



Open Networking Conference Japan 2019

データ中心アプローチはシステムを救う

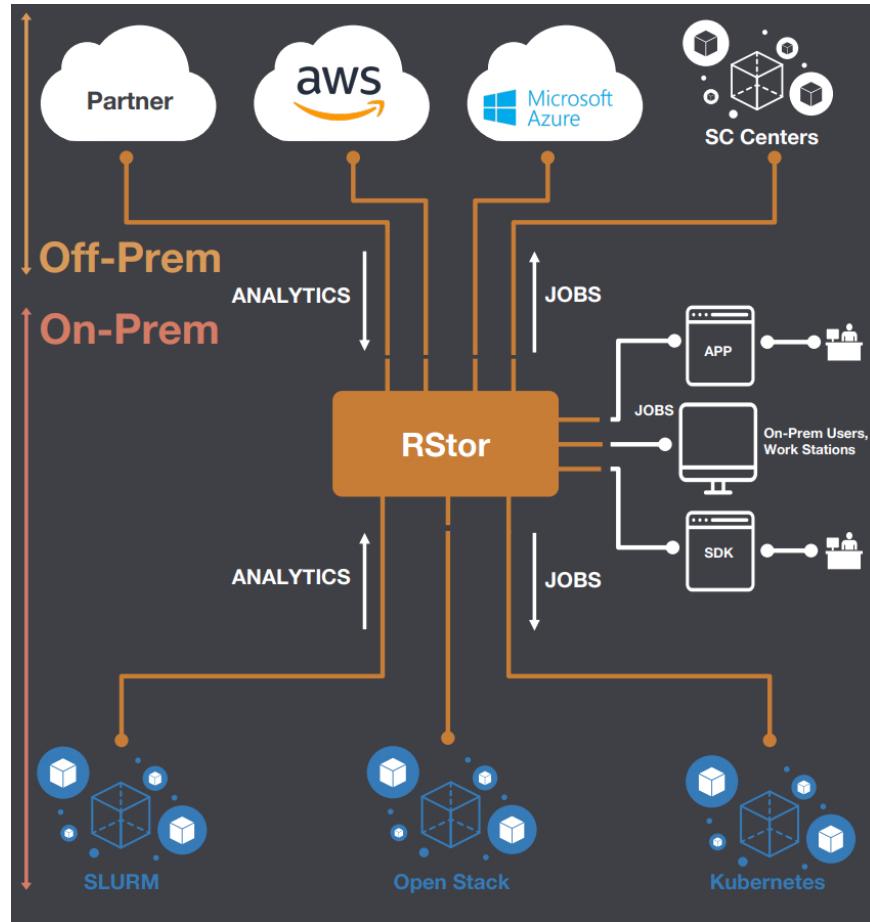
Data Intensive Approach Saves Systems

Miya Kohno (mkohno@cisco.com)

1 November 2019

ネットワークのとてつもない可能性

Network is a Storage !!



Computer Science · Published in IEEE 35th International... 2015 · DOI: 10.1109/icdcs.2015.74

RStore: A Direct-Access DRAM-based Data Store

Animesh Trivedi, Patrick Stuedi, +3 authors Thomas R. Gross

Distributed DRAM stores have become an attractive option for providing fast data accesses to analytics applications. To accelerate the performance of these stores, researchers have proposed using RDMA technology. RDMA offers high bandwidth and low latency data access by carefully separating resource setup from IO operations, and making IO operations fast by using rich network semantics and offloading. Despite recent interest, leveraging the full potential of RDMA in a distributed environment remains a challenging task. In this paper, we present RDMA Store or RStore, a DRAM-based data store that delivers high performance by extending RDMA's separation philosophy to a distributed setting. RStore achieves high aggregate bandwidth (705 Gb/s) and close-to-hardware latency on our 12-machine testbed. We developed a distributed graph processing framework and a Key-Value sorter using RStore's unique memory-like API. The graph processing framework, which relies on RStore for low-latency graph access, outperforms state-of-the-art systems by margins of 2.6 – 4.2× when calculating Page Rank. The Key-Value sorter can sort 256 GB of data in 31.7 sec, which is 8× better than Hadoop TeraSort in a similar setting. [LESS](#)

[VIEW ON IEEE](#)

[SAVE TO LIBRARY](#)

[CREATE ALERT](#)

[CITE](#)

- 超高速分散DRAMストアにより、RDMAを分散環境で実行

<https://www.semanticscholar.org/paper/RStore%3A-A-Direct-Access-DRAM-based-Data-Store-Trivedi-Stuedi/805b7d313543543a0a45b9647db3469d1be3f167>

<https://www.datacenterknowledge.com/design/cisco-backed-rstor-aims-change-data-center-architecture-we-know-it>

ネットワークのとてつもない可能性

Network is a Camera !!

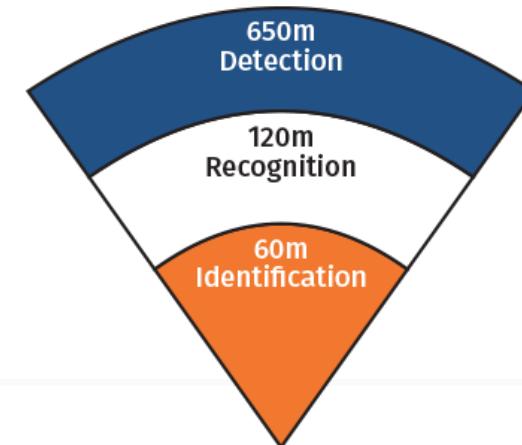
Technology that see beyond - 見えないものを見る

The heart of Aqueti imaging systems is our breakthrough parallel supercamera technology, integrating our innovative optical, electronic and software architectures. Conventional digital cameras form a digital image by capturing the analog image created by a single lens. One lens, one image. Our cameras form images by digitally combining data from an array of micro-cameras. These "images" become interactive digital streams that our technology allows multiple users to manipulate. You can control the spatial resolution, field-of-view and time window to meet your application needs.

Our processing architecture enables clients to manage the data and integrate with other systems for storage, analytics, video streaming and more. We power these complex operations with our Aqueti Camera Operating System (ACOS) and application programming interface, the Mantis API. Working together, this software and operating environment take Aqueti's advanced imaging technology to the next level, allowing users to give the resulting media purpose and meaning.



Mantis 70



- 60m先を識別
- 120m先を認識
- 650m先を検出

<https://www.aqueti.com/technology/>

ネットワークのとてつもない可能性

Network is a Sensor !!



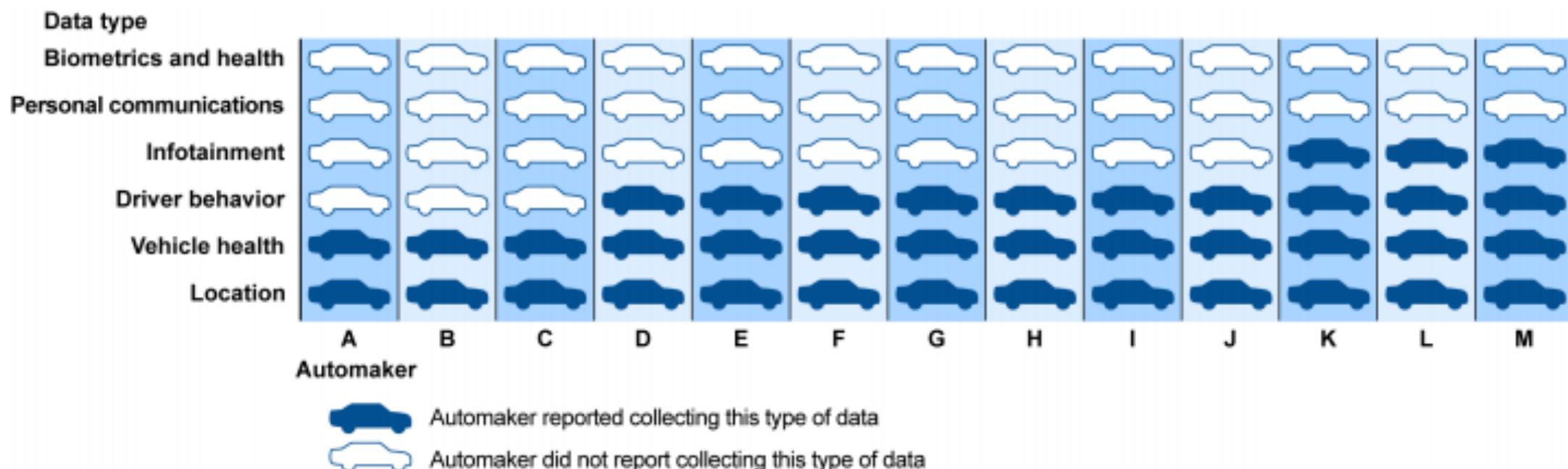
- 人間の注意力には限界がある
- しかし、生体は、人間が知覚していないなくても、様々な反応をしている（瞳孔、呼気、に含まれるCO₂の量、などなど）
- それらの反応をセンサーで捕捉し、伝達したりフィードバックする技術により、人間同士の関係を良くしたり、ひいては人間の機能を向上したりできるのではないか。（Neuroprasticity!）

<https://www.poppycrum.com/speaking>

車両から収集するデータの種別

米国監査院(US GAO)による自動車メーカー13社に対する調査

1. 位置情報
 2. 車両の各種状態
 3. 運転者の振る舞い



Source: GAO analysis of interview responses. | GAO-17-656

「データ中心」に考える

- ・収集するのは誰?
　車両運用管理サービス企業, 医療サービス団体, 公共団体, その他…
- ・何のためのデータ?
　安全のため、最適化のため、高付加価値化のため…
- ・どうやって収集する?
　通信手段、通信頻度、必要通信特性（リアルタイム性..）
- ・どこで収集し、どこで蓄積し、どこで分析する?
　Cloud, Enterprise DC, Regional DC, Edge Computing, …

車両から収集するデータの特性

Application	V2X	Fleet operator	3 rd Party	Public Operator	Safety or Value-added	Hi/lo Data volume	Time-sensitive (*not low latency)	Frequency of exchange	Primary comms path	Secondary comms path
ITS Safety	V2V, V2I	No	No	Yes	Safety	Low	Yes	High	DSRC/ C-V2X	
Predictive Health Maintenance	V2C	Yes	No	No	Value-added	Low	No	Periodic	Wifi	Cellular
Navigation Services	V2C	No	Yes	No	Value-added	High	Yes	Periodic	Cellular	Wifi
Advanced Driver Assistance Services	V2C	No	Yes	No	Value-added	High	No	Periodic	Wifi	Cellular
Real-time logistics planning	V2C	Yes	Yes	No	Value-added	High	Yes	Streamed	Cellular	Wifi

データ中心アプローチとは – Application Systems

- ・ アプリケーションは、「演算指向(Compute-Intensive)」ではなく「データ指向(Data Intensive)」であることにより、信頼性、スケーラビリティ、保守性を獲得した。 [*]

「データシステム」としての抽象化

信頼性

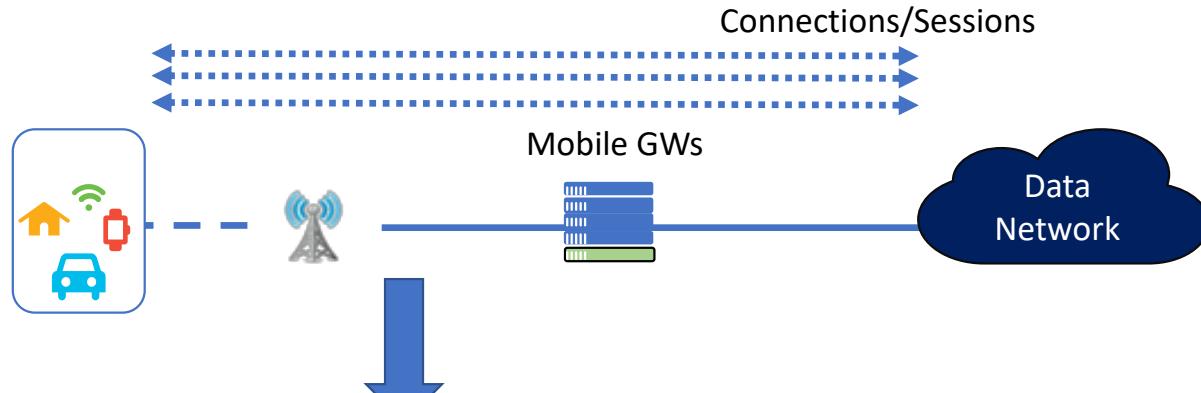
スケーラビリティ

保守性

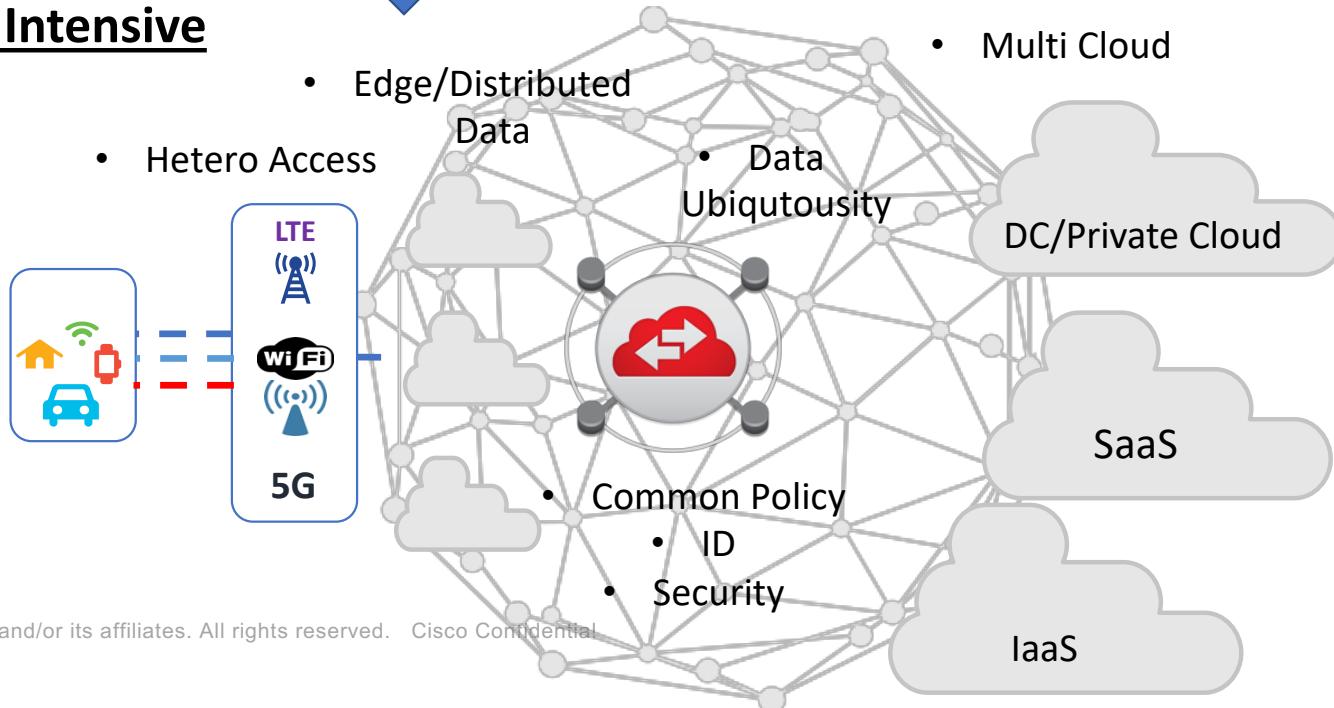
- ・ 何か問題が発生しても正しく動作し続ける
- ・ ハードウェア・ソフトウェア・ヒューマンエラーに対する 耐障害性
- ・ 負荷とパフォーマンスの計測
- ・ 負荷の増大に対してシステムが対応できる
- ・ 運用性
- ・ シンプル性
- ・ 進化への対応

データ中心アプローチとは – Network Systems

Connection Intensive



Data Intensive



- Connection Intensive

エンドポイントを規定し、
そのエンドポイントに対し
接続パスを提供する

- Data Intensive

データの収集・配信、およびその
処理・活用とフィードバック、
そしてセキュリティ保護の観点から
アーキテクチャを検討する

→ データは遍在する
マルチクラウド
分散・エッジクラウド

→ 通信手段に依存しない
ヘテロアクセス

→ Scale Free

データ中心アプローチがシステムを救う?!

「データシステム」としての抽象化

信頼性

- ・何か問題が発生しても正しく動作し続ける
- ・ハードウェア・ソフトウェア・ヒューマンエラー・**通信エラー**に対する耐障害性

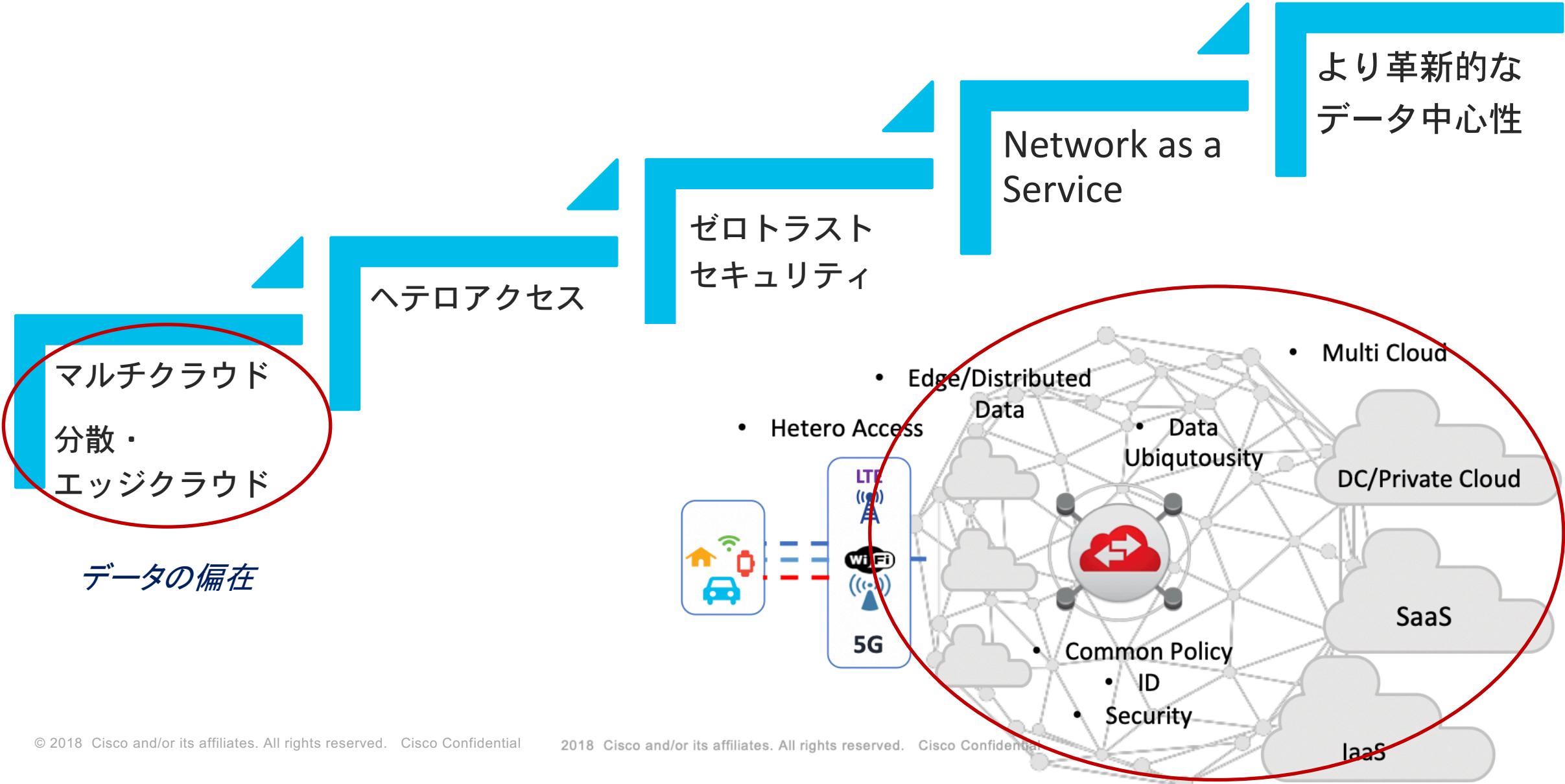
スケーラビリティ

- ・スケーリングネックになる
 - コネクション
 - セッション
 - トンネル
 - ステートを極力排除する
- ・負荷の増大に対してシステムが対応できる

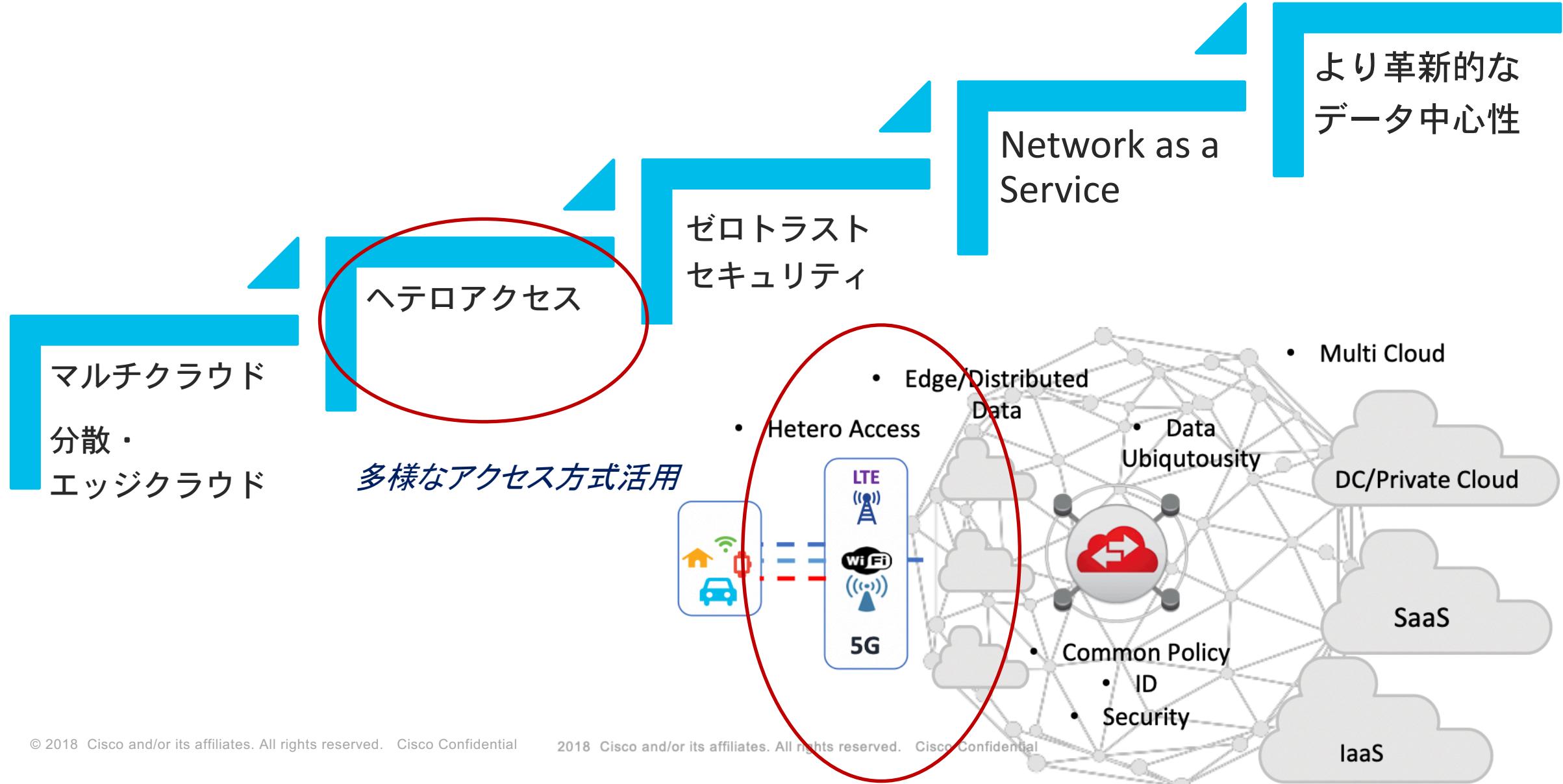
保守性

- ・運用性
- ・シンプル性
- ・進化への対応

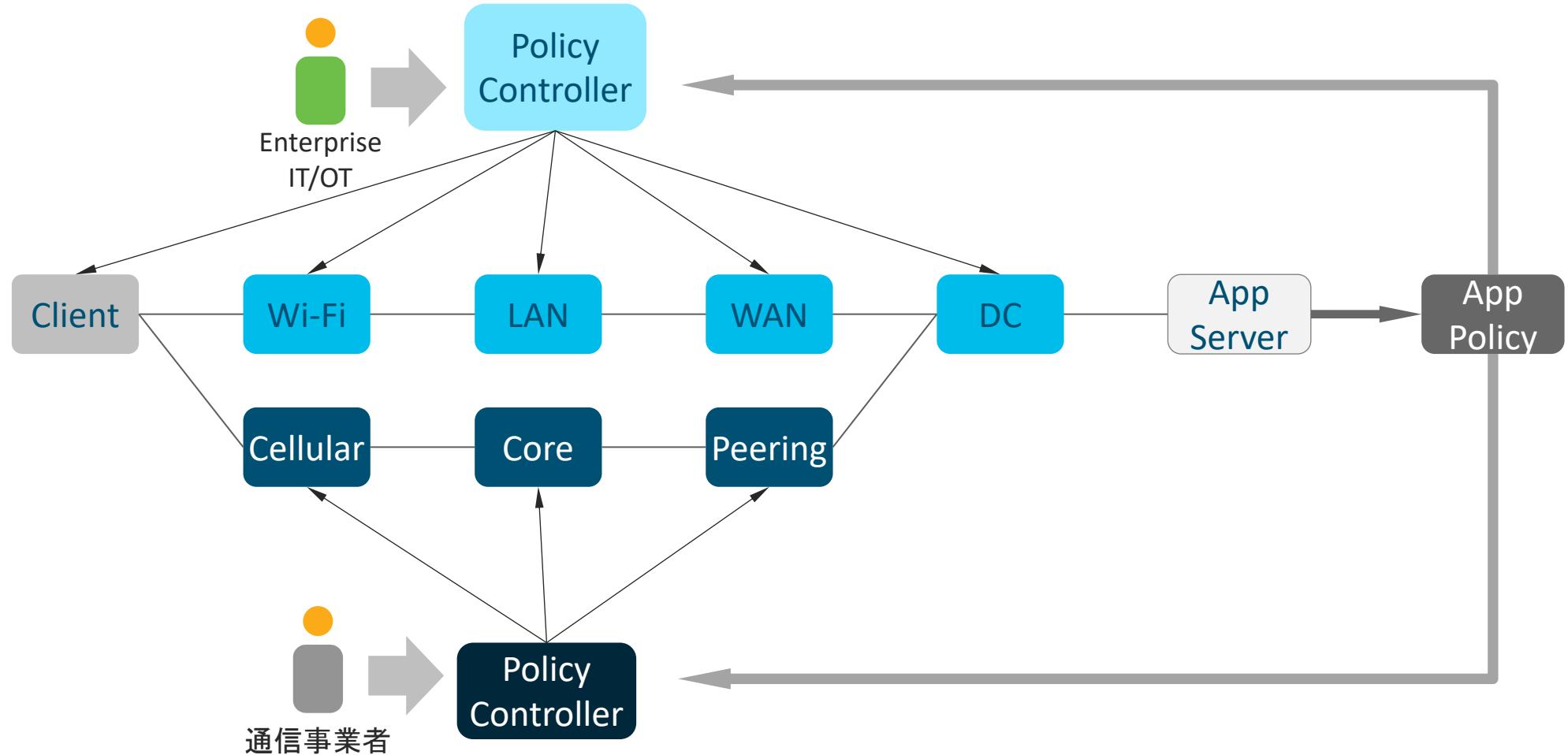
データ中心アーキテクチャへの道筋



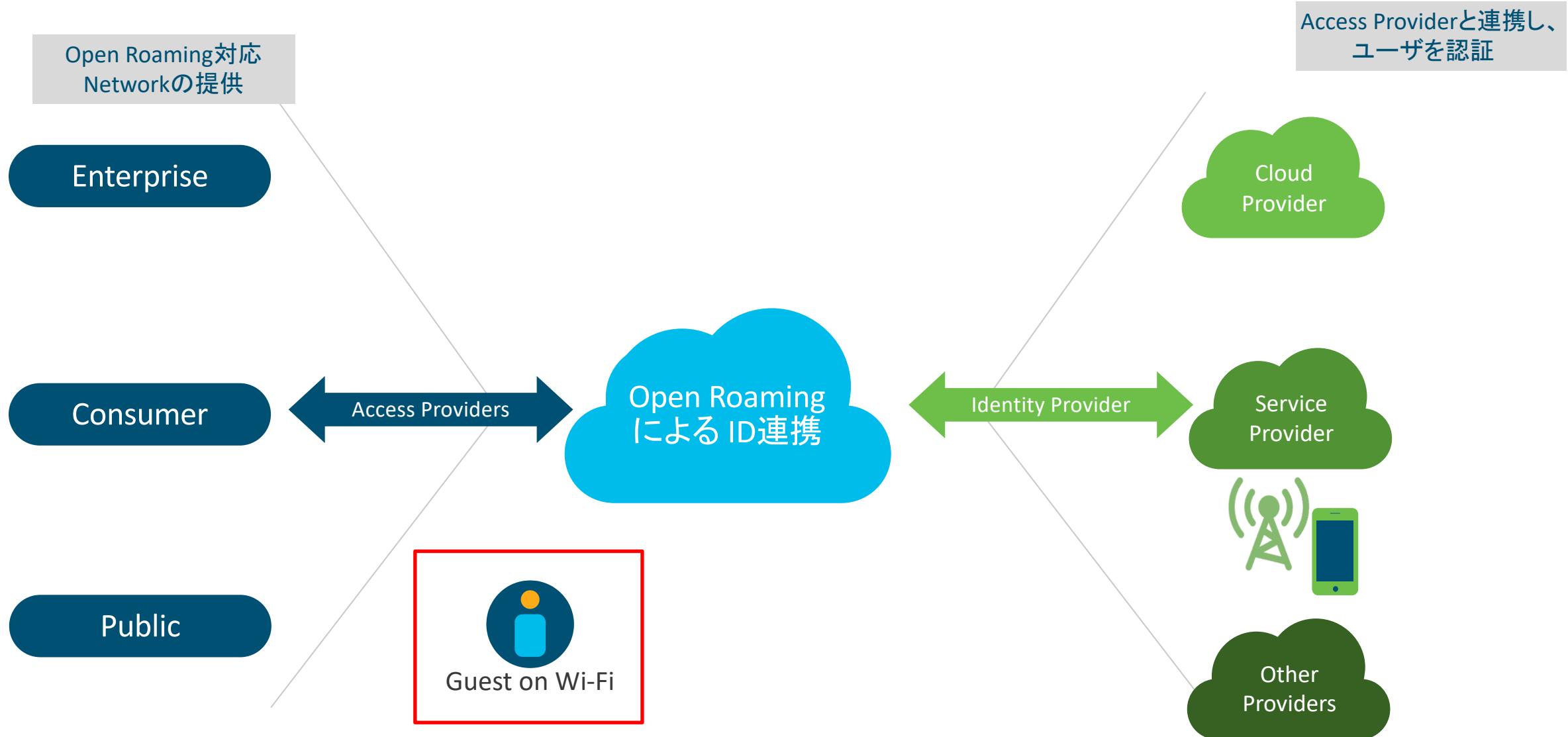
データ中心アーキテクチャへの道筋



ヘテロアクセスのためのpolicy framework



Open Roaming



Open Roamingによる、ヘテロアクセスの利便性向上

Identity Providers



- ・ プラットフォーマーとしての、ユーザへの利便性提供
- ・ アクセスプロバイダとの連携による、フットプリントの拡大

More engagement

Access Providers



- ・ 「データ」のオーナーシップを取り戻す
- ・ 顧客への、良いQoEの提供
- ・ ロケーション分析の提供
- ・ インフラの価値提供(対コモディティ化)

Better experience

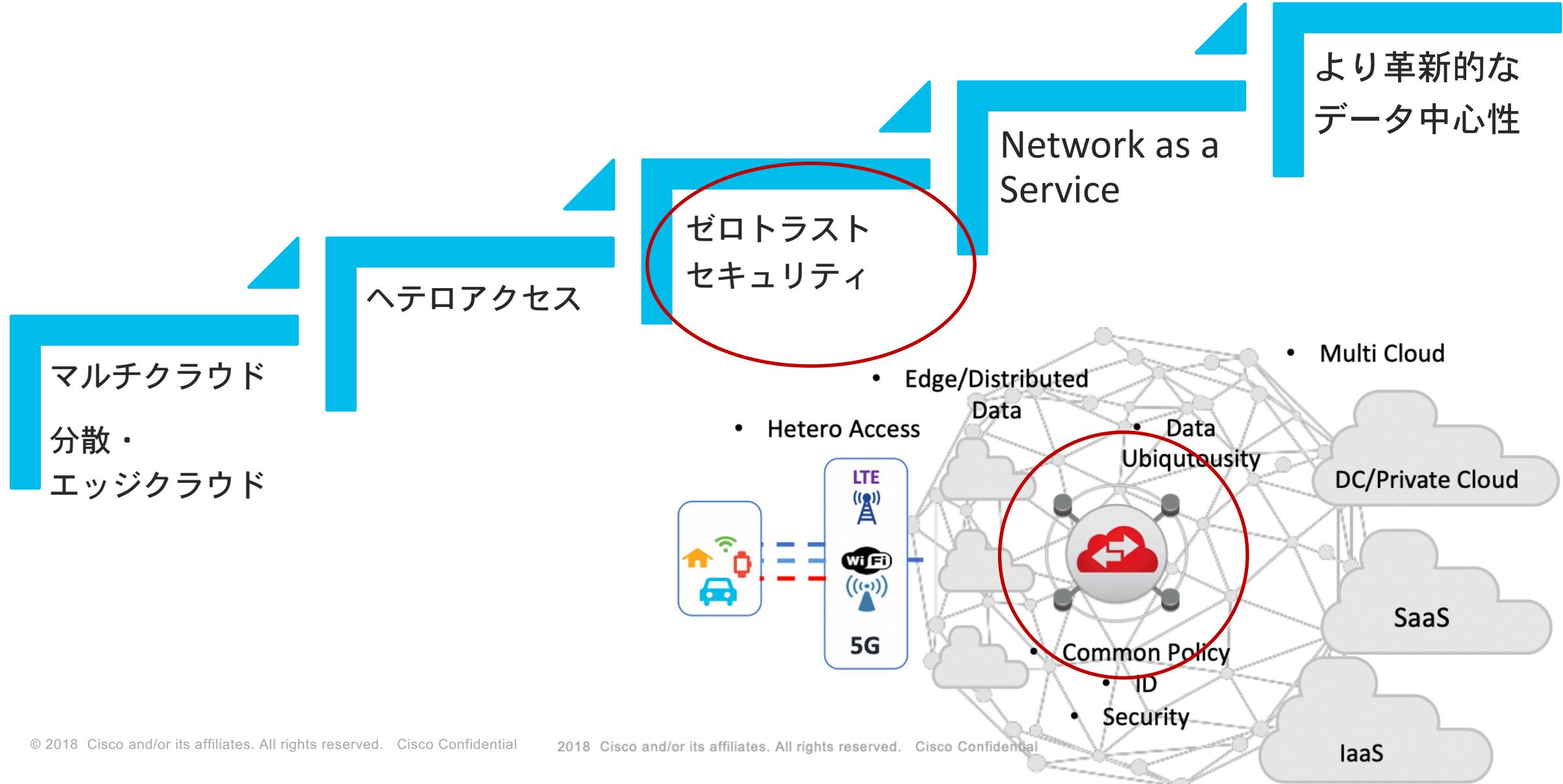
End Users



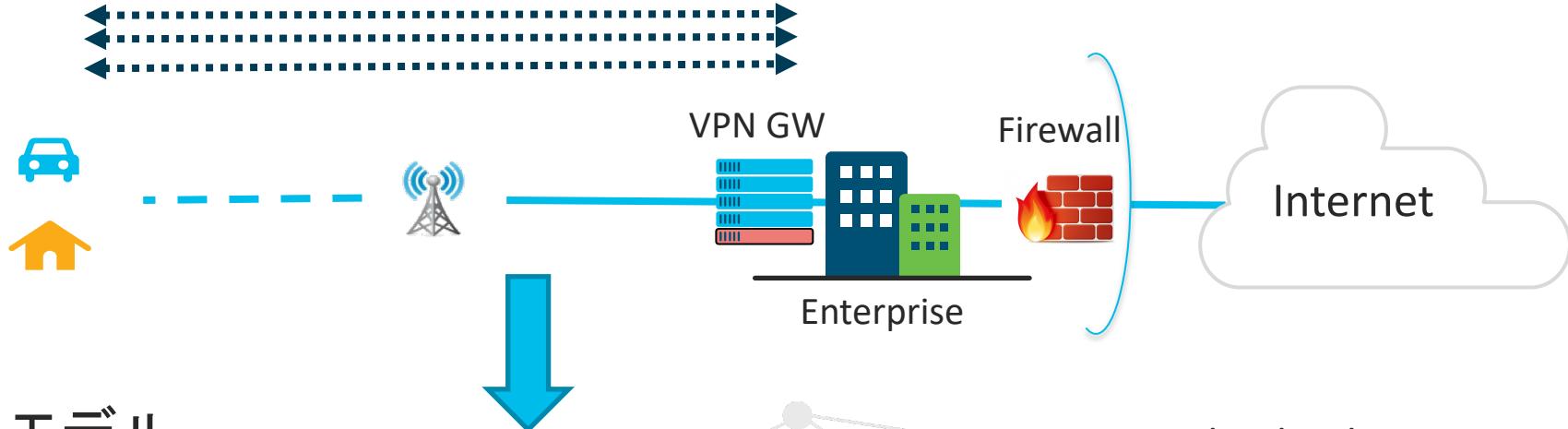
- ・ モバイルエクスペリエンスの向上
- ・ 複数のRadio技術(Public/Private, Cellular/Wifi, etc..) のシームレスな活用

Seamless mobility

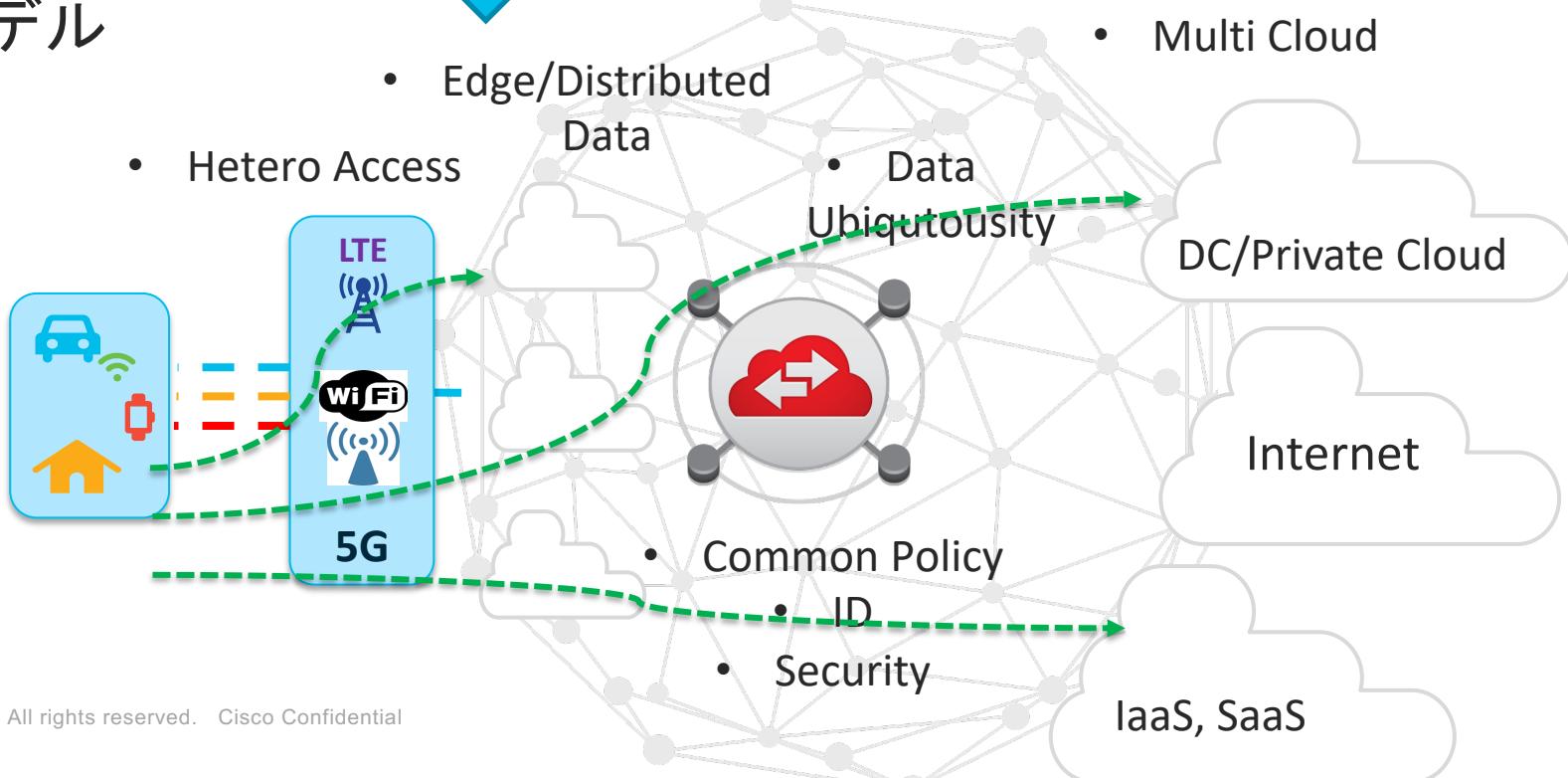
データ中心アーキテクチャへの道筋



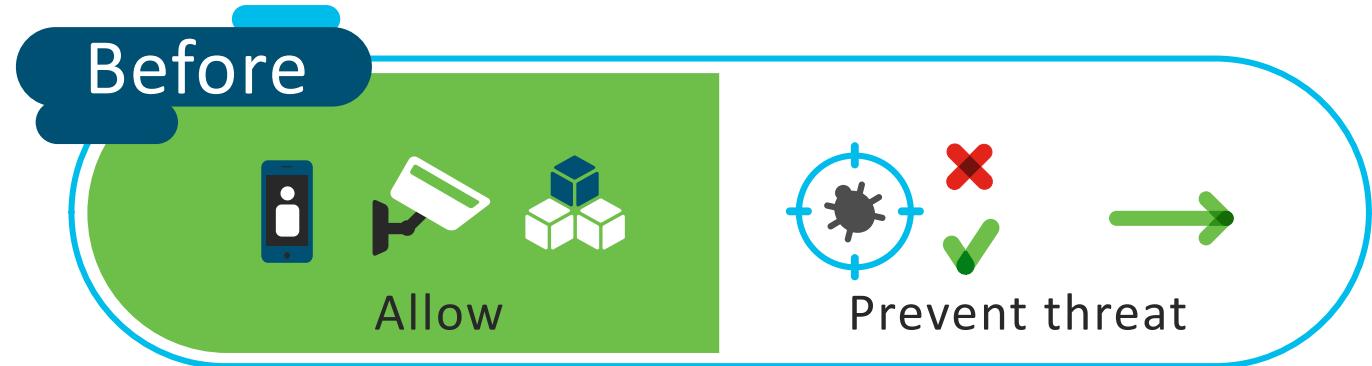
これまでのセキュリティモデル(境界防御型)



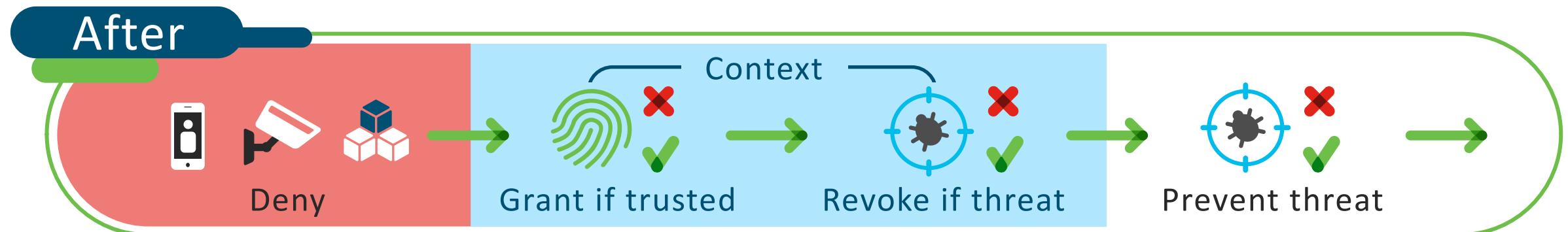
ゼロ・トラストモデル



ゼロ・トラストモデル



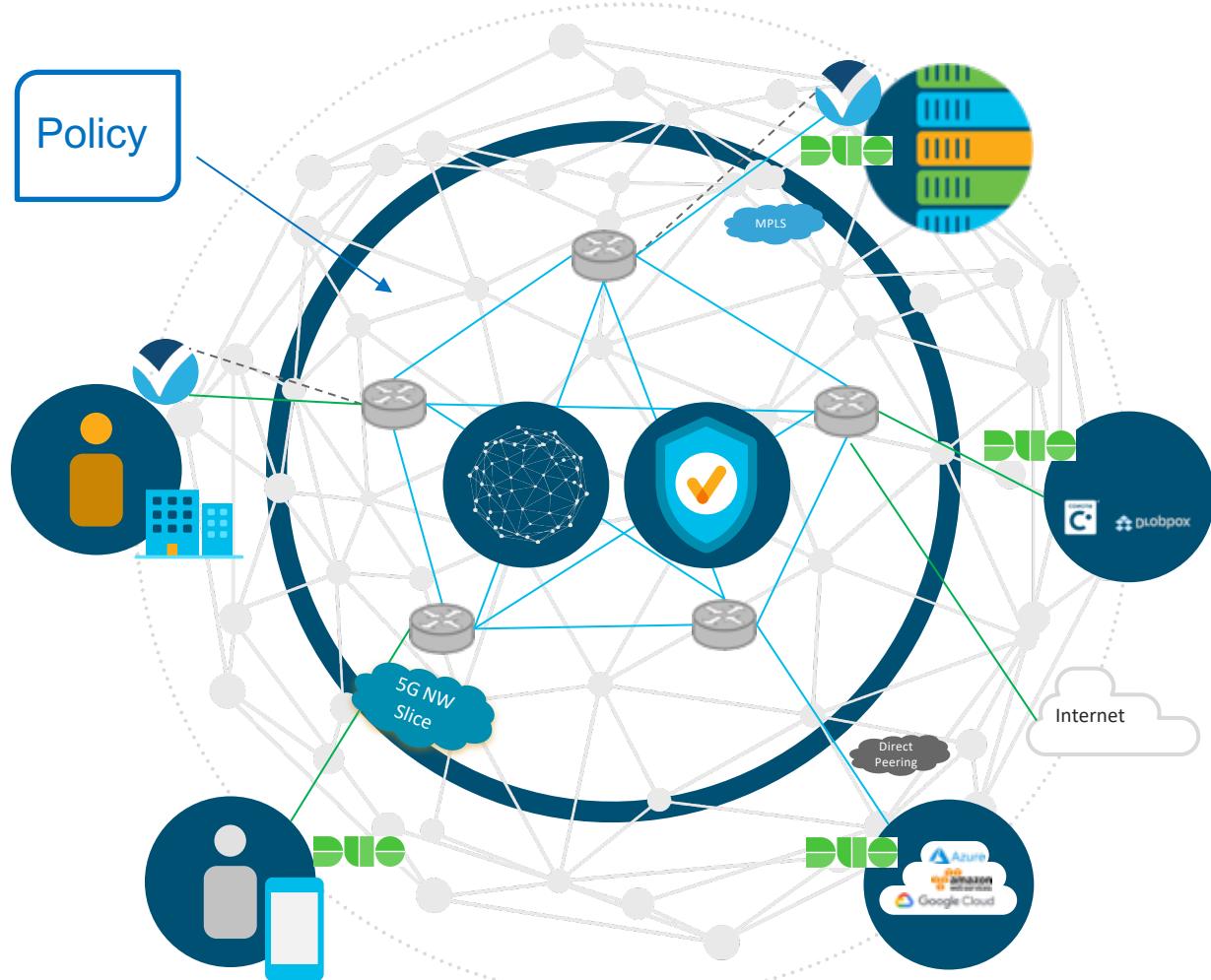
[https://www.oreilly.co.jp/
books/9784873118888/](https://www.oreilly.co.jp/books/9784873118888/)



Network Agent

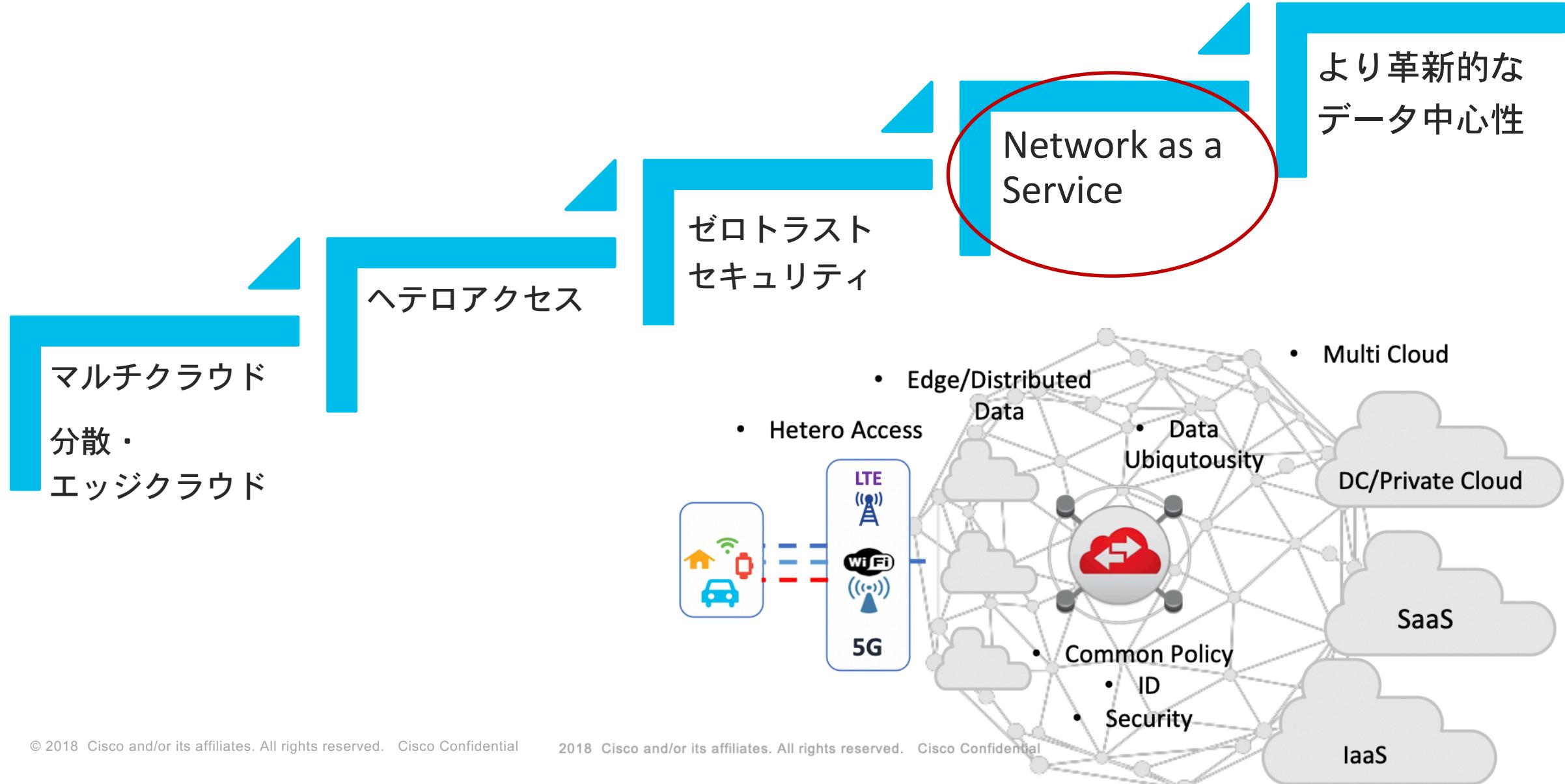
Trust engine & Data Store

Secure Internet Gateway



- 迅速なサービス提供、リソース弹性
- 強い信頼性とアイデンティティ管理
- VPNなしのリモートアクセス

データ中心アーキテクチャへの道筋



Network as a Service

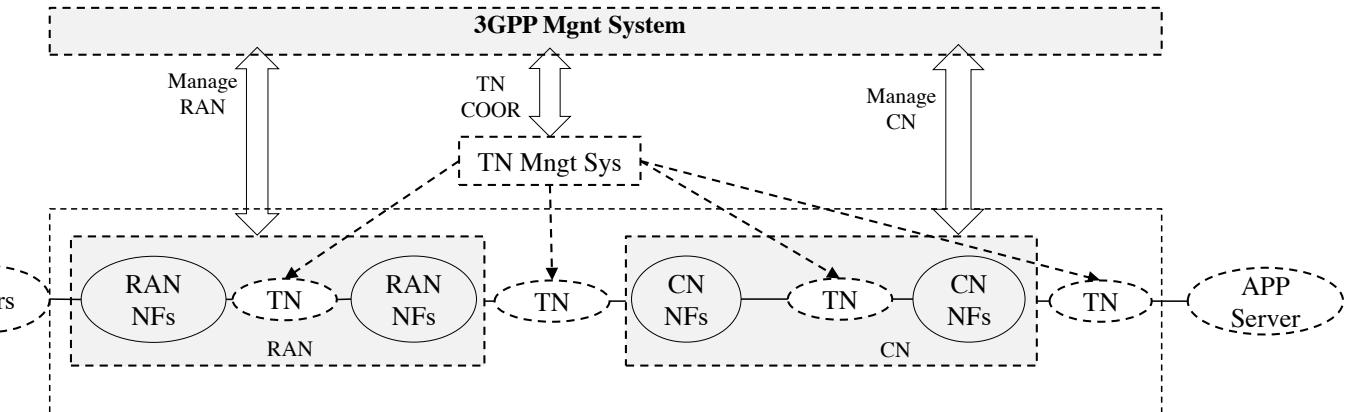
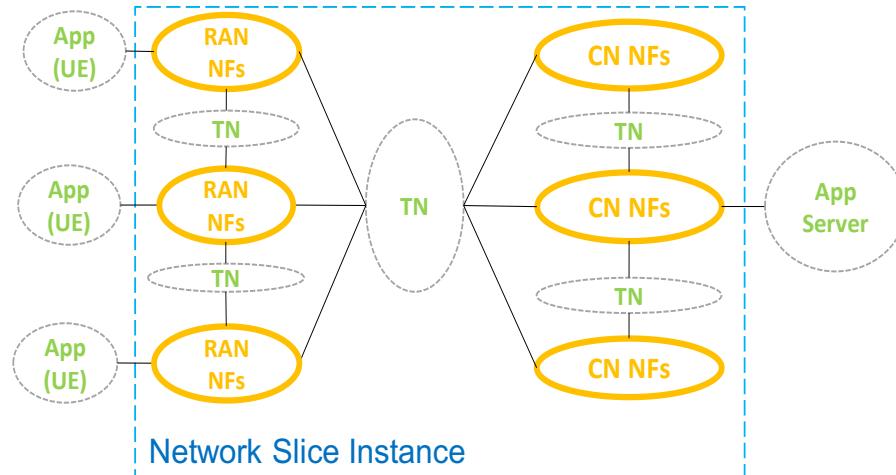
データ中心アーキテクチャに適合するように、より柔軟にネットワークを使えるようにする

- 異なるサービス要請に対応するNetwork Slicingの提供
- Network Slicingを柔軟に使えるようなControl Point/APIの提供
- VPNaas/Mobile SD-WANの提供

Mobility End-to-End Slice

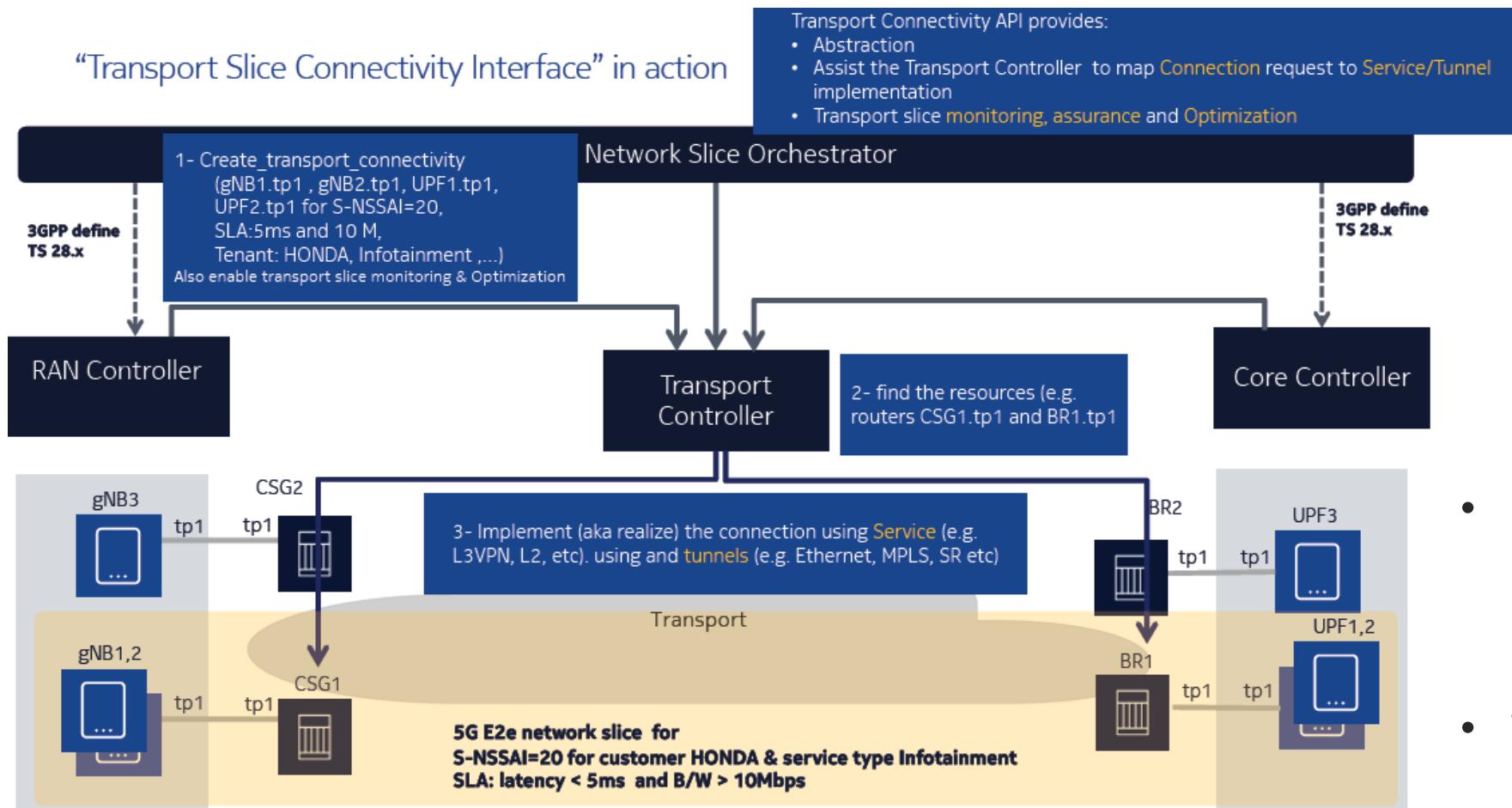


Source : 3GPP TS 28.530



- RAN Slice, CN (Mobile Packet Core Network) Sliceと、TN (IP Infrastructure)をどうマッピングするか
- IP Infrastructureにおけるスライスをどうエンジニアリングするか

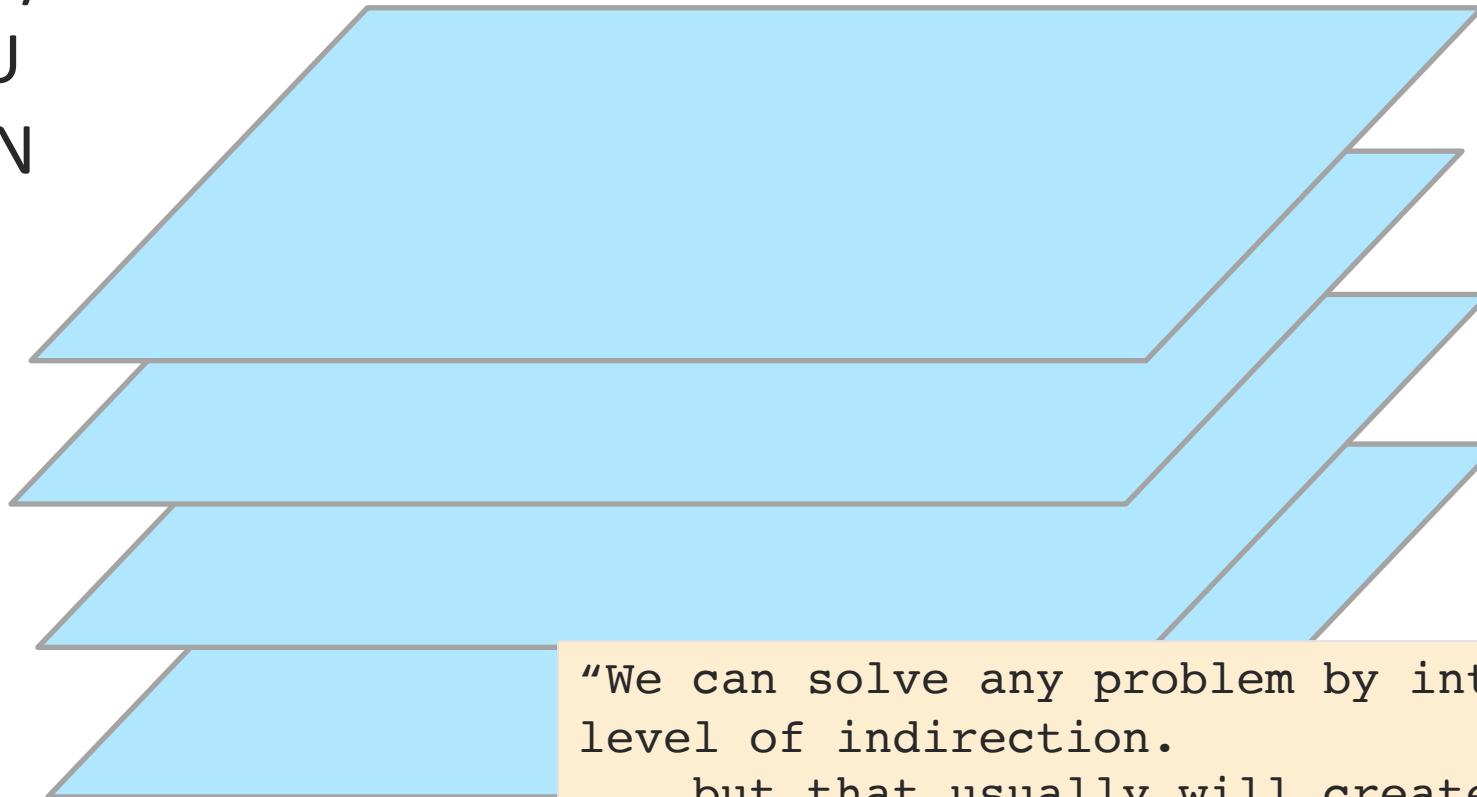
従来の発想で考えれば... draft-roku-5g-transport-slice



- Network Slice Orchestratorが、RAN, TN, CNそれぞれのcontrollerに指令を出す
- TNでは、L2/L3 VPNやTunnelを使ってTN sliceを提供する
- GTP-U tunnelはTNの上でOverlay

しかしレイヤが増えるとコネクションも増える..

- MPLS
- PPPoE/L2TP
- GTP-U
- VXLAN
- NSH
-



"We can solve any problem by introducing an extra level of indirection.
...but that usually will create another problem."

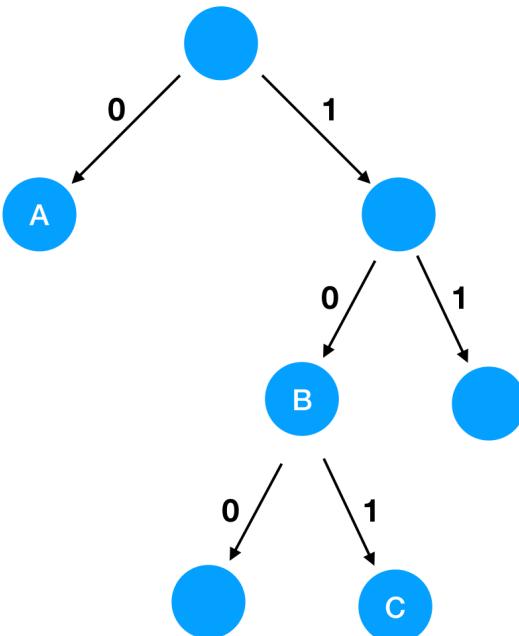
– David J Wheeler

そもそも IP (Internet Protocol) とは

Longest Prefix Match (Not Exact ID Match) !

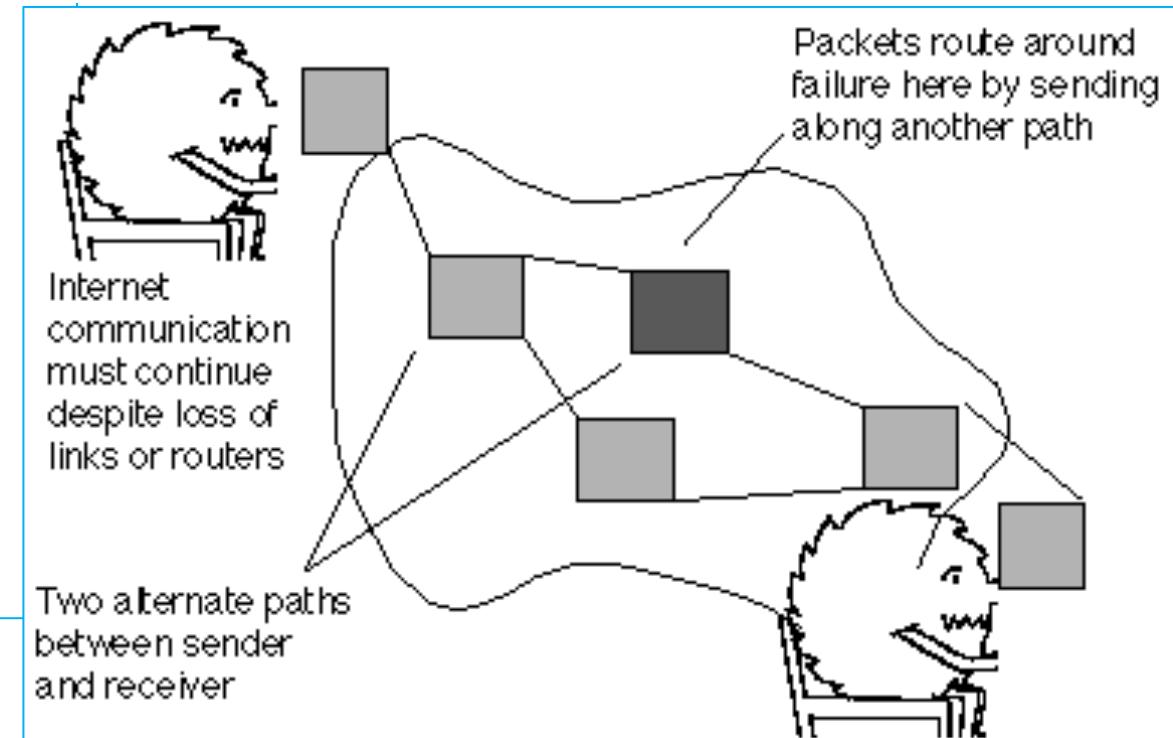
- より Data Intensive
- より Scalable
- ScaleFree!

Prefix	Router
0*	A
10*	B
101*	C



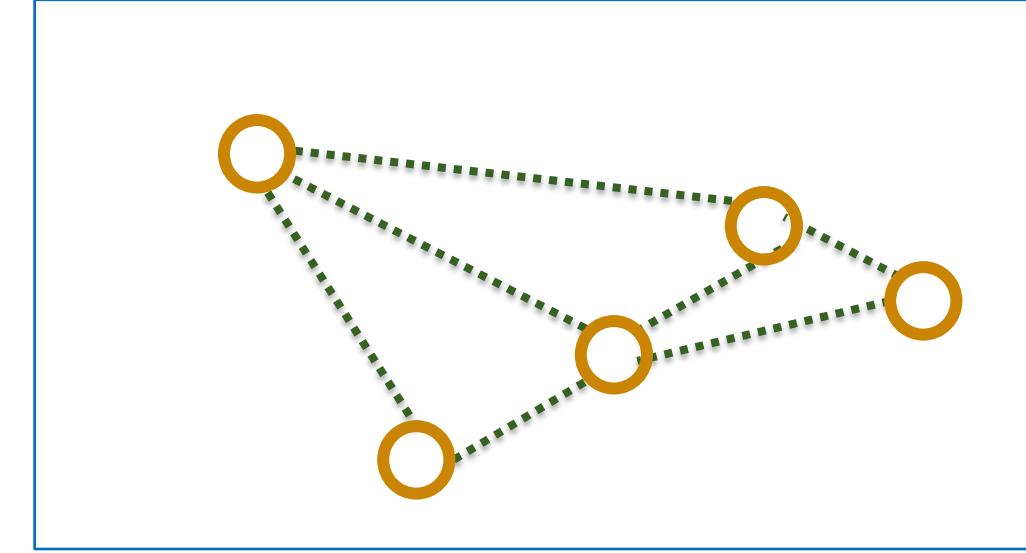
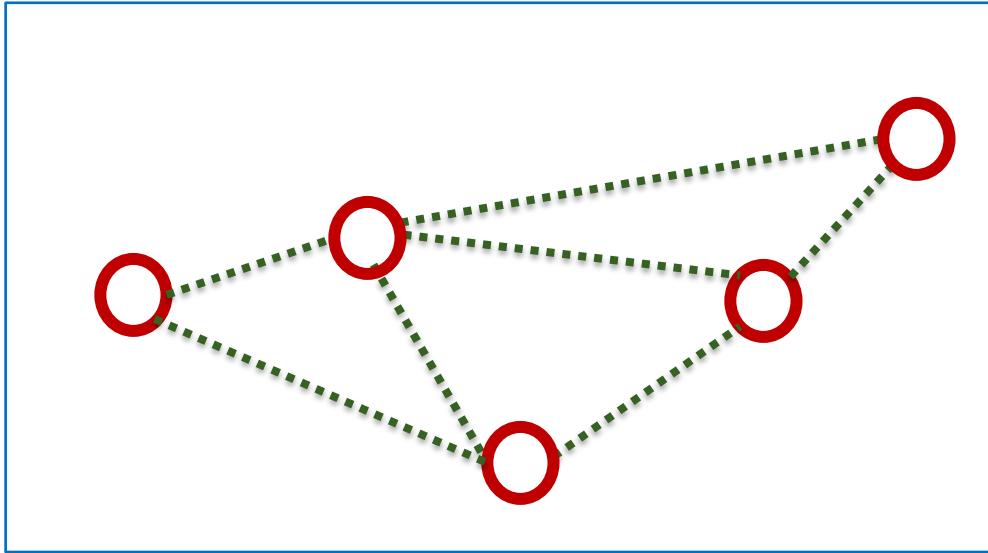
<https://www.lewuathe.com/longest-prefix-match-with-trie-tree.html>

Connection-less/Datagram !!

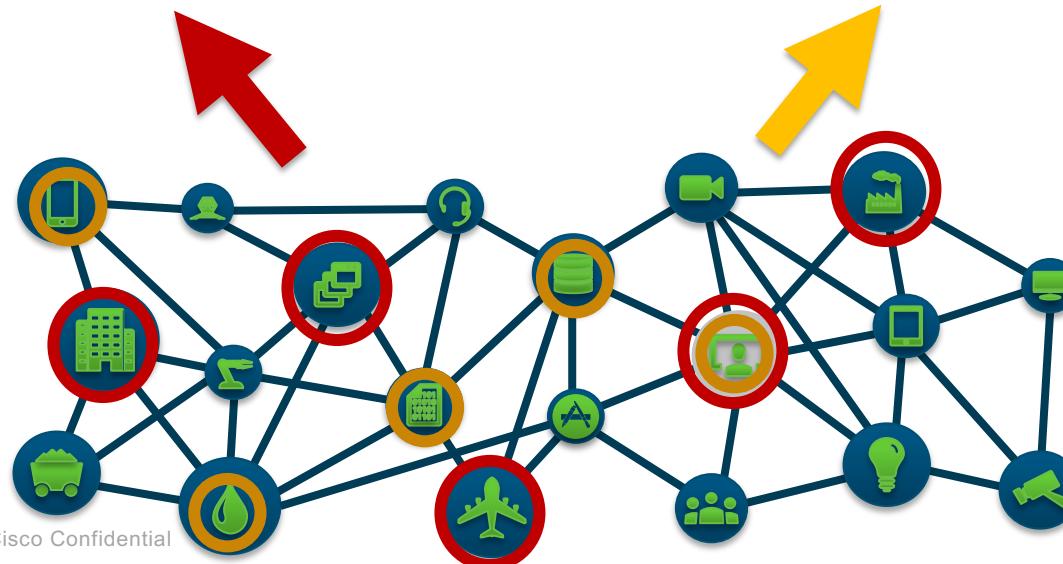


<https://www.erg.abdn.ac.uk/users/gorry/course/intro-pages/datagrams.html>

分離のために無闇にレイヤを使わない



Declarative
Networking



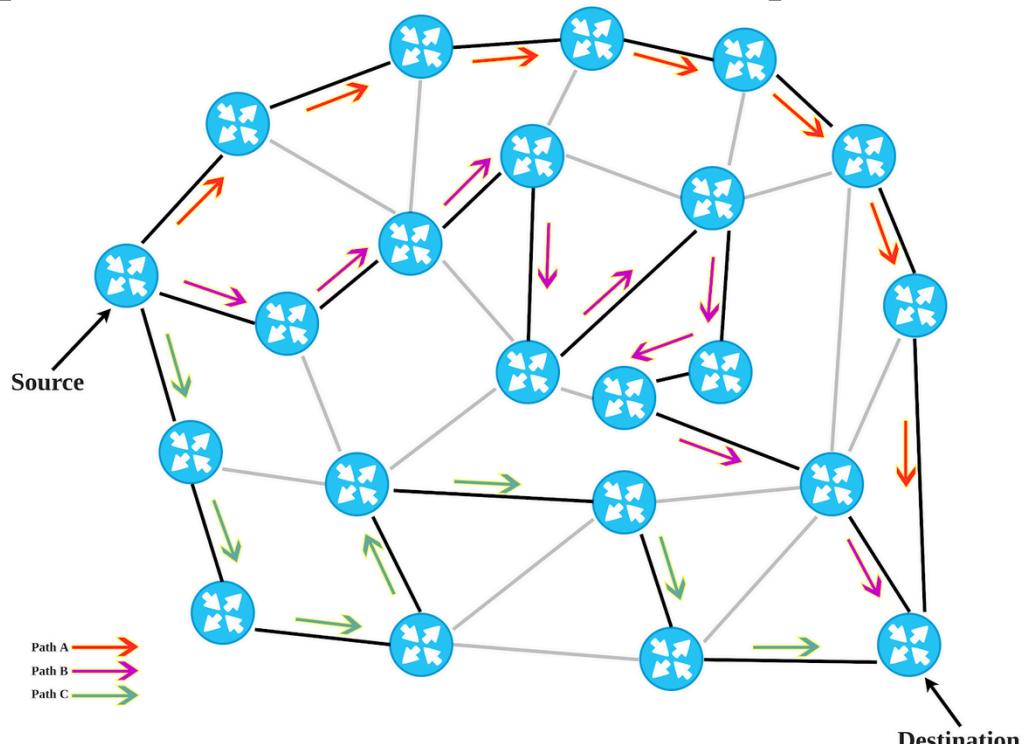
Declare RED !

Declare Yellow !!

Traffic Engineering >> Explicit Route

Traffic Engineeringは、必ずしも“明示的な経由ホップ指定”を必要としない！！！

[誤解を生むかもしれない図]

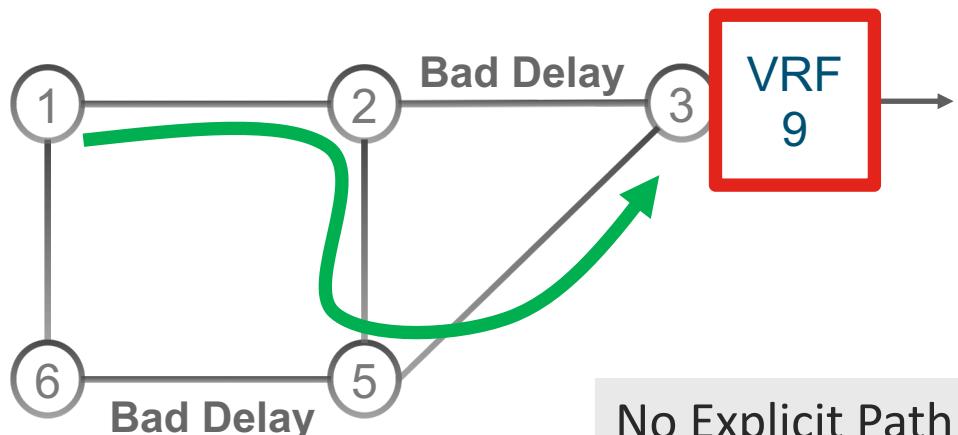


<https://learningnetwork.cisco.com/blogs/vip-perspectives/2018/03/23/introduction-to-segment-routing>
© 2018 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

[実際]

Network Program: D:3:V(9)

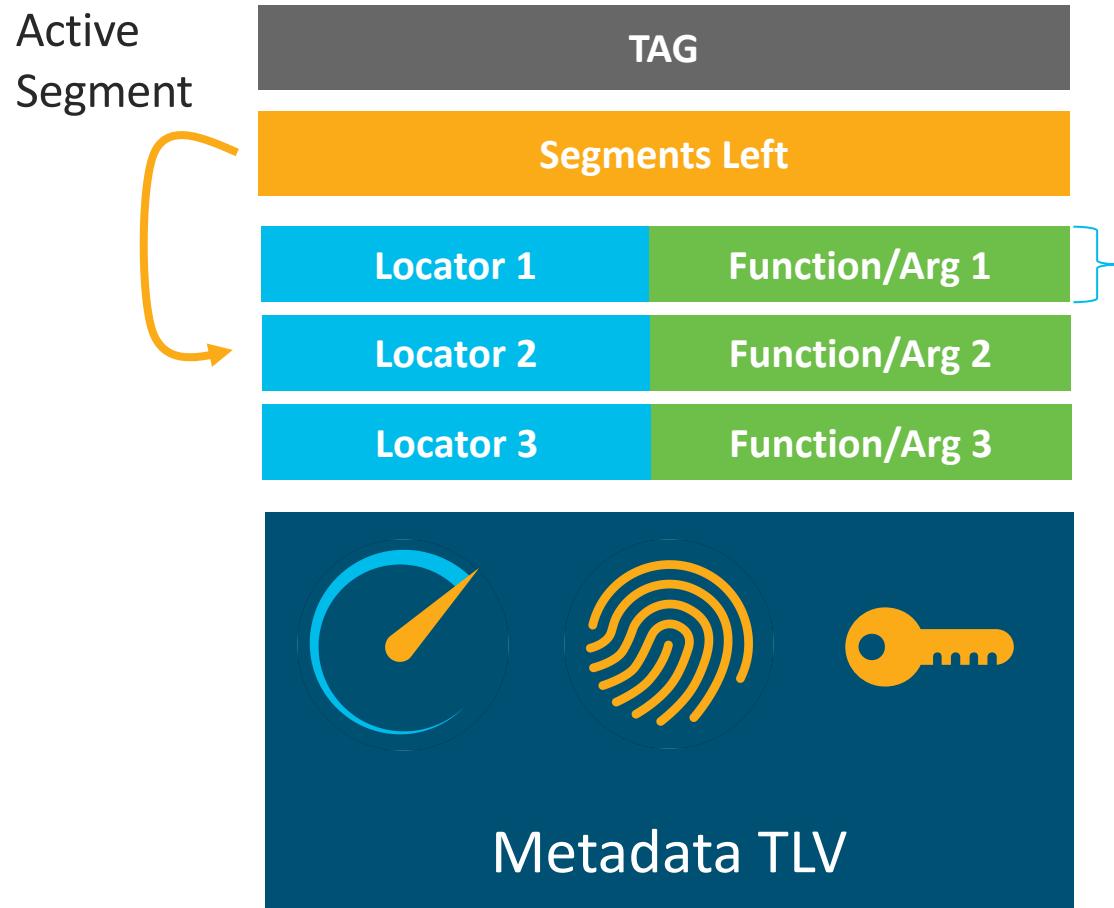
D: locator block is associated with Low Delay Flex-Algo



No Explicit Path!
No SID/SRH!!

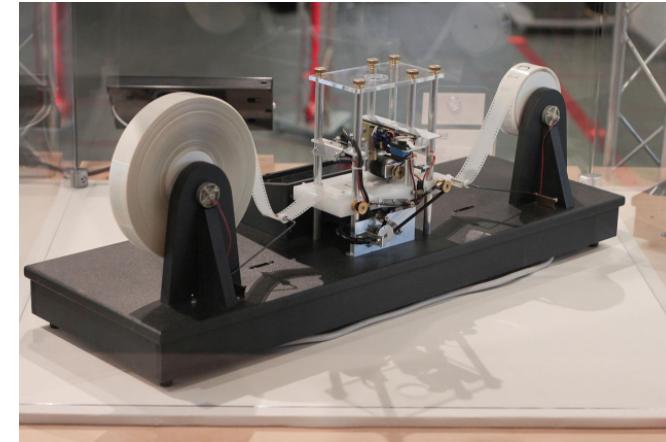
- Network Programming
 - Flex Algo
- 次ページで！

SRv6 Network Programming



IPv6 128 bit
Address space

- Locator
- Function
- Argument



Turing

- Declarative (not imperative) SDN
- Network as a Computer
- In-Network Computing

Flex AlgoによるNetwork Slicing

- 各ノードは、自らが属するFlex-Algo番号を広報する

Nodes 0/9 は Algo 0, 128, 129に属する

Nodes 1/2/3/4はAlgo 0, 128に属する

Nodes 5/6/7/8はAlgo 0, 129に属する

- 各ノードは、Algo毎にprefixを広報する

Node 2は、下記の通りPrefixを広報する

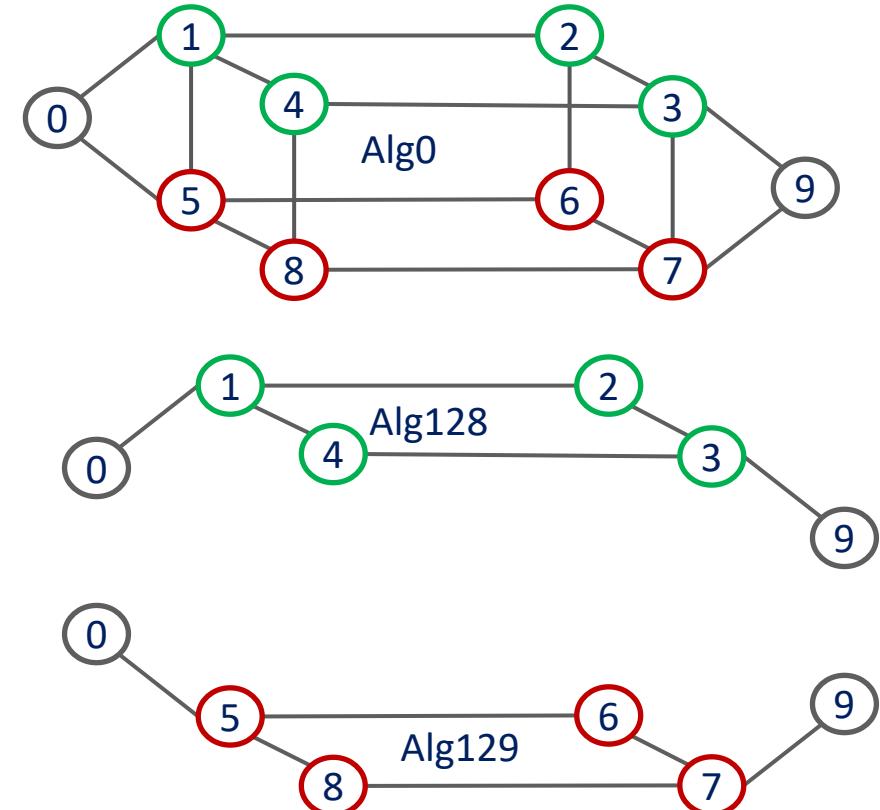
Prefix SID 16002 for Algo 0

Prefix SID 16802 for Algo 128

- ノードNは、Flex-Algo Kについてパス計算を行う

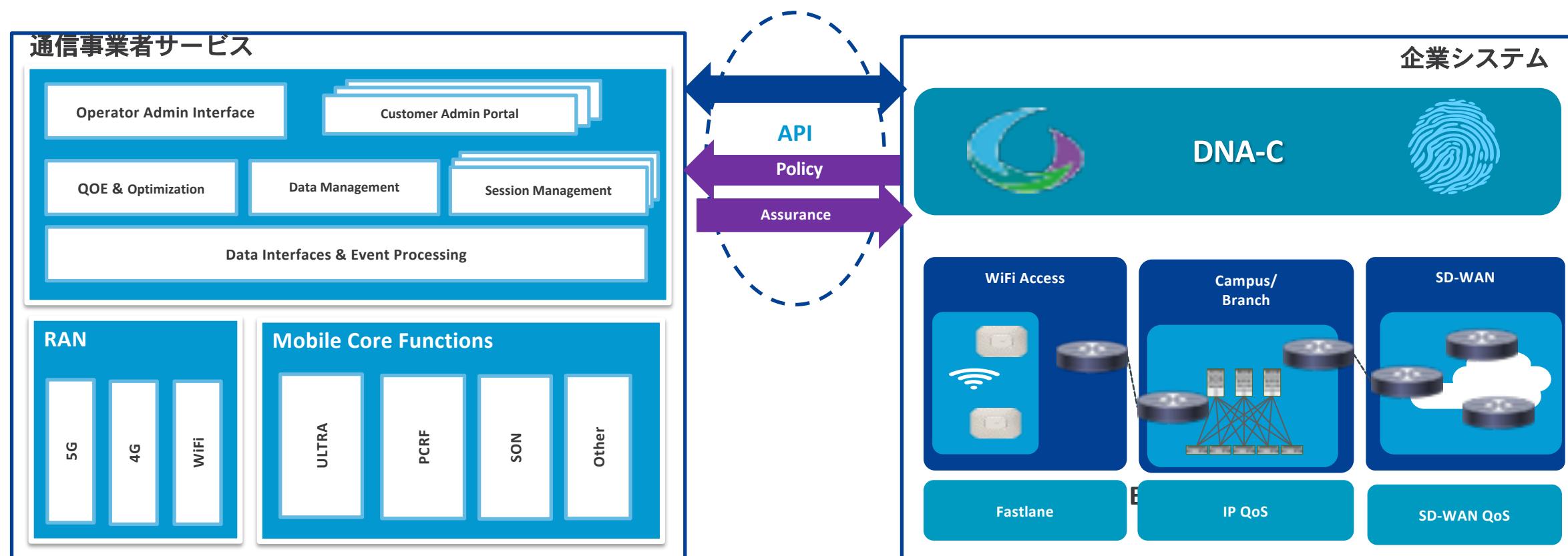
Algo Kにより排除されたリンクを除外し、SPF計算を行う

- 転送テーブルにprefixをインストールする

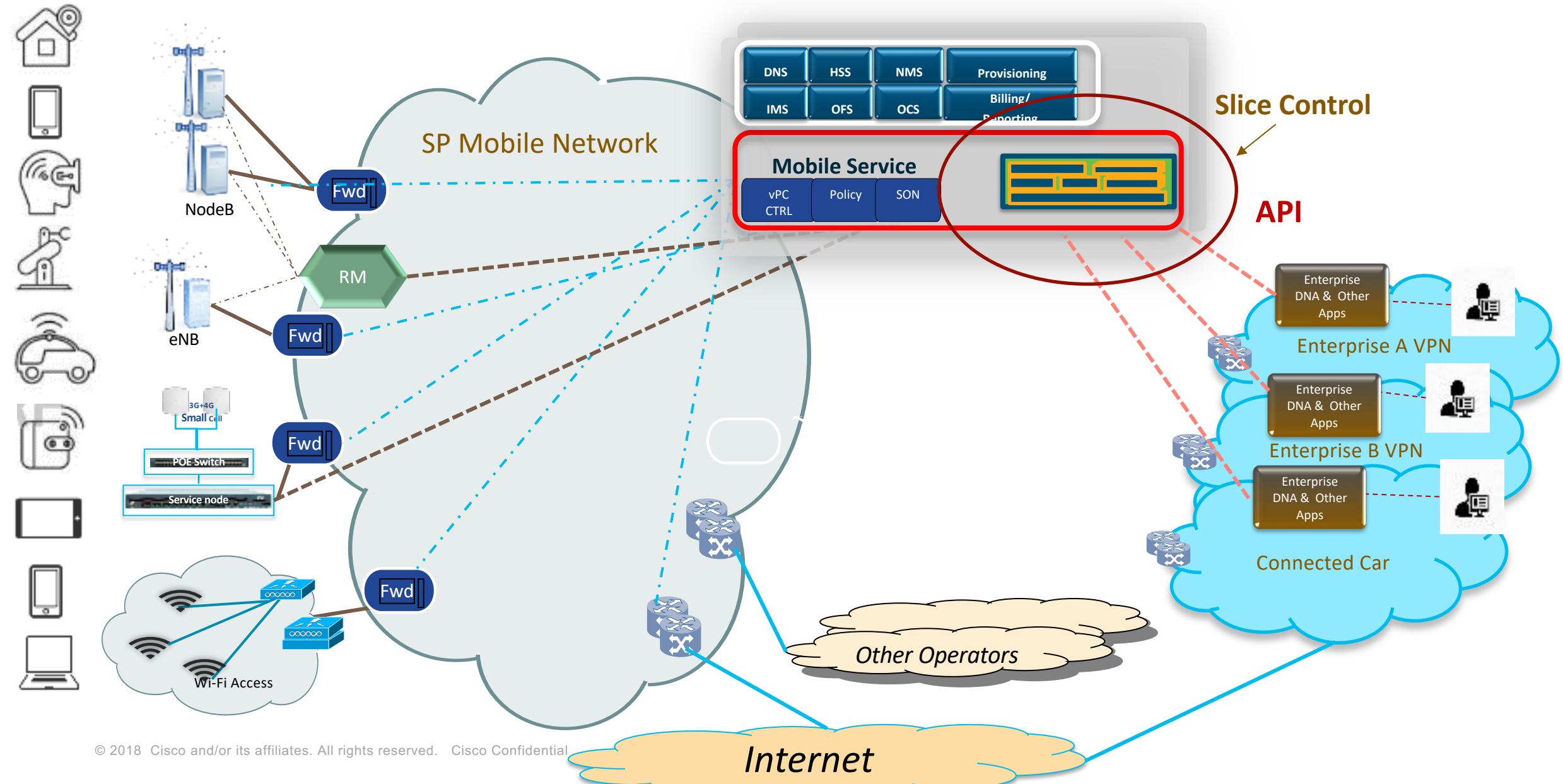


Network Slicingを柔軟に使えるようなControl Point/APIの提供

- ・ 通信事業者のモバイルサービス（スライス）を、企業に対し、API提供する
- ・ 通信事業者のモバイルサービスと企業ネットワーク間で、セキュリティやネットワークポリシーを連携させる



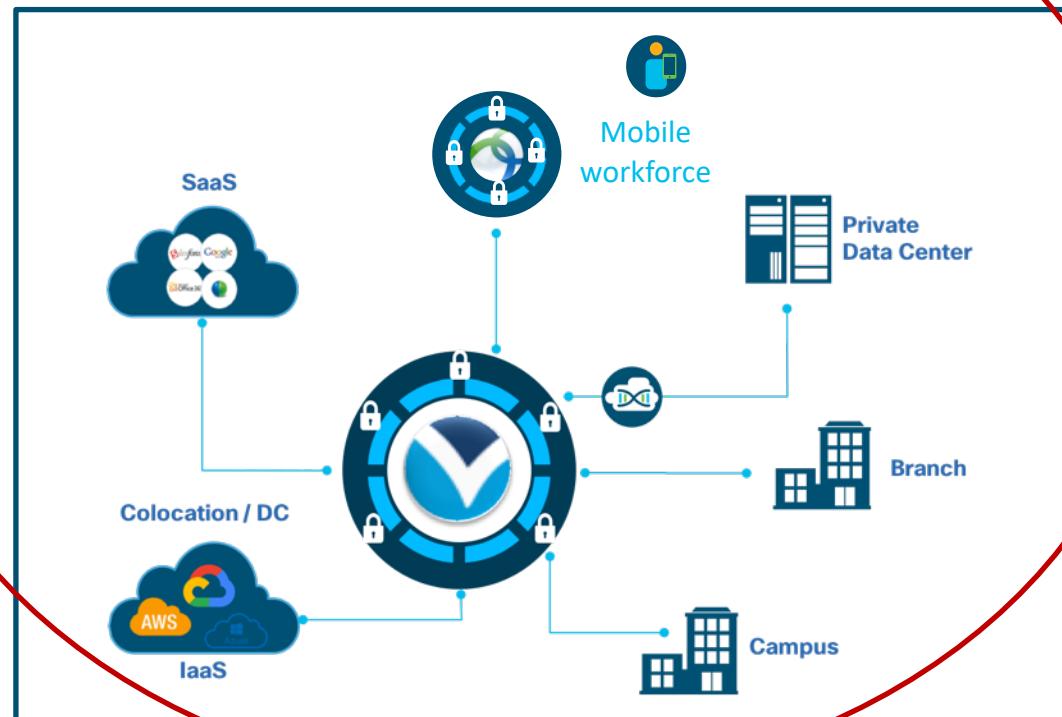
Network Slicingと、そのAPI提供



VPNaas/Mobile SD-WANの提供

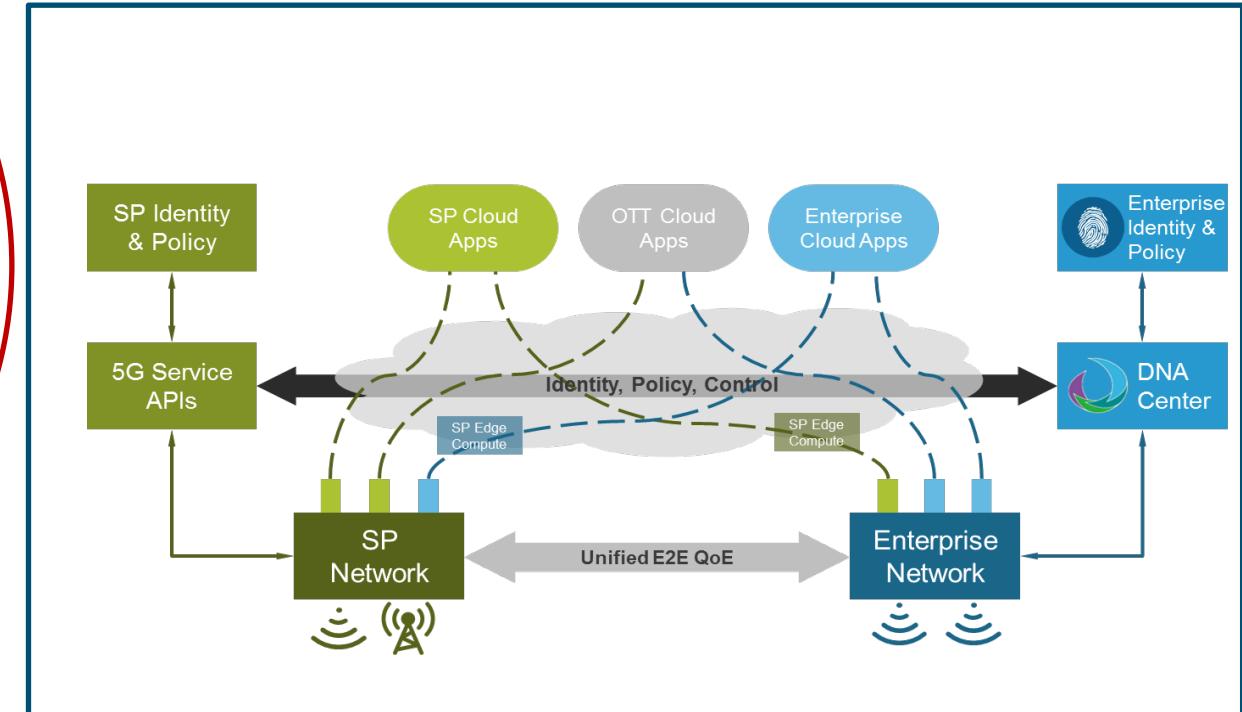
(1) Overlay Approach - Mobile SD-WAN

SD-WANを活用し、Secure Overlayを構成する

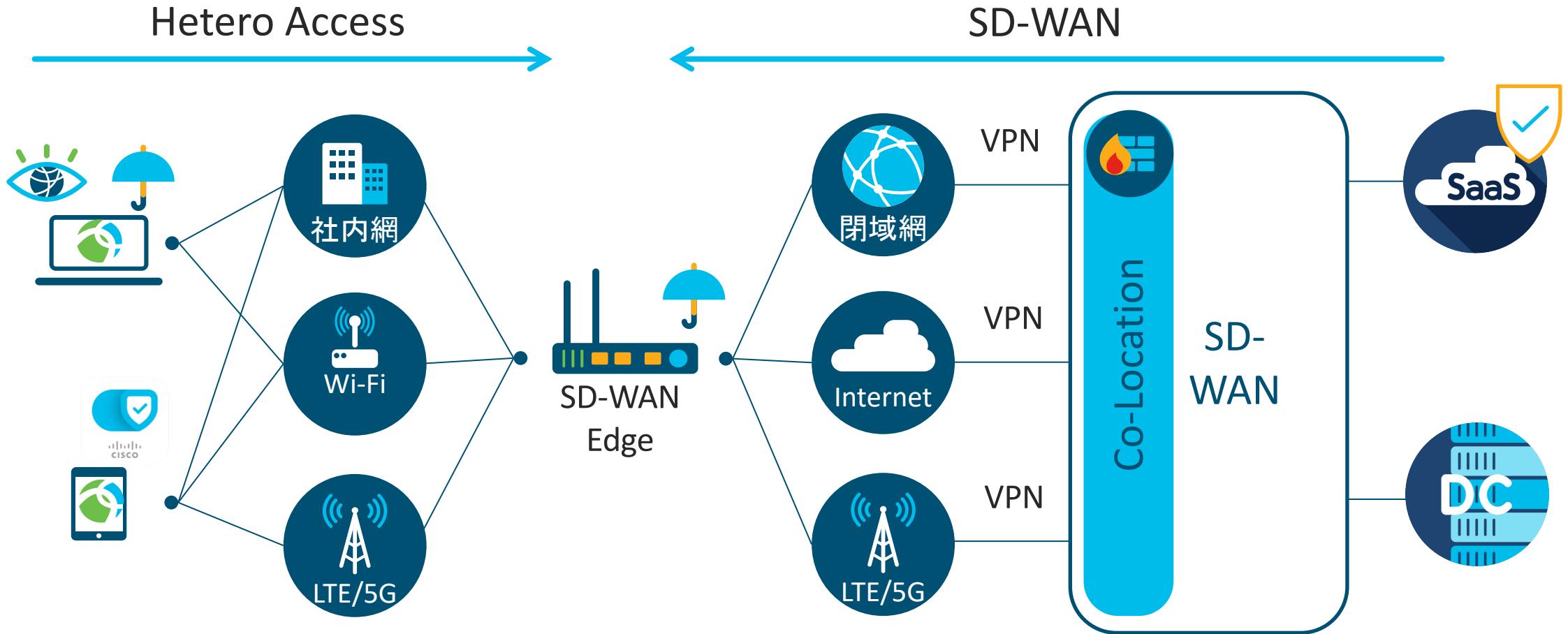


(2) Native Approach - SP API Exposure

通信事業者が提供する「ネットワークスライス」を、企業がAPIにより制御活用する



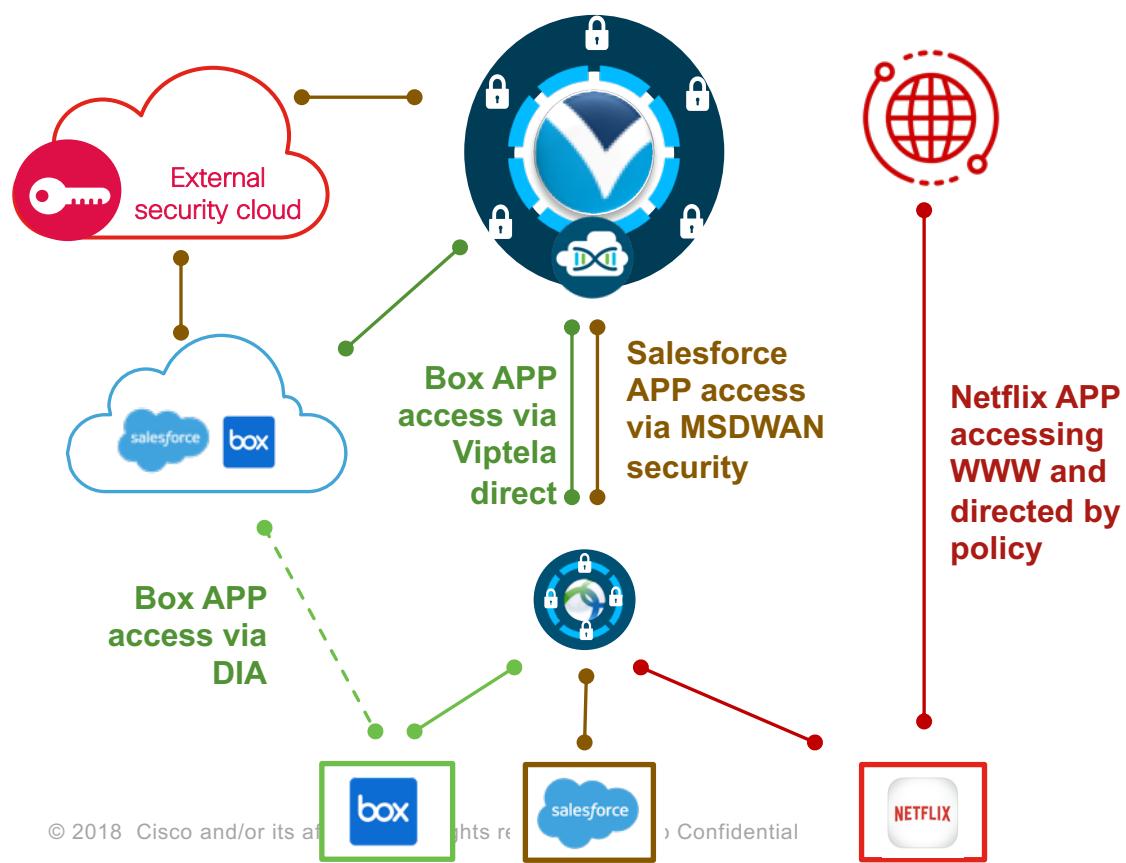
Mobile SD-WAN



Mobile SD-WAN の主な機能

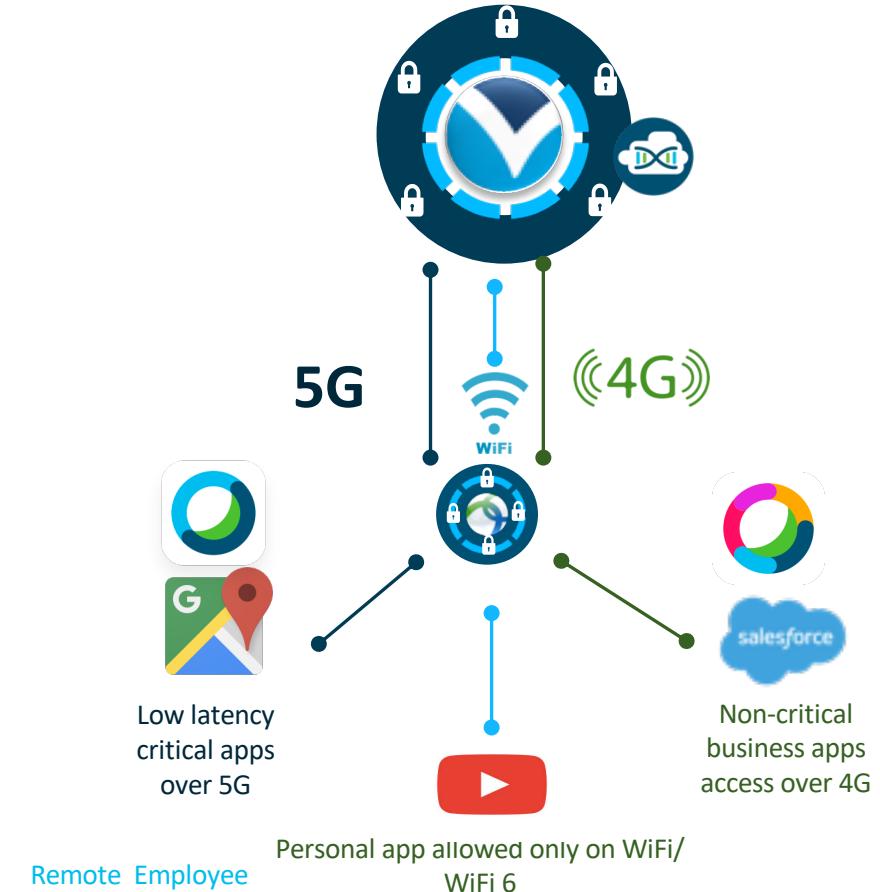
アプリケーション毎のポリシーと可視化

アプリケーション,FQDN,サービス毎のトラフィック制御
Micro Segmentation

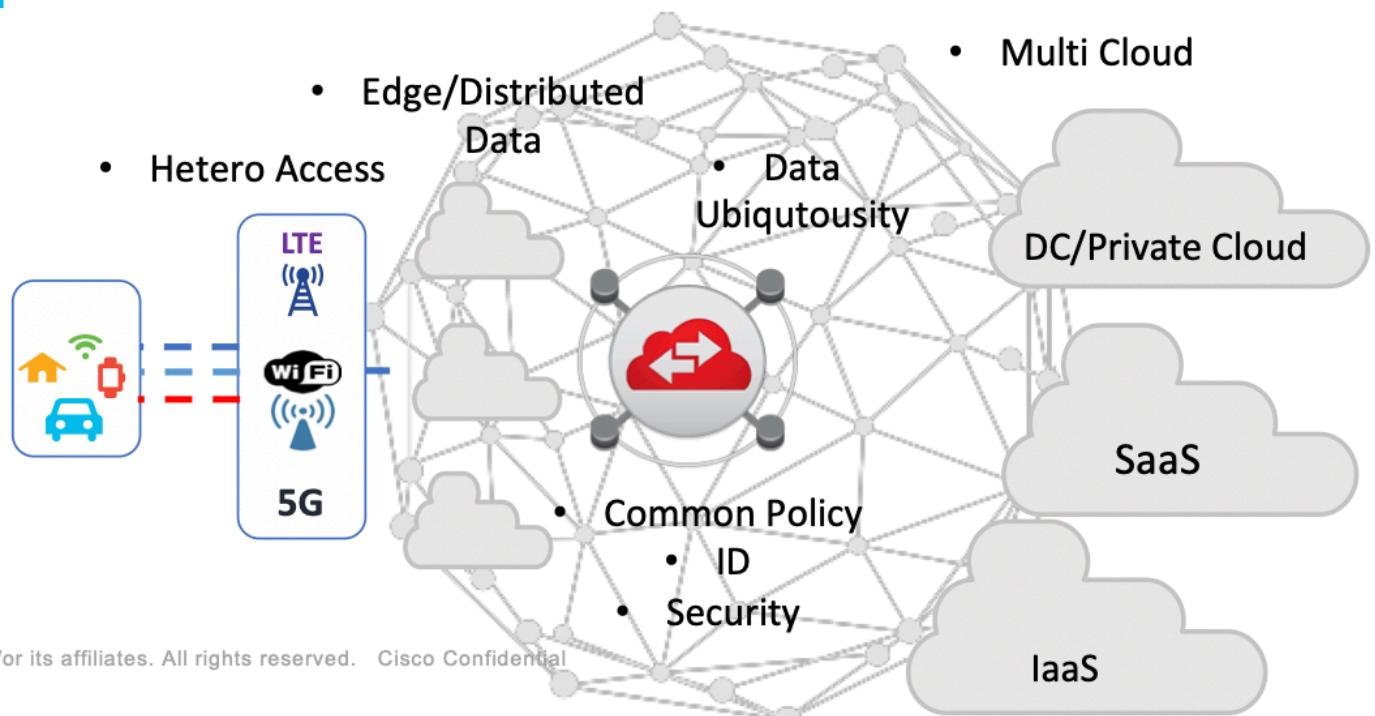
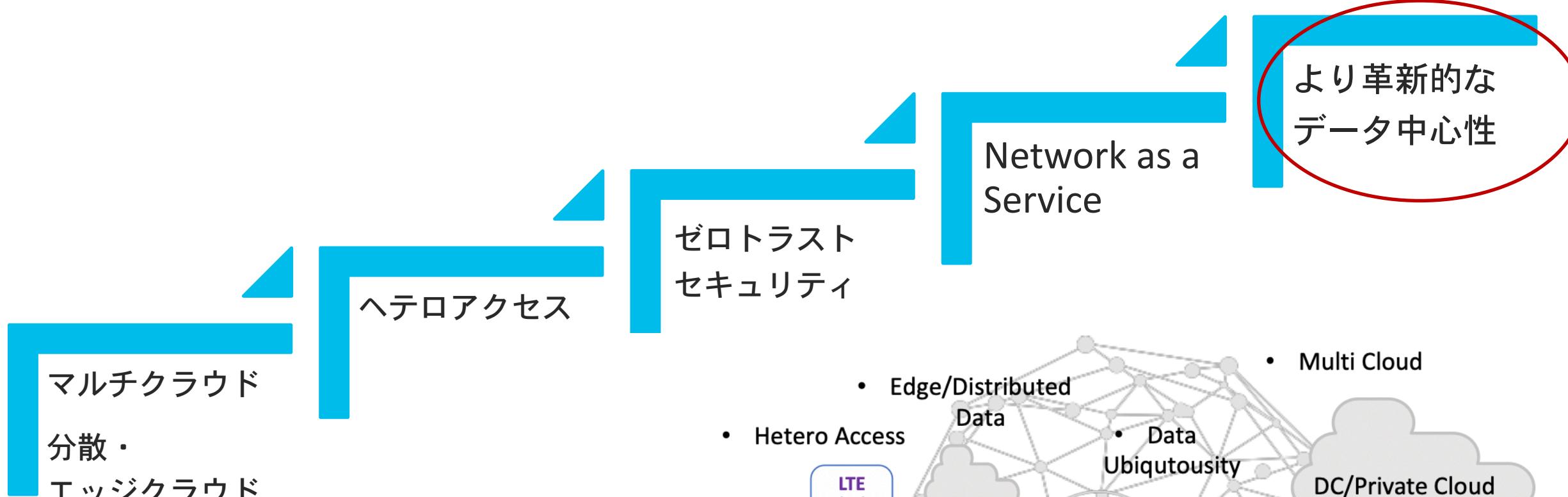


マルチネットワークアクセス

アプリケーションのポリシーに基づき、
複数のアクセス（LTE, WiFi, WiFi6, 5G）を使い分ける

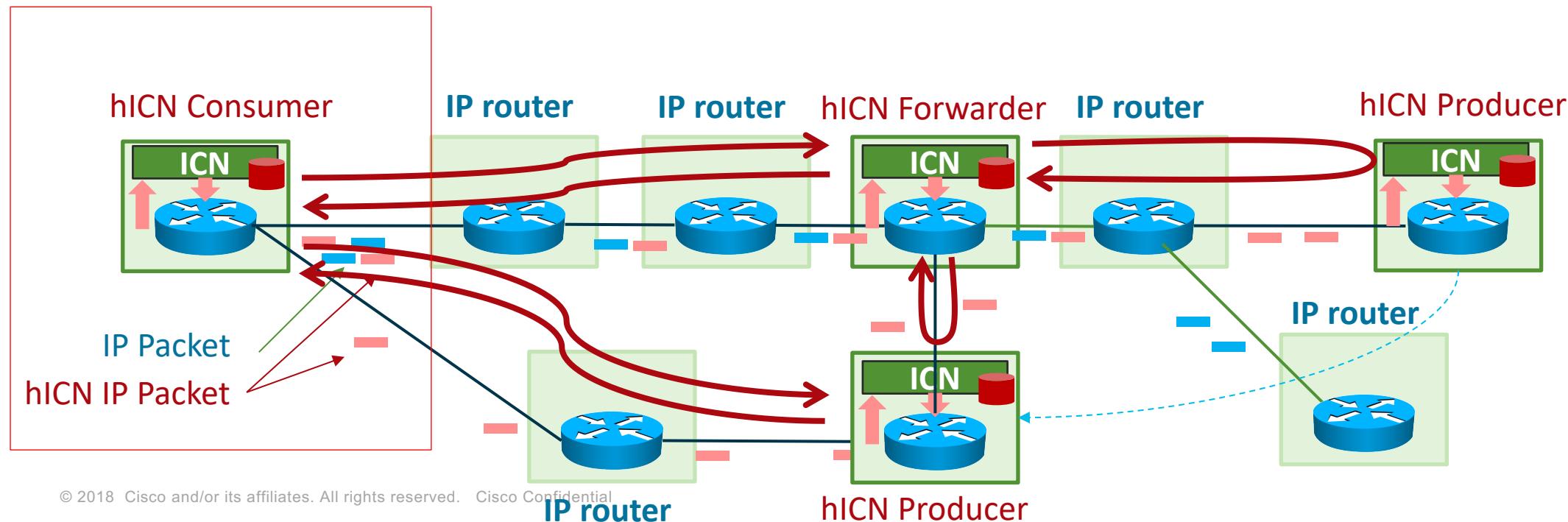


データ中心アーキテクチャへの道筋

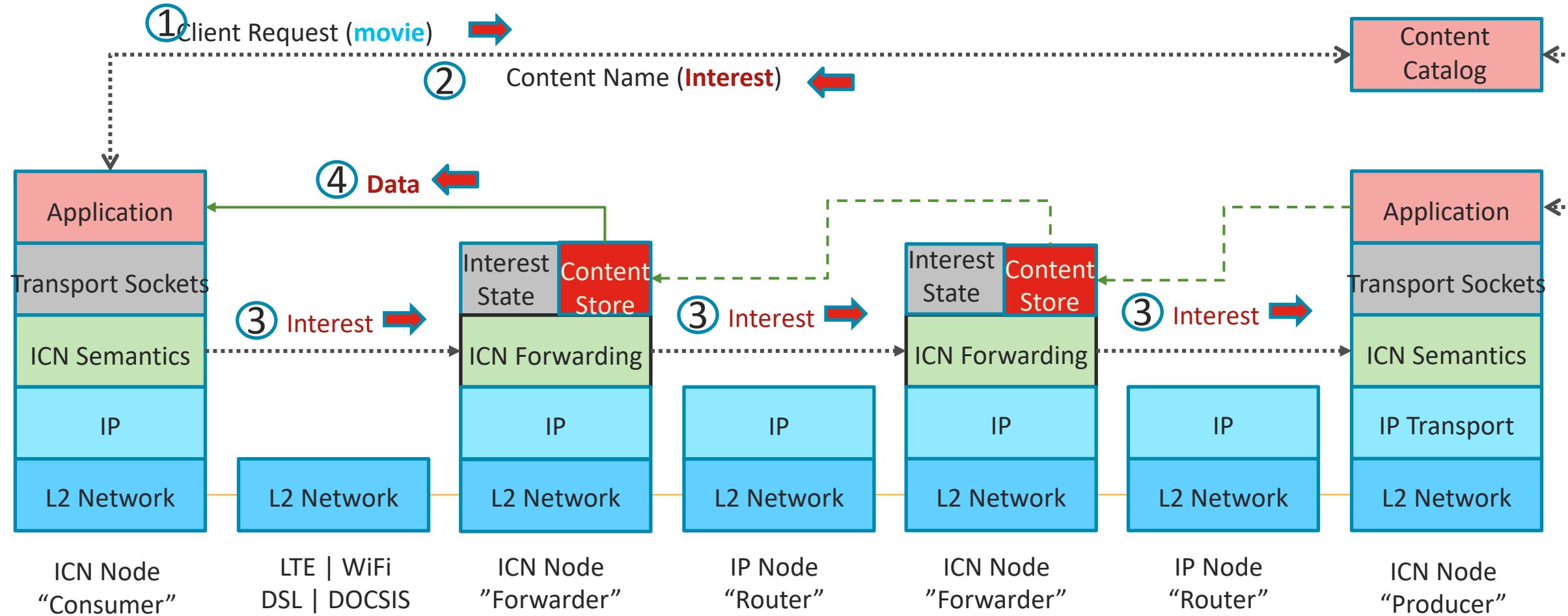


より革新的な「データ中心性」の例

- Hybrid ICN (Information Centric Networking)
 - 宛先アドレスでなく、"Named Data", "Content Name"に基づきルーティングする
 - コンテンツ最適性、Anchor-less Mobility、Multi-Path Accessを実現
 - ICNでは共存普及が困難なため、IPv6アドレススペースを使って"Nameを表現"
--> Hybrid ICN, SRv6のFunctionとして定義可能



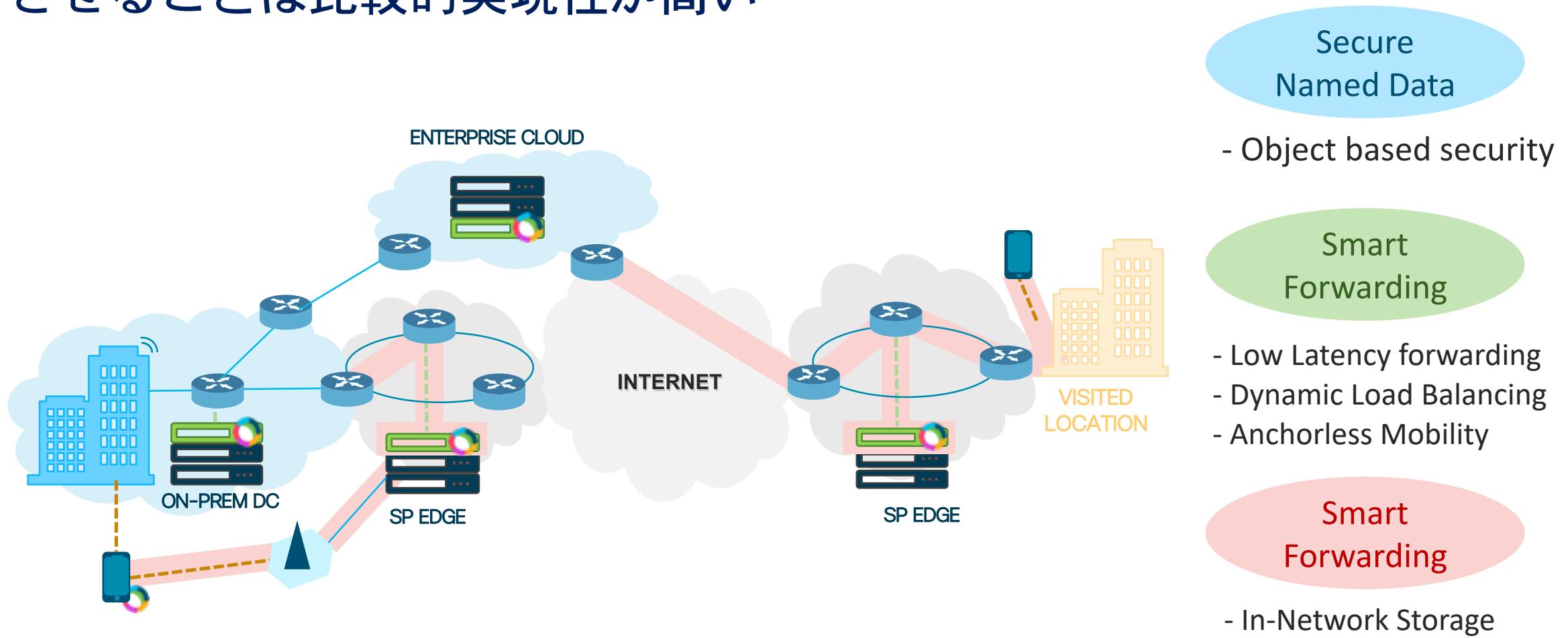
Hybrid ICN 動作概要



- ・コンテンツ・トラフィックの最適化
- ・セッション確立なし・コネクションレスのため、本質的にモビリティをサポート
- ・データ自体の保護（改ざん防止・著作権）

ネットワークスライシングとの連携

hICNのような新たなパラダイムも、ひとつのスライスとして動作させることは比較的実現性が高い



まとめ

- ・ネットワークも「コネクション中心」から「データ中心へ」
 - データの収集・配信、およびその処理・活用とフィードバック、そしてセキュリティ保護の観点からアーキテクチャを見直そう
 - ディジタル時代におけるプラットフォームシステムとしての、スケーリング・保守性・信頼性を獲得しよう
- ・データ中心アーキテクチャへの道筋

