

The background of the slide features a light blue and white hexagonal pattern. Some hexagons are filled with a light blue color, while others are white. A network of thin blue lines connects the vertices of the hexagons, with small blue dots at the connection points, resembling a circuit or data network.

ARISTA

# ウルトライーサネットの謎解き

-Demystifying Ultra Ethernet-

# UECの目的と結成メンバー

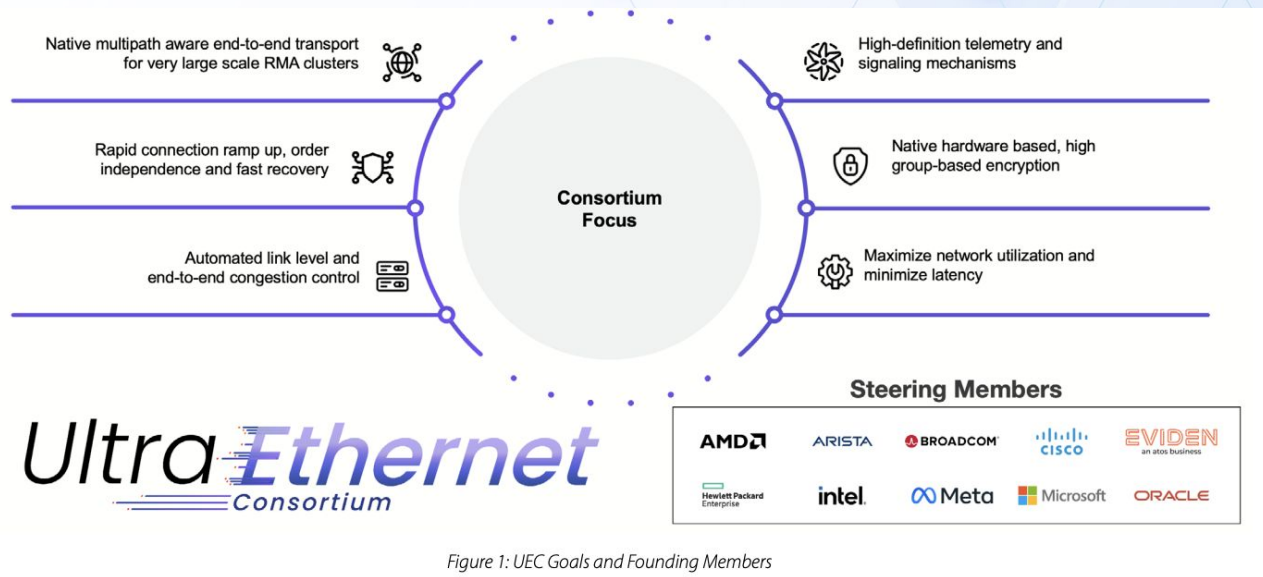
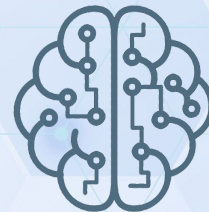












Figure 1: UEC Goals and Founding Members

- AI/MLは初期段階では従来のネットワークとは独立して開発 / 運用されていた
- 現在はあらゆるビジネスで不可欠となっていて、共通の技術パラダイムが必要
- ウルトラ・イーサネット・コンソーシアムは、AIとHPCのニーズに応えるためイーサネットを強化することを目的として設立された標準化団体である。
  - 100社以上の加盟企業と1000人以上

# 誰がAI/MLをどこで要求してるか



産業		使用用途
企業		カスタマーサポート 文章作成 自動要約
製造業		品質管理 メンテナンス予測
医療		診断サポート 文章解析
金融		予測分析 クレジットスコアリング
Webゲーミング		コンテンツ作成 カスタマイズされたエクスペアリング

産業		使用用途
小売業		商品レビュー分析 需要予測
エネルギー		エネルギー効率向上 予測メンテナンス
交通・運輸		輸送の最適化(2024年問題) 交通予測
農業		作物管理 病虫害予測
教育		個別指導 教材作成

# Ultra Ethernet Transport (UET)

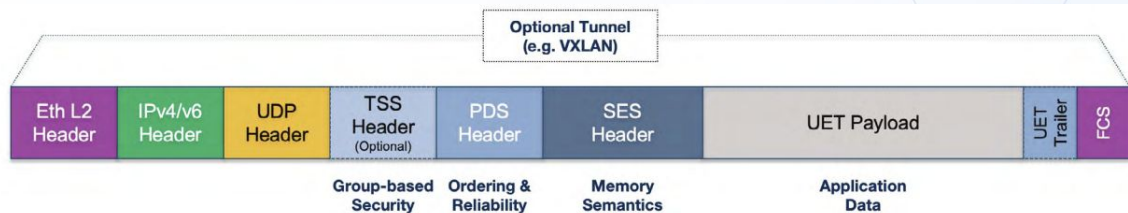
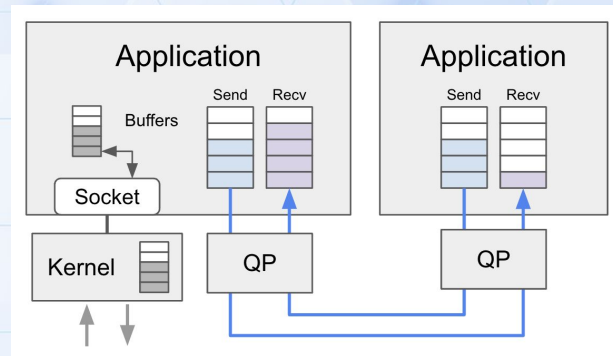
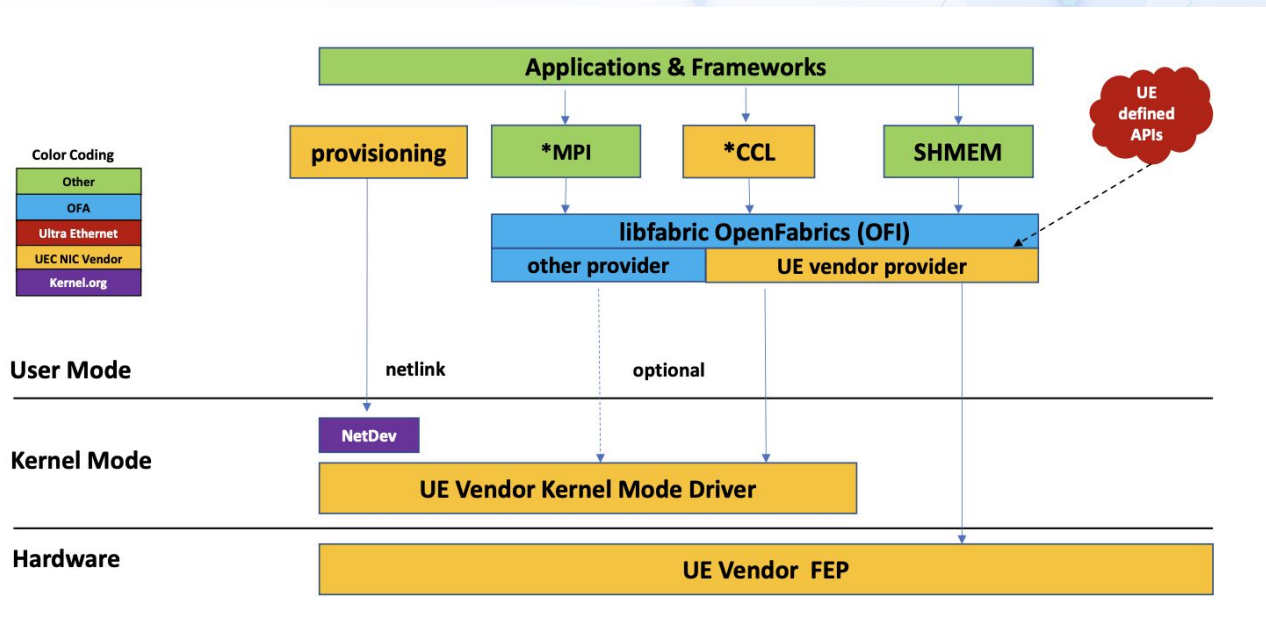


Figure 2: UET Packet Format



- UETの中核はRDMAの再構築
- 現在400Gbps/800Gbpsでネットワーク上でバースト通信する
  - ワークロードを分散出来る
  - 複数のアクセラレーターにまたがる並列計算を実現
- UET(Ultra Ethernet Transport)では標準のイーサネットおよびIPに最新のトラフィック分散技術と輻輳制御を追加

# Native Libraries



- Open Fabrics Allianceによって標準化され、HPC環境で広く実装されている成熟した汎用APIであるlibfabric 2.0を使用
- 多様なRDMAを処理する技術を単一の中央インターフェースに集約する事が出来、異なるシステムやアクセラレータアーキテクチャ間でのアプリケーション移植が容易になる



# Traffic Forwarding

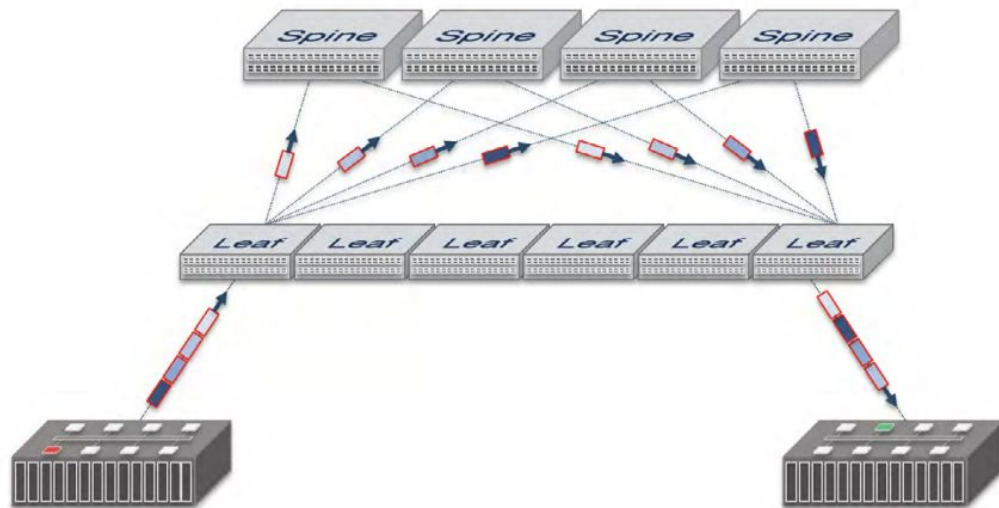


Figure 3: Packet Spraying Overview

- フローベースのトラフィック分散からNIC毎のパケットスプレイに
- NICは順不同で来たパケットを再構築してアプリケーションに渡す
- 個々の損失パケットの再送信が可能になる

# Congestion Management

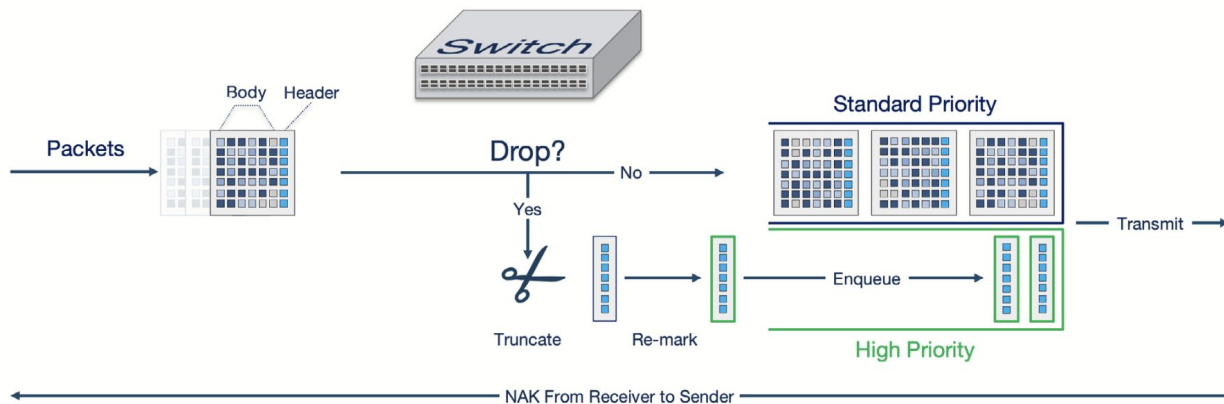
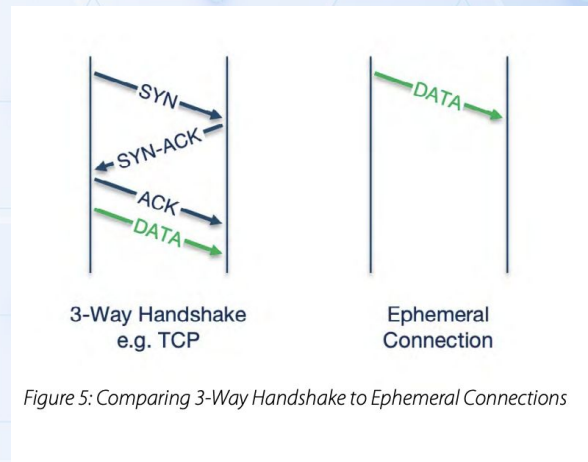


Figure 4: Packet Trimming Mechanism

- 従来はシーケンス番号でパケットの取りこぼしを検出 go-back N
  - 1つのパケット損失の為に多くのパケットが再送され非効率
- UECではパケットが輻輳状態のスイッチに到達するとパケットを廃棄する代わりにパケットを最小に切り込む。
- 切り込まれたパケットは優先キューに入れられ、宛先はこれを見て NAKの形で送信元に戻す
- これは2つの大きな役割を果たす
  - ドロップした事を明示的に通知
  - 輻輳通知となり、転送レートを落とし、輻輳パスを通らない様に迂回する

# Advanced Connection Setup and Host-based Flow Control

- そもそも輻輳を回避する事がより良い方法になる。UECには“Ephemeral Connections” (儚い接続)と2つの輻輳制御スキームを導入してる
- エフェメラル・コネクションは、データが流れ始めるまでのラウンドトリップ・ハンドシェイクの遅延をなくすことで、高速な接続開始を可能にする。明示的な終了は必要無い





# Advanced Connection Setup and Host-based Flow Control

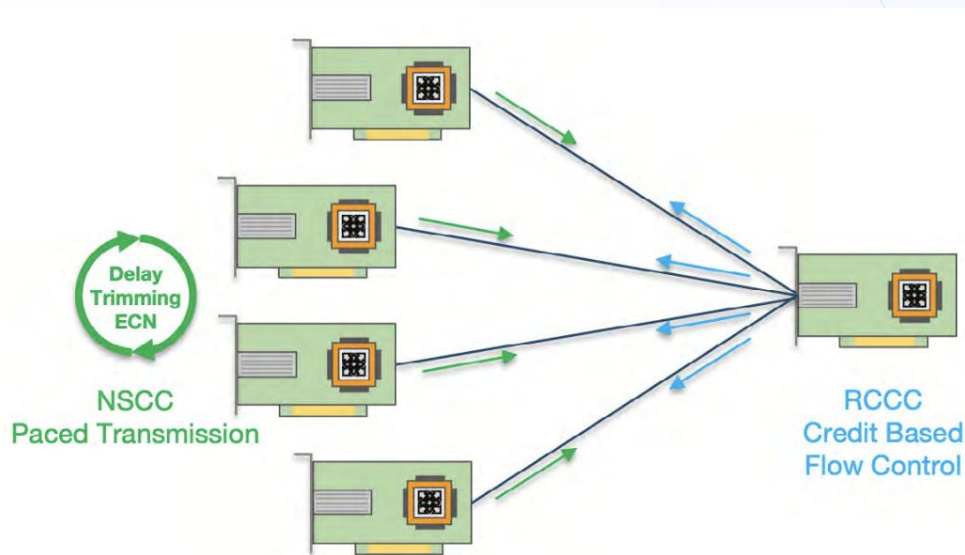
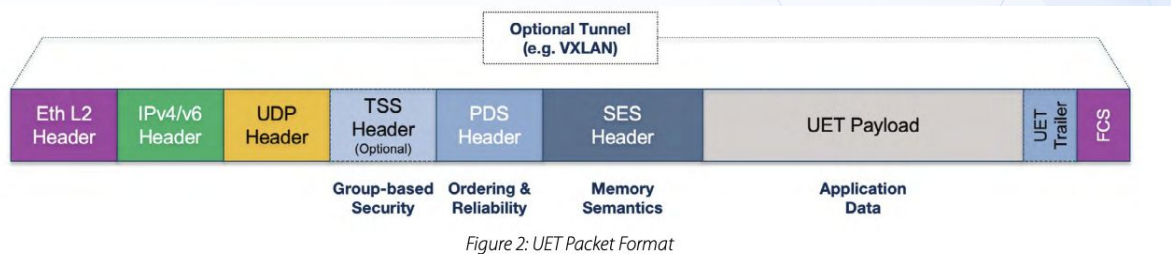


Figure 6: NSCC and RCCC Operation

- Network Signal Congestion Control (NSCC)は送信者ベースの輻輳回避手法で下記を利用し送信ペースを配分する
  - ネットワーク遅延
  - トリミングパケット
  - ECN通知
- Receiver Credit Congestion Control: RCCC)は受信者ベースの輻輳回避。各受信者がクレジットを作成し、全ての送信者に公平に割り当てる。これによりTCPインキャストの様な事を排除出来る。
- NSCCとRCCCは独立して動作する事が出来る

# Security



- UETではセキュリティは第一の目的
- AES-GCM、量子計算機暗号(PQC), Key Derivation Functions (KDF)、リプレイ防止などの実績のある技術を活用したオプションのエンドツーエンドの暗号化と認証が、UETホスト間で動作
- 1つのグループ・キーは、例えば1つのテナントが運用するすべてのXPUなど、1つのjobの全メンバーで共有される。

# Additional Future Capabilities

- LLR(Link Level Retry)

各スイッチ・ポートに小さなバッファを実装することで、個々のリンク単位で再送信を行うメカニズム。

- Credit-Based Flow Control (CBFC)

受信スイッチが予約できるスペースと同数のパケットを正確に要求することができる。

- これらの新機能は、スイッチング・シリコン内で新たな論理設計を必要とし、将来の次世代システムで利用可能になる。

# 参考

- Demystifying Ultra Ethernet
  - <https://www.arista.com/assets/data/pdf/Whitepapers/Demystifying-Ultra-Ethernet-WP.pdf>
- Ultra Ethernet™ Specification v1.0
  - <https://ultraethernet.org/wp-content/uploads/sites/20/2025/06/UE-Specification-6.11.25.pdf>
- The Ultra Ethernet Consortium Launches Specification 1.0
  - <https://www.youtube.com/watch?v=jfC-1u8BR4Y>



# ARISTA

# Thank You

[www.arista.com](http://www.arista.com)