

NOKIA

MPLS JAPAN

DC Networkingにおける DisaggregationとオープンNOS

NOKIA

鹿志村康生

yasuo.kashimura@nokia.com



最新のデータセンターネットワークの要件

高スケール
高信頼性

アプリケーション
がエッジ、地域、
コアに分散

コントロール

必要に応じた自由な
カスタマイズとツールの適用

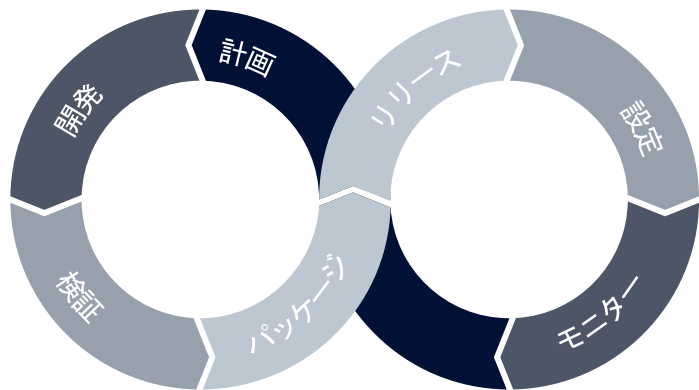
スピード感

オペレーションの
シンプル化と自動化

CAPEXとOPEXを最小限に抑えながら、これらすべてを実行

次世代のDCネットワークキングへの道

DevOpsの考え方をネットワークキングへ - NetOps

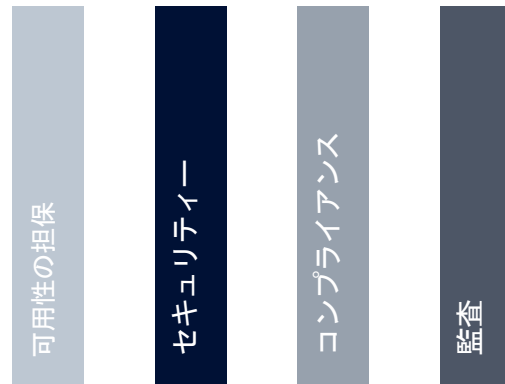


DevOps

- アジリティー
- コーディングスキル
- 可視化
- ……



ネットワーク
オペレーター



日々のネットワーク
オペレーション

Nokia データセンターファブリックソリューション

オープン・スケーラブル

Fabric services system
オートメーション及びオペレーションツールキット









Intent-driven design and operations 	Digital sandbox for true emulation 	Cloud-native integration model 
-----------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------

Cloud-native, Kubernetes-based extensible foundation

Service Router Linux
ネットワークオペレーションシステム(NOS)






Open, scalable telemetry framework 	Open-source CLI plugins (Python) 	NetOps Development Kit (NDK) 
Standard Linux kernel 	Ground-up, model-driven foundation 	Resilient, field-proven protocol stacks SR OS 

本日の
フォーカス

7250 IXR, 7220 IXR
データセンター
ネットワーキング
プラットフォーム



Based on merchant silicon 	Common hardware and software design 	Full range of DC applications 
-------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------

DCネットワークにおけるNetOpsを実現するために “オープン化”

オープン/スケラ
ブルな telemetry
framework



オープンソース
CLI プラグイン
(Python)



NetOps
Development
Kit (NDK)



ソフトウェア / NOS

gNMI/gRPC



Nokia SROSベース
プロトコル

OS

モデルドリブン
CLI



プロセス間通信
iDB



ハードウェア / Whitebox

Linux
カーネル



Broadcom
Trident, Tomahawk,
Jericho



Box/モジュラー
デザイン



HW 抽象化
- XDP



DCネットワークにおけるNetOpsを実現するために “オープン化”

オープン/スケラ
ブルな telemetry
framework



オープンソース
CLI プラグイン
(Python)



NetOps
Development
Kit (NDK)



ソフトウェア / NOS

gNMI/gRPC



Nokia SROSベース
プロトコル

OS

モデルドリブン
CLI



プロセス間通信
iDB



ハードウェア / Whitebox

Linux
カーネル



Broadcom
Trident, Tomahawk,
Jericho



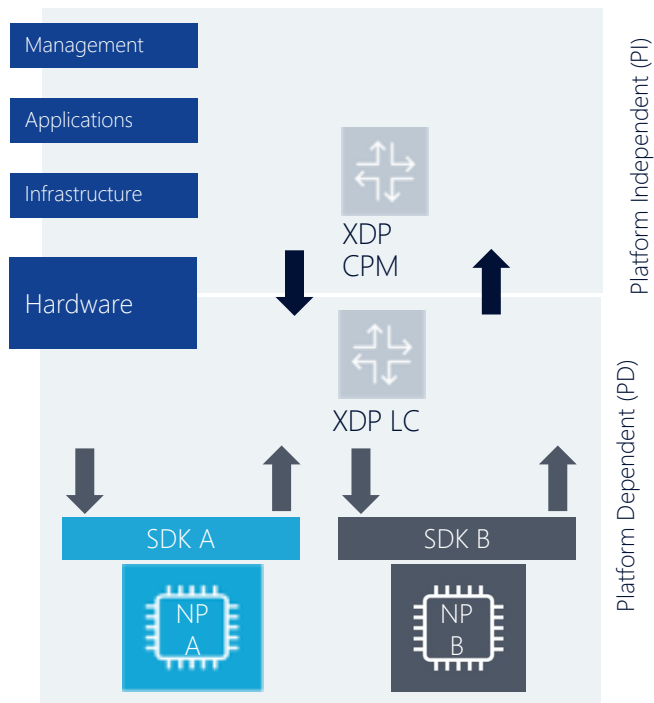
Box/モジュラー
デザイン



HW 抽象化
- XDP



様々なハードウェアへの柔軟な対応 迅速な導入を可能にするオープンなインフラ: eXtensible Data Path (XDP)



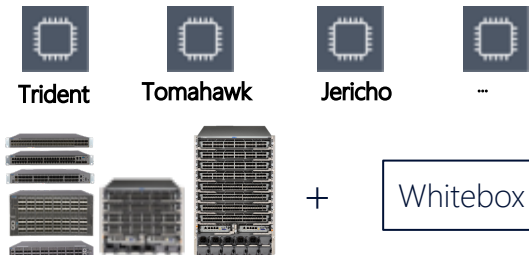
シリコン側のSDKの差分をXDP抽象化レイヤで吸収することによりSRLinuxソフトウェア側への影響を極小化

拡張性

- シリコンSDKを共通のソフトウェア命令セットに抽象化することにより、ハードウェアをSR Linuxのコントロール/管理プレーンと完全に分離
- 新しいシリコンの市場投入までの時間が短く、XDPより上位のアプリケーションでは変更が不要
- アタッチメントメカニズムに関係なく、ACL、QoS、その他のインターフェイスポリシーの均一な設定および動作モデル
- ONIEをサポートを含むホワイトボックス対応Ready

シリコンの多様性

ターゲット市場/用途に応じて異なるシリコンを使用したプラットフォームを活用可能



XDPによるControl planeの処理

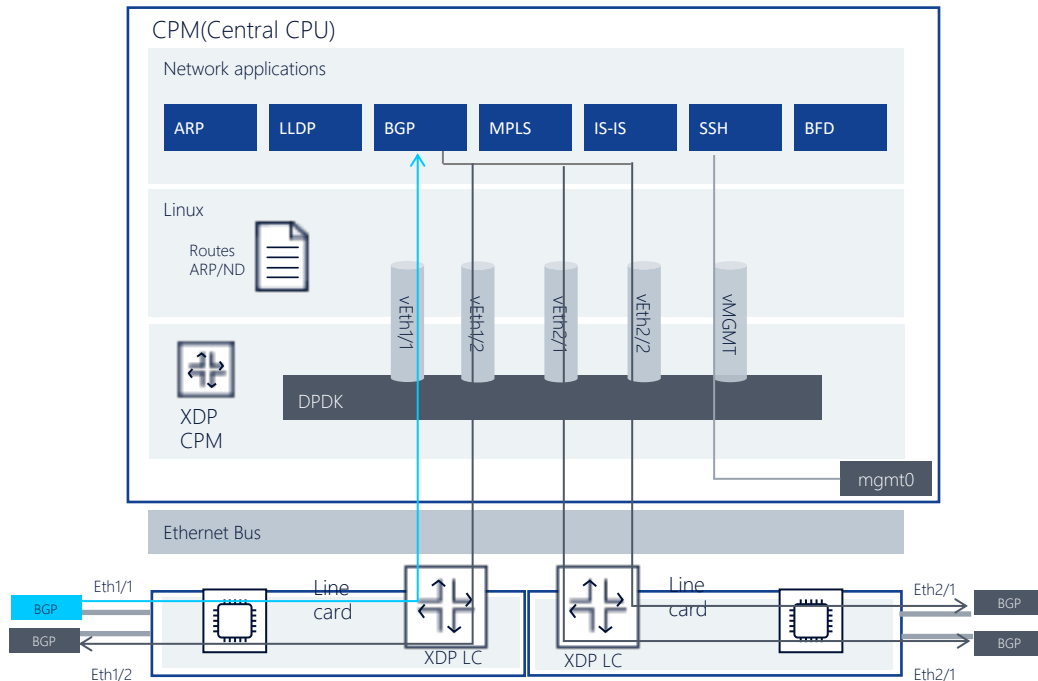
Rx 及び Tx の内部転送

Rxの内部転送

1. コントロールプレーンパケットがXDP LCで受信されると、まずそのパケットがLocalで処理すべきものかどうかを判別する
2. Local処理するものではないと判別した場合には、そのパケットはICC(Inter Card Communication)フレームとしてIngressインタフェース/ネットワークインスタンス等の情報と共にシリアル化されXDP CPMへ送られる
3. XDP CPMはICCフレームをデシリアライズし、該当のvEthパスへ送る

Txの内部転送

1. アプリケーション(eg. BGP)からコントロールパケットが送信されると、Linuxにより該当のvEthパスからXDP CPMに送信される。XDP CPMはパケットをICCフレームにシリアル化しXDP LCに送信する
2. XDP LCはICCフレームをデシリアライズし、ASICヘッダを付加する。ASICはパケットを該当のインタフェースに送信する



DCネットワークにおけるNetOpsを実現するために “オープン化”

オープン/スケーラブルな telemetry framework



オープンソース CLI プラグイン (Python)



NetOps Development Kit (NDK)



ソフトウェア / NOS

gNMI/gRPC



プロトコルスタック

SR OS

モデルドリブン CLI



プロセス間通信 iDB



ハードウェア / Whitebox

Linux カーネル



Broadcom Trident, Tomahawk, Jericho



Box/モジュラーデザイン

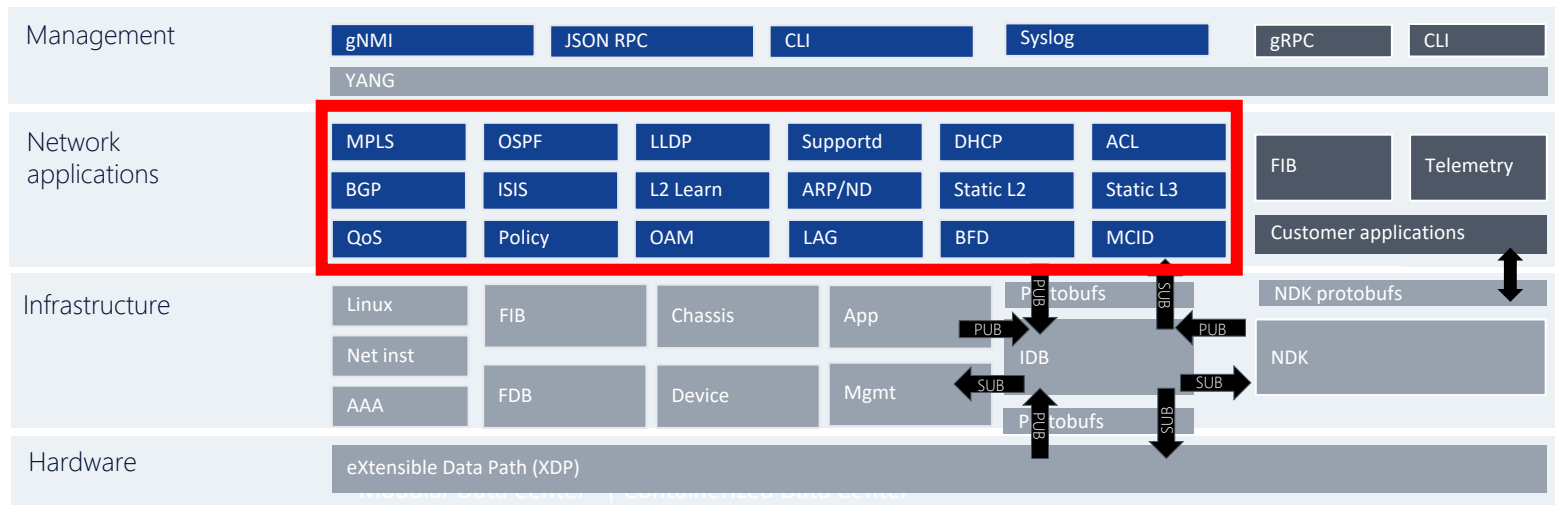


HW 抽象化 - XDP



プロトコルスタック & CLI

Open Infrastructure Enabling Boundless Possibilities



スケールと信頼性

- 20年以上フィールドで鍛えられてきたNokia SR OSのプロトコルスタックとネットワークサービス
- 100万台以上のルーターが1300以上のIPネットワークに導入

モデルドリブンマネジメント

柔軟な書式
具備しているテレメトリーの活用によるプログラマビリティ

オープンインフラストラクチャ

- 柔軟な ASIC/HWプラットフォームのサポート
- ユーザー独自のアプリケーションをシームレスにネイティブアプリケーションと統合可能

DCネットワークにおけるNetOpsを実現するために “オープン化”

オープン/スケーラブルな telemetry framework



オープンソース CLI プラグイン (Python)



NetOps Development Kit (NDK)



ソフトウェア / NOS

gNMI/gRPC



プロトコルスタック

SR OS

モデルドリブン CLI



プロセス間通信 iDB



ハードウェア / Whitebox

Linux カーネル



Broadcom Trident, Tomahawk, Jericho



Box/モジュラーデザイン

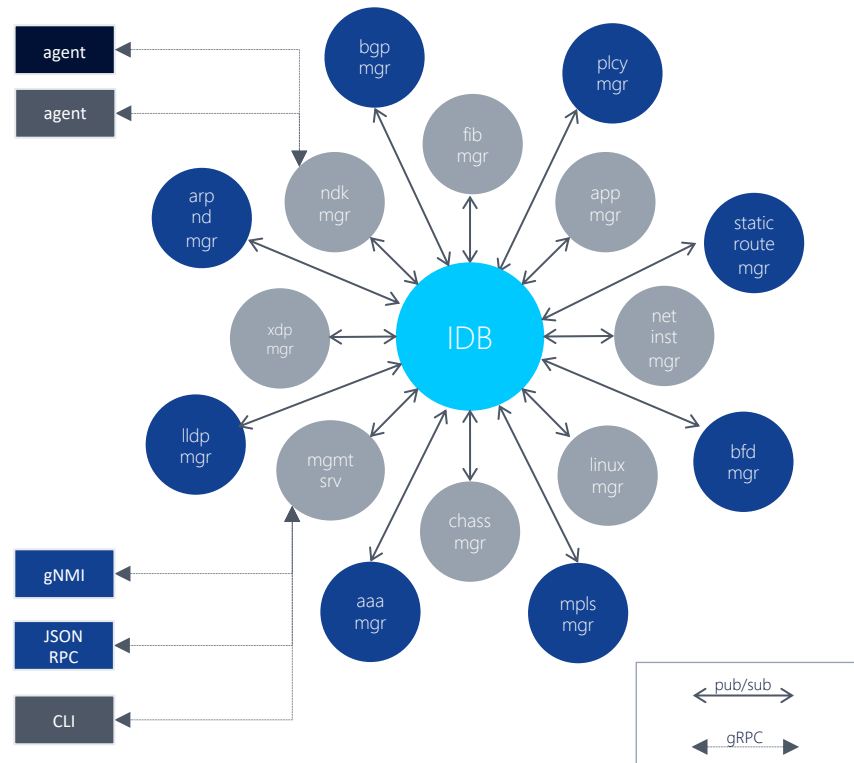


HW 抽象化 - XDP



プロセス間通信 : Impart Database(インパートデータベース/IDB) NOSに最適化

- ライトウェイト DB, Redisスタイル, pub/subベース
- DBはTopicによって分かれる
- それぞれのTopicは1つのアプリケーションのみによって所有される (Cross-Topicでの書き込み無し)
- アプリケーションはIDBへのSessionをOpenしTopicをSubscribeする, また他のアプリケーションのために自身が所有するTopicに対してMessageをpublisする
 - エージェントが直接振る舞いを操作することも許容, eg. 独自のRouteをpublisしFIB manageにそのRouteを選択させる
- アプリケーションのStateはアプリケーション内で持ち, 他のアプリケーションが興味のあるStateのみIDBにpublisする
 - IDBを簡素化, 不要なDBの膨張を防止, IDBはData payloadやStructureは気にしなくて良い
 - Telemetryが常にアプリケーションからStream
- DBの永続性は不要, Speedを重視, データストアというよりもキャッシュのイメージ
- パワフルなNDKとともに提供, オペレーターは内部システムイベントを元にアプリケーションでイベントハンドリングを実行可能



DCネットワークにおけるNetOpsを実現するために “オープン化”

オープン/スケラ
ブルな telemetry
framework



オープンソース
CLI プラグイン
(Python)



NetOps
Development
Kit (NDK)



ソフトウェア / NOS

gNMI/gRPC



プロトコル
スタック

SR
OS

モデルドリブン
CLI



プロセス間通信
iDB



ハードウェア / Whitebox

Linux
カーネル



Broadcom
Trident, Tomahawk,
Jericho



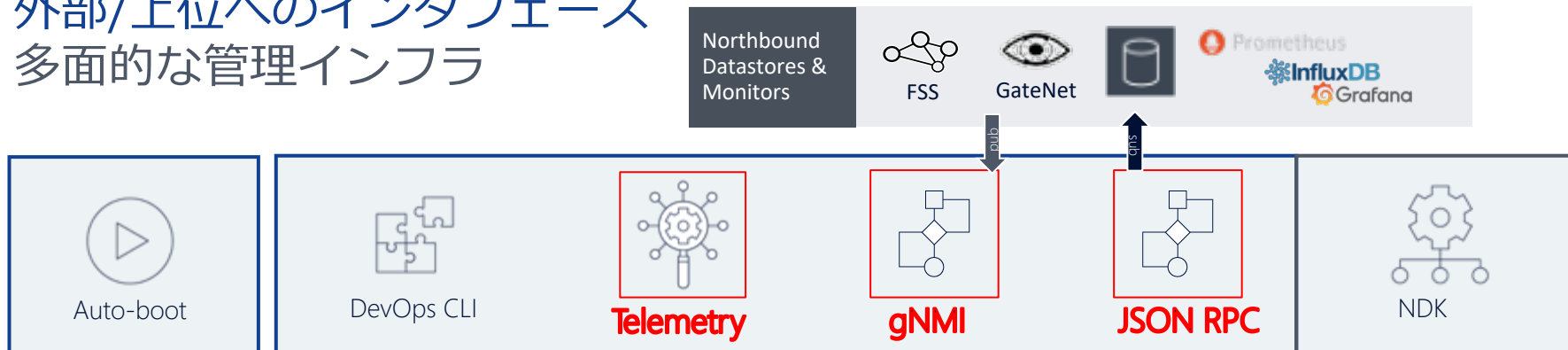
Box/モジュラー
デザイン



HW 抽象化
- XDP



外部/上位へのインタフェース 多面的な管理インフラ



ZTP

- カスタマイズ可能なPython実装
- v4/v6
- HTTP/s
- Out-of-band
- In-band

ONIE(Open Network Install Environment)

- Pythonベース、完全にオープンソース
- オペレーターによる拡張を無制限に可能にするプラグイン
- エイリアス、ユーザーごとの環境

- ゼロから構築されたネイティブなTelemetryアーキテクチャー
- 拡張性と堅牢性
- Auto-telemetryフレームワーク

- ストリーミングインターフェース
- gRPC/protoベース
- 統合された堅牢なマネジメント、設定とTelemetryのためのプロトコル

- 設定とState取得のためのプログラマブルJSONインターフェース
- CLIコマンドをサポート
- Nokia拡張(commit confirmed, validate, etc)

- NetOps Development KitはSR Linuxのインフラストラクチャをサービスとして公開
- アプリケーションはトランザクション構成、ストリーミング、テレメトリを活用可能

DCネットワークにおけるNetOpsを実現するために “オープン化”

オープン/スケーラブルな telemetry framework



オープンソース CLI プラグイン (Python)



NetOps Development Kit (NDK)



ソフトウェア / NOS

gNMI/gRPC



プロトコルスタック

SR OS

モデルドリブン CLI



プロセス間通信 iDB



ハードウェア / Whitebox

Linux カーネル



Broadcom Trident, Tomahawk, Jericho



Box/モジュラーデザイン

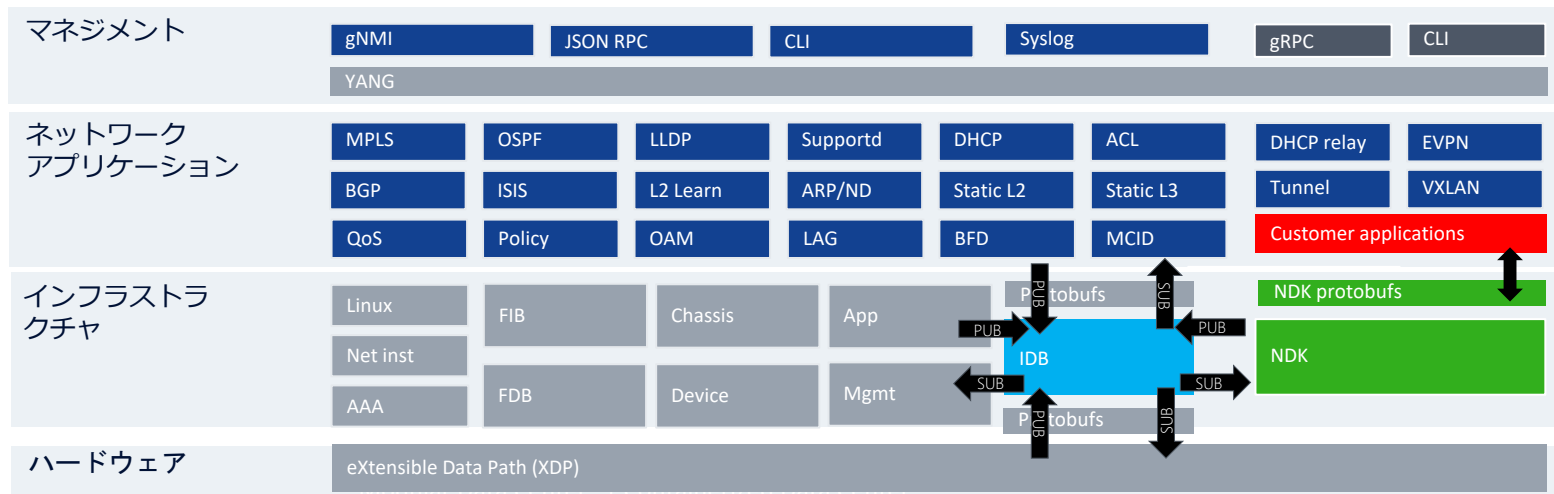


HW 抽象化 - XDP



NDKによるオープンな開発

無限の可能性を実現するオープンなインフラストラクチャ



NDKによるオープンな開発

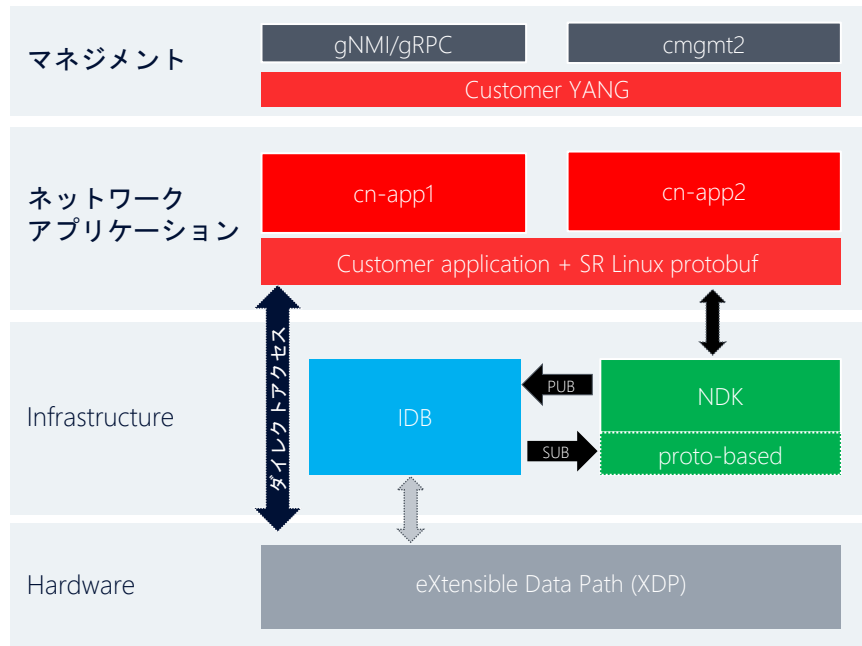
無限の可能性を実現するオープンなインフラストラクチャ

Proto-based NDK

- ランゲージニュートラル
- C++, C#, DART, GO, JAVA, Python, Ruby
- On-box 又は Off-boxでのアプリケーション
インテグレーション

Native integration

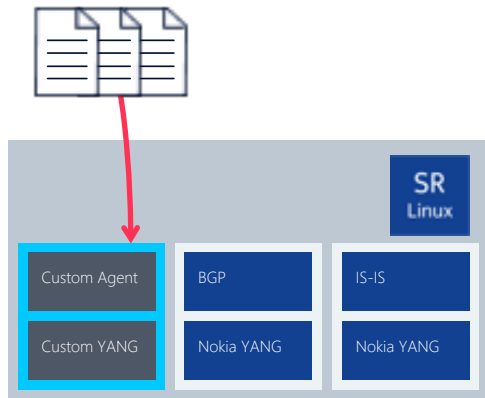
- どのNorthboundインタフェースからでもアクセス可能
- アプリケーション/エージェントからデータプレーンへのダイレクトアクセスも可能



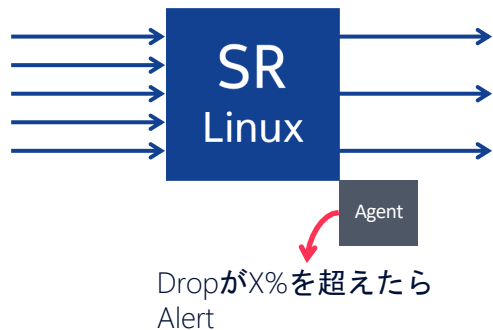
シンプルなエージェントアプリケーション例

トラフィックロ検知

ZTPの際にエージェントをインストール



エージェントは継続的に
IN/OUTのTraffic量をモニター

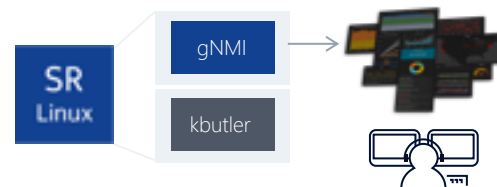
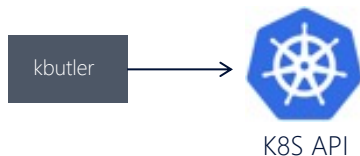


SR Linuxは他のアプリケーションと同様にエージェントのデータをgNMIでストリーム



カスタムアプリケーション例: kbutler

コンピュータ側とネットワーク側の情報をMapしモニター



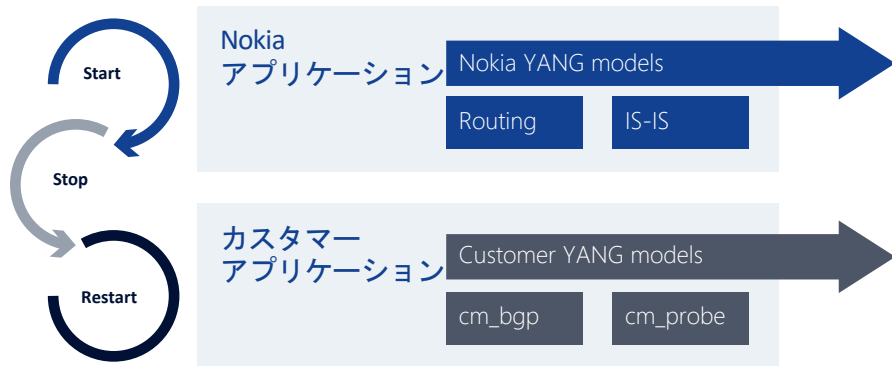
- Kbutlerは K8SサービスAPIをサブスクライブし新しいサービスのexposeや既存サービスの変更をモニター
- どのアプリケーションがどのWorker nodeでホストされているかのViewを得るため

- Kbutlerは SR LinuxのNDK APIをサブスクライブしFIBの変化をモニター
- Kbutlerはサービスアプリケーションが動いているWorker nodeに対して、サービスの外部公開IPアドレスのFIBエントリが存在し、Next-hopがWorker nodeになっていることを監視

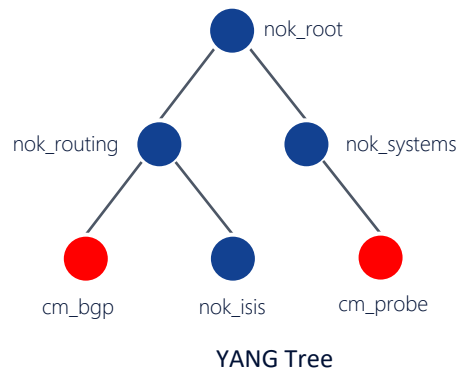
- Kbutlerはオペレーション状況をSR Linux無いでレポートし何か問題・不一致が発生した場合にはすぐにAlertを上げる
- Kbutlerはコンテキストに基づいたモニタリング、アラート、トラブルシューティングを提供

Full stack integration

Lifecycle management, YANG, Python CLI



- : Nokia SRLinuxオリジナルCLI
- : ユーザー独自CLI



ライフサイクルマネジメント

- カスタマーアプリケーションは "first class citizen" として管理され 障害時のレジリエンシやシンプルな音ボーディングが提供される

YANG インテグレーション

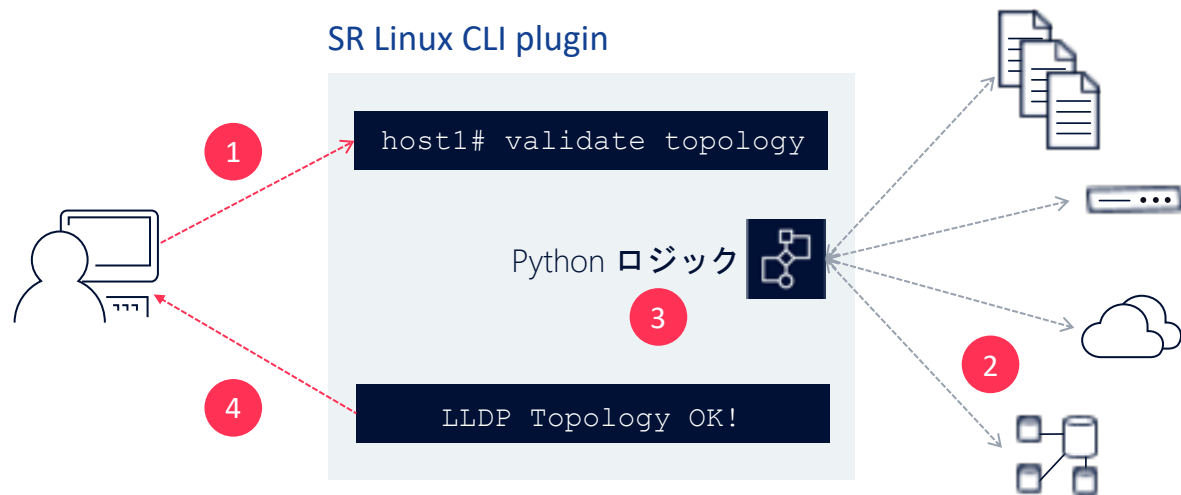
- 完全なYANGインテグレーションにより Native CLI, gNMI, JSON-RPC – config w/persistency, state, stats and show commanが使用可能
- いかなるObject, いかなるInterface

Pythonベース CLI

- オープンソース
- 複数のYANG model/pathからユーザーが必要なOutputを表示されるカスタム showコマンド
- プログラマブル i.e. コンディショナルステートメント

カスタムCLIの例

Topology Validation



1. カスタムCLIコマンドを実行
2. CLIプラグインがトポロジーDBから接続しているべきLLDP Neighborの情報を取得
3. DBから取得したLLDP Neighborの情報とLocalのLLDP Neighbor stateを比較
4. 結果をOutputとして表示

まとめ

- NW機器ベンダーが考えるオープンNOSについてご紹介させて頂きました。
 - ハードウェアへのDependencyを可能な限り少なくし、今後のDisaggregationへ対応
 - 内部・外部アプリケーションへのオープン/スケーラブルなテレメトリ
 - NDK・CLIプラグインでの自由なユーザーアプリケーション開発
 - TypicalなNWアプリケーションは実績のあるソフトウェアコンポーネントを使用可能

NOKIA