

JANOG54 Meeting

生成AI向けパブリッククラウドサービス をつくってみた話

2024年7月4日
さくらインターネット株式会社
井上 喬視
高峯 誠
平田 大祐



1. 自己紹介
2. 生成AI向け基盤について
3. GPU/データセンターインフラについて
4. 導入したネットワークについて
5. トラブルシューティング
6. 会場の皆さんとの「話しあい」 （議論）



名前： 高峯 誠（たかみね まこと）

配属： さくらインターネット株式会社
クラウド事業本部 プラットフォーム部

経歴：

2024年2月～ 高火力 PHYメンバー
2021年10月～ ネットワークチーム
2007年4月～ 入社

JANOG歴：

JANOG32 「運用チームの作り方」
JANOG39 「JDCCフューチャーセッション(出張編)」



名前：井上 喬視（いのうえ たかし）

配属： さくらインターネット株式会社
クラウド事業本部 クラウドサービス部

経歴：

2023年5月～ 高火力 PHYメンバー

2015年5月～ 入社 以降専用サーバーサービスに携わる

JANOG歴：

JANOG28 初参加 / 無線LAN構築
以後だいたい参加



名前：平田 大祐（ひらた だいすけ）

配属： さくらインターネット株式会社
クラウド事業本部 プラットフォーム部

経歴：

2024年2月～ 高火力 PHYメンバー 兼任

2019年6月～ ネットワークチームメンバー

2019年4月～ 新卒入社

JANOG歴：

学生時代JANOG41に人生初参加
登壇は今回が初

生成AI向け基盤について

- 2011年11月 ● さくらインターネット自社運営の石狩データセンターを開所
- 2016年9月 ● さくらの専用サーバ 高火力シリーズの提供開始
現行提供GPU：NVIDIA V100
- 2020年7月 ● さくらの専用サーバ PHYの提供開始
- 2021年10月 ● さくらのクラウド 高火力プランの提供開始
現行提供GPU：NVIDIA V100
- 2023年6月 ● 特定重要物資「クラウドプログラム」の供給確保計画に関する経済産業省の認定 (※)
- 2024年1月 ● 生成AI向けクラウドサービス ベアメタルシリーズ「高火力 PHY」の提供開始
現行提供GPU：NVIDIA H100

(※) AI向けの高度な電子計算機提供に対する、経済産業省による認定及びNEDOからの助成金交付事業

大規模構築が可能な自社運営の石狩データセンターと経験豊富な運用人材

2016年からGPUインフラの構築運用とサービス事業の実績

専用サーバ PHYのネットワーク基盤を活用することにより最短での提供開始が可能

生成AI向けインフラ基盤に対する要望

生成AIの事業化をしたい

GPU間で高速な通信をしたい

大規模なクラスタを組みたい

早く学習を始めたい

最新のGPUを使いたい

早く学習モデルを作って試したい

たくさんのGPUが欲しい

安価なサービスで使いたい

大量のデータセットを用いるため
高速なストレージもほしい

生成AI向けインフラ基盤の規模

GPU基数

NVIDIA
H100
2000+

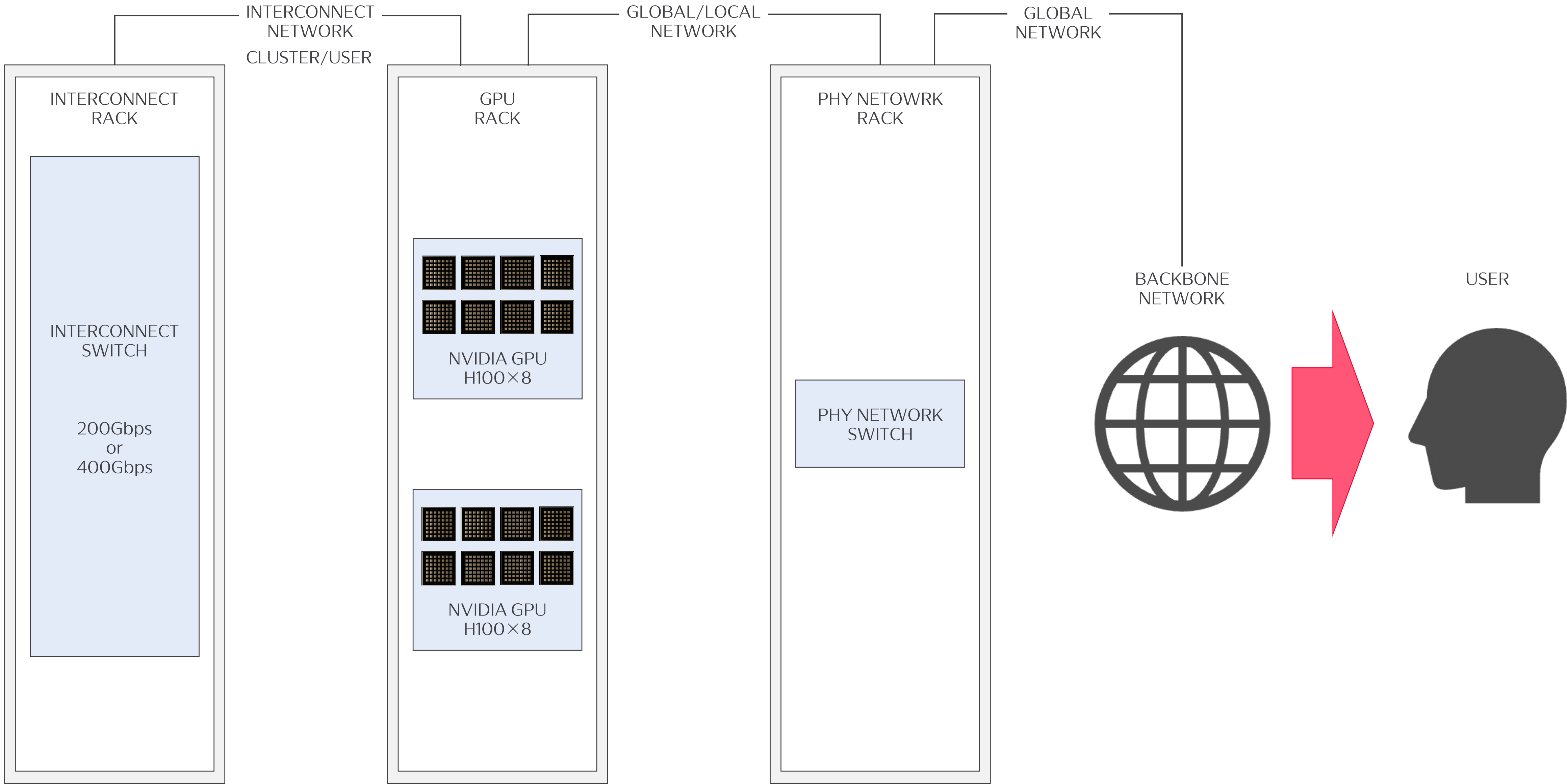
インターコネクト

200Gbps × 4
or
400Gbps × 4

計算処理能力

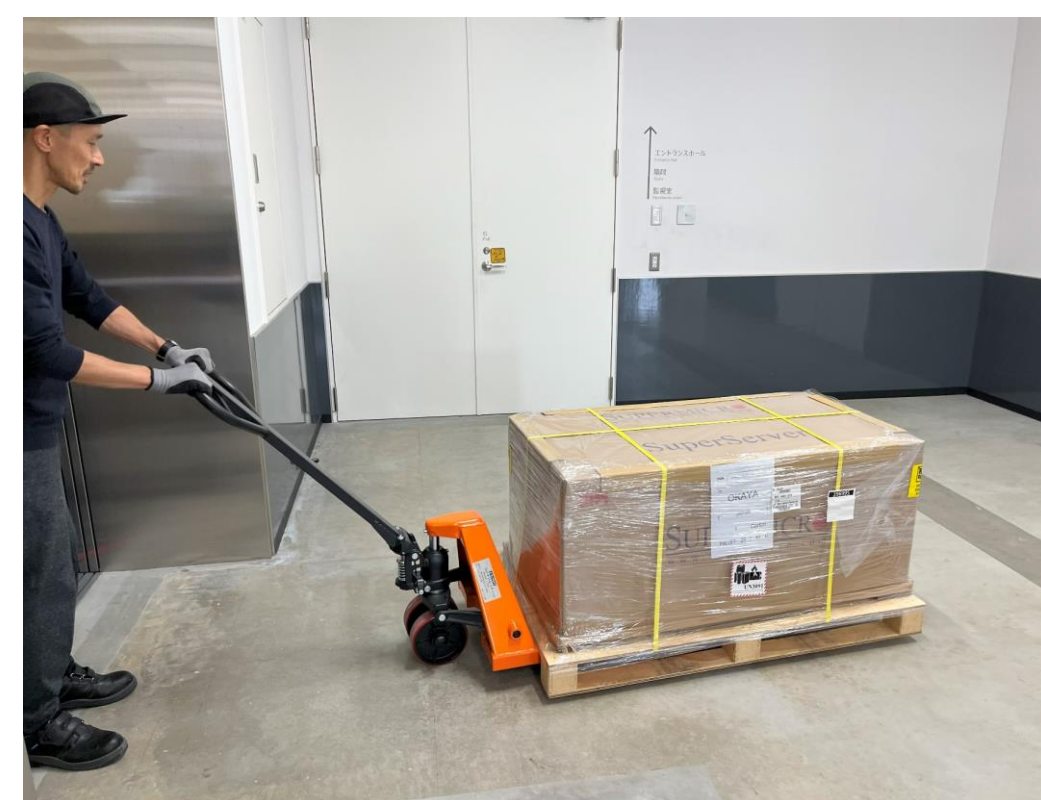
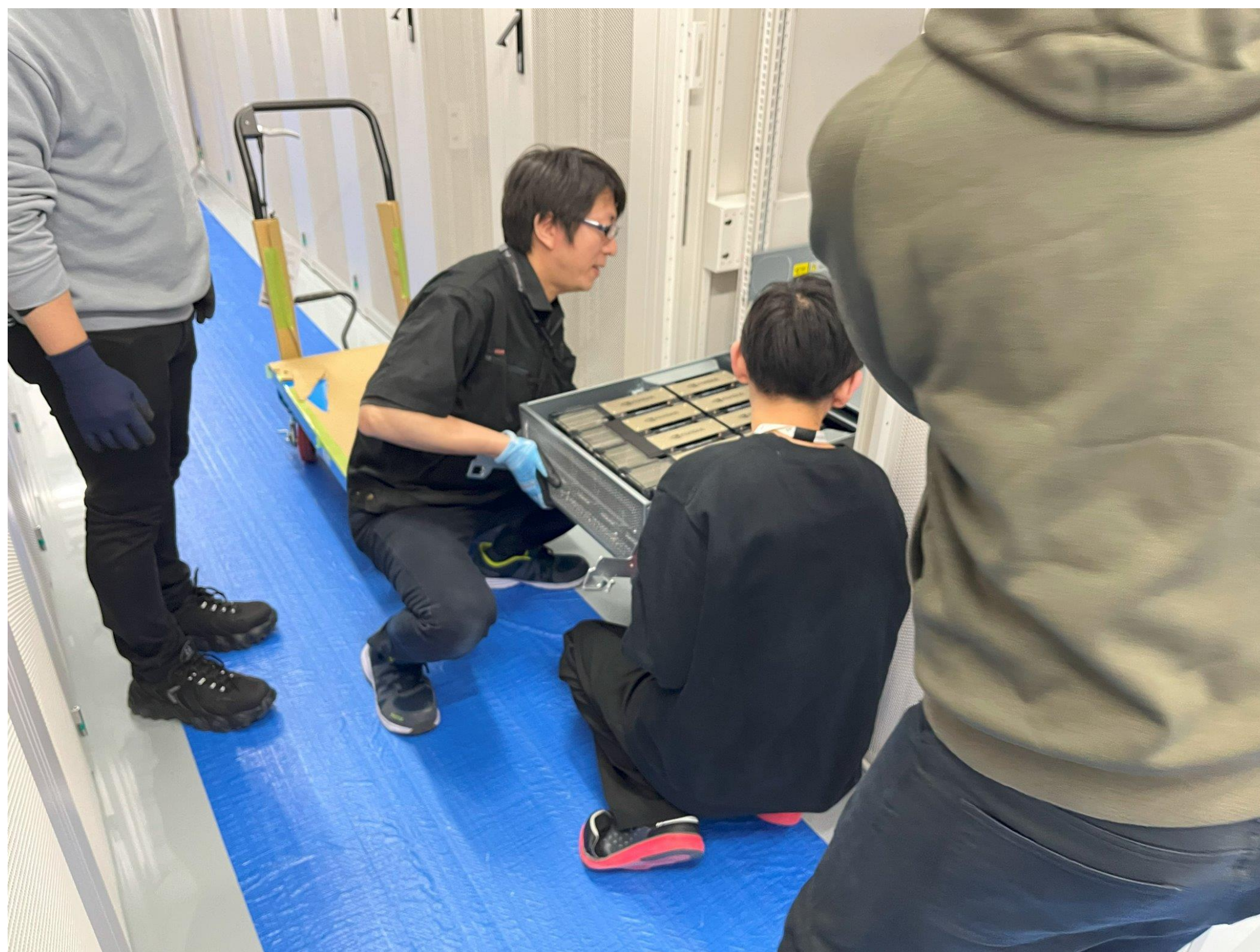
2EFlops+
(エクサフロップス)

(※) 計算処理能力は半精度で計算



GPU/データセンターインフラについて

さくら社員でGPUサーバを構築しました



構築期間 **2週間**

人数 **30人前後**/日

台数 約**200台**

さくら社員でGPUサーバを構築しました



- ラック設置の前に仮置きする場所が必要
- 100kg over/台なので、設置前にパーツ類を外し作業負担を軽減
- 開梱だけで半日程度

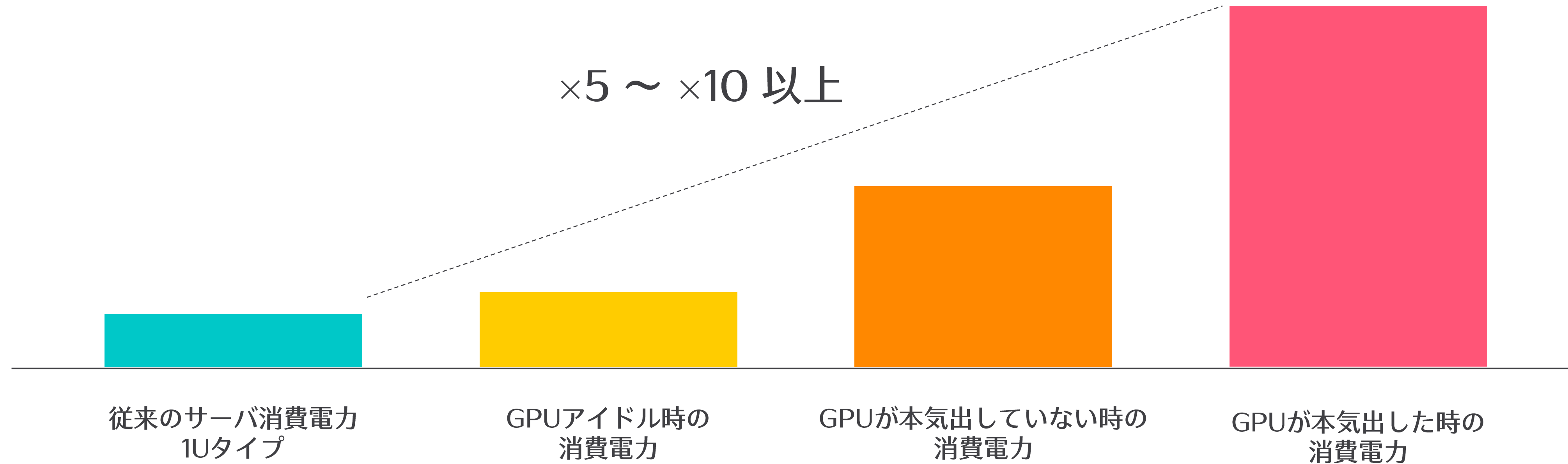
さくら社員でGPUサーバを構築しました



フォーマンセル

- 4人1組で設置、構築
- ハンドリフターは必須、人力だとケガする可能性も

消費電力



- 電力値は平均値ではなく、GPUサーバやインターコネクトスイッチのデータシートから最大値で見る
- ラック単体の消費電力設計も必要だが、サーバルーム全体の消費電力も見る
- クラスタを組んでいる場合、複数GPUサーバの電力負荷が同時に爆発増加する可能性あり

(※) 比率は発表者比の経験則
(※) GPU700Wが8基搭載されている場合

冷却能力



- 部屋全体を冷やすのではなく、アイルコンテインメントで冷たい空気と温かい空気を分離する空調方式が必要
- 排熱量が多い筐体の周囲ラックには何も置けないことも
- それでも複数GPUサーバやシャーシ型スイッチを冷却することは難しく、ラック扉に加工や水冷式の検討が必要

騒音対策

120デシベル	聴覚に異常	飛行機エンジンの近く
110デシベル		自動車のクラクション
100デシベル	ひどくうるさい	GPUサーバがあるサーバールーム
		電車が通るガード下
90デシベル		騒々しい工場内
80デシベル	うるさい	一般的なサーバールーム
		電車社内
70デシベル		騒々しい街頭
60デシベル	普通	乗用車内、一般的な会話
50デシベル		静かな事務所
40デシベル	静か	深夜の街、図書館内

- ただでさえ大声で会話する必要があったが、人間が会話できる場所ではなくなった

騒音対策



- サーバルームには耳栓とイヤーマフ着用の標識

騒音対策



耳栓
MOLDEX製



イヤーマフ



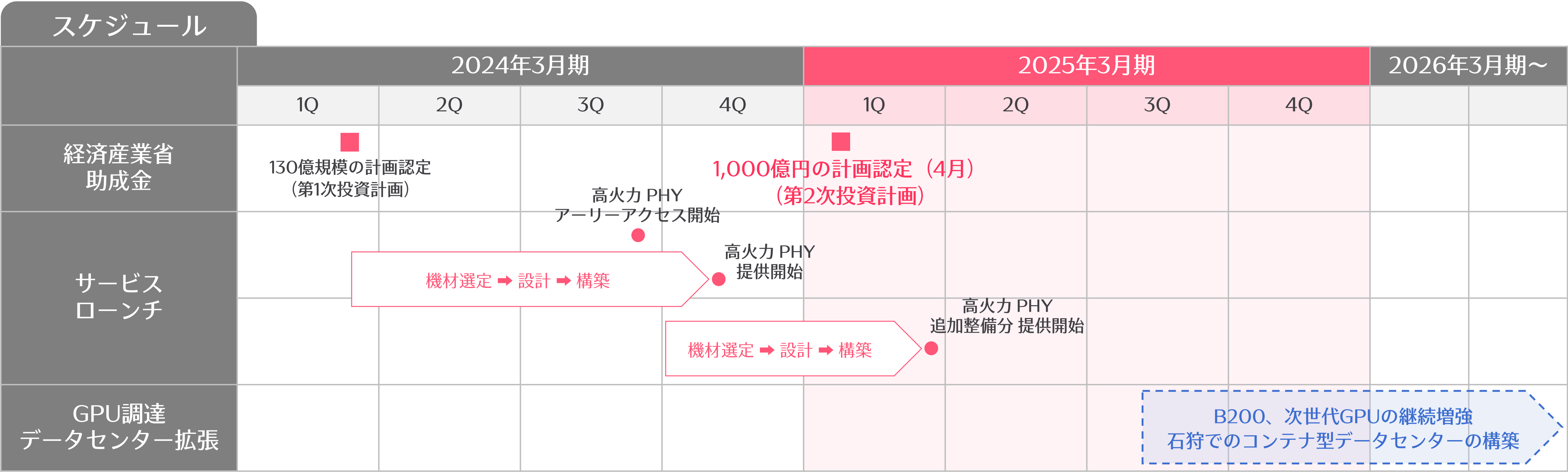
Bluetooth対応高性能インカム
SENA製 (※) バイク向け

導入したネットワークについて

準備を開始してからサービスリリースまでの期間はおよそ半年間

とにかくスピード感重視

- すぐに機材の選定、調達、設計、構築と進めていく必要があった
 - ・ サービスのリリース目標日が決まっていた
 - ・ 前提となる知識も十分ではない状況からのスタート
 - ・ コロナ禍が明けてきていたとはいえ、各ベンダーのリードタイムもまだ回復しきっていなかった



さくらの専用サーバー PHY のサービス基盤を利用してスピードアップ

満たせるもの

- コントロールパネルやドキュメントといったサービスとしての仕組み
- インターネットアクセスやローカル通信をユーザーに提供するためのネットワーク基盤
- サーバーやネットワーク機器のゼロタッチプロビジョニング基盤
- サーバーやスイッチを接続するための配線仕様

足りないもの

- GPUサーバー同士を接続するための**ロスレスネットワーク**
- GPUサーバーのNICやトランシーバーの機器仕様
- GPUサーバー、ロスレスネットワーク用装置の収容・配線設計



さくらの専用サーバ PHYとその裏側 ～物理ホスティング環境構築自動化の取り組み～（前編）

<https://knowledge.sakura.ad.jp/26536/>

さくらの専用サーバ PHYとその裏側 ～物理ホスティング環境構築自動化の取り組み～（後編）

<https://knowledge.sakura.ad.jp/26550/>

できる限りシンプルであること

- とにかく時間が無いので複雑な構成を避けたい
- 運用の負荷も減らせるように、シンプルなアーキテクチャが望ましい

マルチテナンシー

- パブリッククラウドサービスとして提供するためマルチテナンシー要件は必須
 - InfiniBandベースのファブリックではなくEthernetベースでの設計に
- マルチテナンシーを実現する技術においてもシンプルなものが望ましい

公平性

- パブリッククラウドサービスであるため、できる限りユーザーに対して公平なサービスを提供したい

シャーシスイッチによるシンプルなアーキテクチャを採用

さまざまな機種/構成を比較し検討を行った結果、Arista Networks 7800R3を採用

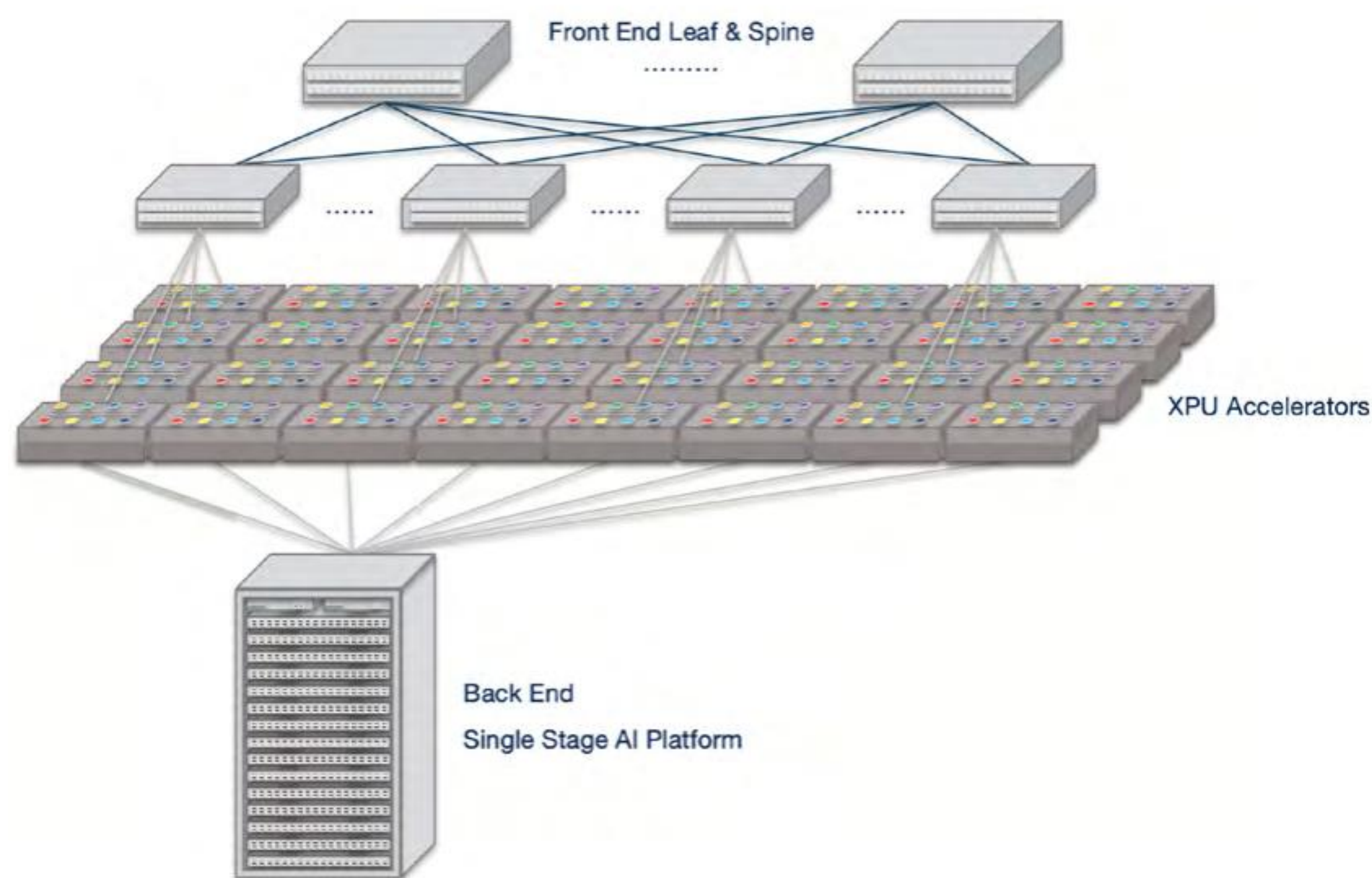


Figure 11: Moderate-scale AI Center, Single Stage based on one Arista 7800R modular chassis

シャーシスイッチ採用の利点 ①

● 1台で大容量

サービスとして提供するすべてのサーバーを収容するに必要十分なファブリックを有しており、かつスパインスイッチが不要なのでトランシーバーやケーブルなどの構築に必要なパーツ点数が削減できた

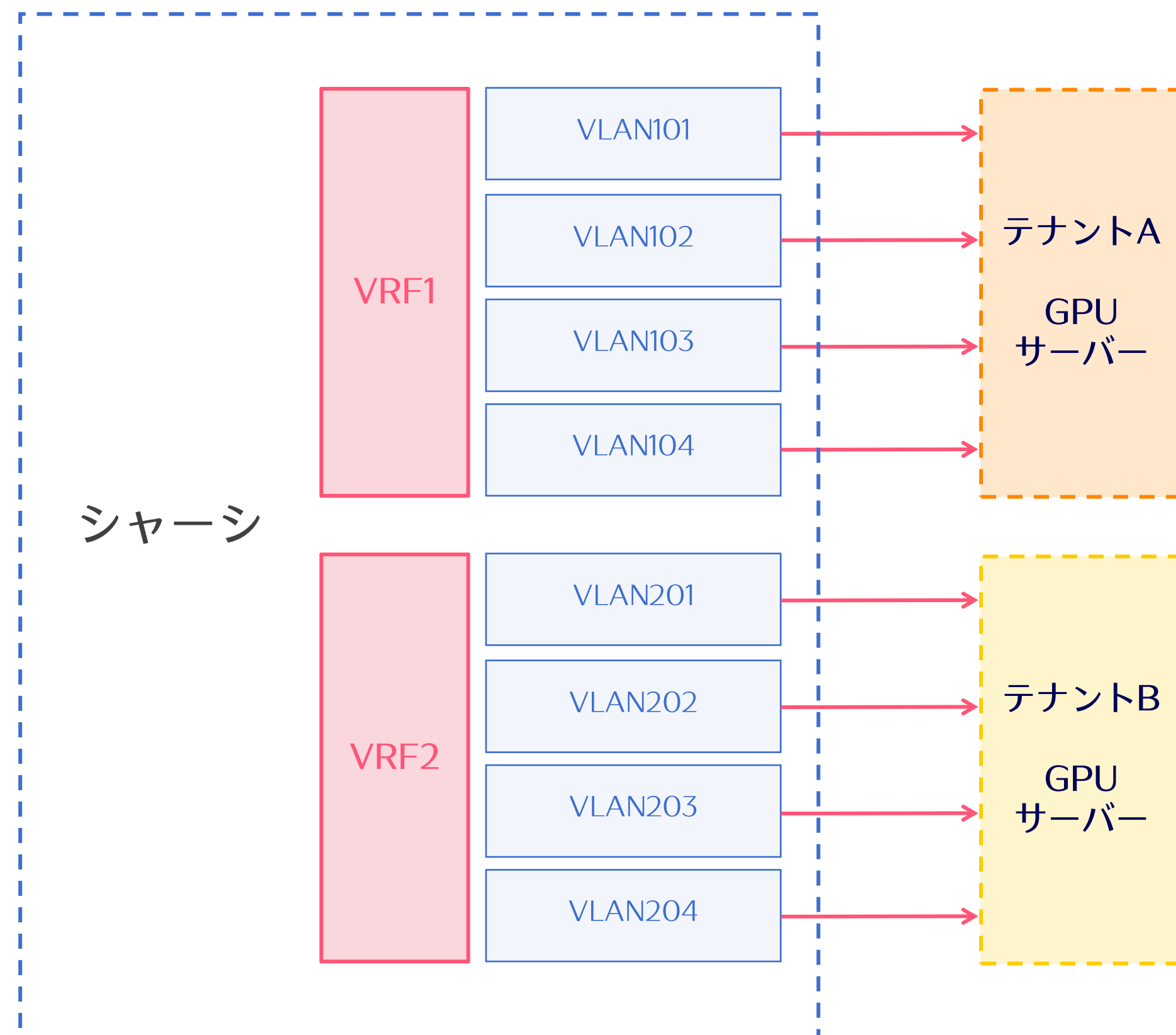
● VOQベースのファブリック

ラインカードがリーフ、ファブリックモジュールがスパインに相当する構成で、シャーシ内部はセルスプレーによるスイッチングにより高効率（ノンブロッキング）で、公平性に優れる、入力されるパケットもVOQベースの転送によりロスレス、アウトオブオーダーも発生しない

<https://www.arista.com/assets/data/pdf/Whitepapers/AI-Network-WP.pdf>

シャーシスイッチによるシンプルなマルチテナンシーの実現

EVPN/VXLANを使わない、VRFとVLANのみによるマルチテナンシー



シャーシスイッチ採用の利点 ②

- シンプル構成

シャーシスイッチは単純に1台のスイッチとして扱うことができるため、マルチテナンシーを実現する機能としてはVRFとVLANのみでよく、シンプルな構成で実現できる

- 検証時間の削減

VRFとVLANは普段からよく利用している機能であるため、EVPN/VXLAN + RoCEv2と比べてマルチテナンシー機能の検証にかかる時間が大幅に削減できた

まずは搬入経路の確保から

搬入日前日に雪が降ったため、社員が朝から除雪



Arista 7816R3のラック設置について

ハードウェアのインストール方法は動画などもあり詳細にまとめられていた

*Arista 7800R3 Series
Modular Data Center
Switches*



ラックマウントについて

- ラック設置する際は推奨のツールがある
ベンダー推奨のツールを探してみたが、国内では取り扱っている業者が無い様子（スペックダウンモデルであれば取扱い有り）
- 自力でのラック搭載は不可と判断
メーカー推奨の方法が取れないことから自社でのラックマウント作業は困難であると判断、専門の業者に依頼することに

<https://www.arista.com/assets/data/pdf/Datasheets/7800R3-Quick-Look.pdf>

スイッチ納品時の様子



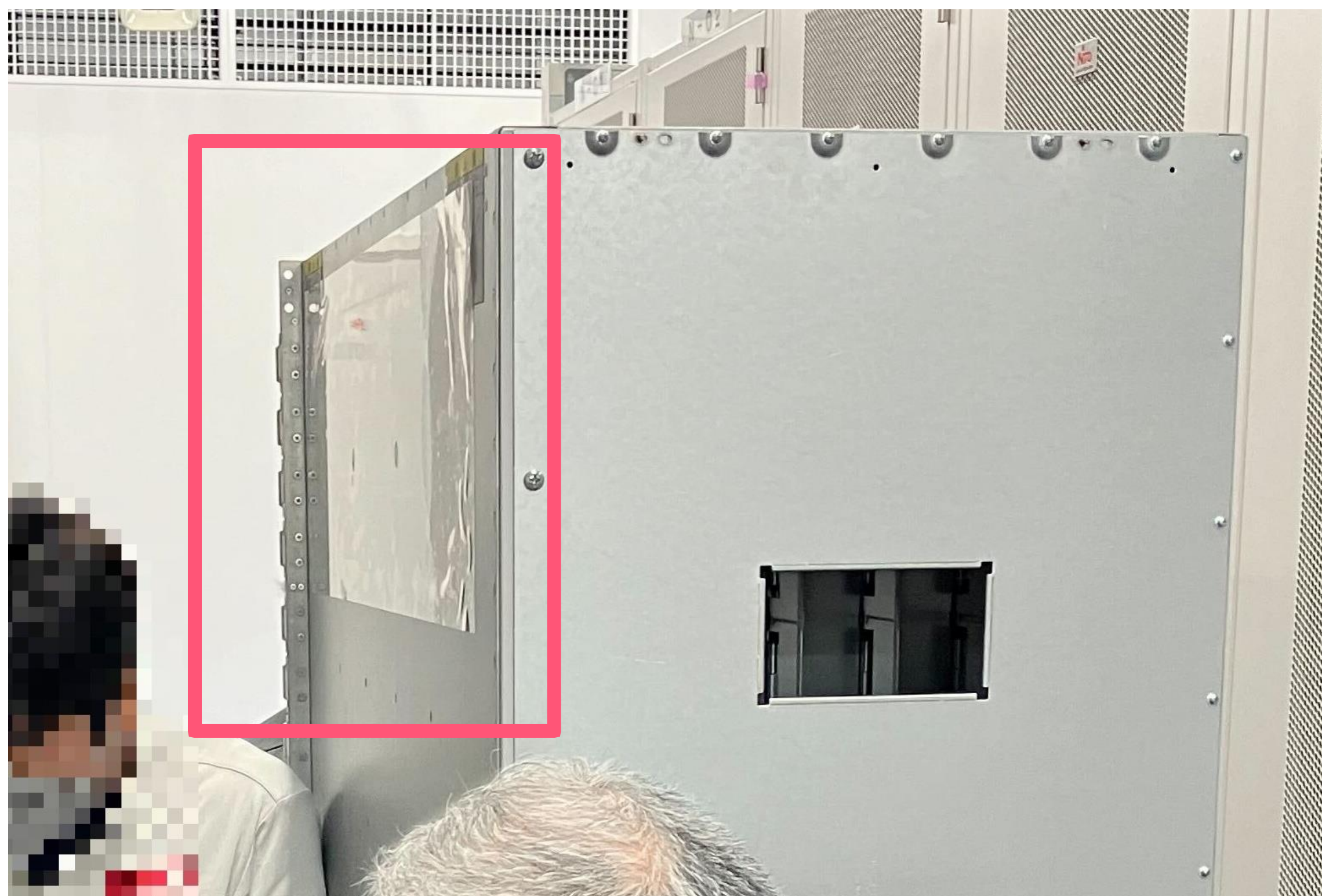
スイッチをラックに設置する際の様子

32Uで200kgぐらいあるため、ラックに納めるのも一苦勞



スイッチをラックに設置する際の注意点

一度ラックに納めるとそう簡単には取り外せない



ラックマウント作業での失敗談

- ラック側面保護シートの剥がし忘れ

ラックマウント後に気づき、シートを剥がそう思ったが、シートの粘着力が思いのほか強く、引っ張っても剥がせない…

- ケージナットの落脱

ラックにネジ止めする際に力を入れすぎてしまい、ネジでケージナットを押し込んで落としてしまった
ケージナットを取り付けることができないので諦めることに
(他のネジ穴でラックにしっかりと固定されていることは確認済)

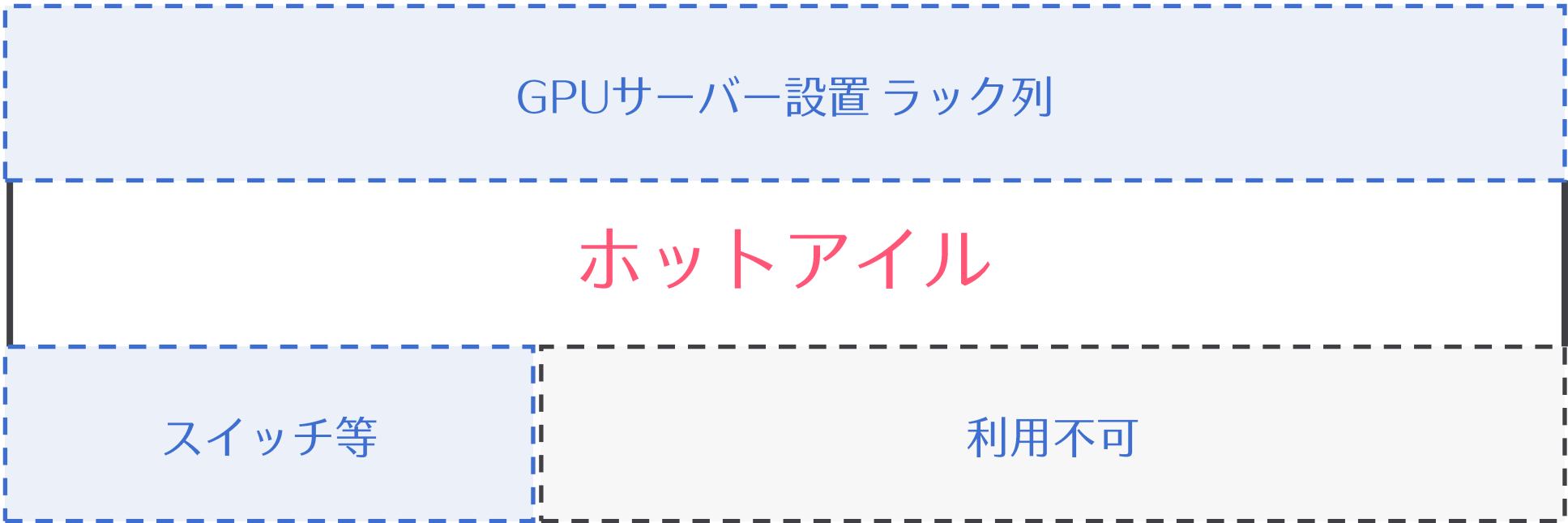
シャーシスイッチとサーバー どうつなぐか問題

各サーバーラックにMMF MPO-12パッチパネルを敷設しスイッチまで配線

コールドアイル



コールドアイル



コールドアイル

設置場所による課題

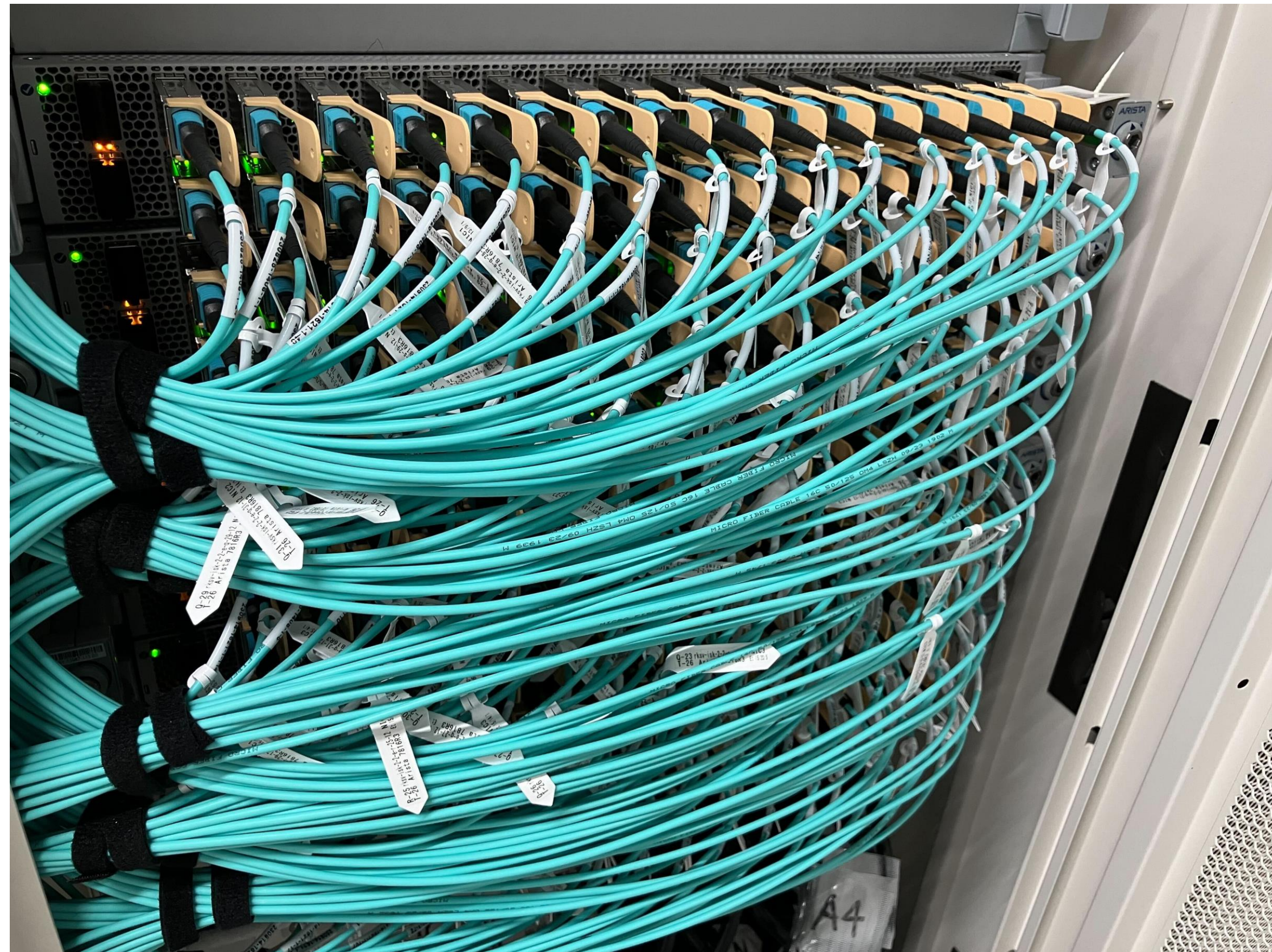
- ラックの電力仕様
ラックの電力仕様より、2ラックで3台のサーバーを設置する構成
- アイルやラック列をまたぐ配置
スイッチを設置するラックからサーバーを設置するラックの距離は配線距離にし40m～50mほど、AOCやDACは利用不可

パッチパネルとシャーシスイッチが1ラックに収まりきらない

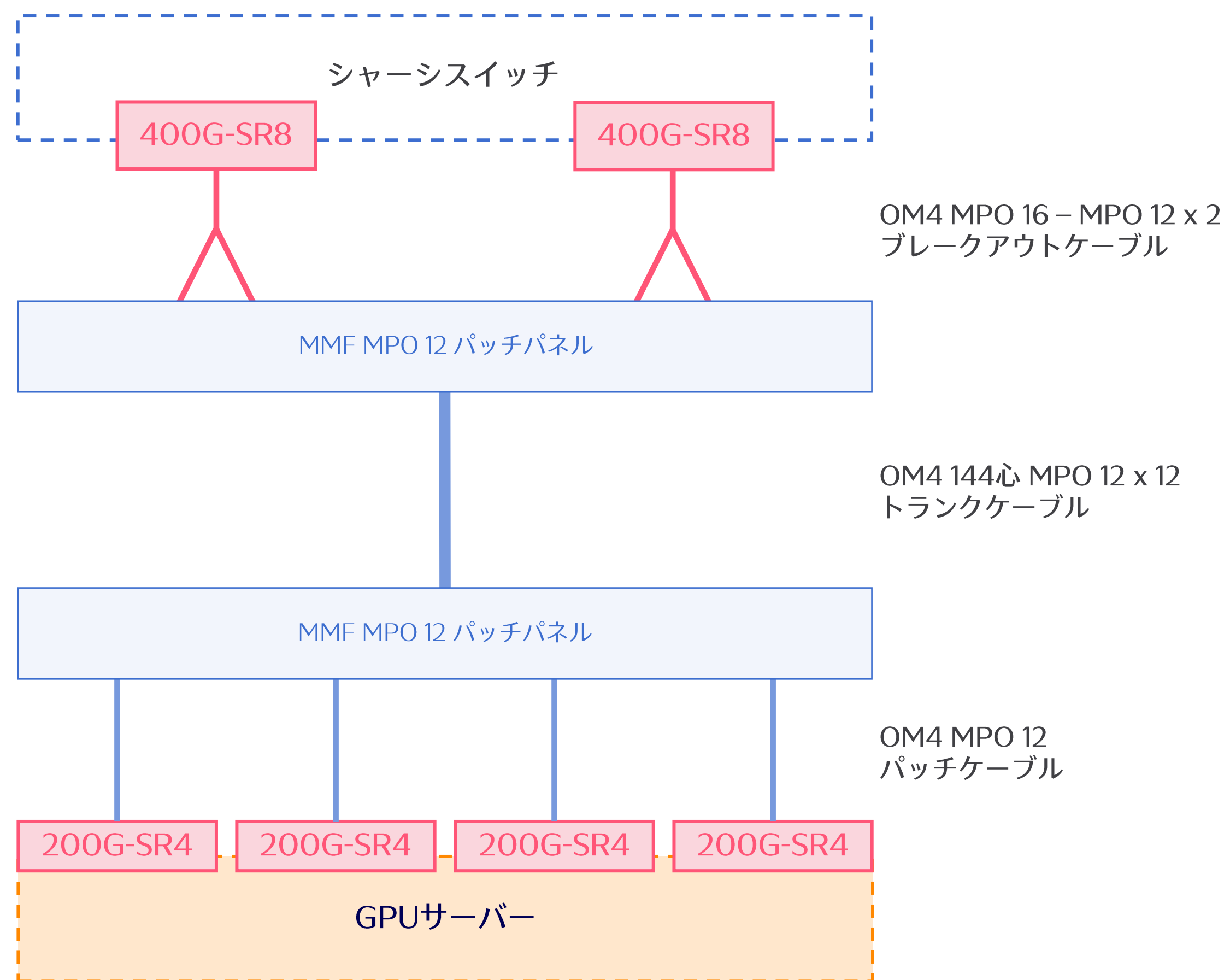


隣のラックから横に流して配線する方式に

※設置した環境が600mmラックなので狭かった



初回リリース時にはGPUサーバー1台あたり 200Gbs x 4 のネットワークを提供



配線のポイント

● MMF & ブレークアウトケーブルを利用

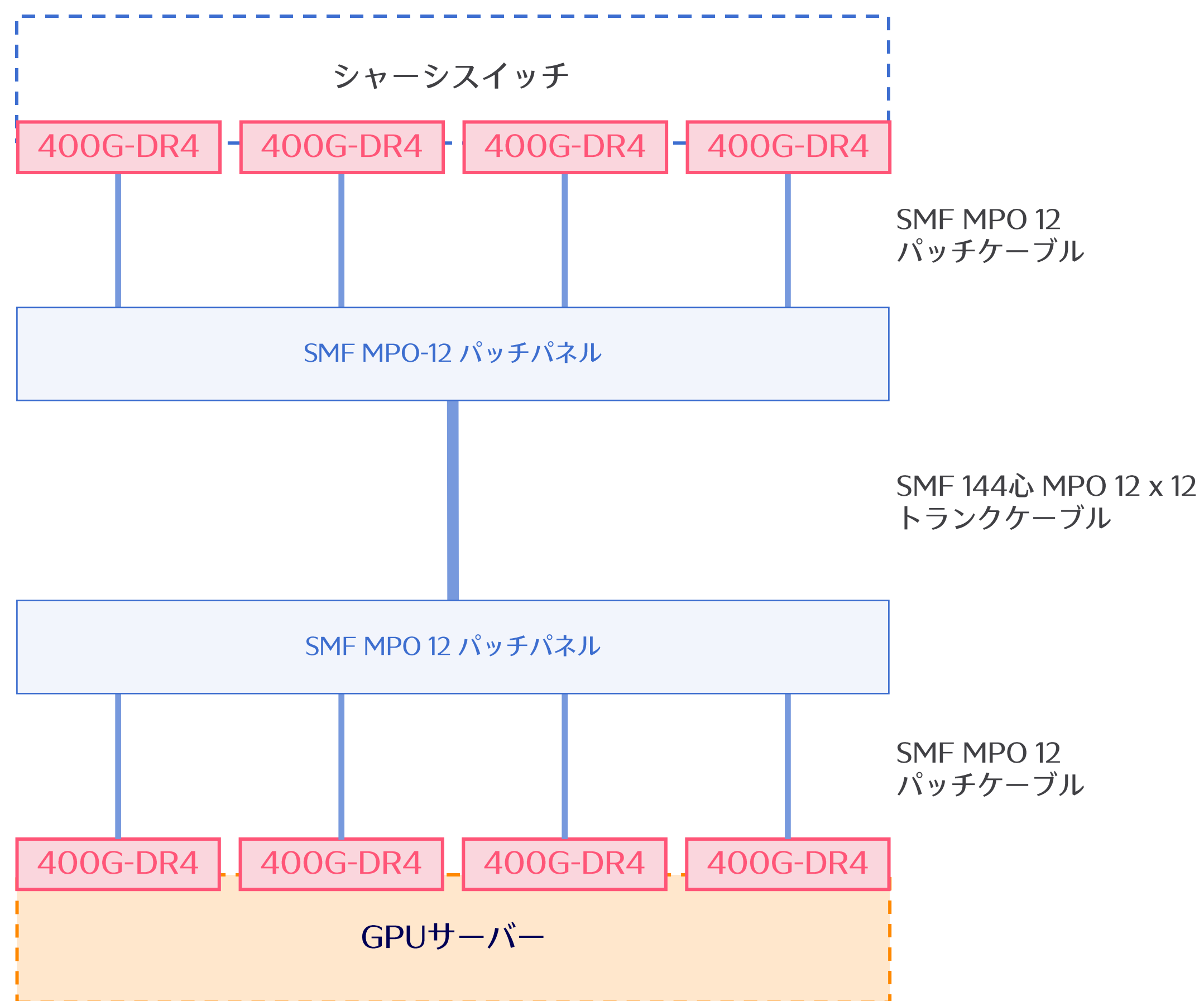
納期 & コスト & トランシーバー種別など複数の要素を勘案した結果、400G-SR8 をブレークアウトして 200G-SR4 x 2 として利用する構成を採用



● やや扱いにくかった

トランシーバー不良などの障害時の切り分け
 100本超のブレークアウトケーブルを高密度に使うことになる
 MPO 16の取り扱いが難しい
 ケーブルクリーナーがMPO 12と違うので持ち替え大変
 MMF MPOの品質に関する問題（後述）

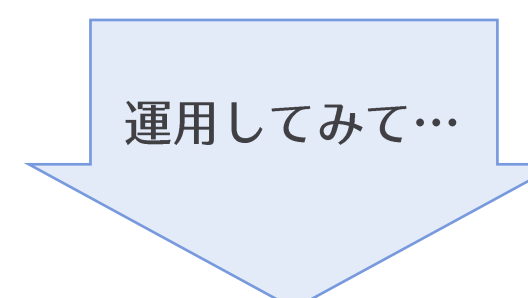
追加整備分では1サーバーあたり 400Gbs x 4 のネットワークに変更



配線のポイント

● SMF MPO-12配線に変更

- ・ 初回構築時より半年ほど経過した結果、サードパーティ品のOSFP-RHSがいろいろ選べるようになった
- ・ 各種部材の納期が大幅に改善した



● いまのところ問題なし

配線もシンプルになり、400G-DR4も安定している

両隣のラックをパッチパネル専用ラックとして利用して配線する設計に



サーバー側では基本的にサードパーティ製のトランシーバーを利用している

200G-SR4

QSFP56 (MMF)

何社かの製品を比較した結果、BERの測定結果に**差がある**ことがわかった

→ BERの結果がより良いベンダーの製品を採用

2つの種類があり、PC研磨MPOケーブル対応品か、APC研磨MPOケーブル対応品か

→ 高火力 PHY ではPC研磨対応品を採用した（NVIDIAの純正トランシーバーはPC研磨のもの）

当然ながら研磨方式が異なるケーブルを接続すると通信不良となる

ラベルに記載が無ければ判別は難しいので要注意

400G-DR4

OSFP-RHS (SMF)

何社かの製品を比較した結果、BERの測定結果にそこまでの差がみられなかった

→ 納期や価格などで採用ベンダーを決定

```
$ sudo mlxlink -d mlx5_0 -m -c -e
Module Info
-----
Identifier : OSFP
Compliance : IB NDR,400GBASE-DR4
Cable Technology : 1310 nm EML
(途中略)
Vendor Name : SAKURA
Part Number : OSFP-400G-DR4
```


MMFにするかSMFにするか

- MMFのほうがトランシーバーも含めて比較的安価だがハマりところも多い
- 基本的にはSMFの規格を採用していくことになる（と思われる）

光配線をどうするか

- MPO-16はパッチパネルの敷設が難しい、扱いにくさはあるものの800G配線しようとする当面は必要になってきそう
- もしMMF/MPOで配線するのであればLow-Lossタイプの採用を検討したほうがよい（パッチパネルも同様）

トランシーバーどうするか

- 400Gであれば今のところ400G-DR4が無難と思われるが、LC2芯の400G-FR4はありかも？
- LPO対応トランシーバーもちらほら見えてきたし、今後はさらに悩むことになりそう。あと800Gはどうする？
- フォームファクタはもうしばらく悩みそう（ConnectX-8は OSFP224-RHS?）

トラブルシューティング

怪奇！ 1分だけ動作するトランシーバ

- 構築完了後にリンクアップしない区間があったため、抜去、清掃、再接続の手順を実施
- リンクアップした。よし事務所にもどろう！
- リンクダウンしてる… なんで…
- もう一回見に行く、抜き差しする、リンク上がる、しばらく見つめる(1分ほど)で落ちる！なんでー

```
2024-05-27T13:57:01 Fru: %FRU-6-TRANSCEIVER_INSERTED: A transceiver for interface Ethernet5/4 has been inserted. manufacturer:
2024-05-27T13:57:27 Ebra: %LINEPROTO-5-UPDOWN: Line protocol on Interface Ethernet5/4/1 (NIC1), changed state to up
2024-05-27T13:57:28 Lldp: %LLDP-5-NEIGHBOR_NEW: LLDP neighbor with chassisId and portId "enp25s0np0" added on interface Ethernet5/4/1
2024-05-27T13:57:50 Ebra: %LINEPROTO-5-UPDOWN: Line protocol on Interface Ethernet5/4/1 (NIC1), changed state to down

2024-05-27T15:09:54 Fru: %FRU-6-TRANSCEIVER_INSERTED: A transceiver for interface Ethernet5/4 has been inserted. manufacturer:
2024-05-27T15:10:22 Ebra: %LINEPROTO-5-UPDOWN: Line protocol on Interface Ethernet5/4/1 (NIC1), changed state to up
2024-05-27T15:10:23 Lldp: %LLDP-5-NEIGHBOR_NEW: LLDP neighbor with chassisId and portId "enp25s0np0" added on interface Ethernet5/4/1
2024-05-27T15:11:30 Ebra: %LINEPROTO-5-UPDOWN: Line protocol on Interface Ethernet5/4/1 (NIC1), changed state to down
```


怪奇！ 1分だけ動作するトランシーバ

- このような中途半端な壊れ方するトランシーバには初めて遭遇しました…
- このような一見良さそうに見えて実はダメみたいなパターンつらい…
- 中途半端に頑張らず、壊れるときは潔く完全に壊れてほしい…



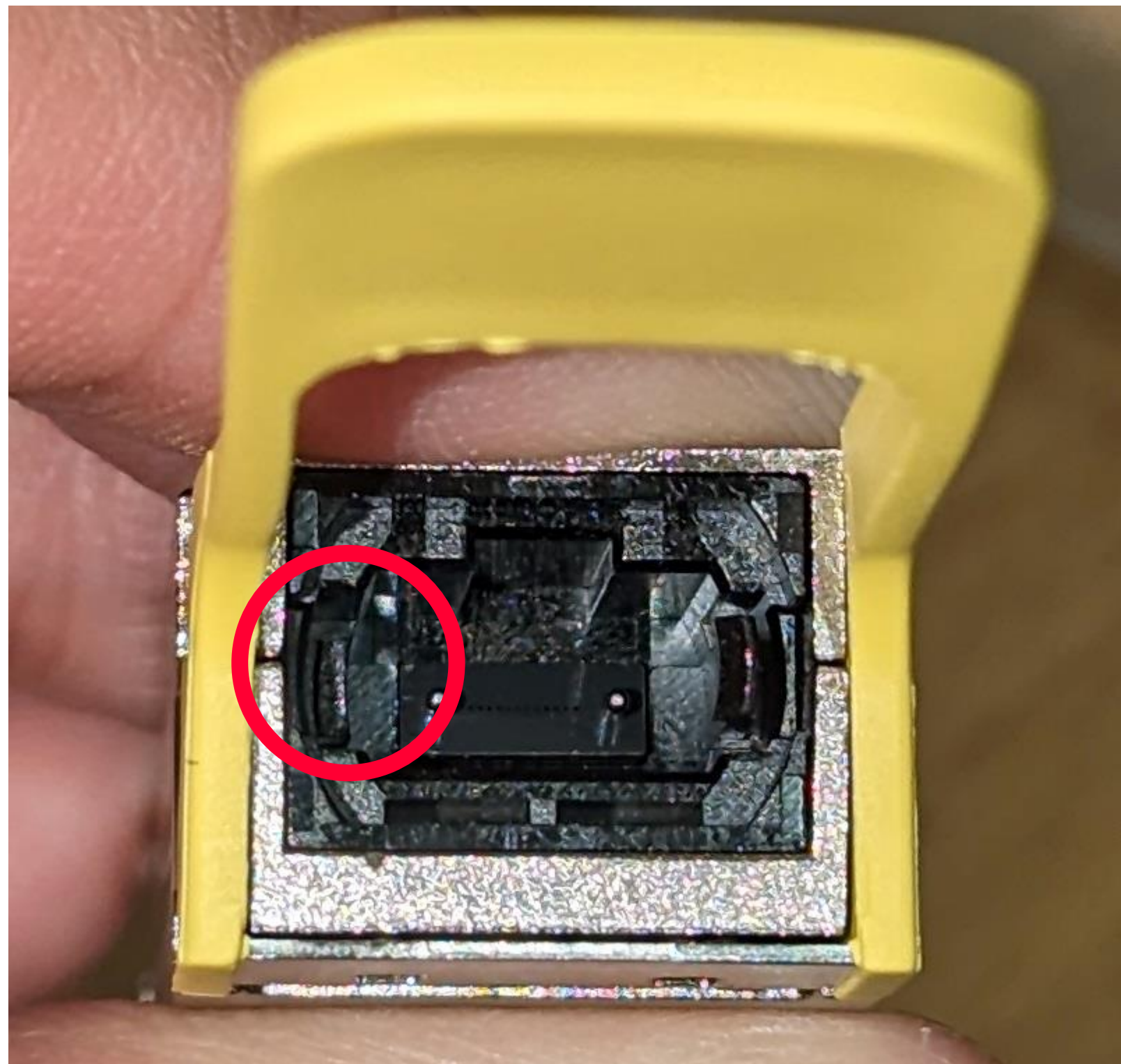
まさかの！ 壊れちゃったトランシーバ

- 構築作業の際には左記のようにトランシーバを挿入後清掃を実行
- シャーシサイズによっては比較的単調な清掃作業が連続500回以上発生
- 配線完了後にミスが見つかり再配線、その結果清掃回数は想定の倍に



まさかの！ 壊れちゃったトランシーバ

- そんな作業の中、クリーナ挿入時にクリーナとMPOの爪が接触しそのまま破損



まさかの！ 壊れちゃったトランシーバ



- 破損を確認後、構築したシャーシのトランシーバをすべて確認
- 基本的な対策としては垂直を意識して挿入、力加減を2段階に分ける？
- 再発防止を考えようにも手作業であるがゆえ、完全には難しい
- 物理的にこうならないような仕組みが実現できるとみんな幸せ？
- みなさんも作業中にネットワーク機器等を壊してしまった経験ありますか？

MPOケーブルの品質問題



- 一部エリアではMPO-MMFと200G-SR4を用いて構築
- 構築後、一部区間で通信品質が悪いことが判明
場合によってはリンクフラップも発生する状況
- 端面検査機でケーブルを検査した結果、NG判定

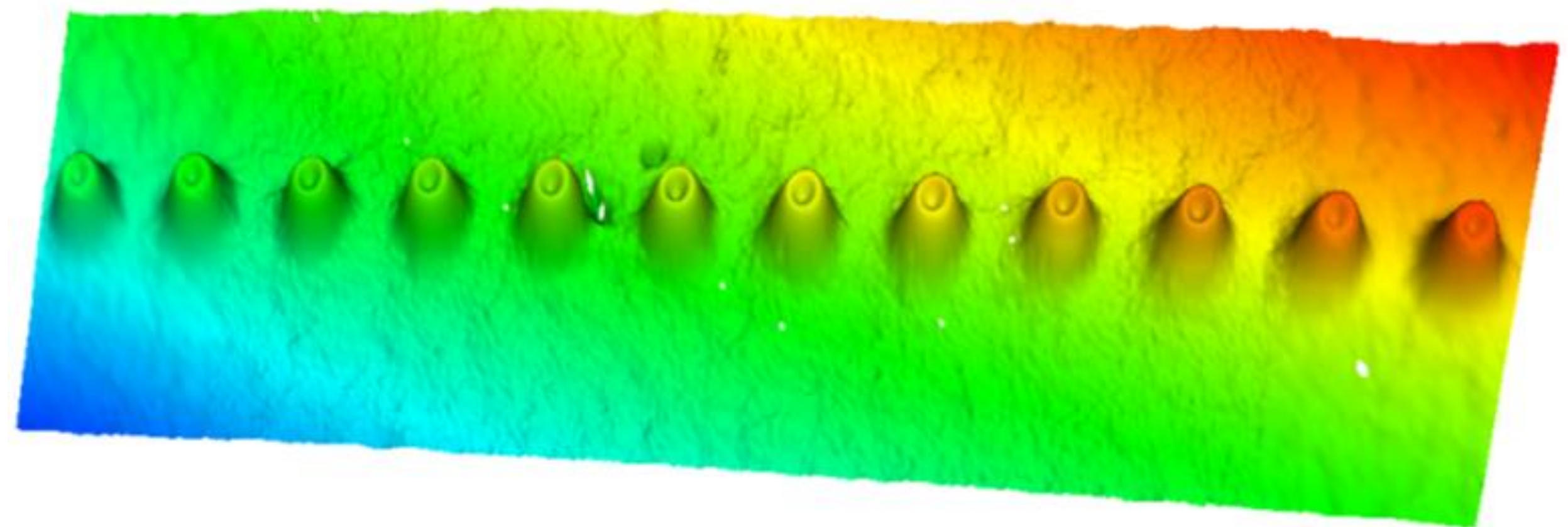
MPOケーブルの品質問題



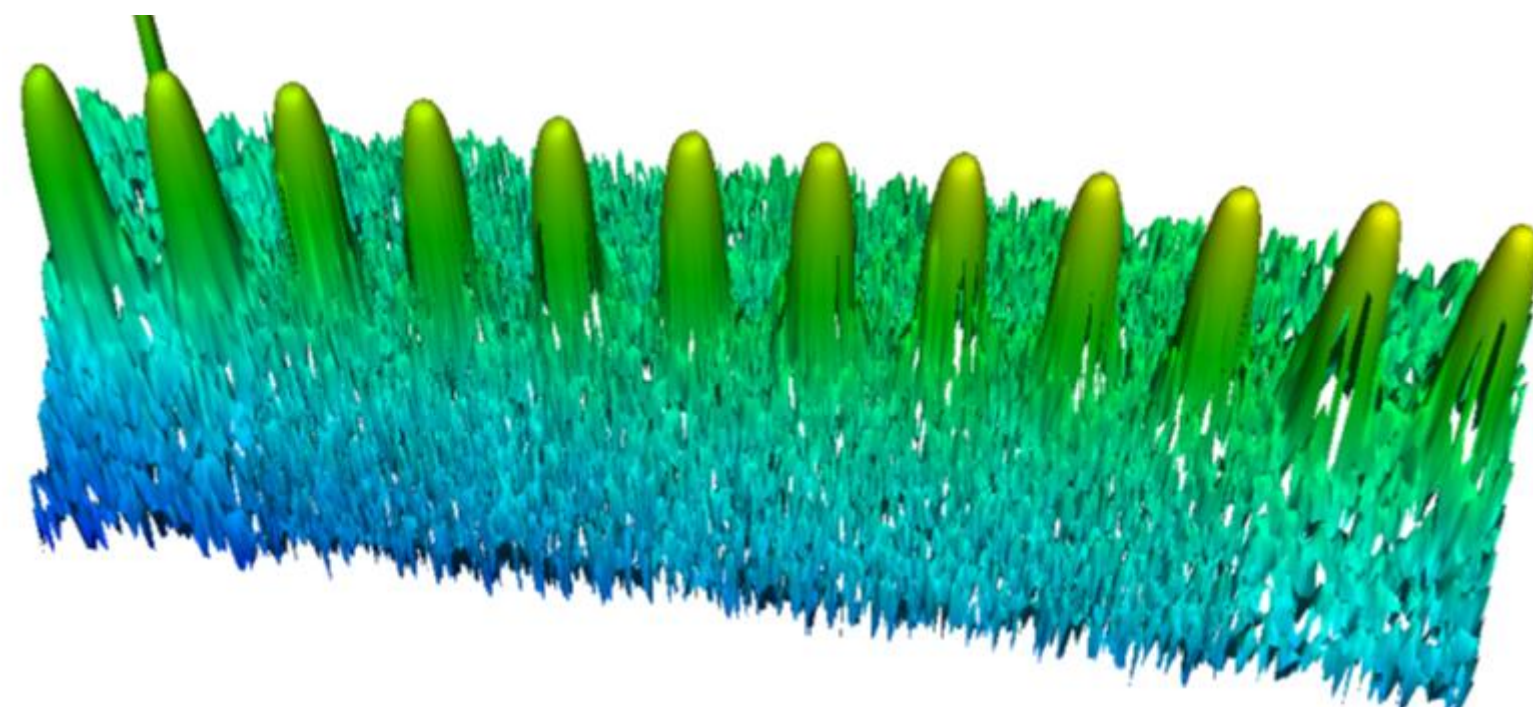
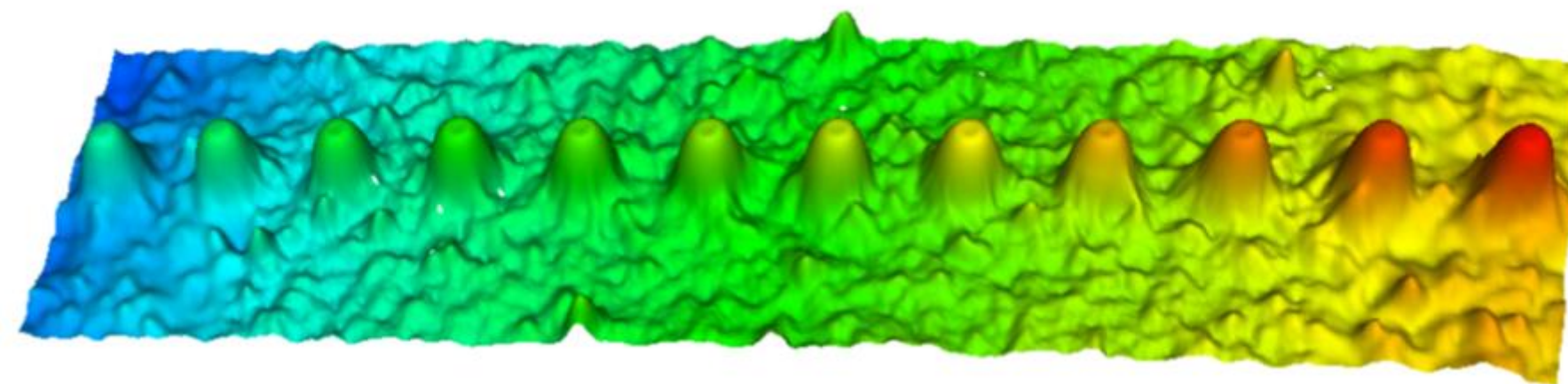
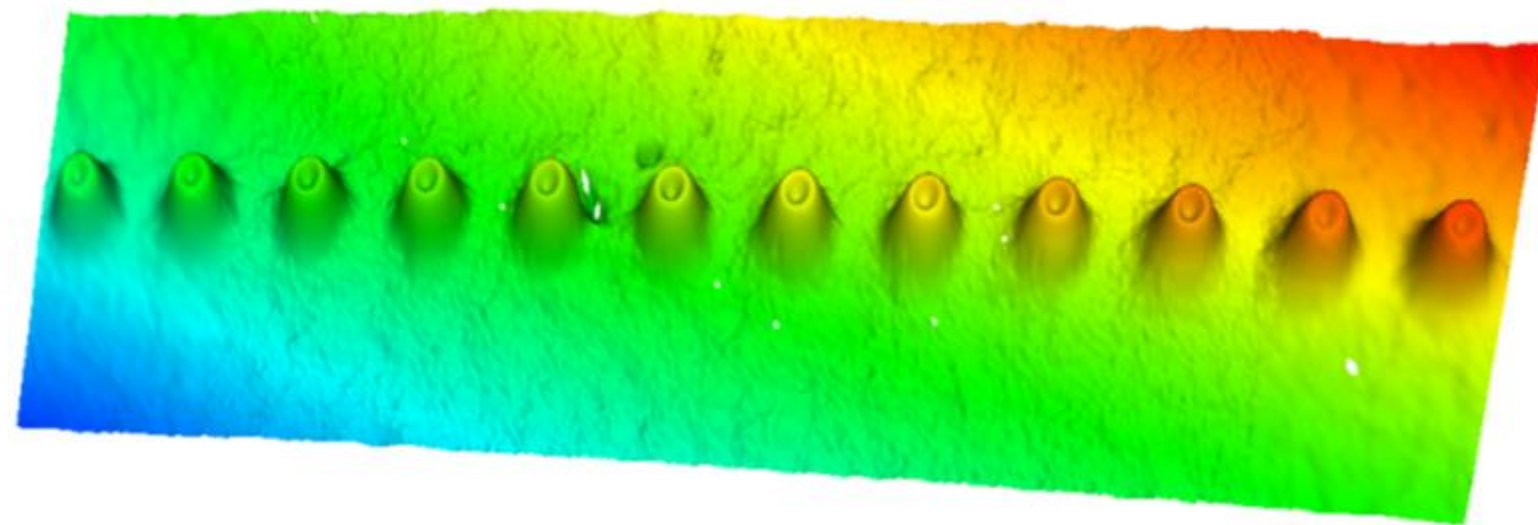
MPOケーブルの品質問題



- 品質低下が発生していた区間はコアディップが目立った
- コアディップとはMPOコネクタの製造工程で研磨する際にコア部分が凹んでしまい出来る空洞で、これが原因で反射率が変わり信号品質に影響を与える
- 左図でもコアの中心が凹んでいる



MPOケーブルの品質問題



- NTT AT社様のご協力のもと、弊社で購入した3社のファイバを調査いただいた
- 弊社が最初に購入した上段のものはコアディップが大きく、これが通信品質低下の原因とみられる
- 中段は上段に比べ、ばらつきも小さく良品に見える
- 下段は最上位品質、コアディップ無し！
- このクラスではケーブル品質も要検証！

データセンター関連

- ・ ラックの設計や配置で工夫したことがありますか？（特に高負荷環境）
- ・ 電源設備やPDUまわりで苦労した経験談などありますか？
- ・ 大規模構築のときはどのような体制で実施していますか？

ネットワーク運用関連

- ・ 大量の配線をどのように設計していますか？ 何か工夫していることがありますか？
- ・ サードパーティ製トランシーバー使っていますか？選定 / 検証ってどうしていますか？
- ・ ケーブルの選定などで気を付けていることはありますか？
- ・ 水冷ラックのネットワーク装置どうしますか？

その他にかあれば

さらなるGPU基盤の確保と提供

GPU基盤拡張イメージ

2 EFlops+

18.9 EFlops

2027年度

生成AI向けインフラ基盤の高度化

- 1. コンテナサービスの実装
- 2. インターコネクトの広帯域化
- 3. 次世代GPU「B200」の調達
- 4. 生成AI基盤に特化したコンテナ型データセンター
- 5. 水冷システム

コンテナサービスの実装

高火力 ^{ドック} DOK

高性能なNVIDIA GPUを低価格で提供

生成AIや機械学習に最適なGPUクラウドサービスです

リリースキャンペーン実施中

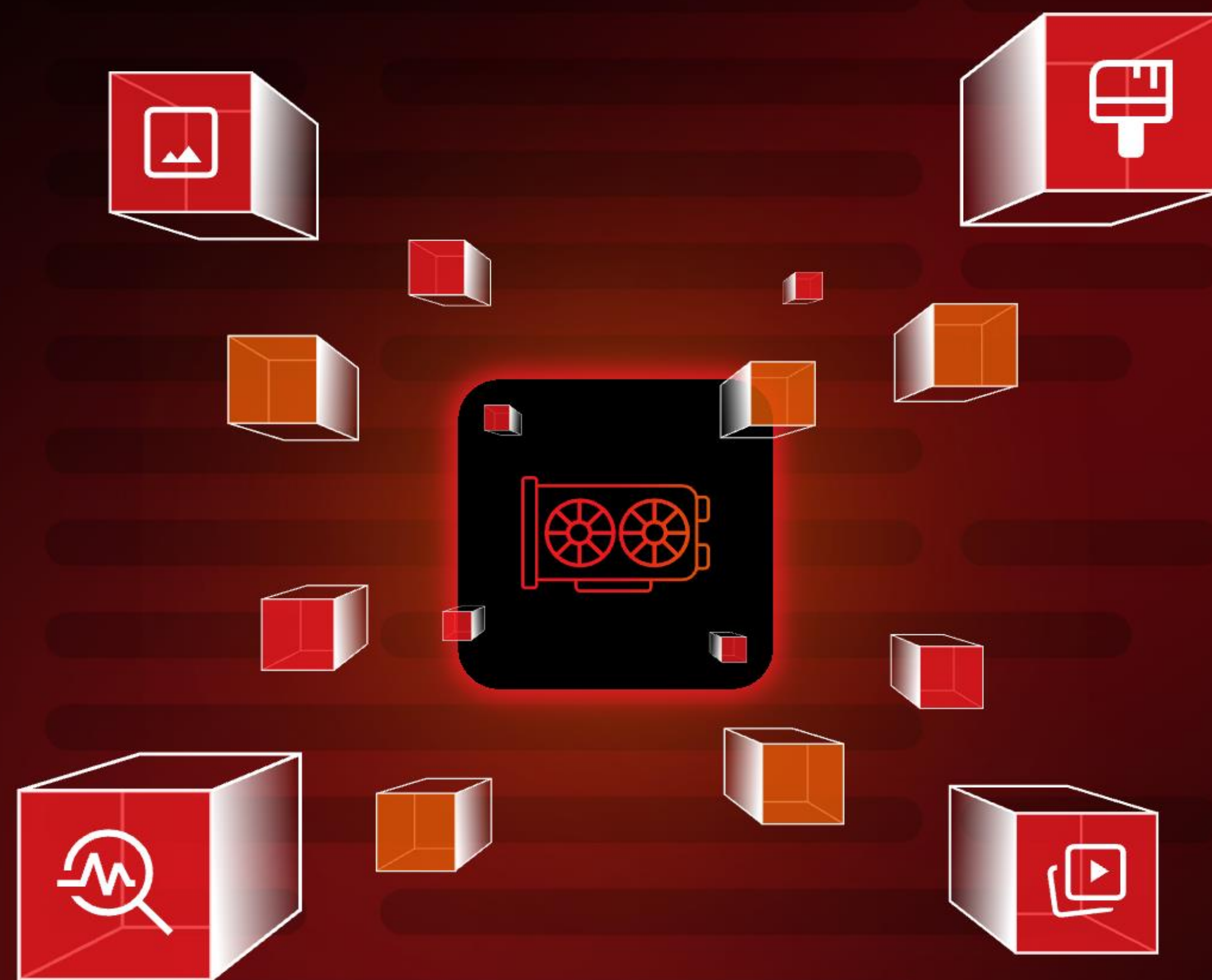
最長3ヶ月間、サービス利用料が無料

さらに、さくらのクラウドクーポンも付いてくる！

申込受付期間：2024年9月20日まで



お申し込み方法
はこちら



(※) H100は近日提供予定

