CISCO Connect Toronto

October 30, 2019

You make **possible**

cisco Connect Toronto

Cisco Connect Toronto 2019

Core Concepts of 5G

Humberto J. La Roche, Ph.D.
Principal Engineer – Service Provider CTO Group
October 30, 2019

Contents

1 What's the deal with 5G?

2 Virtualization will drive the 5G RAN

3 ... and speaking of transport

4 A 5G mobile core dedicated to services

5 Why the 5G network needs end-to-end automation

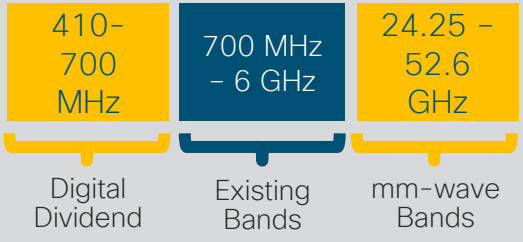
6 Industry Verticals - IIoT

6 Summary

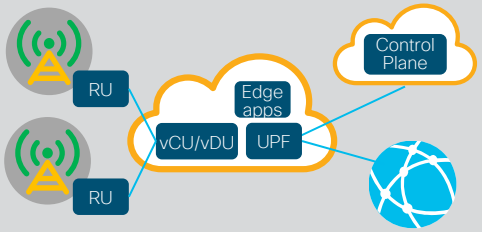
What's the Deal with 5G?

5G New Radio – The Highlights

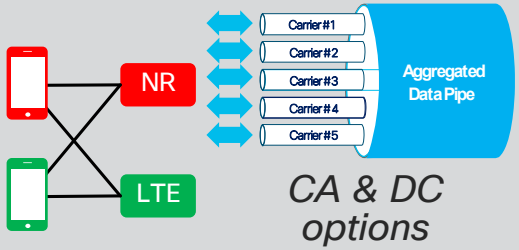
Expanded Spectrum: NR new bands



RAN Decomposition: multi-cell features

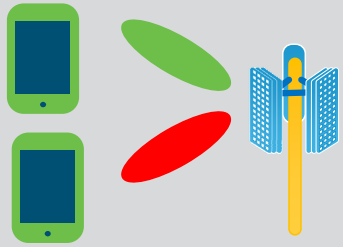


Higher Data rates with Multiband connectivity

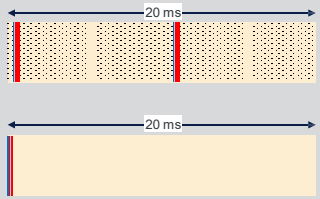


CA & DC options

Massive MIMO & Beamforming

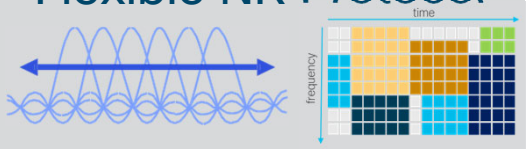


Lean Frame Design



Eliminating LTEs Cell-Specific RS gives capacity & efficient power

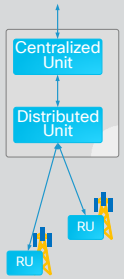
Flexible NR Protocol



- Mixed numerology: Flexible TTI,
- Bandwidth parts - mixed numerology within a carrier (RAN slicing)

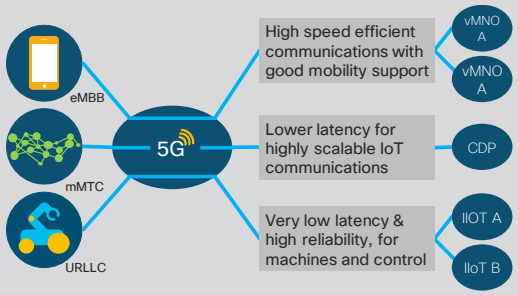
5G System – The Highlights

Decomposed 5G-RAN

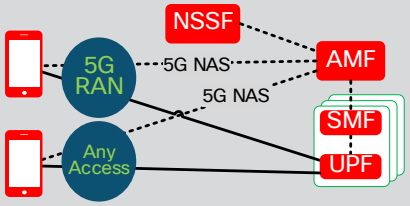


- Virtualization is natural
- CU + DU can be collocated & NR fronthaul has reduced latency restrictions vs. LTE

E2E Network Slicing

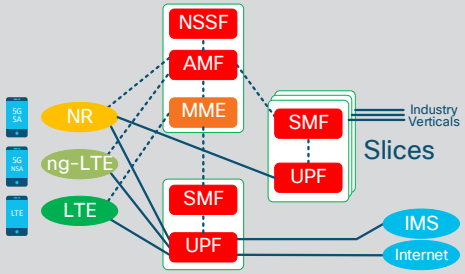


Any-Access

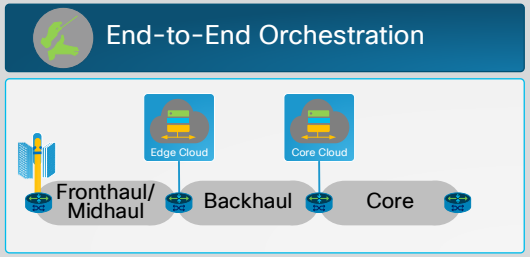


- Access-Indifferent NAS
- Configures device policy

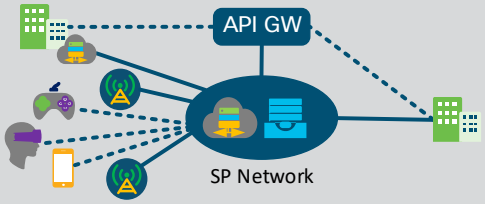
Decomposed 5G Core



E2E Automation



Enablement for B2B



- Industry Verticals
- Network as a Service

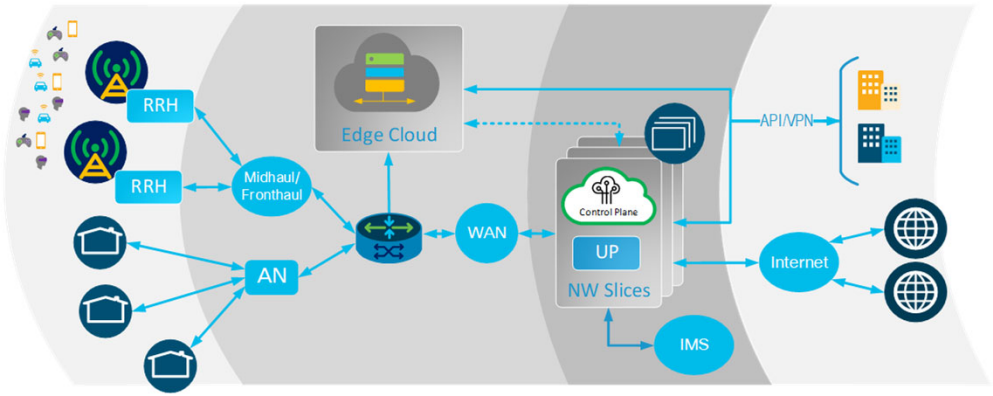
… as a Consequence, the Infrastructure Edge Emerges!

Decomposition
Network functions are separated control/signaling and user/data for optimization of resources

Disaggregation
Software-centric solution using COTS hardware which can be procured separately

Infrastructure Convergence
Fixed and mobile networks to share a common 5G core-based infrastructure for efficient operational practices

These shifts apply to LTE as well!



The 5G system promotes the emergence of an edge infrastructure that combines decomposed subscriber management from a converged core with the data plane (access node – e.g. DSLAM/OLT) of wireline as well as upper layers of the 3GPP radio stack

3GPP NR Use Cases

Enhanced Mobile Broadband (eMBB)

- 10-20 Gbps peak
- 100 Mbps whenever needed
- 10000x more traffic
- Macro and small cells
- Support for high mobility (500 km/h)
- Network energy saving by 100 times

Massive Machine-Type Communications (mMTC)

- High density of devices ($2x10^5$ - 10^6 per km^2)
- Long range
- Low data rate (1 - 100 kbps)
- M2M low cost
- 10 years battery
- Asynchronous access

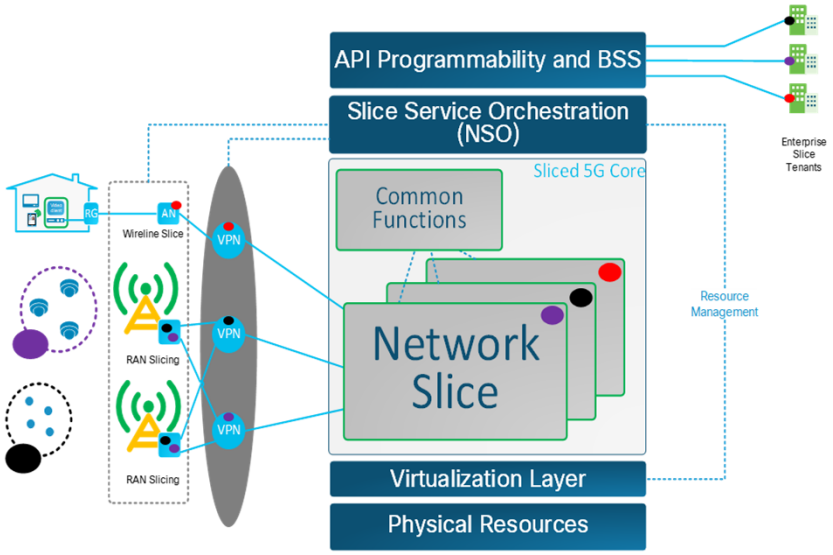
Ultra Reliable Low Latency Communications (URLLC)

- Ultra responsive & mission critical
  - <1 ms air i/f latency
  - 5 ms E2E latency
- Ultra reliable and resilient (better than 10^{-5} PLR)
- Low to medium data rates (50 kbps - 10 Mbps)

Source: Recommendation ITU-R M.2083

Network Slicing

Network Slicing is fundamentally an end-to-end **partitioning of the network resources and network functions** so that selected applications/services/ connections may **run in isolation** from each other and for a **Specific Business Purpose**



Benefits

- Each separable business operation can be efficiently and reliably run on a network slice

Infrastructure orchestration manages the complexity driven by the requirements of each slice

- Each slice can have its own MANO
- Individual OSS environment

New service introductions are fast

- Slicing has a significant reduction in regression testing cycles
- Easier to debug configs limit “collateral damage”
- Isolation eliminates effects of rogue applications (E.g. M2M)
- Smaller failure groups imply no single “too big to fail” node

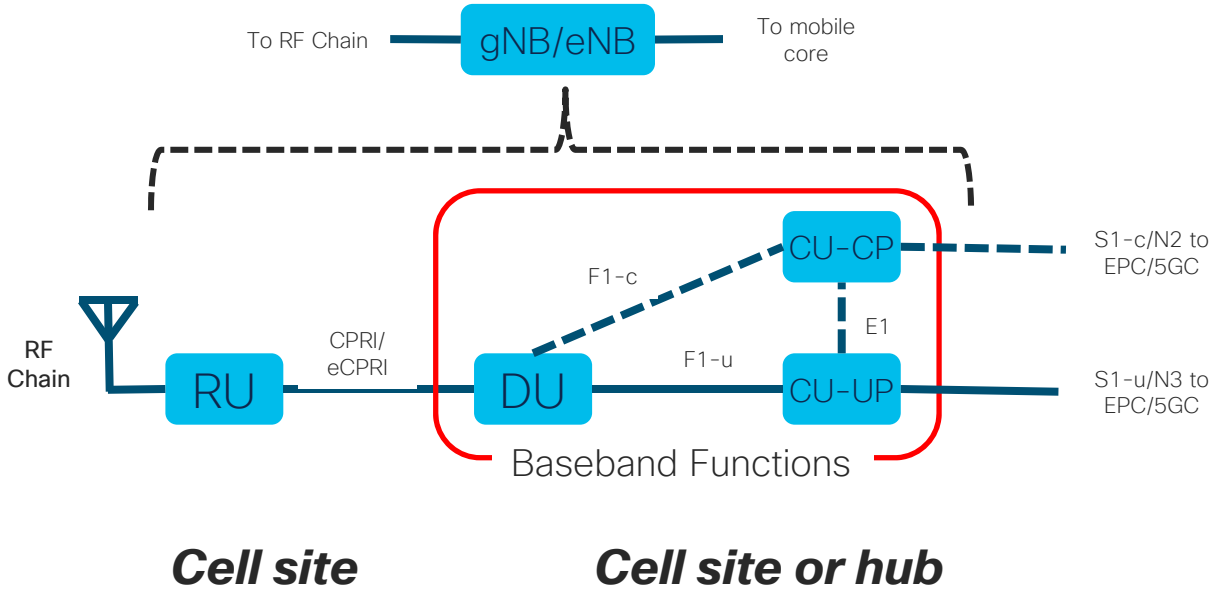
Virtualization will drive the 5G RAN

Decomposition of eNB and gNB

RU – Radio Unit or Remote Unit converts digitized radio to analog baseband waveform

DU – Distributed Unit synchronously schedules UL/DL radio frames

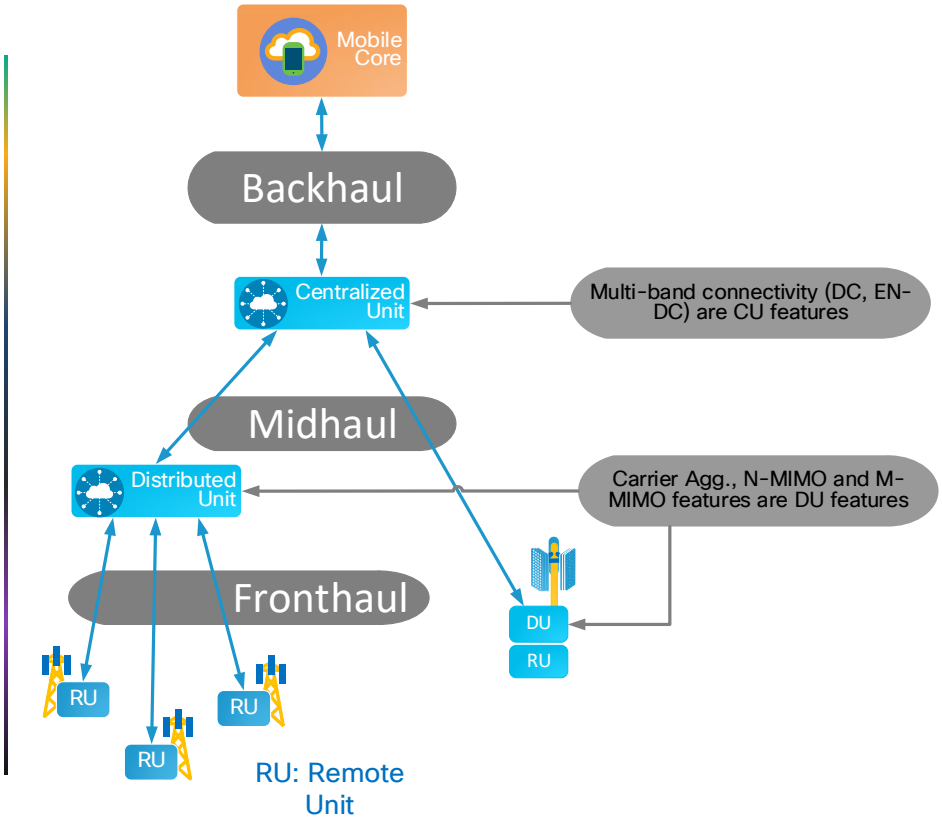
CU – Centralized Unit handles RAN-core signaling and radio control plane while providing packet interface to backhaul transport



RUs can be integrated into the antenna system (an AAS) or can be stand-alone supporting purely analog chains

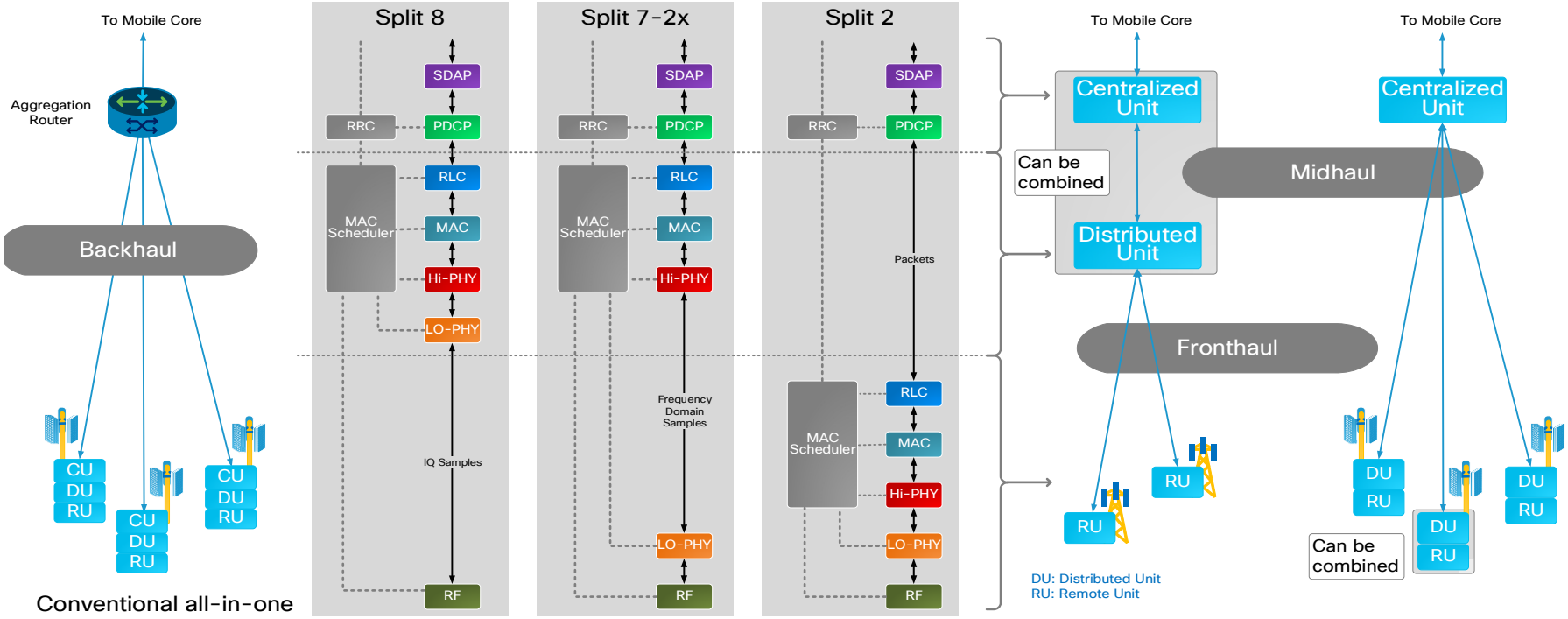
Cloud RAN is a Systems Approach to Supporting Licensed Radio Across Multiple Bands

- It is inherently based on disaggregation of SW and HW
- Why?
  - Improves the radio interface through **cooperative radio benefits** (examples: cooperative transmission/reception, eICIC, integrated X2 HOs, …)
  - Employs a **pooling concept** which reduces power consumption off busy hour
    - Radio stacks are pooled across a service area
    - Unused densification nodes can be shut-off for power savings (require network reconfiguration)
  - Lowers CAPEX and OPEX
    - Can transform the whole RAN ecosystem through decomposition of functions via open interfaces promoting an open competitive ecosystem
    - Defines the major elements through software functions that can cost-effectively run on COTS compute as a cloud platform



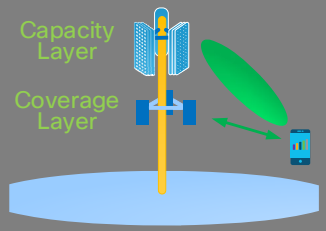
https://www.itu.int/dms_pub/itu-t/opb/tut/T-TUT-HOME-2018-PDF-E.pdf

Functional Splits Mapped to C-RAN Functions



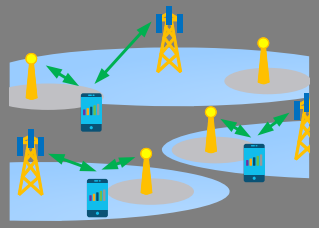
Three **Basic RAN Splits** that determine the IP Transport Architecture

Cloud RAN Facilitates Multiple Antenna Deployments via Centralized Control Points



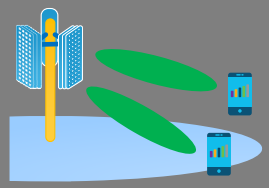
Multiband Connectivity

- Two or more bands deployed cooperatively at same cell (Carrier Agg., DC, EN-DC)
- Requires multi-transceiver UEs



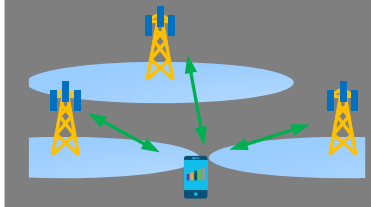
Small Cell Densification

- Supplemental hotspot coverage
- Multi-band connectivity is a bonus benefit



Massive Multi-user MIMO

- Multi-antenna panels support concurrent users via beamforming
- Increases aggregate capacity of cell



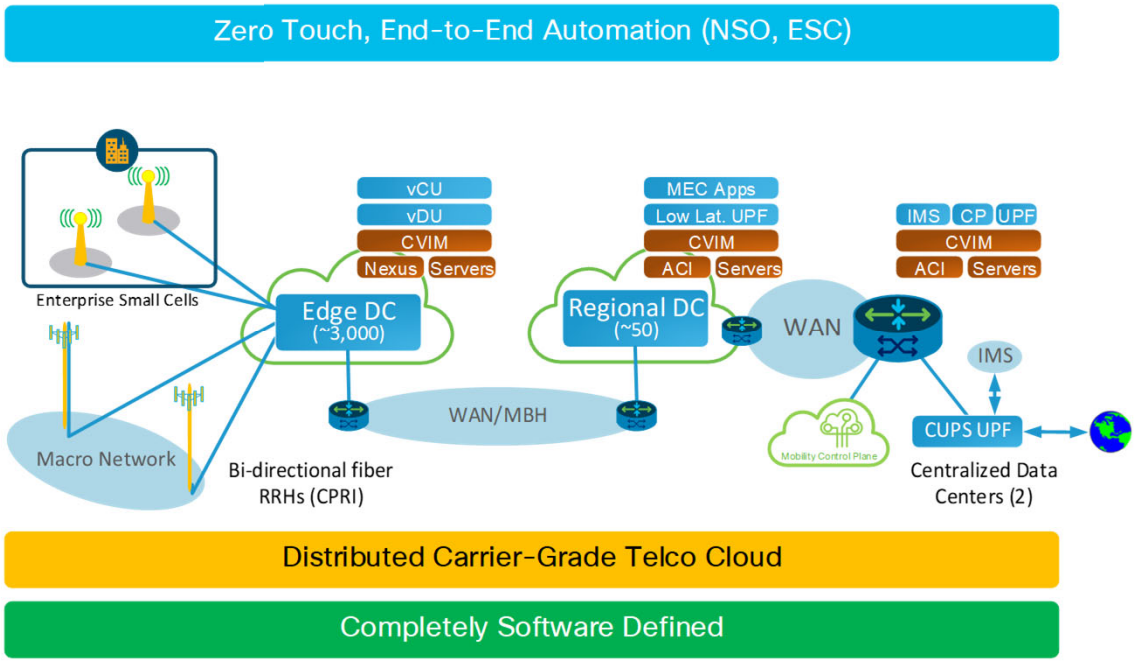
Network MIMO

- Coordinated & synchronized Tx/Rx from multiple cell sites
- Requires perfect fronthaul to meet stringent latency requirements

Outcome

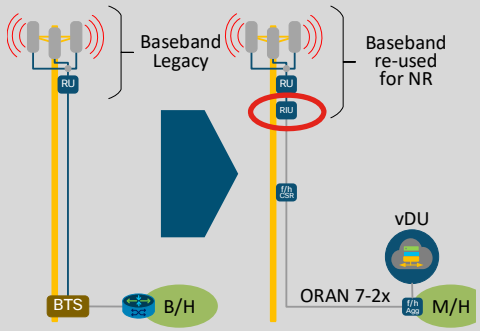
Fully Virtualized Telco Cloud with vRAN & Edge Computing

Rakuten



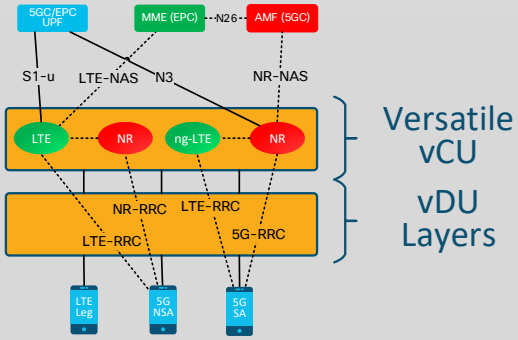
- Open, decomposed, and virtualized RAN
- Edge computing for enhanced experience
- New business models including B2B Monetization
- End-to-end closed loop automation
- Cisco Capabilities: IP transport, automation software, Virtualization platform, Mobile core, WiFi, Services

For “Brownfields”: New Capabilities = Solutions



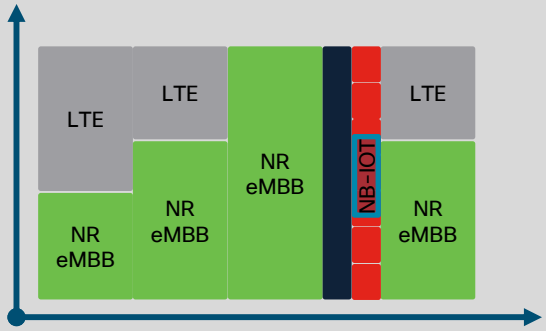
Objective: preserve tower mounted RF chain when evolving to NR and attain “slim cell site”
Solution: a Lo-PHY appliance implementing programmable digitized waveform conversion

Preserve the value of deployed equipment on towers



Objective: Align core evolution to RAN evolution
Solution: versatile vCU supports multiple signaling for registering with MME or AMF and dual connectivity bearers

Align RAN and Core in evolution from EPC to NSA to SA

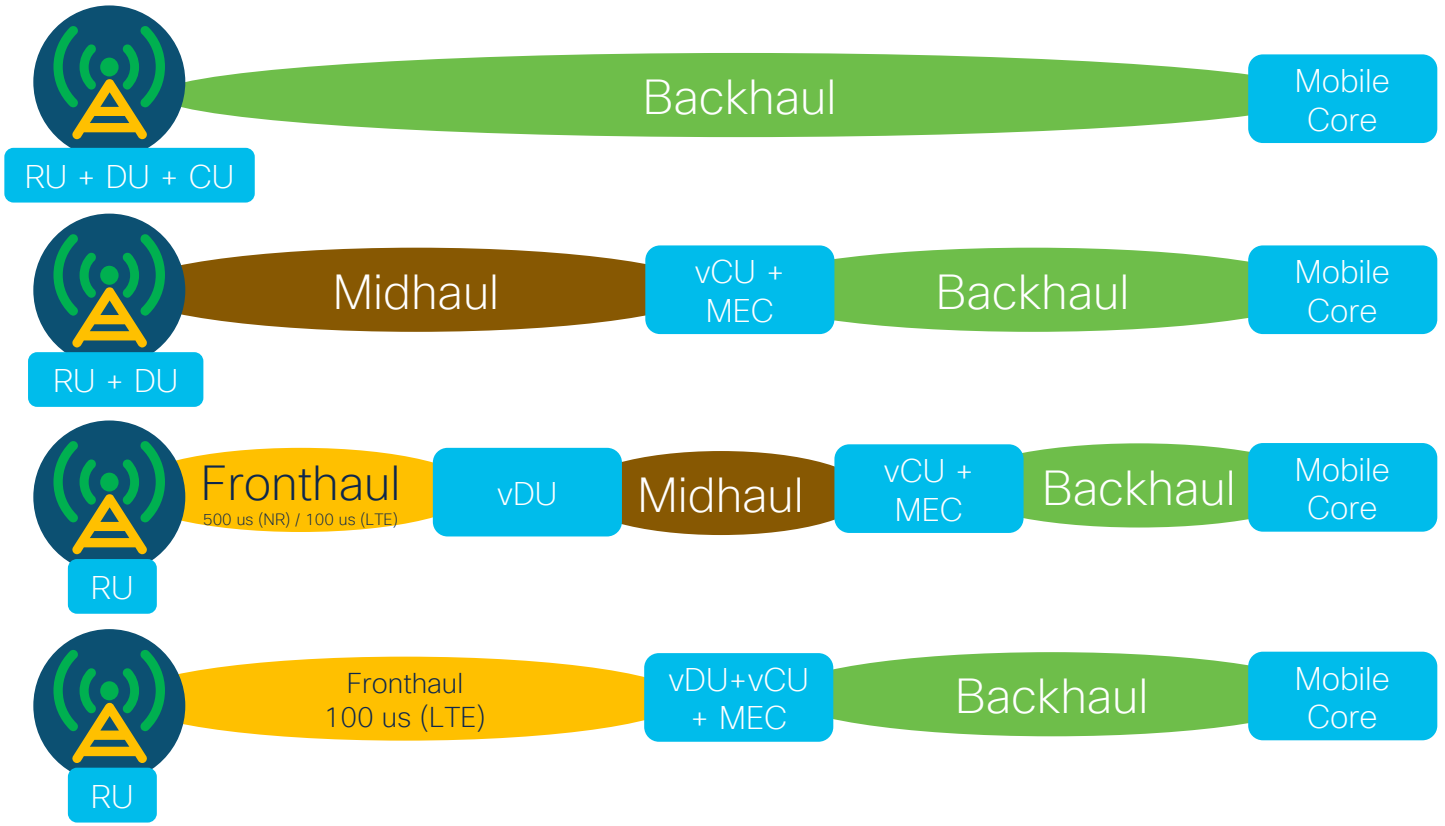


Objective: Ensure coexistence and sharing of 5G NR and LTE on same bands
Solution: Use NR flexible slot-based framework including DSS & BWP to manage transition from LTE to NR

Optimize Spectrum utilization + support legacy and NR UEs

… and speaking of transport

5G RAN Transport

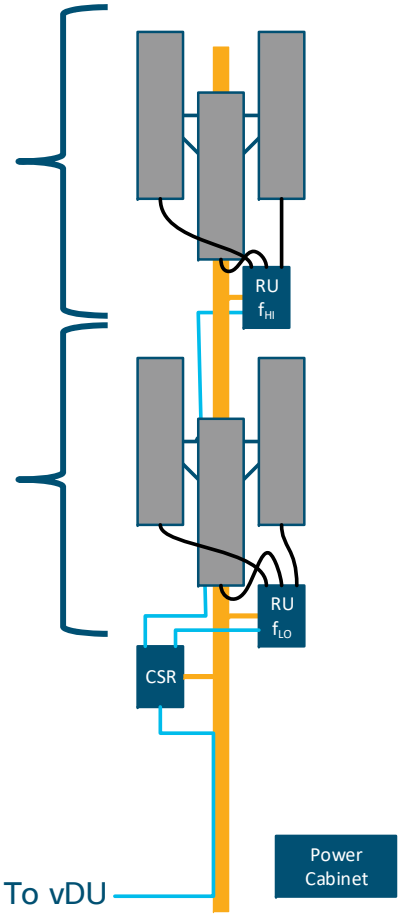


Example: A Lean Cell-Site Design

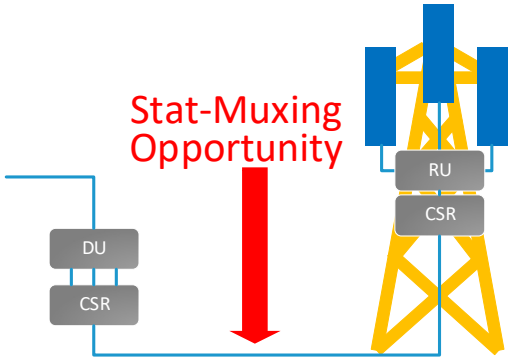
- 3 FDD bands
- Tower mounted RRUs and remote baseband with Option 7-2x fiber interconnect
- The cell-site router is tower-mounted as well in IP65 enclosure; it is required as a way to breakout to each RU
- Only power pad (from tower provider) at base of cell tower

• RF Design

- Two antennas: one for 600 MHz and another for 2GHz
- 4 x 4 MIMO

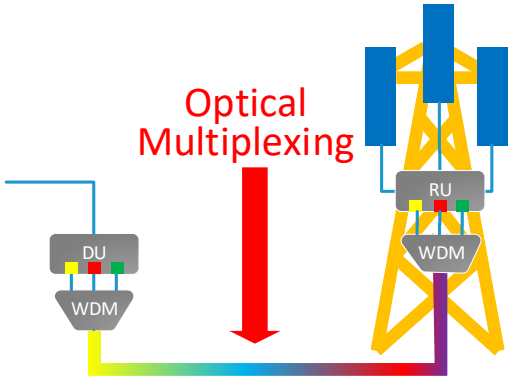


Active Antennas Systems for Split 7-2x
Stat-Muxing is Possible



Electrical Multiplexing

- Load is variable dependent on radio resources used
- Oversubscription is possible. When peak load exceeds allocation, RU can hide impact to prevent RLF
- Latency requirements



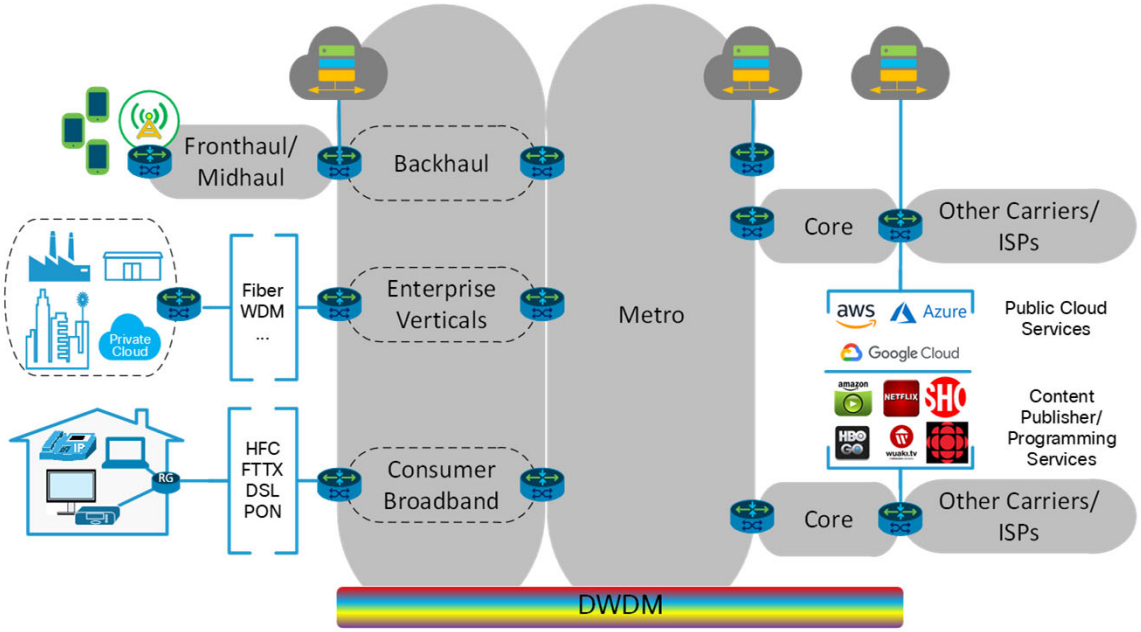
Optical Multiplexing

- No stat-muxing
- Can use direct optics from DU into transport system but expensive anyway of extended-TEMP requirement on outside plant optics

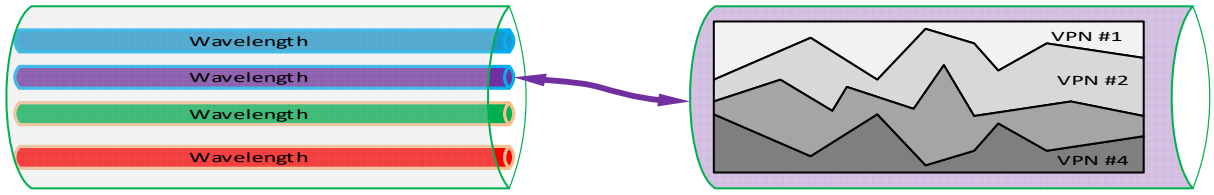
Services Require a Converged Transport Fabric Integrated with the Cloud/5G/Enterprise

The Problem: Converged Network Fabric is more than just 5G ➔ Need a modern fabric

- Multiple connectivity needs include enterprise and wireline
- 5G support
  - midhaul/fronthaul
  - Network slicing
- Enterprise/Private, Telco, and Public Clouds
- Content publishing and video onboarding
- Modern automation architecture

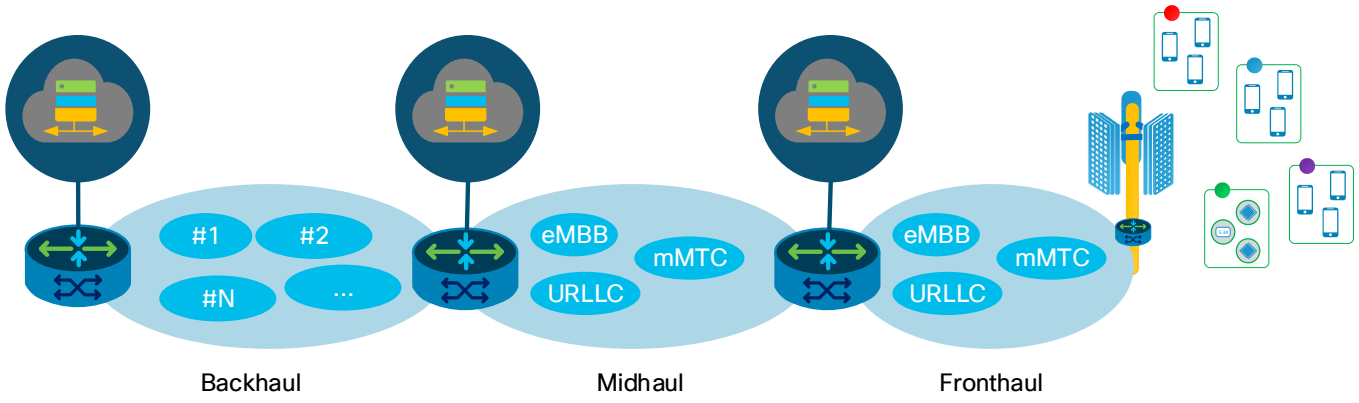


Sharing is the Key to Cost Containment in Network Slicing



Fiber provides common transport and WDM provides separation

VPNs provide statistical multiplexing efficiencies and separation of traffic



Network slicing support across domains is essential

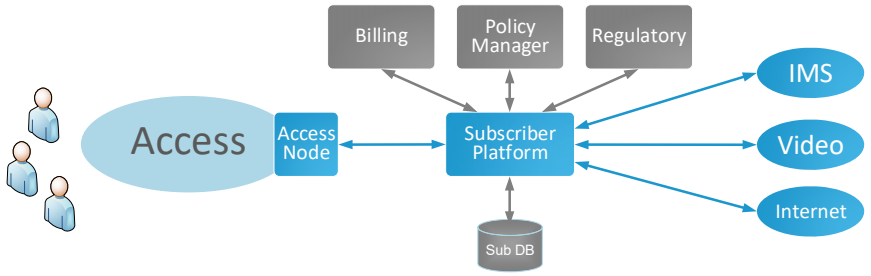
A 5G Mobile Core Dedicated to Services

Decomposition as a Trend

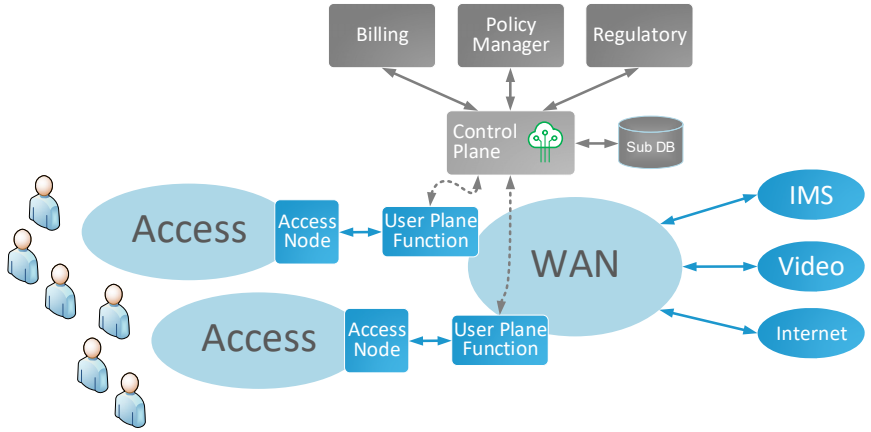
Monolithic subscriber access platforms combine control plane functions and user plane functions into one box

Decomposed subscriber access platforms decompose control plane functions and multiple user plane functions

- Control elements can be deployed in the cloud for transaction scalability
- User plane elements are fairly generic and can be deployed anywhere the IP network allows enabling edge compute and superior scalability for capacity



Conventional Subscriber Management

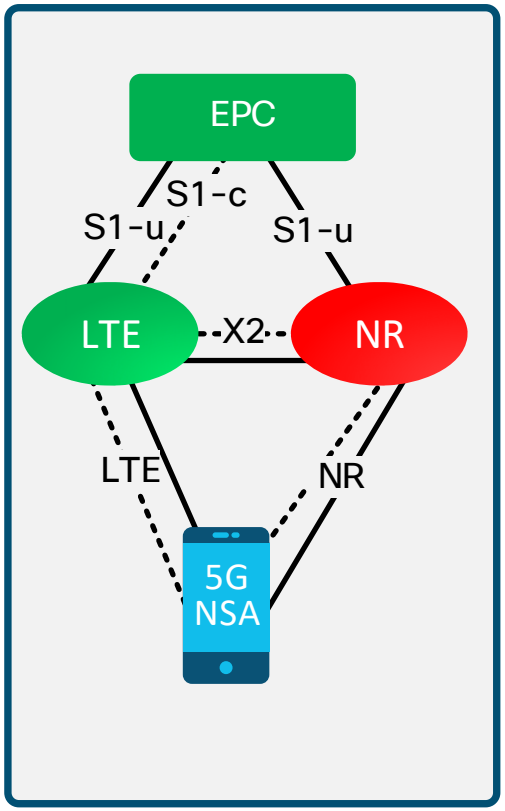


Disaggregated Subscriber Management

NR Non-Stand Alone Mode

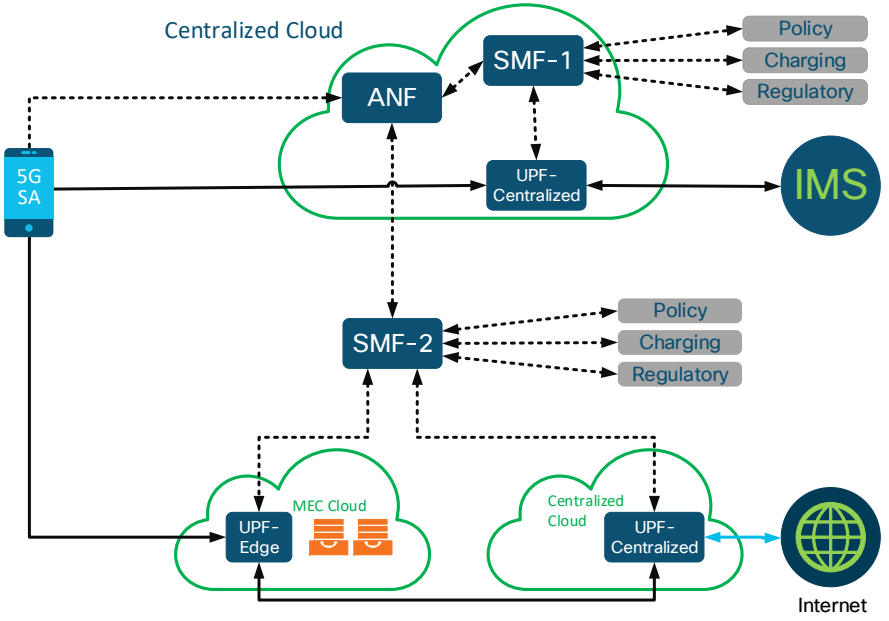
- Intended to address initial deployments of NR in an LTE context through re-use of EPC
- UEs register with an MME; all signaling to core is based on S1-MME
- NR is supported via a co-sited or remote gNB functions
- gNB is a “slave” to eNB “master”
- LTE provides “stability” to NR coverage, often because NR is in higher frequencies
- EN-DC is defined in TS-37.340 “Multi-connectivity”

Option 3x Dual Connectivity



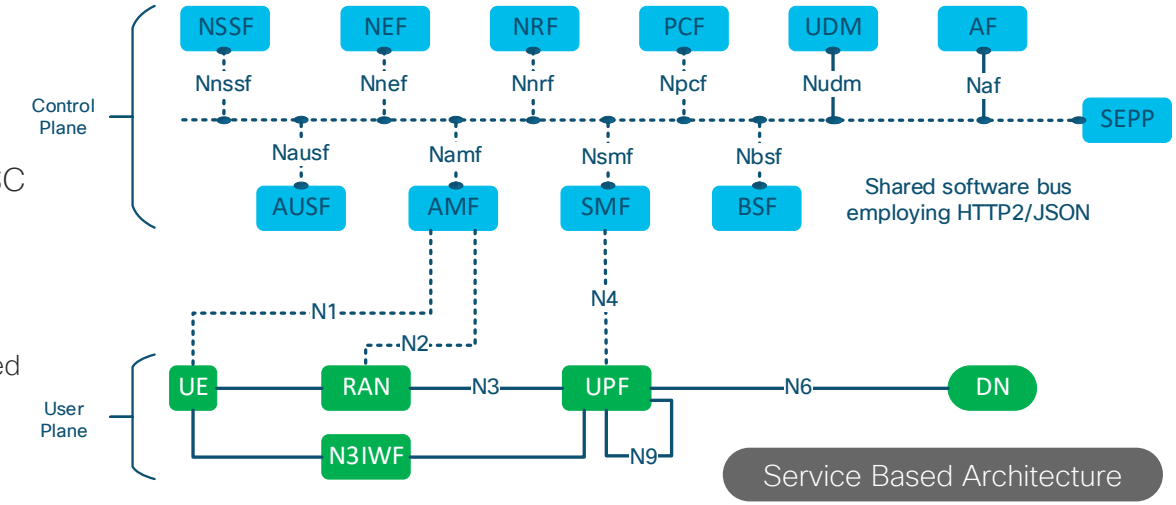
Why Decompose into Control + User Planes? …Flexibility!

- User plane (UP) and control plane (CP) can scale separately
- Option of VNF/CNF or PNF implementation as required
- UP and CP can evolve independently
- Decoupled UPs can be placed anywhere IP network allows
- Centralized CP can provide for centralized capabilities (charging, policy, regulatory)
- Better suited for network slicing



5G Core Architecture

- 5GC “fixes” many of the 4G limitations (SSC modes, SMF per PDU)
- CUPS is designed as part of the architecture from day one
  - MEC, high bandwidth, LBO etc. can be delivered easily and at low cost
- New core is designed for cloud and convergence
  - Services based = rapid innovation, Personalized Internet and web like experiences are much easier to deploy and automate
- Profiles and rules are much easier to push to the UE
  - Dynamic connections, location services, multi access are the norm

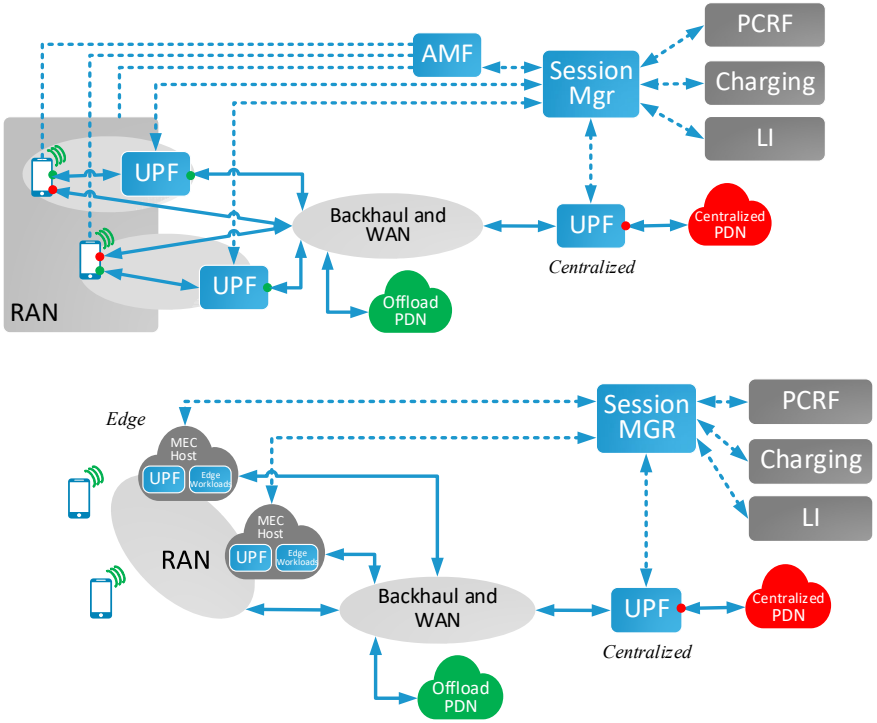


AF	Application Function	**NSSF**	Network Slice Selection Function
AUSF	Authentication Server Function	**PCF**	Policy Control Function
AMF	Access and Mobility Management Function	**SMF**	Session Management Function
DN	Data Network	**UDM**	Unified Data Management
BSF	Binding Support Function	**UPF**	User Plane Function
NEF	Network Exposure Function	**UE**	User Equipment
NRF	NF Repository Function	**N3IWF**	Non-3GPP Interworking Function

A Decomposed Mobile Core Supports Session and Service Continuity Modes

- The 5GC supports SSC modes
  - Conventional IP address preservation or
  - IP address changes on re-anchoring
- This flexibility allows
  - Support for app-based mobility where useful
  - And network-based mobility when required (e.g. VoIP)

TS 23.501

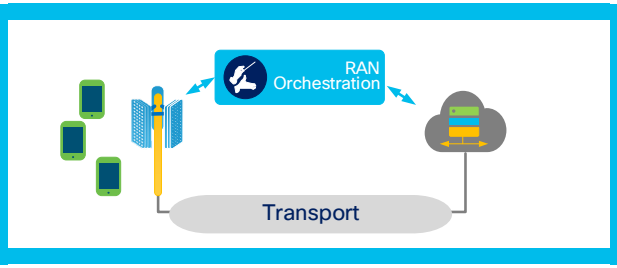


Take-aways

- Decomposition of the mobile core has **huge** advantages including
  - Separately scalability of control and user plane functions
  - Flexibility in choosing user plane functions (virtual or physical)
  - Flexibility in placing the user plane functions anywhere the IP network allows
- The initial 5G core is really a 4G core
  - This is the NSA (Non-Standalone Architecture)
  - The RAN is in charge of presenting S1 interfaces to the core through PDCP dual connectivity
- Evolution to a 5GC is based on decomposition
  - Services-based 5GC architecture has huge advantages for control plane scalability and offers radical simplification of the messaging

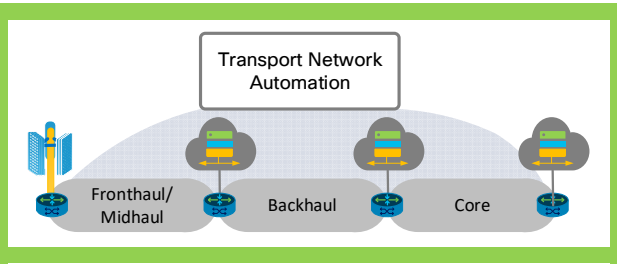
Why the 5G Network Needs End-to-end Automation

Domains for the 5G Automation Architecture



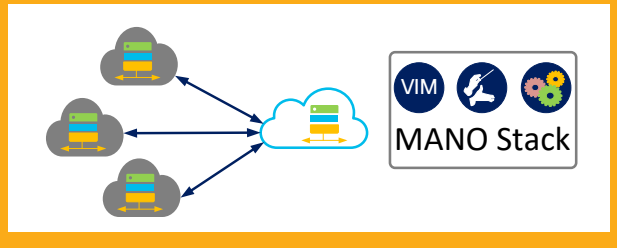
Radio Access Network

- Model-driven Yang and Netconf to configure RRH and DU/CU radio functions
- Follow the recently published XRAN specification designed to promote multi-vendor RRH deployment



Transport Network

- Segment Routing / EVPN – managed with model-driven Yang and Netconf applied end-to-end to transport
- Multi-domain including Data Center
- Network control includes WAN design tools, a path computation element suitable to SR, and model-driven config



Cloud Virtualization

- Common MANO stack across core/edge and service/ infrastructure clouds.
- Supported on open source model (KVM, OpenStack, etc.)
- Integrated with SDN network control; the same as used in the transport network, based on SR

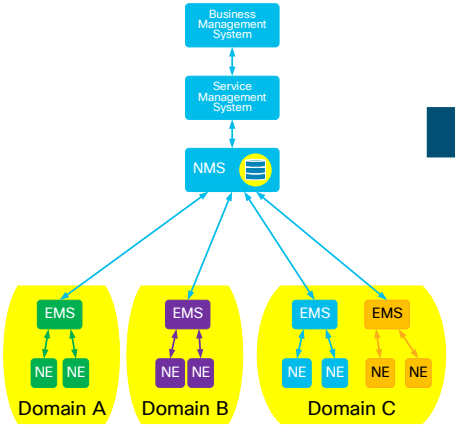
Service Modeling **Changes the Game** for Automation

In a traditional TMN hierarchy, the NMS has the network database and monolithically supports all FCAPS functions

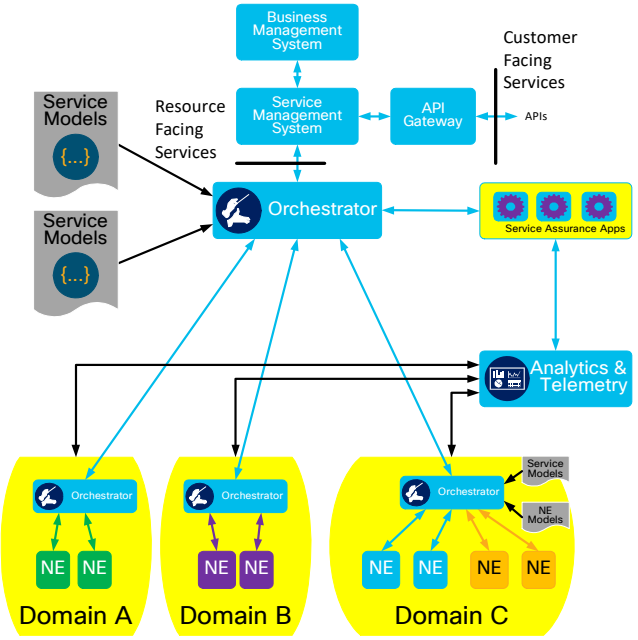
In a modern management system, functions are decomposed and specialized (e.g. separate configuration and fault/performance)

Configuration management uses equipment models (Yang) which drives open protocols (Netconf)

The decomposition is open and flexible allowing specialized systems and open source



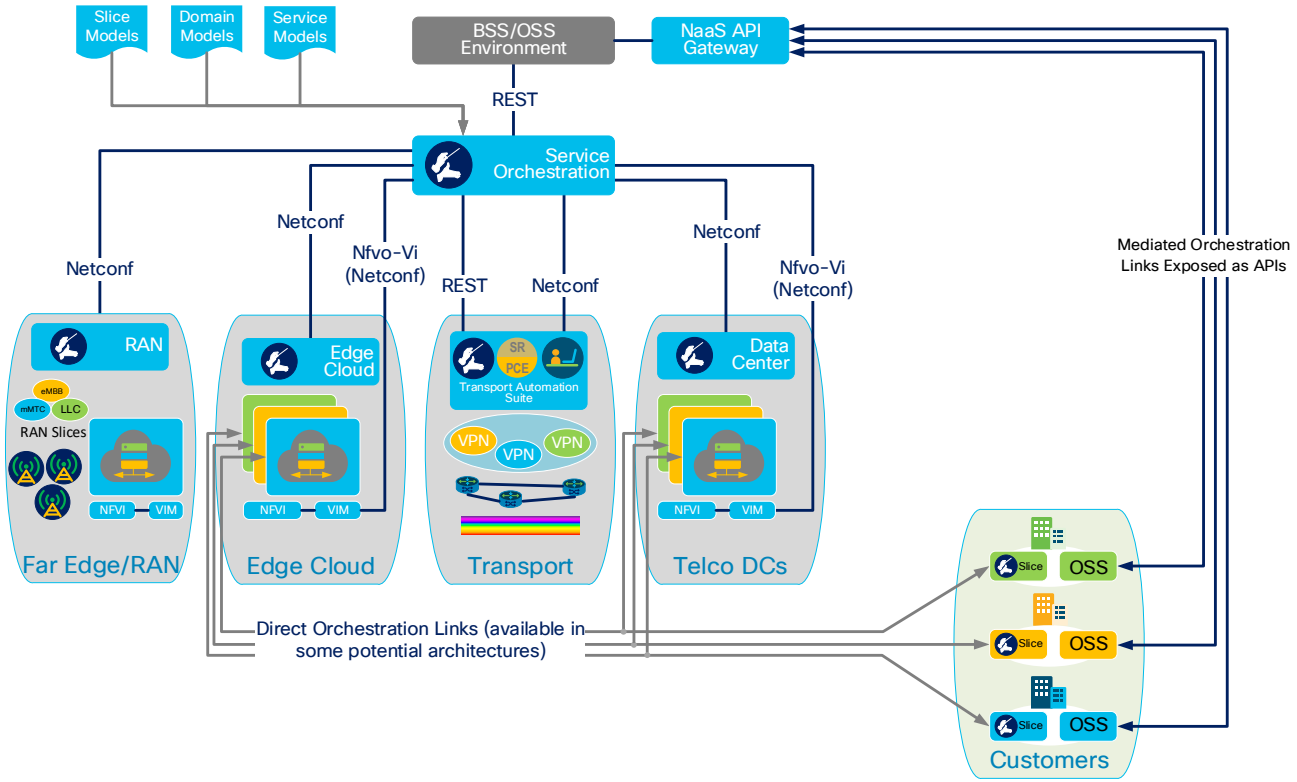
Traditional TMN Hierarchy



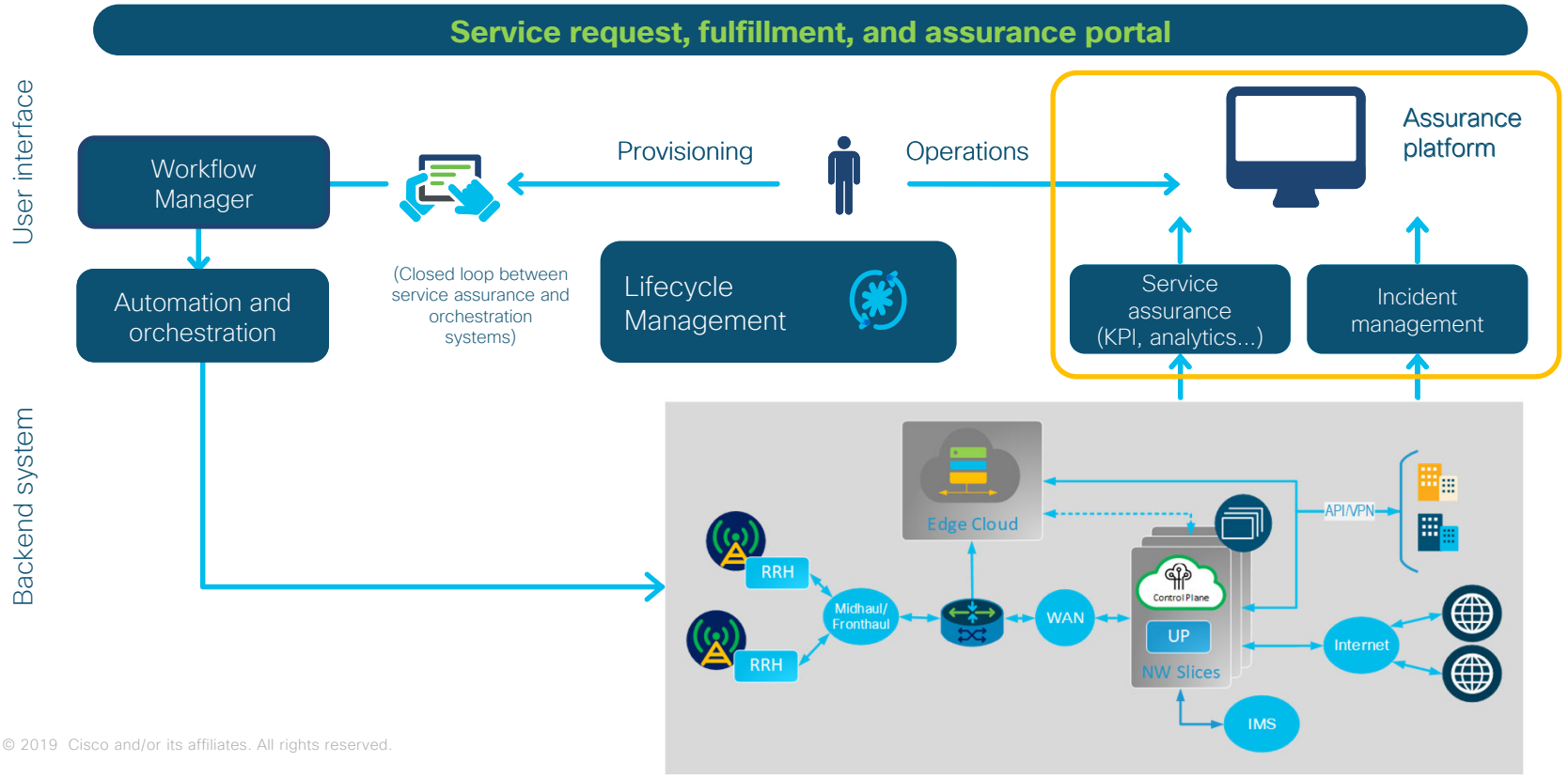
Orchestration in modern infrastructure

An End-to-End Orchestration Stack for 5G

- We **need** automation for the 5G *Network-as-a-Service* model to be viable
  - The 5G network must allocate resources as needed to enterprises (via an API)
  - Among resources allocated is an infrastructure slice which includes NFV functions
- Higher order e2e orchestrator consumes:
  - Abstracted models of multiple domains
  - slice definition models
  - Service definitions that can be instantiated on demand



5G Network Slicing



Industry Verticals IIoT

Private Radio as an MNO Offer

- These are networks that are dedicated to a special use cases which could be a constituency of human users or of “things” (3GPP refers to them as “non-public”)
- Examples of use cases include: industrial automation, venues, ports, logistics, venues, campuses, utilities, hospitals
- Radio may be limited a restricted geographic area or may require the wide area
- The use case is shaped by requirements for bandwidth, latency, availability, resiliency, data privacy.

Operator advantages | spectrum | economies of scale | expertise
Operator challenges | domain expertise | automation | API model

Technology Elements – Private Networks

Network slicing: provides partitioning of resources with isolation and makes it possible for them to be consumed by the enterprise

Hosting environment: enables enterprise customers to deploy workloads within the operator network

API interfaces: to enable the enterprise with simplified control of the network; an API model requires an automated network

Multicloud-enabled: allows the operator network and the enterprise workloads to seamlessly be managed across multiple clouds

Taxonomy of Edges

A use case for the edge, relative to the value, will fit among one or more of the following categories

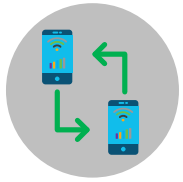
Latency reduction: Improves QoE or enables a service

Personalized user experience through edge computation

Security services through edge control of traffic

Data reduction (UL or DL) can address OPEX/CAPEX profile

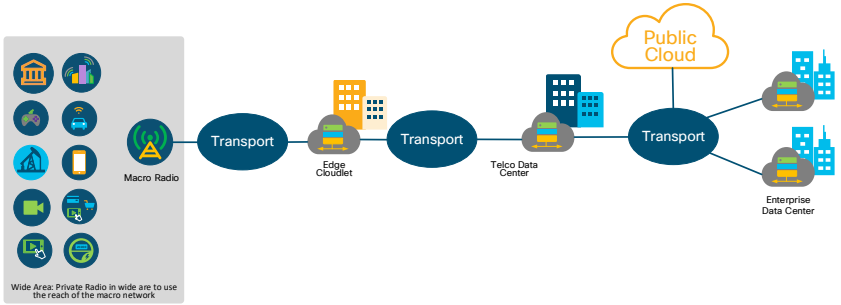
Infrastructure use cases improve/ enable the delivery of services

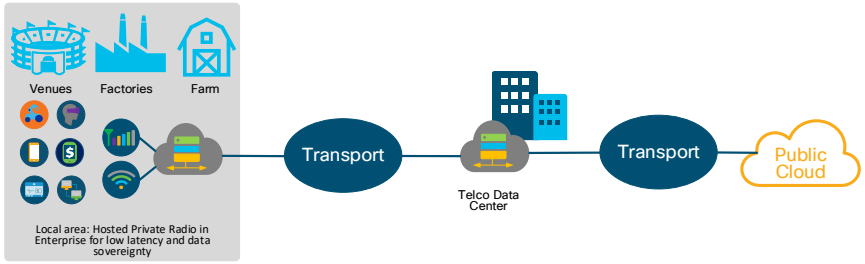


Peer to peer communications facilitated by the edge

Edge Use Cases: need to be tested for value against above

Two Use Cases Families by Location of the Radio





Wide Area Use Cases

- The enterprise uses the macro network as an extension of its own network by consuming an MNO product that allows it to do so
- Operator uses network slicing, defined end-to-end, to partition the domain used by the enterprise customer
- Workload placement optionally includes edge cloud for lower latency

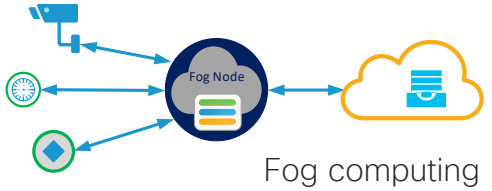
Local Area Use Cases

- The enterprise deploy a radio on premises for a specific use case.
- The need for the radio is localized to the enterprise
- Use cases are enterprise: factory automation, medical, corporate campuses & require domain expertise
- Private radio (licensed or unlicensed) is a significant part of the operator opportunity

Select Use Cases for Network Edge

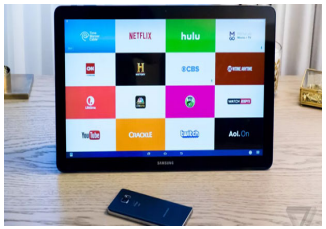
Mobile Video

First-person shooter games



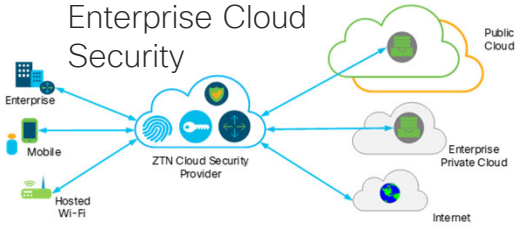
Fog computing

OTT Video on tablets and smart TVs



Mobile gaming

Enterprise Cloud Security



Any-Screen Video	Gaming and Virtual Reality	Services to Business
• CDN caching with multicast to unicast conversion • Programming convergence on app; QoE benefit	• Low latency • Delivered via SaaS model • Requires domain expertise and partners	• Designed for enterprises to outsource their security • Low latency • Campus networks

Select Use Cases for Enterprise Edge

Robotic Control

Advanced Inventory Management

Just-in-time logistics

Power plant operation

AR for jet engine maintenance

Industry 4.0	Warehouse Control	Monitoring and Maintenance
• Reconfigurable requires mobility of robots • 99.9999% (6-9) availability and packet loss ratio ~10^{-6}	• Low power tags & RFID integration to wireless • Data locality • Multicloud	• Sensor & actuator low-latency control loops • Distributed control • Mission critical

Target Customers

Network Operator Infrastructure

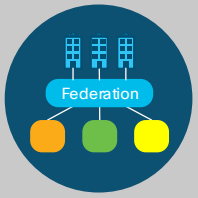
Target improved capabilities for the network orgs of operators to deliver on better CAPEX, OPEX, and QoE (example: 5G)

Mass Market Providers

Mass market use cases (Streaming video, gaming, AR, etc.) target consumer. Delivery is via B2B2C concept Examples: media companies, gaming

Enterprise Verticals

Enterprise verticals target edge use cases for improved operations. Example: municipal, auto vendors, IoT consumers, oil & gas, industry 4.0, …



Federations – Service Aggregators

Businesses that present a common interface to others. Example: CDN providers, public cloud providers, SaaS.

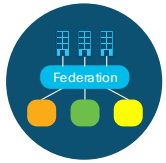
Achieving the commercial objectives requires "Network-as-a-Service"

Target B2B Consumption

Edge is about services to other businesses, internal or external, and whether ultimately targeting consumers, businesspeople or "things"

Support multi-tenancy

To ensure that each business customer has the desired level of isolation and "no shared fate" resiliency



Use a Federation

To address enterprise requirements for global "edge" footprint with a common "presentation" API interface catering to developers

Enable Multicloud

Enterprise customers of edge computing will operate "transaction intensive" workloads in the public cloud alongside edge workloads

Service Providers Can Play Multiple Roles in the Continuum

Network-as-a-Service Options

Colocation

Monetize the footprint and the IP connectivity

Bare Metal as a Service

Monetize compute and storage

IaaS/ CaaS

Monetize workload deployment at edge

PaaS / FaaS

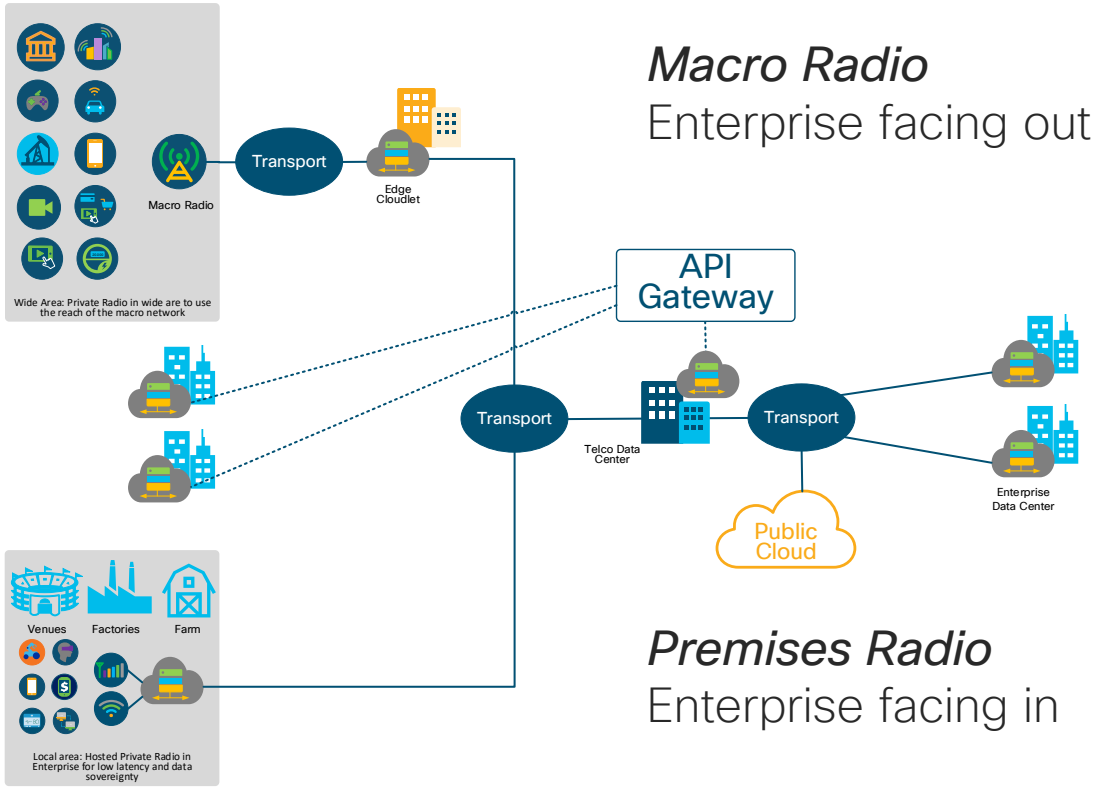
Monetize a development environment

SaaS

Monetize the use case

Comprehensive Automation Architecture for 5G

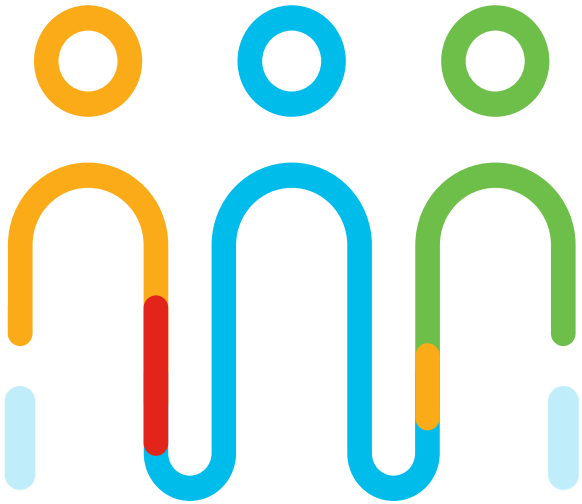
- Automation defines models for “slices” that can be allocated to specific use cases
- Automation also enables enterprise control of “protected” resources within the network
- This is an example of “Network-as-a-Service” in the context of 5G



Macro Radio
Enterprise facing out

Premises Radio
Enterprise facing in

Summary



Summary

- 5G end-to-end networking is rich in new technologies that broaden the appeal of the network
  - System architecture enhancements allowing edge compute, virtualization, and automation
  - RAN enhancements supporting new techniques for fast access
- While maintaining a focus on the consumer market, 5G also enables industry verticals
  - New opportunities for enterprise services & new product development
  - Powerful architectures at lower cost per bit

5G Optimizes Spectrum utilization + Supports New Use Cases


CISCO