

TechDAY 2018
2018年11月22日

ホワイトボックス・スイッチの期待と現実



25th
Anniversary

株式会社インターネットイニシアティブ
ネットワーク本部 SDN開発部
白崎 博生

コスト削減

- 調達コスト（機器,ライセンス等）
- オペレーションコスト

必要機能を 独自開発

- API
- OpenFlow
- 独自プロトコル／独自機能

ベンダー ロックイン 排除

- ベンダー独自仕様の排除
- ユーザー主導によるライフサイクルコントロール

とはいえ...「IIJ」では

- 検証は継続していますが、ホワイト・ボックススイッチをサービス環境では使用していません
- 既存のメーカー製スイッチの置き換えを目的に採用することもないと思います

ホワイト・ボックス スイッチは**使えない子なのか!?**

- ✓ ネットワーク機器ではなくサーバ機器として使ってみる

ホワイト・ボックススイッチとサーバ機器を組み合わせ、
ひとつのコンピュータリソースとしての利用を検討中

- ✓ DataCenter as a Serviceならぬ「**Rack as a Service**」を実現

✓ 思ったよりもコストが高い

- 有料NOSを搭載するとメーカー製品よりも高いことも
- 独自開発NOSの開発費（初期開発費と開発体制維持費）
⇒コストメリットを出すにはある程度の展開ボリュームが必要

コスト

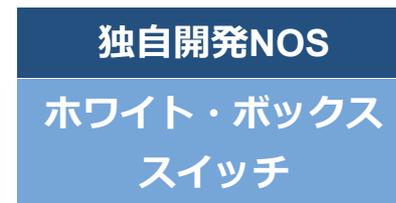
有料NOSを搭載すると
メーカー製品と同じか、
高いことも...



展開台数が少ないと
独自開発NOSの開発費
負担が大きい



(台数が**少ない**場合)



(台数が**多い**場合)

✓ 保守体制はゼロから構築

- 有料NOSの場合はOSメーカーが一次対応はしてくれるが、レスポンスは??
- 独自開発NOSの場合は、問題箇所の切り分けから自分たちで行う

✓ 意外と製品寿命が短い

- 筐体のEoS (End of Sales) までおおよそ3年??
「お代わり」したくてもできない可能性がつかまとう
- 特定メーカーのASICを搭載する機種を突然製造しなくなるときもある
- ホワイトボックス・スイッチメーカーの製品ロードマップ全体の見直しも突然おきる

- ✓ ASIC がなくなることもある
 - ・ ファツ！？
- ✓ メーカー製スイッチと接続できないこともある
 - ・ A社製スイッチとつないでみた

A社スイッチ QSFP+ 40GbE	ケーブル	WB スイッチ QSFP+ 100GbE	リンク アップ
	純正 DAC 40GBASE-CR4 3m		NG
	純正 DAC 40GBASE-AOC 15m		OK

～ 使えないムードが漂う ～

- ✓ ASIC がなくなることもある
 - ファツ！？

- ✓ メーカー製スイッチと接続できないこともある
 - A社製スイッチとつないでみた

A社スイッチ QSFP+ 40GbE	ケーブル	WB スイッチ QSFP+ 100GbE	リンク アップ
	純正 DAC 40GBASE-CR4 3m		NG
	純正 DAC 40GBASE-AOC 15m		OK
純正 40GBASE-SR4	MPO (MPO/MTP)	不明 40GBASE-SR4	OK
	サードパーティDAC 40GBASE-CR4		OK

} 追加判明

✓ ASICのSDKはASICメーカーに直接コンタクト

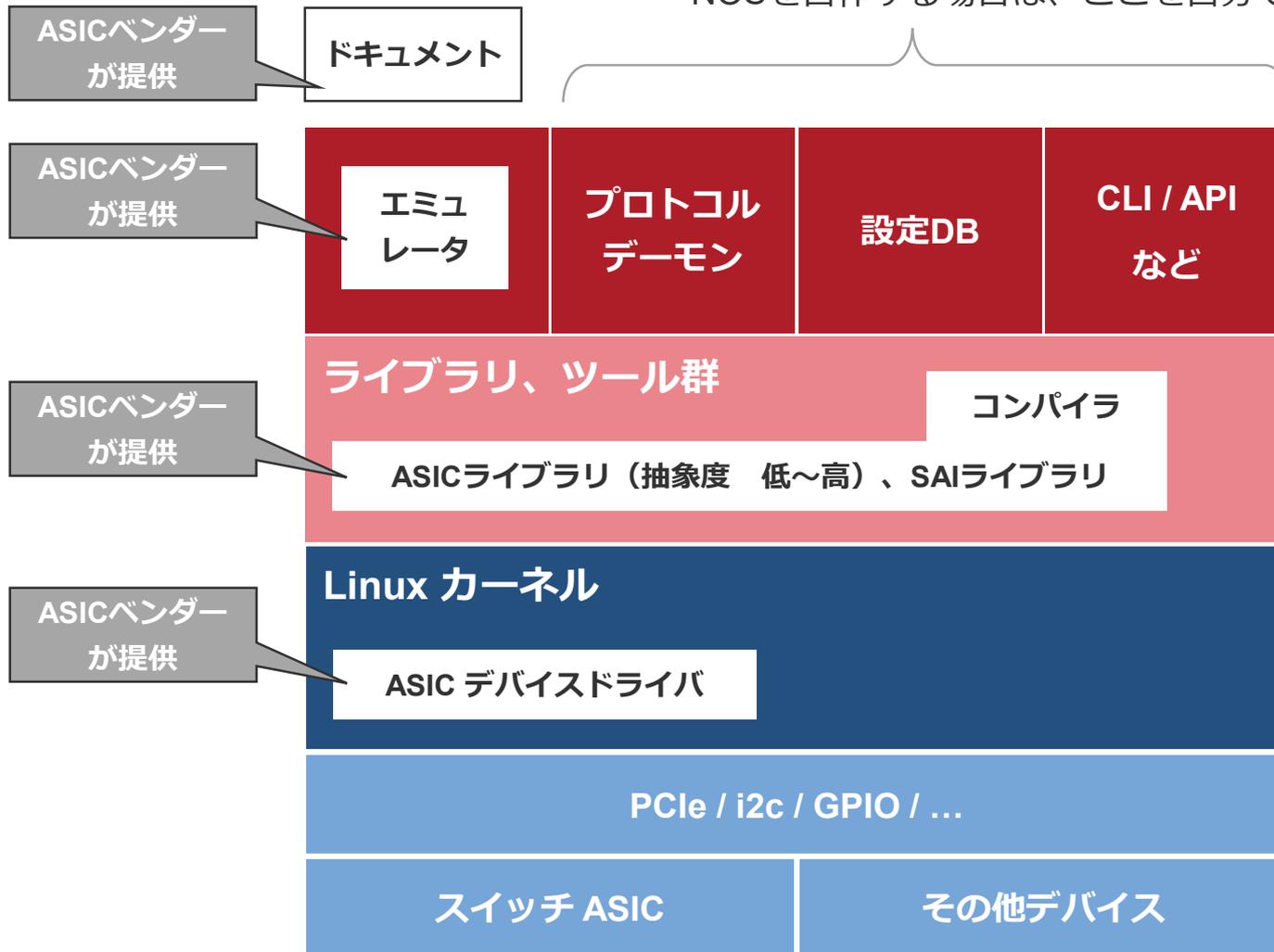
- NDAとSLAの締結が必要
 - ⇒SDKの深さ毎にライセンスが必要
- SLAの内容は「製品販売」が前提
 - ⇒自社開発・自社利用の場合は頭をひねって解釈する
- 内容修正に応じる会社・応じない会社
 - ⇒修正できない場合はさらに頭をひねる
- ドキュメントが少ない
 - ⇒問い合わせベースの開発作業
- 無料だったり、実質無料だったり、有料だったり

✓ ASICメーカー毎に異なるSDK API

- ここにもベンダーロックインが！
 - Broadcom : OpenNSL
 - Cavium : OpenXPS
 - Mellanox : OpenEthenet
 - Barefoot : P4
- 学習コストが半端ない
 - ⇒ . . . といってもフレームワークみたいなものですが
- SAI (スイッチ抽象化インターフェース) という標準APIもありますが . . .

SDKで提供されるもの

NOSを自作する場合は、ここを自分で用意する



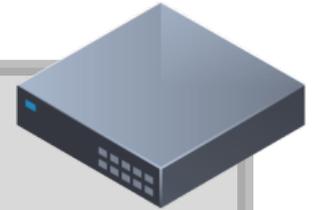
※ 古いgccではコンパイルできなかつたり、新しいgccではコンパイルできなかつたり・・・

✓ 開発は難しくないけど簡単でもない

- LinuxだけどLinuxじゃない
- 決まったロジックだけを実装するのは難しくない
⇒Pythonでも書ける
- 「何でもできるようにする」のは難しい！
 - APIで操作できることは非常にプリミティブ (だけど多い)
 - 利用するASIC機能の取舍選択とプロファイル設定
⇒ASICの得意(特異)な機能
 - 「モデル」の定義 (←ここが一番難しい)
 - ユーザインタフェースの設計
⇒コマンド引数、設定ファイル等
- ASICは意外といろいろやってくれる
- ASICは意外なことをやってくれない
- SONiCやOpen Network Linuxを改造するのがお手軽？

- **某大手メーカー（一体型スイッチメーカー）の対応**

- ✓ **一通りのAPIを実装**している ✨
- ✓ **Linuxインターフェースを解放**している ✨
- ✓ Ansibleのような**プロビジョニングツール**が使える ✨
- ✓ ユーザーの**リクエストを実装**してくれる ✨



ホワイト・ボックス スイッチをスイッチとして使う

必要性が薄れてきた . . .

Q. ホワイト・ボックススイッチは「ネットワーク機器」か「サーバ機器」か？

ネットワーク機器



- 見た目（ポートがたくさんある）
- パケットをフォワードする



サーバ機器



- ネットワークプロセッサを載せた

ネットワーク機能が超強力な
アクセラレータを載せたPCサーバ

IIJのアプローチ

ホワイト・ボックス スイッチを
「**サーバ機器**」として扱ってみよう！



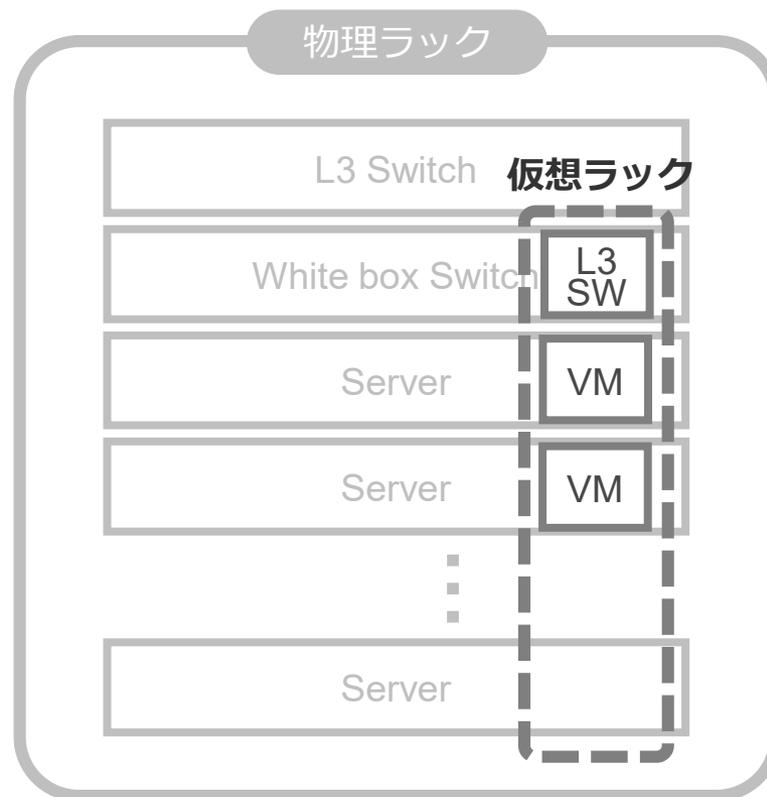
- ✓ ホワイト・ボックス スイッチとサーバ機器を組み合わせて、ひとつのコンピュータリソースとしての利用を検討中
- ✓ **DataCenter as a Service**ならぬ、**Rack as a Service**を実現
 - ⇒物理ラック内に收容される様々な機能を持った物理リソースを細分化し、必要な分だけ論理的なラックにまとめてリソース提供

例えば・・・

- 仮想 L3 Switch + Kubernetes, ESXi
- フィルタリング
- LB+NAT を ホワイト・ボックス スイッチ に置けばDSR要らず

つまり

- サーバ側で実装すると面倒なことを ASICにオフロードしよう
- ASICの方が速いし

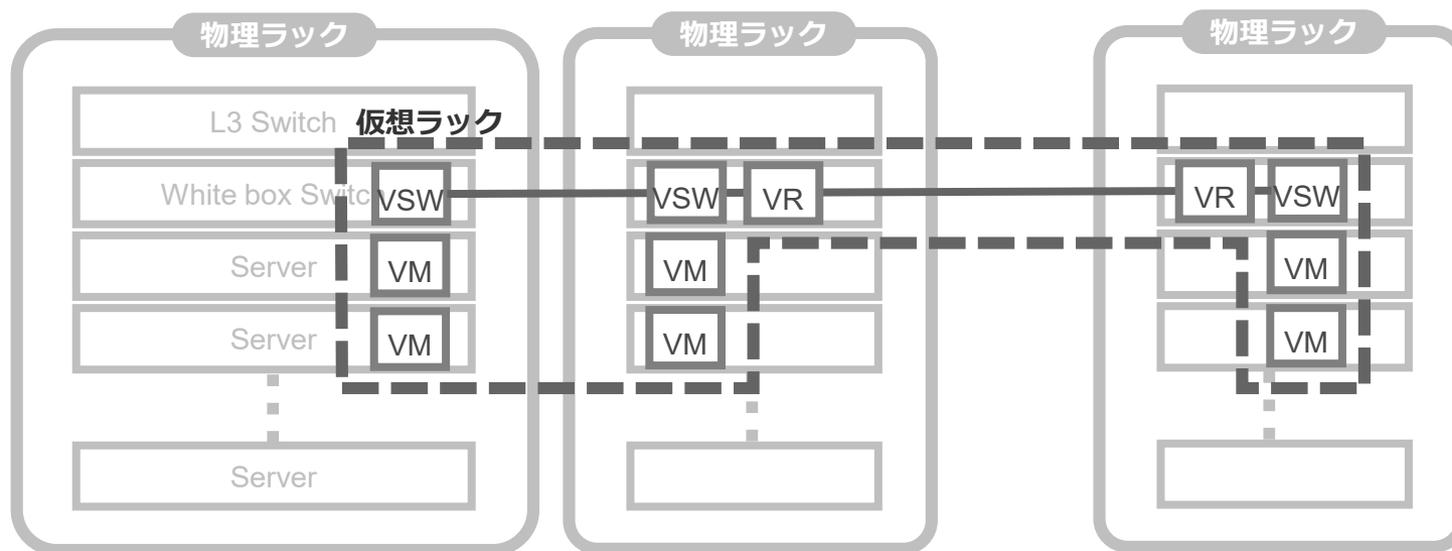


✓ 仮想L3スイッチで行うこと

- ✓ 仮想ルーター、仮想スイッチ
- ✓ トンネル終端（回線系サービス等との接続）
- ✓ NAT、NAPT、フィルタ
- ✓ ロードバランサー
- ✓ 帯域制御

プロトタイプ開発中

- VM上にVRを実装すると遅い
- VM上に機能分散すると複雑
- ASICの有り余るパワーで解決



- ✓ **正直なところ未知**
- ✓ **考えることはまだたくさんある**
- ✓ **開発効率**
- ✓ **ホワイトボックス・スイッチ vs スマートNIC**
- ✓ **運用体制**
 - ✓ 壊れ方が分からない（未経験）
 - ✓ 監視項目、監視パラメータ
 - ✓ 遠隔操作（IPMI がないよ）
 - ✓ 交換機材、修理窓口、修理期間、交換作業要員
- ✓ **逃げ先を用意しておくこと**
 - ✓ 別のASIC、別メーカーの箱
 - ✓ 選択肢は少ない



日本のインターネットは1992年、IIJとともにはじまりました。以来、IIJグループはネットワーク社会の基盤をつくり、技術力でその発展を支えてきました。インターネットの未来を想い、新たなイノベーションに挑戦し続けていく。それは、つねに先駆者としてインターネットの可能性を切り拓いてきたIIJの、これからも変わることのない姿勢です。IIJの真ん中のIはイニシアティブ
————— IIJはいつもはじまりであり、未来です。

本書には、株式会社インターネットイニシアティブに権利の帰属する秘密情報が含まれています。本書の著作権は、当社に帰属し、日本の著作権法及び国際条約により保護されており、著作権者の事前の書面による許諾がなければ、複製・翻案・公衆送信等できません。本書に掲載されている商品名、会社名等は各会社の商号、商標または登録商標です。文中では™、®マークは表示していません。本サービスの仕様、及び本書に記載されている事柄は、将来予告なしに変更することがあります。