

JANOG52

# NWの拡張 ～広帯域化で解決を目指して100Gから400Gへ切り 替えの挑戦～

日本電気 井垣恵太

NECネットワークスアイ 大橋憲昭

# 自己紹介

氏名：井垣一恵太 中原 一彦

- 主な業務
  - NW,NFVI
- 経歴
  - 2018～NEC ネットエスアイ入社
  - 2023～NEC ネットエスアイ入社
- 趣味
  - NBA、バスケットボール（スラムダンク）観戦、スカレー始めました

氏名：大橋 憲昭

- 経歴
  - 2009～NEC ネットエスアイ入社
  - 海外キャリア・ベンダーのお仕事をいくつか
  - ここ数年はNECCIメイン
- 趣味
  - 旅行



# はじめに

## ◆NECCIとは

- NECが提供するクラウドサービスのプラットフォーム
- 関東Rは3年間の運用でLeaf SW 80台規模に拡大・昨年1月に関西Rもオープン

詳しくは昨日のJANOG発表  
「IP-Closにサービスもストレージも通してみた」で・・・

↑ NEC Cloud IaaS 3つの特長

- 安全・安心なクラウド**
  - 安全・安心なデータセンター/国内2リージョン提供によるBC/DR実現
  - 安全・安心を担保する公的認証を取得
- 多様なニーズに対応するラインナップ**
  - フロント業務から基幹システムまでカバーするサーバラインナップ
  - 様々なシステムとの連携を実現するネットワークサービス
  - ホワイトボックス化されたクラウド基盤
- クラウド運用を効率化**
  - 日本の企業・法人向けに進化したセルフサービスポータル
  - ICT予算の管理負担を軽減するシンプルな料金体系
  - ハイブリッドクラウド環境の運用を効率化する運用支援機能

© NEC Corporation 2023

<https://www.janog.gr.jp/meeting/janog52/wp-content/uploads/2023/06/janog52-stor-oishi-pre.pdf>

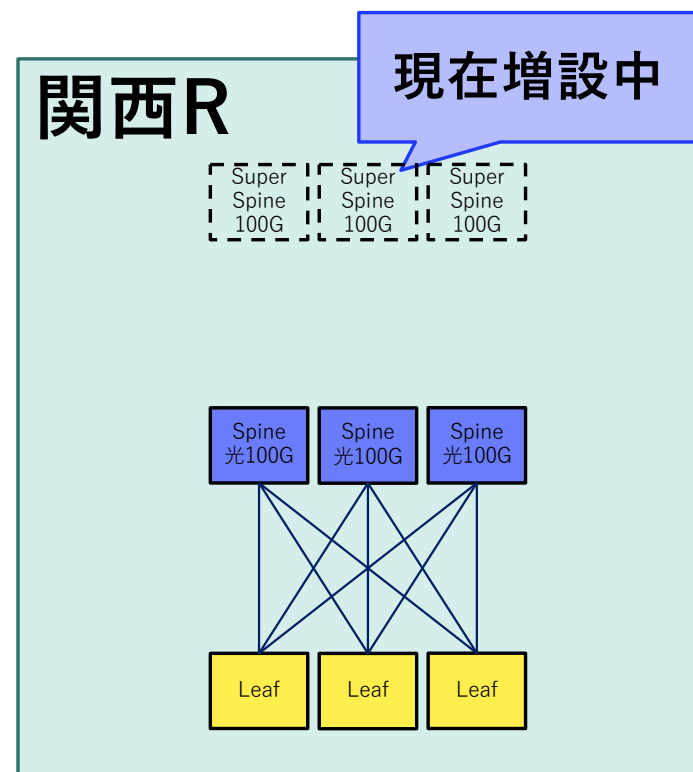
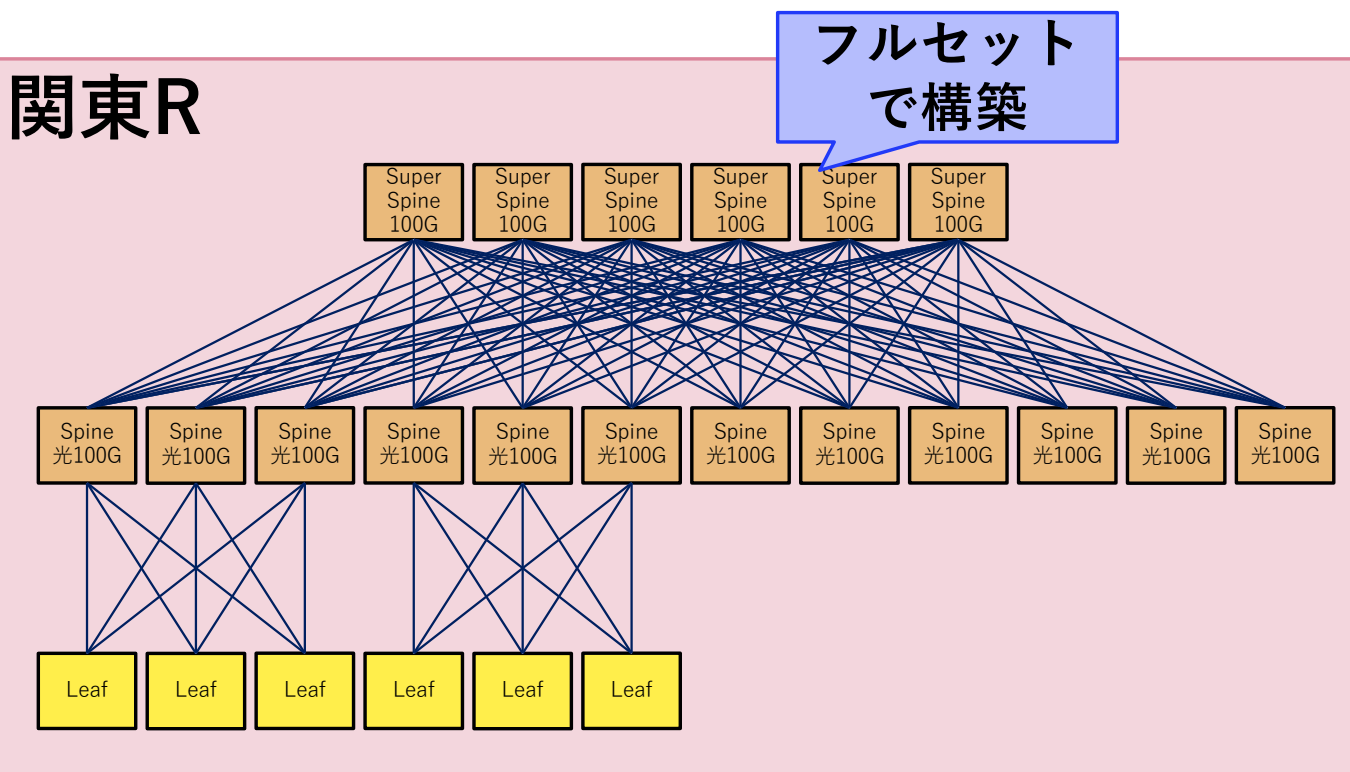
## ◆拡張方法を検討中

- サービスの拡大・追加に伴いバックボーンNWの拡張が必要に
- 最適な拡張方法とは・・・ **★今回議論させてください**

# はじめに

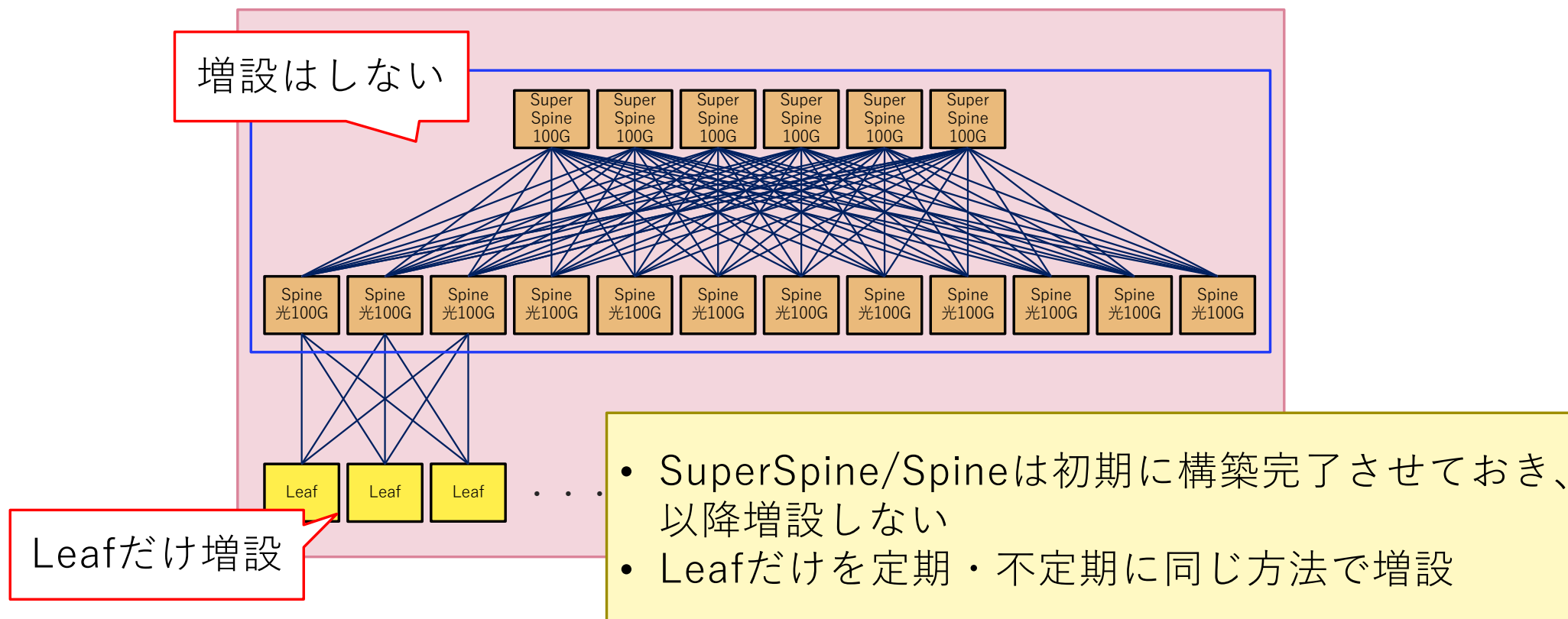
## ◆ 関東Rと関西Rの構成差分

- ベンダーA 100G×64port
- ベンダーB 100G×64port



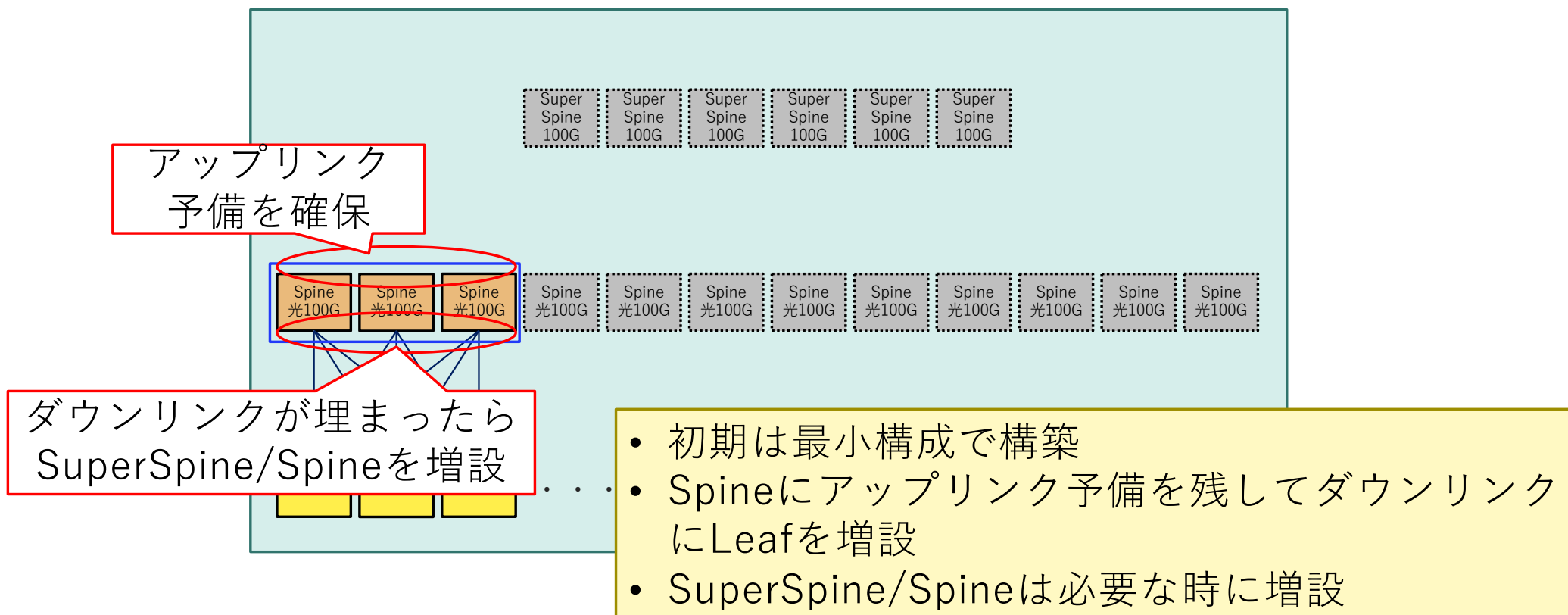
## A. 最初から大規模派

### ◆ 3tear構成を最初に作りきり、Leafだけを増設



## B. スモールスタート派

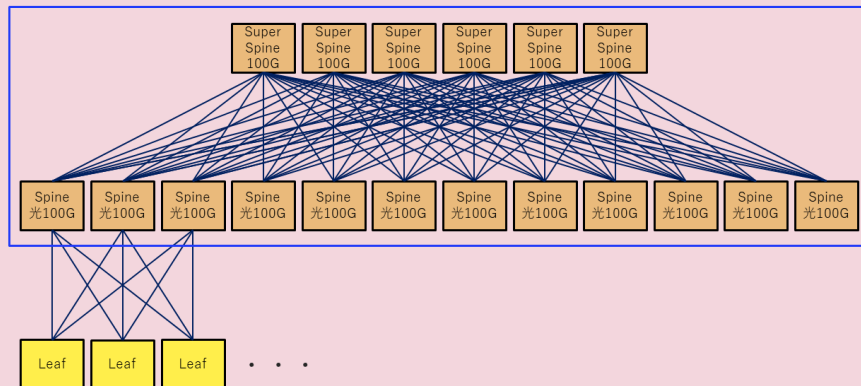
### ◆ 最初はSpine-Leafだけ、後から3tearに変更



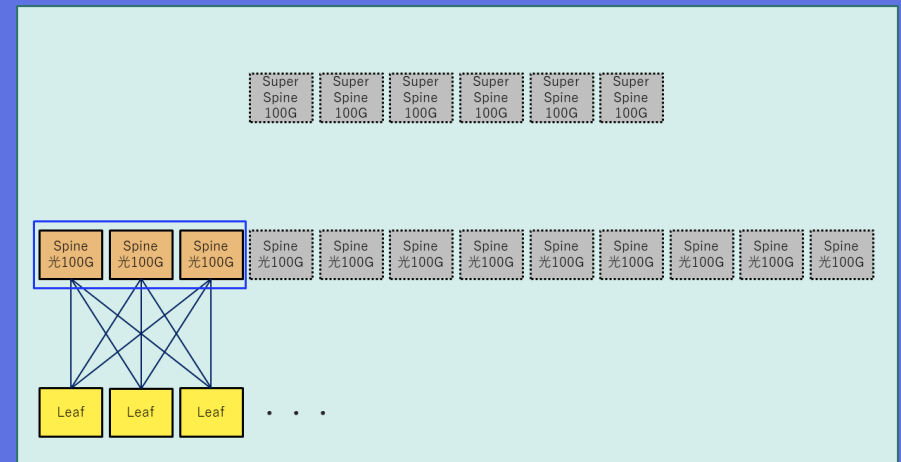
# 全JANOGERに問います

# 導入時の構成はどこまで作るべきか

## A. 最初から大規模派



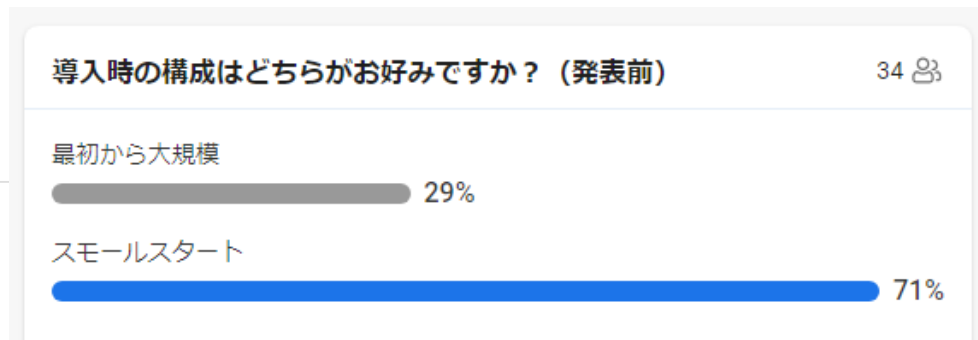
## B. スモールスタート派





# 事前アンケート結果

## ◆結果



A.最初から大規模派

29 %

B.スモールスタート派

71 %

アンケートご協力ありがとうございました。

◆ 立場や観点の違いにより推しアーキテクチャが異なる

- 設計・運用の立場
- 投資側の立場

設計・運用者の立場代表  
A. 最初から大規模派

NESIC 大橋

VS

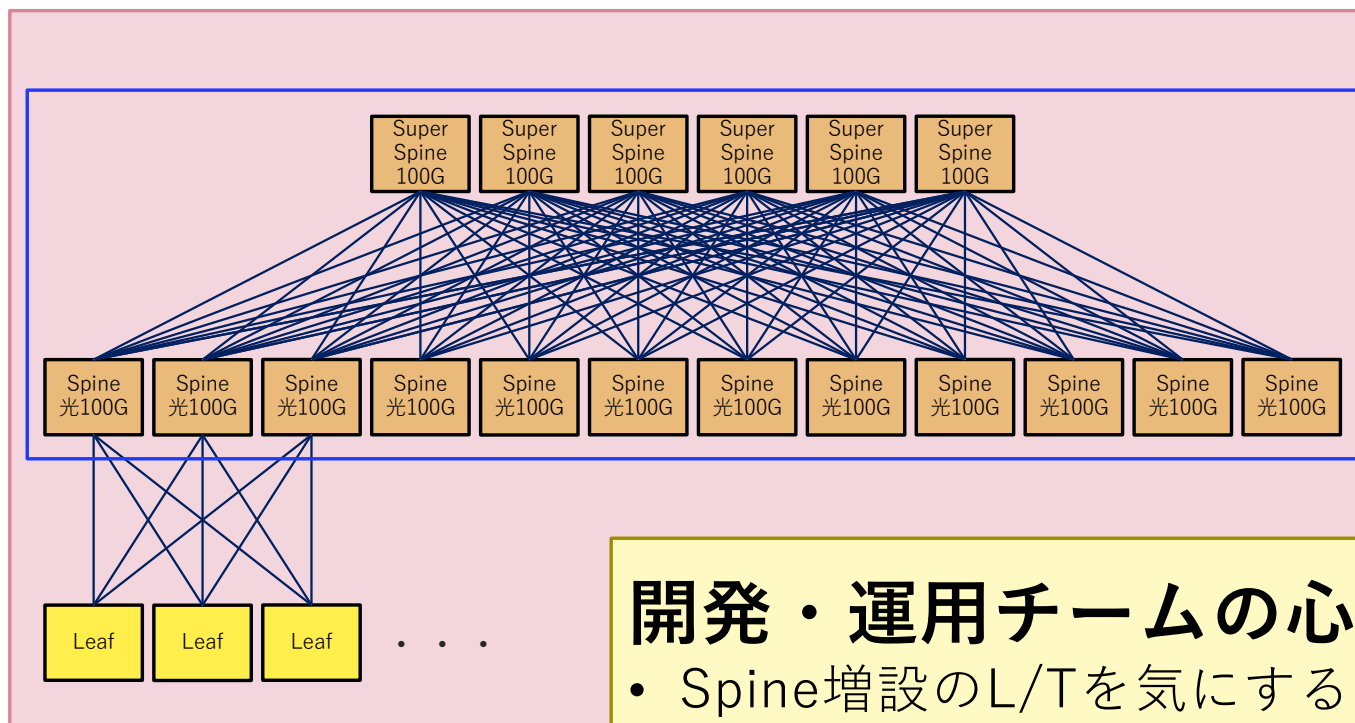
投資側の立場代表  
B. スモールスタート派

NEC 中原

# 各増設方式それぞれの主張

## A. 最初から大規模派

### ◆ 最初から大規模派の意見

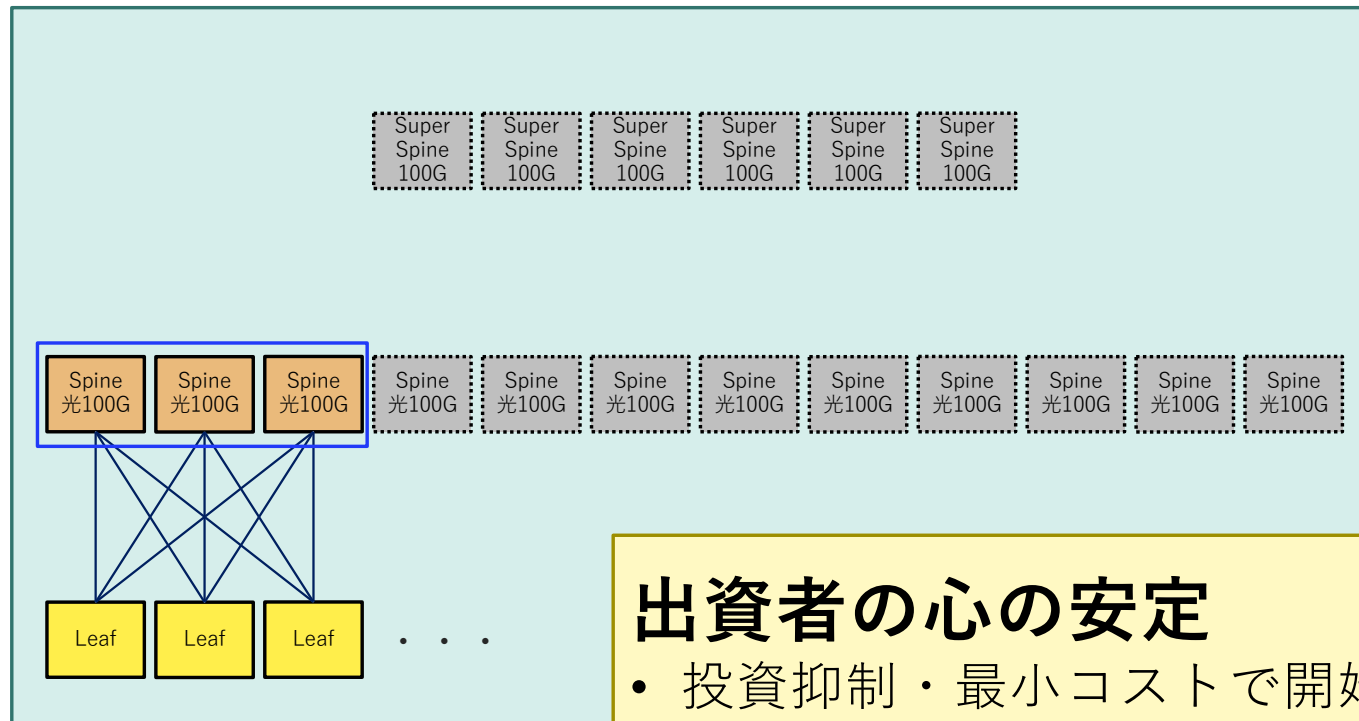


### 開発・運用チームの心の安定

- Spine増設のL/Tを気にする必要がない
- 設計・運用が簡素

## B. スモールスタート派

### ◆ スモールスタート派の意見



### 出資者の心の安定

- 投資抑制・最小コストで開始
- 投資リスク（最悪たたむことも・・・）

最初から大規模は投資が莫大

## スモールスタート派

莫大な初期導入コスト  
(数千万の差)

- SuperSpine/Spine機器費
- SFP購入費
- 工事・構築費

ランニングコスト  
(数百万/年の差)

- 機器保守費
- ラック利用費
- 故障交換・メンテナンス

SDGsの観点  
(無駄な電力・CO2排出)

- トラフィックは流れず光ってるだけ
- 冷却効率が落ちエアコンが頑張る



# 事業継続にはコスト意識を強く持つことが大事

運用費だってバカにならない

最初から大規模派

### 機器費高騰

- スモールスタートでは利かないボリュームディスカウント
- 為替リスク

### 終売

- **ベンダー買収**
- チップ等の部品の終売に伴うEoS
- 最新機種リリース

### 設計・運用費

- 新機種導入の設計コスト
- 多機種運用によるコスト増
- LCM（アップグレード）作業の複雑化

スモールスタートはトータルコストで  
むしろ高くなる可能性も

# 機器が買えない！？

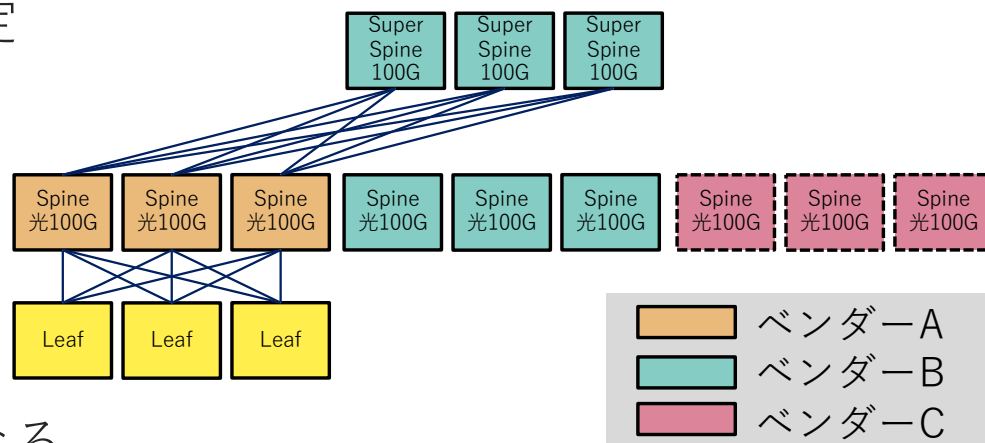
## 最初から大規模派

### ◆ ホワイトボックスの甘い罠

- Spineは必須機能が少ない（アンダーレイでBGPを話すだけ）
- 価格・L/T等を比較してベンダーを選定
- ホワイトボックスSWを選択

### ◆ しかしいざ増設時に、、、

- OSベンダーが買収される
- ホワイトボックスSWに乗せられなくなる。。。



## ベンダー・機種種の再選定からやり直しに。。。

そしてパッチワークに・・・



最初から大規模派

# 運用の簡素化はとても重要

作業の簡素化→自動化へ

### 最初から大規模の課題

- サービス開始前に終局のトラフィック量予測は困難

### スモールスタートのメリット

- 実際のトラフィック量をみながら適切に増設
- 売れ行きを考慮した構成の再検討

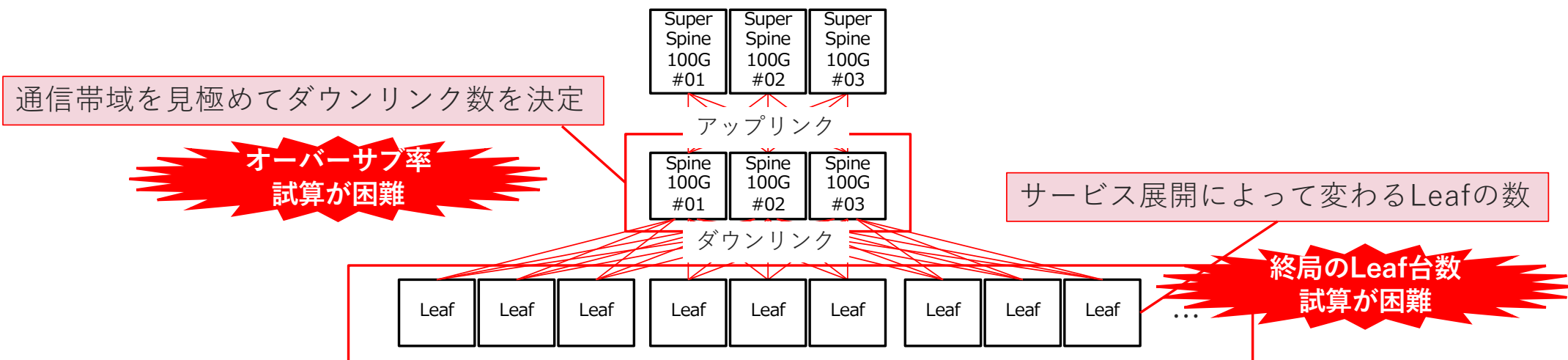
# 予測困難な終局

## スモールスタート派

### ◆ サービス開始前に終局の見極めが必要

- 終局のLeaf台数の見極め
  - サービス展開によって変わるLeaf台数
- 通信量の見極め
  - Spineに割り当てるダウンリンク数に変動

Spine/SuperSpineの  
台数に影響

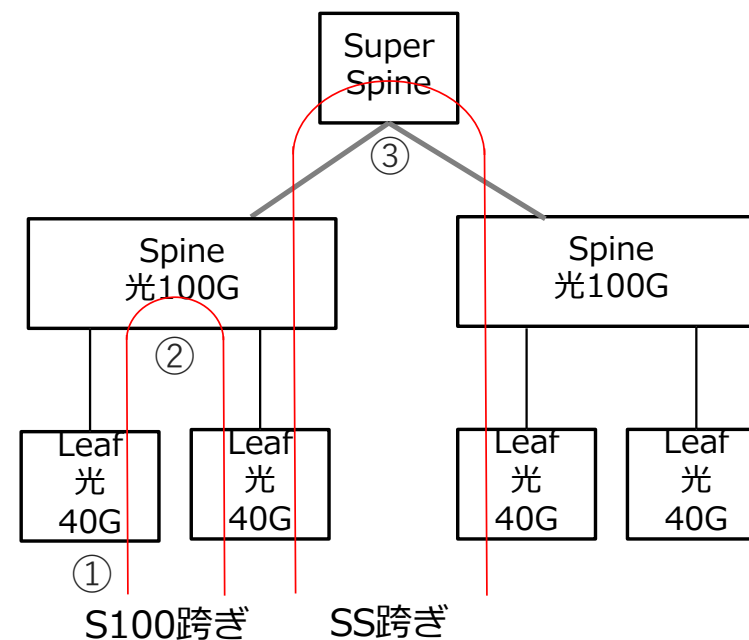


# 予測困難な終局

## スモールスタート派

### ◆ 予測困難なトラフィック量

- ① 収容機器からLeafへのトラフィック量
  - 多様な収容機器全てのトラフィックを予測
- ② Spine跨ぎの通信量
  - Spineを共有するLeafへのトラフィック量の予測
- ③ SuperSpine跨ぎの通信量
  - SpineSpineをまたがるトラフィック量の予測



試算は困難

# 全部仮置き

最悪使い切るか、使いすぎて足りなくなることも。。。

### スモールスタートの課題

- 増設時の構成検討が困難
  - ※欲が出てより良い構成を取りたくなる

### 最初から大規模のメリット

- 作りきっているので増設方法を検討する必要がない
- 運用を考慮した最適な構成を取れる

## 増設時の構成検討が困難

最初から大規模派

### ◆多様な機種からの選定

- 100G、400G多ポートの機種
  - ・右表のような組み合わせの製品がある

ポートスピード	ポート数
400G	64port
400G	32port
100G	64port
100G	48port
100G	32port

### ◆400G規格の考慮

- 400Gで直結する場合は400G-SR8を、100G\*4へBreakoutする場合は400G-DR4を利用する。

伝送タイプ	QSFP-DD製品の説明	リーチ	光コネクタ	波長	光変調	プロトコル
マルチモードファイバ	400G QSFP-DD SR8	OM4で最大100m OM3で最大70m	MTP/MPO-16	850nm	50G PAM4	IEEE P802.3c IEEE 802.3cd
シングルモードファイバ	400G QSFP-DD DR4	パラレルSMFで最大500m	MTP/MPO-12	1310nm	100G PAM4	IEEE 802.3bs

# 増設時の構成検討が困難

## 最初から大規模派

機器	パターン1	パターン2	パターン3	パターン4	パターン5	パターン6
SuperSpine	400G	400G	100G	100G	400G	100G
Spine	400G	400G	400G	400G	100G	100G
Leaf	100G	100G	100G	100G	100G	100G
トポロジ						
メリット	SSpine/Spine台数減 配線本数減	SSpine台数減	Spine台数減	SSpine台数減	SSpine台数減	既存構成と同等
デメリット	Leaf機種制約	Spine收容効率低	Leaf機種制約 SSpine-Spine配線多	Spine收容効率低 SSpine-Spine配線多	Spine機種制約	(特になし)



### ◆ 既存Leafによる制約

- 100GBreakoutでLeaf接続：400G SFPは 400G-DR4  
→Leaf側に使う100G-DRを、NECCI2で導入している既存の機器は対応していない
- QSFP-DDでLeaf接続：QSFP28と互換性があり100Gでの接続も可能だが400Gにするメリットなし

## 既存Leafの制約で、Spineは100G多ポート一択

# 増設時の構成検討が困難

## 最初から大規模派

機器	パターン1	パターン2	パターン3	パターン4	パターン5	パターン6
SuperSpine	400G	400G	100G	100G	400G	100G
Spine	400G	400G	400G	400G	100G	100G
Leaf	100G	100G	100G	100G	100G	100G
トポロジ						
メリット	SSpine/Spine台数減 配線本数減	SSpine台数減	Spine台数減	SSpine台数減	SSpine台数減	既存構成と同等
デメリット	Leaf機種制約	Spine收容効率低	Leaf機種制約 SSpine-Spine配線多	Spine收容効率低 SSpine-Spine配線多	Spine機種制約	(特になし)

## 増設時の構成検討が困難

## 最初から大規模派

◆ 各パターンのメリデメ整理・構成図示すると次頁のようになる。

パターン	評価	寸評
パターン1	×	400G化で想定される構成。集約度が高まるが、Leafで100G-DR非対応のため実現不可。
パターン2	×	400G化するが、Leafは400Gでなく100Gで收容。集約度が低く、無駄が多い。特にオーバーサブ率を高めると顕著。
パターン3	×	パターン1に劣るため考慮外。
パターン4	×	パターン2に劣るため考慮外。
パターン5	○	SuperSpineの集約度を上げるケース。より多ポートの400G採用でSuperSpine台数が減らせる。
パターン6	○	関東R構成と同等、互換性が高い。 集約度を上げるため、100G*128ポートなど採用検討。

最初から大規模派

# Spineへの400G導入は断念

SuperSpineに400Gを入れることも出来たが結局は関東Rと同じ構成に・・・

- ◆最初から大規模に作るのもっと大変な設計が必要になることも

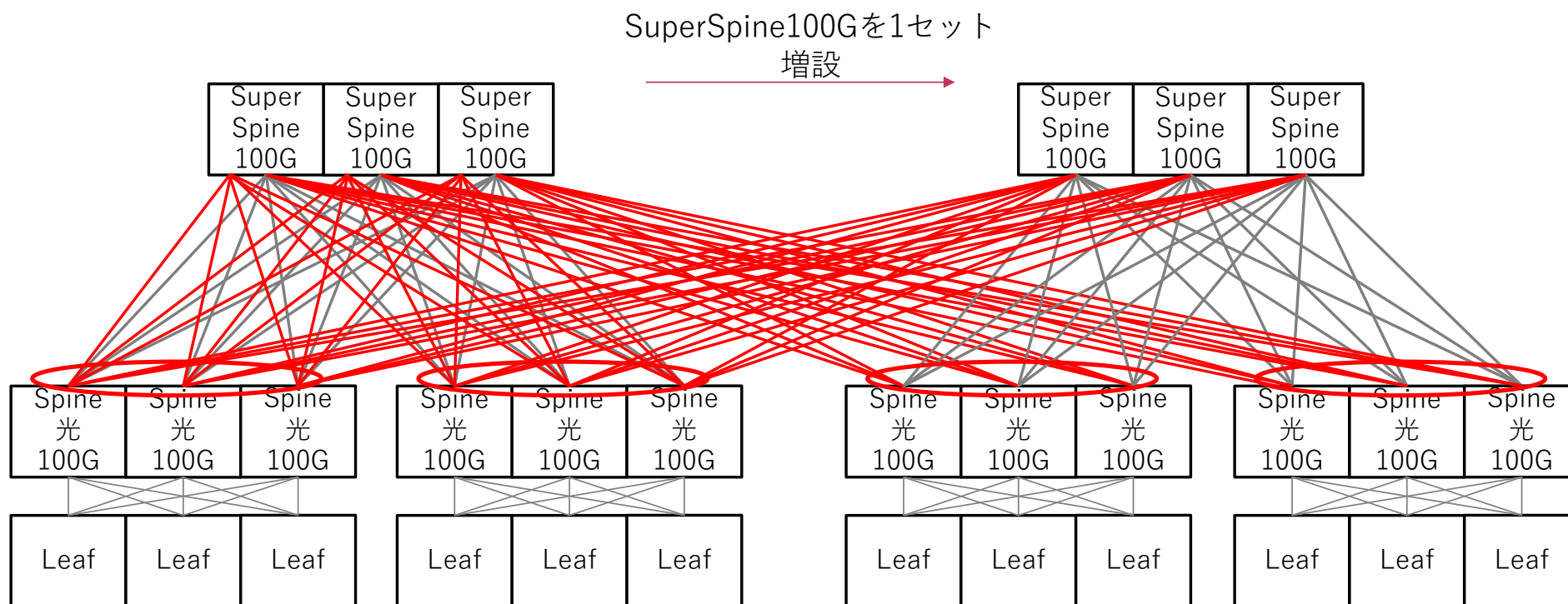
# 終局を見誤ったとき

終局を見誤ると・・・

## スモールスタート派

◆ 初期導入後にさらなる増設が必要になってしまったときの対応が困難

■ 新設SuperSpineに大量の活線挿し替え



# 初期に終局を見誤ると大変なことに・・・

## 工事の観点

- 新規配線
  - 工材コスト：倍必要
- 既存配線撤去
  - 活線との混在リスク



二度とやらないからな！

工事

## 作業の観点

- 差し替え作業
  - 差し換えミス リスク
- 設定変更作業
  - 設定/設計ミス リスク



超危ないよ

運用

DCのフロア/ラック数で大きさは決まる

最初から大規模派

◆フロア・ラックは有限

- 最初にリージョンの大きさを決定し、それに合わせた構成を初期構築が可能
- リージョン間接続を検討



◆NECCI1の実績を考慮した

- 販売計画・サービス開発計画によりある程度予測は可能

終局の見極めは出来ないことではない  
~~い(大変だが)~~



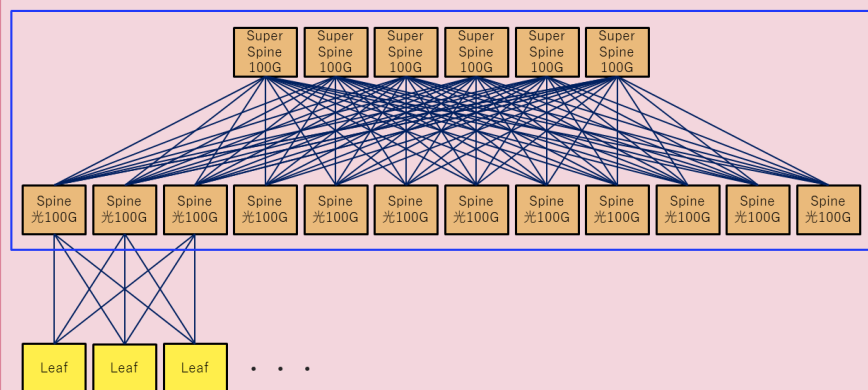
JUDGE

**ご清聴ありがとうございました。**

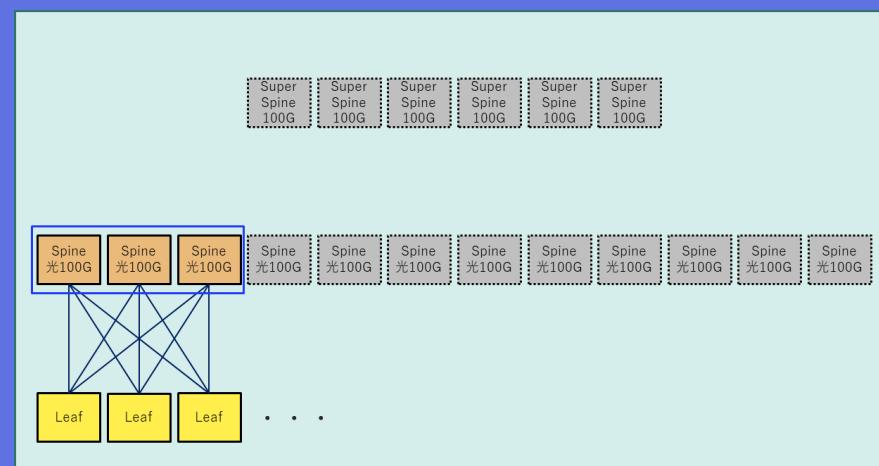
**今回の発表をご聴講いただいた  
皆様のご意見をお聞かせください**

# 最初から大規模派・スモールスタート派 あなたならどっち!?

## A.最初から大規模派



## B.スモールスタート派



Slackの全般チャンネルに送りました発表後slidoアンケートで投票ください

<https://app.sli.do/event/nkk8559DM1KoS6FCqMzrwR/embed/polls/86d78bbe-2f44-429d-85f8-3befc5235ffc>

# まとめ

---

## ◆費用の観点

- 初期投資、保守費、データセンター利用料
- 設計費、運用費

## ◆リスクの観点

- 機器が買えなくなる
- 急遽たたむことになった場合

## ◆設計の観点

- 増設構成
- 終局の見極め、帯域設計

# 議論ポイント

---

## ◆他の観点もあるのではないか

- 例. スイッチベンダーの観点など

## ◆機器のスイッチングコストをどうされているのか

- もう買えないとなったとき
- EoLがきたとき

# 事後アンケート結果

## ◆結果

A.最初から大規模派

XX %

B.スモールスタート派

XX %

\Orchestrating a brighter world

**NEC**