Hewlett Packard Enterprise

Brochure

From Subscriber Data Management to shared data environment

SDM in 5G architecture

Introduction

The telecommunications industry is evolving at an unprecedented pace. Communications service providers (CSPs) are facing several major disruptions. Network function virtualization is rapidly taking place. Heterogeneous networking (HetNet) is becoming the norm. Wi-Fi is blurring differences between fixed and mobile networks. Over-the-top (OTT) providers and the Internet of Things (IoT) are stressing traditional business models and pushing CSPs to transform into Digital Service Providers (DSPs). 5G is emerging with a set of requirements and applications that will change current network architectures.

While these changeovers occur, end users will continue to expect a ubiquitous, seamless access and a lifestyle experience. In other words, DSPs will need to be subscriber centric from the start, for example, build the required knowledge to maximize the customer lifetime value and remain successful in business.

In order to support those transformations, DSPs will expect their Subscriber Data Management (SDM) ecosystem to support them in building a consistent, in context, view of users' behavior, experience, and expectations. The SDM will need to go beyond traditional subscriber network data. It will have to hold new information such as quality of experience (QoE) KPIs, service usage, and derive real-time insights relevant to the entire organization including partners.

This paper focuses on how operator's SDM capabilities need to evolve into shared data architecture available to the entire organization.

From CSP to DSP

CSPs broaden their business orientation toward being DSPs. At the same time, providers face stronger challenges of handling larger than ever volumes of data while revenue and margins decrease. Pressure to reduce OPEX and improve scaling has never been more acute.

DSPs will partner and deliver many third-party applications along with their own services. While keeping the end customer relationship, they will need the ability to quickly integrate and bind users to those new applications. Also, as customer-centric organizations, they will need to settle and entitle, in priority, the most popular and growing applications as they emerge.

Naturally, DSPs will face competition coming from OTT providers and will need to excel by delivering media-intensive and context-aware services with the best possible experience. A DSP will constantly need to understand users' experience, behavior, and expectations, and reshape its infrastructure accordingly.

These underlying permanent business changes underway prompt CSPs to shape their systems with 5G acting as a tailwind.

5G impact

Indeed, 5G by many aspects implicitly mandates adoption of Telco Clouds: a distributed, automated, virtual, and composable environment.

Ultra-low latency: To enable new networked applications such as augmented reality or even remote surgery, 5G expects latency under 5 ms from end to end. This implies to revisit the network architecture especially in terms of where and how elements are deployed. A drastic reduction of the physical distance separating users or user equipment (UE) and core network resources becomes mandatory. The ability to deploy both the core functions and applications' business logic and their associated data as close as possible to end users becomes paramount.

Service deployment time: As part of DSP's transformation, time to market remains not only a competitive advantage but also a survival factor. DSPs need to deploy new applications in a few hours and bind users to them near instantaneously. As over the internet today, popular applications are the ones driving business outcomes and an effective management of long tail of applications combined with a truly user-centric QoE are winning factors. From a user data management perspective, the new SDM environment needs to ease activation, provide seamless access, and make customer knowledge immediately available.

Programmable infrastructure: DSPs need to offer more than a network. They need to provide a composable infrastructure as a set of exposed capabilities that application developers can combine to rapidly build new applications and scale them as their success comes. A new application may require an authentication service with no location and get access to specific customer data. A different application may leverage location information but do not need subscriber's charging information. Customer data itself should become a service accessible through APIs where only authorized applications could get relevant data. Core network functions themselves will need to become elastic, modular, and exposed.

Cloud nativeness

The effectivity of a cloud approach, with functions that are virtual and decomposed, implies a stronger separation of data storage versus processing functions for supporting:

- Fast scaling of processing functions without a need to migrate, move, or save data intermediately
- User profiles and context resiliency during normal lifecycle operations (that is, start, stop, scale, terminate) and also failures (and healing) of virtual network functions should they be at software or infrastructure level. Any failure shall remain very transparent to the end user preserving an unmatched QoE

A clear separation of processing from data storage lightens network functions and leads to multiple benefits:

- **Design of stateless and dataless VNFs** implements the required business logic while leaving the state retained by a data storage capability. Within a stateless virtual network function (VNF), VNF components (VNFCs) have smaller footprints. They can scale out and be placed easily, which leads to more effective cloud resource usage scaled-in, as less business logic instances are needed. Thus, the Telco Clouds deploying those cloud-native VNFs become truly elastic resources, as opposed to a **one-size-fits-all** design paradigm. This modularity also helps to achieve better service availability at less cost through N+1 type deployments. Indeed, the traditional fully redundant architectural choices and the corresponding data replication inside the VNFs are not needed anymore. VNFS can also be easily deployed placed either centralized or at the edge.

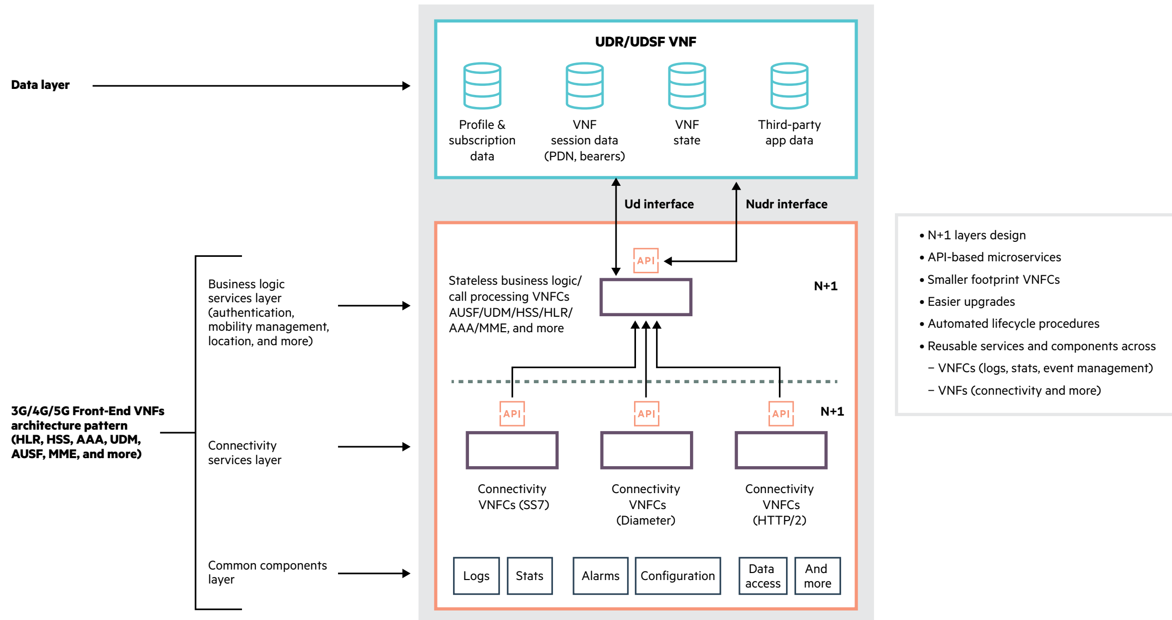

Figure 1. Telco Cloud data layer conceptual view

- **Layered VNF deployment approach** with common core services, such as logs, statics, or events, across VNFCs and common connectivity services (or protocols) used by various business logic types makes VNF more composable and future-proof.
- **Functional upgrades and deployments** of VNF instances are easily completed with just the core logic involved compared to monolith functions with data included. This agility is critical for DSP's ability to introduce services in minutes with no service interruption. Operators could combine existing and new VNFCs to form 5G functions enabling smooth transitions from the current network technology. For instance, Home Subscriber Server (HSS), Authentication Authorization, and Accounting (AAA), and Authentication Server Functions (AUSF) functions may leverage common capabilities such as authentication algorithms. An HSS and User Data Management (UDM) may reuse the same diameter connectivity stack at first step before evolving to the service-based architecture.
- **Independent scaling** with traffic volume enables better network agility and balances load.

VNF data sharing facilitates a cross function customer-centric approach. As VNFs externalize their data and states into a separated storage, other functions can access this data when appropriate. Even contextual information, such as session data, that is traditionally spread across functions can now be shared and accessed. Analytics can be performed on context, in real time, as opposed to remain after the fact as with usual Extract, Transform, and Load method (ETL) and data warehouse approaches. Data may be shared between VNFs, for example, HSS and UDM may share subscription data to enable seamless 4G/5G handovers. Moreover, this VNF data sharing practice helps in managing transition from silos Wi-Fi, to 4G, to broadband networks, to 5G. For instance, starting to consolidate some network state and user context data helps co-existence of future migration to 5G. An operator may choose to first duplicate AAA data in the data environment, transform and simplify the AAA business logic, and finally turn down the legacy AAA.

- **Core network simplification:** Sharing data across functions can decrease signaling traffic between functions and reduce corresponding network sizing.
  - 3G and 4G core networks signaling traffic includes network elements' exchange of data that is anyway present in a databases (UDR, SPR, and more). Those exchanges could be reduced if not eliminated.

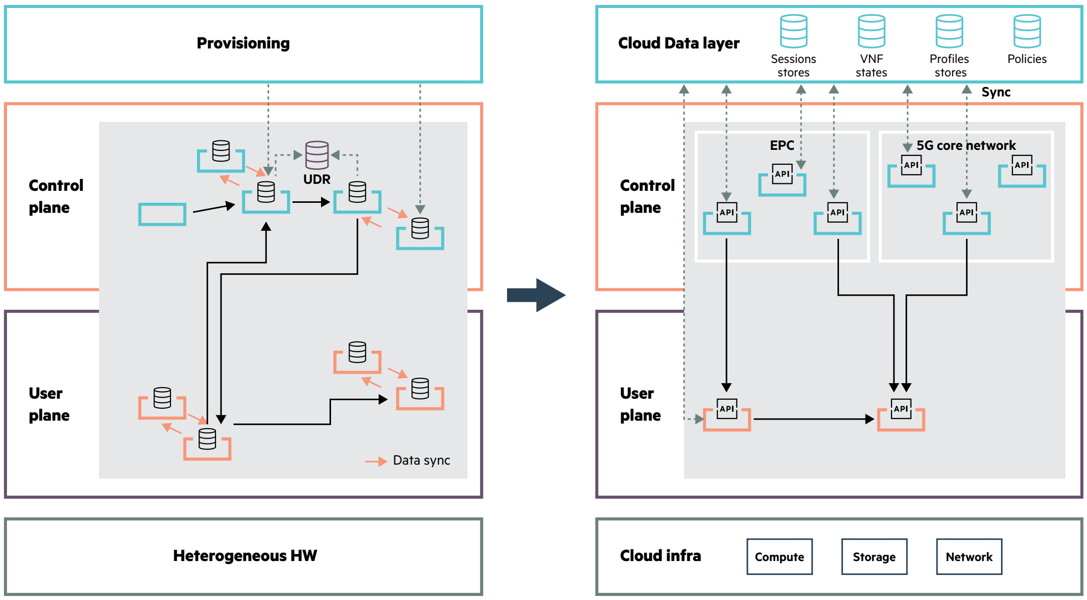

Figure 2. From data duplication to data cloud architecture

– Preservation of users' context into a separated data layer helps VNF recover from failure and avoid signaling storms that can slow down the network and affect the user experience. Today, when a MME handling millions of users' contexts fail, UE need to re-attach to a new available MME generating a storm. With a proper MME VNF state preservation, MME could failover transparently to the UE avoiding the traffic storm and the need to size the network as a function of this peak.

– Easier interoperability between existing legacy, 4G and new 5G functions and better consistency across access networks—For instance, a user QoS policy need to remain consistent for handovers from 4G to 5G.

This architecture paradigm change directly affects the subscriber data environment. Data can be available and accessed by many functions that require throughput performance of millions of TPS and response times within the milliseconds latency range of 5G.

Sharing of user data

Beyond the benefits of shared data for the cloudification of network functions, user data has its own unique characteristics and special challenges:

- **Building a 360-degree view with data residing either over or under the top, in central premises or at the edge for efficiency reasons:** It should include analytical insights that are often based upon statistical techniques that need a large data corpus going beyond the small local clouds. Large-scale data consolidation, federation, and synchronization will be needed to ensure the end-to-end customer view.
- **Exposing to ecosystem:** Application partners should be able to access and query for real-time contextual insights. In some way, a CSP needs a secure, carrier-grade, real-time operational datastore that is able to handle massive profile, usage, real-time context information such as session data.
- **Keeping context:** Handling massive amounts of volatile data while performing analytics and exposing needed user **on-context** insights remains a challenge with current database technologies. Current systems based on traditional **after-the-fact** analytics can't satisfy ultra-reactive 5G applications, such as augmented reality, where analytics must occur timely as data comes.

Naturally, implementation of this data sphere requires significant element of integration with surrounding systems, and is dependent on regular customization and agile development. Key capabilities for that are

- **Open data model:** Flexible and extensible to support a wide variety of data types, structured and unstructured such as binary data.
- **Open interfaces:** To facilitate and foster innovative application developments. Data shall be accessed as a **service** through APIs where developers are independent of underlying database administration complexities.

As VNFs turn stateless and relay on external repositories to hold their state, new kinds of performance challenges are introduced by the unprecedented levels of simultaneous read and write capacities. Efficiencies need to be kept not only with growing data volumes and throughputs but also with new governance complexities and with always increasing actors who create, share, and update user data. The new data environment shall be a new layer in the end-to-end core architecture layer.

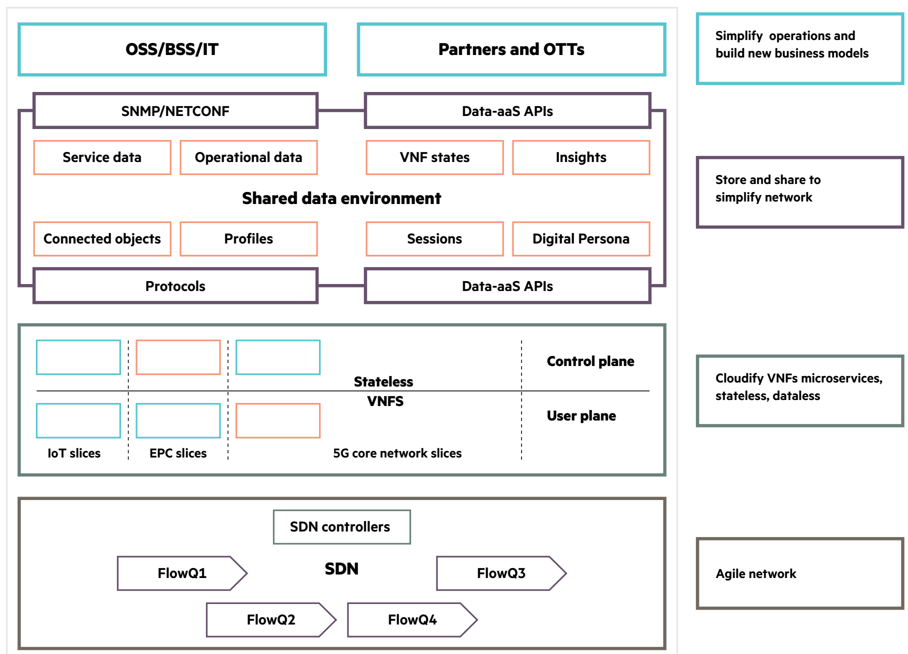

Figure 3. Shared data environment conceptual view

Shared data environment

This shared layer requires combining unique and often contradictory capabilities: openness and security, carrier-grade performances and extreme agility, and low-cost and extreme volumes. For user data or SDM, this is adhering to key architectural principles:

- Adaptation to truly virtual and distributed environments
- Ability to cope with a large variety of workloads and massive performances
- Openness to support current and future function data

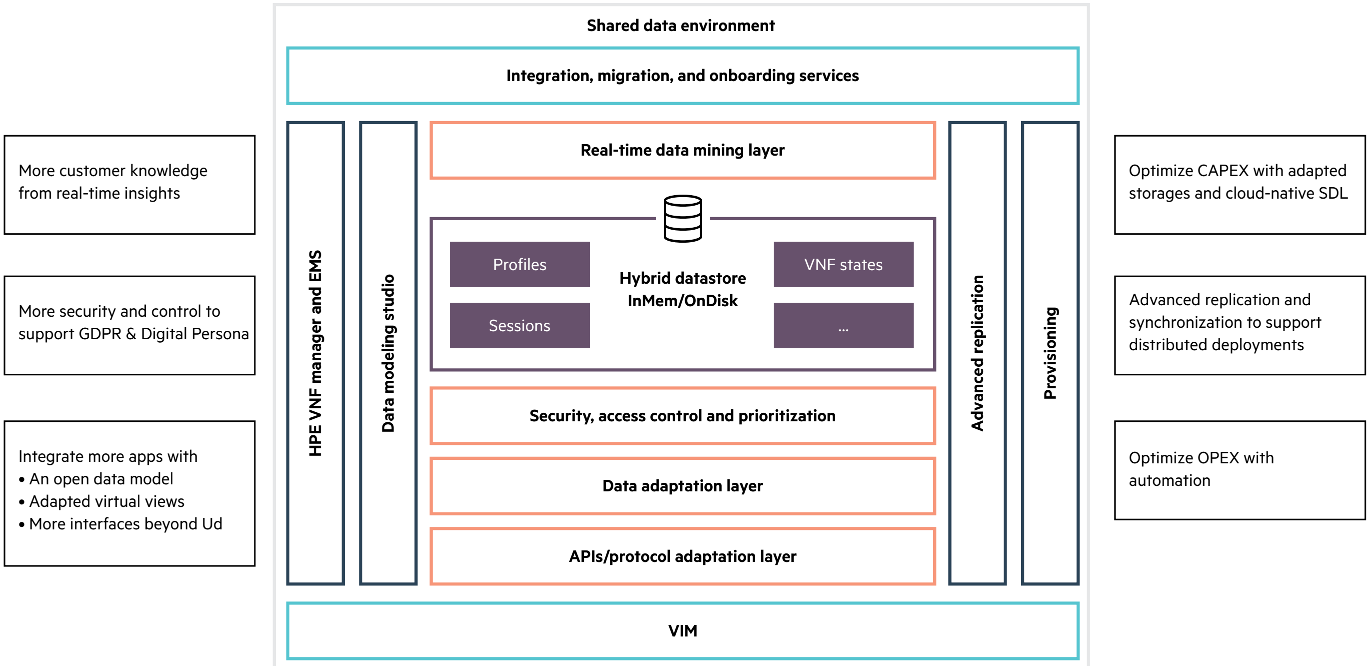

Figure 4. Shared data environment conceptual overview

Several capabilities are key for coping with these requirements:

- **Virtualization:** The functions of SDM itself shall be virtualized and elastic to ensure the best data availability for the evolving infrastructure. Should it be deployed centrally or at the edge, in separate data centers or collocated with Front-End application clouds, it reduces latency and offers required availability options.
- **Hybrid storage:** Taking advantage of various memory and disk technologies to avoid traditional limitations of those two core-computing components. With right combination, highly volatile session states and persistent data can be stored in a single environment, for improved efficiency with reduced CAPEX and OPEX. This architecture also offers readiness to upcoming persistent memory with unprecedented levels of performance and density. Data layer is designed to cope with the millions of TPS needed at the millisecond response time.
- **Data model openness:** As operators transition into DSPs, new applications need to be integrated continuously. A protocol and a data adaptation layers relying on a truly open data model offer unique integration capabilities for current and future functions and is ready for HTTP/REST service-based interfaces. This openness allows all networks (3G, 4G, 5G, broadband, Wi-Fi, and beyond) to co-exist over the same data environment and access data as a unified service, ready to support upcoming 5G functions such as UDR and unstructured UDSF functions as defined in 3GPP.
- **Security and privacy:** As CSP move toward a customer-centric DSP model, customer knowledge will become the new currency while security and privacy will have to be preserved. When compared to having to deal with multiple datastore technologies with different access rights policies, interfaces, and protocols, sharing of data simplifies and makes security management more robust.

Evolving with HPE SDM

Whether you have a siloed environment or have started to deploy a Back-End to Front-End (BE-FE) separation, HPE SDM can help in upcoming transformations. With service layer transformation and digitalization of core services, Hewlett Packard Enterprise is helping CSP to better leverage IMS and LTE and integrate OTT applications to provide immersed end user experiences by pervasively sharing of user profiles, devices, connected objects data, and context.

As an NFV leader, Hewlett Packard Enterprise has transformed its portfolio and is driving multiple CSP migration to Telco Cloud and getting ready for 5G as a catalyst toward a true open DSP environment.

HPE is a proven SDM partner and is well advanced in the implementation of its vision. Our SDM provides some of the key features, and HPE is strongly committed to a long-term, adaptable, and composable data management environment supporting current NFV migrations and future DSP transformations.

About the author

Alain Guigui is Chief Architect of the Actionable Customer Intelligence domain part of the Communication and Media Solutions business at Hewlett Packard Enterprise. He is an expert of telecommunication network and has created the HPE Telecom Analytics business line. He has worked on 4G/LTE as a solution architect and now works on 5G evolution at HPE. He works and acts as a consultant for several major Communication Services Providers around the globe and has participated to several ETSI standardization Committee for several years.

Learn more at
hpe.com/dsp/personalize

a00035556ENW, November 2017