

ネットワークとムーアの法則に関わる10の予測

Shishio Tsuchiya

shtsuchi@arista.com

Agenda

- **このお話の背景**
- 過去のJANOGでの議論とこの話の関係
- 議論

過去のJANOGでもあったホワイトボックス話

- **JANOG34 2014年高松**
 - DCネットワークの新時代? –ホワイトボックススイッチの利用可否–
 - <https://www.janog.gr.jp/meeting/janog34/program/whbox.html>
- **JANOG35.5 2015東京**
 - パネルセッション: 検証してみて感じたホワイトボックススイッチの未来
 - <https://www.janog.gr.jp/meeting/janog35.5/program1>
- ホワイトボックスだけではなくて色んなベンダーがマーチャントシリコンを採用しだした。

そもそもなぜそんな流れになっていたのか

- アンディ・ベクトルシャイムがこんな事を言いだした

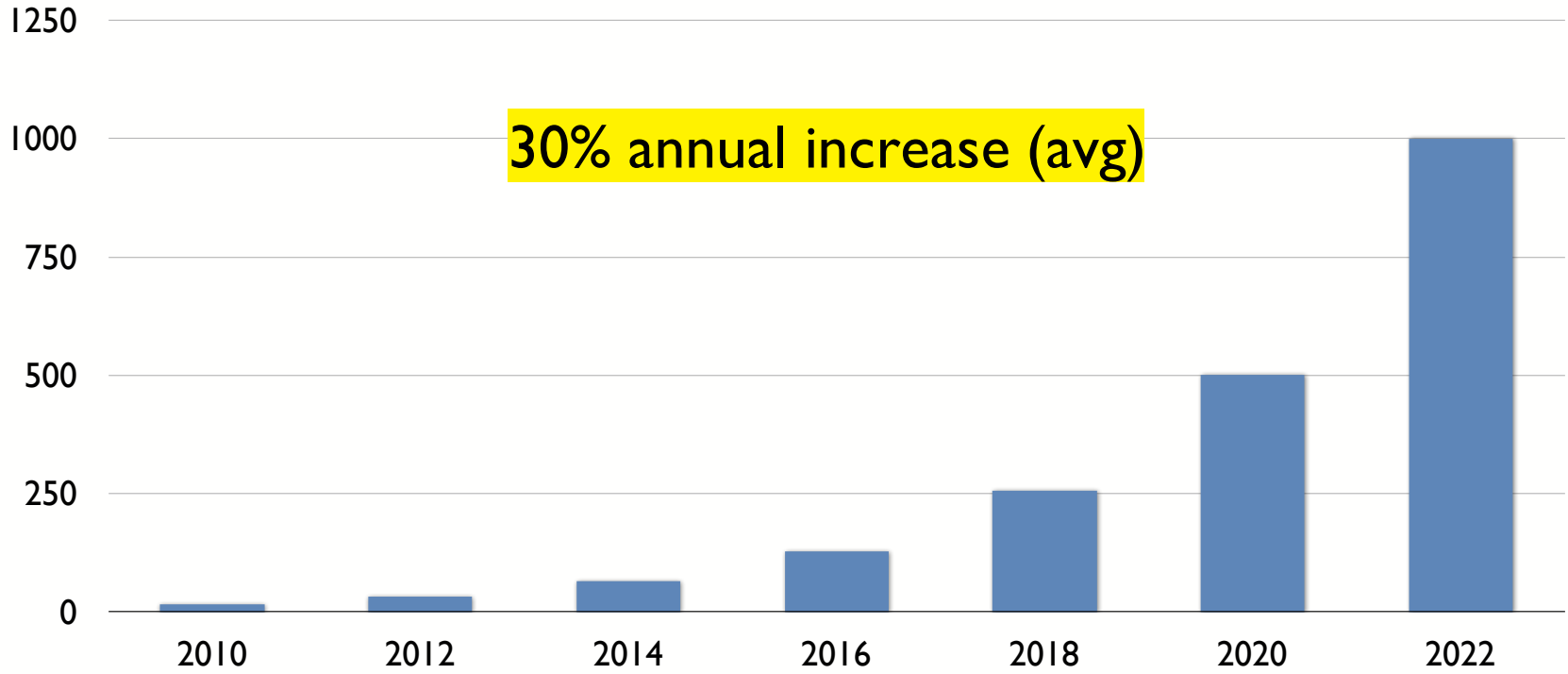
CPUとかメモリがムーアの
法則に乗っ取ってるにネット
ワークっておかしくね？



そもそも誰？

- サン・マイクロシステムズの共同創業者
- Granite Systems設立 ギガビットイーサネットを開発
- 1996年にシスコシステムズに買収
- 1998年にラリー・ページとセルゲイ・ブリンに10万ドルの小切手を渡し、グーグルの最初の出資者に
- 2001年 Opteron搭載サーバーのKealiaを設立
- 2004年サン・マイクロシステムズに買収
- 2008年アристаネットワークス

64ビットCPUコアの処理時間



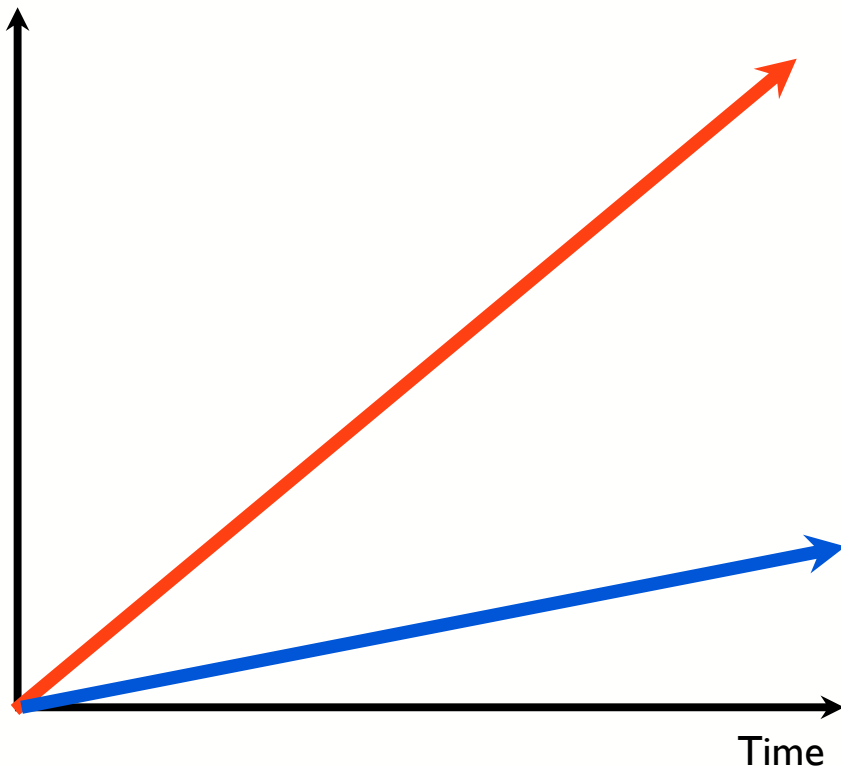
100X Performance Gain from 2010 to 2022

ムーアの法則(1971-2024)

- 集積密度は2年毎に2倍増加
- 1971年～2011年には100万倍の改善が見られた
- 2012年～2023年には別の100倍の改善が見込まれている
- 1971年～2024年には一億倍に拡大される

ムーアの法則とネットワーク

Performance



CPU: $2X/2Y = 64X/12Y$

CPUは2年毎に2倍
12年に64倍の処理能力の
向上が見られた

1GigE-10GigE: $10X/12Y$

1Gbpsから10Gbpsへの移行
12年かかった

3つの問題点

1. 市場における競争原理の欠如

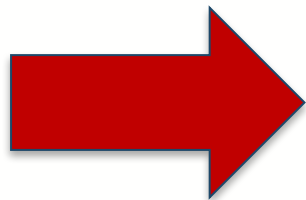
- 独占されている為、早く動く動機が無い

2. パフォーマンス競争力が無いASIC設計フロー

- ASIC設計フローはクロックレートと密度を制限する

3. ASICは古いプロセッサテクノロジーを使用

- 2-3世代前のCPU技術を使う

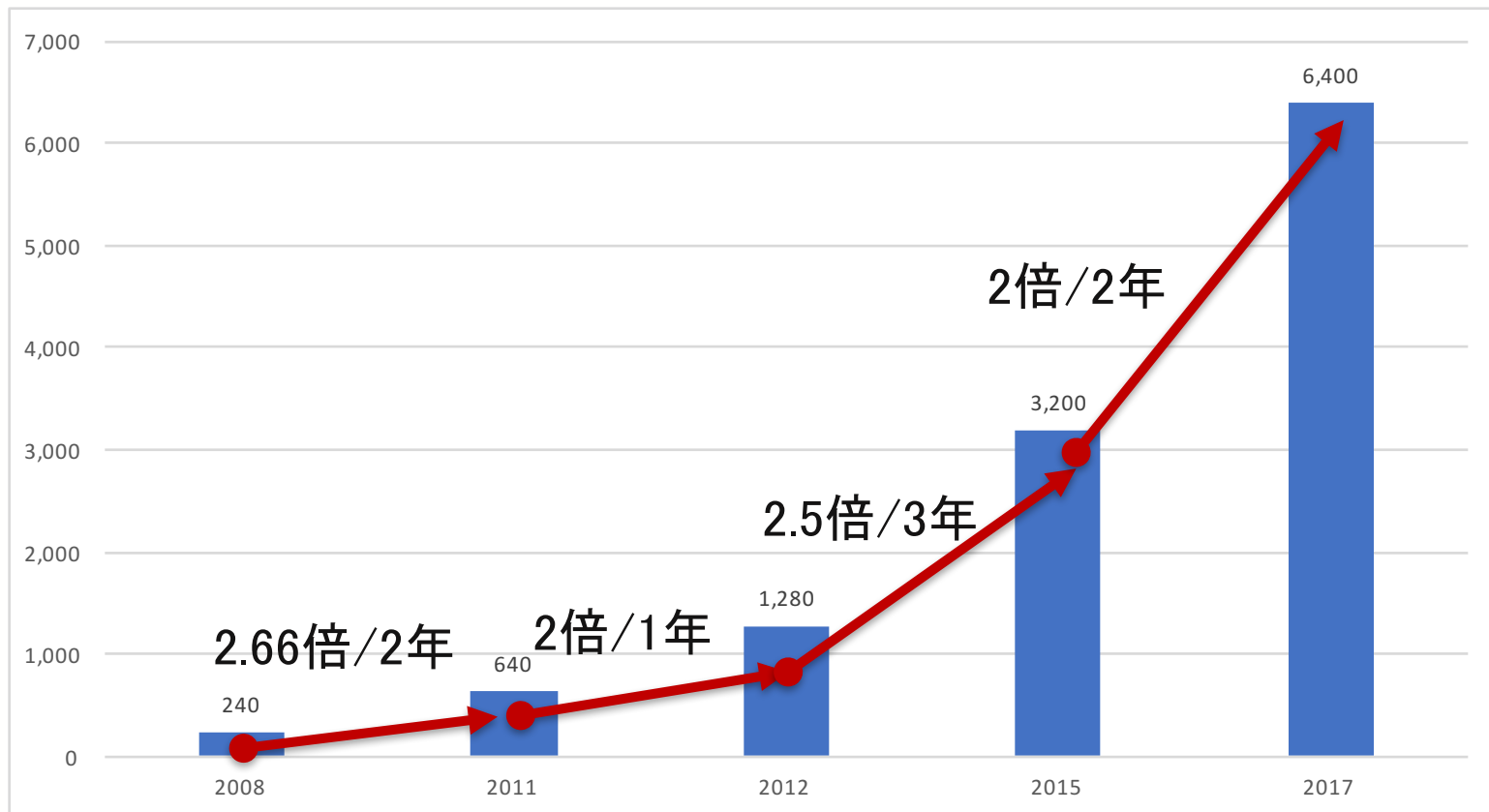


ネットワークの世界にCPU/メモリの様な競争原理を

マーチャントシリコンの登場

- 2008年:超低遅延24ポート10Gbps シングルチップ
- 2010年:大容量バッファVoQ付きの付きの10Gbpsチップ
- 2011年:64ポート10Gbpsシングルチップ
- 2012年:32ポート 40Gbps / 128ポート 10Gbps シングルチップ
- 2013年:大容量バッファ VoQ付きの40Gbpsチップ
- 2015年:32ポート 100Gbps / 128ポート25Gbps シングルチップ
- 2016年:大容量バッファ VoQ付き100Gbpsチップ
- 2017年: 64ポート 100Gbps / 256ポート25Gbpsシングルチップ

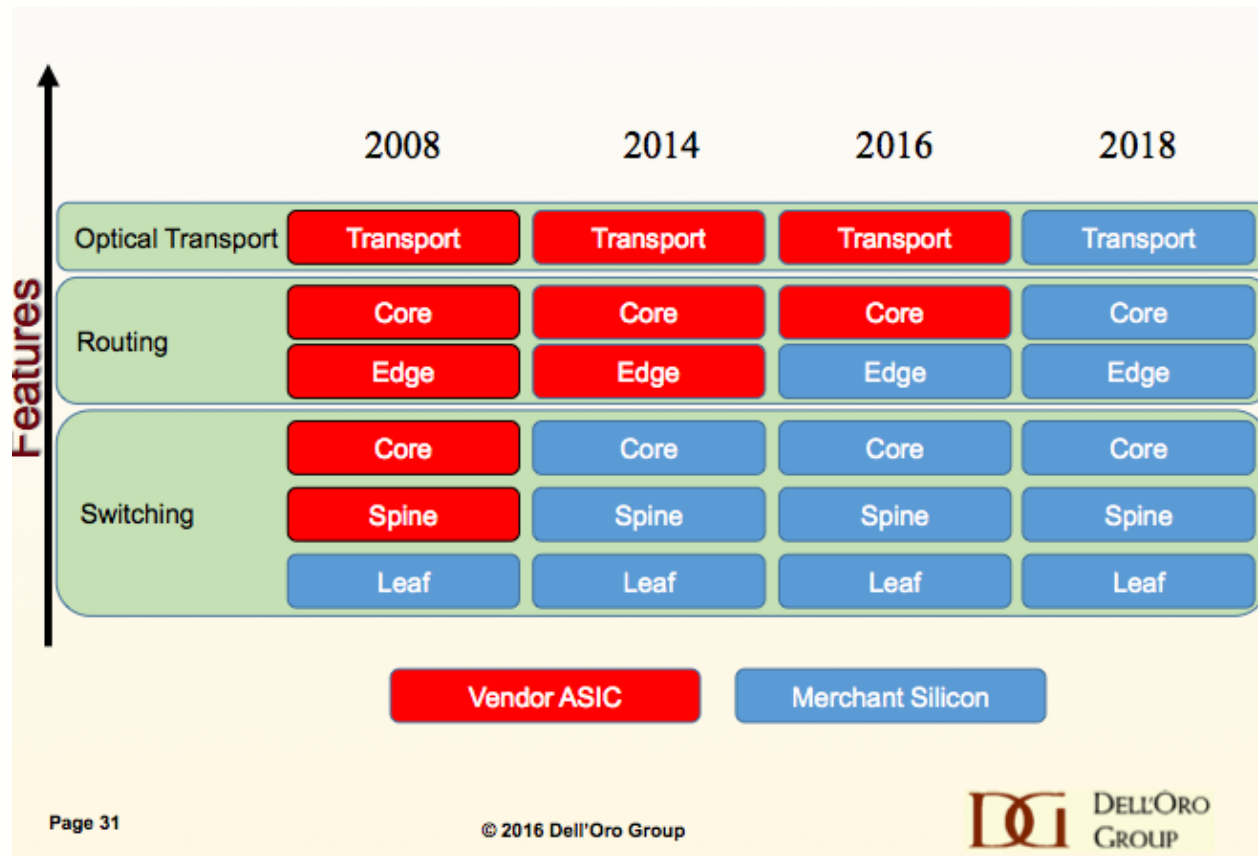
マーチャントシリコンシングルチップ辺りの帯域変化



運用者から見ると何が嬉しいのか？

- ビット単価が安くなる
- bpsあたりの消費電力が小さくなる
- ポート密度が上がる
- 遅延が小さくなる

マーチャントシリコンとベンダー-ASICの関係



ネットワークとムーアの法則に関わる10の予測

- アンディ・ベクトルシャイムは2012年に10の予測を言った

<https://venturebeat.com/2012/10/11/bechtolsheims-10-predictions-about-networking-and-moores-law/>

10の予測

1. ムーアの法則はこれからも生き続ける
2. チップの経済性は変化している。一つの会社が設計者を抱えておく事ができなくなっている
3. マーチャントシリコンベンダーが市場を獲得する
4. カスタムデザインが市場をリードする
5. 最高のシリコン技術を採用する事が重要(最新のチップが安く/早く/小さくなる)

10の予測

6. 製品のライフサイクルは短くなる
7. アーキテクチャーは重要である
8. 柔軟性も重要
9. ビルディングブロックが重要
10. シングルチップデザインになるべき

Agenda

- このお話の背景
- **過去のJANOGでの議論とこの話の関係**
- **議論**

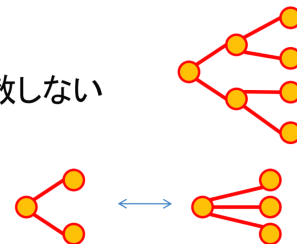
JANOG27仕様と運用

<https://www.janog.gr.jp/meeting/janog27/program/spec.html>

- 運用者と仕様の間係を議論
- 不明確だが多様してる例としてロードシャアリングなどを提示

load sharingに必要な要素

- 網内で多段分散
 - 同じkeyのみで分散しない
- 分散数
 - 奇数、偶数、上限
- 分散状態の維持
 - 経路変動、構成変更
- 筐体内で多段分散
 - IGP、BGP、Link Aggregation



- もしシリコンファミリーが同じが一緒なら、仕様はほぼ一緒 (BRCM-DNX :Petra->Arad->Jaricho)
- ASICとのプログラムを共通化するSAIなどもある

<https://azure.microsoft.com/ja-jp/blog/switch-abstraction-interface-sai-officially-accepted-by-the-open-compute-project-ocp/>

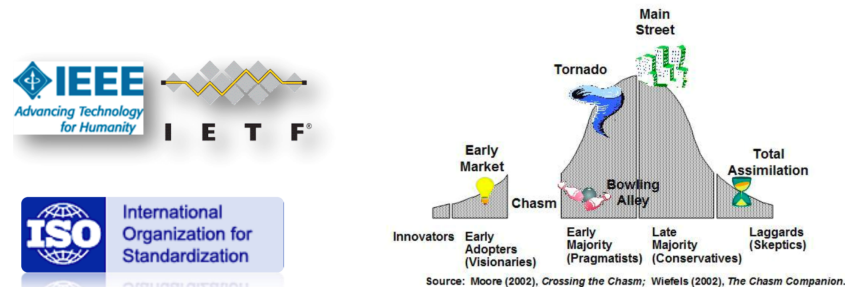
JANOG29 どうする？ どうやる？ データセンター間ネットワーク

<https://www.janog.gr.jp/meeting/janog29/program/dcdr.html>

- VPLS/EVPN/802.1aq/Trill/OTV/FabricPathなどの技術を紹介する中で何を選ぶべきなのかを提示
- 難しいのはマーケットを取るためのベンダーの戦略との機器選定の関係

キャズム理論 Crossing the chasm MAC Routing Trill vs IEEE802.1aq vs Proprietary

Technology Adoption Life Cycle:
Diagnose and adapt as markets evolve



ベンダーはキャズムを超えて、アーリーマジョリティの為に製品を投入する

サービスのライフサイクルを考えると、メインストリームになりうる機器を選ぶ事が必要

JANOG36 TRILL/SPBに見る標準化の課題

<https://www.janog.gr.jp/meeting/janog36/program/trill>

- TRILL/SPBの類似技術を議論
- TRILLに関してはベンダー独自が多い事を言及

独自技術を利用するリスク



Photo Credit: [abardwel cc](#)

ベンダロックイン

- 将来に渡って、同一ベンダの互換ソリューションを利用しなくてはいけなくなるかもしれない

継続性

- その技術が流行らなかつたら、サポートや後継製品もなくなるかもしれない



Photo Credit: [Thomas8047 cc](#)



Photo Credit: [Joshua Tree National Park cc](#)

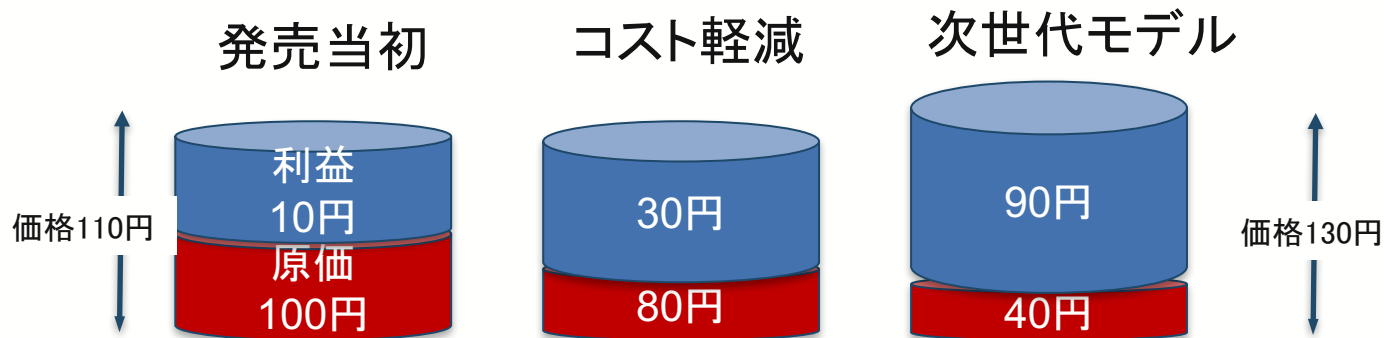
罨 = 障害の頻発

- 利用ユーザが少ないが故に
“こんな障害、お客様が初めてです!”
としょっちゅう言われるかもしれない

DMM.com Labo

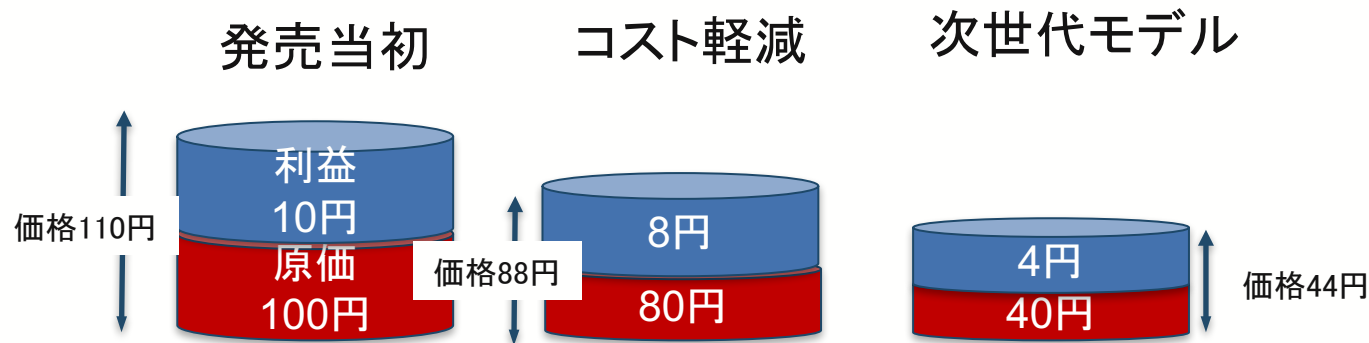
9

では今後は起こる得るのか？



- 製品のコンポーネントが明示されていない場合
- 一般的な製品利益は上記の様なモデルで形成
- コスト削減モデル(価格一緒);市場を取ったあとの次世代モデル:単価据え置き

では今後は起こる得るのか？



- マーチャントシリコンを使っていると原価が見えている為、今までの様な利益率モデルは取りにくい
- 利益を取るための付加価値として独自機能を考えるかもしれない

敵は味方の振りをする

Be ware of enemies disguised as friends.

VXLAN Header:

```
+++++
|R|R|R|R|I|R|R|R|                Reserved                |
+++++
|                VXLAN Network Identifier (VNI) |   Reserved   |
+++++
```

VXLAN-GBP Header:

```
+++++
|G|R|R|R|I|R|R|R|R|D|R|R|A|R|R|R|                Group Policy ID                |
+++++
|                VXLAN Network Identifier (VNI)                |   Reserved   |
+++++
```

Rビットは送信時に0にセット
受信時には無視をする

従来のRビットにGが入り
互換性は無い

- [VXLAN](#)と[VXLAN-GBP](#)の違い
- ヘッダーフォーマットは似ているが根本的な所は違う
- なぜ新フォーマットとして宣言しないか・・・
- 似たようなケースは沢山・・・
- もちろん主流にならないとは言い切れない。But...

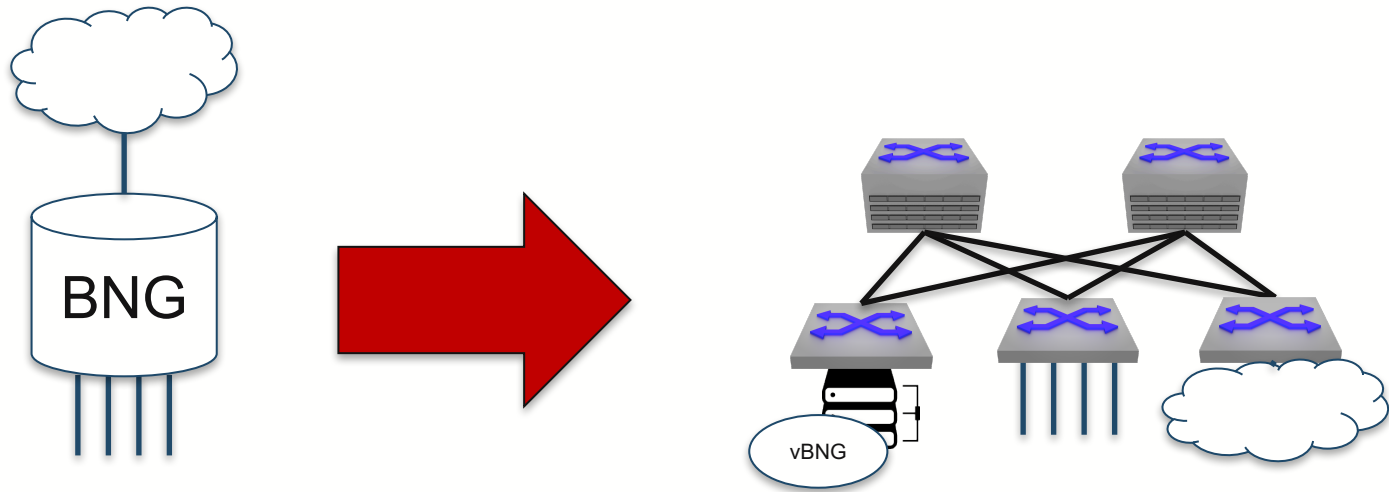
運用者はどうロックインを防ぐのか

- 寄らば大樹の陰
 - 避けられないなら潰れ無さそうな会社による(でも製品が無くならないとは限らない)
- 信頼出来る人を見つける
 - 独自機能を実装しない人を見つける
 - 標準化および実装を絶えず観察している人を味方につける
 - 自分でキャッチアップする
- プログラマブルパイプライン
 - CAVIUM-XPA
 - Broadcom Trident 3
 - P4.org対応の製品など

ライフサイクルが短い製品への対応

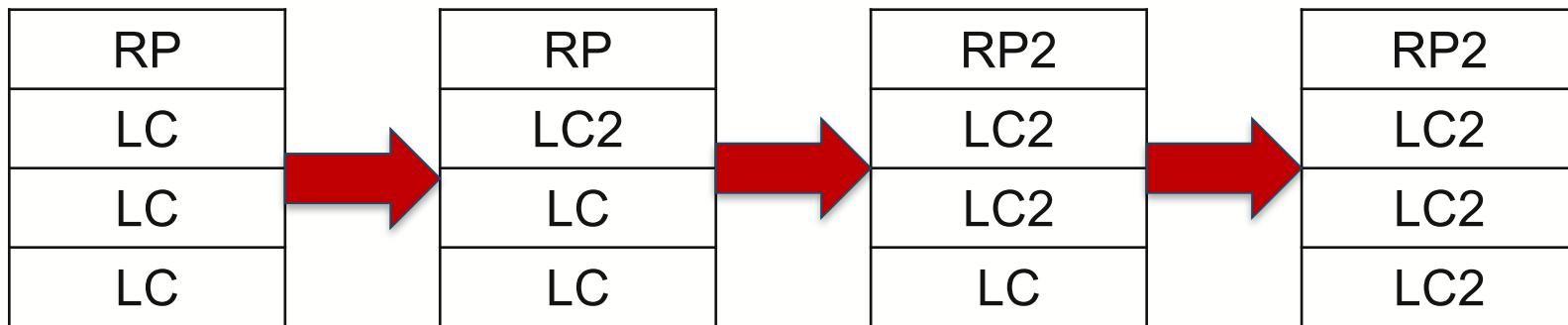
- 影響が少ないデザインにする
- 下位互換性を提供するモジュラー機器を選ぶ
- 落ちない機器を選ぶ

影響の少ないデザインにする



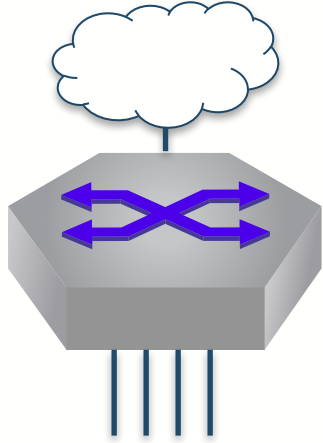
- 加入者および機能を一点に集中してしまう為に影響を受けやすい
- 機能および収容を分散する
- CORD(the Central Office Re-architected as Data Center)などの考え方と同様

下位互換性を提供するモジュラー機器を選ぶ



- 下位互換を提供するモジュラー型機器を選ぶ

落ちない機器を選ぶ



- 高品質の機器を選考する
- もちろんソフトウェアにはサポート期間はあるが、枯れていたら運用し続ける？

まとめ

- 2008年くらいより徐々にネットワークの世界での競争原理が持ち込まれてきた
- JANOGではホワイトボックスを中心に議題が進められていたが、ベンダーでも当然マーチャントシリコンが進められ、市場も変わりつつある
- ガラケーがスマホに変わった様に、ネットワークも変わる可能性がある
- ネットワーク設計など影響があるため是非議論をしたい

Agenda

- このお話の背景
- 過去のJANOGでの議論とこの話の関係
- **議論**



Thank You

www.arista.com