# Transport Network Requirements and Architecture for 5G

Mikael Holmberg
Distinguished Systems Engineer

Metro Networks Architecture 2015

- Increased bandwidth
  - 1G to >100G
  - 100G Accelerates
- Dynamic architectures
  - L0 to L3
  - Fast moving towards SDN architectures
  - Fronthaul/Backhaul/FTTx
  - Cloud based services
- Cost optimized solutions
  - Purpose build hardware
  - White label boxes
  - Management solutions
- Scalability
  - Platform and commercial aspects
- Multi Layer Management
  - Ability to manage L0 to L3
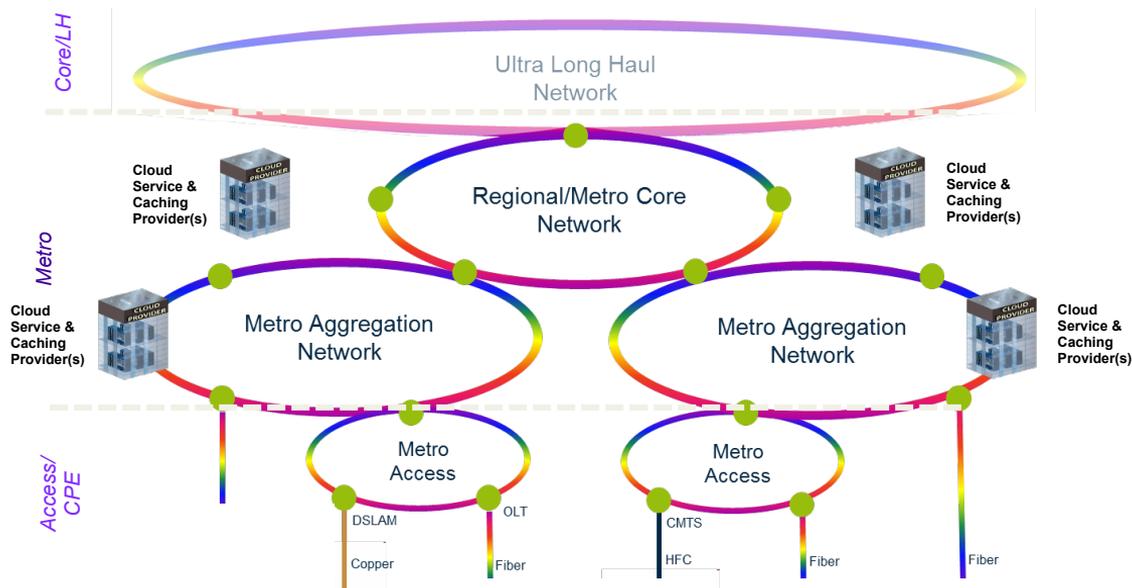
Mobile - 5G Requirements

LTE – min 10ms

4ms

4ms

Core Network

1-2ms

Internet

5-10ms
if in the same country
as the customer

5G service sub-1ms

> 0.5ms

Content
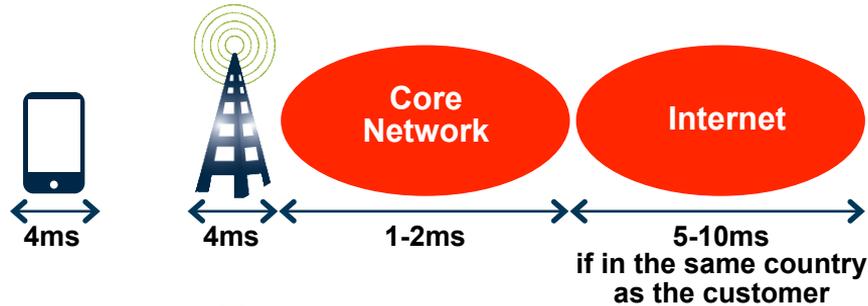
> 0.5ms
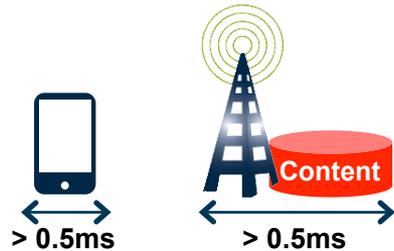
Source: GSMA

- 1-10 Gbps connections to end points in the field (i.e. not theoretical maximum)

- 1 millisecond end-to-end round trip delay (latency)

- 1000x bandwidth per unit area

- 90% reduction in network energy usage

5G Scenarios…

- Key Scenarios to be Addressed by 5G♪

Mission Critical Service
- Full reliability & high availability
- Real-time responsiveness
- On-the-fly coverage scalability for disaster situations

Fixed Broadband
- Next-generation broadband
- Multi-Gbps peak throughputs
- Alternative to costly fibre
- New VAS possibilities for fresh revenue generation

Massive IoT
- Connectivity for a new wave of device types
- High density deployments
- Networks-as-a-Service to meet each service provider's needs
- Robust QoE / QoS management
- New revenue opportunities

Mobile Broadband
- Multi-Gbps peak throughputs
- Universal gigabit connectivity
- Unparalleled mobility support
- New service / application enablement
- Advanced big data analytics

Ambulance

Autonomous driving

Vending machine

CCTV

## Requirements♪

- 10x bandwidth per connection
- Low-ms latency
- Five 9's reliability
- 100% coverage
- >10x connections
- 50Mbps per connection everywhere
- 1000x bandwidth/area
- 10 year battery life
- Reduction in TCO♪

Different Context of the same environment

| Requirements♪	Applications♪	Customer segments♪	MNO biz model♪
• 10x bandwidth per connection • Low-ms latency • Five 9's reliability • 100% coverage • >10x connections • 50Mbps per connection everywhere • 1000x bandwidth/area • 10 year battery life • Reduction in TCO♪	• Enhanced Mobile BB • Connected vehicles • AR/VR • S-UHD/3D Video • Haptics/Sensing • Massive IoT • Remote machine control • Mission critical services • Fixed-wireless access • …	• Consumer • Auto industry • Health • Industry 4.0 • Agriculture • Smart City/Public sector • Smart building • Utilities • Education • Transport • …	• B2C • B2B • B2B2C♪

5G Service Enablers

mmWave System/RFIC/Ant.♪

700 MHz♪ 3 GHz♪ 18♪27♪30 GHz♪

Legacy Bands New Bands

mmWave RFIC♪ Wide Coverage Antenna♪
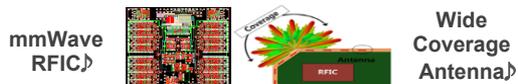
New Channel Coding
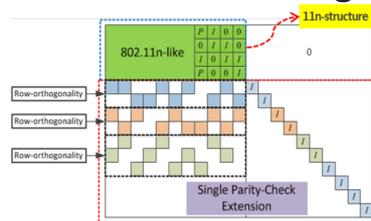
LDPC (Low-Density Parity-Check)

Network Slicing♪

MBB
MTC
UR/LLC

Global Mobility/Delay-Tolerant
Low Latency

5G Core VNF
MBB Server
MTC Server
Central Data Center

5G Core VNF
UR/LLC Server
Local Data Center

< 6 GHz Massive MIMO♪

Half Wavelength♪
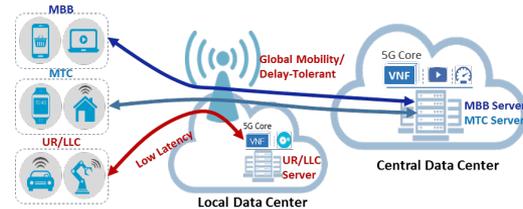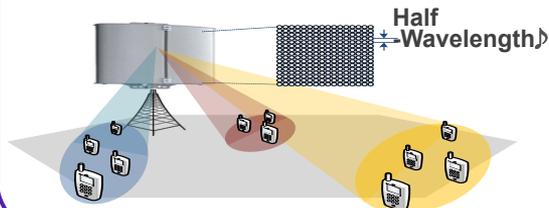
Massive Connectivity (IoT)♪

Grant-based Multiple Access♪

3~4 Step♪
eNB♪
UE♪

Grant-Free Multiple Access♪

1 Step♪
eNB♪
UE♪

Low Latency NW♪

① Radio Information
② TCP Rate Control

Mobile
Data
BS
Server
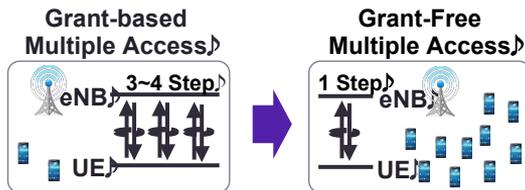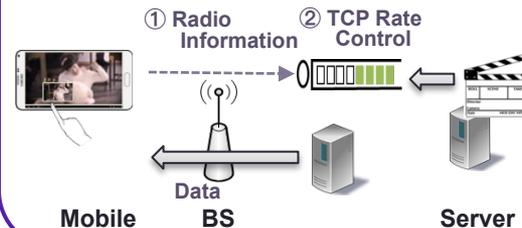
5G Topology Flexibility…?
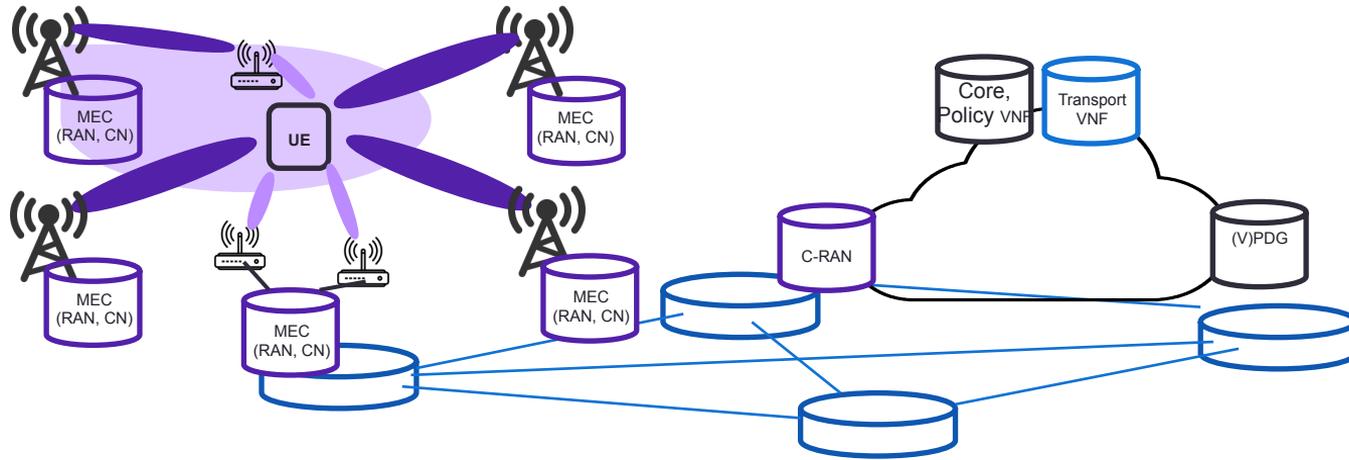
'Softwarisation' of the network

C-RAN – removal of functionality from cell sites to consolidation point in the network

NFV and SDN – enabling flexibility in where functions are deployed and scaled

MEC – pushing Core Network functions and content ingress to cell sites

CP/UP split – decoupling of user plane traffic from control plane functions

5G Network Slicing

Orchestration Layer

RAN Orchestration | CN Orchestration | Transport Orchestration

Inter-orchestration system interface

UE

MEC (RAN, CN)

CN, Policy VNF | Transport VNF | CN, Policy VNF

VNF

C-RAN

(V)UPF

Orchestration (Network **Slicing**)

Enterprise Customer (or SI)

Data Centre | Apps | VNF vCPE

2G, 3G, 4G Slice

NB-IoT, LTE-M slice

Wi-Fi Slice

Fixed Line Slice
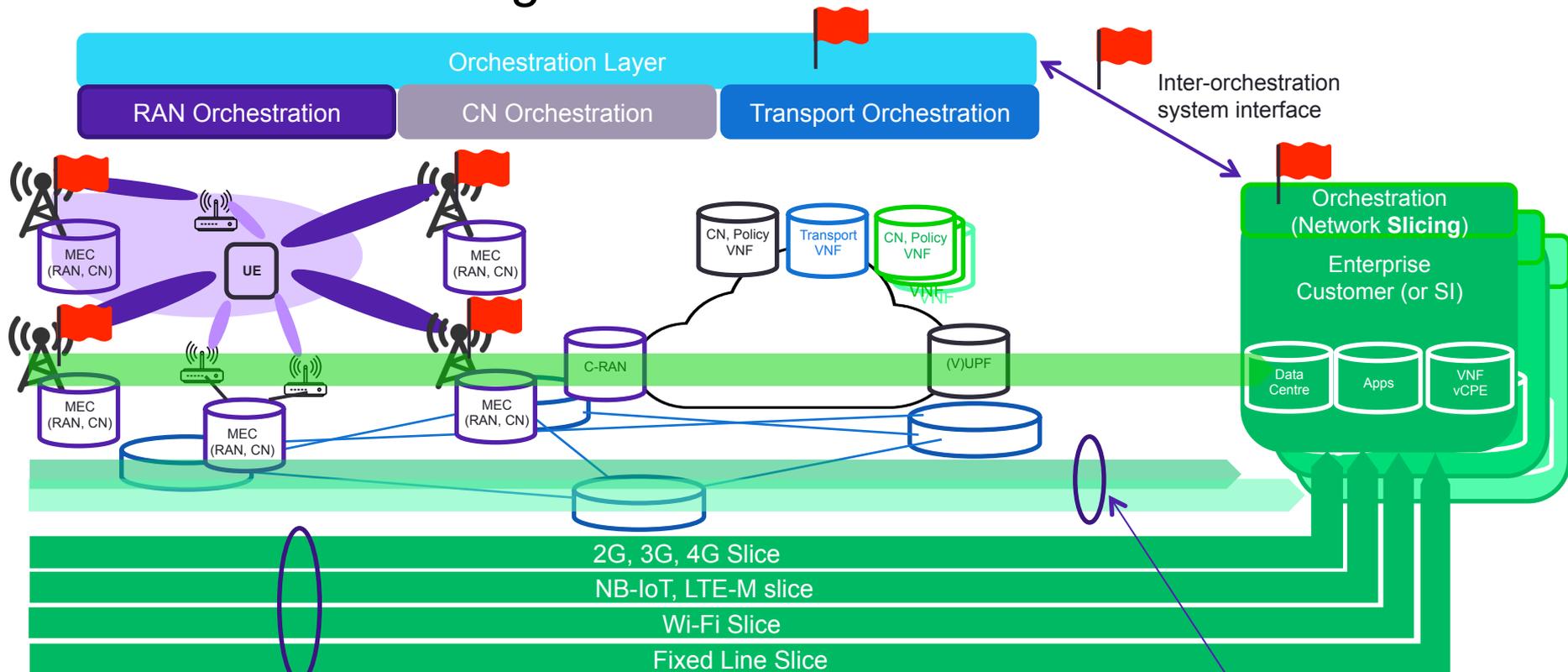
Potentially multiple other network slices per network customer

One (or more) 5G slice per enterprise customer

5G Transport is More Than a Necessary Cost…

- It is an investment in critical assets that enables customers to interact with services

- It is an architecture that will underpin the long-term development of the 5G commercial offer

- Part of a "universal edge" based on packet/optical integration, SDN, and etc, etc.

4G to 5G Network Evolution Challenges

- Latency budget
  - FH: 4G is 75µs to 100µs while ultra low latency requires under 50µs
  - MH: New for 5G @ 1ms to 3ms max
  - BH: Unchanged
- Connectivity
  - 4G using CPRI is P2P and P2MP
  - 5G is Cloud based and will also need MP2MP
- Synchronization
  - 4G using GPS
  - 5G using PTP and SyncE
- Massive MIMO transport rate impact

5G RAN Model

	Backhaul		Midhaul		Fronthaul	
Core	Carrier Ethernet	CU	Low latency Ethernet	DU	RoE/eCPRI	RU
	10ms		1-3ms		100µs or 50µs for URLLC	

4G Central RAN Model

	Backhaul		Fronthaul	
MSC	Carrier Ethernet	BBU	CPRI	RRH
	10ms		100µs	
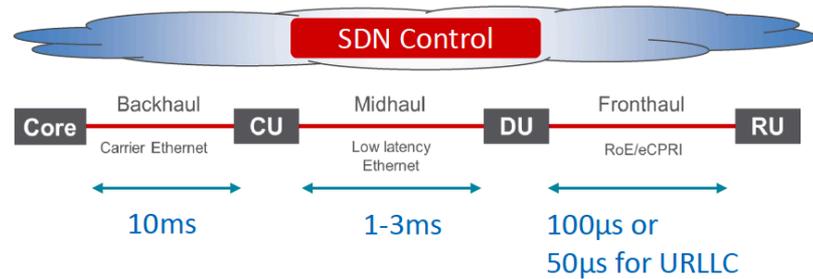
Max latency values based on IEEE 1914.3

TSN (Time Sensitive Networking)

- TSN is part of the 802.1Q family of standards and is designed to provide **deterministic messaging on standard Ethernet networks**

- Important to many industries – e.g. aerospace, automotive, transportation and utilities, and of course, manufacturing – where TSN is emerging as the baseline for real-time networking

- A wide variety of applications and users, contributing scale to the technology relative to a mobile-only standard

- A common X-Haul transport network for 5G RAN, serving fronthaul, mid-haul and backhaul, can therefore be developed using TSN

What is TSN…

- IEEE 802.1CM standard
- TSN enables reliable, deterministic real-time communication over Ethernet (RTCoE)
- Reliable, deterministic RTCoE is achieved using
  - Time synchronization
  - Scheduler shared between network components
  - Frame replication
- Offers
  - Bounded low latency
  - Low delay variation
  - Extremely low loss

Time-Sensitive Networking

TSN Components

Time sync:
Timing and Synchronization

Synchronization

Reliability

Latency

Resource Mgmt

Ultra reliability:
Frame Replication & Elimination
Path Control
Per-Stream Filtering & Policing
Time sync reliability

Bounded low latency:
Credit Based Shaper
Frame Preemption
Scheduled Traffic
Cyclic Queueing & Forwarding
Asynchronous Traffic Shaping

Dedicated resources & API
Stream Reservation Protocol
TSN configuration
YANG
Link-local Reservation Protocol

Zero congestion loss

Source: IETF

Why TSN for 5G RAN?

- Features such as zero congestion loss, bounded ultra-low-latency, hitless 1+1 redundancy, and the ability carry sync accurate to <1µs, make TSN an attractive candidate for 5G RAN transport.

  "802.1CM TSN for Fronthaul" is a collaborative effort by the CPRI and IEEE organizations.

- Two profiles are under development
  - Profile A sends user data (IQ data) as high priority traffic
  - Profile B uses "frame preemption", a TSN feature to prioritize different traffic types (sync, user data, management data, etc.)

5G Transport Platforms and Technology

- Fronthaul @ 5 µsec to 100 µsec
  - TSN
  - WDM active and passive
  - Dedicated dark fiber

- Midhaul @ 1ms
  - OTN
  - WDM
  - Low latency Ethernet

- Backhaul @ 10ms
  - Carrier Ethernet

Optical Transport

| DU | —— | RU |

| CU | —— | DU |

| Core | —— | CU |
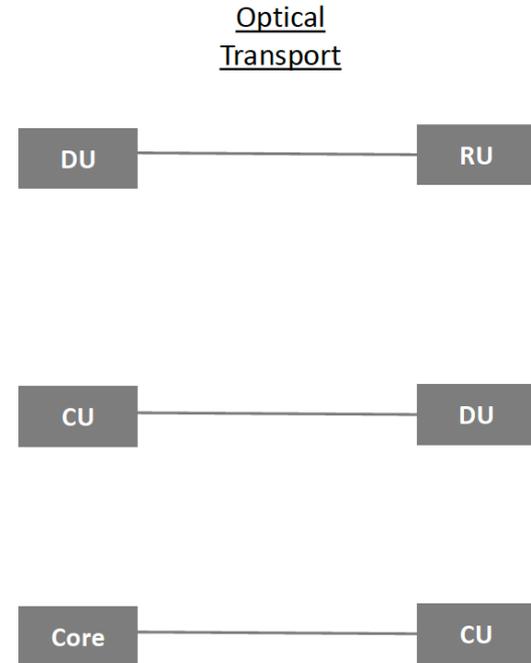
5G Model Driven Automation

- The new metro-edge network will be SDN-based, programmable, and open to multi-domain orchestration
- Automation is especially important at the edge because
  - there will be significantly greater number of functions to manage; and
  - equipment may be deployed in locations that are not staffed
- Model-driven, intent-based configuration is of primary importance to lowering opex
- NETCONF/YANG gaining broad-based support as operators migrate away from CLI-based device management to API-driven devices

xRAN Adopts NETCONF/YANG

YANG based Management-plane for 5G RAN

Source: xRAN Forum
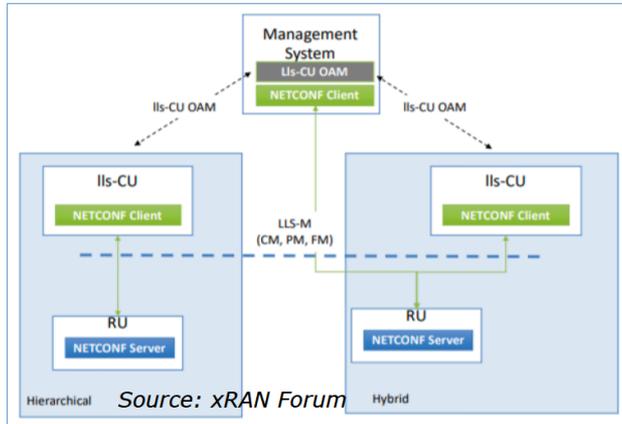
- Programmatically configure and manage RUs (in LLS architecture)
- Simplify integration between CU and RU, even (especially!) in multi-vendor scenarios
- Model the state of 5G RUs, enabling programmatic control of radio resources across a coverage area
- Lays the foundation for cross-domain orchestration of the RAN with other domains that have already adopted NETCONF/YANG

With a full mesh packet network architecture, cross domain control will allow operators to support end-to-end services, with the appropriate resources allocated in the radio, core, and transport.

Overview of Synchronous Ethernet
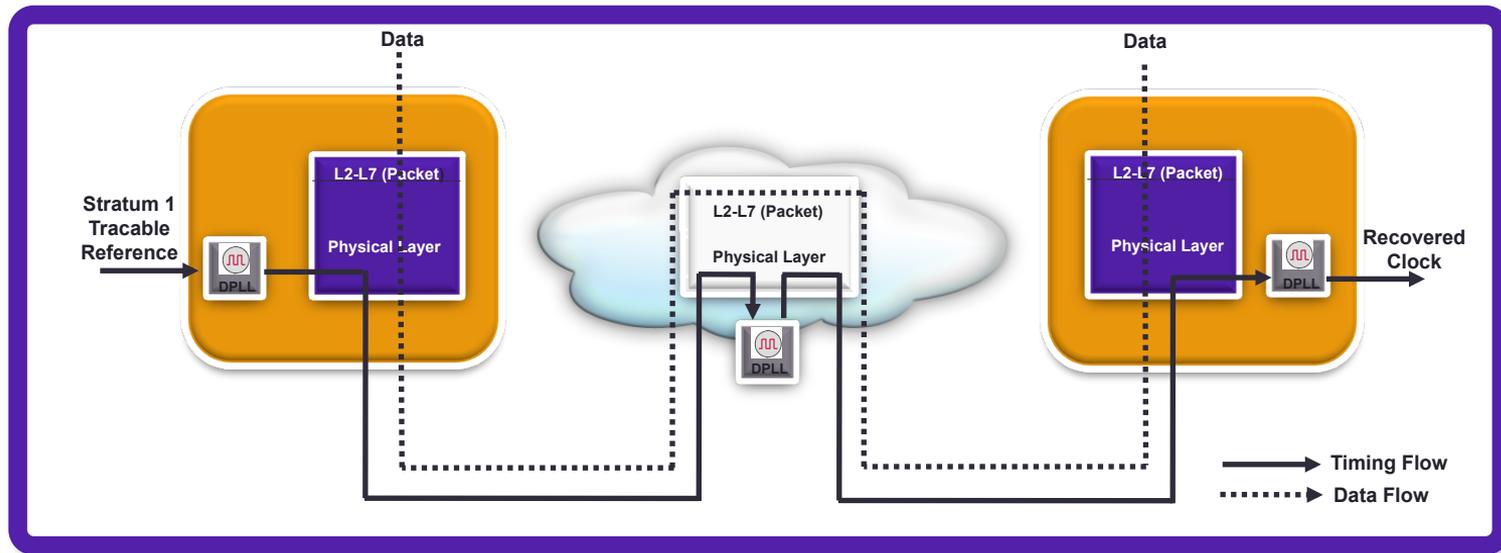
- Use the PHY clock
  - Generates the clock signal from "bit stream"
  - Similiar to traditional SONET/SDH/PDH PLLs
- Each node in the Packet Network recovers the clock
- Performance is independent of network load
- Uses the OAMPDUs (slow path protocol data units) to pass SSM (synchronization status message).

PTP (1588v2) on Ethernet

- The IEEE 1588v2 Precision Time Protocol (PTP) defines a packet-based time synchronization method that provides frequency, phase and time-of-day information with sub-microsecond accuracy. PTP relies on the use of carefully time stamped packets to synchronize one or more slave clocks to a master clock. Synchronous time information is distributed hierarchically, with a grand master clock at the root of the hierarchy. The grand master provides the time reference for one or more slave devices. These slave devices can, in turn, act as master devices for further hierarchical layers of slave devices.
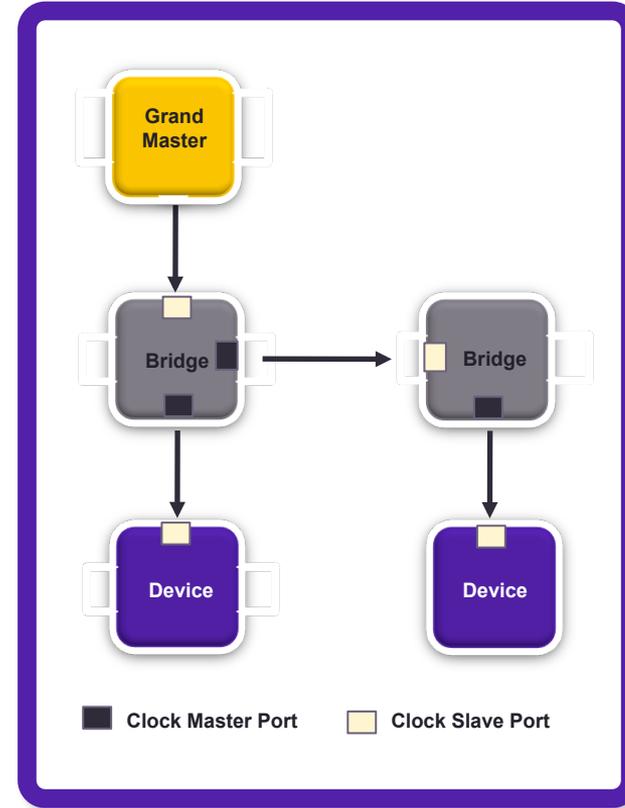
Precision Time Protocol (1588v2)

PTP Overview

- IEEE 1588-2008
  - Also known as Precision Timing Protocol (PTP)
  - 1588-2008 is also referred to as version 2 (v2)
  - Intended to synchronize independent clocks to a high degree of precision on separate nodes over a distributed network
  - Defines how to transfer precise time over networks. It does not define how to recover frequency or high precision time of day

- Message Exchange Pattern
  - A) The master sends a Sync message to the Slave and notes the time t1 at which it was sent.
  - B) The Slave receives the Sync message and noets the time of reception t2.
  - C) The Master conveys to the Slave the timestamp t1 by either
    - Embedding the timestamp t1 in the Sync message. This requires some sort of hardware processing for highest accuracy and precision. This is called as one step clocking.
    - Embedding the timestamp t1 in a Follow_Up message. This is called as two step clocking.
  - D) The Slave sends a Delay_Req message to the Master and notes the time t3 at which it was sent.
  - E) The Master receives the Delay_Req message and notes the time of reception t4.
  - F) The Master conveys to the Slave the timestamp t4 by embedding it in a Delay_Resp message.

Basic Synchronization message exchange
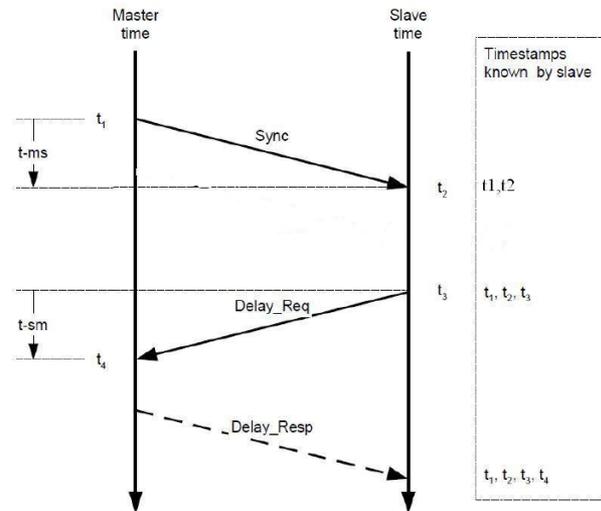
Extreme Product Portfolio

EXOS Series
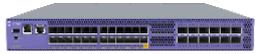
Extreme X460/X460-G2

Extreme X670/X670-G2

SLX Series

SLX 9140

SLX 9240

SLX 9560

Workflow Composer
Powered by StackStorm

Workflow Composer
Automation Suites

Entry Into
Automation

Begin
Customization

Cross-domain
Automation

Embedded Fabric
Automation

python

Management
Center

ANSIBLE

RESTful API
GET PUT POST DELETE

Automation
Suites

Workflow
Composer

Network Automation

Network Automation

IT Automation

5G High Level Network Architecture
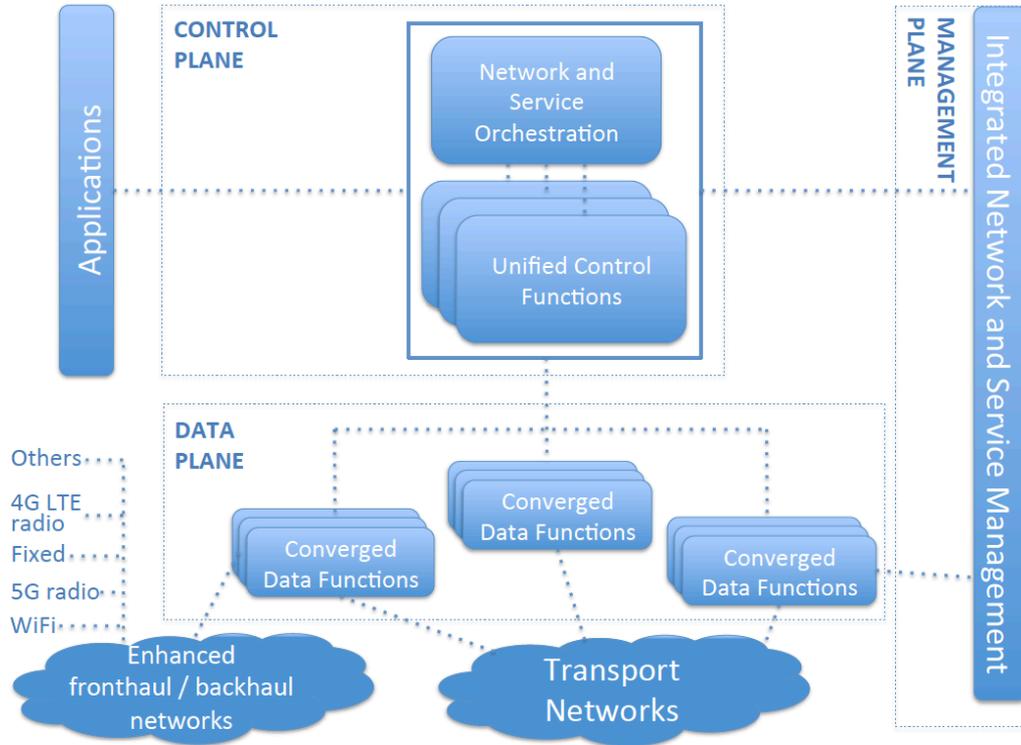
Source: ITU, 3GPP

Thank You

WWW.EXTREMENETWORKS.COM