

輻輳分散とそれを回避する方法

このドキュメントでは、輻輳分散（低速ドレインとも呼ばれる）が、Connectrix B-Series と MDS Series の両方について各タイプの輻輳の重大度を説明するために使用されるメトリックと、輻輳の分散の影響を回避するために取ることができる予防措置について説明します。

2019年5月

輻輳分散とそれを回避する方法 | H17762.2 |

リビジョン

日付	説明
2019年5月	初期リリース

謝辞

このホワイトペーパーは、Dell EMC ストレージ エンジニアリング チームの次のメンバーによって作成されました。

作成者：

Alan Rajapa

Erik Smith

この資料に記載される情報は、「現状有姿」の条件で提供されています。Dell Inc.は、この資料に記載される情報に関する、どのような内容についても表明保証条項を設けず、特に、商品性や特定の目的に対する適応性に対する黙示の保証はいたしません。

この資料に記載される、いかなるソフトウェアの使用、複製、頒布も、当該ソフトウェア ライセンスが必要です。

©2019年5月発行：Dell Inc. その関連会社。All rights reserved.（不許複製・禁無断転載） Dell、EMC、Dell EMC、および Dell または EMC が提供する製品及びサービスにかかる商標は Dell Inc.またはその関連会社の商標又は登録商標です。その他の商標は、それぞれの所有者の商標又は登録商標です。

掲載される情報は、発信現在で正確な情報であり、この情報は予告なく変更されることがあります。

目次

1	はじめに.....	4
2	概要.....	6
	前提条件.....	6
3	輻輳の分散とは？.....	8
4	オーバーサブスクリプションによる輻輳分散.....	11
	アプリケーション ベースライン.....	12
	アプリケーション ベース プロファイル グラフの生成.....	12
4.1.1	Brocade.....	18
4.1.2	Cisco.....	19
	UNISPHERE 輻輳分散アラート.....	20
	まとめ.....	25
5	修正.....	26
	予防.....	26
	帯域幅比.....	26
	帯域幅制限の実装.....	27
6	付録.....	29
	パフォーマンス監視の有効化.....	29
	CONNECTRIX 輻輳分散モニタリング.....	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	オーバーサブスクリプション.....	46
6.1.7	Brocade.....	47
6.1.8	Cisco.....	49

1 はじめに

このドキュメントでは、輻輳分散（低速ドレインとも呼ばれる）が、Connectrix B-SeriesとMDS Seriesの両方について各タイプの輻輳の重大度を説明するために使用されるメトリックと、輻輳の分散の影響を回避するために取ることができる予防措置について説明します。

Dell EMCは、自社製品ラインのパフォーマンスや機能を向上させ強化するための取り組みの一環として、ハードウェアとソフトウェアのリビジョンを適宜リリースしています。そのため、このガイドで説明されている機能の中には、現在お使いのソフトウェアまたはハードウェアのリビジョンによっては、サポートされていないものもあります。製品機能の最新情報については、お使いの製品のリリース ノートを参照してください。

製品が正常に機能しない、またはこのマニュアルの説明どおりに動作しない場合には、Dell EMCの担当者にお問い合わせください。

対象読者

このTechBookは、テクノロジー コンサルタントを含むDell EMCフィールド担当者、およびEMCとホスト デバイスを含むネットワークストレージ環境の取得、管理、運用、設計に携わるストレージ アーキテクトや管理者を対象としています。

関連

ドキュメント

すべての関連ドキュメントおよびリリース ノートについては、<https://dell.com/support>を参照してください。[製品サポート] をクリックし、製品名を入力して、ドキュメントをクリックします。

Dell EMCサポートマトリックス

およびE-Lab

Navigator

相互運用性

最新の情報については、必ずDell EMCサポート マトリックスを参照してください。これは以下のアドレスのE-Lab Interoperability Navigator (ELN) にあります。

<https://www.dell.com/ja-jp/products/interoperability/elab.htm#tab0=2>

問い合わせ先

Dell EMCのサポート情報、製品情報、ライセンス情報は、次に説明しているDell EMCオンライン サポート サイトから入手できます。

注： Dell EMCオンライン サポート サイトを通してサービス リクエストを開始するには、有効なサポート契約が必要です。有効なサポート契約を取得する方法の詳細や、アカウントに関するご質問については、Dell EMC担当営業にお問い合わせください。

製品情報

ドキュメント、リリース ノート、ソフトウェアの更新の他、Dell EMC製品、ライセンス、サービスに関しては、以下のDell EMCオンライン サポートWebサイトをご覧ください（登録が必要です）。

<https://www.dell.com/support>

テクニカル サポート

Dell EMCは、さまざまなサポート オプションを提供しています。

Support by Product : Dell EMCは、次のWebサイトで統合された、製品固有の情報を提供します。

<https://support.dell.com/products>

[Support by Product] を提供するWebページでは、ドキュメント、ホワイトペーパー、アドバイザリー（頻繁に使用されるナレッジベース記事など）、ダウンロードのほか、プレゼンテーション、ディスカッション、関連するカスタマーサポートフォーラムへのエントリー、[EMC Live Chat] へのリンクなど、より動的なコンテンツにもアクセスできます。

Dell EMC Live Chat : Dell EMCのサポートエンジニアとチャットやインスタントメッセージでやり取りできます。

eライセンス サポート ライセンス資格をアクティブ化して、ライセンスファイルを取得するには、お客様宛てにメールで送信したLAC（ライセンス認証コード）レターに記載されている指示に従って、<https://dell.com/support>のサービスセンターにアクセスしてください。

2 概要

このホワイトペーパーの目標は、次のとおりです。

1. 輻輳分散（低速ドレインと呼ばれる）がSAN（ストレージエリアネットワーク）にどのように影響を及ぼすかを説明します。
2. Connectrix-B SeriesとMDS Seriesの両方の製品について、各重大度と輻輳のタイプを説明するために使用されるメトリックを定義します。
3. 輻輳分散の影響を回避するために使用できる予防措置について説明します。
4. 前述の情報を使用して、オーバーサブスクリプションによる輻輳分散を検出、防止、修正する方法を説明します。

前提条件

注：

このドキュメントでは、次のソフトウェアバージョンが使用されていることを前提としています。ステップは、古いバージョンでは異なる場合があります。必要な機能を有効にする方法の詳細については、付録を参照してください。

1. Dell EMC Unisphere for PowerMaxとfor VMAXがインストールおよび実行されており、アレイがパフォーマンス データを収集するように登録されている。
https://www.dell.com/support/products/27045_Unisphere-for-/Documentation/?source=promotion
2. SAN管理GUIがインストールされている。
 - a. Brocadeファブリックの場合：CMCNE（Connectrix Manager Data Center Edition）14.x以降
ダウンロード：
https://www.dell.com/search/?text=CMCNE%2014&searchLang=en_US&facetResource=DOWN
管理者ガイド：
https://www.dell.com/search/?text=CMCNE%2014%20admin%20guide&searchLang=en_US
 - b. Ciscoファブリックの場合：Cisco DCNM（Data Center Network Manager）10.x以降
ダウンロード：
<https://www.dell.com/support/search/?text=DCNM%2010&facetResource=DOWN>
管理者ガイド：
<https://www.cisco.com/c/en/us/support/cloud-systems-management/prime-data-center-network-manager/products-installation-guides-list.html>
3. SANスイッチ ファームウェアは、次のものである必要があります。
 - a. Brocade：Fabric O.S 7.4.1d以降
ダウンロード：
https://www.dell.com/support/search/?text=Brocade%20FOS%20download&searchLang=en_US&facetResource=DOWN

- b. Cisco : NX-OS 6.2(13)以降
ダウンロード :
<https://www.dell.com/support/search/?text=NX-OS%20download>

- 4. 必要なすべてのパフォーマンス監視ライセンスがインストールされている。
 - a. BrocadeにはMAPSライセンスが必要 :
<https://docs.broadcom.com/docs/53-1005239-04>

 - b. CiscoにはDCNM SANサーバー パッケージのライセンスが必要です。
<https://www.cisco.com/c/en/us/support/cloud-systems-management/prime-data-center-network-manager/products-installation-guides-list.html>

 - c. PowerMAXとVMAXには、Dell EMC Unisphere e ライセンスが必要です。詳細については、次のPDFの21ページを参照してください。
<https://www.dell.com/collateral/TechnicalDocument/docu88904.pdf>

3 輻輳の分散とは？

ストレージ アレイとの間でデータを転送するには、すべてのデータを適切なタイミングで宛先に配信する必要があります。これは、SCSI（たとえば、ファイバー チャネル-FCPなど）を使用するブロックベースのストレージ プロトコルの場合に特に当てはまります。このホワイトペーパーではその理由は説明されていません。詳細については、『Networked Storage Concepts and Protocols』Techbookの「Congestion and Backpressure」のセクションを参照してください。

(<https://www.dellemc.com/ja-jp/products/interoperability/elab.htm#tab0=1hardware/technical-documentation/h4331-networked-storage-cncpts-prtcls-sol-gde.pdf>)。

他のネットワーク プロトコルと同様に、FC（ファイバー チャネル）は、幅広い一般的なネットワーク輻輳状況において、データをタイムリーに配信する必要があります。FCで使用されるメカニズムは、バッファツーバッファ フロー制御を使用したフレーム ロスの防止に重点を置いています。このため、FCは「ロスレス プロトコル」とみなされます。

各プロトコルで使用されるフロー制御メカニズムはいくらか異なりますが、FCやその他のロスレス プロトコル（たとえば、DCB Ethernetや Infinibandなど）は、リンクの一方の端で受信側が容量の上限に近づいていることをトランスミッターが判断できるようにすることで、リンクのもう一方の端でのバッファ オーバーフローを防いでいます。そのように判断されると、ポートは、リンクのもう一方の端がさらにデータを受信する準備ができていないことを示すまで、データの転送を停止します。トランスミッターがこの状態になるとフレームを転送できません。このことを「輻輳が発生している」と表現します。トランスミッターで長期間にわたって輻輳が発生する場合、その輻輳がソースに伝播する可能性があります。この現象は、輻輳分散として知られています。その例を次の一連の図で示します。

図1は、輻輳が発生していないSANの例を示しています。ホスト1とホスト2の両方が、アレイへのREADコマンドを実行しています。アレイとホストの両方が16Gbpsで接続されており、十分なISL帯域幅（たとえば32Gなど）があるため、SAN内で輻輳が発生することはありません。

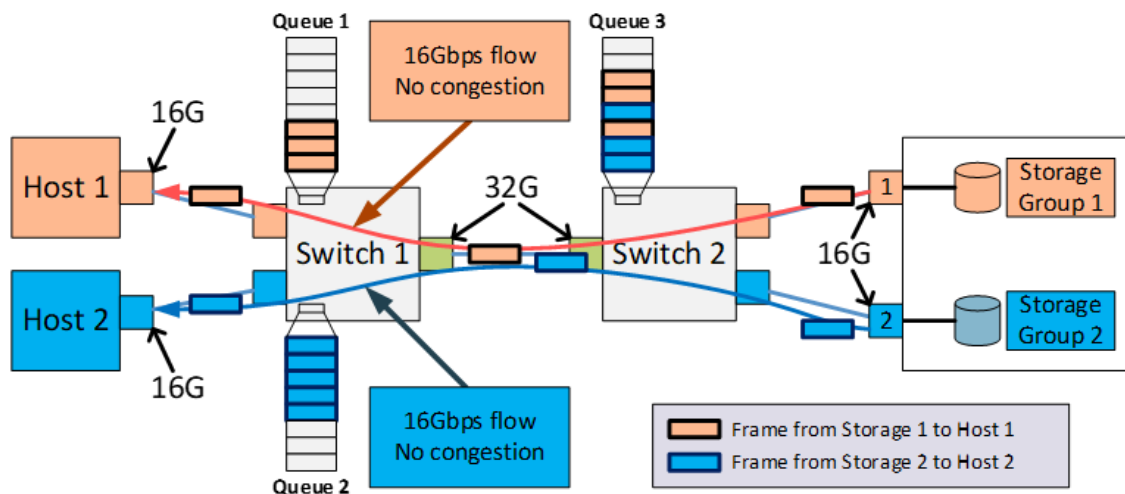


図 1 輻輳なし

図2は、オーバーサブスクリプションにより輻輳分散が発生しているSANの例を示しています。2つの数値の違いは、図3に示すように、ホスト1のインターフェイスが16Gbpsではなく4Gbpsで実行されるように設定されていることだけであることに注意してください。この処理が完了するとすぐに、接続されているHBAの速度（たとえば4G）よりも速い速度でアレイ インターフェイスがデータを転送すると、ホスト1は転送される速度でデータを受信できず、フレームのキューイングという影響がすぐに出ます。キュー1がいっぱいになると、輻輳はデータのソースのほうに分散します。

Error! No text of specified style in document.

ホスト1とホスト2の両方が同じスイッチ間リンク（ISL）を共有しているため、この輻輳はホスト2とストレージ2の間の「無害な流れ」に影響し、スループットは16Gbpsから4Gbpsに低下します。

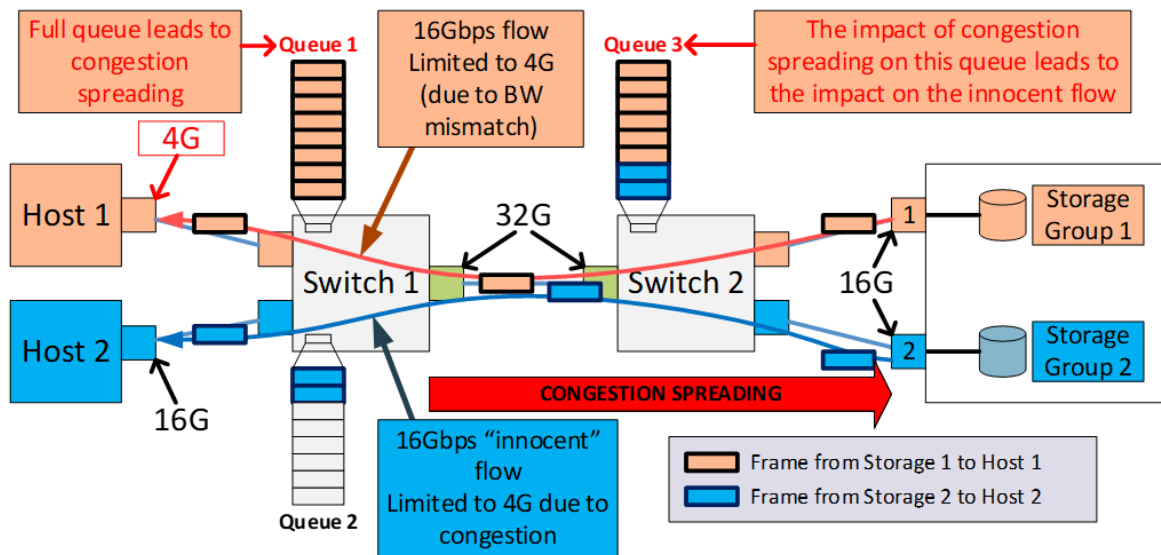


図 2 輻輳

輻輳と輻輳分散に関する追加情報については、『Networked Storage Concepts and Protocols』Techbookの「Congestion and Backpressure」のセクションを参照してください。

(<https://www.dell.com/collateral/hardware/technical-documentation/h4331-networked-storage-cncpts-prtcls-sol-gde.pdf>)。オーバーサブスクリプションは、輻輳分散の潜在的な複数の原因の1つにすぎないということに注意するのは重要です。その他の原因については、以降のセクションで説明します。

• 輻輳率（C比）

輻輳率またはC比は、輻輳分散が発生した場合の検出に役立つ計算値です。たとえば、図3は、4Gbpsの速度でデータを受信できるものの、16Gbpsでデータを転送できるストレージ インターフェイスからデータを受信しているホスト（ホスト1など）を示しています。

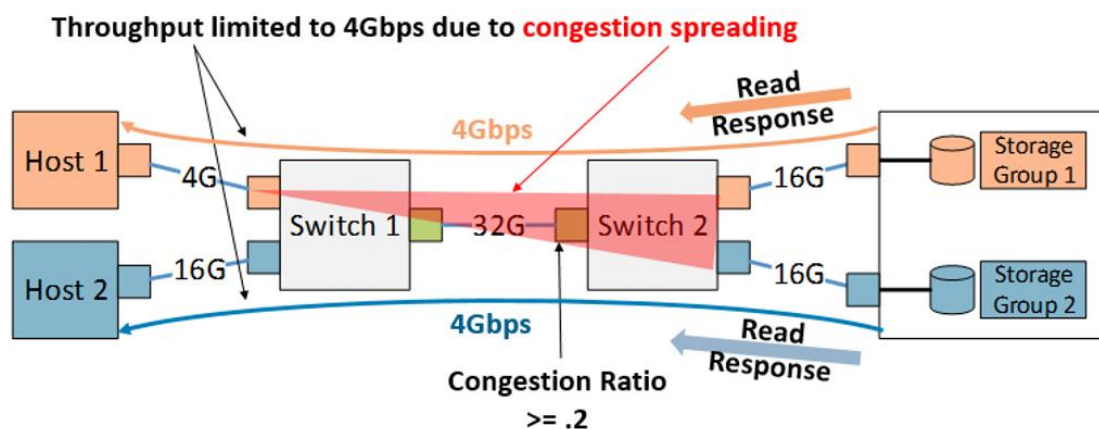


図 3 輻輳率

このような問題を検出してトラブルシューティングを行うことが困難な理由の1つは、スイッチ1の4Gインターフェイスの観点からは、何も問題がないように見えるということです。スイッチ インターフェイスは、リンクが対応できる速度でフレームを転送します。ただし、ストレージはデータをそのリンクが対応できる速度（たとえば16Gbps）で転送しており、アレイが転送する帯域幅が12Gbpsであると（16Gbpsには4Gbps不足）、データをキューイングする必要が生じます。通常、このキューイングはファブリック内で発生し、輻輳分散の原因になります。前述のように、輻輳分散の存在を検出するために使用できる方法の1つとして、輻輳率を計算する方法があります。これを行うには、「ゼロ転送クレジットで費やされた時間」カウンターを使用し、それを「転送されたフレーム」カウンターの値で除算して、数値を得ます（通常は0～1の範囲内）。この数が0.2より大きい場合には、輻輳があります。この数は、インターフェイスごとに計算する必要があります。そのため、この値を確認するにはプロセスをスクリプト化することをお勧めします。

4 オーバーサブスクリプションによる輻輳分散

次のケーススタディーは、オーバーサブスクリプションによる輻輳分散に基づいています。次の図4は、このケーススタディーのトポロジーを示しています。このケーススタディーでは、この問題を検出して発生を防ぐために現在利用可能なツールと手法について説明します。

注：輻輳分散は、検出と解決が非常に困難な問題です。これは主として、現行世代の管理ツールは、問題の解決方法に関する実際的なガイダンスの提供は言うまでもなく、問題の発生を明確に示すこともできないからです。そのため、この問題のトラブルシューティングにおいては、エンドユーザーは問題を理解し、利用可能な限られたデータから結論を引き出すために現在利用できるツールの使用方法を知る必要があります。

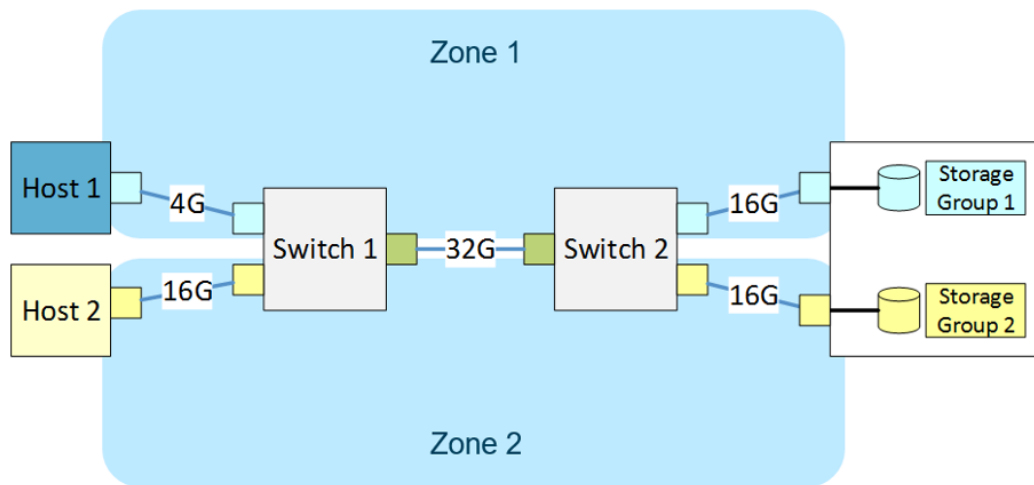


図 4 オーバーサブスクリプションのケーススタディーのトポロジー

- シナリオ：

ユーザー1は、さまざまなブロックサイズ、キューの深さ、I/OパターンでI/Oを実行しているホスト2（16G HBA）で、既存のアプリケーションを実行しています。このアプリケーションは、この環境で長時間実行されてきており、最近まで問題は発生していませんでした。今月の初めに、ユーザー2はテスト目的でホスト1（4G HBA）にアプリケーションをロードすることを決定しました。最初は、パフォーマンスとレイテンシーの点では、環境内ですべてが正常に動作していました。ただし、ユーザー1は最近、アプリケーションに関するパフォーマンスの問題に気付くようになりました。

- トラブルシューティングの概要：

一般的には、問題をトラブルシューティングするには、まず何がどのように実行されるか、および理想的な条件で処理されるにはそれらがどのように構成されるのかを理解する必要があります。SANには、エコシステムを構成する多くの可動部品があります。そのため、SANを構成する次の3つの主要なコンポーネントのプロファイルで構成される環境プロファイルを作成することが非常に重要です。つまり、アプリケーション、SANファブリック、ストレージです。

Error! No text of specified style in document.

環境内のさまざまなコンポーネントでこれらのベースライン プロファイルを作成すると、問題が発生した場合に簡単に特定できるようになります。これらのプロファイルは、1回限りのものではないことに注意してください。問題をトラブルシューティングするだけでなく、さらに拡張する計画を立てるためにも、環境のライフタイムを通じてベースライン データは常に収集する必要があります。

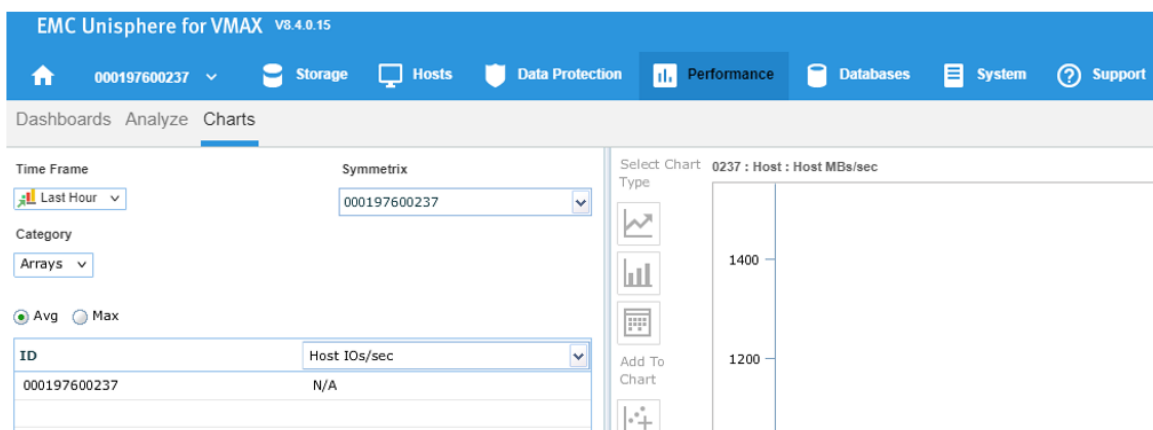
続くいくつかのセクションでは、前述のシナリオで説明したように問題が発生したときに根本原因を特定できるように、ストレージ アレイからこれらのベースライン統計情報を収集する必要があることを示します。

アプリケーション ベースライン

Dell EMC PowerMaxとVMAXを使用すると、パフォーマンス監視を有効にした場合に、履歴を振り返ることができます（この機能を有効にしてから最大で1年分）。これにより、変更を行う前に、アプリケーション プロファイルがIOPとレスポンス タイムの観点でどのような状態であったかを理解できるようになります。このアプリケーション ベース プロファイルがあると、生成されたチャートを使用して、問題がある可能性がある箇所を容易に判断できます。

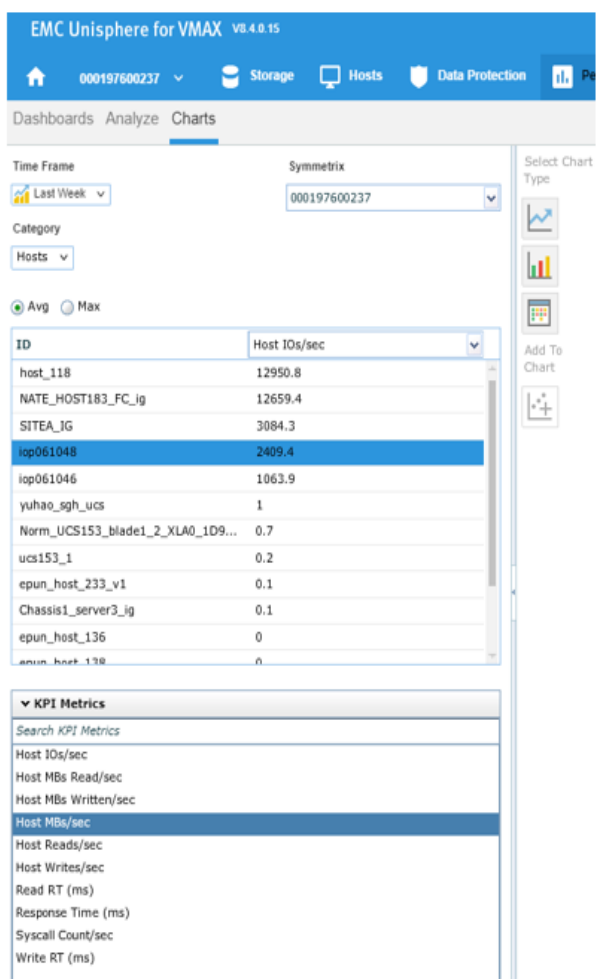
アプリケーション ベース プロファイル グラフの生成

Dell EMC Unisphere で、**[Performance]** > **[Charts]** をクリックします。



[Time Frame] を選択します。これは、パフォーマンスの問題に気付く前の任意の時間にします。**[Category]** ドロップダウンメニューで、**[Hosts]** > **[Hosts]** を選択します。

Error! No text of specified style in document.



対象のホストを選択します。**KPI メトリック**では、7種類の異なるチャートが生成されます。KPI メトリックごとにこのセクションを繰り返します。同時にすべてのメトリックをクリックすると、チャートが1つのグラフに表示されます。

- ホストIO/秒
- ホストMB/秒
- Host Reads/sec
- Host Writes/sec
- 読み取りRT (ミリ秒)
- レスポンスタイム (ミリ秒)
- 書き込みRT (ミリ秒)

図5の「ホストIOとMB/秒」では、ホストIOとMB/秒を確認します。これらのチャートからは、アプリケーションがいつでもどの期間IOを最も実行しておりリンクの全帯域幅を利用しているか、および最も低い点はどこかを確認できます。

注：この凡例では2つのホストが存在しますが、一方のホストではIOが一切実行されていないため、1つのホストのIOのみを表示しています。



図 5 ホスト IO と MB/秒

Error! No text of specified style in document.

図6に示すチャートは「読み取りと書き込み/秒」であり、アプリケーションが生成するIOのタイプの詳細を示します。これらのチャートに基づいて、アプリケーションIOの読み取りと書き込みの割合を判断できます。このケースでは、読み取り/書き込みの観点で、アプリケーションは約70/30であることを確認できます。



図 6 読み取りと書き込み/秒

Error! No text of specified style in document.

図7と図8は、トラブルシューティング時に使用できるおそらく最も役に立つチャートです。また、読み取りと書き込みの間のレスポンス タイムの詳細を示し、アプリケーションで発生しているレイテンシーを把握することができます。これは、レスポンス タイムが急増した場合に、以前のチャートを使用して急増と特定のイベントを関連づけることができるため、パフォーマンスの問題をトラブルシューティングする必要がある場合に非常に役立ちます。

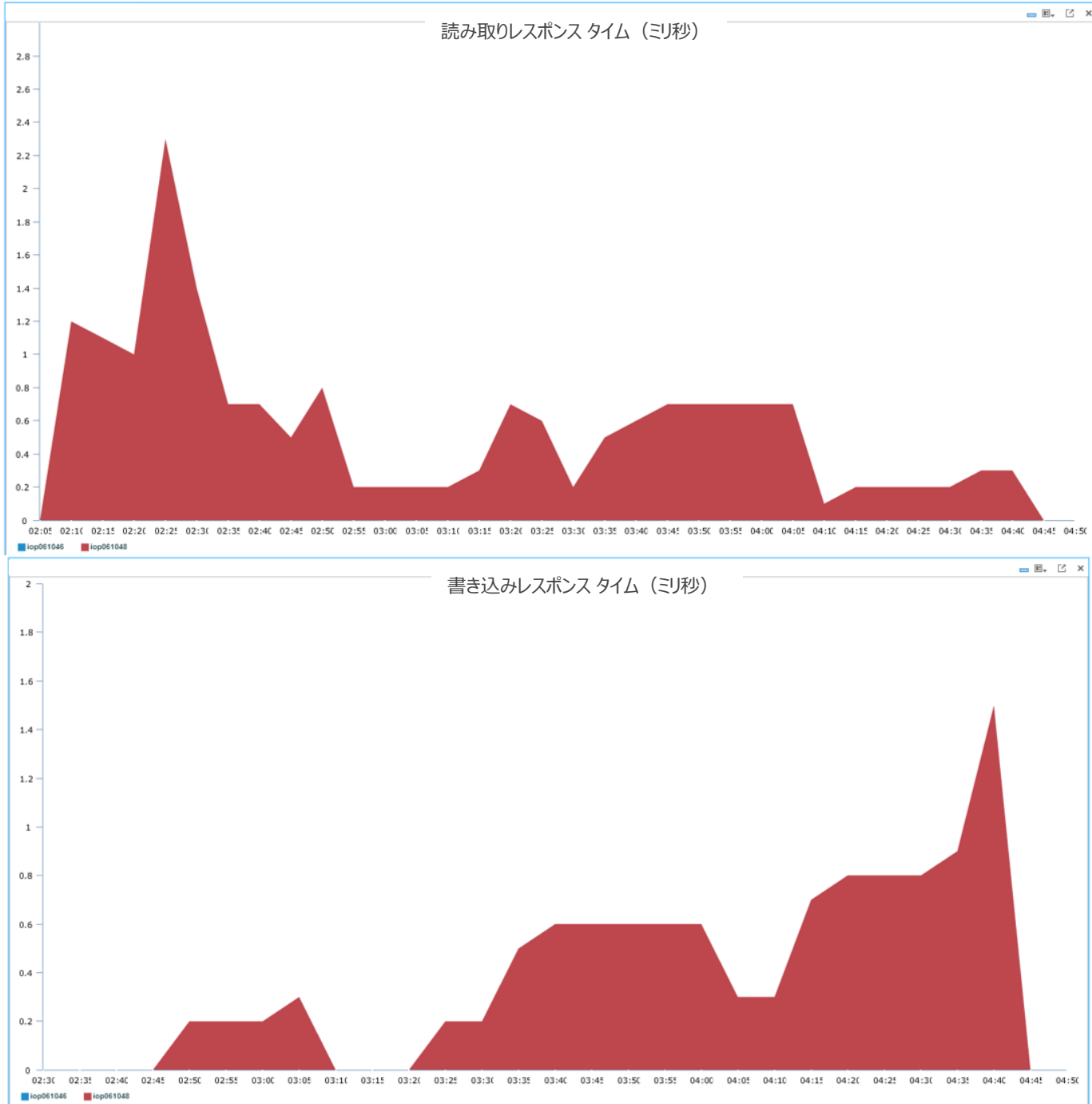


図7 読み取りと書き込みレスポンス タイム (ミリ秒)

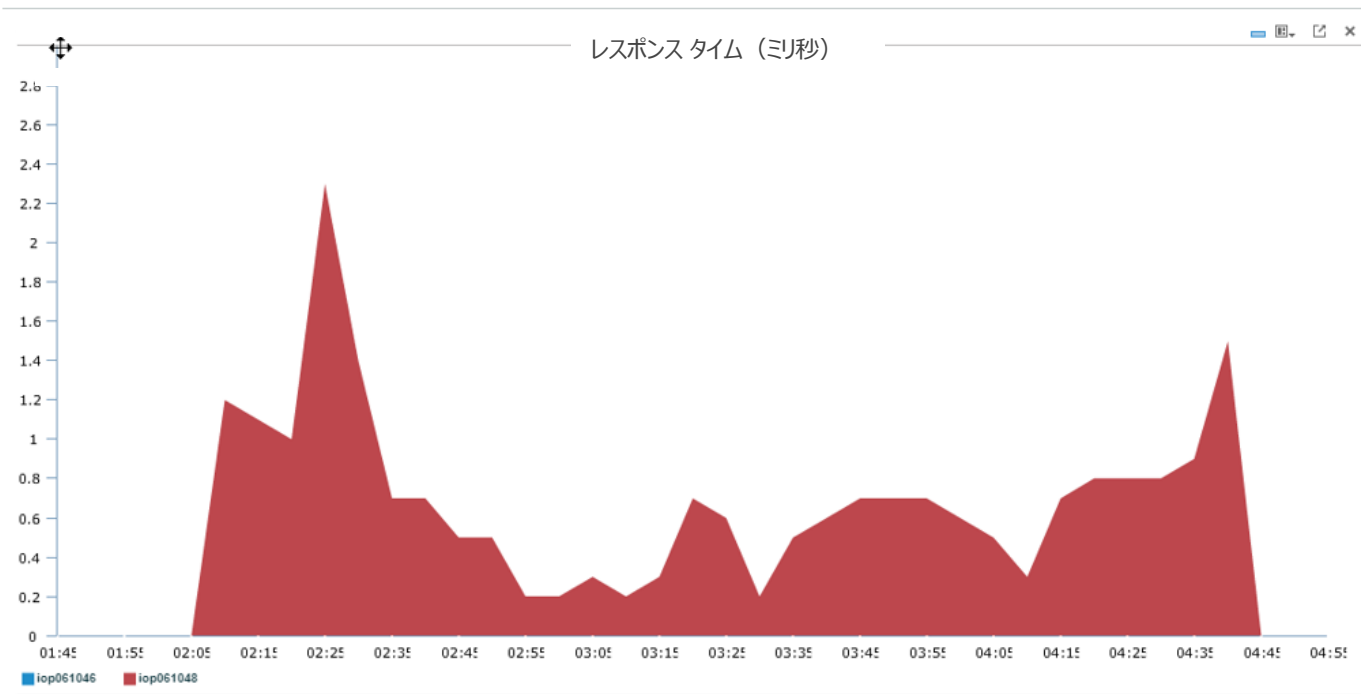


図 8 レスポンス タイム (ミリ秒)

ユーザー1のアプリケーションのアプリケーション プロファイルがあります。読み取り/書き込みは約70/30、レスポンス タイムは平均で約0.7ミリ秒、最大は2.3ミリ秒です。

[シナリオ セクション](#)で説明したとおり、最近環境内に追加したホストでパフォーマンスの問題が生じ始めました。それで、その問題をトラブルシューティングする方法を見てみましょう。

環境内でパフォーマンスの問題が発生しているため、これらのタイプの問題がいつ発生するかを判断するために役立つ、SANで利用可能な機能を実装する必要があります。

平均レスポンス タイムがどのようなものかをアプリケーション プロファイルに基づいて把握しているので、このパフォーマンスの問題が、予期されるレスポンス タイムよりも高いことが分かります。

Connectrix SAN 輻輳分散アラート

このセクションでは、SANスイッチ側から報告された輻輳イベントのタイプを確認します。
動作条件に従って、環境内でこれらの機能の有効化が完了していることを確認してください。

4.1.1 Brocade

1. ダッシュボードに、少なくとも**最上位ポートトラフィック**と**BBクレジットゼロ**があることを確認します。左上隅にあるレンチをクリックすると、それらが追加されていない場合に追加できます。

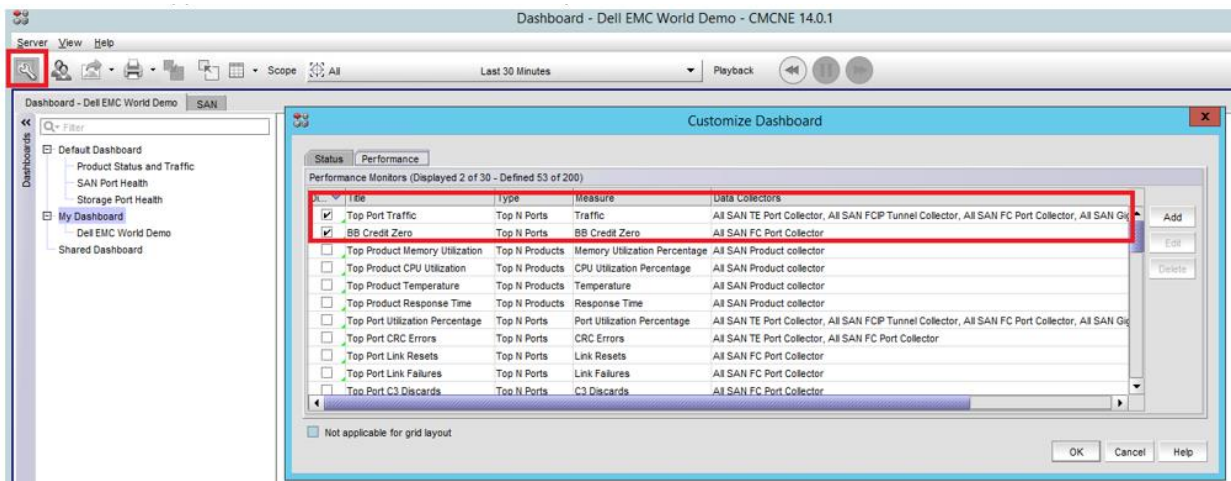


図 9 CMCNE ダッシュボード

2. オーバーサブスクリプションによる輻輳分散が発生すると、図9の「CMCNE Dashboard」の例のように、一般的にはCMCNEダッシュボードに次のアラートが表示されます。
 - a. 高使用率 F ポート
 - b. BB クレジットゼロ

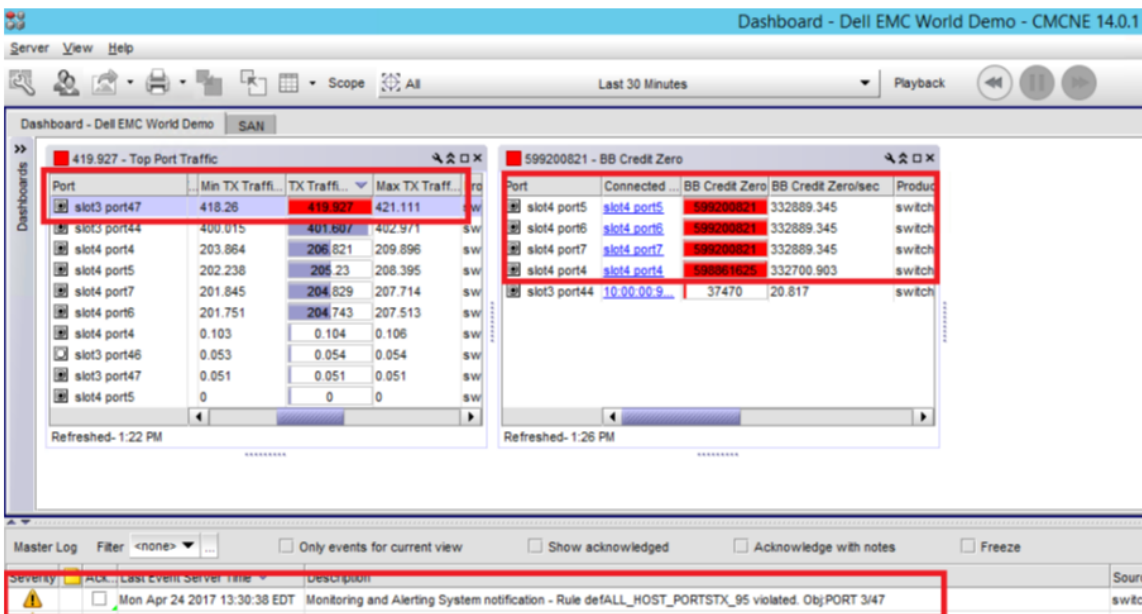
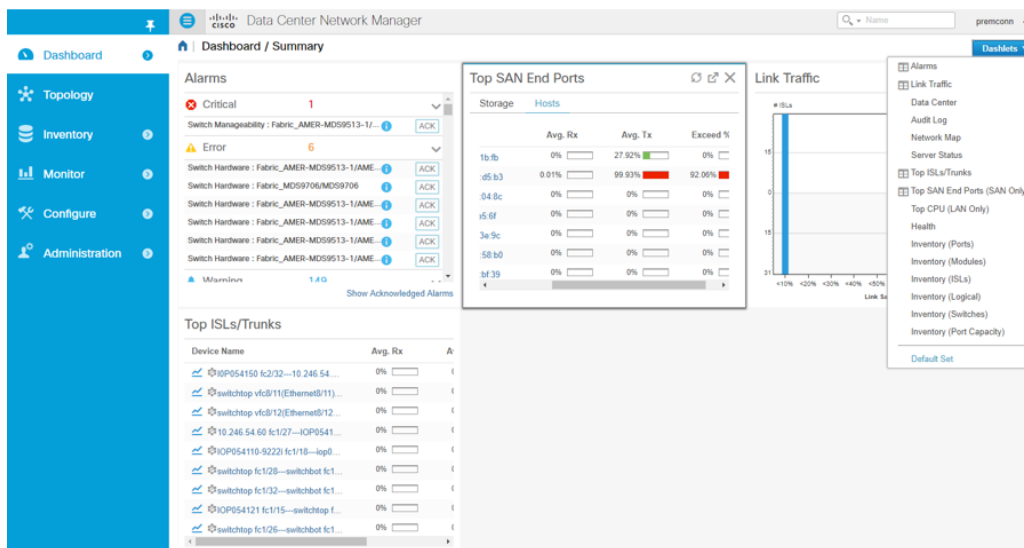


図 10 アラートを表示する CMCNE ダッシュボード

- この2つのイベント（高使用率Fポートと高いバッファツーバッファ クレジットがISL上でゼロになる）を組み合わせると、調査が必要になる潜在的なパフォーマンスの問題があることを示している可能性があります。確認のステップについては、「修正」の章を参照してください。

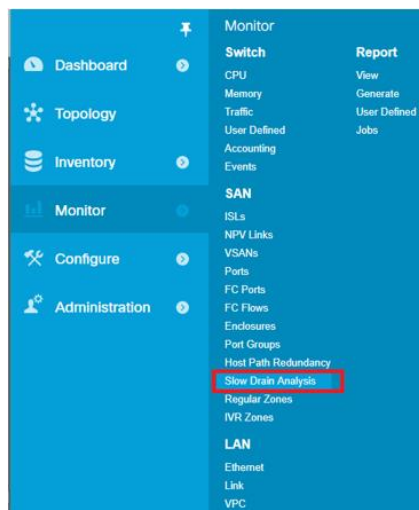
4.1.2 Cisco

- DCNMダッシュボードでは、**最上位のSANエンドポート**をダッシュレットとする必要があります。そのようになっていない場合は、ドロップダウンメニューから追加することができます。**最上位SANエンドポート**には、使用率が90%を超えていると報告されるデバイスが表示されます。DCNMにはデフォルトの閾値が設定されており、デフォルトの使用率を超え始めると、ポートには黄色または赤色のフラグが立てられます。このアラートは、必ずしもSANにパフォーマンスの問題があることを意味するわけではありません。ファブリック内の他のアラートも確認する必要があります。



- 高使用率Fポートが表示される場合は、低速ドレイン解析ツールを実行します。

[Monitor] > [SAN] > [Slow Drain Analysis] をクリックします。



Error! No text of specified style in document.

3. 低速ドレイン解析ツールを10分間実行します。レポートの作成が完了すると、レポートの実行中にTxWaitカウンターが大幅に増加しています。これらのアラートと高使用率Fポートの組み合わせにより、オーバーサブスクリプションによるSANの輻輳が発生していることが示されます。

Interface	Speed	ConnectTo	Type	Level 3			Level 2		Level 1		
				TxCreditLoss	TxLinkReset	RxLinkRe...	TxTimeoutD...	TxDiscard	TxAvg10...	RxB2Bto0	TxB2Bto0
fc2/36 ...	16Gb	IOP054151 fc2/36 (port-channel6)	Switch	0	0	0	0	0	0	41429192	37.3350
fc2/37 ...	16Gb	IOP054151 fc2/37 (port-channel6)	Switch	0	0	0	0	0	0	30150028	28.8393
fc2/38 ...	16Gb	IOP054151 fc2/38 (port-channel6)	Switch	0	0	0	0	0	0	26941217	25.6649

この2つのイベント（高使用率Fポートと高いバッファツープバッファクレジットがISL上でゼロになる）を組み合わせると、調査する必要がある潜在的なパフォーマンスの問題があることを示している可能性があります。確認のステップについては、「[修正](#)」のセクションを参照してください。

UNISPHERE 輻輳分散アラート

このセクションでは、Unisphere for PowerMAXとfor VMAXを使用して、SANスイッチのイベントをストレージアレイに関連づける方法について説明します。

「動作条件」セクションに従って、環境内でこれらの機能の有効化が実行済みであることを確認してください。

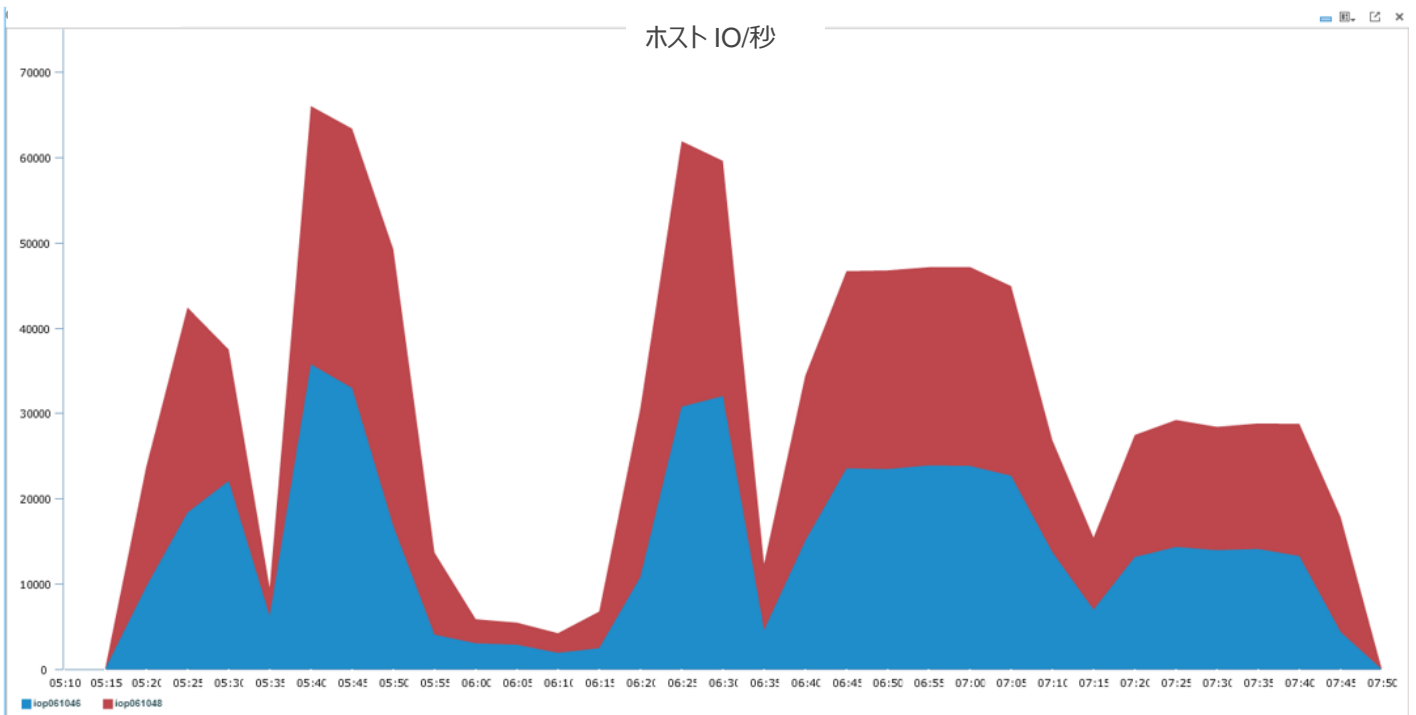
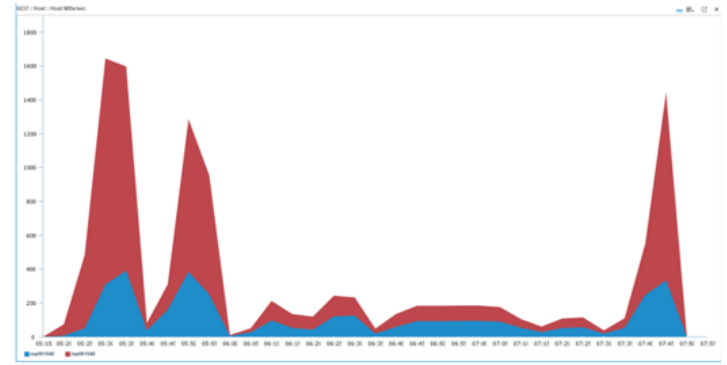
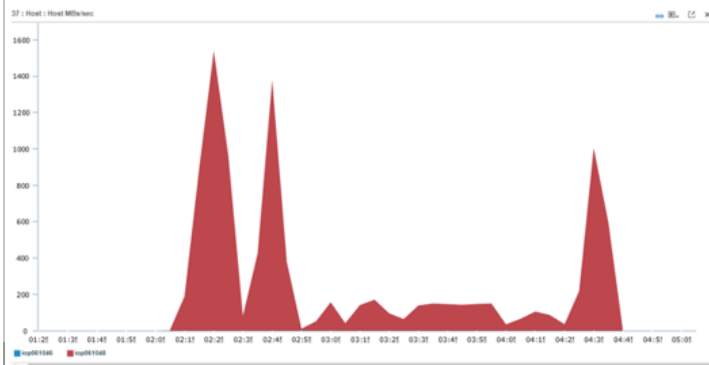
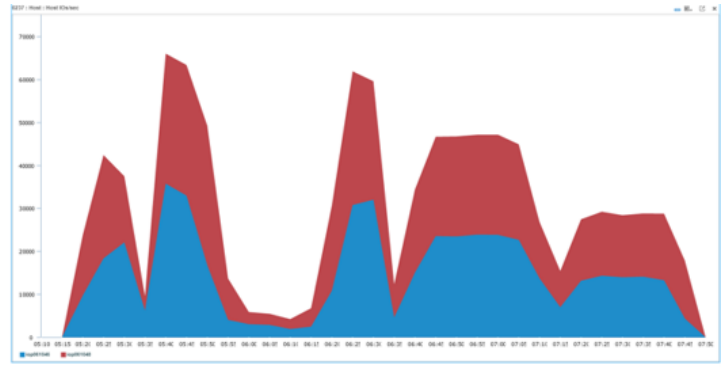
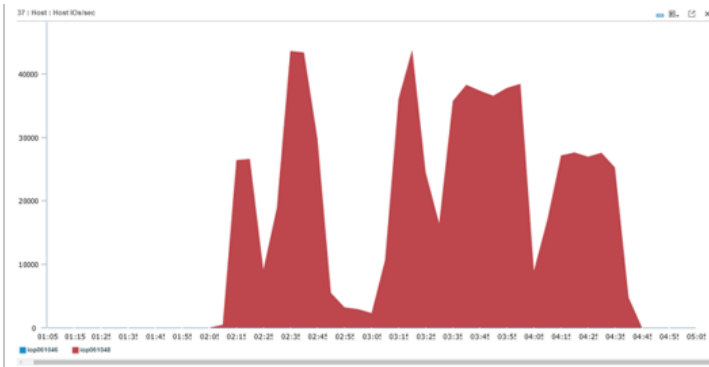
1. 「[アプリケーションベースプロファイルグラフの生成](#)」のステップを使用します。同じ7つのチャートを生成して、ユーザー2のホストをミックスに追加します（これは、パフォーマンスの問題が発生する前の、環境内での最近の変更の1つであったため）。データを確認します。

アプリケーションベースプロファイルが読み取り/書き込みは70/30、レスポンスタイムは平均で約0.7ミリ秒、最大は2.3ミリ秒であることに留意します。

次の図11では、IOとMB/秒を比較しても、問題を示しているようには見えません。実際には、元のベースラインアプリケーションチャートと比較した場合、より多くのIOPを実行しています。

また、ラインレートに近づいているいくつかのポイントがあります（下にハイライト表示されている）。これらのポイントについては後で取り上げます。

Error! No text of specified style in document.



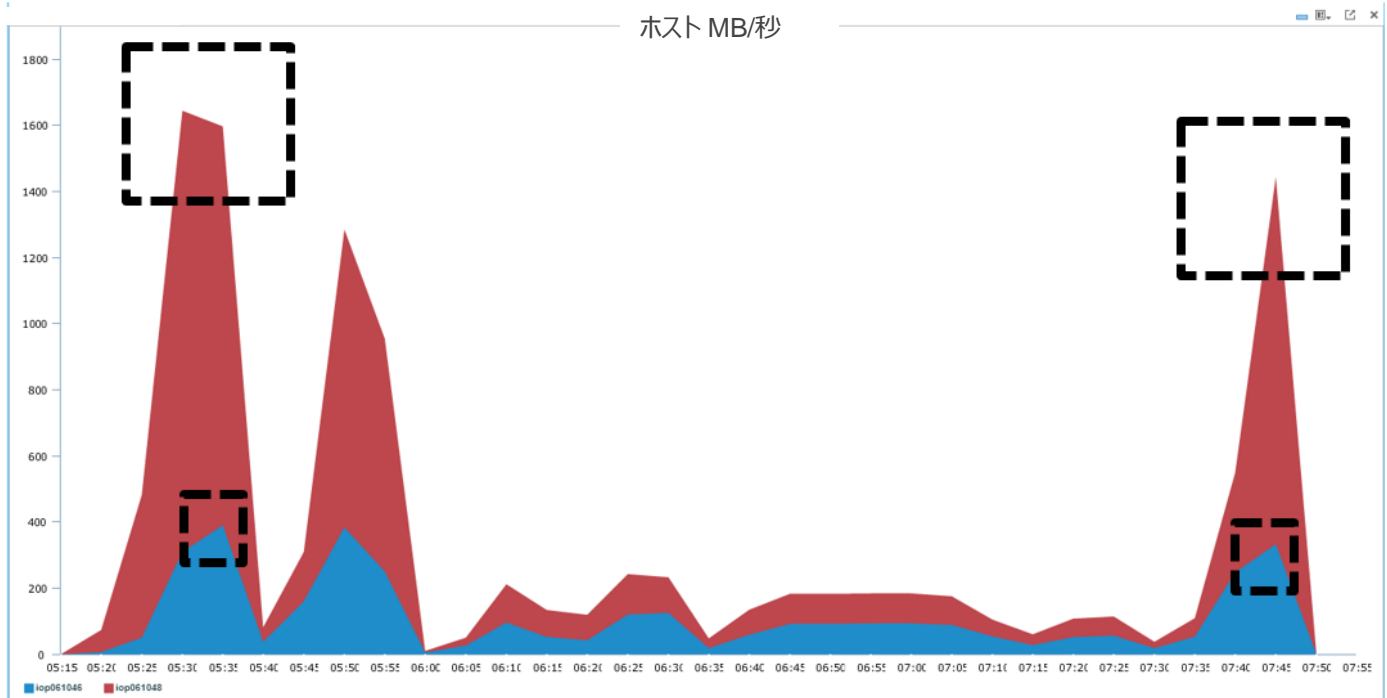


図 11 ホスト IO と MB/秒

図 12 は、2つのサーバー間の読み取り/書き込みの比較を示しています。この時点では、サーバー間の IO プロファイルにはそれほど違いはありません。

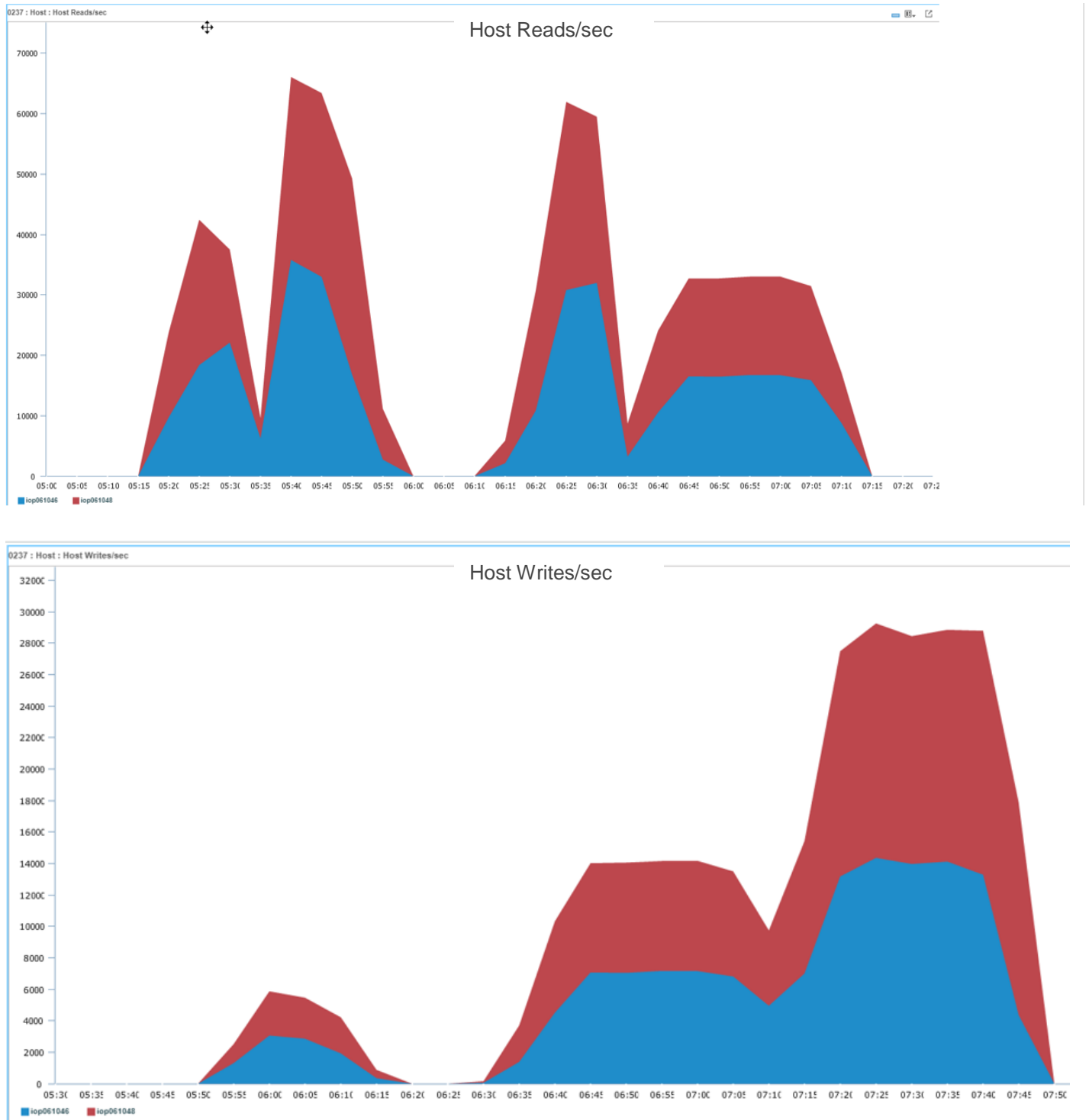
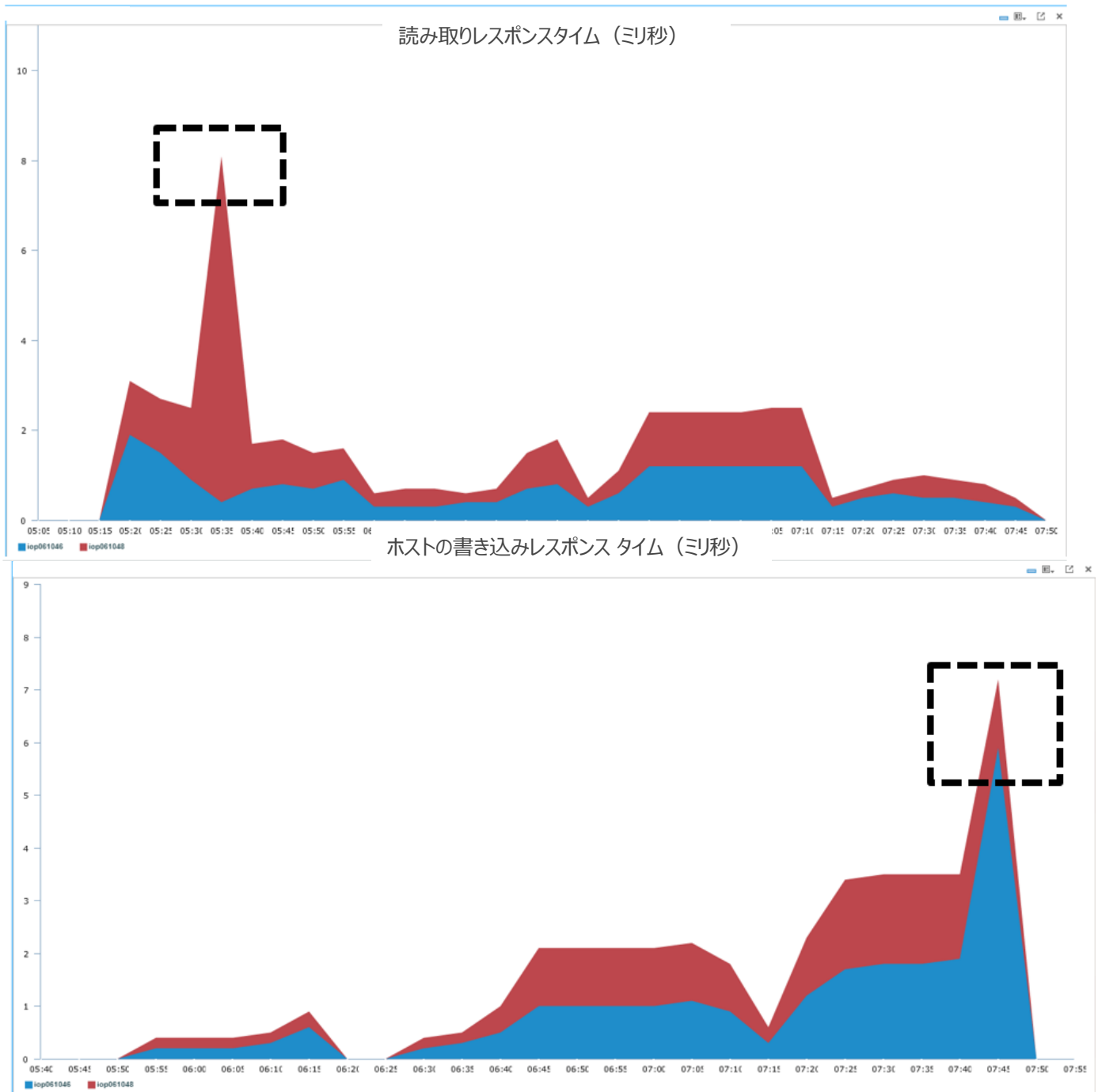


図 12 ホストの読み取りと書き込み/秒

Error! No text of specified style in document.

図13は、最も役立つ情報を示しています。アプリケーション プロファイルでは、レスポンス タイムは平均で約0.7ミリ秒、最大は2.3ミリ秒であることに留意します。次のチャートでは、レスポンス タイムに大きな急増があります。レスポンス タイムは8ミリ秒で、全体的な平均レスポンスタイムも増加しています。

図11に戻ると、これらの高いレスポンス タイムは、両方のサーバーがライン レートに近づいたときに関連しています。通常、ファイバー チャネルでは、リンクが飽和状態になるには大規模なブロックIO（128Kより大きい）が必要です。



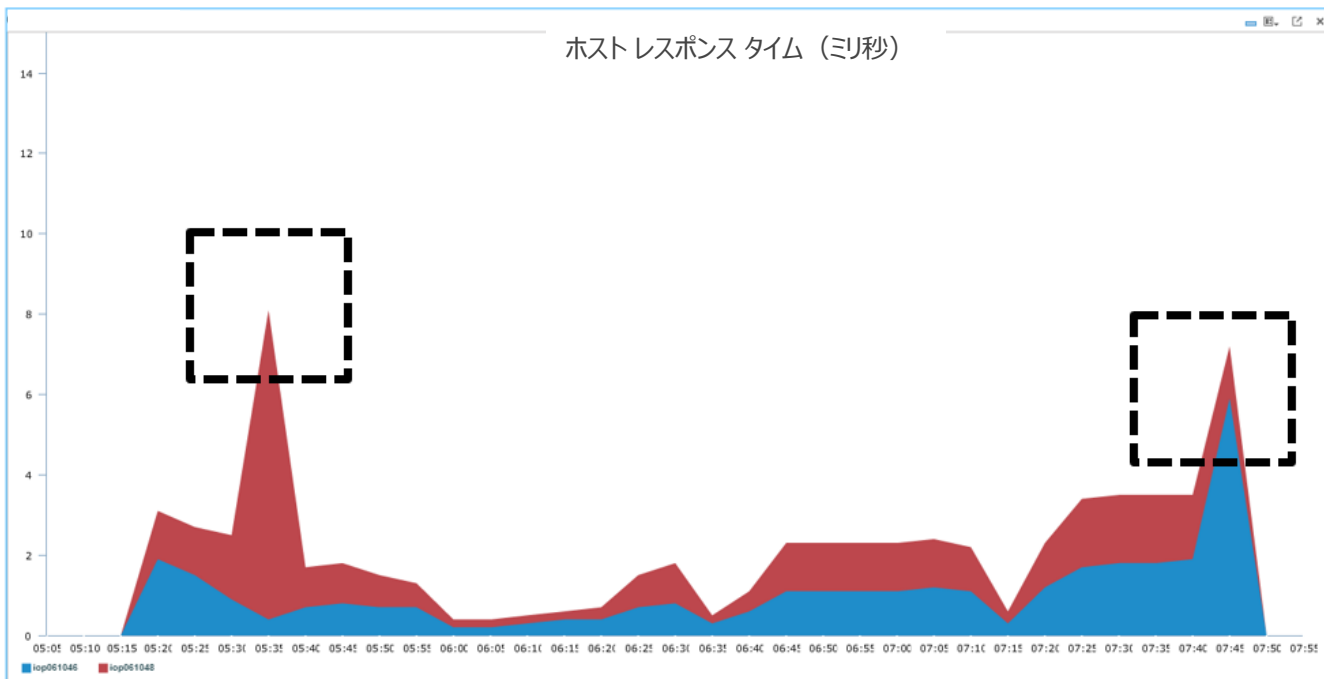


図 13 ホストの読み取り/書き込みレスポンス タイム (ミリ秒)

まとめ

このケース スタディーでこれまで説明したすべての情報を要約します。

- **Connectrix SAN :**

1. SANは、ゼロになるバッファ間クレジットの数が多いと報告しています。
2. F ポートではトラフィック使用率が高くなっています。

- **Dell EMC PowerMAX と VMAX :**

1. リンク使用率がフルのときの高いレスポンス タイム

前述のとおり、帯域幅の不整合による輻輳は、現在利用可能なツール セットで検出したり確認したりするのは非常に困難です。ただし、前述のアラートに基づいて、帯域幅の不整合や大きなブロックの読み取り/書き込みにより問題を推定できます。これは、リンク使用率がフルのときのレスポンス タイムが高いことを示しています。

この問題を検出するもう1つの方法は、**輻輳率**を使用することです。現時点では、環境内でこれを手動で計算する必要があります（またはスクリプト化することもできます）。C比が0.2を超えると、SAN環境で発生するバックプレッシャーが原因で輻輳が発生します。C比は、低速ドレインを最初に示したものです。

5 修正

予防

この特定のケーススタディー（オーバーサブスクリプションによる輻輳分散）については、この問題が発生しないようにするために、ご使用の環境に導入できるオプションがいくつかあります。

帯域幅比

- SANを確認する場合は、低速で実行されているデバイスを特定し、アプリケーション トラフィック プロファイルのタイプを把握することができます。帯域幅の不整合があるというだけでは、必ずしも問題があることを意味するわけではありません。
- ファブリックをエンド ツー エンドで確認して、すべてのエンド デバイスが同じリンク速度で実行されていることを確認します。
- ISLで十分な帯域幅を確保します。確かな経験則として、ISLの合計帯域幅は、可能であればファブリック内のストレージ帯域幅の合計以上にしてください。
- 以下の図14に示すように、すべてのコンポーネントをエンド ツー エンドでアップグレードすることで、SAN全体を最新化できます。このアプローチの欠点は、大規模な環境では、エンド ツー エンドからのゼロ オーバー サブスクリプションが現実的ではないということです。さらに、非常にコストがかかる可能性があります。したがって、特定のホスト、スイッチ、ストレージ接続のアップグレードのみに注意を集中してください。

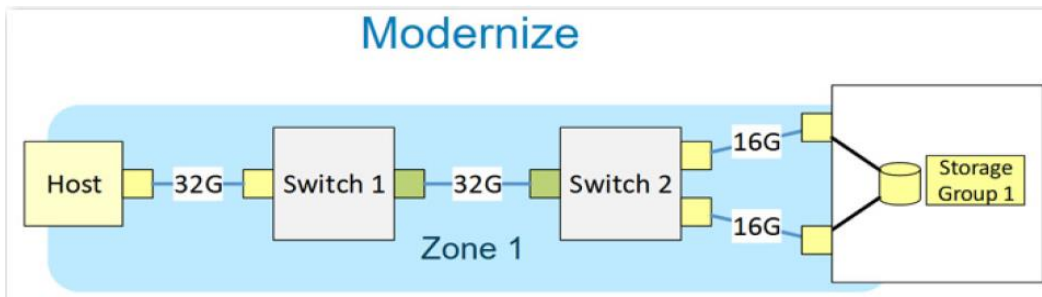


図 14 最新化

Error! No text of specified style in document.

図 15 のように、もう 1 つの方法は再ゾーニングです。

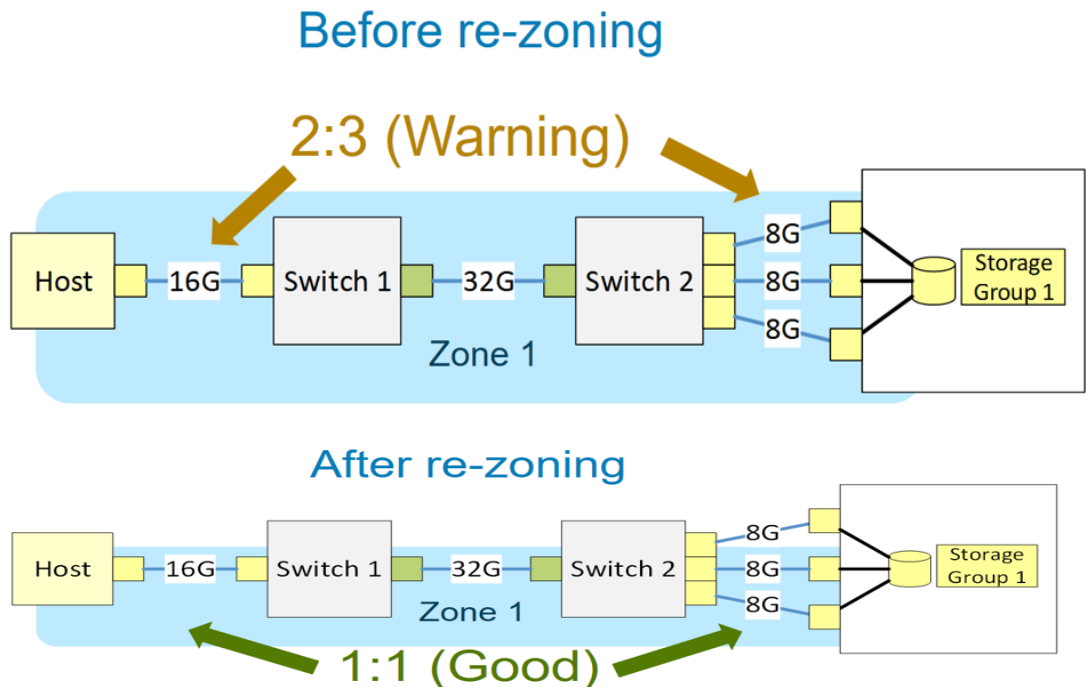


図 15 再ゾーニングの前後

帯域幅制限の実装

Dell EMC VMAXとDell EMC Unityプラットフォームでは、ストレージグループ（VMAX）またはLUN（Unity）の帯域幅制限が作成されます。前述のケーススタディーでは、オーバーサブスクリプションによる輻輳分散がありました。帯域幅制限を実装すると、次の図16に示すようにパフォーマンスが回復しました。これは、ストレージグループのUnisphereを通じて直接実行できます。

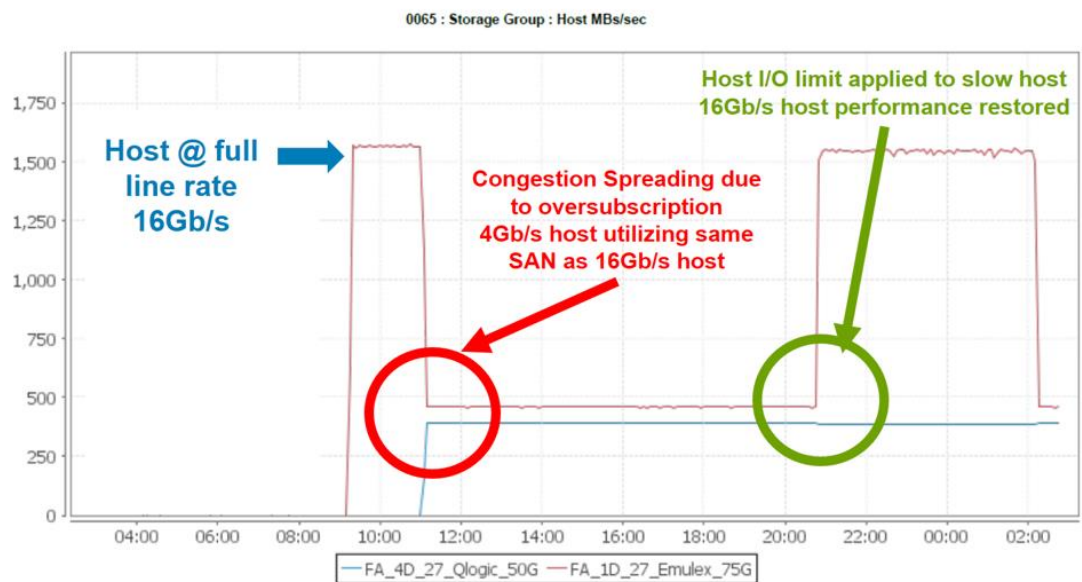


図 16 適用されるホスト I/O の上限値

I/Oの上限値を使用する場合、クラスターではこれがうまく機能しないことに注意することは重要です。たとえば次の図17を参考に見ましよう。バックプレッシャーを引き起こしている4Gbホストにホスト制限が適用されると、アレイは4Gbに送信されるデータ量の制限を（設定されたIO制限に基づいて）開始します。これによりバックプレッシャーの問題はまとめて解消され、他のフローはフルラインレートで動作するようになります。

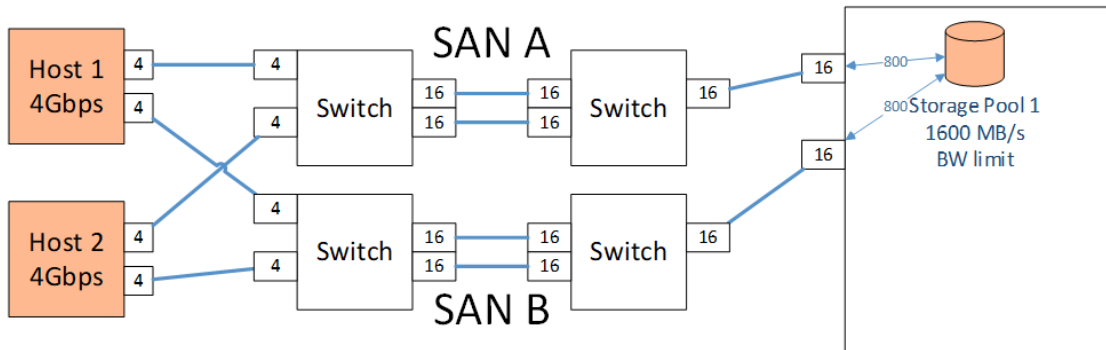


図 17 クラスターでの I/O の上限値

この例では、4Gb/秒で実行する2つのホストがクラスター内にあります。これらはクラスター内にあるため、どちらのホストも各ファブリックを介してボリュームにアクセスします。つまり、帯域幅I/Oの制限として1600MB/秒（各FAに対して800MB/秒）を設定する必要があるということです。ただし、このアプローチでは、1つのHBAが800MB/秒すべてを利用するのを防ぐことはできません。

- 分離

この問題を回避するもう一つの方法は、高速トラフィックから低速トラフィックを分離し、専用ISLを使用することです。これを実現するには、仮想ファブリック（Brocade）またはVSAN（Cisco）を作成します。

これについては図18で示しています。このアプローチの欠点は専用のポートが必要だということですが、これにより低速のトラフィックが高速のトラフィックに影響を与えることはなくなります。Brocadeで仮想ファブリックを有効化するには、スイッチ全体を再起動する必要があるため、ダウンタイムが必要になります。Ciscoの別のVSANにポートを移動する場合、移動元のエンドデバイスのみが影響を受けます。

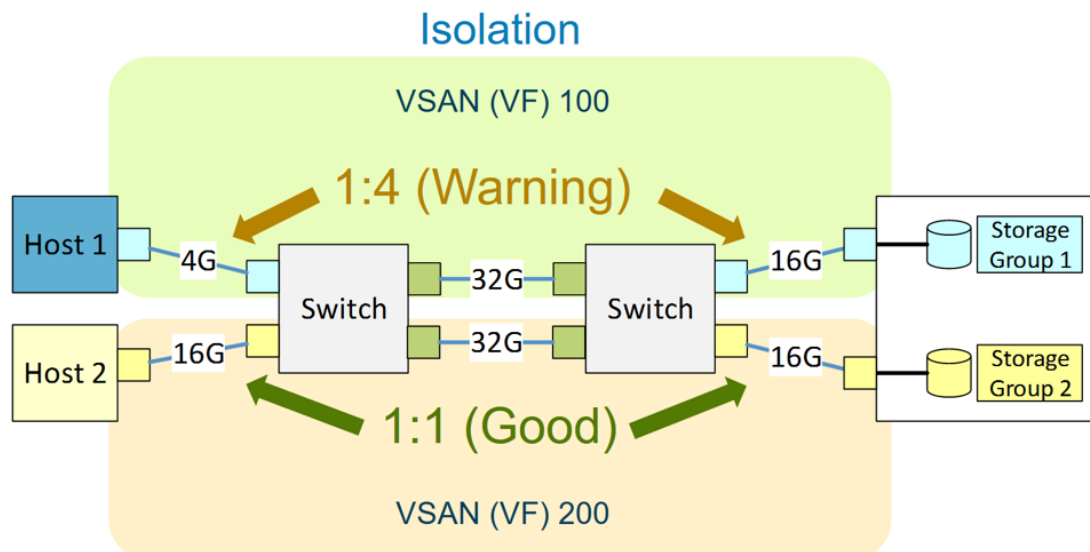


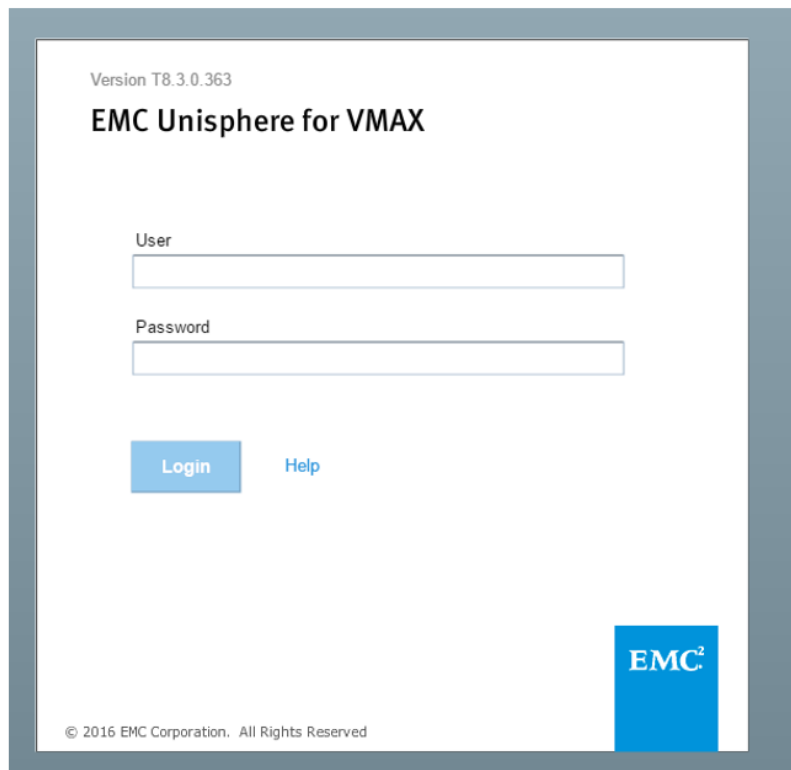
図 18 分離

6 付録

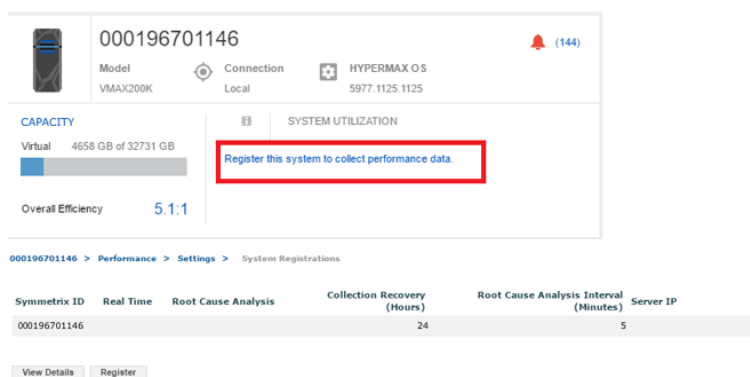
パフォーマンス監視の有効化

このセクションでは、Unisphere for VMAXでのパフォーマンス監視を有効化してデータを確認するステップについて説明します。

1. Unisphere GUI にログインします。

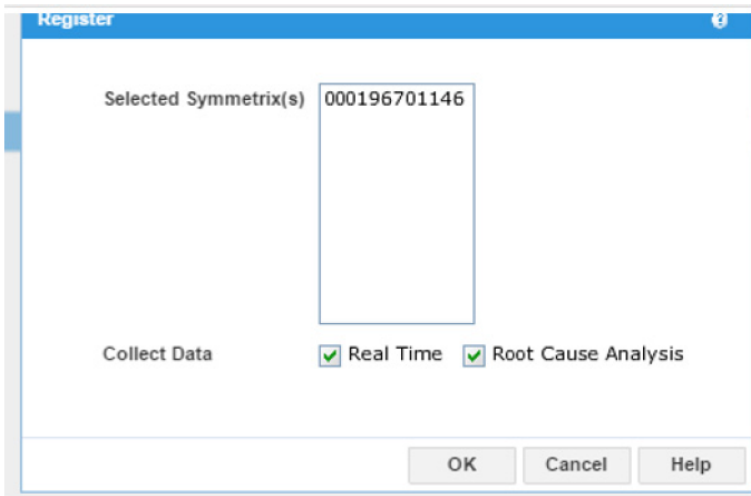


2. アレイがパフォーマンス データを収集するように登録されていることを確認します。登録されていない場合は、アレイを登録します。



Symmetrix ID	Real Time	Root Cause Analysis	Collection Recovery (Hours)	Root Cause Analysis Interval (Minutes)	Server IP
000196701146			24	5	

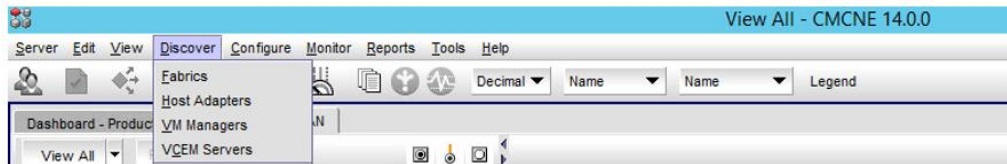
Error! No text of specified style in document.



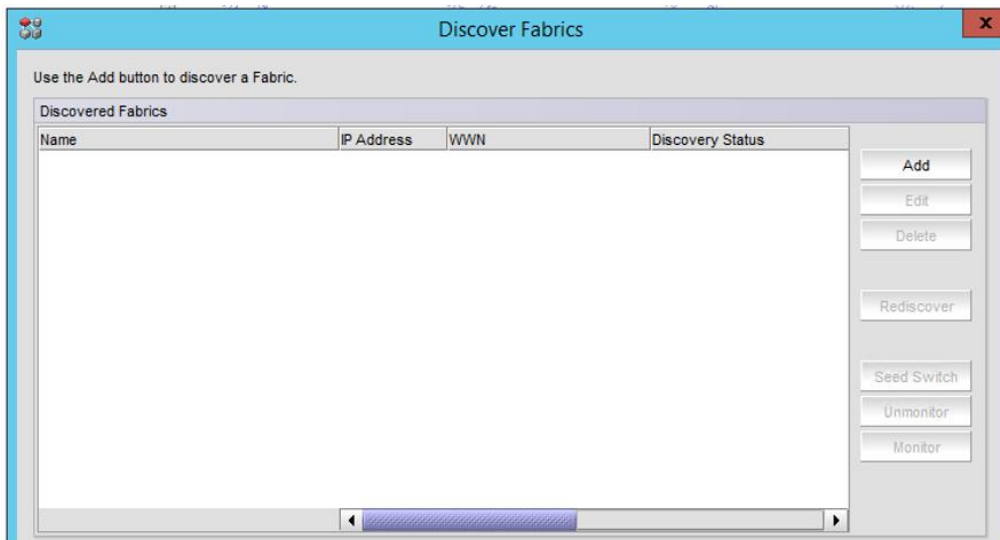
6.1.1 Brocade

- ファブリックの検出

1. CMCNE Serverにログインし、**[Discover]** > **[Fabrics (SANnav?)]** をクリックします。

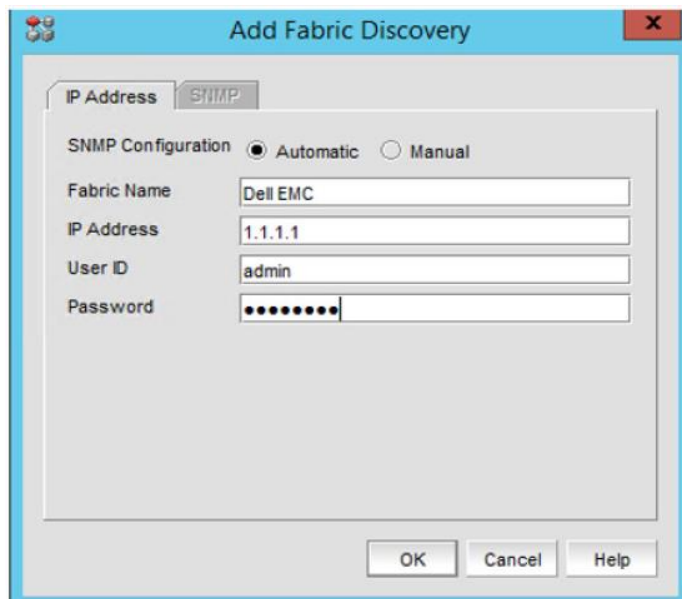


2. 新しいウィンドウで、**[追加]** をクリックします。



3. ファブリック内のいずれかのスイッチに必要な情報を入力します。CMCNE は、そのファブリック内のすべてのスイッチのユーザー名とパスワードが同じであると見なし、すべてのスイッチを自動的に検出します。

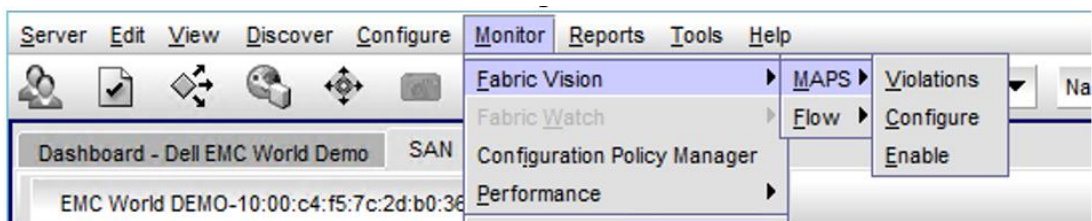
Error! No text of specified style in document.



4 その他のすべてのファブリックでこのセクションを繰り返します。

- マップと FPI の有効化

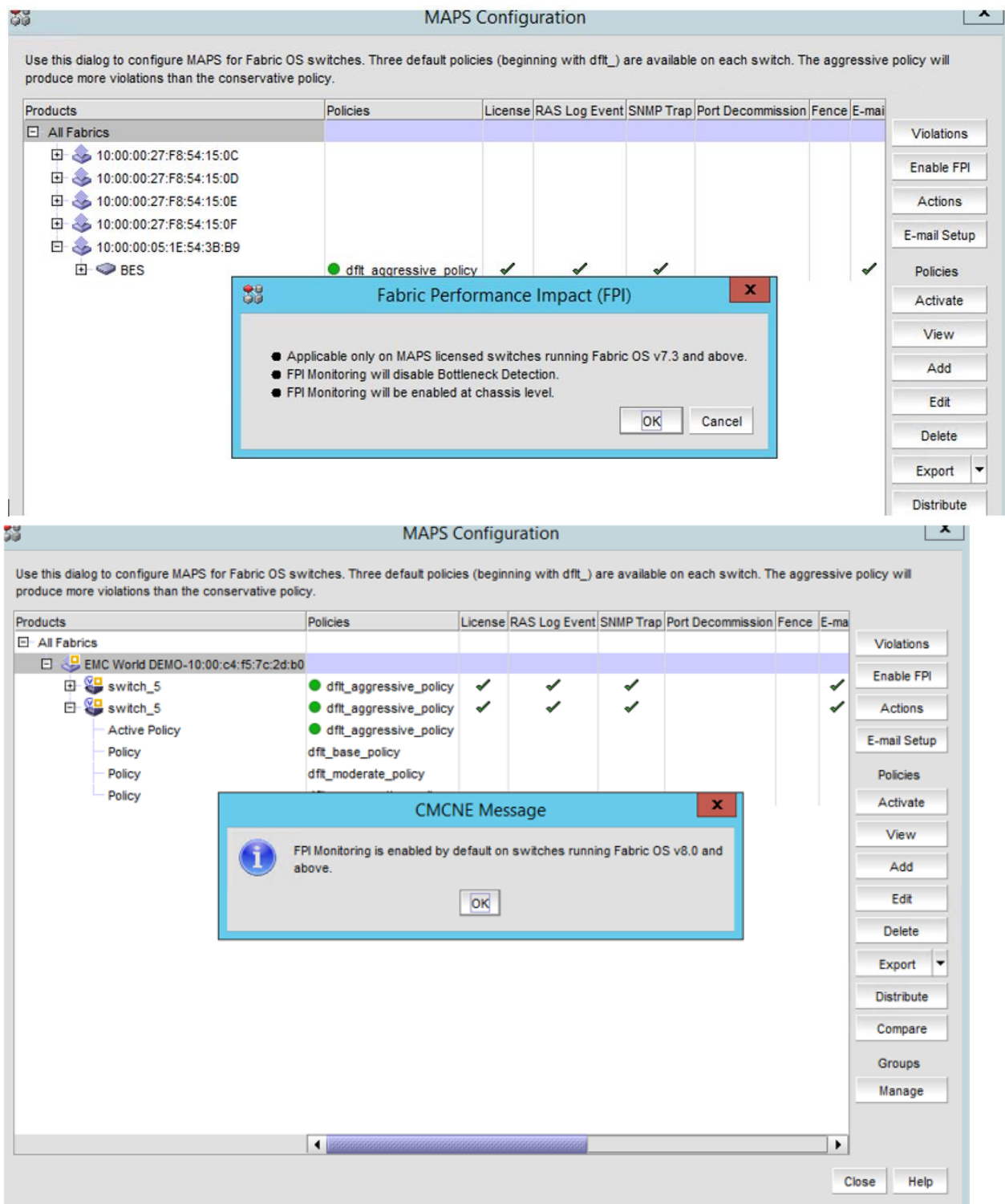
1. [Monitor] > [Fabric Vision] > [MAPS] > [Configure] をクリックします。



2. ファブリックをハイライト表示し、FPI を有効にします。

注：FPI は、FOS 8.0 以降を実行するスイッチではデフォルトで有効になっています。

Error! No text of specified style in document.



3. このメニューから、ファブリック内の各スイッチを構成し、目的のマップ ポリシーを設定することができます。

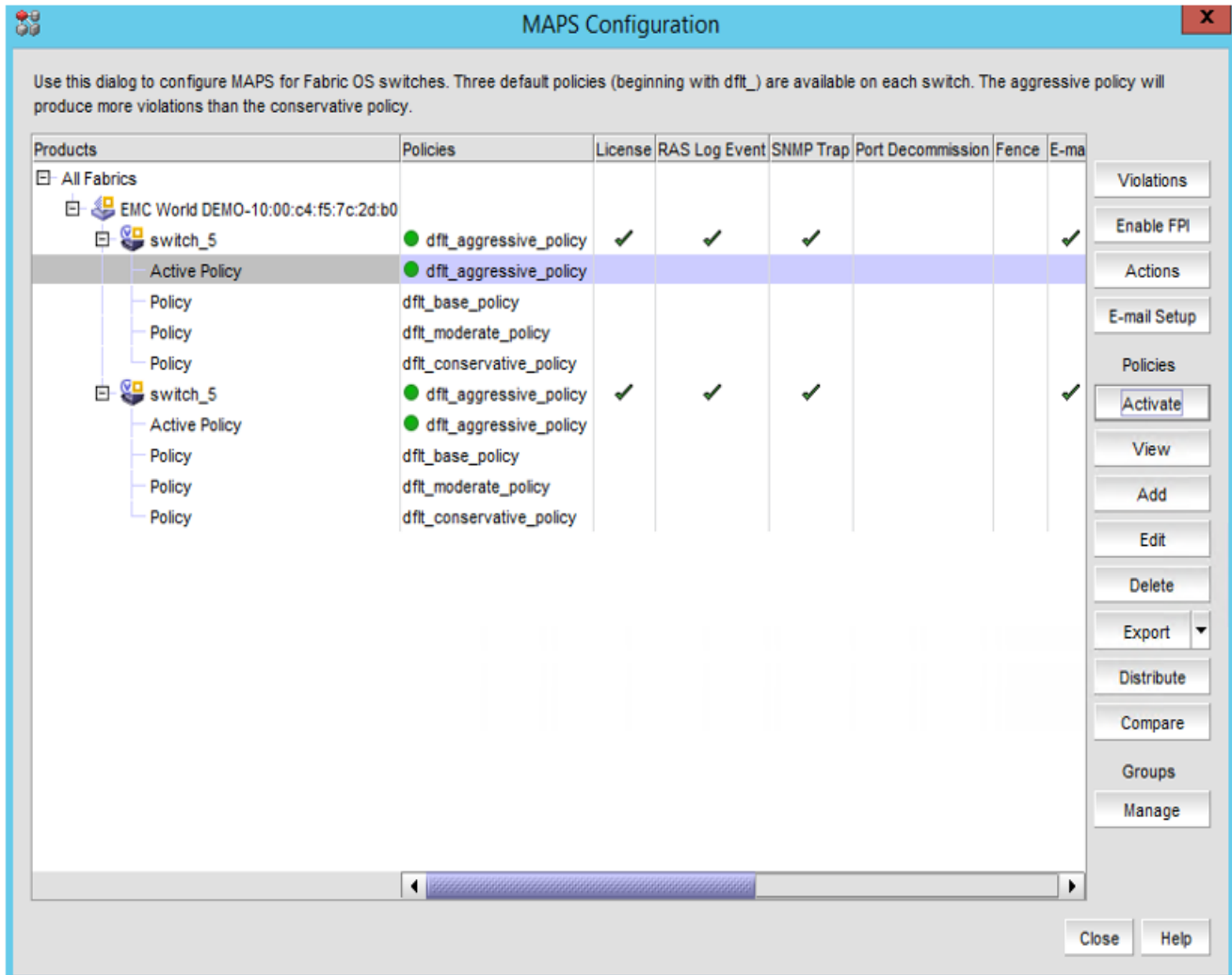
注： CMCNE は、クローンを作成して編集できる事前定義済みのポリシーを提供します。デフォルトのポリシーは編集できません。各ポリシーと設定の詳細については、『MAPS 管理ガイド』を参照してください。

Error! No text of specified style in document.

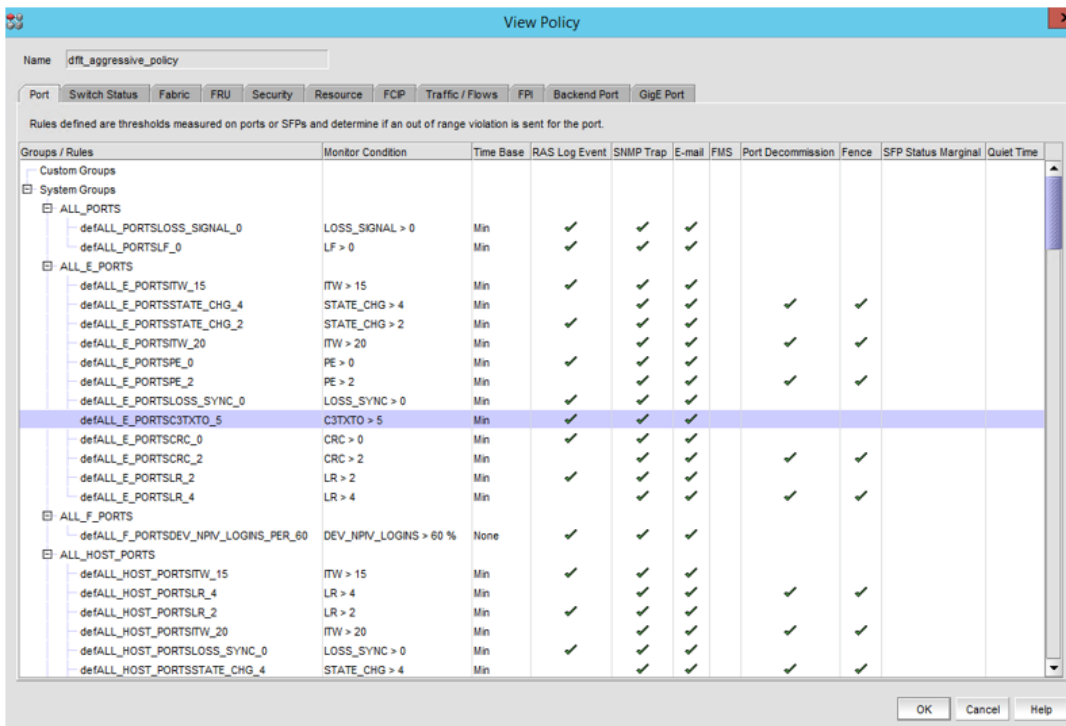
このケースでは、デフォルトの積極的なポリシーをアクティブ化します。これを行うには、**dflt_aggressive_policy**をハイライト表示し、**[Activate]** をクリックします。このステップは、ポリシーを有効化するファブリック内のすべてのスイッチで繰り返す必要があります。現在のところそれをファブリック全体で一括して有効にすることはできません。

まず積極的なポリシーをアクティブ化することで、ファブリック内の問題をすぐに把握できます。その後、受け取るアラートが多すぎる場合は、調整したり他のポリシーを使用したりできます。

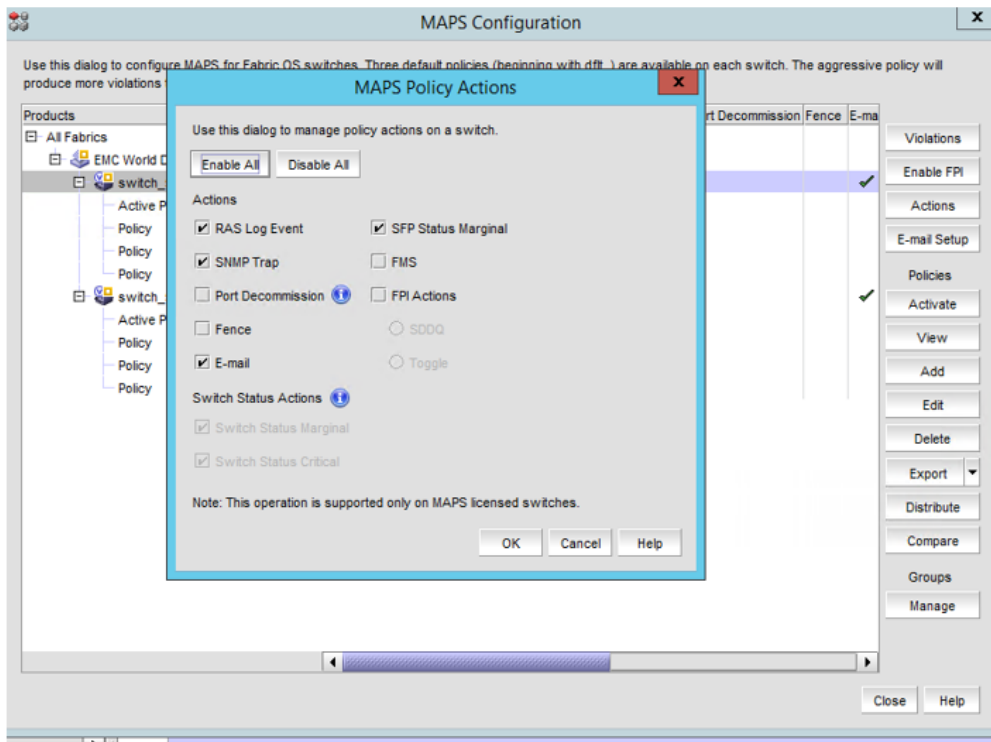
[View] をクリックすると、各イベントの閾値を確認できます。



Error! No text of specified style in document.

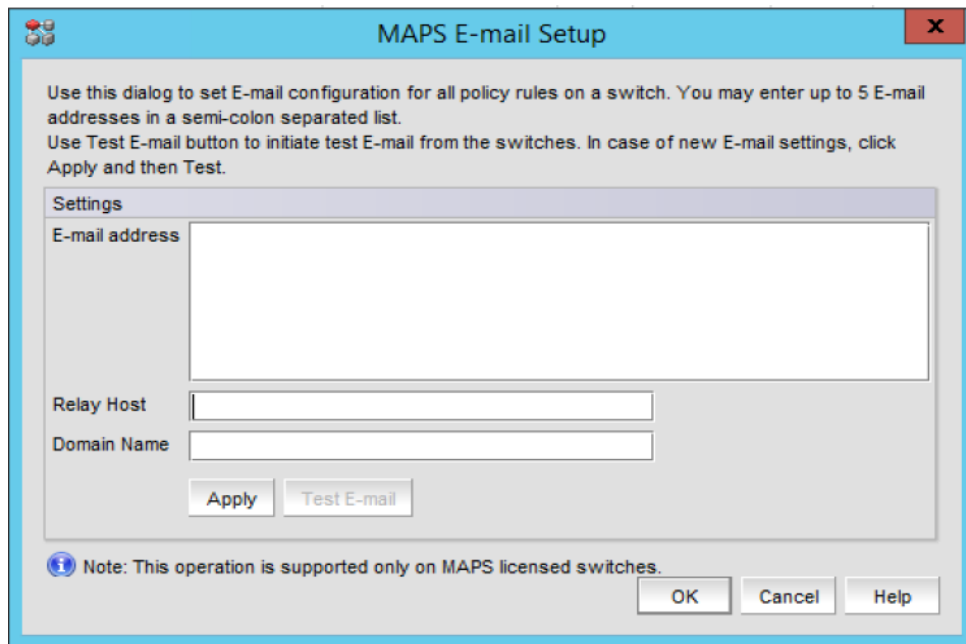


4. スイッチをハイライト表示し、[Actions] をクリックします。ここから、輻輳分散の場合に取ることができるアクションを決定できます。オーバーサブスクリプションによる輻輳分散の具体的なケーススタディについては、メールと RAS ログ イベントのチェックがオフになっていることを確認するだけで十分です。



5. メールでアラートを受信する場合は、次のようにします。[E-mail Setup] をクリックして、該当するフィールドに入力します。

Error! No text of specified style in document.

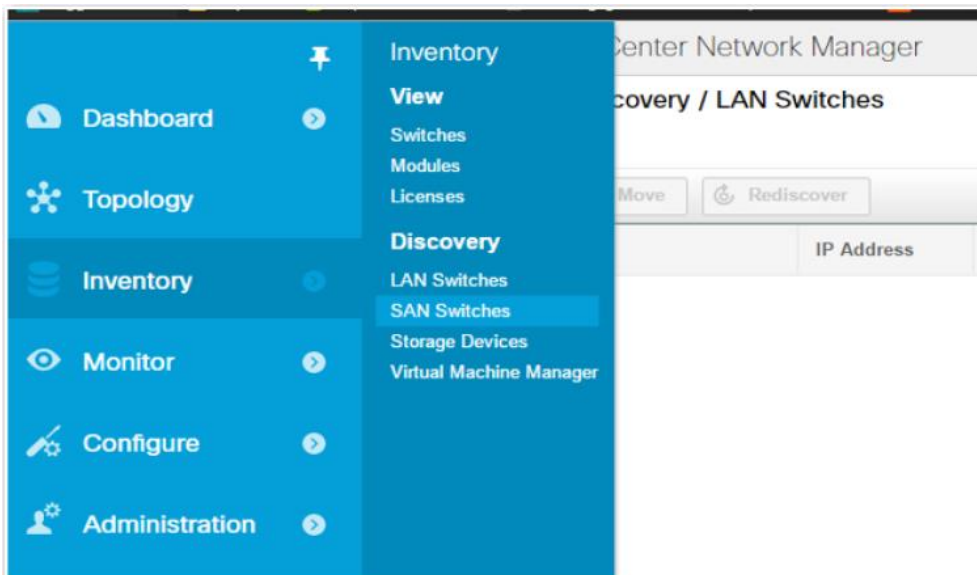


6. ファブリック内の**すべてのスイッチ**で、これらのステップを繰り返します。

6.1.2 Cisco

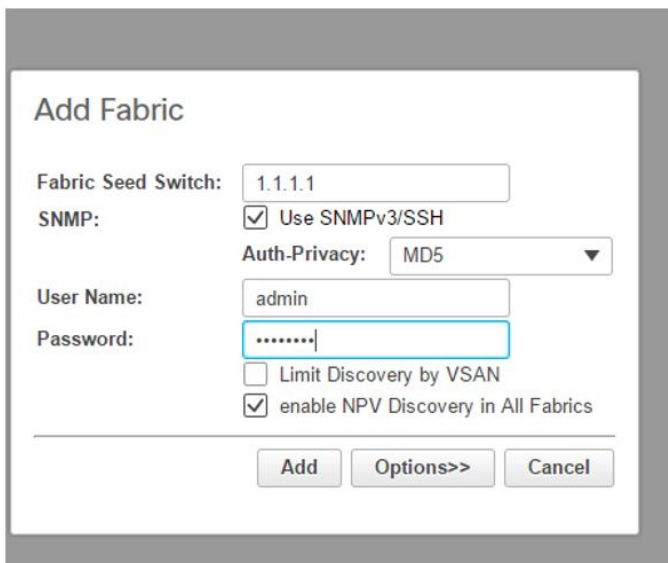
- ファブリックの検出

1. DCNM にログインし、[Inventory] > [Discovery] > [SAN] > [Switches] をクリックします。



2. 新しいウィンドウで、**プラス記号 (+)** をクリックし、ファブリック内のいずれかのスイッチに必要な情報を入力します。

Error! No text of specified style in document.



The image shows a configuration window titled "Add Fabric". It contains the following fields and options:

- Fabric Seed Switch:** A text input field containing "1.1.1.1".
- SNMP:** A checked checkbox labeled "Use SNMPv3/SSH".
- Auth-Privacy:** A dropdown menu currently set to "MD5".
- User Name:** A text input field containing "admin".
- Password:** A text input field containing ".....".
- Limit Discovery by VSAN:** An unchecked checkbox.
- enable NPV Discovery in All Fabrics:** A checked checkbox.

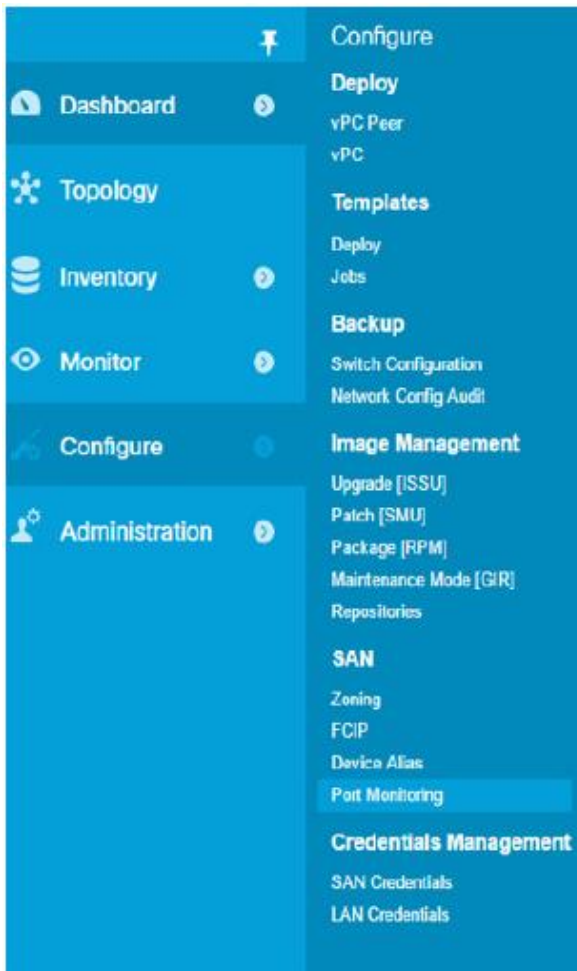
At the bottom of the dialog are three buttons: "Add", "Options>>", and "Cancel".

3. その他のすべてのファブリックでこのセクションを繰り返します。

Error! No text of specified style in document.

Cisco ポートモニタリング (PMON) の有効化

1. **[Configure]** > **[SAN]** > **[Port Monitoring]** をクリックします。



2. デフォルト プロファイルを選択し、スイッチに対して [Push] をクリックします。

The screenshot shows the Cisco Data Center Network Manager interface. The left sidebar contains navigation options: Dashboard, Topology, Inventory, Monitor, Configure, and Administration. The main content area is titled 'Configure / SAN / Port Monitoring'. At the top, there are buttons for 'default', 'Save', 'Save As', 'Delete', and 'Push to switches'. The 'Push to switches' button is highlighted with a red box. Below the buttons is a table with 12 columns: SI No, Counter Description, Rising Thresh..., RisingEvent, Falling Thres..., FallingEvent, Poll Interval, Warning Thre..., Port Guard, and Monitor ?. The table lists 16 different error counters and their configurations.

SI No	Counter Description	Rising Thresh...	RisingEvent	Falling Thres...	FallingEvent	Poll Interval	Warning Thre...	Port Guard	Monitor ?
1	Link Loss	5	Warning	1	Warning	60	0	false	true
2	Sync Loss	5	Warning	1	Warning	60	0	false	true
3	Signal Loss	5	Warning	1	Warning	60	0	false	true
4	Invalid Words	1	Warning	0	Warning	60	0	false	true
5	Invalid CRC	5	Warning	1	Warning	60	0	false	true
6	Tx Discards	200	Warning	10	Warning	60	0	false	true
7	LR Rx	5	Warning	1	Warning	60	0	false	true
8	LR Tx	5	Warning	1	Warning	60	0	false	true
9	Timeout Discard	200	Warning	10	Warning	60	0	false	true
10	Credit Loss Reco	1	Warning	0	Warning	1	0	false	true
11	Tx Credit Not Available (%)	10	Warning	0	Warning	1	0	false	true
12	Rx Datarate (%)	80	Warning	20	Warning	60	0	false	true
13	Tx Datarate (%)	80	Warning	20	Warning	60	0	false	true
14	Tx Slowport Oper Delay (msec)	50	Warning	0	Warning	1	0	false	true
15	TxWait (%)	40	Warning	0	Warning	1	0	false	true
16	State Change	5	Warning	0	Warning	60	0	false	true

3. すべてのファブリックを選択し、[Push] をクリックします。





The screenshot shows the 'Push to switches' dialog box. It has a 'Policy:' dropdown set to 'default'. Below it, 'Port Type:' has three radio buttons: 'trunks/Core', 'access-port/Edge', and 'all', with 'all' selected. The 'Scope:' section shows a tree view of the network hierarchy. Under 'Data Center', 'SAN' is expanded, and three fabric entries are listed: 'Fabric_AMER-MDS9513-1', 'Fabric_AMERGen2MDS9509', and 'Fabric_I0P054150'. The 'Fabric_I0P054150' entry is highlighted with a blue selection bar. At the bottom, there is a note: 'Note: Features not supported by the NX-OS version of the switch will be ignored.' and two buttons: 'Push' and 'Cancel'.

注 : この IP アドレスは意図的に削除されました。

Error! No text of specified style in document.

Push to switches Result

Policy: default
Port Type: All

Total 2     ▼

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

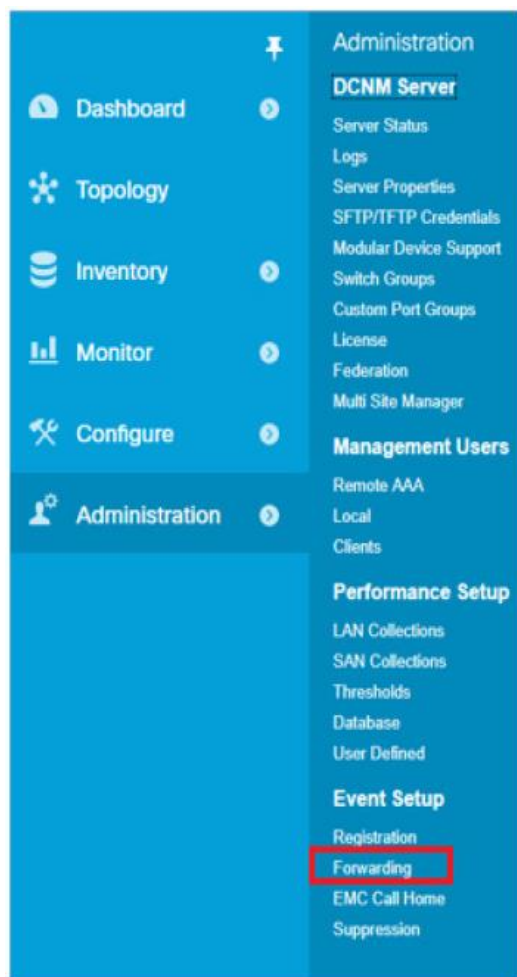
Done

4. Cisco MDS では、SNMP または Syslog 経由でアラートを受信できます。両方の構成については、次の構成ガイドを参照してください。

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

5. メールのホームを構成するには（オプション） **[Administration]** > **[Event Setup]** をクリックします。

Error! No text of specified style in document.



6. プラス記号 (+) をクリックし、受信者のメール アドレスを入力し、[Add] をクリックします。

A screenshot of the 'Add Event Forwarder Rule' dialog box. The dialog contains the following fields and options:

- Forwarding Method: E-Mail Trap
- Email Address:
- Forwarding Scope: Fabric/LAN Port Groups
- Scope:
- VSAN Scope: All List
- Source: DCNM Syslog
- Type:
- Storage Port Only
- Minimum Severity:

At the bottom right, there are two buttons: 'Add' and 'Cancel'. The 'Add' button is highlighted with a light blue background.

Error! No text of specified style in document.

- SMTP Server 情報と送信者のメール アドレスを入力し、**[Apply and Test]** をクリックして、メールを受信したことを確認します。

The screenshot shows the 'Administration / Event Setup / Forwarding' configuration page. It includes fields for 'SMTP Server' (1.1.1.1), 'From' (dcm@dell.com), 'Start' (07/10/2017 09:59 AM), and 'End' (07/10/2017 09:59 AM). There are 'Apply' and 'Apply and Test' buttons. Below the form is a table for 'Event Forwarder Rules' with a 'Selected 0 / Total 3' indicator.

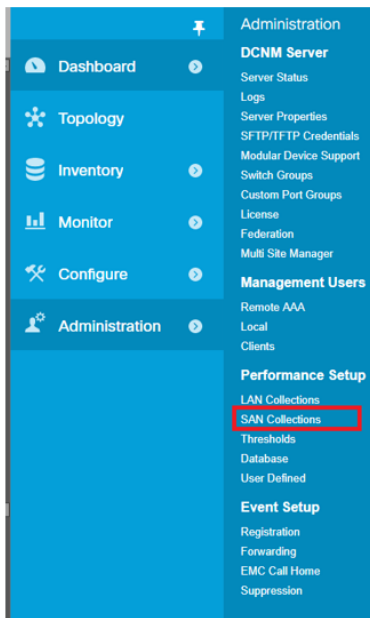
- パフォーマンス監視が実行されていることを確認します。**[Administration] > [Server Status]** をクリックします。**Performance Collector**が実行していることを確認します。実行していない場合は、**[Play]** ボタンをクリックして開始します。

The screenshot shows the 'Administration / DCNM Server / Server Status' page. The left sidebar contains navigation options: Dashboard, Topology, Inventory, Monitor, Configure, and Administration. The main content area shows a table with the following data:

DCNM Server	Actions	Service Name	Status
localhost		Database Server	Running
localhost		Search Indexer	Last updated: 2017-06-30 14:30:01
localhost		Performance Collector	Running. Collecting 118 entities. 100% response in last hour. last DB update: 2017/06/30 14:55
10.241.213.177		SMI-S Agent	Running

- [Administration] > [Performance Setup] > [SAN] > [Collections]** をクリックします。

Error! No text of specified style in document.



10. パフォーマンス統計情報を収集するファブリックがチェックされていることを確認します。Performance Collectorサービスが再起動します。

The screenshot shows the Cisco Data Center Network Manager interface. The page title is 'Administration / Performance Setup / SAN Collections'. Below the title is a table with the following data:

	Name	ISL/NPV Links	Hosts	Storage	FC Flows	FC Ethernet
1	<input checked="" type="checkbox"/> Fabric_0P054151	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
2	<input checked="" type="checkbox"/> Fabric_MDS9706	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

参考資料

Brocade MAPS 構成ガイド :

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

Cisco 低速ドレイン デバイスのホワイトペーパー :

<http://www.cisco.com/c/dam/en/us/products/collateral/storage-networking/mds-9700-series-multilayer-directors/whitepaper-c11-737315.pdf>

VMAX ホスト I/O 制限機能に関する一般的なリファレンス :

<https://community.emc.com/thread/188068?start=0&tstart=0>

Ezfiio I/O ツール

<https://github.com/earlephilhower/ezfiio>

輻輳分散の重大度

輻輳分散メトリックはすべて重要ですが、次のセクションでは、イベントが発生する率によって、各イベントが環境に及ぼし得る影響が劇的に変わる可能性があることを示します。さらに問題を複雑にしているのは、BrocadeとCiscoの両方が異なる重大度分類法を使用しているという事実です。そのため、次のDell EMC固有の分類法を使用して、それらを各スイッチ タイプにマッピングすることになります。

6.1.3 Dell EMC

- **タイプ 1 :**
 - 輻輳率が 0.2 以上になっている。
 - フレームロス（破棄）またはリンクリセット
- **タイプ 2 :**
 - 輻輳率が 0.2 以上になっている。
 - フレームロス（破棄）またはリンクリセットなし
- **タイプ 3 :**
 - 輻輳率が 0.2 以上になっている。
 - フレームロス（破棄）およびリンクリセット

6.1.4 Brocade

- **軽度**
 - わずかなクレジット遅延
 - わずかなキューレイテンシー（10 ミリ秒未満）
 - フレームロス（破棄）またはリンクリセット

- 中程度
 - 中程度のクレジット遅延
 - 中程度のキュー レイテンシー（10～80 ミリ秒）
 - フレーム ロス（破棄）またはリンクリセットなし
- 重大
 - 大きいクレジット遅延
 - 大きいキュー レイテンシー（80 ミリ秒を超える）
 - フレーム ロス（破棄）および一部のリンクリセット

6.1.5 Cisco

- レベル 1 : レイテンシー
 - 残りのクレジット数の減少または短期間のクレジット使用不可期間
 - 破棄、再送信、リンクリセットはなし
- レベル 2 : 再転送
 - クレジット使用不可期間の長期化
 - 再送信に先行する、輻輳ドロップ タイムアウトまたはクレジットなしドロップ タイムアウト*により、フレームが破棄される（リンクリセットはなし）。
- レベル 3 : 極端な遅延
 - クレジット使用不可期間の延長（F ポートの場合は 1 秒、E ポートの場合は 1.5 秒）
 - リンクのリセットまたはポート フラップ

輻輳分散用語のクロス リファレンス

メトリックと重大度を組み合わせて使用することで、さまざまなタイプの輻輳分散イベントを特定することができます。前述のセクションと同様に、BrocadeとCiscoの両方についての別個のセクションがありますが、BrocadeとCiscoの両方がオーバーサブスクリプションという用語を使用しているため、このセクションではまずその概要を説明します。

6.1.6 オーバーサブスクリプション

オーバーサブスクリプションとは、簡単に言えば、「システム上の潜在的な需要がその需要を満たすシステムの容量を超えている」という状態のことです。この例としては、幹線道路システムがよく知られています。すべての住民が（ハリケーンからの避難などで）車を同時に運転することを突然決定した場合、交通が停止することがあります。

FC SANの場合は、帯域幅（BW）比率でオーバーサブスクリプションを考慮すると便利です。たとえば、図3に示すように、ホスト1（4Gbps）とストレージ1（16Gbps）の間の帯域幅（BW）比率は1:4です。したがって、ホスト1は4:1のオーバーサブスクリプション状態ということです。これを、ホスト2（16Gbps）とストレージ2（16Gbps）の間の帯域幅比率1:1と比較します。ホストと、ホストがアクセスするストレージの両方で32Gbps ISLを使用することを検討してください。ホスト2とストレージ2の間のオーバーサブスクリプションがないことを確認できます。このケースでは、ホスト2とストレージ2はオーバーサブスクリプション状態ではありません。

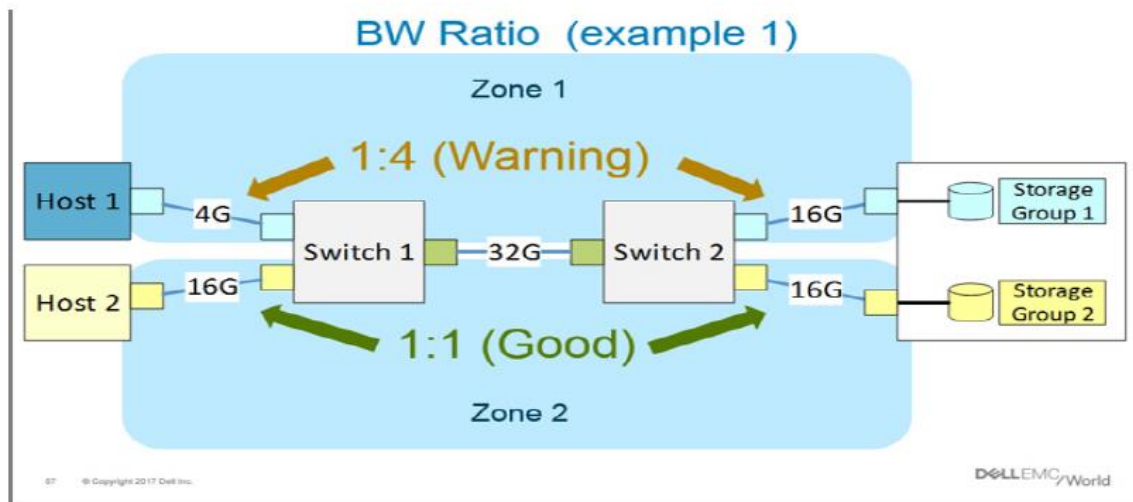


図 19 帯域幅比率 - 例 1

オーバーサブスクリプションを計算するときは、図19で示すとおり、考慮されているインターフェイスの帯域幅（BW）を追加することで帯域幅比率を計算することに注意してください。一見すると、16Gbps HBAで8Gbpsのストレージにアクセスしているように見えますが、実際には3つのストレージ インターフェイスがあるため、16Gbps HBAは24 Gbpsのストレージにアクセスしています。このため、ホストは3:2のオーバーサブスクリプション状態です。

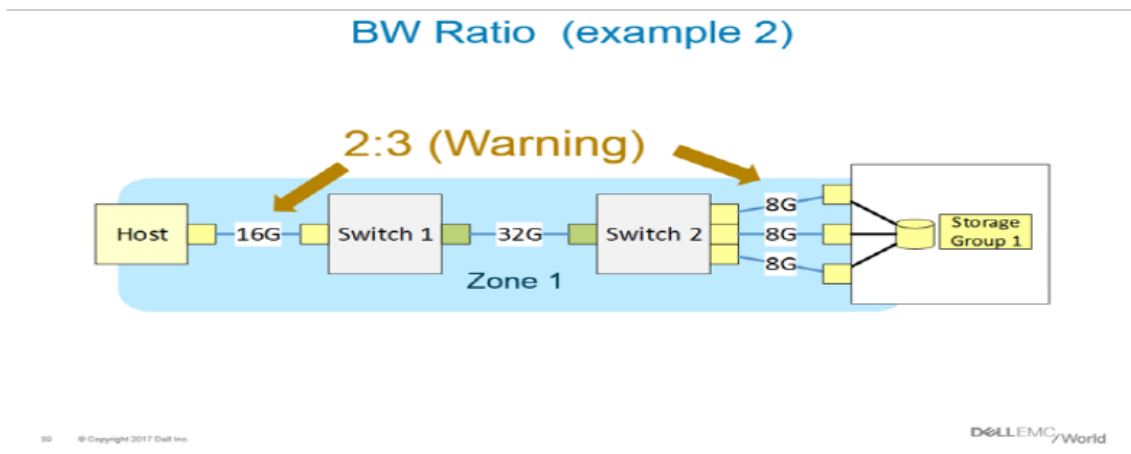


図 20 帯域幅 (BW) 比率 - 例 2

前述の2つの例では、ISLの帯域幅 (BW) は、エンド デバイスでサポートできる帯域幅 (BW) の量以上になります。このようなケースは通常はありません。図20に示すように、ホストは実際には3:4のアンダーサブスクリプション状態ですが、ISLは16Gbpsしかないため、エンド デバイスと使用されるISLの間にオーバーサブスクリプションがあり、ISLは3:2のオーバーサブスクリプション状態であると言えます。

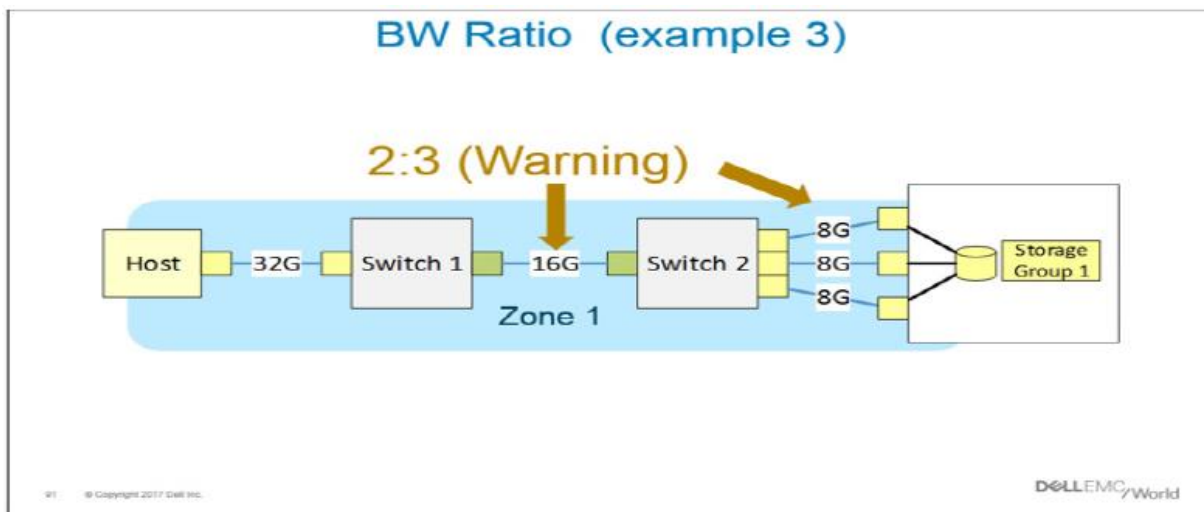


図 21 帯域幅 (BW) 比率 - 例 3

6.1.7 Brocade

Brocadeは、次の3つの異なるクラスの輻輳分散イベントを定義しています。

- オーバーサブスクリプション

前のセクション (上記) で定義されているとおりです。

- 誤動作するデバイス

ライン レートを維持するのに十分な速度でクレジットを解放しないエンド デバイスまたは ISL。たとえば、エンド デバイスが 16Gbps のリンク速度をネゴシエートし、16Gbps でデータを受け取れる速度でクレジットを返すことができない場合、誤動作するデバイスが生じる可能性があります。これらのタイプのデバイスは、「低速ドレイン」とも呼ばれます。ポートは輻輳分散の影響を受けているため、ドライバーの問題や ISL の場合などのさまざまな理由により、デバイスの誤動作が発生する可能性があることを指摘することは重要です。

- クレジット損失

クレジット損失シナリオとは、何らかの理由（通常は不定期なビット エラーなど）で、特定のリンク上の1つまたは両方のデバイスが、実際に実行されたよりも少ない転送クレジットしかないと判断していることを意味します。このような状況の1つの原因として、R_RDYを破壊するビットエラーがあります。これが頻繁に発生すると、時間の経過とともにパフォーマンスが低下し始めデータを転送するSANの能力がゆっくりと低下します。この問題の詳細については、KB 464245（「Bit Errors and their impact」）を参照してください。

Brocade 輻輳分散用語のクロス リファレンス

Brocadeの場合、すべての要素をまとめると、次のBrocade固有の輻輳分散用語のクロス リファレンスになります。

原因	軽度	中程度	重大
オーバーサブスクリプション¹	<ol style="list-style-type: none"> 1. デバイス ポートの帯域幅は広い 2. ISL ポートで短いクレジット レイテンシー 3. フレーム ロスまたはリンクリセットはなし 	<ol style="list-style-type: none"> 1. デバイス ポートの帯域幅は広い 2. ISL ポートで中程度のクレジット レイテンシー 3. ISL ポートで 10～80 ミリ秒のキュー レイテンシー 4. フレーム ロスまたはリンクリセットはなし 	<ol style="list-style-type: none"> 1. デバイス ポートの帯域幅は広い 2. ISL ポートでの長いクレジット レイテンシー 3. ISL ポートで 80 ミリ秒を超えるキュー レイテンシー 4. アップストリーム (ISL) ポートでのフレーム ロス (キュー レイテンシーは 220～500 ミリ秒) 5. リンクリセットなし
デバイスの誤動作	<ol style="list-style-type: none"> 1. デバイス ポートとアップストリーム ISL ポートでの短いクレジット レイテンシー 2. デバイス ポートとアップストリーム ISL ポートでの 10 ミリ秒未満のキュー レイテンシー 3. フレーム ロスまたはリンクリセットはなし 	<ol style="list-style-type: none"> 1. ISL ポートとアップストリーム ISL ポートでの中程度のクレジット レイテンシー 2. デバイス ポートとアップストリーム ISL ポートでの 10～80 ミリ秒のキュー レイテンシー 3. フレーム ロスまたはリンクリセットはなし 	<ol style="list-style-type: none"> 1. デバイス ポートとアップストリーム ISL ポートでの長いクレジット レイテンシー 2. デバイス ポートとアップストリーム ISL ポートでの 80 ミリ秒を超えるキュー レイテンシー 3. デバイスまたはアップストリーム (ISL) ポートでのフレーム ロス (キュー

			レイテンシーは 220～500 ミリ秒) 4. ISL ポートでのリンクリセット (クレジットストールは 2 秒超)
クレジット損失 ²	1. ポートで短いクレジットレイテンシー 2. ポート、またはポートのアップストリームでの 10 ミリ秒未満のキューレイテンシー 3. フレームロスまたはリンクリセットはなし	1. ポートで中程度のクレジットレイテンシー 2. ポート、またはポートのアップストリームリンクでの 10～80 ミリ秒のキューレイテンシー 3. フレームロスまたはリンクリセットはなし	1. ポートでの長いクレジットレイテンシー 2. ポート、またはポートのアップストリームでの 80 ミリ秒を超えるキューレイテンシー 3. ポート、またはポートのアップストリームでのフレームロス (クレジットストールは 220～500 ミリ秒) 4. ポート、またはポートのアップストリームでのリンクリセット (クレジットストールは 2 秒超)
¹ オーバーサブスクリプションにより重大な輻輳が発生するのは非常に稀です。 ² クレジット損失の原因は主に、ITW、CRC、または他のシグナル関連の問題による送信エラーです。			

6.1.8 Cisco

Ciscoは、次の2つの異なるクラスの輻輳分散イベントを定義しています。

- **オーバーサブスクリプション**

前述のとおりです。

- **クレジットの枯渇**

ラインレートを維持するのに十分な速度でクレジットを解放しないエンドデバイスまたは ISL。たとえば、エンドデバイスが 16Gbps のリンク速度をネゴシエートし、16Gbps でデータを受け取れる速度でクレジットを返すことができない場合、誤動作するデバイスが生じる可能性があります。これらのタイプのデバイスは、「低速ドレイン」とも呼ばれます。ポートは輻輳分散の影響を受けているため、ドライバーの問題や ISL の場合などのさまざまな理由により、デバイスの誤動作が発生する可能性があることを指摘することは重要です。

- **Cisco 輻輳分散用語のクロスリファレンス**

Ciscoのすべての要素をまとめると、次のCisco固有の輻輳分散用語のクロスリファレンスになります。

輻輳タイプ	レベル - 1 : レイテンシー	レベル - 2 : 再送信	レベル - 3 : 極端な遅延
オーバーサブスクリプション	<ol style="list-style-type: none"> 1. エンド デバイス ポートでのリンク使用率は高い 2. エンド デバイス ポートで B2B クレジットの枯渇はなし 3. ISL への輻輳分散 4. フレーム ロスまたはリンクリセットはなし 	<p>オーバーサブスクリプションにより再送信または極端な遅延が発生するは非常に稀</p>	
クレジットの枯渇	<ol style="list-style-type: none"> 1. エンド デバイス ポートでのリンク使用率は低い 2. 残りのクレジット数の減少または短期間のクレジット使用不可期間 3. ISL への輻輳分散 4. 破棄、再送信、リンクリセットはなし 	<ol style="list-style-type: none"> 1. エンド デバイス ポートでのリンク使用率は低い 2. クレジット使用不可期間の長期化 3. ISL への輻輳分散 4. 再送信に先行する、輻輳ドロップ タイムアウトまたはクレジットなしドロップ タイムアウト*により、フレームが破棄される (リンクリセットはなし) 	<ol style="list-style-type: none"> 1. フレームがエンド デバイスに送信されない 2. クレジット使用不可期間の延長 (F ポートの場合は 1 秒、E ポートの場合は 1.5 秒) 3. ISL への重大な輻輳 4. リンクのリセットまたはポートフラップ
<p>*デフォルトの構成 : 輻輳ドロップ タイムアウト - 500 ミリ秒、クレジットなしドロップ タイムアウト - なし 構成可能なオプション : 輻輳ドロップ タイムアウト - 100~500 ミリ秒、クレジットなしドロップ タイムアウト - 1~500 ミリ秒 推奨構成 : 輻輳ドロップ タイムアウト - 200 ミリ秒、クレジットなしドロップ タイムアウト - 50 ミリ秒</p>			