FMA Cloudification: Methods and Architecture Patterns

A Technical Paper prepared for SCTE by

Jim Huang
Principal Solutions Architect
Telecom Industry Business Unit, Amazon Web Services
2795 Augustine Drive, Santa Clara, CA 95054
jimhuan@amazon.com

Jeff Finkelstein
Chief Access Scientist
Cox Communications
Atlanta, GA
Jeff.Finkelstein@cox.com

Table of Contents

Title **Page Number**

List of Figures

Title **Page Number**

List of Tables

Title **Page Number**

1. Introduction

Over the past decade, Cable operators and communication service providers have been striving towards Distributed Access Architecture (DAA)[1], which disaggregates Converged Cable Access Platform (CCAP) systems for cable network scaling, operation simplification and cost reduction, as well as resource and space savings at the HFC headend or hub location. DAA deployment also provides a platform for cable operators to continually integrate virtualization elements into their networks for service velocity.

The Flexible MAC Architecture (FMA) specification[2] released by CableLabs in 2020 is the latest undertaking for DAA evolution. FMA compartmentalizes DAA management plane, control plane, and data plane functions with APIs for vendor product interoperability. It accommodates DAA Remote PHY (R-PHY) and Remote MAC PHY (R-MACPHY) system architectures for flexible DAA solution variants. FMA access technologies include xPON and 4G/5G wireless for evolving last-mile access architectures. The FMA specification further calls out software defined networking (SDN) and network function virtualization (NFV) for future implementation with DAA.

Today, FMA is undergoing vendor product implementation and multi-vendor interoperability demonstrations. But what will FMA NFV look like end to end? How will the FMA NFV be implemented? Will FMA be further optimized for a cable operator’s service agility and cost reductions? These questions are all open for exploration and development.

This paper proposes FMA cloudification to advance FMA along its evolutionary path. Many communication service providers (CSP) in the telecommunication industry have recognized Cloud as a mechanism for lowering operator’s total cost of ownership (TCO) and gaining operations efficiency. Some CSPs have started migrating their network infrastructures as well as IT workloads to Cloud. Challenges in supporting virtualized network functions (VNFs) in Cloud are often around real-time constraints, high availability, and operation automation. This is particularly true for DOCSIS like products.[3] Our approach to FMA cloudification is to properly map FMA functions across cloud region, cloud edge, and on-premise environments based on their latency tolerance characteristics and access network technology requirements. Specifically, we allocate FMA control plane functions such as MAC Manager and management plane functions such as Operations Support System (OSS) in the cloud region and control functions of auxiliary cores at the cloud edge, leaving latency-sensitive Remote MAC Devices as well as last-mile access networks AS IS in the field. With this MFA cloud architecture, CSPs can focus their capital and resources on the last-mile technology and deployment into a particular geographic area while simplifying network provisioning and management processes with the cloud.

In the following, we briefly review the FMA architecture in Section 2. Section 3 details our approach to FMA cloudification by introducing a Cloud Continuum model and positioning FMA functions along the cloud continuum based on timing analysis and FMA architecture. In section 4, we describe how the cloud FMA can be realized through a cloud infrastructure and address FMA high availability and deployment automation using cloud-native services. We further demonstrate the FMA cloudification with cloud-native services for an FMA use case – streaming telemetry in Section 5. We conclude the paper in section 6.

2. Flexible MAC architecture overview

The FMA concept began after a number of discussions among technical staff from Cox Communications and the AT&T Foundry as they worked on what became known as "CORD" (Central Office Reimagined as a Data center). FMA was originally called "HERD" (Head-End Reimagined as a Data center), but as it transformed from an architecture for simply managing remote MACPHY devices to an agile service delivery platform that included its original intent of Remote MACPHY Device (RMD) interoperability, the decision to rebrand it was made.

FMA was designed from the beginning with the idea of all components being able to work in a physical and virtual environment. The legacy back-end software, MAC manager, SDN and DOCSIS controllers, and auxiliary cores, would be designed to be placed at the point in the compute network that was consistent with the architectural demands of each individual operator. The interfaces between FMA components were standardized which would allow for the decomposition of functions and their strategic placement in the operator network or public cloud as desired.

The core objectives of the FMA working group are as follows:

1. Define interfaces between a management entity and DOCSIS MAC network element such that the management entity from Vendor A may be interoperable with a DOCSIS MAC network element from Vendor B.
2. Define a DOCSIS MAC network element that contains all necessary DOCSIS MAC layer functions such that a CCAP core network element is not needed for data plane forwarding of customer data traffic.
3. Define a DOCSIS MAC network element for both physical network functions (PNFs) and virtual network functions (VNFs) used for data plane forwarding.
4. Define an architecture such that typical CCAP functions such as management plane, L2/L3 control plane, and DOCSIS control plane may be separate from the DOCSIS MAC network element.
5. Define an open standard interface between operator BSS, OSS, NMS, and orchestration with the management entity or entities which may include management plane, L2/L3 control plane, and DOCSIS control plane used to operate the DOCSIS MAC network element.
6. Leverage previous cable industry specifications as appropriate as well as consider the transition to next generation approaches used in the areas of software defined networking (SDN), network function virtualization (NFV), VNFs, and PNFs, new protocols, new data models, and new telemetry methods as desired by the R-MACPHY MSO steering committee and working group.

Towards the objectives, the FMA working group developed and published the first FMA System Specification in 2020 and since has made two revisions.[2] A detailed view along with the connections between the FMA components can be seen in Figure 1 below.

The FMA architecture will evolve in three phases.

Phase 1 – Support key components necessary to bring RMDs into existing operator networks without causing disruption to legacy EMS and NMS platforms

Phase 2 – Allow for continued use of Remote PHY devices in the FMA network management paradigm

Phase 3 – Add support for the Remote MAC Core concept with physical and virtual MAC support

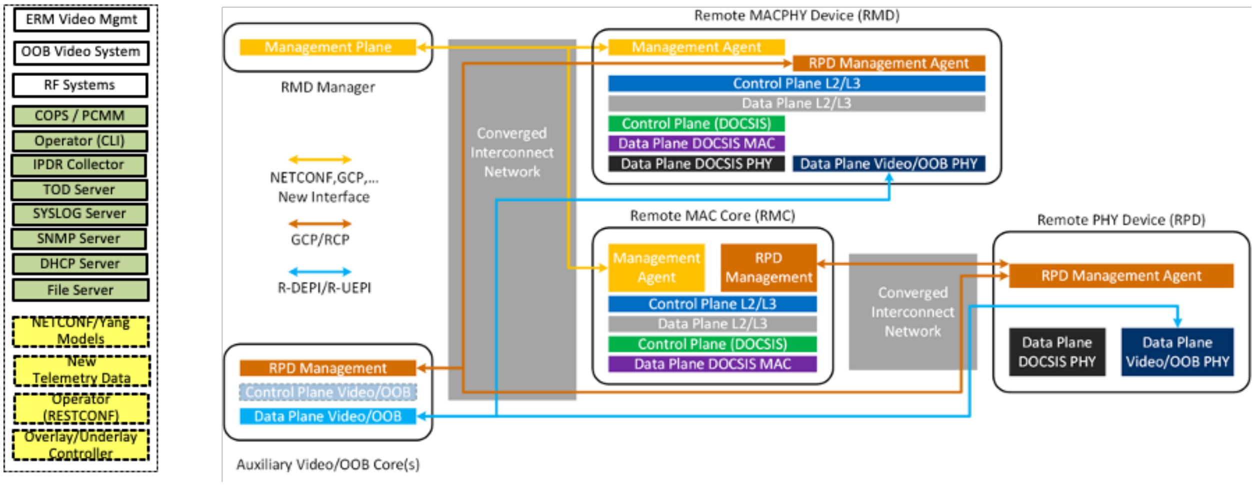

Figure 1 – FMA Details

Figure 2 shows the FMA Phase 1 reference architecture with components and their functions and interfaces. For details, refer to the FMA System Specification[2].

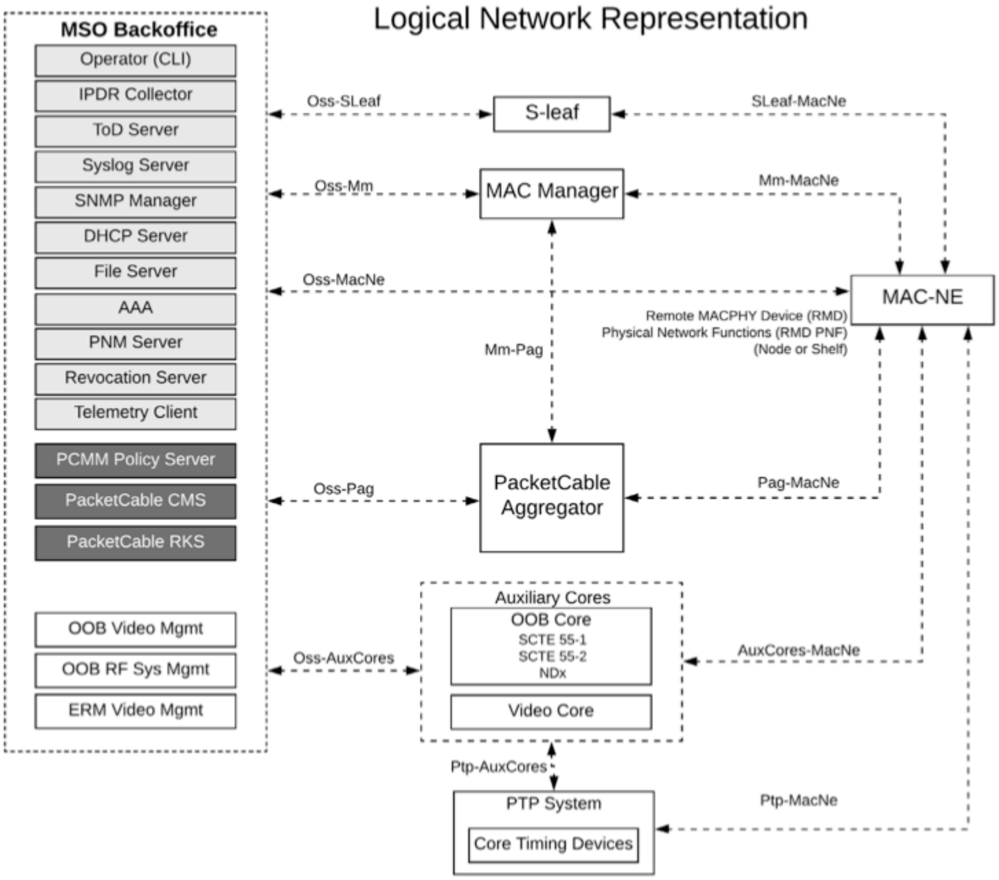

Figure 2 – FMA Phase 1 Reference Architecture

This paper introduces how to take Cloud advantages for business agility, operation efficiency, modern application services, and cost savings by running the components of the Flexible MAC Architecture in a cloud environment. The paper focuses on the FMA Phase 1 reference architecture for cloudification.

3. Approach to FMA cloudification

The FMA cloudification builds on a cloud infrastructure model, called Cloud Continuum. Typical cloud workloads such as web applications are centered around Cloud Region which consists of multiple inter-connected data centers. The Cloud Continuum extends the cloud region to cloud edge for time-sensitive network functions and applications. It enables multi-access edge computing (MEC) for network operators' business growth. The Cloud Continuum is further out to the far edge with cloud services like IoT for managing physical devices in the field.

Figure 3 shows the Cloud Continuum model with Cloud Regions, Cloud Edge, and Cloud Far Edge. They are inter-connected and have the same "look and feel" for cloud infrastructure services such as networking, compute, storage, CLI commands, and deployment automation.

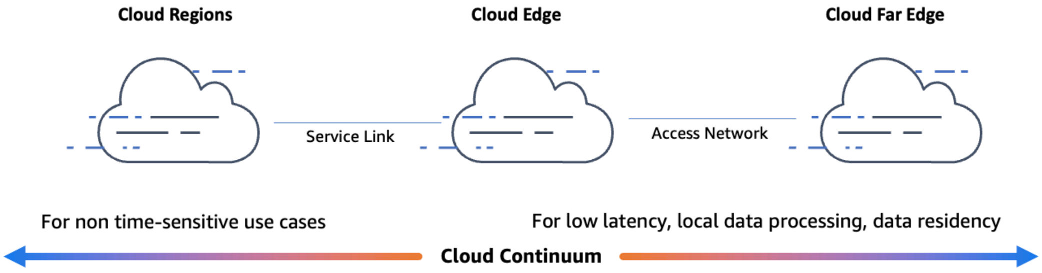
Cloud Regions
Cloud Edge
Cloud Far Edge
Service Link
Access Network
For non time-sensitive use cases
For low latency, local data processing, data residency
Cloud Continuum

Figure 3 – Cloud Continuum: Region, Edge, and Far Edge

With the Cloud Continuum model, FMA is overlayed across the cloud region, cloud edge, and cloud far edge in on-premises environments. This is accomplished through analyzing the real time characteristics and operation locality of FMA systems[2] and accordingly allocating FMA functions along the Cloud Continuum and remote MAC network elements (MAC-NEs). Table 1 summarizes the analysis with respect to the degree of latency tolerance per major FMA functions, where High means the latency requirement is greater than 50ms roughly, Medium is in-between 20ms - 50ms, Low in-between 5ms - 20ms, and Ultra Low is less than 5ms. It also shows the operation locality of an FMA system or function, i.e., Backoffice, Headend, Hub, or MAC-NE.

Given the timing constraints and operation locality, the FMA functions are now allocated across cloud region, cloud edge, access network, and MAC-NEs. Figure 4 illustrates the FMA System Architecture[2], where the upper portion shows the management and control plane functions allocated in Data Center, Headend, Hub, and Optical Node and the lower portion depicts the data plane.

The FMA System Architecture can be cloudified in a cloud architecture pattern shown in Figure 5. The FMA functional allocation and architecture cloudification scheme is as follows.

Table 1 – Analysis of FMA Function Timing Characteristics and Operation Locality

FMA System	FMA Functionality	Plane	Latency Tolerance*	Operation Locality
MSO Backoffice	OSS - IPDR Collector, Syslog Server, SNMP Manager, DHCP Server, File Server, AAA, PNM Server, Telemetry	Management	Medium - High	Backoffice
	PacketCable management - PCMM Policy Server, PacketCable CMS, PacketCable PKS	Management	High	Backoffice
	Auxiliary Core Management - OOB Video Mgmt, OOB RF Sys Mgmt, ERM Video Mgmt	Management	High	Backoffice
MAC Manager	Manage MAC-NE through FMA protocols	Management	High	Backoffice or Headend
PacketCable Aggregator	Scale PacketCable Multimedia Backoffice element to RMDs	Management	High	Backoffice or Headend
OOB Core	Manage and control legacy set-top boxes (STBs) as RMDs in substitution of SCTE 55-1 and SCTE 55-2 OOB Cores	Management Control	Medium Low	Headend or Hub
NDx Core	Support Narrowband Digital Forward (NDF) and Narrowband Digital Return (NDR) capabilities	Control Data	Low Low	Headend or Hub
Video Core	Provide video EQAM processing functions, except video PHY and QAM-related processing in RMD	Control Data	Low Low	Headend
PTP System	Clock synchronization across RMDs	Control	Low	Headend
MAC-NE (RMD)	DOCSIS MAC and the upper layer protocols, QAM, digital to/from analog conversion for RF/Ethernet or PON transmission	Management Control Data	High Low Ultra low	Hub or Node

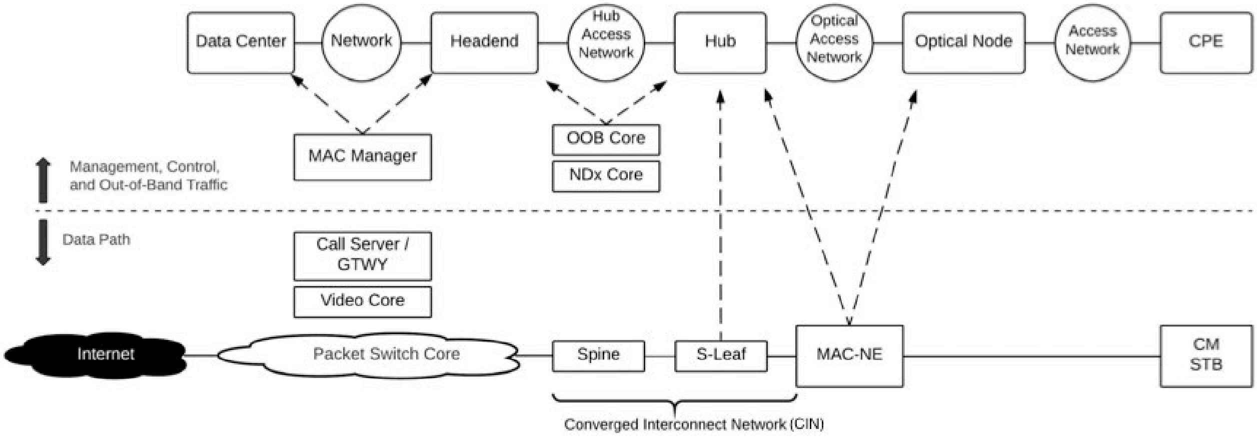

Figure 4 – FMA System Architecture

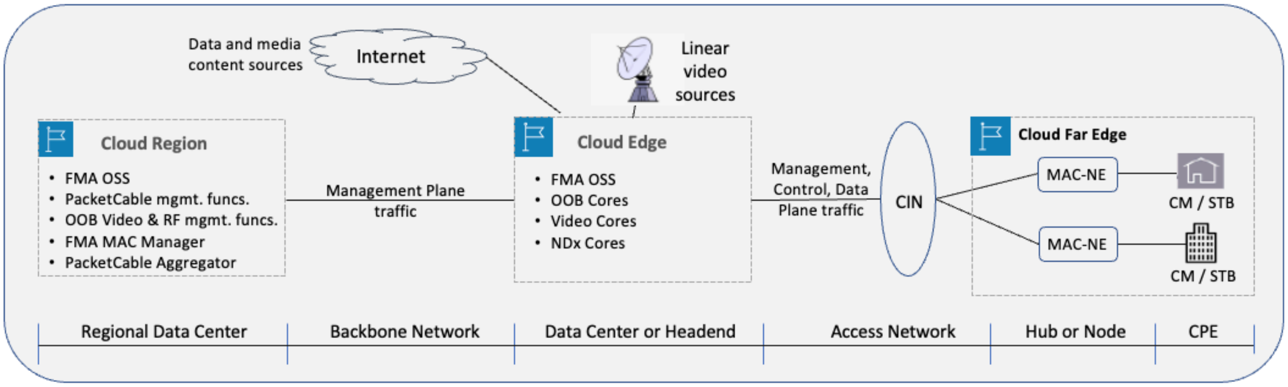

Figure 5 – FMA Cloudification Architecture Pattern

FMA Management Plane consists of FMA OSS functions, MAC Manager, PacketCable Aggregator, PacketCable management functions, and OOB Video & RF management functions. They are High in latency tolerance. Mirroring Multi-System Operator (MSO) Backoffice at regional data centers, MAC Manager, PacketCable Aggregator, most of OSS functions, and PacketCable management functions are placed in the cloud region. The OSS functions such as streaming telemetry that deal with a large amount of data can be allocated at the cloud edge to localize data processing, filtering, and analytics. MAC Manager and PacketCable Aggregator can run at the cloud edge as well because of the operation locality.

FMA Control Plane involves auxiliary Cores, OOB Core, NDX Core, and Video Core, under the FMA Phase 1 Reference Architecture shown in Figure 2. Since they are Low in latency tolerance, they run at the cloud edge. The FMA control plane functional entity resides in MAC-NE for processing L2/L3 control plane traffic. For example, the MAC-NE DOCSIS Control Plane functional entity processes control plane traffic from the PacketCable Aggregator. Hence, the FMA control plane functional entity in MAC-NE is allocated at the cloud far edge.

FMA Data Plane is comprised of data plane functions of MAC-NE (Remote MACPHY Device) as well as NDx Core and Video Core systems. As physical device, MAC-NE performs RF conversion with Ultra Low data process latency outside of cable plant. In addition, MAC-NE may run IoT functions for Proactive Network Maintenance (PNM). Hence, MAC-NE is at the cloud far edge. NDx Core and Video Core data plane functions are Low in latency tolerance and thus placed at the cloud edge.

FMA Networks are comprised of Packet Switched Network Core and Converged Interconnect Network (CIN). The Network Core is for communications between FMA data centers and headend systems as well as Internet traffic. CIN consists of Spine switches and Secure Leaf switches (S-Leaf), connecting MAC-NE (RMD) instances in a hierarchical structure. In the cloud environment, the infrastructure backbone network and SD-WAN in-between the cloud region and cloud edge replace the FMA Network Core, thus relieving FMA operators from the core network management. CIN remains as it is in the field for network technology as well as vendor deployment flexibility.

FMA MAC-NE (RMD) and CPE are vendor and customer choices. Cloud IoT can be added to RMD and CPE for MSO to provide advanced services such as device control, fault diagnostics, or preventive network maintenance which the FMA System Specification[2] describes.

In summary, the FMA cloudification builds on the cloud continuum infrastructure with FMA functions allocated in the cloud region, cloud edge, and cloud far edge according to their timing characteristics and operation needs.

4. Cloud infrastructure and services for FMA cloudification

This section describes how the FMA cloudification can be realized in cloud. It uses Amazon Web Services (AWS) as an example of a cloud infrastructure. Please note that similar architectures can be built with other cloud providers.

4.1. Cloud edge services and selection

The cloud continuum is comprised of different types of cloud edge technology. AWS provides three types of cloud edge services: Outposts, Wavelength, and Local Zones as shown in Figure 6. Though all provide cloud compute, storage, and networking services for real-time, short-latency or high-throughput virtualized network functions (VNFs) and applications, they are architected for different use cases. MSOs or cable operators need to examine and select a right type of cloud edge technology and services for FMA cloudification.

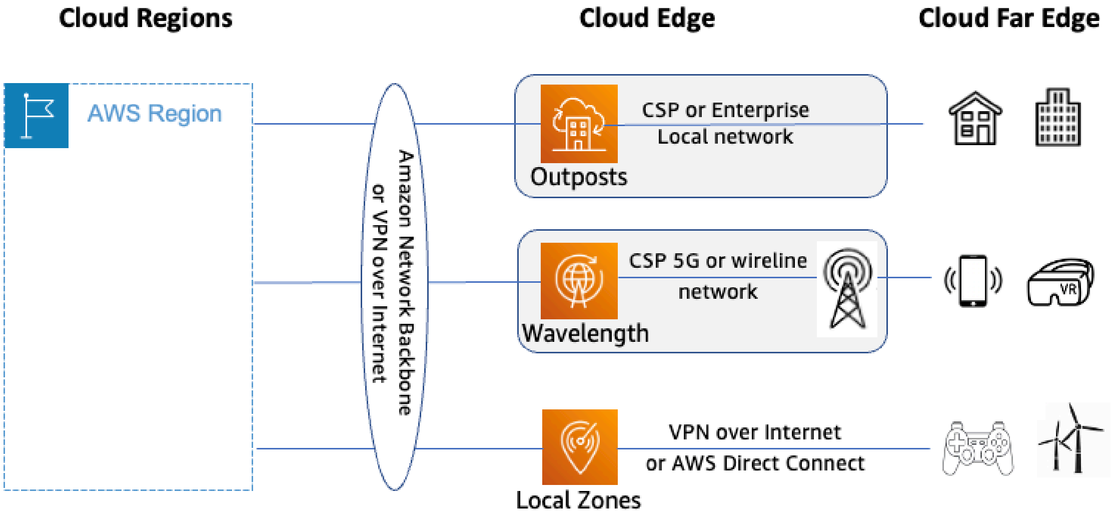
Cloud Regions
Cloud Edge
Cloud Far Edge
AWS Region
Amazon Network Backbone or VPN over Internet
CSP or Enterprise Local network
Outposts
CSP 5G or wireline network
Wavelength
VPN over Internet or AWS Direct Connect
Local Zones
VR

Figure 6 – Types of Cloud Edge Services

Table 2 summarizes the characteristics and use cases of these three types of cloud edge services. Although both AWS Outposts and Wavelength are embedded in CSP networks, the former is dedicated to an operator or enterprise, thus called "private edge cloud," whereas the latter is shared by different enterprises, organizations, or users, called "public edge cloud." That is, the Outposts is for private use and the Wavelength is for public. AWS Local Zones is fully managed by AWS with end-user connectivity through CSP wireline or wireless services or enterprise's direct connection. Hence, the private edge cloud service (Outposts) is suitable for FMA in MSO or cable operator networks. MSO deploys and manages FMA systems and functions in the cloud.

Table 2 – Comparisons of Cloud Edge Services

Type of Cloud Edge Service	Cloud Edge Architecture & Deployment location	Use Case	Type of CSP	Management Responsibility
Outposts	On-premises networks or in CSP networks	• Private Multi-Access Computing (e.g., smart factory, healthcare operations, real-time ML and analytics) • Private Mobile Network • MSO network	• Telecom operators • MSOs	• Cloud infra. & cloud services: AWS • Local network: CSP or on-premises operator
Wavelength	In CSP networks, especially 5G mobile networks	• Public Multi-Access Computing (e.g., intelligent vehicles, real-time retails)	• Mobile network operators	• Cloud infra. & cloud services: AWS • Local network: CSP
Local Zones	In AWS infrastructure	Public Multi-Access Computing (e.g., real-time gaming, interactive live video streams)	N/A	• Cloud infra. & cloud services: AWS • Local Zones connectivity: CSP or enterprise

4.2. FMA cloud architecture

Figure 7 depicts an FMA cloud architecture with AWS Region and Outposts-based cloud edge integrated with CIN and MAC-NEs at the far edge.

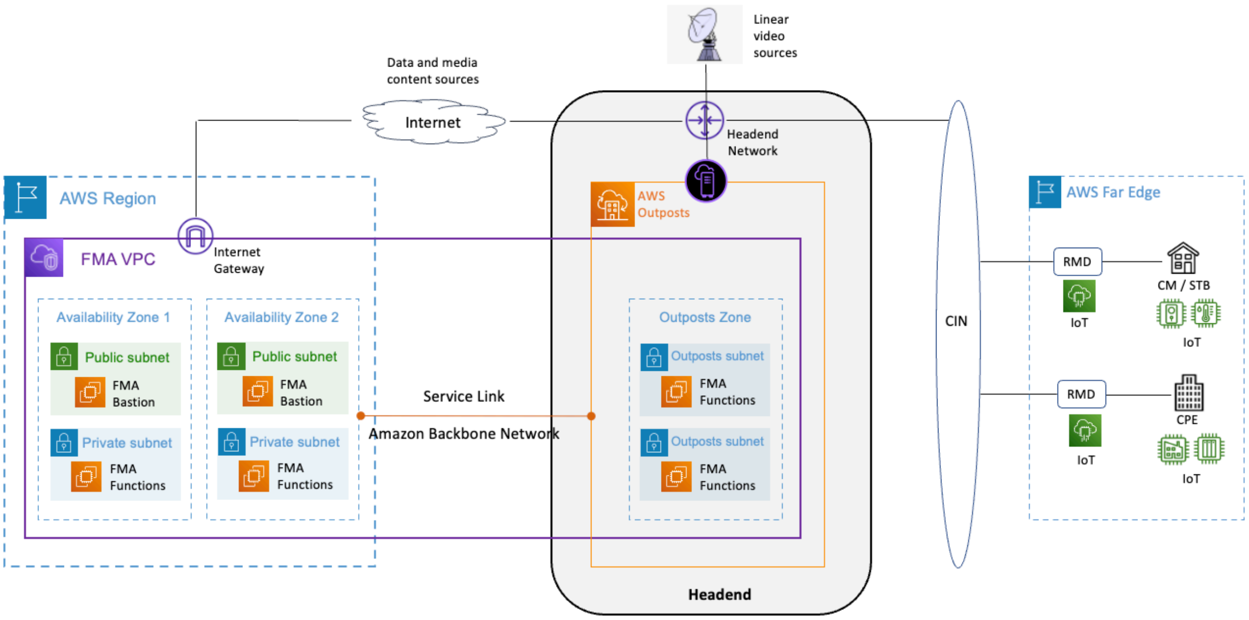

Figure 7 – FMA on AWS Cloud

The FMA cloud portion is comprised of four building blocks:

- *AWS Region* – is a physical location, which is comprised of multiple Availability Zones for high availability and high scalability. Each Availability Zone has multiple datacenters housed in physically separated facilities with redundant power, networking, and connectivity. A region hosts the high latency-tolerance FMA management plane functions such as OSS and PacketCable management.

- *AWS Edge Outposts* – are cloud edge platforms in three form factors, 42-RU, 2-RU, and 1-RU. The platforms provide cloud-native services such as Elastic Cloud Computing Service (EC2) for virtual machines and Elastic Kubernetes Service EKS) for containers. The Outposts run FMA functions that are low latency tolerance (e.g., auxiliary Cores), require intense local data processing and analytics (e.g., OSS telemetry), or need to be operated at headend (e.g., MAC Manager and PacketCable Aggregator).
- *AWS Virtual Private Cloud (VPC)* – closely resembles a traditional network in a data center, with the benefits of using the scalable cloud infrastructure. For the FMA cloudification, a VPC is extended from a region to one or more edge locations where headend resides. Interconnecting the region and the edge is an AWS Service Link over the Amazon backbone network.
- *AWS Cloud Edge Network* – provides L2 and L3 connectivity with the headend network. It serves FMA traffic between AWS Outposts and three places: CIN for RMD, Internet for data, voice, and media content, and local equipment for linear video.
- *AWS Cloud Far Edge* – provides AWS IoT services on RMD and CPE with device connectivity over CIN to the cloud edge and cloud region. Since the subject of applications such as IoT is out of the FMA scope, the cloud far edge is not discussed below.

This FMA cloud architecture serves the traditional MSO Backoffice, headend, and their Packet Switch Network Core as a whole and leaves to MSO operators the flexibility and choice for CIN connectivity and the last-mile access network technologies (e.g., DOCSIS Remote PHY, DOCSIS Remote-MACPHY, EPON, GPON). The managed AWS infrastructure services (e.g., VPC, Outposts, EC2, and EKS) enable MSO operators to focus on the FMA functionality for technology innovation and business growth.

4.3. High Availability

High Availability (HA) is essential in cable services. Along the cloud continuum, HA can be achieved for FMA systems and functions using cloud infrastructure services. Figure 8 illustrates multiple HA solution patterns in the cloud region and at the edge with AWS as an example.

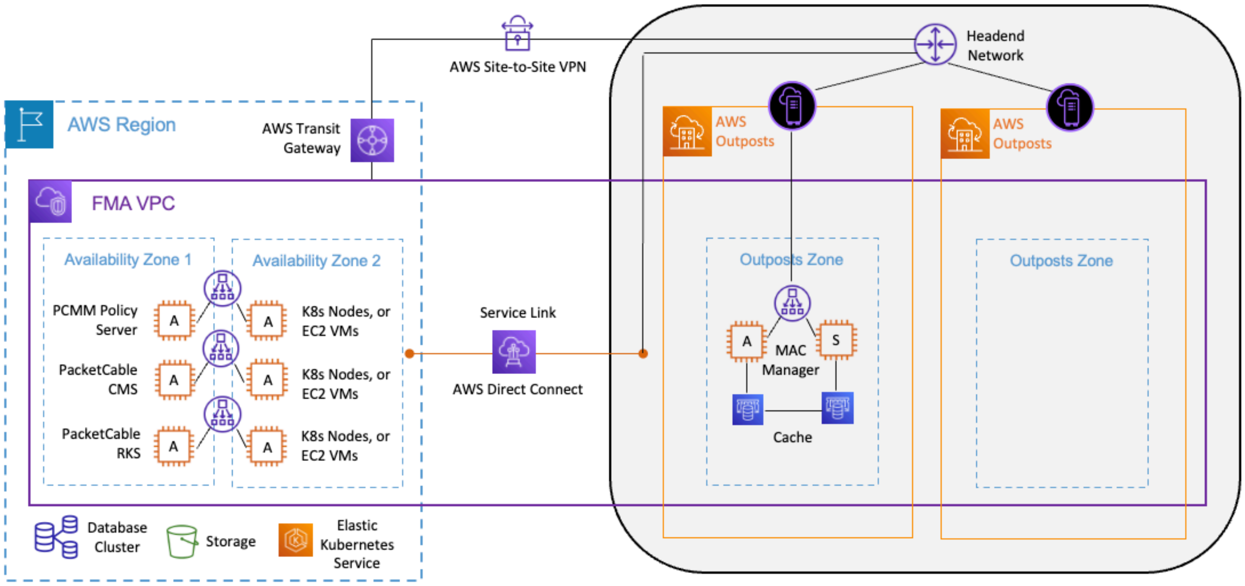

Figure 8 – HA Solution Patterns

Active-Active server HA solution pattern is based on a cloud service mechanism which runs two or more Kubernetes (K8s) nodes with containers in a K8s cluster, or two or more Virtual Machines (VM) instances, across two or more Availability Zones (AZ). The containers or VM instances, are load-balanced through Load Balancers. It is illustrated with FMA PacketCable management functions running in the Active-Active mode in the AWS Region in Figure 5. Failure of an entire Availability Zone, a VM instance, a K8s node, or physical server hosting the VM instance or K8s node will lead to re-routing the traffic to the active resources. Combined with the AWS Auto Scaling service, the Active-Active HA solution delivers the required resource capacity by automatically replenishing lost resources (VM instances or K8s nodes) after the failover.

Active-Standby server HA solution pattern uses two sets of resources (VM instances or K8s nodes) in two AZs in an AWS region or on an Outposts platform, with one set in the Active mode and the other Standby. When the Active fails, the Standby takes over. To simplify the Active-Standby HA solution, the application running on the VM instances or K8s nodes should be stateless with its state data stored in a cache or storage. Figure 5 illustrates the Active-Standby HA model in an AWS Outposts at the edge hosting MAC Manager. For higher availability, two Outposts platforms can be used and the applications are configured for failover over local network paths.

Network HA is supported for the Service Link between the cloud region and the edge with an encrypted set of VPN connections through AWS Direct Connect-based private or/and public connectivity or/and Internet-based public connectivity. In addition, the FMA may add a redundant out-of-band AWS Site-to-Site VPN link over Internet or AWS Direct Connect in-between the cloud region and the edge.

Database, cache, storage, and file system HA are provided as cloud-native services managed by AWS. For example, Amazon Relational Database Service (Amazon RDS) supports High Availability with one standby or two read standbys. AWS ElastiCache provides Redis replication groups within an AZ or across multiple AZs with automatic failover. Amazon Simple Storage Service (S3), an object storage, provides HA and durability through across-AZ replication; S3 can be enabled for cross-region replication as well. Amazon FSx and Amazon EFS file systems can also be deployed across multi-AZs for HA. The FMA cloudification will benefit from these and many other cloud-native database, cache, and storage services.

4.4. Operation automation

The FMA cloudification can utilize cloud services to ease FMA system deployment, provisioning, upgrade, and event monitoring and handling through automation, thus reducing MSO OpEx and in turn TCO. Figure 9 illustrates a procedure for FMA cloud operation automation. It represents an FMA cloud life cycle with AWS services as example.

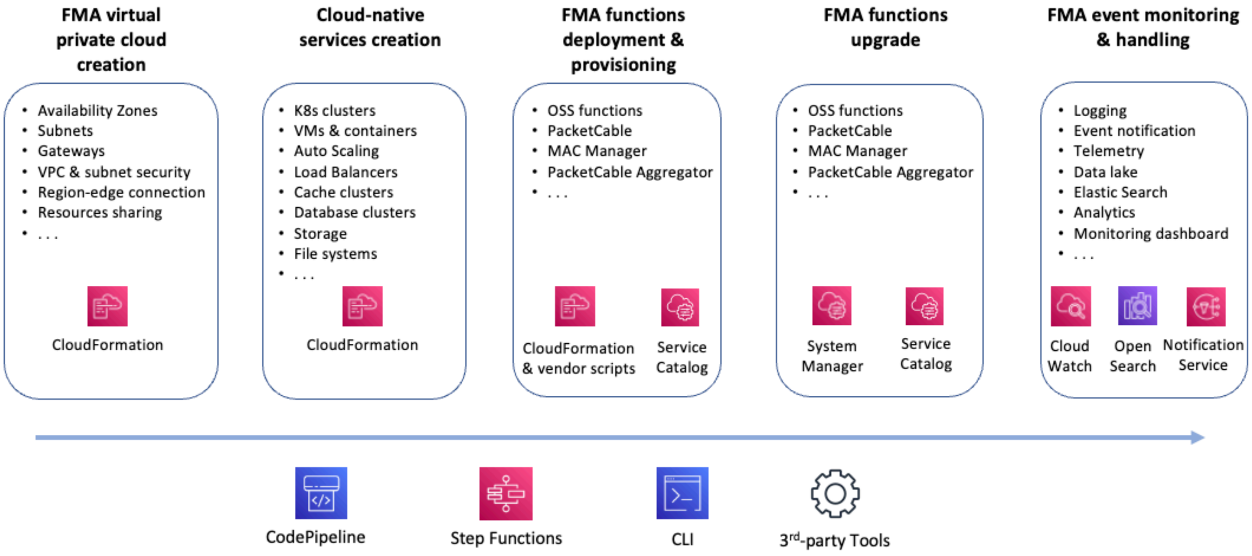

Figure 9 – FMA Cloud Operation Automation

- *FMA AWS VPC creation* – sets the network and security environment for the cloud continuum automatically by AWS CloudFormation templates.
- *AWS cloud-native services creation* – establishes AWS services used by the FMA cloud automatically by AWS CloudFormation templates.
- *FMA functions deployment and provisioning* – are performed in combination of AWS CloudFormation, container services, 3rd-party tools, or scripts with vendor container or VM images. AWS Service Catalog is used to manage products deployed in AWS.
- *FMA functions upgrade* – is part of product life cycle in the FMA cloud. Among different tools, AWS System Manager is a cloud-native service for product update. AWS Service Catalog is used to track the product upgrade.
- *FMA event monitoring & handling* – can be supported by AWS cloud-native services such as AWS CloudWatch, OpenSearch, Analytics, and Simple Notification Service (SNS). These services can implement or integrate some of FMA OSS functions.

The FMA cloud deployment, provision, and upgrade procedures can be driven by AWS CodePipeline and Step Functions with other 3rd-party operation automation tools.

The above content uses AWS cloud infrastructure and services as an FMA cloudification example. Similar solutions can be built with other cloud providers.

5. Cloud FMA use case exercise

The previous sections establish a cloud FMA reference architecture with solution building blocks and provide methods for FMA high availability and deployment automation in the cloud. This section demonstrates how an FMA function – Streaming Telemetry can be implemented across MAC-NE and CIN on premises and MAC Manager and Network Support Services (NSS). It uses AWS cloud infrastructure and services as an implementation example. Similar implementation can be built with other cloud providers.

The streaming telemetry is a PUSH-based mechanism to transport monitored network status data from network elements to external data collectors. Instead of the traditional data PULL model, FMA as well as CCAP utilize the PUSH model to stream data to backoffice applications in near real time[4][5]. Figure 10 shows the MAC-NE dial-out streaming telemetry sequence diagram in the FMA OSS Interface Specification[4].

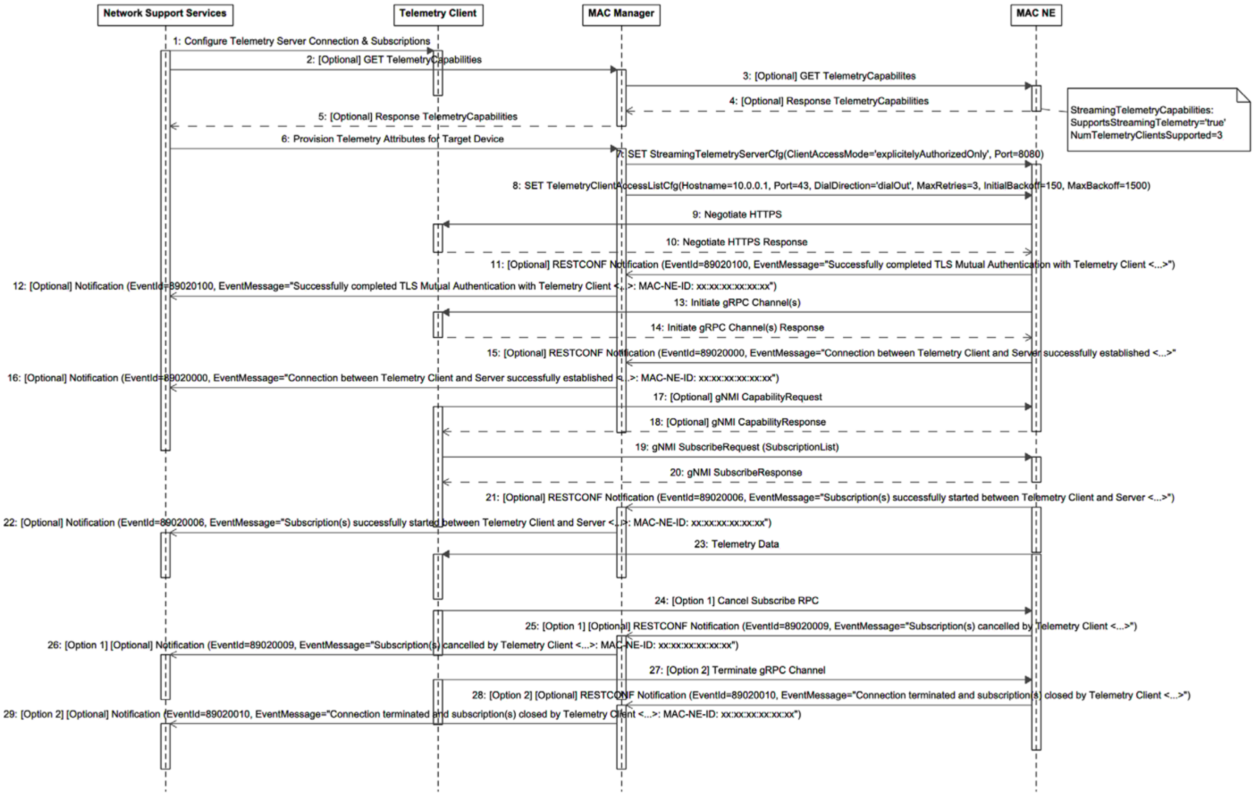
Network Support Services
Telemetry Client
MAC Manager
MAC NE
1: Configure Telemetry Server Connection & Subscriptions
2: [Optional] GET TelemetryCapabilities
3: [Optional] GET TelemetryCapabilites
4: [Optional] Response TelemetryCapabilities
StreamingTelemetryCapabilities:
SupportsStreamingTelemetry='true'
NumTelemetryClientsSupported=3
5: [Optional] Response TelemetryCapabilities
6: Provision Telemetry Attributes for Target Device
7: SET StreamingTelemetryServerCfg(ClientAccessMode='explicitelyAuthorizedOnly', Port=8080)
8: SET TelemetryClientAccessListCfg(Hostname=10.0.0.1, Port=43, DialDirection='dialOut', MaxRetries=3, InitialBackoff=150, MaxBackoff=1500)
9: Negotiate HTTPS
10: Negotiate HTTPS Response
11: [Optional] RESTCONF Notification (EventId=89020100, EventMessage="Successfully completed TLS Mutual Authentication with Telemetry Client <...>")
12: [Optional] Notification (EventId=89020100, EventMessage="Successfully completed TLS Mutual Authentication with Telemetry Client <...>: MAC-NE-ID: xx:xx:xx:xx:xx:xx")
13: Initiate gRPC Channel(s)
14: Initiate gRPC Channel(s) Response
15: [Optional] RESTCONF Notification (EventId=89020000, EventMessage="Connection between Telemetry Client and Server successfully established <...>"
16: [Optional] Notification (EventId=89020000, EventMessage="Connection between Telemetry Client and Server successfully established <...>: MAC-NE-ID: xx:xx:xx:xx:xx:xx")
17: [Optional] gNMI CapabilityRequest
18: [Optional] gNMI CapabilityResponse
19: gNMI SubscribeRequest (SubscriptionList)
20: gNMI SubscribeResponse
21: [Optional] RESTCONF Notification (EventId=89020006, EventMessage="Subscription(s) successfully started between Telemetry Client and Server <...>")
22: [Optional] Notification (EventId=89020006, EventMessage="Subscription(s) successfully started between Telemetry Client and Server <...>: MAC-NE-ID: xx:xx:xx:xx:xx:xx")
23: Telemetry Data
24: [Option 1] Cancel Subscribe RPC
25: [Option 1] [Optional] RESTCONF Notification (EventId=89020009, EventMessage="Subscription(s) cancelled by Telemetry Client <...>")
26: [Option 1] [Optional] Notification (EventId=89020009, EventMessage="Subscription(s) cancelled by Telemetry Client <...>: MAC-NE-ID: xx:xx:xx:xx:xx:xx")
27: [Option 2] Terminate gRPC Channel
28: [Option 2] [Optional] RESTCONF Notification (EventId=89020010, EventMessage="Connection terminated and subscription(s) closed by Telemetry Client <...>")
29: [Option 2] [Optional] Notification (EventId=89020010, EventMessage="Connection terminated and subscription(s) closed by Telemetry Client <...>: MAC-NE-ID: xx:xx:xx:xx:xx:xx")

Figure 10 – MAC-NE Dial-out Streaming Telemetry Sequence Diagram[4]

Figure 11 depicts the streaming telemetry cloudification with cloud-native services on AWS as an example. An AWS Region hosts a Cable operator's backoffice, where the NSS, MAC Manager, and Telemetry Clients are deployed in the FMA VPC subnets. These FMA components can run as VMs on Amazon Elastic Compute Cloud (EC2) or as containers on EC2 orchestrated by either Amazon Elastic Container Service (ECS) or Amazon Elastic Kubernetes Service (EKS). According to the dial-out streaming telemetry protocol, the components communicate with each other through the subnet traffic routing within the VPC. The telemetry servers in MAC-NEs are connected to the telemetry clients and NSS in the backoffice in the cloud region via Amazon Direct Connect over AWS network backbone. AWS Site-to-Site VPN over Internet is a backup link for connection high availability. Note that the telemetry clients can be implemented at the cloud edge as well if preprocessing, filtering, or analysis of telemetry data should take place locally first.

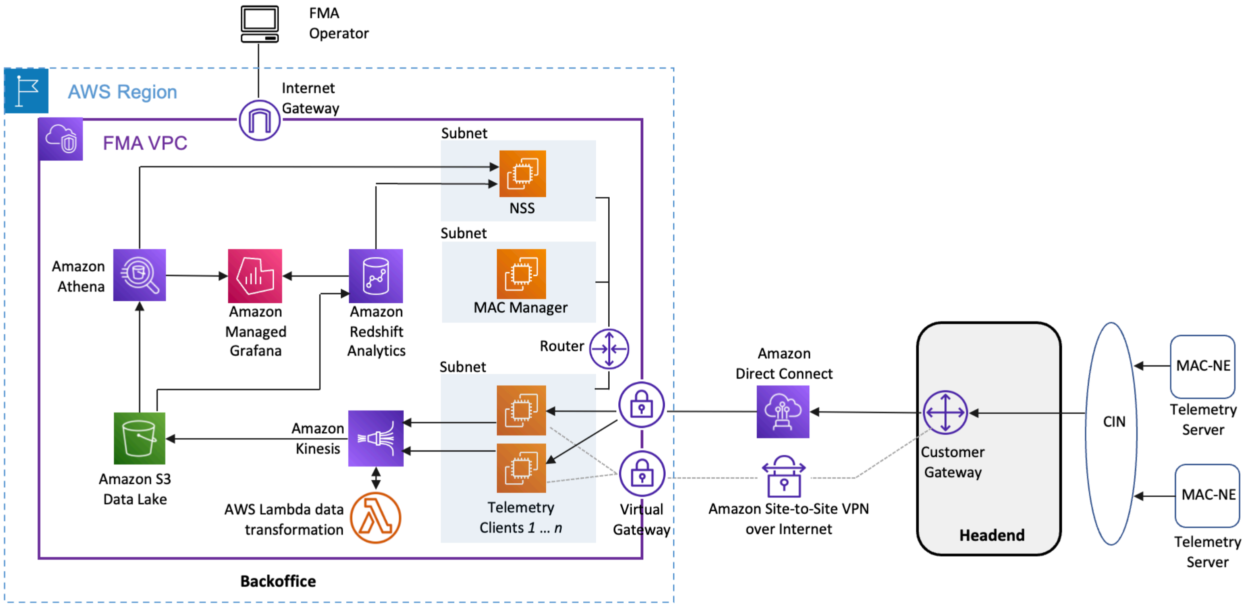

Figure 11 – Streaming Telemetry in Cloud FMA

In addition to the streaming telemetry implementation, the cloud provides several salient services for advanced OSS operations as illustrated in Figure 11.

- *Data lake* – is nowadays a common service provided by cloud service providers to store, process, and secure large amounts of structured and unstructured data. For example, AWS Simple Cloud Storage (S3) is a durable and scalable data storage to host a variety of OSS data for recording, troubleshooting, and predictive equipment maintenance. The telemetry clients further stream the received data to S3 using Amazon Kinesis Data Firehose service.
- *Analytics* – is used by advanced OSS operations for network event correlation, root cause analysis, and outage prevention. Many types of analytics services are provided by cloud service providers. As an example, Amazon Redshift is a fully-managed data warehouse service that can analyze petabytes of telemetry and other OSS data efficiently. Coupled with Amazon Machine Learning service, Amazon Redshift ML makes it easy for data analysts to create, train, and apply machine learning models.
- *Dashboard* – is an essential function of network observability. Grafana is a widely used open-source observability tool. For instance, with Amazon Managed Grafana in the cloud, cable operators can analyze and visualize telemetry and OSS metrics, logs, and traces and configure alerts for OSS event notifications. The Grafana service uses Amazon Athena as data source to access the telemetry and other OSS data stored in the data lake S3.

Note that there are other methods or cloud services for implementing the Streaming Telemetry function for cable operator’s data networks, e.g., [6], which are not out of the scope of this paper.

6. Conclusions

The telecom industry has started its journey from network function virtualization to cloudification in order to lower the total cost of ownership, increase operation efficiency, leverage modern application services, and achieve business agility. FMA cloudification is the next undertaking by communication service providers and multi-access operators to further modernize the Distributed Access Architecture.

This paper shows that the FMA can be cloudified along the cloud continuum across the cloud region, edge, and far edge with FMA systems and functions allocated based on their timing characteristics and operation locality. At the same time, the FMA cloudification enables the operators to retain the flexibility of access network technology choices and operations.

It is shown through the cloud infrastructure and services that the private edge cloud is more suitable for cable operator's latency-sensitive FMA functions than other types of cloud edge technology. The virtual private cloud extends from a cloud region ("FMA regional datacenter") to a private cloud edge ("FMA headend"), simplifying FMA OSS operations. Traditional cable HA capabilities can be achieved by the cloud infrastructure and services, including network redundancy, multiple availability zones, load balancing across containers and virtual machines, and database, cache, storage, and file system redundancies. Cloud-native services can not only implement FMA functions but also bring them to modern OSS operations as illustrated through the streaming telemetry use case.

It is time to implement FMA in the cloud by applying the FMA cloudification approach and leveraging the cloud infrastructure and services.

Acknowledgements

The authors would like to thank Dr. Jennifer Andreoli-Fang for her feedback on the writing style for SCTE Cable-Tec Expo and her valuable comments on the paper content.

Abbreviations

CCAP	Converged Cable Access Platform
CIN	Converged Interconnect Network
EMS	Element Management System
FMA	Flexible MAC Architecture
gNMI	gRPC Network Management Interface
MAC-NE	MAC Network Element
MSO	Multi Service Operator
NSS	Network Support Services
RMD	Remote MACPHY Device
VPC	Virtual Private Cloud

Bibliography & References

[1] Web publication, *Distributed Access Architecture*, CableLabs, https://www.cablelabs.com/technologies/distributed-access-architecture.

[2] CM-SP-FMA-SYS, *Flexible MAC Architecture System Specification*, CableLabs, January, 2022, https://www.cablelabs.com/specifications/CM-SP-FMA-SYS.

[3] John Chapman and Tong Liu, *Unleash the Power of Cloud Computing for CMTS*, SCTE 2021, Atlanta, GA, October, 2021. https://www.nctatechnicalpapers.com/Paper/2021/2021-unleash-the-power-of-cloud-computing-for-cmts.

[4] CM-SP-FMA-OSSI-I02-220602, *FMA OSS Interface Specification*, CableLabs, June, 2022, https://www.cablelabs.com/specifications/CM-SP-FMA-OSSI.

[5] CM-SP-CCAP-OSSI, DOCSIS 4.0, *CCAP Operations Support System Interface Specification*, CableLabs, June, 2022, https://www.cablelabs.com/specifications/CM-SP-CCAP-OSSIv4.0.

[6] R. Harlin, A. Moustafa, and A. Kalawat, *How Comcast uses AWS to rapidly store and analyze large-scale telemetry data*, August, 2021, https://aws.amazon.com/blogs/big-data/how-comcast-uses-aws-to-rapidly-store-and-analyze-large-scale-telemetry-data/.