	dflt_base_policy							VICW
Policy	dfit_moderate_policy							Add
Policy	dit_conservative_policy							Edit
								Delete
								Export
								Distribute
								Compare
								Groups
								Manage

	l'in																	
lame	dfl_aggressive_	policy																
Port	Switch Status	Fabric	FRU S	ecurity	Resource	FCIP. T	raffic / F	lows	FPI	Backend Po	irt GigE F	Port						
Rules	defined are thresh	olds measu	ured on port	s or SFPs	and determin	e if an out of	range v	iolation is	sen	for the port.								
roups	Rules				Monitor C	ondition	1	Time Bas	e R	AS Log Event	SNMP Trap	E-mail	FMS	Port Decommission	Fence	SFP Status Marginal	Quiet Time	T
Cust	om Groups										-							P
Syst	em Groups																	1
E.	ALL PORTS																	1
	detALL_PORT	SLOSS_S	GNAL 0		LOSS_SK	SNAL > 0		Min		1	1	1						8
	defALL_PORT	SLF_0	1994-07 7 5-8		LF > 0			Min		1	1	1						
	ALL_E_PORTS	100																l
	defALL_E_PO	RTSITW_1	5		ITW > 15			Min		1	1	1						I
	defALL_E_PO	RTSSTATE	CHG_4		STATE_C	HG > 4		Min			1	1		1	1			l
	defALL_E_PO	RTSSTATE	CHG_2		STATE_C	HG > 2		Min		1	1	1						I
	defALL_E_PO	RTSITW_2	0		ITW > 20			Min			1	1		1	1			I
	defALL_E_PO	RTSPE_0			PE > 0			Min		1	1	1						I
	defALL_E_PO	RTSPE_2			PE > 2			Min			1	1		1	1			I
	defALL_E_PO	RTSLOSS	SYNC_0		LOSS_SY	NC > 0		Min		1	1	1						l
	defALL_E_PO	RTSC3TXT	0_5		C3TXTO >	5		Min		1	1	1						1
	defALL_E_PO	RTSCRC_0)		CRC > 0			Min		1	1	1						I
	detALL_E_PO	RTSCRC_2	2		CRC > 2			Min			1	1		1	1			I
	defALL_E_PO	RTSLR_2			LR > 2			Min		1	1	1						I
	defALL_E_PO	RTSLR_4			LR > 4			Min			1	1		1	1			Ľ
E	ALL_F_PORTS																	l
	defALL_F_PO	RTSDEV_N	PIV_LOGN	S_PER_60	DEV_NPN	LOGINS >	50 %	None		1	1	1						l
Ð	ALL_HOST_PORT	s																l
	defALL_HOST	_PORTSIT	W_15		ITW > 15			Min		1	1	1						1
	defALL_HOST	PORTSLA	R_4		LR > 4			Min			1	1		-	1			ľ
	defALL_HOST	PORTSLA	R_2		LR > 2			Min		1	1	1						l
	detALL_HOS	PORTSIT	W_20		ITW > 20			Min			1	1		1	1			I
	defALL_HOST	_PORTSLO	DSS_SYNC_	0	LOSS_SY	'NC > 0		Min		1	1	1						Į
	defALL HOST	PORTSST	TATE CHG	4	STATE C	HG > 4		Min			1	1		1	1			Ŀ

4. Evidenziare uno switch e fare clic su **Actions**. Da qui è possibile decidere le azioni da intraprendere in caso di diffusione della congestione. Per il nostro caso di studio specifico di diffusione della congestione dovuta a sottoscrizione in eccesso, è sufficiente selezionare le opzioni **email** e **RAS log event**.

9		MAPS Configuration			
Use this dialog to configure produce more violations	MAPS for Fabric OS switches	Three default noticies (beginning with dflt.) are availab MAPS Policy Actions	le on each switch. The aggr	essive	policy will
Products	D. MARS-305 3	1 (A. 1977)	rt Decommission Fence	E-ma	
E All Fabrics	Use this dialog to manage po	icy actions on a switch.			Violations
E 😌 EMC World C E	Enable All Disable All			1	Enable FPI
Active P	Actions				Actions
- Policy	RAS Log Event	SFP Status Marginal			E-mail Setup
- Policy - Policy	SNMP Trap	E FMS			Policies
E 😂 switch_	🗌 Port Decommission 🕚	FPI Actions		1	Activate
- Active P	Fence			1	View
- Policy Policy	E-mail			1	VIEW
Policy				17	Add
	Switch Status Actions				Edit
	Switch Stabis Marginal				Delete
	Switch Status Critical				Export
	Note: This operation is suppo	rted only on MAPS licensed switches.			Distribute
		OK Cancel Help			Compare
L					Groups
					Manage
				•	
				-	
				CI	ose Help

5. Se si desidera ricevere avvisi via e-mail, fare clic su E-mail Setup e compilare i campi appropriati.

Use Test E-mail t Apply and then T	utton to initia est.	te test E-mail fi	rom the switc	nes. In case of n	iew E-mail settin	gs, click
Settings						
L-mail address						
Relay Host						
Domain Name						
	Apply	Test E-mail				

6. Ripetere questi passaggi su TUTTI gli switch nel fabric.

6.1.2 Cisco

- Discovery del fabric
- 1. Eseguire il login a DCNM e fare clic su Inventory > Discovery > SAN > Switches.

	Ŧ	Inventory	Center <mark>N</mark> etwork Manager
Dashboard	۲	View Switches	covery / LAN Switches
📩 Topology		Licenses	Move & Rediscover
		Discovery LAN Switches SAN Switches	IP Address
Monitor	٥	Storage Devices Virtual Machine Manager	
no Configure	۲		
Administration	۲		

2. Nella nuova finestra, fare clic sul **segno più (+)** e immettere le informazioni richieste per uno degli switch nel fabric.

Add Fabric			
Fabric Seed Switch:	1,1,1,1		1
SNMP:	Use SNMP	v3/SSH	÷
	Auth-Privacy:	MD5	
User Name:	admin		1
Password:			1
	☐ Limit Disco ✓ enable NP\	wery by VSAN / Discovery in	All Fabrics
	Add	Options>>	Cancel

3. Ripetere questa sezione per tutti gli altri fabric.

Abilitazione del monitoraggio delle porte Cisco (PMON)

1. Fare clic su **Configure > SAN > Port Monitoring**.

		Ŧ	Configure
🕥 Dashbo	ard (>	Deploy vPC Peer
Topolo	٩V		vPC Templates
S Invento	ny 🤅		Deploy Jobs
Monitor		>	Backup Switch Configuration Network Config Audit
🔥 Configu	ire 🤇		Image Management
⊥ ° Admini	stration (>	Patch [SMU] Package [RPM] Maintenance Mode [GIR] Repositories
			SAN
			Zoning FCIP
			Davico Alias
			Port Monitoring
			Credentials Management
			SAN Credentials
			LAN Credentials

			enen	Data Center Network	M	enager						O_{k} +	amin
6 01	ishboird	•	Config default *	ture / SAN / Port Monito	ning	Push to us	intee.						Last Updaled: 2017/95/15-17-84
🛠 То	pology		SI No	Counter Description		Rising Throut	Aluingtower	Falling Inves	fallingtown	Poll Interval	Warning Thes	Port Gearth	Monitor 7
🕿 Inv	watery	•	4	Link Sone	0	\$	Variety	1	Harring	01	1	false	tue
-			7	Type Lose	91	5	Warring		thereing .		1	tabe	The State
() M	10000	۰		Invalid Viseda		1	Warring		Juning		1	Select :	10.0
tim	1		E	Invalia CRC		4	Vianing	1	thinking .	61	1	halte	tue
A C0	ontidure	•	4	To Doubth		300	Tiering	10	Taring			50+	114
-	ALAN ALAN ALAN AL	~	P. (LR.Ra		4	Vianing	1	maning	01	1	Salver	2016
* A0	brunchration	•	1	LR Ta			Tianing	3	Taring	65		Sile:	114
			1	Timenat Discard		200	Vianing	10	maning	01	1	fahe .	5140
			10	Credit Loss Reco		4	Tiering		Thering .	1		tite.	110 -
			11	Tx Credit Not-Available (N)		10	Vianing		maning	1.		false .	true
			12	Ro Delarette (%)		81	Tiering	20	Thering .	-		tite.	2140
			13	To Detarola (%)		80	Vianing	28	thaning	01	1.	Salan.	
			34.	To Steepert Oper Delay (resect	10	94 -	Tiaring		Warring :	1		tite.	11.0
			86.	forware (%)		40	Vianing	÷	thaning	E.	1	Salan .	1040 ·
			W	State Charge		4	Tianing		Taning .			take :	11.0

2. Selezionare il profilo predefinito e fare clic su Push to switches.

3. Selezionare tutti i fabric e fare clic su Push.

ort Type:	trunks/Core access-port/Edge all
	💌 🗹 🗁 Data Center
	👻 🗹 🗁 SAN
	Fabric_AMER-MDS9513-1
	Fabric_AMERGen2MDS9509
C	Fabric_10P054150
Scope:	
Note: Featur	es not supported by the NX-OS version of the switch will be ignored.
	Durb Concel

Push to switches Result

Policy: default Port Type: All

		10		-	
Switch Name	IP Address	Status			
AMER-MDS9513-1		Success			
AMERGen2MDS9509		Success			

4. Con Cisco MDS, è possibile ricevere gli avvisi tramite SNMP o Syslog. Fare riferimento alla seguente Configuration Guide per la configurazione di entrambi:

http://www.cisco.com/c/en/us/support/storage-networking/mds-9000-nx-os-san-os-software/products-installationand-configuration-guides-list.html

5. Per configurare la home per l'e-mail (opzionale), fare clic su Administration > Event Setup.

6. Fare clic sul segno più (+), specificare l'indirizzo e-mail del destinatario e fare clic su Add.

Forwarding Method:	E-Mail Trap
Email Address:	dell_emc@dell.com
Forwarding Scope:	Fabric/LAN Port Groups
Scope:	All Fabrics 🔻
VSAN Scope:	All O List
Source:	DCNM Syslog
Type:	All 🔻
	Storage Port Only
Minimum Severity:	Emergency 🔻

7. Completare le informazioni relative al server SMTP e all'indirizzo e-mail del mittente, quindi fare clic su **apply and test** per confermare di aver ricevuto l'e-mail.

Crushie	SMTP Server:	1.1.1.1		From:	dom@del.cam		Apply	Apply and Test *	
300000	Stort:	37/10/2017 09 58 AM	- 223	End:	07/10/2017 89-59 AM	10	Apply		

 Verificare che il monitoraggio delle prestazioni sia in esecuzione. Fare clic su Administration > Server Status. Accertarsi che Performance Collector sia in esecuzione. In caso contrario, fare clic sul pulsante play per avviarlo.

9. Fare clic su Administration > Performance Setup > SAN > Collections.

10. Assicurarsi che i fabric per i quali si desidera raccogliere le statistiche delle prestazioni siano selezionati. Il servizio Performance Collector verrà riavviato.

Riferimenti

Configuration Guide di Brocade MAPS: <u>http://www.brocade.com/content/html/en/configuration-guide/fos-80x-maps/GUID-426E1CD4-3763-</u> <u>419D-9D54-91F824F463EB-homepage.html</u>

White paper sul dispositivo a flusso lento Cisco: <u>http://www.cisco.com/c/dam/en/us/products/collateral/storage-networking/mds-9700-series-multilayer-</u> <u>directors/whitepaper-c11-737315.pdf</u>

Informazioni generali sulle funzionalità di limitazione degli I/O dell'host VMAX: <u>https://community.emc.com/thread/188068?start=0&tstart=0</u>

Ezfio I/O tool

https://github.com/earlephilhower/ezfio

Severity della diffusione della congestione

Anche se le metriche di diffusione della congestione sono importanti, come illustrato nella sezione di seguito, la velocità con cui gli eventi si verificano può mutare radicalmente l'impatto di ciascun evento sull'ambiente. Per complicare ulteriormente l'assunto, occorre tenere presente che Brocade e Cisco utilizzano uno schema di categorizzazione diverso. Di conseguenza, utilizzeremo il seguente schema di categorizzazione specifico di Dell EMC e li mapperemo a ciascuno dei tipi di switch come mostrato di seguito:

6.1.3 Dell EMC

- Tipo 1:
- o Rapporto di congestione maggiore o uguale a 0,2
- o Nessuna perdita di frame (eliminazioni) o reimpostazione dei link
- Tipo 2:
- Rapporto di congestione maggiore o uguale a 0,2
- o Perdita di frame (eliminazioni), ma nessuna reimpostazione dei link
- Tipo 3:
- Rapporto di congestione maggiore o uguale a 0,2
- Perdita di frame (eliminazioni) e reimpostazione dei link

6.1.4 Brocade

- Lieve
- o Ritardo minimo del credito
- Latenza ridotta della coda (minore di 10 ms)
- Nessuna perdita di frame (eliminazioni) o reimpostazione dei link

• Moderato

- o Ritardo medio del credito
- Latenza media della coda (10 ms 80 ms)
- o Perdita di frame (eliminazioni), ma nessuna reimpostazione dei link

Grave

- o Ritardo rilevante del credito
- Latenza rilevante della coda (maggiore di 80 ms)
- Perdita di frame (eliminazioni) e reimpostazione di alcuni link

6.1.5 Cisco

- Livello 1: Latenza
- o Riduzione del numero di crediti rimanenti o indisponibilità di breve durata del credito
- Nessuna eliminazione, ritrasmissione o ripristino dei link
- Livello 2: Ritrasmissione
- o Durata prolungata dell'indisponibilità del credito
- I frame vengono eliminati (ma nessuna reimpostazione del link) a causa del timeout di connessione dovuto alla congestione o all'esaurimento del credito che determina la ritrasmissione.
- Livello 3: Ritardo estremo
- o Durata prolungata dell'indisponibilità del credito (1 sec per la porta F, 1,5 secondi per la porta E)
- o Reimpostazione dei link o flap delle porte

Riferimento incrociato per la terminologia di diffusione della congestione

Le metriche e i livelli di severity possono essere combinati e utilizzati per identificare i diversi tipi di eventi di diffusione della congestione. Come per la sezione precedente, è presente una sezione separata sia per Brocade che per Cisco, ma dal momento che sia Brocade che Cisco utilizzano il termine "sottoscrizione in eccesso" (oversubscription), questa sezione inizierà con una panoramica di questo termine.

6.1.6 Sottoscrizione in eccesso

La sottoscrizione in eccesso è semplicemente una condizione in cui "la domanda potenziale su un sistema supera la capacità del sistema di soddisfare tale domanda". Un esempio noto a tutti è il sistema autostradale. Se tutti improvvisamente decidessero di guidare la propria auto contemporaneamente (ad esempio in caso di evacuazione forzata a causa di un uragano), il traffico si bloccherebbe.

Nel caso di una SAN FC, è utile pensare alla sottoscrizione in eccesso in termini di rapporto di larghezza di banda. Ad esempio, come mostrato nella Figura 3, il rapporto di larghezza di banda tra Host 1 (4 Gbps) e Storage 1 (16 Gbps) è di 1:4. Pertanto, possiamo dire che Host 1 è sottoscritto in eccesso con un rapporto di 4:1. Confrontare questa situazione con il rapporto di larghezza di banda tra Host 2 (16 Gbps) e Storage 2 (16 Gbps), che è 1:1. Tenere presente che sia gli host che lo storage cui questi accedono utilizzerà un ISL a 32 Gbps, pertanto non si verifica alcuna sottoscrizione in eccesso tra Host 2 e Storage 2. In questo caso diciamo che Host 2 e Storage 2 non sono affetti da sottoscrizione in eccesso.

Figura 19 Rapporto di larghezza di banda - Esempio 1

È importante notare che, quando si calcola la sottoscrizione in eccesso, come mostrato nella Figura 19, il rapporto di larghezza di banda viene calcolato aggiungendo la larghezza di banda delle interfacce prese in considerazione. A prima vista, si potrebbe pensare che abbiamo un HBA a 16 Gbps che accede a uno storage a 8 Gbps, ma dal momento che ci sono in realtà tre interfacce di storage, abbiamo un HBA a 16 Gbps che accede a 24 Gbps di storage. Di conseguenza l'host è sottoscritto in eccesso con un rapporto di 3:2.

Figura 20 Rapporto di larghezza di banda - Esempio 2

Nei due esempi precedenti, il rapporto di larghezza di banda dell'ISL era sempre maggiore o uguale alla quantità di larghezza di banda supportata dai dispositivi finali. In genere, però, non è così. Come mostrato nella Figura 20, l'host è interessato da una sottoscrizione in difetto con rapporto di 3:4, ma poiché l'ISL è solo di 16 Gbps, esiste una sottoscrizione in eccesso tra il dispositivo finale e gli ISL che verranno utilizzati, tanto da poter affermare che gli ISL sono oggetto di una sottoscrizione in eccesso con rapporto 3:2.

6.1.7 Brocade

Brocade definisce tre diverse classi di eventi di diffusione della congestione:

Sottoscrizione in eccesso

Come definito nella sezione precedente (sopra).

• Dispositivo malfunzionante

Un dispositivo finale o un ISL che non rilascia credito con una rapidità sufficiente per sostenere la velocità di linea. Ad esempio, se un dispositivo finale ha negoziato una velocità di link di 16 Gbps e non è in grado di restituire credito a una velocità tale da consentire la ricezione di 16 Gbps di dati, si parla allora di dispositivo malfunzionante. Questi tipi di dispositivi sono anche indicati come "a flusso lento". È importante sottolineare che un dispositivo potrebbe non operare correttamente per una serie di motivi, tra cui un problema del driver oppure, nel caso di un ISL, perché la porta sta subendo gli effetti della diffusione della congestione.

Credito perso

Uno scenario di credito perso significa che per un qualche motivo (ad es. errori generalmente occasionali di bit), uno o entrambi i dispositivi su un determinato link credono di avere meno credito di trasmissione di quanto non ne abbiano realmente. Una possibile causa di questa situazione potrebbe essere un errore di bit che danneggia un R_RDY. Se ciò accade abbastanza spesso, le prestazioni inizieranno a peggiorare nel tempo e ridurre lentamente la capacità della SAN di trasportare i dati. Questo problema viene esaminato in modo più dettagliato nel KB 464245 (Errori di bit e relativo impatto).

Riferimento incrociato per la terminologia di diffusione della congestione Brocade

Per Brocade, mettendo assieme tutte le possibili variabili si ottiene il seguente riferimento incrociato per la terminologia di diffusione della congestione specifica di Brocade.

Causa	Lieve	Moderato	Grave
Sottoscrizione in eccesso ¹	 Larghezza di banda	 Larghezza di banda elevata	 Larghezza di banda elevata sulla porta del
	elevata sulla porta del	sulla porta del dispositivo Latenza del credito media	dispositivo Latenza del credito elevata sulla porta ISL Latenza della coda superiore a 80 ms sulla
	dispositivo Latenza del credito ridotta	sulla porta ISL Latenza della coda tra 10 m	porta ISL Perdita di frame nella porta a monte (ISL)
	sulla porta ISL Nessuna perdita di frame	e 80 ms sulla porta ISL Nessuna perdita di frame	(indica la latenza della coda non compresa
	o reimpostazione dei link	o reimpostazione dei link	nel range 220 ms - 500 ms) Nessuna reimpostazione del link:
Dispositivo malfunzionante	 Latenza del credito ridotta sulla porta del dispositivo e sulla porta ISL a monte Latenza della coda inferiore a 10 ms sulla porta del dispositivo e sulla porta ISL a monte Nessuna perdita di frame o reimpostazione dei link 	 Latenza del credito media sulla porta del dispositivo e sulla porta ISL a monte Latenza della coda tra 10 ms e 80 ms sulla porta del dispositivo e sulla porta ISL a monte Nessuna perdita di frame o reimpostazione dei link 	 Latenza del credito elevata sulla porta del dispositivo e sulla porta ISL a monte Latenza della coda superiore a 80 ms sulla porta del dispositivo e sulla porta ISL a monte Perdita di frame sulla porta del dispositivo o sulla porta a monte (ISL) (indica la latenza della coda non compresa nel range 220 ms - 500 ms) Reimpostazione dei link su una porta ISL (indica lo stallo del credito per più di 2 secondi)
Credito perso ²	 Latenza del credito ridotta	 Latenza del credito media	 Latenza del credito elevata sulla porta Latenza della coda superiore a 80 ms sulla porta
	sulla porta Latenza della coda	sulla porta Latenza della coda tra 10 ms	o a monte della porta Perdita di frame sula porta o a monte dalla porta
	inferiore a 10 ms sulla	e 80 ms sulla porta o sul link	(indica lo stallo del credito per 220 ms - 500 ms) Reimpostazione dei link sulla porta o a monte
	porta o a monte della porta Nessuna perdita di frame	a monte della porta Nessuna perdita di frame	della porta (indica lo stallo del credito per più di
	o reimpostazione dei link	o reimpostazione dei link	2 secondi)

¹La congestione grave dovuta alla sottoscrizione in eccesso è un'evenienza rara o estremamente rara.

²Le cause di perdita di credito sono in genere da ascrivere a errori di trasmissione quali ITW, CRC o altri problemi relativi al segnale.

6.1.8 Cisco

Cisco definisce due diverse classi di eventi di diffusione della congestione:

• Sottoscrizione in eccesso

Come definito in precedenza.

• Esaurimento del credito

Un dispositivo finale o un ISL che non rilascia credito con una rapidità sufficiente per sostenere la velocità di linea. Ad esempio, se un dispositivo finale ha negoziato una velocità di link di 16 Gbps e non è in grado di restituire credito a una velocità tale da consentire la ricezione di 16 Gbps di dati, si parla allora di dispositivo malfunzionante. Questi tipi di dispositivi sono anche indicati come "a flusso lento". È importante sottolineare che un dispositivo potrebbe non operare correttamente per una serie di motivi, tra cui un problema del driver oppure, nel caso di un ISL, perché la porta sta subendo gli effetti della diffusione della congestione.

• Riferimento incrociato per la terminologia di diffusione della congestione Cisco

Per Cisco, mettendo assieme tutte le possibili variabili si ottiene il seguente riferimento incrociato per la terminologia di diffusione della congestione specifica di Cisco.

Tipo di congestione	Livello - 1: Latenza	Livello - 2: Ritrasmissione	Livello - 3: Ritardo estremo
Sottoscrizione in eccesso	 Utilizzo elevato dei link sulla porta del dispositivo finale Nessun esaurimento del credito B2B sulla porta del dispositivo finale Diffusione della congestione verso gli ISL Nessuna perdita di frame o reimpostazione dei link 	La ritrasmissione o il ritardo estremo rara o estremamente rara.	dovuto alla sottoscrizione in eccesso è un'evenienza
Esaurimento del credito	 Utilizzo ridotto dei link sulla porta del dispositivo finale Riduzione del numero di crediti rimanenti o indisponibilità di breve durata del credito Diffusione della congestione verso gli ISL Nessuna eliminazione, ritrasmissione o ripristino dei link 	 Utilizzo ridotto dei link sulla porta del dispositivo finale Durata prolungata dell'indisponibilità del credito Diffusione della congestione verso gli ISL I frame vengono eliminati (ma nessuna reimpostazione del link) a causa del timeout di connessione dovuto alla congestione o all'esaurimento del credito che determina la ritrasmissione 	 Nessun frame viene trasmesso al dispositivo finale Durata prolungata dell'indisponibilità del credito (1 sec. per la porta F, 1,5 sec. per la porta E) Grave congestione verso gli ISL Reimpostazione dei link o flap delle porte
*Configurazione predefinita: timeou timeout) = off	t frame persi per congestione (congest	tion-drop timeout) = 500 ms, timeout f	rame persi per mancanza di credito (no credit-drop

Opzione configurabile: congestion-drop timeout = 100 - 500 ms, no-credit-drop timeout = 1 = 500 ms

Configurazione consigliata: congestion-drop timeout = 200 ms, no-credit-drop timeout = 50ms