

**VIOPSO6 WORKSHOP** 

ここまで身近に!10Gbpsを越えるネットワークの世界

#### **Everything over Ethernet**

~その要件と技術~



ブロケード コミュニケーションズ システムズ

### 本日の内容

- 既存のデータセンタ内プロトコルとその特徴
- Everything over Ethernet:その前提条件
  - Datacenter Bridge, TRILL
- FCoE, RoCE
- 課題、注意点、その他
- ・まとめ



### 既存の Datacenter 内プロトコルとその特徴

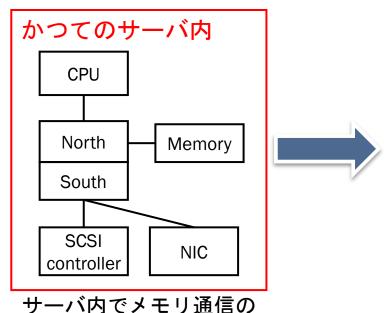


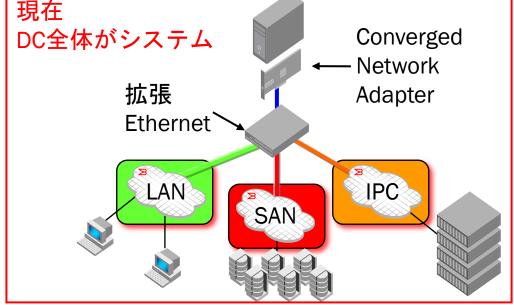


### System Area Network の特徴

ロスレス性はなぜ必要か?

- RDMA, I/O の特徴
  - メモリ間データ転送をベースに、バス(SCSIやメモリチャネル)を経由して DMAでデータをブロックで読み書きする
    - メモリ、ディスクの Read/Write ではロスレスは暗黙の仮定





SAN/IPC のロスレスを実現

損失は無いと仮定

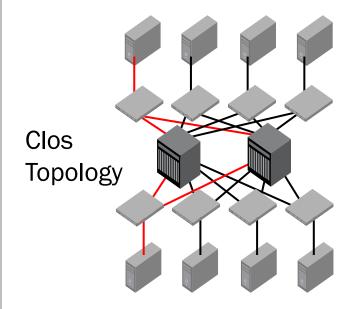
#### Datacenter 内のEthernet以外のネットワーク

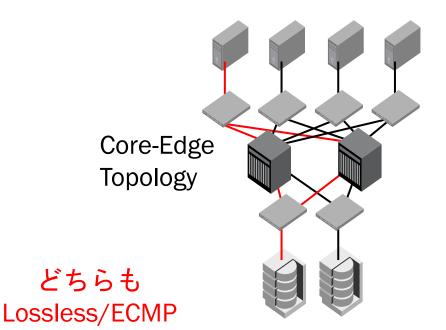
が備わったネットワーク

次世代 Ethernet が実装しなくてはならない技術

- InfiniBand
  - 低遅延(300ns)、広帯域 (10Gbps/Lane)

- Fibre Channel
  - 低遅延(450ns)、広帯域 (16Gbps/Lane)



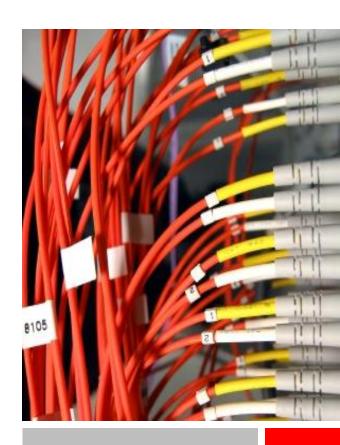




### 今日の DC ネットワークでの挑戦

#### 通常の Ethernet の限界

- Layer 2 性能と信頼性の向上
  - 等価コストマルチパス技術の実現
  - ロスレス
  - ブロードキャスト トラフィック の最適化
  - ・ MAC 学習の最適化
  - 高速収束
- 自動化による管理オーバヘッドの削減
  - 自律的ファブリック形成
  - 自動的発見、構成
  - VM 移動性への対応
- シンプル化
  - 次ページ参照



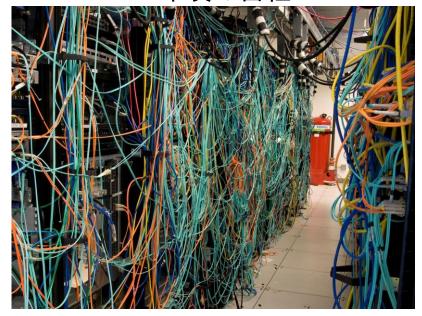
**Data Center** 



#### Ethernet にまとめることの利点

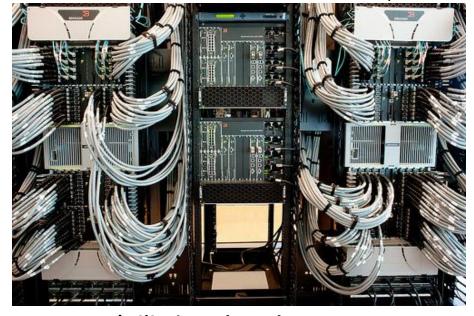
Brocade の自社データセンタ

2008年頃の自社 DC



仮想化により、複雑性が増大

2010年8月オープンの新社屋/DC



広帯域+すっきり+ 今までと変わらない運用



### **Everything over Ethernet**

その前提条件



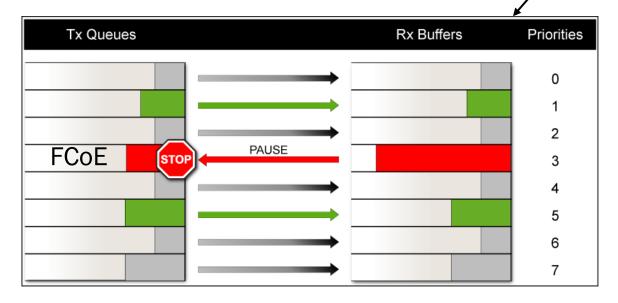


#### ロスレス性の実現

#### Data Center Brdinging 技術

- "Data Center Bridging"とはデータセンタ向けに拡張された Ethernet
- ・ 下記の複数の標準を含む
  - Priority-based Flow Control (802.1Qbb)
  - Enhanced Transmission Selection (802.1Qaz)
  - Data Center Bridging eXchange (802.1Qaz)

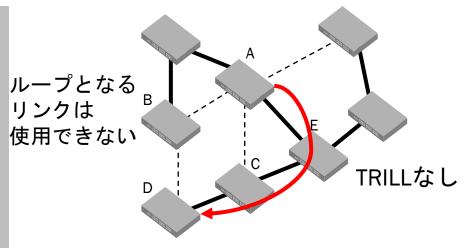
10G リンクを仮想的に 8本のチャネル(Priority) に分割し、フロー制御



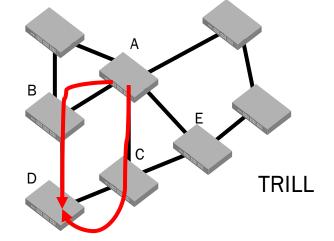


#### Equal Cost Multi-Path の実現

Transparent Interconnection of Lots of Links



等価コスト パスをバラ ンスさせて 使用可能



- 何故 TRILL が必要なのか?
  - 従来のEthernetでは実現できないマルチパスの実現 (redundant path failover)
  - 最短経路へのルーティング
  - •<u>ルーティングなどのゼロ設定</u>
  - FCoE などのレジリエント構成に必 要

・ 注意事項で説明



# Fibre Channel over Ethernet (RDMA over Converged Ethernet)

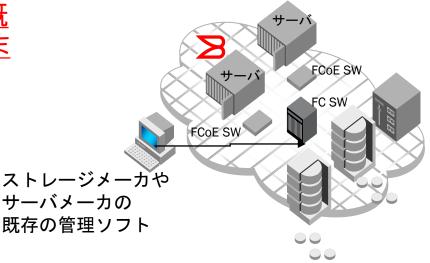


#### FCoE の利点

- 従来の FC のインフラがそのまま使用できる
  - 既存の設備投資がそのまま生 きる
- 従来の管理スキームがそのまま使用できる
  - ストレージメーカの提供する既存の管理ソフトウェアがそのまま使用できる

#### ・ 新技術ではない

- iSCSI のような新しい技術と違い iqn や SLP/iSNS などを知らなくても構成できる
- そもそも IP は zero 設定ではない





#### **Fibre Channel over Ethernet**

- DCB 上にマッピングした上位 プロトコルのひとつ
  - ANSI T11 FC-BB-5 v2.0 で標準化
- FC フレーム(2112 bytes)を Ethernet フレームでカプセル 化
  - ジャンボフレーム(2500bytes)
  - FC フレームと Ethernet フレー ムは1:1でマッピング
  - RFC3643 Fibre Channel Frame Encapsulation が標準として使用されている

FC-4

Upper layer (SCSI-3)

FC-3

FC Generic Service

Link Service
BLS/ELS

DCB Layer-2

DCB Layer-2

64b/66b encoding, others

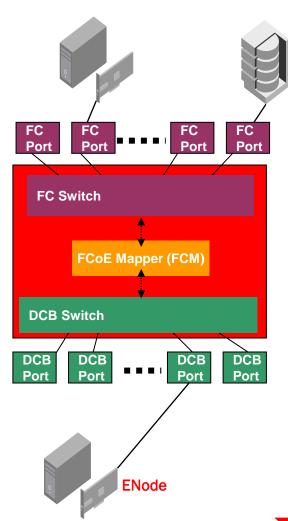
10Gbps Ethernet phy



#### FCoE の構成と言葉

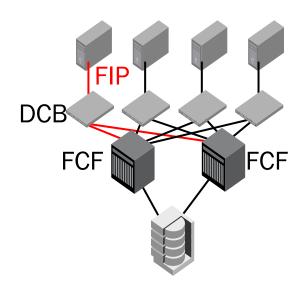
- FCoE Node (ENode)
  - DCBのインターフェースを持ちFCノードとして振舞うことが可能な統合ネットワークアダプタ (CNA)がこれに当たる
- FCoE Forwarder
  - FCスイッチとして動作することができるDCBスイッチを表す
- FCoE Mapper (FCM)
  - ファイバチャネルフレームを イーサネットフレームにカプ セル化/非カプセル化を行う 機能

FCoE Forwarder (FCF)



#### **FCoE Initialization Protocol**

- ・ FIP の役割
  - DCB ファブリックにおけるFCFの発見
  - DCB ファブリックに対するFCファブリックサービスの提供
  - FCoE ファブリックパラメータの交換 ⊾

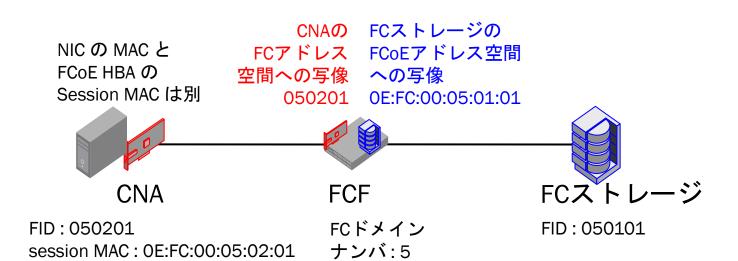


- Discovery Phase
  - ENode による FCF の発見
  - FCoE VLAN の発見
  - アドレッシング方式の提供
- Login Phase
  - ENode と FCF の関連付け
  - 既存の FLOGI, FLOGI\_ACC, LOGO をそのまま使用
- Data Transfer Phase
  - PLOGI/PRLI
  - その他のすべての FC プロトコルフレーム
    - ELS, FC4 ULP



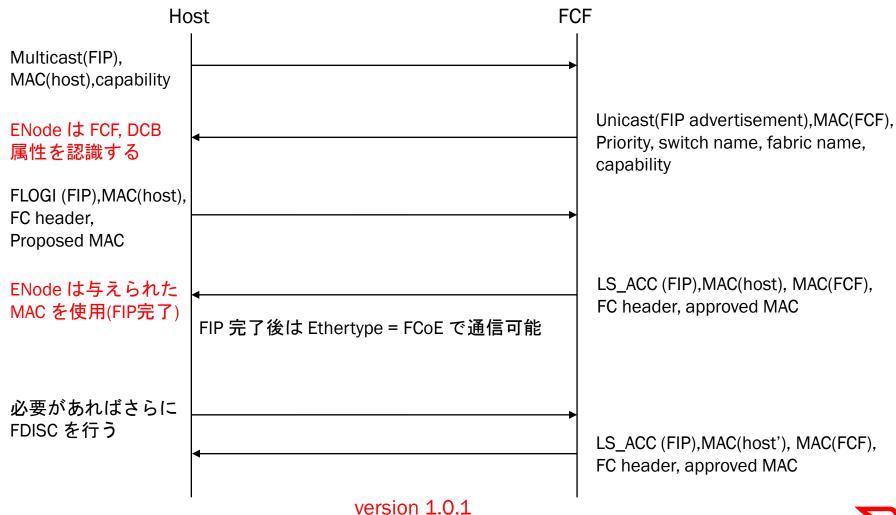
#### FCoE アドレッシング

- DCB ファブリック内では MAC アドレスベースで通信するため、旧来の FC のアドレスだけでは通信できない
  - そのため FCoE では MAC アドレスを動的にアサインする仕組みが必要
    - Fabric Provided MAC Address (FPMA) によりアドレスを確保
      - MAC アドレスの上位3バイトは "OxOE, OxFC, OxOO"
      - MAC アドレスの下位3バイトは FID (FC fabric 内のアドレス)

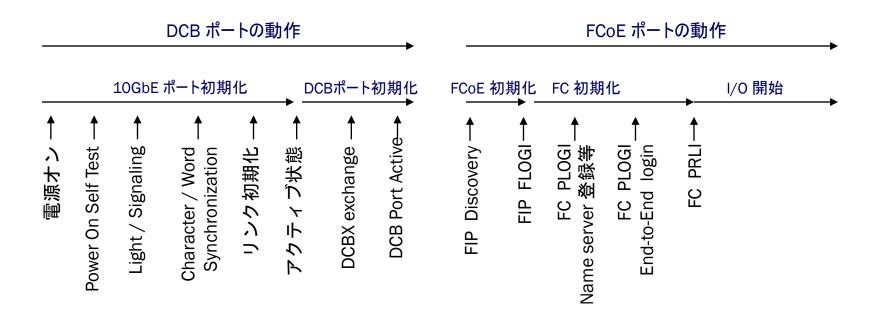




#### FIP の手順



## DCB/FCoE のリンク初期化の流れ





#### FCoE の事例

#### 2010年11月09日

#### プレスリリース

#### ビットアイル、ブロケードのコンバージド・ネットワークソリューションを採用し、 クラウドサービス向けネットワーク共通基盤を模築

DCB/FCoE対応スイッチと統合ネットワーク・アダプタにより、LANとSANを統合。 既存の環境を生かしながら、将来の拡張への柔軟な対応を支援するネットワークを実現

ブロケードコミュニケーションズ システムズ株式会社(本社:東京都千代田区、代表取締役社長:青葉 雅和)は本日、大手インターネット・データセンター(IDC)事業者の株式会社ビットアイル(本社:東京都港区、代表取締役社長兼CEO:寺田 航平)が、同社クラウドサービスのより効率的かつ安定的な運用を可能にする統合型共通ネットワーク基盤の構築に、ブロケードのDCB(データセンター・ブリッジング)/FCoE(ファイバーチャネル・オーバー・イーサネット)対応スイッチであるBrocade 8000と統合型ネットワーク・アダプタBrocade 1000 CNAを採用し、先進的な次世代コンバージェンス技術を業界に先駆けて本格導入したことを発表しました。

ビットアイル社では、2009年9月からクラウドサービス「Cloud ISLE」を開始。以来、インターネット・サービス事業者をはじめとする多くの企業ユーザに対して、豊富なオプションとともに柔軟なサービスを提供してきました。サーバ仮想化のさらなる導入・拡大、およびユーザトラフィックの増加といった状況を受け、より安定性に優れ、高信頼なサービスを効率的に提供できる環境を整えるべく、2010年春、同社では「Cloud ISLE」サービスを支えるネットワーク基盤の刷新について検討を開始しました。

新たなネットワーク基盤を構築するにあたってビットアイル社が要件として掲げていたのは、既存のファイバーチャネル資産を生かしながら、一方でシステム全体の効率化、複雑性の排除、管理の簡素化を実現できることでした。そして、ファイバーチャネルの信頼性と性能を維持しながら、LANとSANを統合してネットワークをシンプル化できる新しい技術「DCB/FCoE」の採用を決定。具体的なDCB/FCoE対応スイッチの選定過程では、仮想サーバ環境でのベンチマークおよび運用性テストの結果や、ファイバーチャネルでの圧倒的なリーダーシップと実績が考慮され、2010年10月、プロケードのDCB/FCoE対応スイッチBrocade 8000の導入が決定しました。これは、プロケードが提供する統合型ネットワーク・アダプタとの連携による、統合DCB/FCoEソリューションだからこそ実現可能な総合的な高パフォーマンス、安定性、および信頼性も高く評価された結果です。

ビットアイル社では、ブロケードの統合DCB/FCoEソリューションによりSANとLANのサーバ・インタフェースを統合。少ないデバイス数でシンプルなネットワークを実現しました。

#### ビットアイルにおけるブロケード統合DCB/FCoEソリューション導入のメリット:

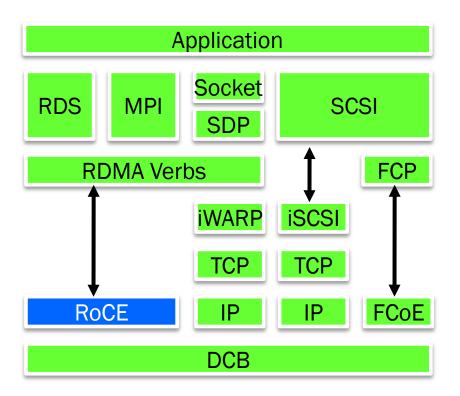
- ・LANとSANを統合することで、ケーブリング、デバイス数を減らし、コストと消費電力を削減
- ・ファイバーチャネルの信頼性・性能を損なうことなく、ネットワークをシンプル化し、かつ運用を効率化
- ・スモールスタートから、必要に応じて柔軟な水平型の展開・拡張が可能

#### 株式会社ビットアイル マーケティング本部事業推進部 部長 高倉 敏行氏談

「ブロケードが提唱するネットワーク・アーキテクチャ「Brocade One」は、当社が掲げるオープンかつマルチベンダーによるエコシステム、および投資の保護といった要件に沿っています。中長期的なブロケードの技術ロードマップでは、DCB/FCoE技術のさらなる革新が明確に示されており、ファイバーチャネル製品での安定稼働や品質といったこれまでの実績への信頼感とともに、今回のブロケードのソリューション導入を決定づけました」

ビットアイル社では、ブロケードのDCB/FCoEソリューション導入した新しい基盤で稼働する「サーバーオンデマンドNEXT」サービスを、本年12月10日に提供開始します。

### RDMA over Converged Ethernet (ロッキー)





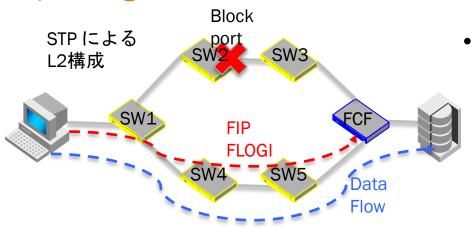
# 課題、注意点、その他



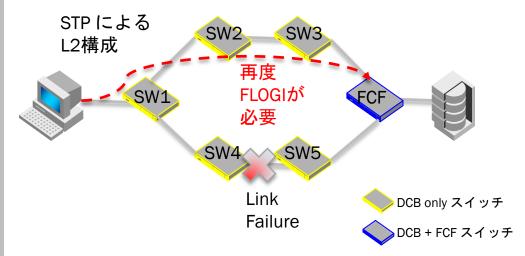


#### FCoE と ECMP の関係

#### Spanning Tree の問題点



- Link Failure が起こった場合、上の Path(SW1/SW2/SW3/FCF)に non-disruptive に切り替わるか?
  - そもそもマルチパスドライバからみてマルチパスになっていないため、 冗長パスとして設定すらできない
  - だから初めから Multi-Path が可能なトポロジが必要



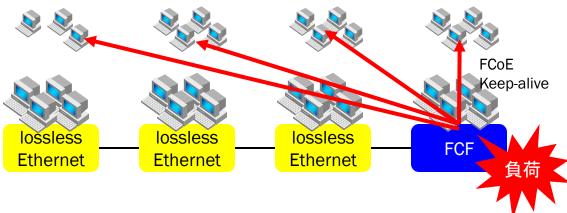


#### FCoE 環境でのカスケードの危険性

FCoE Keep-Alive のスケーラビリティとマルチホップ FCoE

DCB only SW と FCoE SW の混在 仮想サーバ

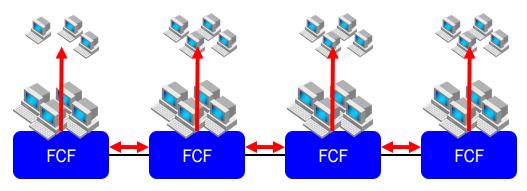
物理サーバ



FCF:FCoEに対応したスイッチ Lossless Ethernet:単にDCB対応しただけのスイッチ

FCoE SW に よる構成 仮想サーバ

物理サーバ



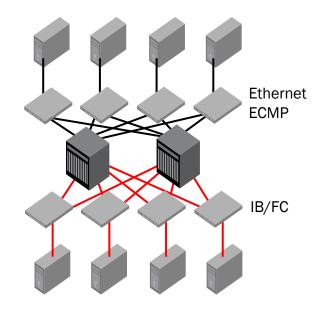
すべてのスイッチにFCoE forwarder がないとネットワークを拡張できない

#### マルチプロトコルサポートと ECMP

• Ethernet に IB/FC を統合すると すると、ルーティングプロトコ ルは何を使用する?

	ルーティングプロトコル
TRILL/SPB	IS-IS
InfiniBand	SubnetManager の実装次第 (MINHOP/UPDN等)
Fibre Channel	Fabric Shortest Path First

・ 実は標準上は未解決





# まとめ





#### **Everything over Ethernet**

- Ethernet によるデータセンタ内インタコネクトの集約の前提条件
  - DCB
  - TRILL
- XoE のメリット
  - 共通: LAN/SAN/IPC が1本のワイヤに統合できるため、TCOを削減可能
  - FCoE: 従来の FC 基盤が活かせる
  - RoCE: Ethernet でも IB と同じことができる
    - ただし同じ性能というわけではない
- 注意点
  - とりわけ FCoE には multi-path やLink Level Service にまつわる注意点が多いので、接続性のサポートは要確認







