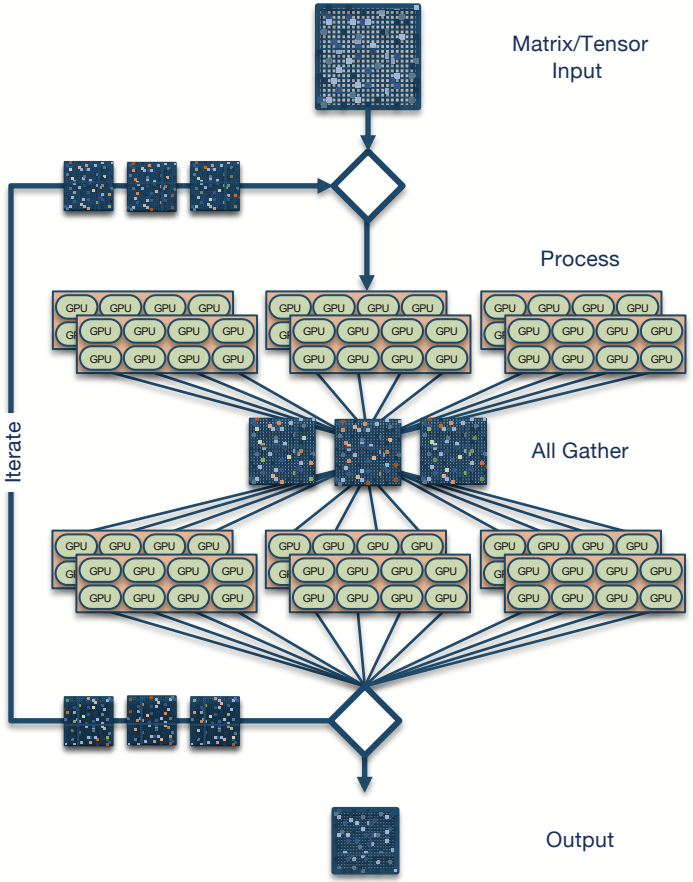
分散イーサネットリンクとAIセンター

Shishio Tsuchiya
shtsuchi@arista.com

AI/MLネットワーク要件

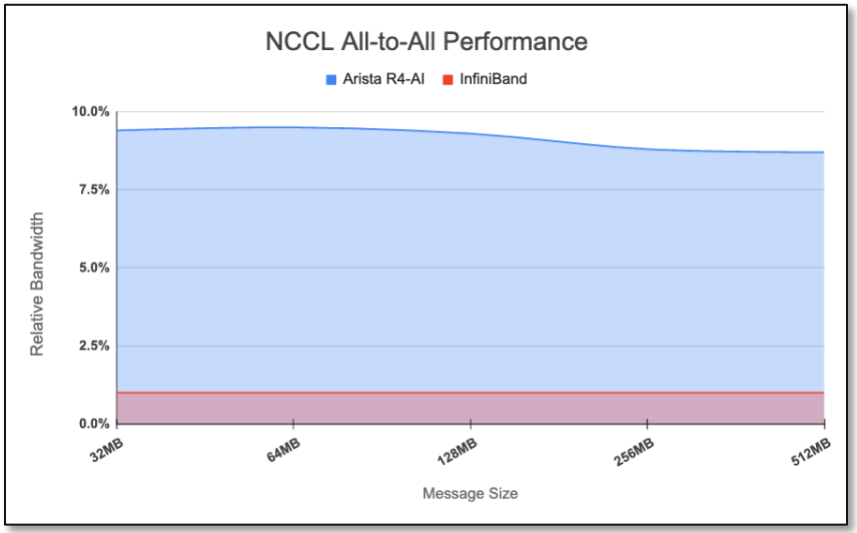
- スケールアウト要件
  - 100 -> 1000アクセラレーター->100K アクセラレーター
- ノンブロック高帯域
  - 400G->800Gのコンピュートノードへの接続
  - オーバーサブスクライブをせず、回復力のあるフェイルオーバー
- 難しいトラフィックプロファイルへの対応
  - 大きなメッセージサイズ
  - 少ないエントロピーの長時間フロー
  - 同期性の高いバースト
- ロスレスフォワーディング
  - RoCEとDCQCNをサポートするPFCとECN
  - 輻輳を起こさないロードバランシング
  - 輻輳解消のためのバッファリング

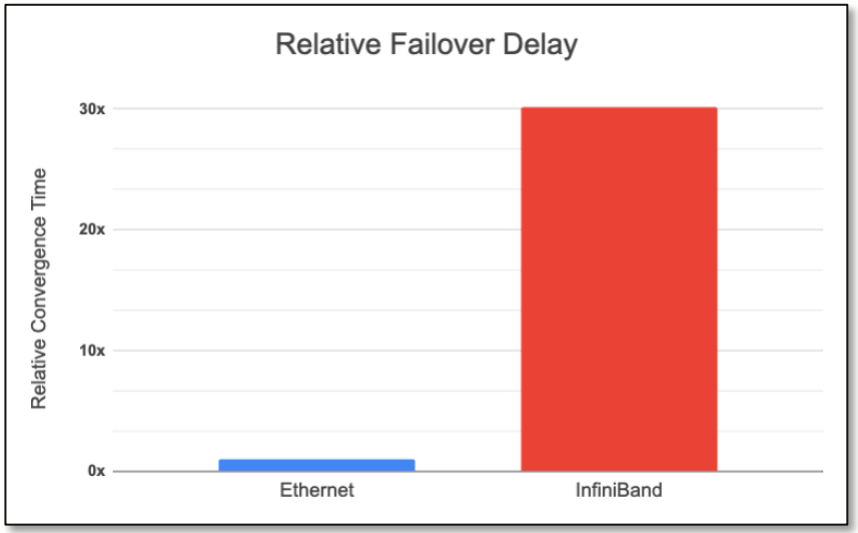


AIトレーニングのワークロードはパフォーマンス低下や遅延の影響を受けやすく相関性が高い

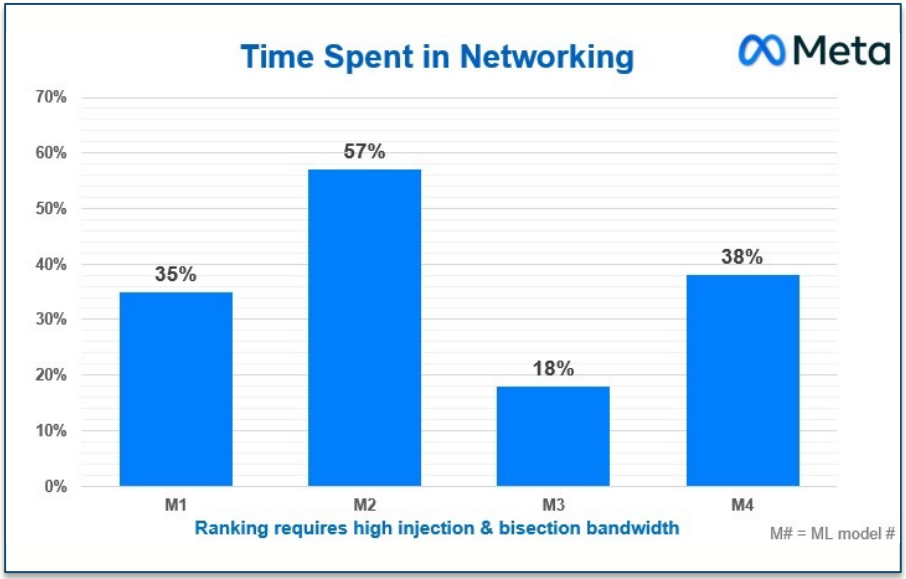
最高のパフォーマンスを実現するイーサネット＆IPソリューション

- 最適化されたクラスタパフォーマンス
  - InfiniBandを8-10%上回る性能
  - 一般的なイーサネット・プラットフォームを最大65%上回る性能
- 効率の最大化
  - 大規模ネットワークでは障害は避けられない
  - 高速フェイルオーバーでジョブを30倍高速に再開





AI/MLバックエンド・ネットワークの要件



Component	Category	Interruption Count	% of Interruptions
Faulty GPU	GPU	148	30.1%
GPU HBM3 Memory	GPU	72	17.2%
Software Bug	Dependency	54	12.9%
Network Switch/Cable	Network	35	8.4%
Host Maintenance	Unplanned Maintenance	32	7.6%
GPU SRAM Memory	GPU	19	4.5%
GPU System Processor	GPU	17	4.1%
NIC	Host	7	1.7%
NCCL Watchdog Timeouts	Unknown	7	1.7%
Silent Data Corruption	GPU	6	1.4%
GPU Thermal Interface + Sensor	GPU	6	1.4%
SSD	Host	3	0.7%
Power Supply	Host	3	0.7%
Server Chassis	Host	2	0.5%
IO Expansion Board	Host	2	0.5%
Dependency	Dependency	2	0.5%
CPU	Host	2	0.5%
System Memory	Host	2	0.5%

Table 5 Root-cause categorization of unexpected interruptions during a 54-day period of Llama 3 405B pre-training. About 78% of unexpected interruptions were attributed to confirmed or suspected hardware issues.

Data courtesy of Meta

ネットワークのパフォーマンスと信頼性はJobのパフォーマンスに不可欠

ARISTA

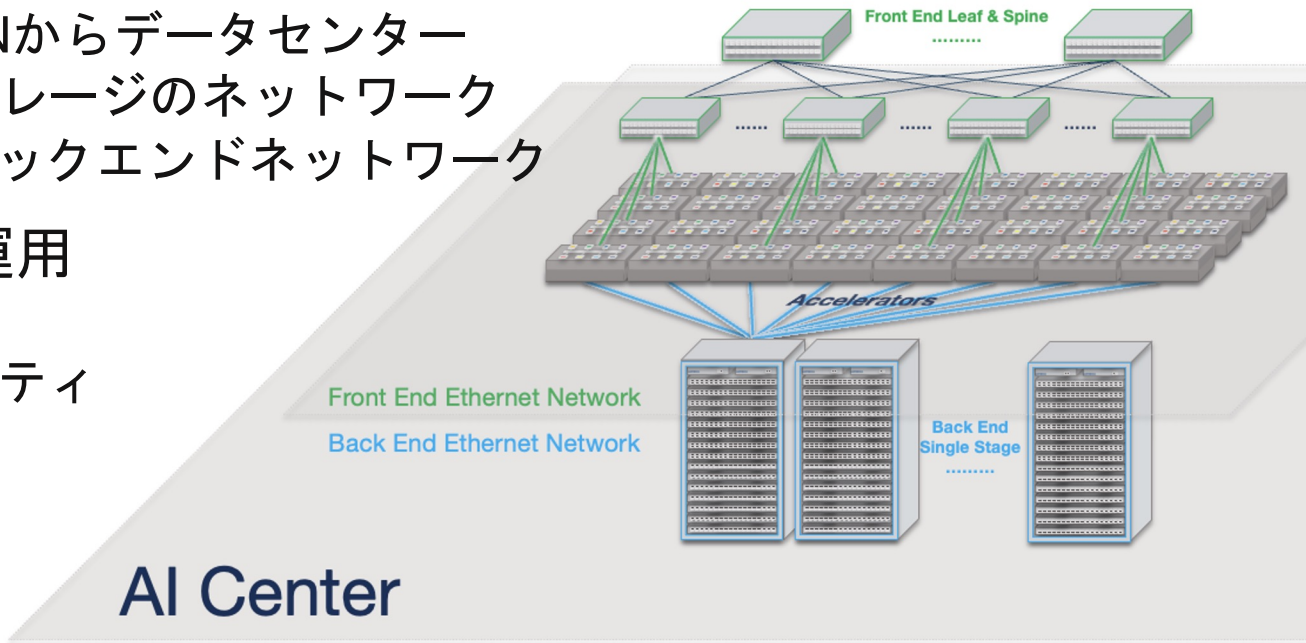
AIには専用のバックエンドネットワークが必要

	従来のコンピュート	AI センター
ホストあたりの接続性	10, 25, 50 または 100G	200, 400 から 800Gへ
設計理念	オーバーサブ	ノン・ブロッキングまたはオーバー・プロビジョニング
最適なトポロジー	2 または 3 Tier	1 または 2 Tier
128kプロセッサあたりの帯域幅	1 Petabits/Sec	100+ Petabits/Sec
トラフィック・パターン	ランダム	高度に同期 All-to-All, All-Reduce
フロー持続時間	短期間	長期間
輻輳の影響	軽微	極めて重要
ネットワークの最適化	高度に最適化	一般的

ネットワークパフォーマンスと信頼性はJobのパフォーマンスに不可欠

AIセンター

- End to End の単一技術パラダイム
  - キャンパス、WANからデータセンター
  - フロント側、ストレージのネットワーク
  - AI学習と推論のバックエンドネットワーク
- 単一のツールと運用
  - 構築、運用
  - ツールやセキュリティ
- 投資保護
  - 拡張への経済性
  - オープン、標準
  - 再配備

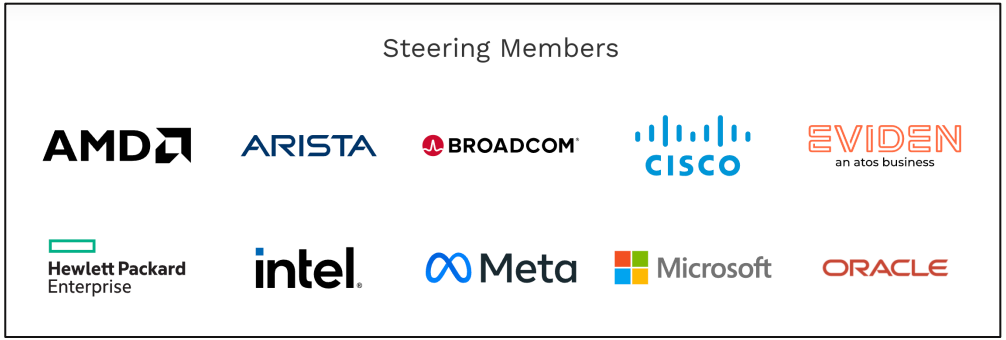


Ethernet - コストとパフォーマンスを最適化したAIインフラを実現する鍵

ARISTA

Ultra **Ethernet** Consortium: UEC

https://ultraethernet.org/



ARISTA

UECでは下記の改善を提供

- マルチパスとパケット・スプレーイング
- 柔軟な順序のパケットデリバリー
- 近代的な輻輳緩和メカニズム
- エンド・ツー・エンド・テレメトリー
- 大規模(100万エンドポイント)、安定性、信頼性

Ultra Ethernet

特徴	InfiniBand	Ethernet/RoCEv2	Ultra Ethernet
プライマリRFMインターフェース	IB Verbs	IB Verbs	libfabric https://github.com/ofiwg/libfabric
スケーラビリティのあるコントロールプレーン			
複数経路のパケットスプレイ			
フロー制御	Credit-base	PFC/ECN	Dynamic Multi-Path
スケジュール化されたFabric			
E2Eドロップ再生			
トランスポート暗号化			
マルチベンダーエコシステム			

Arista次世代 AI Etherlink ポートフォリオ

- 3 — AI最適化新800GbEシステムの選択肢を最大化
- ∞ — アクセラレータ、ワークロード、NICに依存しない
- AIワークロード向けにクラス最高のパフォーマンス
- 5nm — 最新世代の5nmジオメトリー 2倍の密度と25%の消費電力削減
- -50% — 50%以上の消費電力削減 LPOとDACケーブル

10台規模から10万台規模のアクセラレータまでAIをスケール

ARISTA

2024–5: 次世代シリコン

高機能

Trident4

4倍高いパフォーマンス

最大12.8Tbps & 132MBバッファ

プログラマブル・パイプライン

最大な帯域幅

Tomahawk5

2倍のパフォーマンス

最大51.2Tbps & 165MBバッファス

ケールアウトと高Radix

クラウド・グレード・スケール

Jericho3

800G + 高いパフォーマンス

14.4Tbpsおよび5.4Bpps

ディープ・バッファと拡張性

一貫した高性能と拡張性のあるEOS

Arista Etherlink AI ポートフォリオ

シングル構成

ボックス & シャーシ

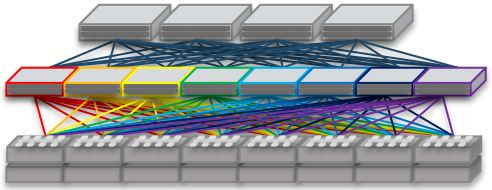
51.2T ➡ 460.8T

64-576 x 800G ノード

128-1152 x 400G ノード

低コスト、低消費電力、低複雑

リーフ・スパイン



2層、3層

リーフ・スパインや平面

数ペタbpsバイセクション帯域

DLB、PFC、ECN、

数万ノード

分散 Etherlink スイッチ

シングルホップ・インターコネクト

完全スケジュールド・ロスレス

効率性100%、セル・スプレイ

数ペタbpsバイセクション帯域

数万ノード

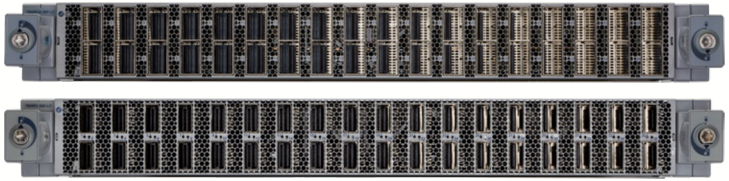
アクセラレータやNICに依存しない、オープン標準、スマートAI機能

ARISTA

7800R4 36ポート800G Jericho3ラインカード

7800R4C-36PE

AIクラスタ向け – J3-AI



7800R4-36PE and -36DE

フル機能– J3

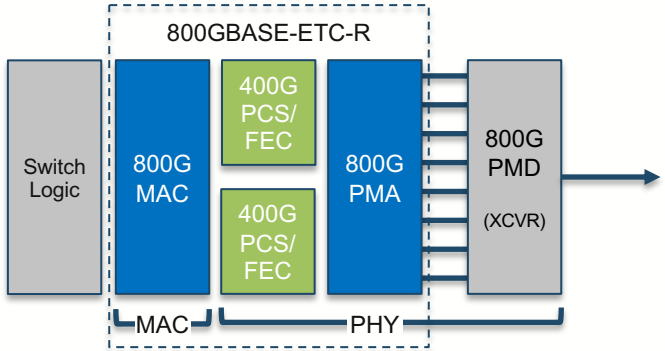
- OSFP800
- 32G HBM搭載28.8 Tbps/10.8 Bpps
- コンピュート最適化の低遅延(-4µsec)パイプライン
  - 従来のDC機能からパイプラインを削減

- OSFP800 or QSFP-DD800
- 32G HBM**搭載**28.8 Tbps/10.8 Bpps
- **プログラマブル・パイプライン、複数のオプション**
  - **データセンター**
  - **ルーティング + 暗号化**
  - **大規模ルーティング + 暗号化**

Gbpsあたりの消費電力は、7800R3比で最大 65% 減少

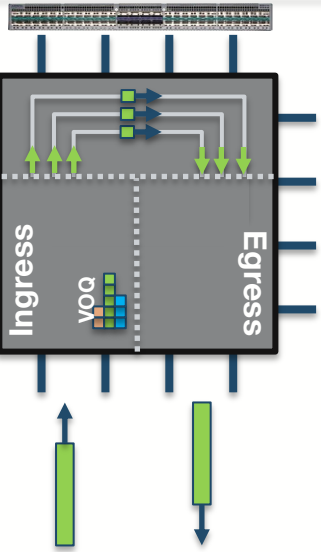
800Gイーサネットの現状

- 必要な技術
  - 100G SerDes – サポート済（TH4、7060DX5-32など）
  - 800Gトランシーバー – サポート済(OSFP800 /QSFP-DD800)
  - 800G MACレイヤー – 次世代シリコン

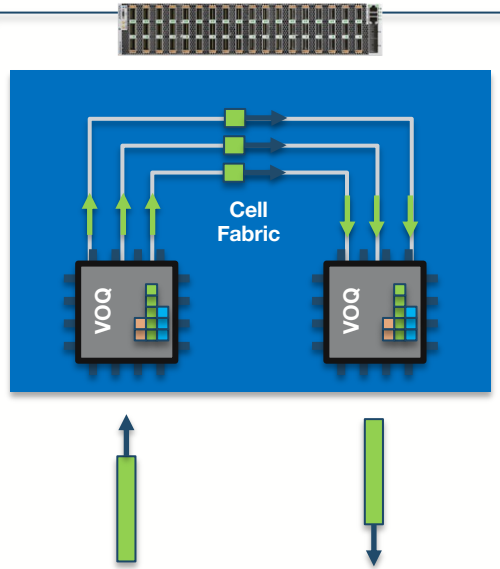


標準化	ステータス	Description	Electrical Interfaces	チップセット
IEEE 802.3ck	2022年9月	100G lanesを使った 100, 200 & 400 GbE	100G-1, 200G-2, 400G-4	TH4-100, TH5, J3ファミリー
Ethernet Technology Consortium 800GBASE-ETC-R	2020年10月	100G lanesを使った 800G	800G-8	TH5, J3ファミリー
IEEE 802.3df	2025年目標	200G Laneを使った200, 400,800,1.6 T GbE 100G Lanesを8/16使った800 and 1600 Gbit/s	800G-8, 1600G-16 200G-1, 400G-2, 800G-4, 1600G-8	将来

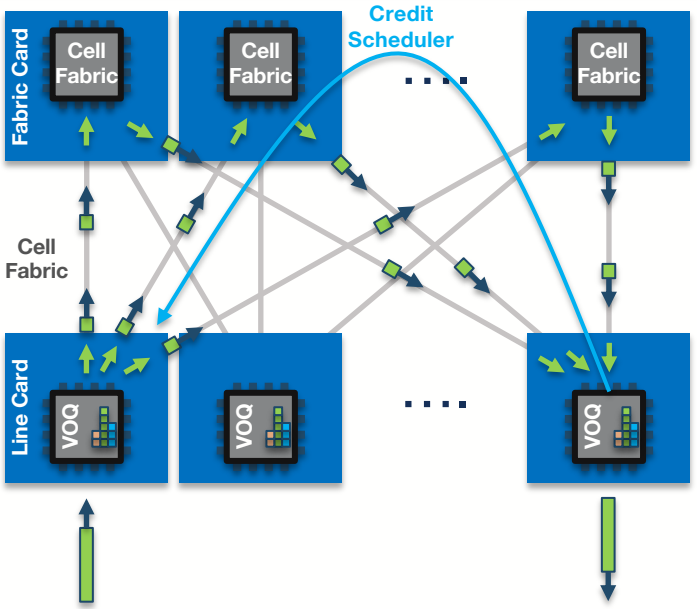
Broadcom DNX/Jerichoを使ったシステム例



シングルチップ固定型



マルチチップ固定型



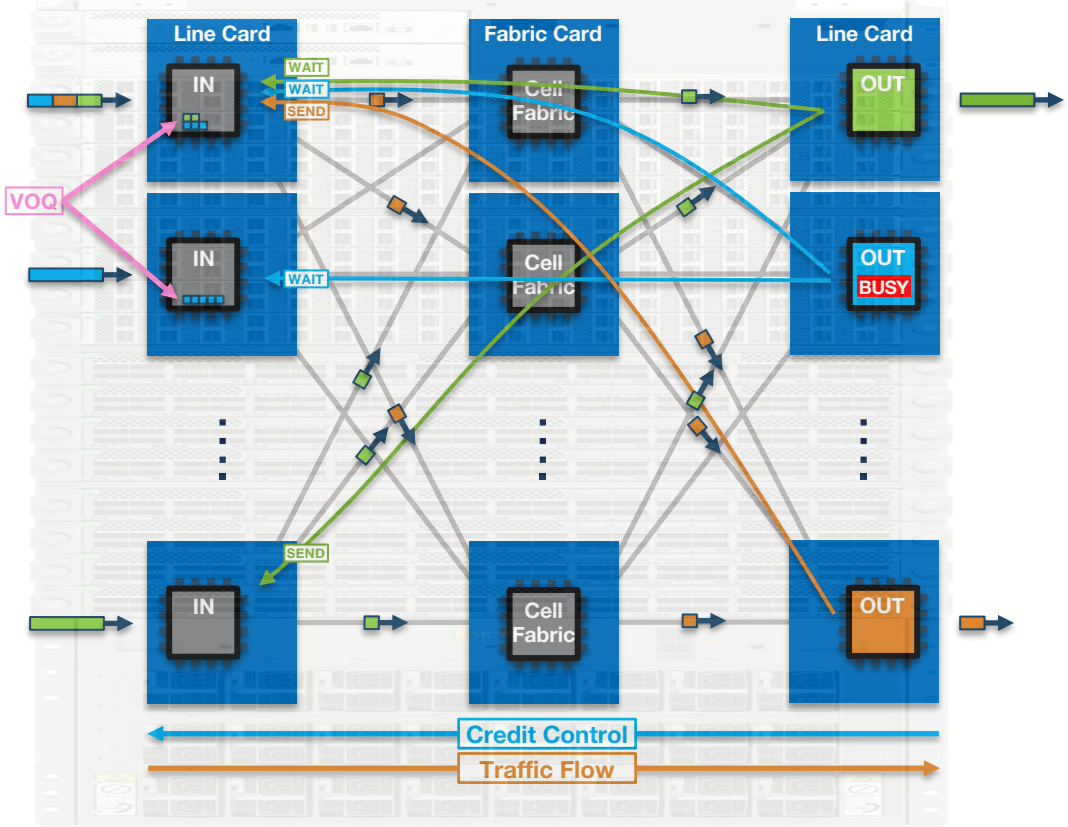
モジュラーシステム

固定システムとモジュラーシステムは共通のVOQスケジューリングを利用する

ARISTA

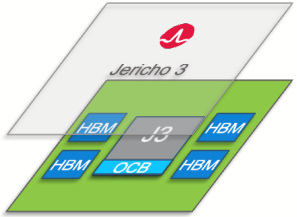
J3+Ramon3モジュラーシステム VOQ & Cell Fabric

- Ingress・パケット・バッファによるインキャストの防止
  - チップあたりの大容量バッファ（16GB）
  - ラインカードの追加に応じてリニアにスケールする
  - 増加するEgressのオーバーサブスクリプションを解消
- バッファを仮想キューに分割
  - トラフィックラスごとのVOQ
  - シングルキューのHOLBの排除
  - 公平なシステム全体のQoSを可能にする
- Egressポート別にスケジューリングされるトラフィック
  - TXが利用可能な場合のみファブリックに送信
  - ビジー時は分散されたVOQに保持
  - ファブリック内で衝突なし
- Cellベースのファブリック
  - 100%効率的なCell Spray
  - ハッシュの問題なし
  - インターフェイス速度の不一致を分離

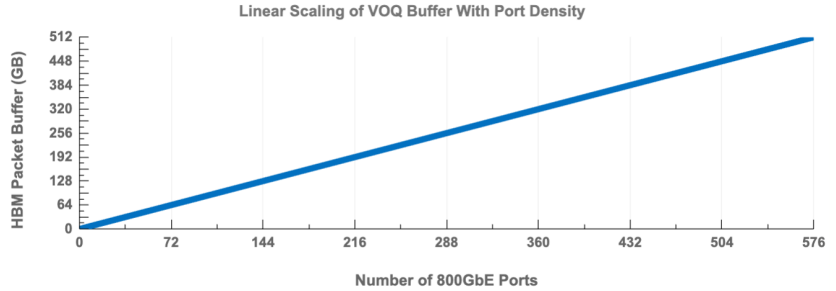


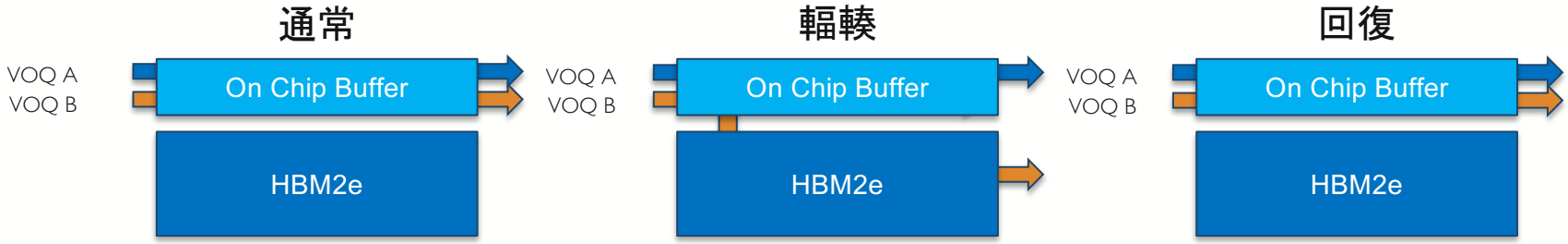
独自のアーキテクチャで数千ポートまで拡張 – ボトルネックを解消

実証済みのアーキテクチャ – 分散バッファ



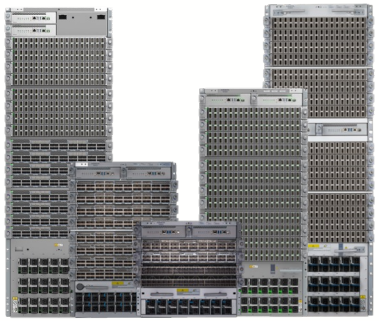
- 先進的なディープ・バッファ・アーキテクチャ
  - バースト吸収のための高速パスオンチップバッファ
  - 輻輳用大容量オンパッケージHBM2eバッファ
- VOQバッファはポート密度でリニアにスケールする
  - 各Jericho3はバッファとVOQスケールを追加する
- ダイナミックバッファの動作
  - 通常時の低レイテンシーオンチップバッファ(128MB)
  - 輻輳したフローおよび回復時にHBM2e(16GB)へのシームレスな移行が可能





R4 AIスパインをスケールアウト

7800R4 AI Spine



ロスレス完全スケジュールドVOQ
効率性100%のトラフィック・スプレイ
統合された冗長性と回復力
AIワークロードに最適化されたパイプライン

スケール
アウト

7700R4 Distributed Etherlink Switch

ロスレス完全スケジュールドVOQ
効率性100%のトラフィック・スプレイ
統合された冗長性と回復力
AIワークロードに最適化されたパイプライン

共通アーキテクチャ – AIスケールアウトへ再パッケージ

ARISTA

7700R4 分散イーサリンクシステム

- シングル構成での分散スイッチシステム
  - １システムで32Kアクセラレータまでスケール
- AIワークロードへ最適化
  - 効率性100%のVOQ/セルアーキテクチャ
- 機能豊富なEOSとCloudVision
  - シングルネットワーククラスタとして管理
  - 独立したデバイスアップグレードと交換
- 標準のEthernet接続
  - Linear-drive Passive Optics により、50%以上の消費電力削減

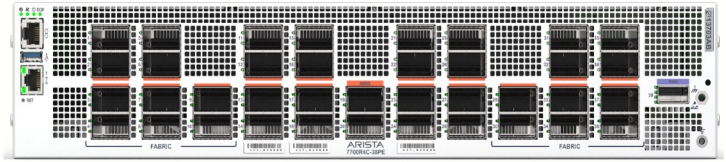
DESはAIのためのスケールアウト・ネットワーキングの最も純粋な形である

ARISTA

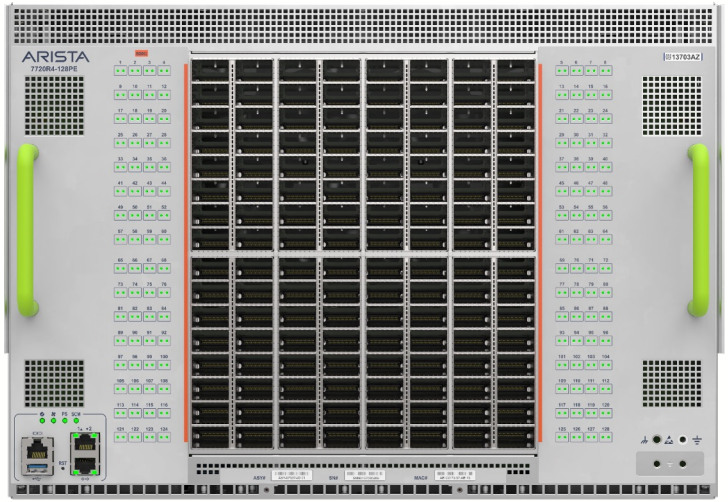
DESスマートネットワーキングでAIの利点を引き出す

利点	説明	インパクト
アクセラレータに依存しない	あらゆるXPU、あらゆるワークロード、あらゆる業種 (クラウド、ホスティング、エンタープライズ)	将来性があり、柔軟性があり、ロックインがない。
NICに依存しない	DESはロスレスで、完全にスケジューリングされ、順序を保持します。	大規模クラスタにおけるNIC最適化による大幅なコストと電力の節約。
UEC対応	UECの有無にかかわらず動作	将来性、柔軟性 - 待つ必要なし
チューニング不要	DLB、ECNなどの必要なし。 箱から出してすぐに100%効率的	L/Sの効率を80%以上にするには、大規模なチューニングが必要。時間を節約し、xPUへの投資を最大化する。
10%の冗長性を内蔵	最初から10%のオーバープロビジョニングで設計されている。	無駄なポート、中途半端なスイッチ、おかしなケーブル配線がない
高速ハードウェアフェイルオーバー	100msのリンク障害検出とリルート	低速プロトコルのフェイルオーバーなし、サブネットマネージャーなし
LPOに対応	全てのポートでLinear Drive Pluggable Opticsサポート	リーフスパインリンクで50%以上の電力削減。 ~32kクラスターで300kWh
AIのための賢い機能	可視性、高度なトラフィック管理、NICの統合	クラスタパフォーマンスの深い理解と容易なトラブルシューティング

分散イーサリンクスイッチ–構成単位



DCS-DL-7700R4-38PE/J3-AI
分散 Leaf
2U, 38 x 800Gポート
AC または DC
18x800ホスト向け 14.4Tbps
20x800 ファブリック向け 16.0Tbps



DCS-DS-7720R4-128PE/R3x2
分散 Spine
7U, 128 x 800G ポート
AC または DC

7700R4スケーラビリティ

スケール例	1152 x 800G ホスト	2304 x 800G ホスト	4608 x 800G ホスト	4608 x 800G 以上のホスト
Spineノード数	10	20	40	Super Spineを含め800まで
Leafノード数	64	128	256	1500以上
Leaf-Spine 相互接続	16 Tbps (Spine毎1600G)	16 Tbps (Spine毎800G)	16 Tbps (Spine毎 400G)	16 Tbps (Spine毎400G)

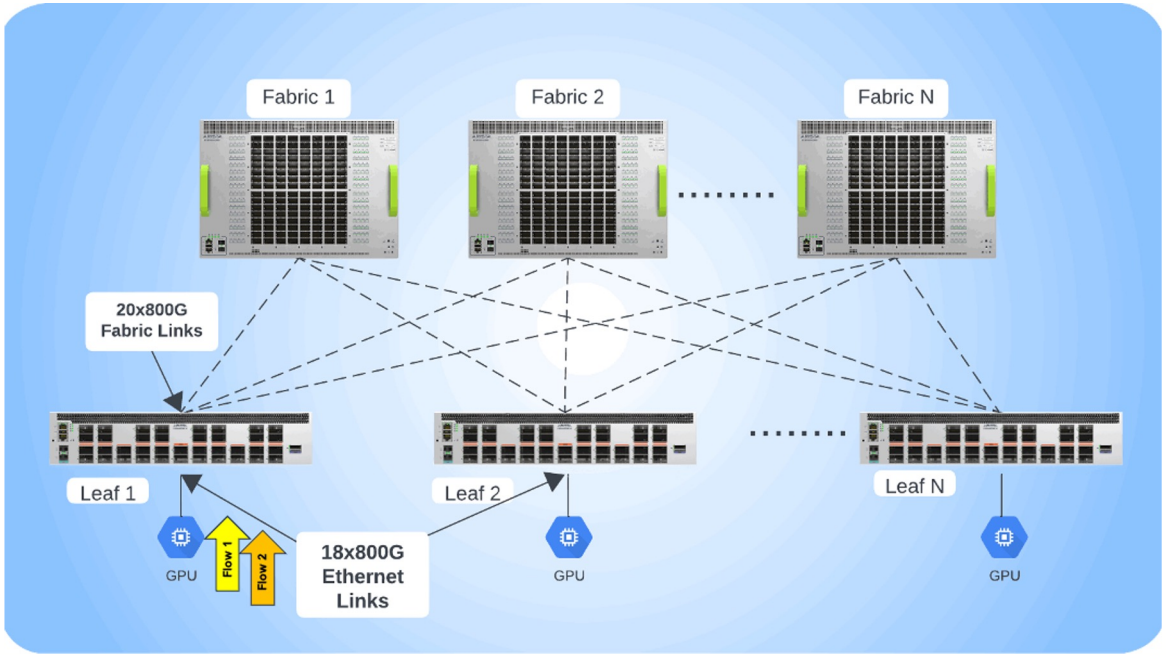
分散スイッチ/DESコントロールプレーン

- 集中コントローラもデータベースもない
  - すべてのデバイスはEOSを実行し、独立して管理され、アップグレードされる
  - ZTPはクラスタのブートストラップに使用可能
  - DES 固有のコマンドを追加した EOS CLI による設定
- プラグ &プレイ接続モデル
  - ファブリックリンクは設定不要
  - リーフとスパイン間にコントロールプレーンがない
  - ファブリックはすべてのリーフ間で論理的なL2接続を提供
- PFC、ECNはホスト側ポートで通常通り動作

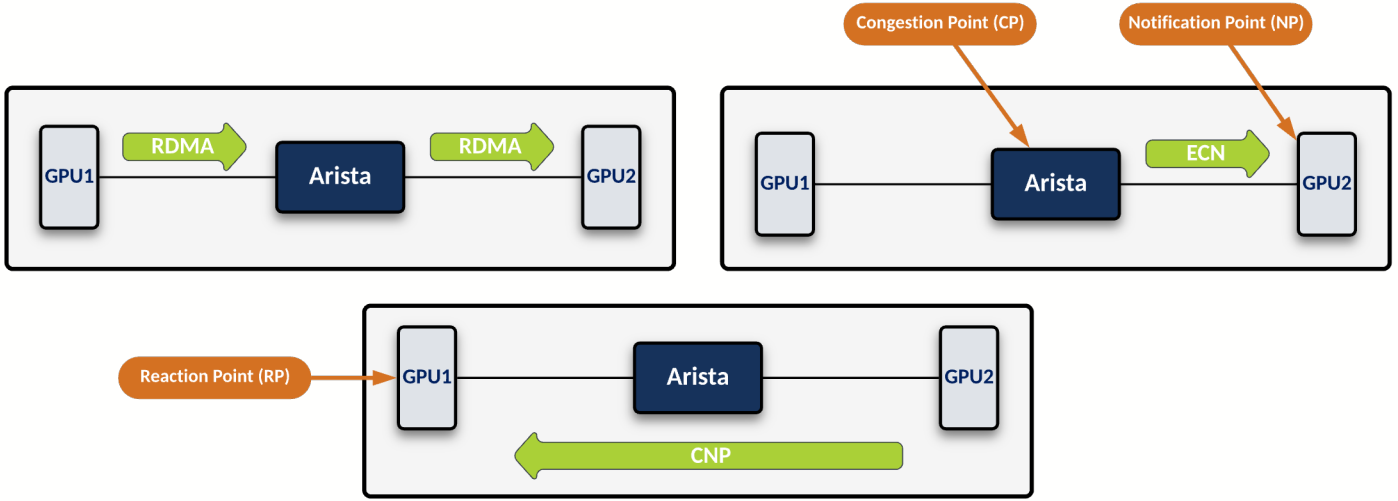
DES 利点
スケーラビリティと耐障害性に優れたコントロールプレーン
YES
プラグアンドプレイゼロコンフィグトポロジー
YES
自動化が容易な共通EOS
YES
Pay as You Grow展開
YES

分散スイッチパケットフロー

- XPU 1がパケットをleaf1に転送
- Leaf1(Ingress)はルート検索を行い、宛先Leaf#を決定
- イーサネット・パケットはVOQのingress leafでキューイング
- ハードウェアPeerツーPeerのCredit要求と提供がファブリック全体のトラフィックをスケジューリング
- パケットはセルにスライスされ、ファブリックリンクを横切ってすべてのファブリックチップ（またはプレーン）に散布される。
- 宛先Leaf（egress）はセルをパケットに再構成し、宛先XPUに転送する。



DCQCN/Data Center Quantized Congestion Notification

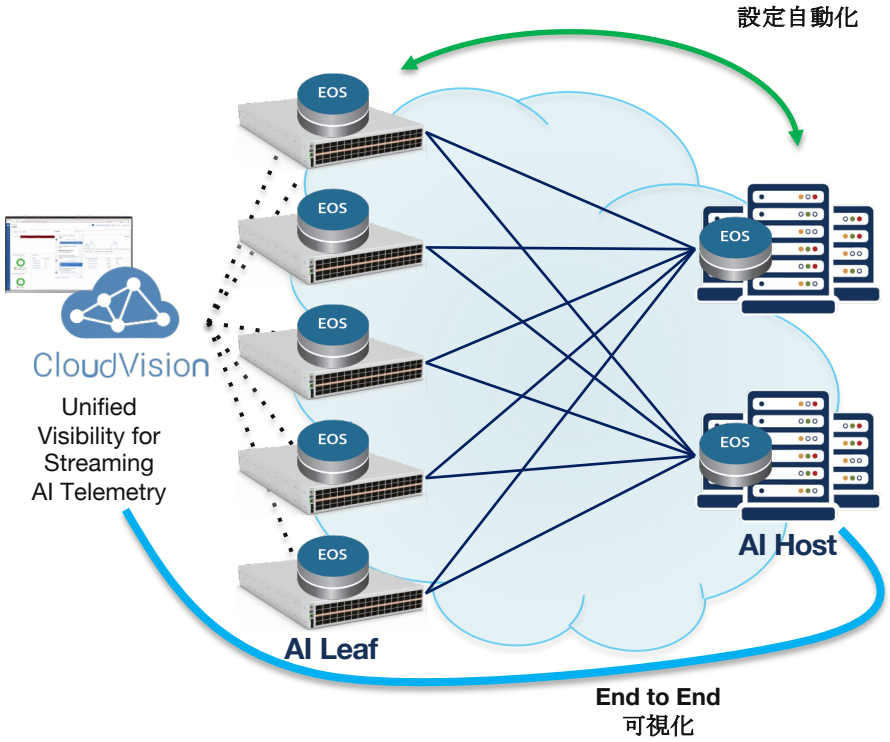


- Congestion Point(CP):ECNをマークしたスイッチ
- Notification Point(NP):ECNをマークしたパケットを受け取って、CNPを送信する受信GPUのNIC
- Reaction Point(RP):NPからCNPを受け取り、バッファオーバーフローさせないようにレートを調整

DCQCNではスイッチ/GPUノード連携し適切に
動く必要がある ECN/PFC/DSCPなど

Arista EOS AI エージェント: ネットワークとホストの協調

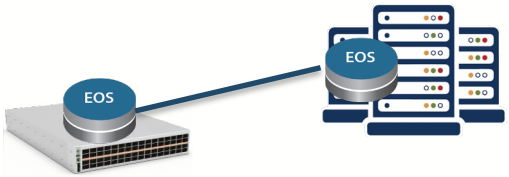
- Arista EOS Host AI エージェントが自動ネットワーク設定
- ネットワークトポロジー、GPU、CPU・NIC (Intel, ARM, or SmartNIC)に非依存
- 迅速なクラスタのデプロイ
- NICの誤設定を防ぎ、AIネットワークを最適化
- ネットワークとホストにまたがるAIテレメトリの統一された可視性

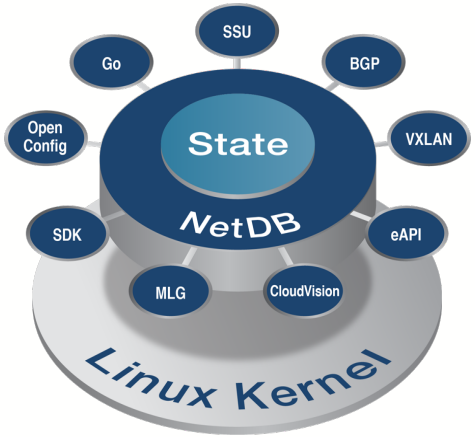


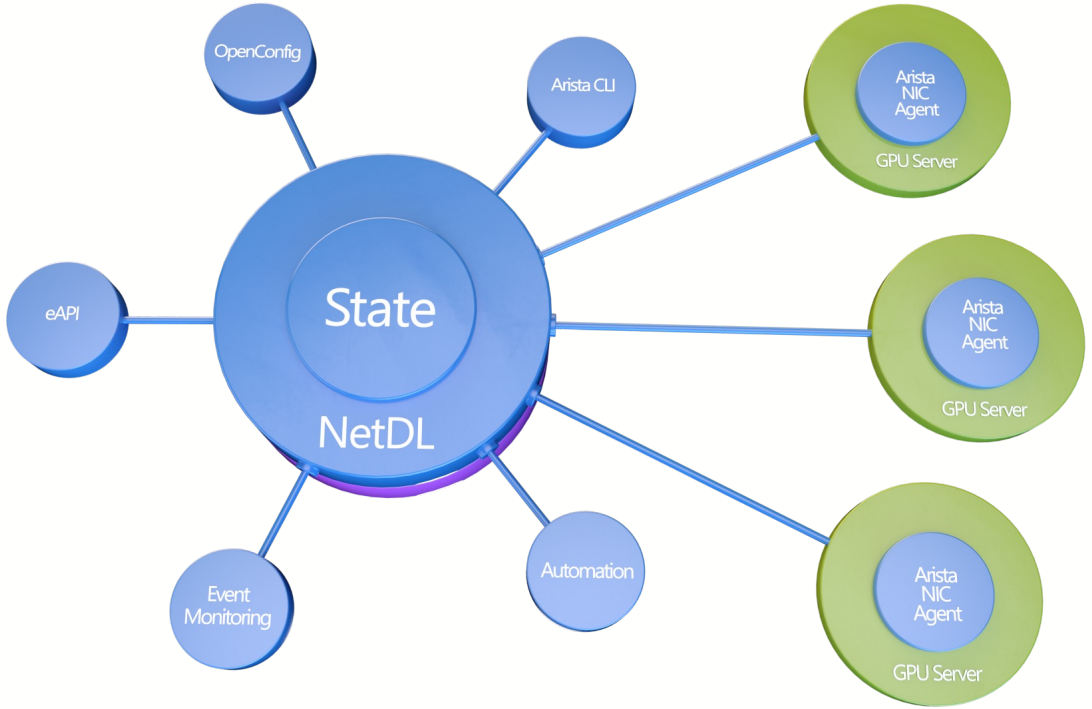
AIの最適化と保証 エンド・ツー・エンドの制御と可視性

ARISTA

Extend EOS with AI Agent





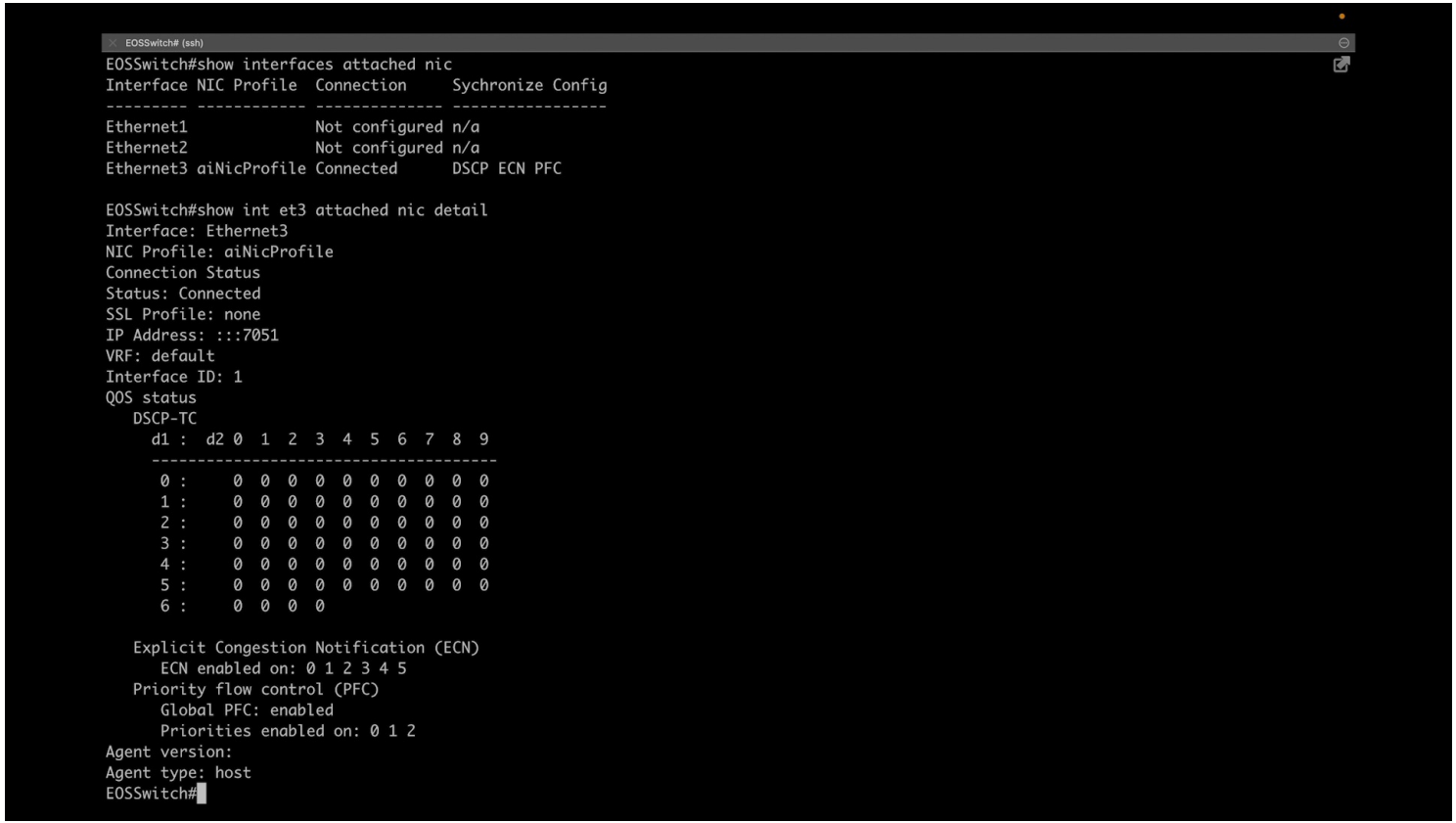


ARISTA

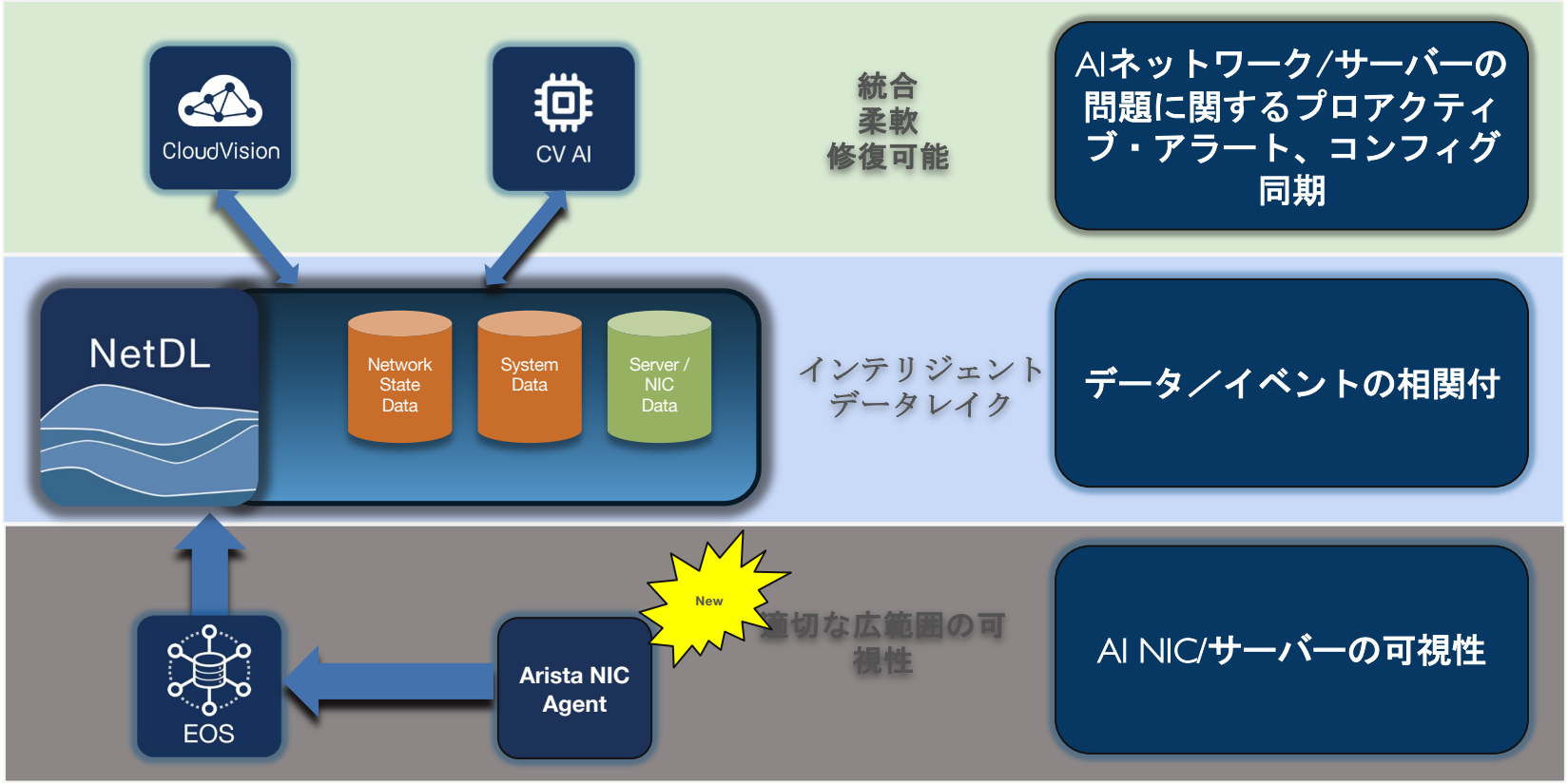
NICの自動検出

自動化されたコンフィギュレーションの一貫性

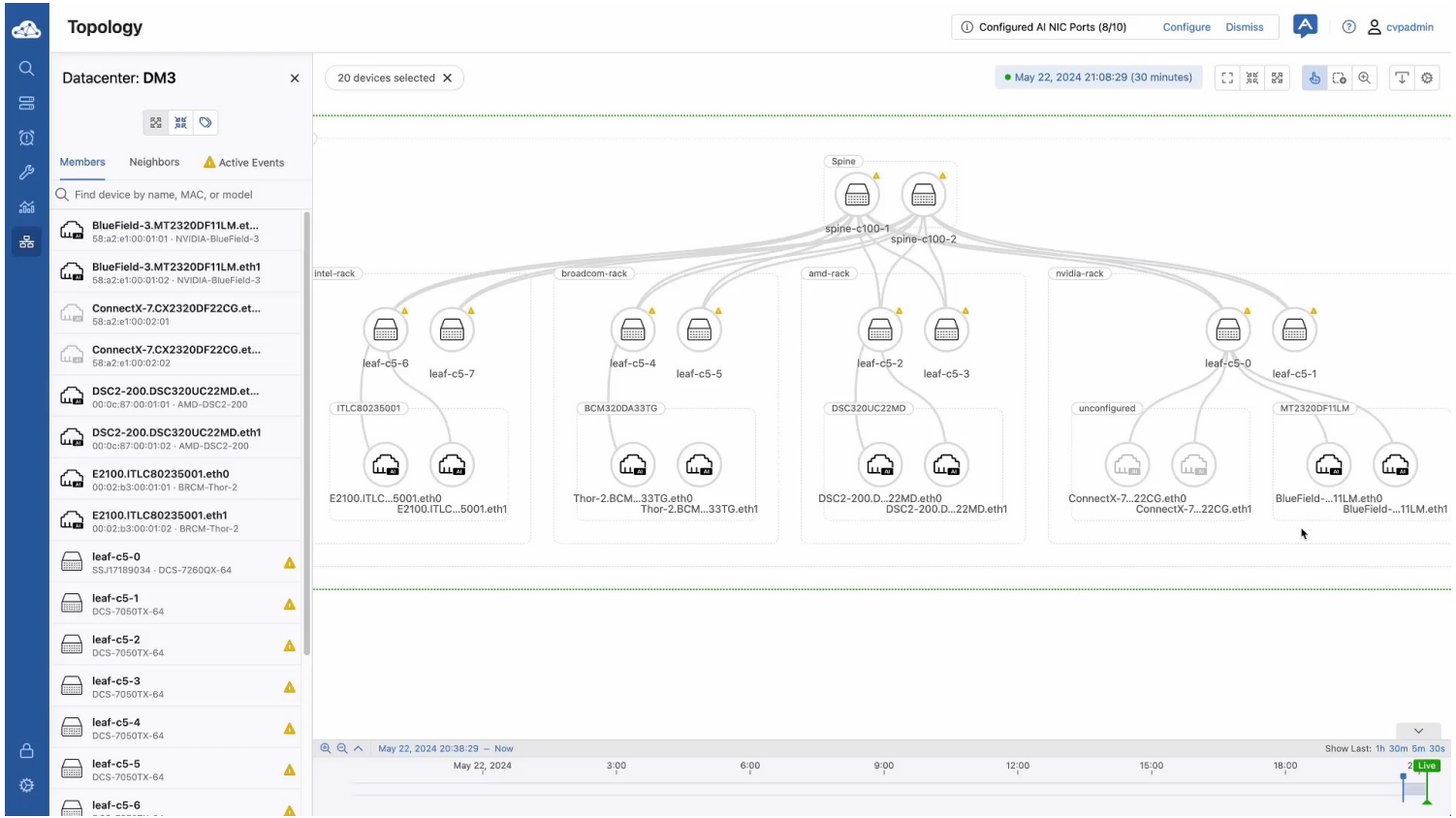
統合されたAIステータスレポート



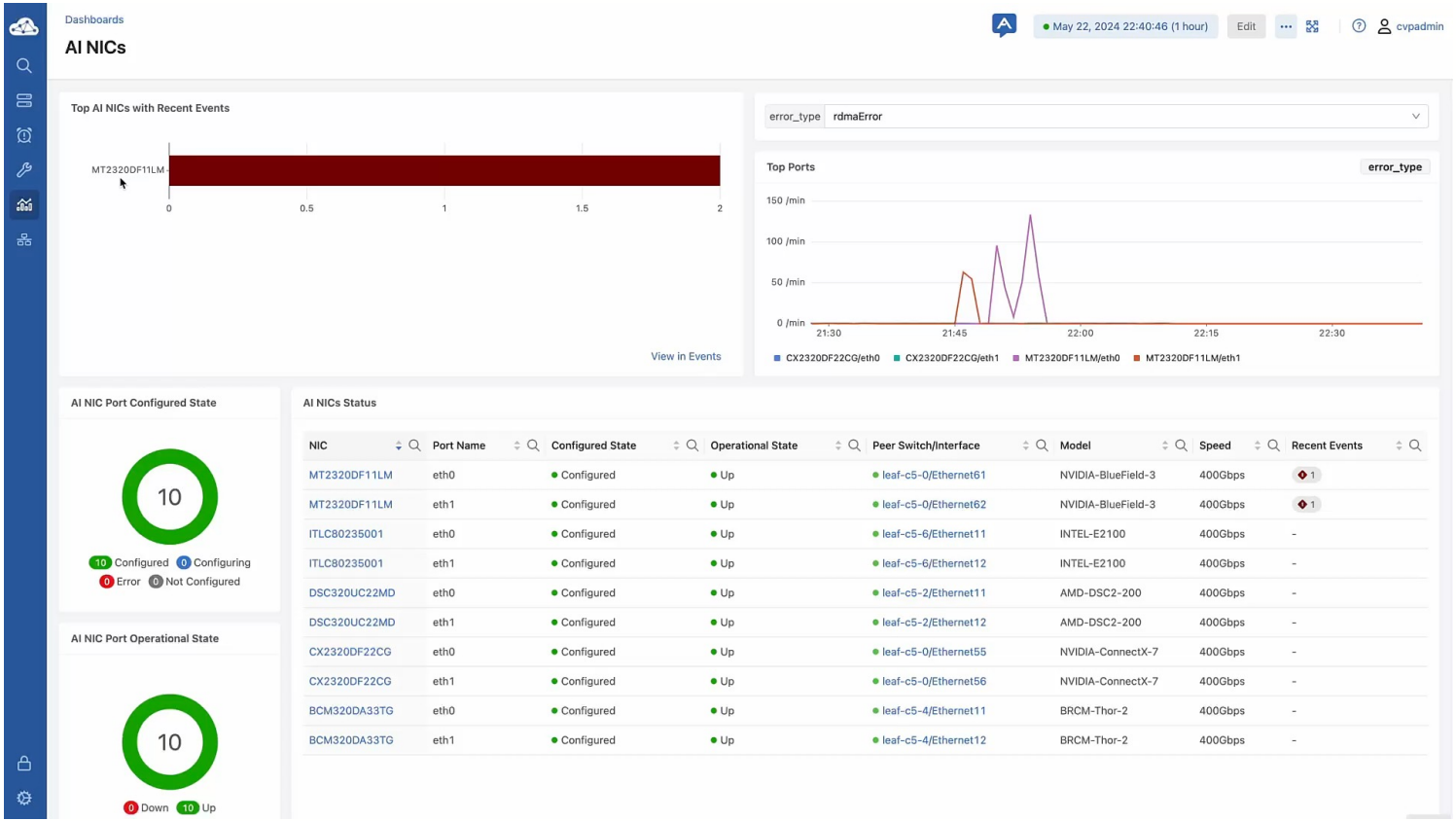
CloudVision AI



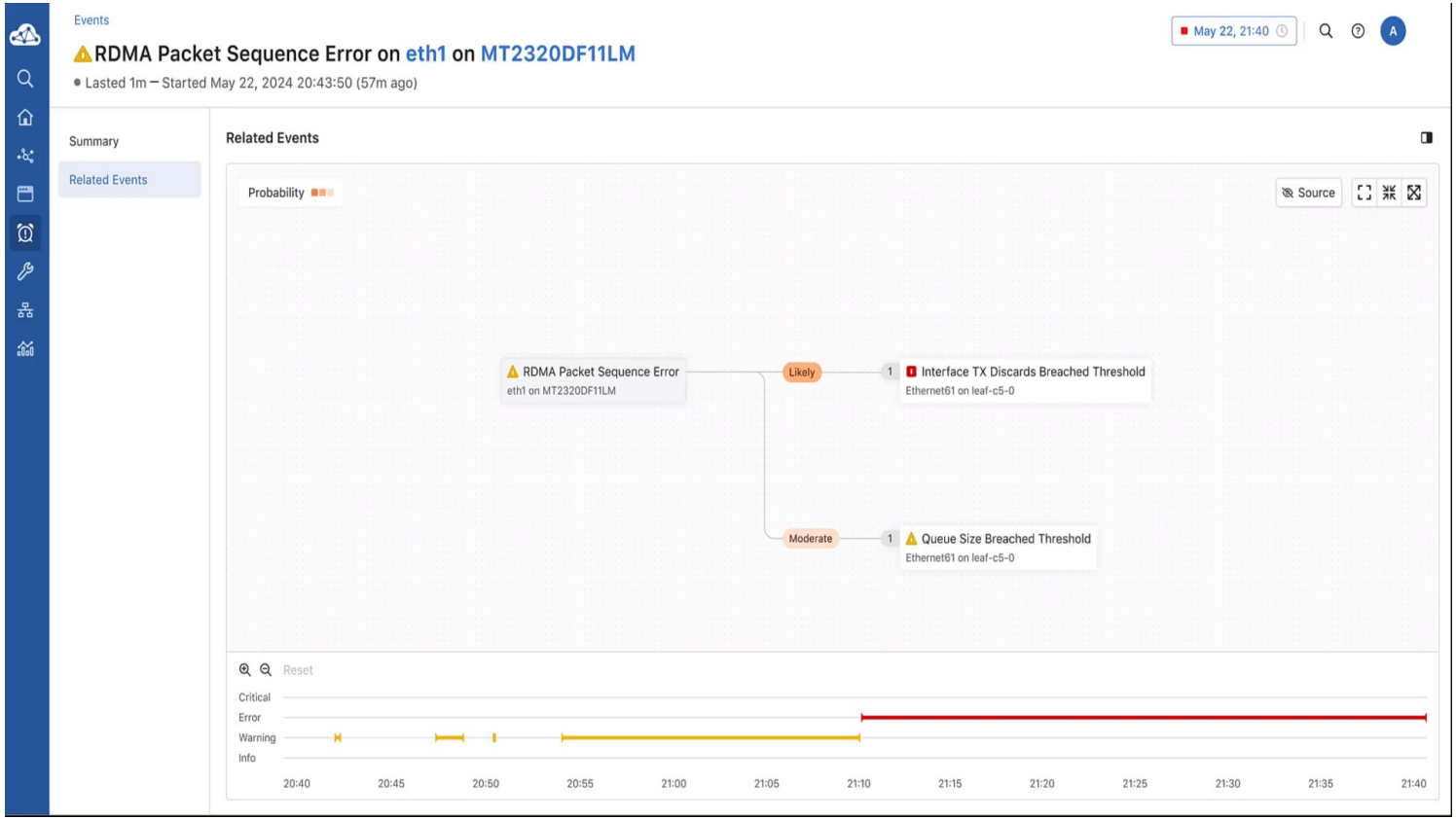
単一の管理画面



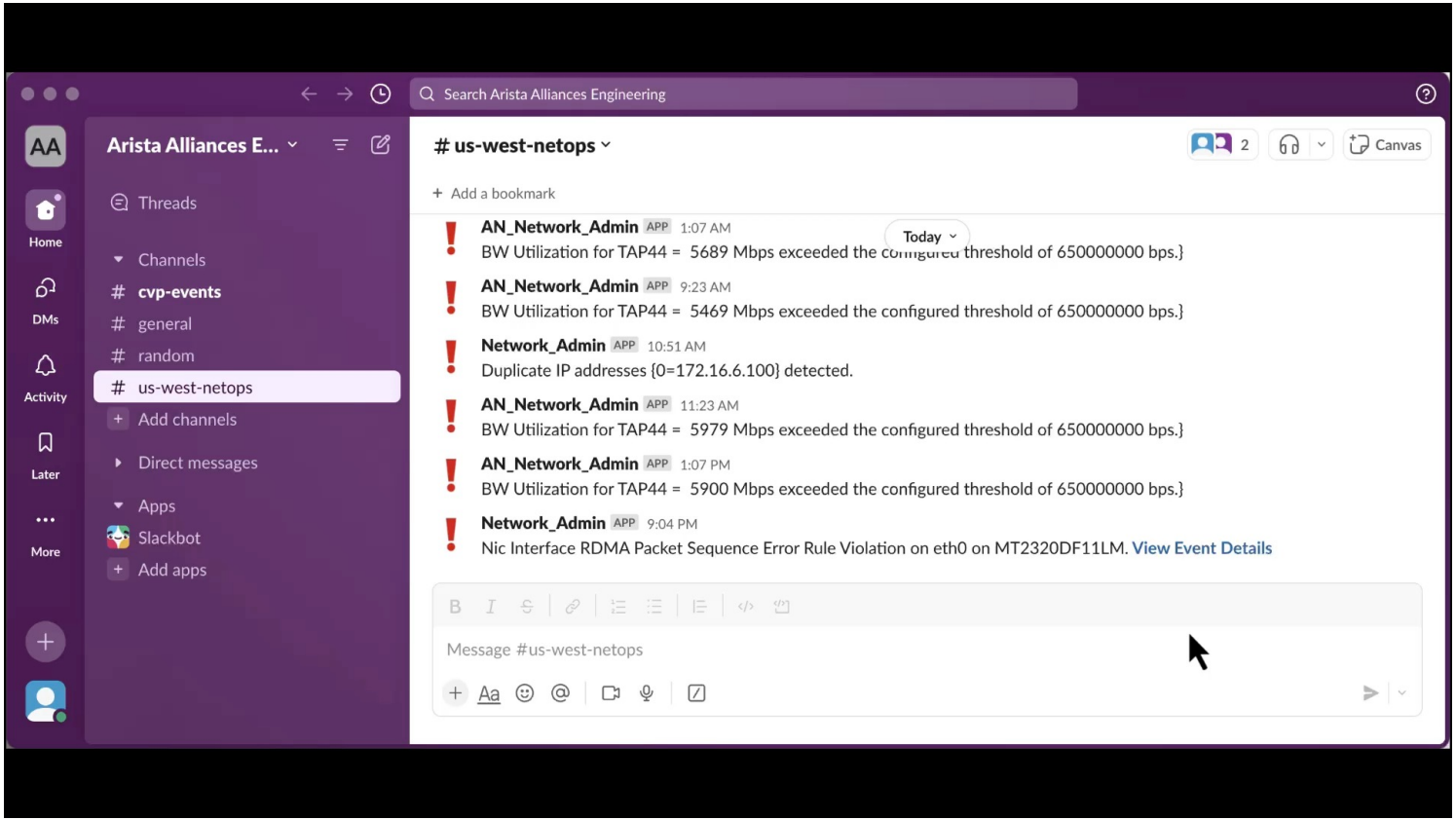
全体的なデバイスの統計



イベントの自動相関付け



自動化されたメンテナンスアップデート(Slack)



まとめ

- AIネットワークでは高度なロスレスアーキテクチャーが必要
- 現状のイーサネットの課題を克服する為にUECが現在標準化推進中
- スケジューリング化された大容量ファブリックおよび効率的なパケットスプレイを実現する分散イーサネットスイッチおよびエンドツーエンドのテレメトリーを実現するAI Agentを紹介した