

Why Problems With Service Chaining Are Stalling NFV





August 8, 2018 Fergus Wills

Has <u>NFV</u> promised too much and delivered too little? It has been over <u>six years since</u> <u>ETSI</u> set bold standards for the technology. It was meant to usher in an era that would replace physical and software-based appliances with virtual functions, allowing <u>services</u> to be restructured and redesigned around the network and the subscriber. Solutions would be interoperable, working seamlessly across equipment from multiple vendors in a cooperative environment. The reality has been rather different, and the NFV journey has been slower and more problematic than expected.

A number of mobile operators have found service function chaining (SFC) particularly bumpy. Not so much in the case of virtual evolved packet core (vEPC) or virtualized infrastructure management (vIM), where NFV can be deployed without SFC, but more in the case of GiLAN. (GiLAN refers to the domain between the GGSN/PGW in the operator network and the internet.)

An example of this is mobile video traffic management — a key function for managing high-demand, rapidly increasing traffic for mobile. In this case the service chain would involve transferring end-user meta data on the control plane and dynamically applying this to the user plane functions — all within tight time constraints — to both add value and add a network management point capable of handling application traffic dynamically. The reality of the inherent problems of doing this in real-time has led to inertia, slowing the evolution of a solution that meets the demand of today's traffic, both in diversity and capacity.

A Virtualized Headache

Problems within service chains have come to epitomize the problems with NFV. When it comes to deployments, there are significant restrictions on the number and variety of functions in a service chain. This leads to either remaining with legacy, physical network

functions vendors or increasing the number of silos, which is a shame as the NFV vision was meant to break down these two barriers. Frustratingly, this can lead to increased costs as the operator transforms fixed physical infrastructure into a software-based, dynamically switched model. It turns out this is easier said than done.

So, why are these restrictions manifesting today? At a macro level it could be attributed to the maturity of both the infrastructure and the functions at the same time.

Here are three physical limitations that have become apparent for a number of mobile operators:

- Scaling in a heterogeneous landscape: When scaling functions, the reality of the physical infrastructure has to be considered from a deployment perspective. Unfortunately, in a heterogeneous solution environment, operators can have three or four vendors offering different services. In an all-IP traffic management example, one vendor would offer parental control, another antivirus, another optimization, etc. They are not likely to work seamlessly on the front end and to run smoothly, invariably requiring multiple physical components. Each box adds extra latency to the overall traffic management, which will result in poor Quality of Experience (QoE). Currently, the time spent in the mobility path is 5 to 10 milliseconds for GiLAN services. This means that when deploying multiple services, if the user plane traffic has to transition between various physical commercial off-the-shelf (COTS) hardware, then this delay automatically increases and can result in poor QoE. Poor QoE not only leads to poor scores in network speed tests, but ultimately contributes to churn. In 5G that time could drop to 1 millisecond.
- Control plane and signaling metadata: To execute a specific function on the service chain, the transfer of metadata to the lightweight directory access protocol (LDAP) store from the classifier and head-end is often a basic requirement. This is usually the case when the function is policy based for the signaling store to deliver a subscriberspecific service. This starts with policy reference and network identity but it can rapidly expand.
 - There are a variety of techniques for this, starting with network service headers (NSH). Owed to the plethora of equipment and protocols from multiple vendors, there is incompatibility between functions and the volume of metadata that needs to be managed. All of this can lead to significant inefficiencies, especially when the network needs to transmit every packet. As such, the basic rules of the service chain must be agreed on between all vendors; and metadata should be cached so that it can be communicated when changes occur. This is a critical design change that vendors must implement.
- Switching rules and multi-tenancy: Operators have made significant efforts to define chaining rules and their scalability in projects like Open vSwitch (OVS). With open source, there is room for improvement to handle both the volume of rules as well as the changes needed to make to those rules. This can result in either a simplified switching framework, a reduced number of functions in the chain, or siloed service chains with no multi-tenancy.

As NFV deployments continue to evolve, what three strategies can operators adopt to mitigate these challenges? First and foremost, the industry needs to foster an environment of collaboration between vendors, mobile operators, and working groups to advance virtual GiLAN services.

• Smarter scaling: In the case of virtual GiLAN, increased flexibility in the definition of virtual network function (VNF)-components would enable better mapping to physical hardware — which could allow all components to be on one physical COTS blade — in turn reducing latency for transferring data. For example, in a 40 core blade system, eight may be assigned to hypervisors, and the remaining 32 could be divided by four, eight, or 16 for the core. Using more standard sizing helps operators prepare for any

failovers and for hardware planning. Carriers can also benefit from a wider choice of off-the-shelf components.

- Collaborative control and signaling: This is an area where working groups can play a key role. NSH is today's prevalent common mechanism, but it does not suit all environments and can overload the payload significantly. In two respects, standardization and interoperability can play a role.
 - Firstly, in the case of NSH, operators have experienced vendor lock-in where each domain player and switching provider have their own flavor which has adversely reduced the choices for network providers.
 - The second area, is the need for new thought leadership when the data exchange between functions needs to be more efficient and expandable. The work the industry is doing with vector packet processing (VPP) in this area may be a way forward.
- Switching rules and multi-tenancy: Some operator groups function together and
 utilize internal orchestration within each function. This is counterintuitive. In mobile
 data traffic management, a granular deconstruction of all functions from transport
 optimization and switching to parental controls will examine packets and payloads
 for application-level analysis into separate VNF components. This leads to packet,
 flow, and session data being reconstructed at multiple points in order to enforce the
 policy rules.

At its best this is inefficient, with complex switching and unnecessary hops; and at its worst this increases latency and leads to poor QoE and subscriber churn. The operator should take a view on the composition of services actually required and determine how they could be more tightly integrated.

Is NFV For Real?

NFV was meant to break down silos — instead the limitations of development are perpetuating them. Virtualization was supposed to reduce opex and capex —instead it adversely impacted some operator's bottom lines. No wonder operators want to know whether anyone out there is actually deploying the "real" NFV.

The good news is that the path to "real" NFV is maturing. There are practical difficulties for orchestration, scaling, and latency. This is limiting the efficient rollout of flexible service chains today. Is this something that can be overcome? It can be with the caveat that vendors have to recognize that the promise of NFV implies a deeper degree of flexibility. So, now we have to collaborate with a broader spectrum to discuss the practical realities of NFV today. With 5G on the horizon, there's never been a better time for the industry to change how they do things. The hunt for "real" NFV continues.

CONTRIBUTED ARTICLE DISCLAIMER

Statements and opinions expressed in articles, reviews and other materials herein are those of the authors; not the editors and publishers.

While every care has been taken in the selection of this information and reasonable attempts are made to present up-to-date and accurate information, SDxCentral, LLC cannot guarantee that inaccuracies will not occur. SDxCentral will not be held responsible for any claim, loss, damage or inconvenience caused as a result of any information within this site, or any information accessed through this site.

The content of any third party web site which you link to from the SDxCentral site are entirely out of the control of SDxCentral, and you proceed at your own risk. These links are provided purely for your convenience. They do not imply SDxCentral's endorsement or association. The copyright and any other intellectual property right any third party content belongs to the author and/or other applicable third party.



About Fergus Wills

Fergus Wills is Director of Product Management for Openwave Mobility and has over 16 years of experience in telecoms, primarily on large scale server-side mobile internet data products. Holding a BSc in Information Technology from Queen's University Belfast, he has been involved in several industry bodies on mobile data.