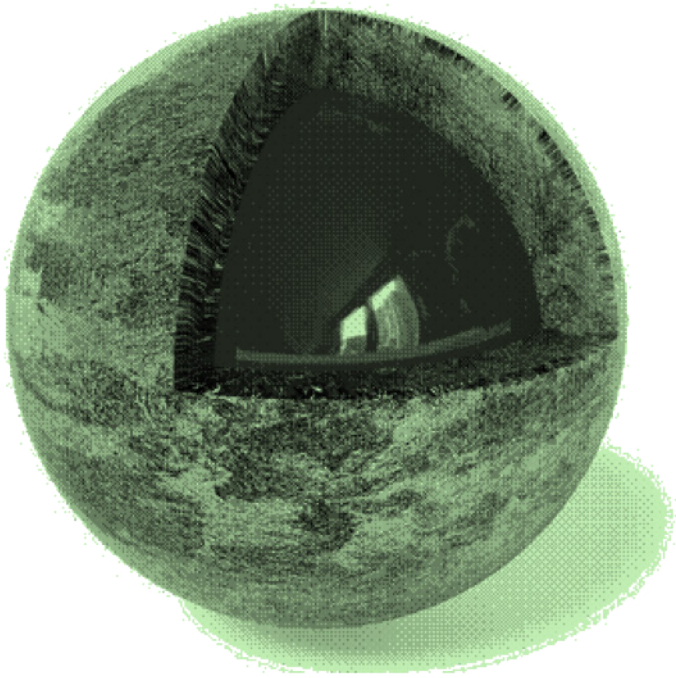
2016

The Business Case for Service Automation in Mobile Networks



Commissioned by:

{Core Analysis}
Patrick Lopez

affirmed®

redhat

Mellanox TECHNOLOGIES

{Core Analysis}

Contents

Figures

Tables

About This Report

Abstract – This white paper, commissioned by Affirmed Networks, Mellanox and Red Hat, examines the business benefits that a virtualized, cloud-ready, automated service provisioning environment can bring to mobile operators.

Accelerate and Reduce Costs of Existing Services

The use cases in this section are composite examples of commercial deployments from individual case studies and their corresponding project plans. The products deployed are a virtual EPC with collapsed Gi LAN functionality and a stateful Automated Configuration Management Platform that records and automates third-party physical and virtual element configuration.

EPC and MVNO

Legacy EPC	EPC Virtualized + Automation
$3.1 million	$0.78 million
19.5 months	9.75 months
Add MVNO	**Add MVNO with Automation**
$48,000	$1,610
2 months	2 days

Sponsored Data

Legacy Zero Rating	Zero Rating with Automation
$9,660	$1,610
12 days	2 days

Parental Control

Legacy	Virtualized + Automation
$0.9 million	$0.38 million
8 months	4 months
Subsequent Service	**Subsequent Service with Automation**
$64,400	$1,610
2 months	2 days

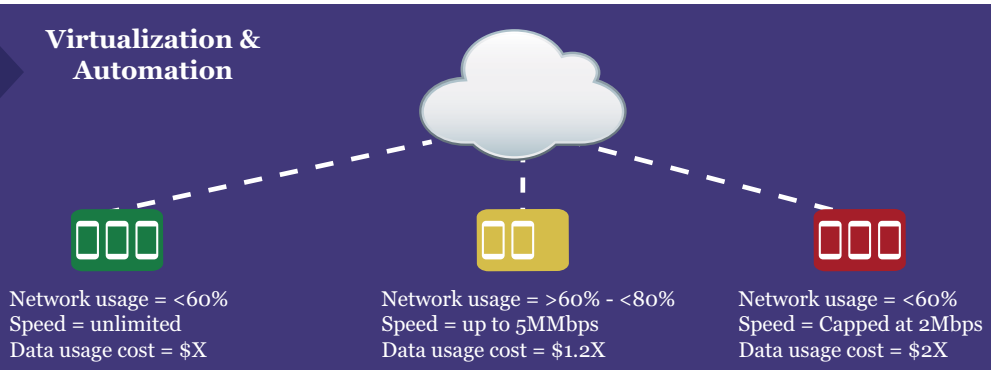
Enterprise VPN as a Service

Legacy VPNaaS Enterprise Onboarding	Virtualized + Automation VPNaaS Enterprise Onboarding
$12,075	$1,610
2 weeks	2 days

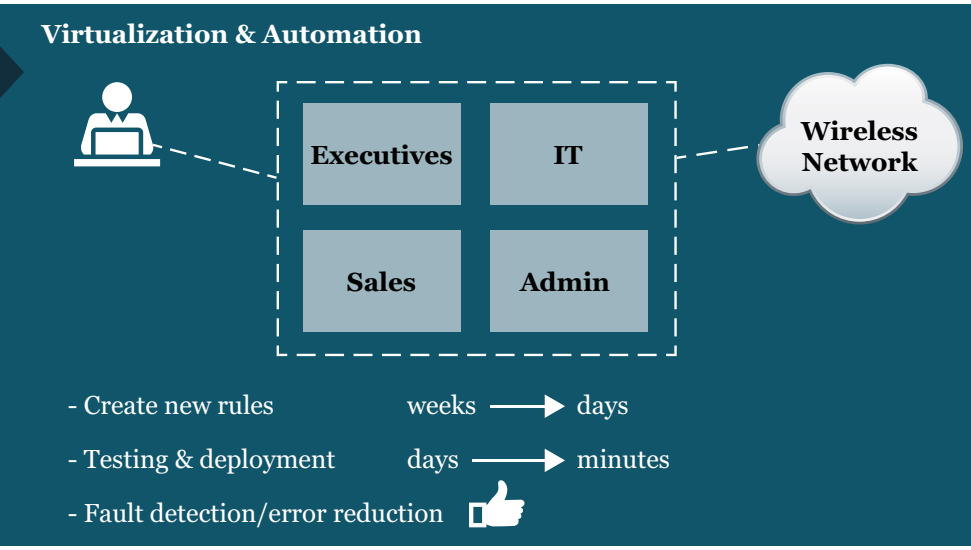
New Services Enabled by Automation

One has to look at the Automated Configuration Management Platform (ACMP) as a best practice that accelerates and multiplies service opportunities. The real benefit of automating service provisioning in virtualized and hybrid environments is the opportunity to launch services that were otherwise too daunting or too expensive to launch.

Demand Congestion or Event-Based Pricing



Enterprise Traffic Management Self-Service



Blockbuster/ Niche Services

EXECUTIVE SUMMARY

Mobile network operators (MNOs) are reducing the cost of delivering services through Network Functions Virtualization and Software-Defined Networking (SDN).

Today, service creation and management rely on many proprietary elements, platforms, interfaces and protocols. Service creation is lengthy (typically 12 to 18 months), error prone (requiring two to three months of planning and testing) and expensive (costing millions of dollars in direct and indirect costs). Most of the delays, costs and errors are attributable to the fact that legacy systems require manual configuration through proprietary, non-interoperable interfaces. To address this challenge, cloud-native vendors are introducing virtualized network functions (VNFs) that provide highly modularized, auto-scalable software to replace many of the legacy network elements in a dynamic, flexible and cost-effective manner. VNFs have been shown to accelerate the time-to-market for new networks or new services such as Voice over LTE (VoLTE) or Machine-to-Machine (M2M) communications, resulting in savings of 50% or more. Yet to truly boost top-line growth through service agility—the ability to create, change, scale or tear down a service—MNOs need the ability to integrate existing legacy systems with Virtual Network Functions (VNFs) and software-defined elements in an automated fashion. This study introduces the concept of a new function to enable service agility: The **Automated Configuration Management Platform (ACMP)**.

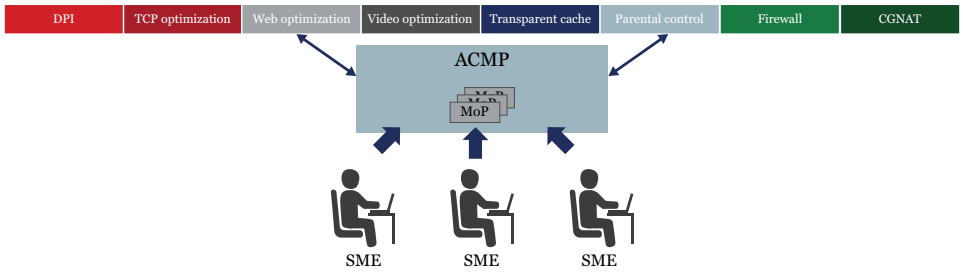


Figure 1: Automated Configuration Management Platform

ACMP translates Methods of Procedures (MoPs) into configuration files that are pushed over open APIs, proprietary interfaces, protocols and Command Line Interface (CLIs) to any legacy or virtualized network function. An Automated Configuration Management Platform compounds virtualization and automation benefits to achieve shorter time to revenue, lower risk and higher agility of service creation. This study quantifies the impact by examining the following use cases:

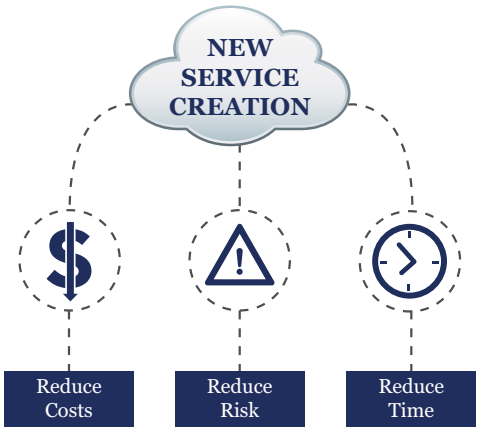
- Provisioning of new Enterprise customer for VPN-as-a-Service – operational savings of 76%-87%;
- Launch of a Mobile Virtual Network Operator (MVNO) – operational savings of 92%-97%;
- Provisioning of multiple sponsored data or zero rating plans 83% - 97%; and
- First to market services such as demand or congestion-based charging, niche and micro-services or enterprise traffic management self-services.

INTRODUCTION

Mobile network operators (MNOs) are currently under pressure to grow their network capacity and services while simultaneously containing costs. Mobile data traffic is steadily increasing at a 45-55% CAGR[1]. Revenue from voice and data, however, remains flat as MNOs try to attract new 4G customers and compete against free or ad-sponsored over-the-top (OTT) services.

MNOs are unlikely to generate more revenue from existing voice, data and text services, as OTT providers such as Skype and WhatsApp have undercut the market price for those services by using free or ad-based revenue models. Instead, MNOs must look to new service creation to generate revenue, while at the same time reducing the cost, risk and time-to-market for these services.

The mobile industry recognizes that new service creation is costly and inefficient in a network that relies on monolithic, proprietary, rigid, specialized appliances. Under the current network environment, any new service requires months or years of planning, deployment and testing before it comes to fruition. Traditional networks have to undergo massive configuration changes even for the introduction of seemingly trivial new capabilities.



Software Defined Networking (SDN) and Network Function Virtualization (NFV) are two very important architectural frameworks aimed at simplifying and improving network flexibility. Most MNOs have trialed or already deployed commercial SDN and NFV products in their networks.

SDN and NFV, by themselves, do not have enough of an impact in achieving positive business cases in legacy environments. Legacy multi-vendor integration is slow, painful and error prone and hinders service acceleration. Automation is the missing link that allows rapid service iteration in a legacy / hybrid environment. An Automated Configuration Management Platform (ACMP) can unlock this situation by creating repeatable, high velocity automated integration recipes.

Virtualization and automation allow MNOs to think about service creation in new ways:

- What if launching a new service took hours or days instead of months?
- What if it costs thousands of dollars, rather than millions, to launch a new service or update an existing one?
- What if there were virtually no sunk costs when a new service failed, and resources could be easily re-allocated?

The Challenge of Creating New Services in a Legacy Mobile Network

A mobile network basically consists of the Radio Access Network (RAN), backhaul and core network. The core is the heart of the network, where subscriber information is kept and where telco services are created and maintained. Within the core, the GiLAN is the collection of network functions that create and maintain data services.

Traditionally, core networks rely on the Policy and Charging Rules Function (PCRF) to provision and manages rules for the handling of data traffic and services in the network. A PCRF may enforce rules such as virus checking, adult content filtering, web optimization, video optimization, caching, throttling, etc. Each element typically polls the PCRF sequentially to understand what rules must be applied to the traffic.

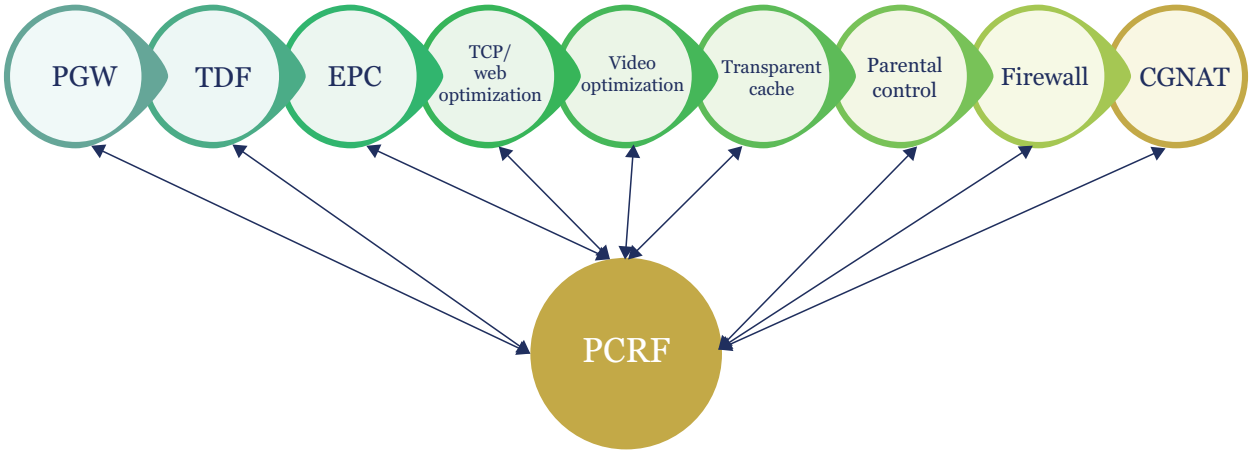


Figure 2: Traditional core network service creation

This architecture is costly and complex, and results in risky and lengthy service creation processes. For example, in a 3G network, data traffic often ends up traversing many elements—GGSN/SGSN, PGW, SGW, load balancers, DPI, cache, adult filtering, TCP/web/video optimization, firewall, etc.—each of which has to be told whether or not to inspect the packet or perform any manipulation before passing it to the next hop. Even a minor change in an existing service requires configuration changes for each network element that sees the traffic. In a 4G network, the elements may be different but the complexity is the same.

The result of this additional layer of complexity has a profound impact on service creation in a mobile network:

- Any new service requires major changes to network architecture, configuration and integration across a host of proprietary elements, resulting in long and expensive rollouts;

- Innovation is impeded, as new services represent a significant risk in terms of time/cost investment with manual intervention causing many possible errors;
- Even "me too" services are delayed by long evaluations and business cases from many layers of the MNO's organization.

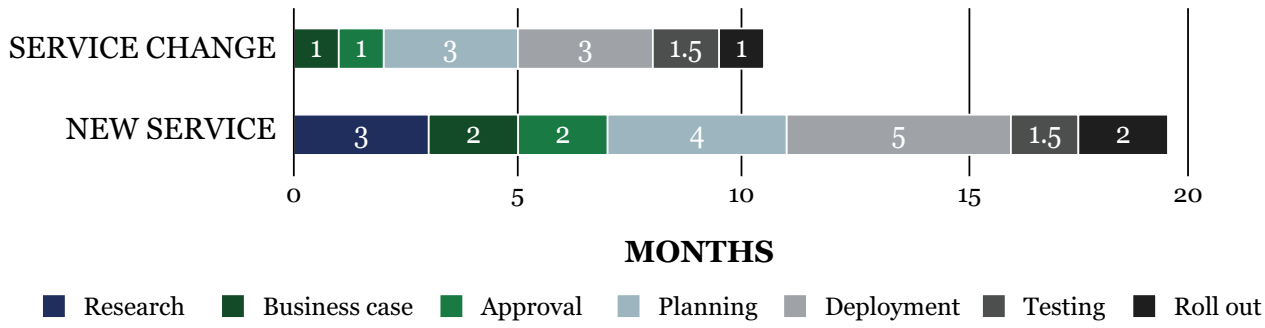


Figure 3: Typical duration in months of new service introduction and service change in wireless networks

In a legacy network, the timeframe for launching or altering services can be 18 to 24 months, with a cost model running from hundreds of thousands to millions of dollars. If new service rollouts or service changes among MNOs are rare, it's no wonder; the cost, delay and impact of launching new services can have a negative effect not only on the network but also the MNO's bottom line.

Can Virtualization Alone Increase Service Velocity?

A fully virtualized core network can dramatically increase an MNO's ability to reduce their cost structure and react to traffic spikes in an elastic and seamlessly scalable manner. In determining whether virtualized mobile networks can lead to higher service velocity, however, let us examine the most advanced NFV commercial deployments today, which feature vendor solutions for virtual Evolved Packet Core (vEPC) and virtual IP Multimedia Subsystems (vIMS).

vEPC with integrated GiLAN services is a great candidate for increasing service velocity. vEPC can provide a microservices architecture where policy, charging and service function chaining are now efficient and cost effective. vEPCs are popular for launching new MVNOs, enterprise customers or as an EPC alternative for new and non-traditional applications like network overlays for IOT communications.

vIMS is an attractive option to enable self-contained Voice over LTE (VoLTE) and voice-over-wifi services. The caveat with vIMS is that, in order to achieve cost reduction and service acceleration, a single-vendor solution with some degree of proprietary development is usually required.

vEPC and vIMS are successful in fulfilling agility and elasticity promises when deployed stand-alone

in a greenfield environment but this impact is reduced in a hybrid network where multi legacy vendor integration is required.

SDN and NFV open source and standardization bodies are trying to resolve this issue from opposite ends. They both refer to orchestration as the means to resolve service integration, but this term covers different realities in each instance.

Neither SDN nor NFV currently allow for automated, fast-paced service creation. Furthermore, because the standard interfaces with legacy networks have not yet been defined in many cases, an SDN/NFV environment implies a greenfield implementation or heavy customization to be commercially rolled out.

Choosing a single-vendor approach to virtualization has been shown to save MNOs time and money. When deployed in a hybrid multivendor environment, however, the cost and project duration of virtualization increase significantly.

SDN and NFV are intended to help MNOs reduce their dependency on vendors, not increase it. So, how do we enable service velocity without relying on single-vendor, proprietary approaches?

MNOs that are struggling to create a positive business case for new services face many challenges. If they cannot show significantly lower costs and higher revenues, the exercise rarely moves from trial to commercial deployment. While the CAPEX and OPEX reduction case for virtualization is well understood by most MNOs, the revenue part is less obvious.

Revenue growth hinges on either usage increase of existing paying services or launch of new services. Current projections show a decrease in revenue for existing traditional mobile services (voice, messaging, browsing) even though the traffic keeps increasing. The only way to significantly increase revenue is in new service launches. We have looked at the challenges associated with service creation in traditional networks; let's now examine what conditions would enable operators to launch profitable new services.

- If the cost of creating a new service drops from millions of dollars to thousands, the operator no longer needs an addressable market of millions of potential users to realize its break-even targets, creating more opportunities for new services.
- Likewise, agile service creation allows MNOs to quickly decommission unprofitable services in an efficient and non-disruptive manner.
- Services targeted to small, loyal communities can be extremely profitable, especially in markets where there is limited or no existing competition.
- If the cost to onboard new wholesale customers, either MVNO or enterprises is effectively negligible, what is the impact on the ecosystem?

In all these instances, the missing piece to realize agile, elastic service creation in multi-vendor hybrid environments is automation.

Virtualization + Automation: A Winning Combination

Automation is a great time-to-market accelerator when launching new services. Cloud enablement, application creation and service provisioning are the three main areas that benefit from automation with service provisioning providing the largest benefit.

1. "CLOUD ENABLEMENT"

The network operator must cloud-enable its data center(s) using VMware, OpenStack or a similar cloud Operating System platform. This operation requires manual planning and design, but configuration and deployment can be fully automated.

2. CREATE APPLICATIONS

Once the cloud OS is deployed, applications can be created using physical or virtualized instances of elements of the traditional GiLAN (e.g., EPC, IMS, DPI, browsing gateways). The deployment of VNFs can be automated using an orchestrator but requires some proprietary integration today. The most velocity is observed with single-vendor deployments.

3. PROVISION SERVICES

Once VNFs have been created and assembled together, they can be configured using NETCONF/YANG or other methods. This configuration is important because it is the step where the policies are provisioned for the handling of traffic under certain conditions. This operation today is largely manual and responsible for most of the service creation's lengthy process.

Challenges with Provisioning in Hybrid and Multi-vendor Environments

The biggest problem to this approach occurs when network operators try to create new services in a hybrid physical/virtual environment. Configuration management is a highly proprietary operation and not many physical applications support NETCONF/YANG or a configuration agent (puppet, Chef…). This leads to manual configuration and testing of both the physical network functions and their links to other physical or virtual functions.

The main reason for long duration in service introduction is too many network elements from too many vendors. Methods of Procedures (MoPs) provide a means to capture in great detail what needs to be done, on which element, at what time and in which sequence in order to enact the configuration change. But MOPs are usually word documents ranging from a few pages to several hundred. They are slow, require long planning time, are susceptible to manual mistakes and are not very scalable. When new vendors, new elements or even upgrades to an existing element are introduced, MoPs usually need to be rewritten and tested before being enacted. As a rule of thumb, for every service change, a MoP is created for each network element that is impacted.

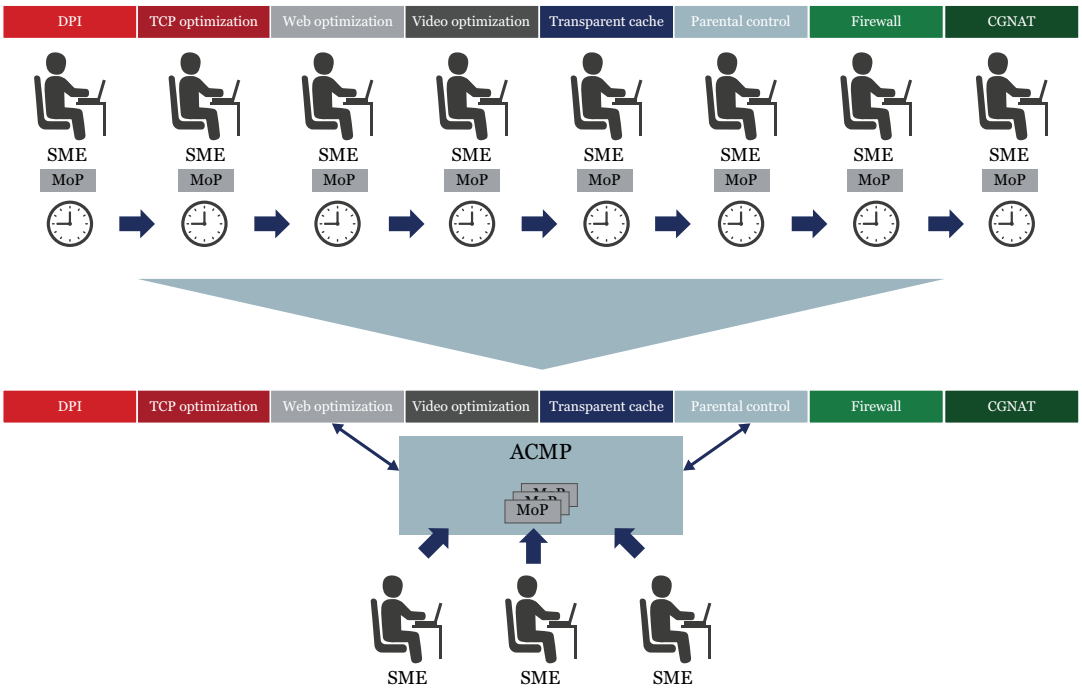


Figure 4: From iterative to automated service provisioning

Clearly, network operators are interested in automating service provisioning and testing, but the barriers to attain this goal are many:

- There is no standard interface or flow for system and service configuration
- Vendors have multiple CLIs, GUIs and other interfaces that require subject matter expertise for elements configuration
- Multiple protocols are used depending on the network element, the vendor or the intent (NETCONF, SOAP/XML, SNMP, SSH/CLI, REST, etc.)
- Most network functions don't support Configuration Agents (Chef, Puppet, etc.)
- Operators have subject matter experts (SMEs) to perform these manual operations, but not the programming resources to automate them
- Programmers often have limited understanding of the network elements and network services
- SMEs lack the programming skills to create complex, automated scripts for service provisioning in multivendor environments

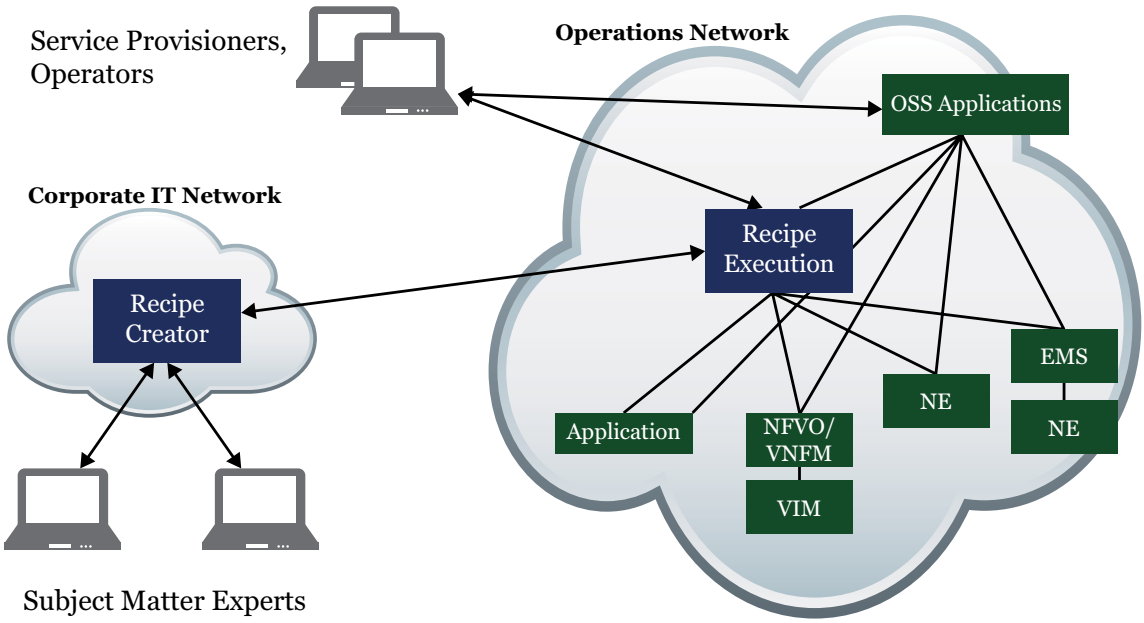


Figure 5: Automation Platform

What Can an Automated Configuration Management Platform do to Enable Rapid Service Creation?

Service provisioning errors often arise from the manual execution of repetitive, complex processes. An effective way to reduce these errors is to record successful deployments and automate their subsequent re-deployments. It stands to reason that enabling fast, error-free, service creation or modification, without the need for programming, in multi-vendor physical and virtual environments would help tremendously in terms of service velocity and agility.

Service Automation: Recipe Creation

Network operators can simplify and accelerate service creation by using an abstraction layer to create specific configurations based on best practices that can then be pushed out to all appropriate network elements at once. This can be achieved with a platform that has the ability to record all of the steps in a MoP, including the configuration artifacts, in its native form (CLI, configuration file…) using a native interface (NETCONF, SOAP/XML, SNMP, SSH/CLI, REST, etc.). Each method for each element and vendor could then become a modular ingredient from which the network operator could create a service using a specific "recipe."

Service Automation: Recipe Execution

Once recorded, these MoPs would be stateful and could be re-deployed at will from a centralized configuration manager that would translate the recipe back into specific configuration files pushed individually and in the right sequence to the proper network element. The provisioning of each

ingredient for each device would still be a manual operation, but it would only need to take place once, with automation handling any future changes.

Using an automation platform not only dramatically increases service velocity, but also provides other business advantages:

- Direct OPEX savings by reducing time to deploy, test, edit and launch as much as 90%;
- Reduced risk through the use of a single master configuration for subsequent deployments;
- Extremely agile re-allocation of resources to meet shifts in demand, turning down of services, etc. with minimal impact to the network.

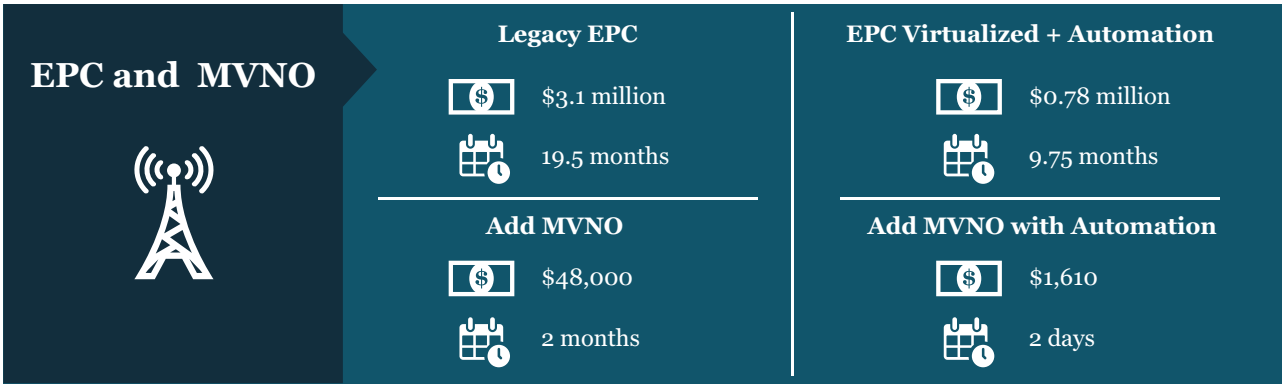
Ultimately, automating service creation processes leads to more revenue by creating new services around microsegments, reducing the cost of service creation, enhancing response to new threats/ opportunities and accelerating time to market.

Accelerate and Reduce Costs of Existing Services

The use cases in this section are composite examples of commercial deployments elaborated from individual case studies and their corresponding project plans. The use cases are anonymous for confidentiality purposes, but the magnitude in savings and revenue generation is preserved.

The products deployed are a virtual EPC with collapsed Gi LAN functionality and a stateful Automated Configuration Management Platform that records and automates third-party physical and virtual element configuration.

1. Greenfield Mobile Network Operator (MNO) or Mobile Virtual Network Operator (MVNO)



In this example, a large tier one operator network looks to host a new MVNO. For our illustration, we will focus exclusively on the deployment of the virtual Evolved Packet Core (vEPC), together with the integration of legacy network elements including the AAA server, HLS, DNS, Router, OCS and OFCS from a representative panel of vendors including Ericsson, Cisco and others. The vEPC includes the policy control and routing function (PCRF) as well as web and video proxy/optimization.

BEFORE VIRTUALIZATION + AUTOMATION: 19+ MONTHS AND $3+ MILLION

This operator had already implemented MVNOs in the past. Their baseline project plan from planning to lab to commercial deployment--was more than 19 months. The deployment required 20 Full Time Equivalent (FTE) personnel from the operator and 10 from the various vendors involved for the duration of the project, generating an OPEX cost of $3,139,500.

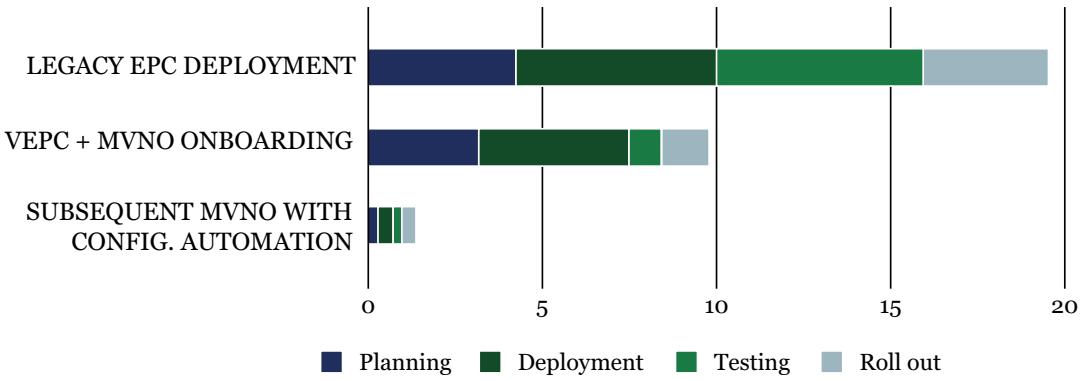


Figure 6: MVNO launch project plans

AFTER VIRTUALIZATION: NEW MVNO IN NINE MONTHS

In this case, the deployment of the vEPC for a new MVNO across four production sites took a little over nine months. The deployment involved nine separate MoPs that were recorded and executed sequentially. It entailed the replacement of existing systems by a virtualized EPC and automation system. The 50% reduction in time was due primarily to the ease of deploying COTS hardware instead of dedicated appliances and the automation of VNF instantiation, both of which simplified planning and deployment (from specialized commissioning, site study, footprint, heat dissipation, cabling, switching and load balancing to COTS racks, switches and servers) and testing (from point to point, unit, system and service testing to automated programmable testing).

AFTER VIRTUALIZATION + AUTOMATION: FROM TWO MONTHS TO TWO DAYS

The most dramatic time savings occurred in subsequent MVNO deployments. Here, integration points, interfaces and MoPs for configuration had been recorded in the Automated Configuration Management Platform (ACMP) during the first deployment. In this case, a MoP had been recorded into recipes

for each of the PCRF, P Gateway, Firewall, DNS, DPI and HLR/HSS elements. The "templatizing" of IMSI ranges to APN, new authentication, policies and charging models allowed the operator to simply modify existing recipes in the ACMP and automatically push the configuration files to the appropriate elements/functions. The result showed dramatic acceleration, with a new MVNO instance created, deployed in lab, tested and rolled out commercially in only two days!

	New legacy EPC	Virtualized EPC + new MVNO	Legacy Subsequent MNVO	Virtualized + Automated Subsequent MVNOS
Duration	19.5 months	9.75 months	2 months	2 days
Effort (vendor man/day)	3900	975	60	2
OPEX (vendor)	$3,139,500	$784,875	$48,300	$1,610

Table 1: Legacy EPC & MVNO

Beyond the obvious cost benefit, implementing an automated service provisioning platform enables new levels of organizational agility. The ability to launch a new MVNO in two days with only one FTE allows operators to create a wholesale operation and multiply these types of deployments.

2. Sponsored Data: URL / Domain Zero Rating

Not every task that operators have to perform is time consuming, inordinately complex or resource intensive. Sponsored data and cooperation with content providers require light integration and the ability to identify and isolate specific traffic origin.

Zero rating of a URL or Internet domain is a good example of a mundane, recurring task that wouldn't necessarily register as a large cost individually, but collectively represents a hindrance to service innovation because of maintenance costs. Operators routinely perform these operations as new agreements are made with content providers for permanent or temporary service promotions.

Zero rating a service entails creating a class of service with a list of domains/URLs associated with specific subscriber groups, which end up impacting configuration of many core network elements such as PCRF, charging functions, HSS, EPC, etc. While the operation itself is not very complicated, the complexity arises from having to coordinate the configuration of MoPs across various devices and interfaces.

This would be acceptable if it was a relatively low occurrence, but operators end up provisioning these type of rules on a nearly daily basis as new services are launched or new service URLs change over time. Maintaining the list of URLs with special actions (e.g., zero rate, premium, throttle, ban) becomes problematic when thousands of URLs are provisioned. In relatively small networks with a low number of data centers (i.e., two or three), the effort to zero rate a URL might be considered trivial. For tier one operators with massive and geographically distributed data centers, however, even trivial changes can be costly. For this use case, the operator has a medium-size network with two data centers, each with a P Gateway and PCRF that need to be configured.

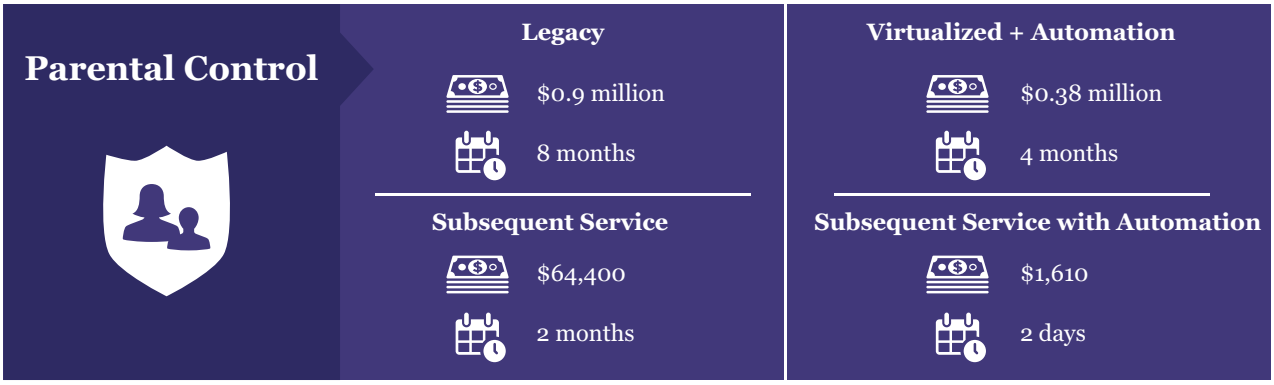
	Legacy Zero rating	Zero rating with ACMP
Effort (man days)	12	2
OPEX	$9,660	$1,610

Table 2: Legacy Zero Rating

AUTOMATION YIELDS 97% OPEX SAVING

Using an ACMP does not save very much time for the first incidence, but the savings are more apparent as subsequent URLs are introduced or services modified. The only configuration change is a new rule created to provision the URLs, domain and wildcards. The testing and deployment is then automated, which drastically reduces the operational costs, typically associated with the review of test procedures and results. An added benefit occurs in the error suppression introduced by the automation and the programming of the configuration files across vendor elements.

3. Parental Control



An intermediate use case in terms of effort is the introduction of a new customer plan. Parental control and security services have lately gained interest as more teens access Internet services from mobile devices—one study found that the customers who purchased Parental Control or Anti-Malware services improved service "stickiness" and reduced the annual churn rate by 2.4%[2].

Most new product offerings from vendors centered around security and control of data traffic on mobile devices are fairly unidimensional, providing white/blacklists of apps and domains for underage users.

This simplicity of features is due in many cases to the extraordinary complexity of integrating legacy systems to fit within an adult filtering framework.

The challenge for network operators is to enable a set of holistic rules that support parental control; for example, creating a rule that, upon attempt to access a blocked site, would display a warning, send an alert to the parent/control authority or bar user access. That set of instructions traditionally impacts the Deep Packet Inspection (DPI) function, the browsing gateway, the PCRF, the adult filtering database and the subscriber profile. It is possible to enable such a service if all these elements are provided by a single vendor in a single platform or after much integration. Additionally, integration can extend beyond the network to IT systems if credit card verification or other proof of age are required.

Multiple vendors lead to multiple provisioning interfaces. Factor in a potential list of thousands of web sites, domains and apps to be updated on a monthly basis, and this use case quickly becomes too complex or onerous to deploy. As a result, operators either look at new vendors to collapse many functions in order to deploy these types of services, or once deployed in multi-vendor environments, keep them fairly static and simple. In many cases, adult filtering tools end up managing blacklists for an arbitrary age group.

AUTOMATION REDUCES COMPLEXITY, INCREASES AGILITY

In the example of parental control, an automated configuration management platform has a multiplier effect on agility. The platform can record the individual sub-procedures to:

- Provision DPI to classify and flag data traffic based on white/gray/blacklist by looking up the adult content database;
- Provision the PCRF to look up subscriber profiles in the HSS and create rules to alert and drop traffic that is gray/blacklisted for the underage subscriber;
- Send HTTP traffic to the browsing gateway to look up in the PCRF;
- Alert and stop traffic in the gray/blacklist for underage viewers.

Once recorded, these procedures can be blended into recipes to create any form of automated access control based on subscriber, content or destination profile. The system automation takes care of abstracting specific vendor configurations, their sequence and interface, and provides an automated service configuration tool that has practically no incremental cost, delay due to complex rules or provisioning errors.

	Legacy Parental Control	Virtualized Parental Control	Legacy Subsequent Service	Virtualized + Automated Subsequent Service
Duration	8 months	4 months	2 months	2 days
Effort (vendor man/day)	480	160	80	2
OPEX (vendor)	$901,600	$386,400	$64,400	$1,610

Table 3: Legacy Parental Control

In our example, we presume the existence of legacy devices from a variety of vendors, a modularized virtual browsing gateway, filtering logic and DPI, with the ACMP recording provisioning methods for each platform. For legacy configuration management, the creation of a new class of services has to be provisioned for the browsing Gateway, DPI, PCRF and HSS over two sites; this can easily take eight months and exceed $900,000 in budget. The collapse of functions and replacement by a virtualized browsing gateway/filtering/DPI function yields a 57% cost saving in OPEX and a 50% reduction in project time—from eight months to four months. This is a good result, but an operator would not necessarily replace their browsing gateway/filtering function by a virtualized version just for these savings.

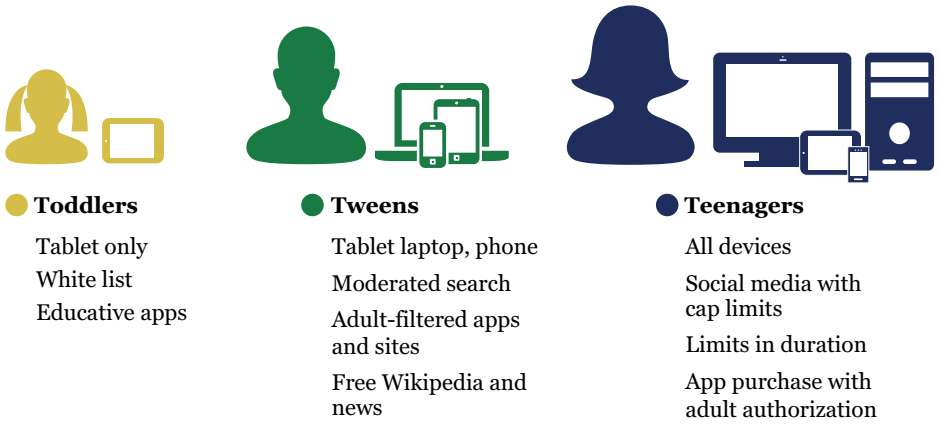


Figure 7: Example of service segmentation enabled by automated service provisioning

The benefits show themselves more plainly after the integrated systems are used for new service launches. If service provisioning is drastically reduced, there is an incentive to segment the user base further by creating different classes of services for different age groups. For example, we can imagine a class for toddlers and young children who only have access to whitelisted educational apps on tablets, while tweens could access Wikipedia and moderated search services and teenagers would have access to social media sites such as YouTube and facebook but not to banking, gambling, etc.

AUTOMATION LEADS TO 95% ACCELERATION

After the creation of this first class of service, the next iteration sees drastic acceleration due to the automation of service provisioning in the ACMP. In this case, only variables (e.g., location of list of domains, age groups, etc.) are updated in the recipes, while the mechanisms and interfaces remain the same and the testing and deployments are automated. In our legacy example, a new class of service becomes a scaled-down version of the first deployment, yet still requires over eight weeks to deploy, as configuration is stateless in the legacy platforms. The savings when using the ACMP is remarkable, reducing service creation time from eight weeks to two days with an OPEX cost in the $1,600 range.

4. Enterprise VPN as a Service

Virtual Private Network (VPN) as a Service is an increasingly popular use case for cloud deployment. Until recently, VPN deployments required specialized hardware and applications (e.g., firewalls, proxies, NATs) at each enterprise site to enable the safe and secure exchange of data between these points. The deployment of hardware-based VPNs is expensive for setup and management and lacks flexibility as new locations are rolled out or new protocols emerge.

VPN as a Service is growing in popularity, fueled by the flexibility and agility granted by cloud deployments. In this use case, operators are able to create VPNs on demand by dynamically allocating cloud resources to enterprises customers in real time. Yet even in a virtualized cloud environment, VPN as a Service still requires cooperation and integration between several Layer 3 VPN vendors.

	Legacy VPNaaS Enterprise Onboarding	Virtualized + ACMP VPNaaS Enterprise Onboarding
Duration	2 weeks	2 days
Effort (vendor man/day)	10	2
OPEX (vendor)	$12,075	$1,610

Table 4: Legacy Enterprise VPN as a Service

FROM TWO WEEKS TO TWO DAYS

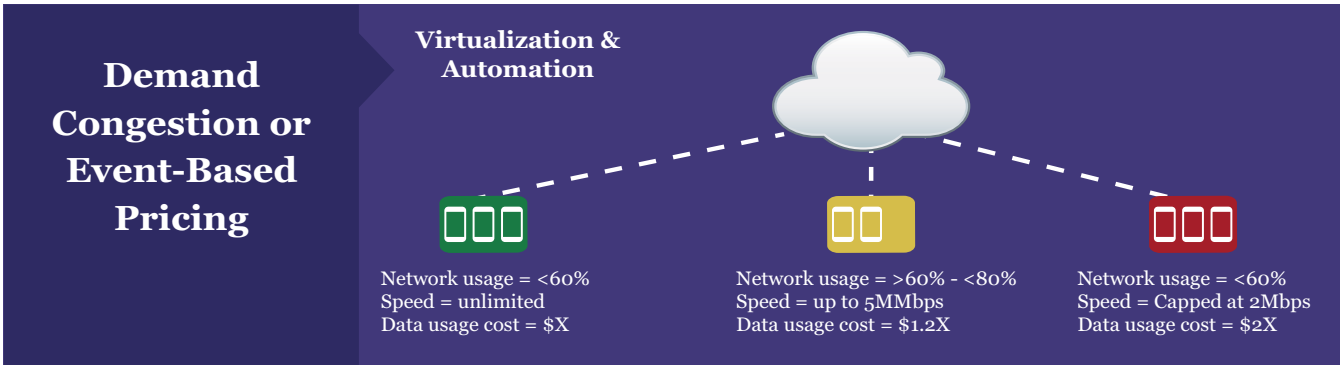
In this scenario, all of the network elements are already deployed and configured for VPN as a Service; we are just looking at onboarding a new enterprise client. A typical project takes, on average, two weeks to complete and is essentially the same set of MoPs updated for the specific enterprise, with minor variations. In a legacy environment, operator supervision is necessary to control and test the MoP implementation. The ACMP again compresses these timelines considerably, from two weeks to two days, with OPEX savings over 85% due to testing automation.

New Services Enabled by Automation

Providing a programmable platform for automating service configuration can dramatically reduce the costs, time-to-market and errors associated with repetitive, multi-interface processes. Yet these benefits are only one side of the equation. It is difficult to build a business case on service agility or revenue acceleration alone. One has to look at the ACMP as a best practice that accelerates and multiplies service opportunities. The real benefit of automating service provisioning in virtualized and hybrid environments is the opportunity to launch services that were otherwise too daunting or too expensive to launch.

In this example, a modularized EPC with distributed P Gateway functions would be able to collect real-time analytics from the mobility core and apply traffic and charging rules accordingly. The costs and timelines for such a project would be comparable to those shown in Greenfield s. Interestingly, once this type of system is deployed, it introduces the flexibility and speed to create services that were not possible before.

1. Demand, Congestion or Event-Based Pricing



Until recently, operators' charging plans have been fairly static and the main variable has been traffic. This is due mostly to the fact that the provisioning of price plans and conditions require coordination across many network elements in the GiLAN, while voice and text traffic are known quantities and predictable as a function of demographics and customer acquisition. The emergence of data services and, more specifically, social media and video streaming have drastically changed this paradigm. It has become more difficult to plan traffic growth at the macro level as viral campaigns and new OTT service launches have caused data traffic to be increasingly unpredictable.

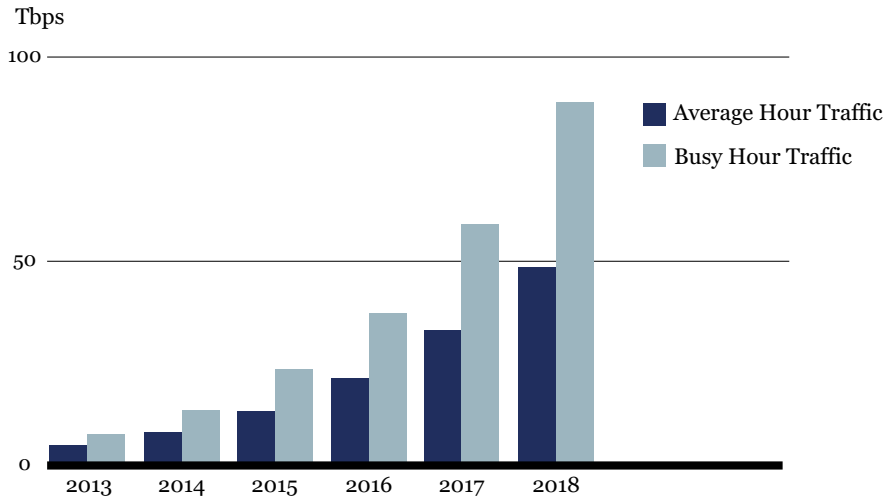


Figure 8: Ratio busy / average hour traffic in wireless networks[3]

As the graph above illustrates, the disparity between peak and "normal" traffic is growing and makes network conditions less predictable and network capacity planning more difficult and expensive. Until recently, network capacity planning looked at the peak capacity needed in a given year and tried to meet this demand with the addition of network capacity (e.g., spectrum, coverage, penetration, traffic management). As the difference between the average and peak usage increases, it becomes difficult to justify investment for what will essentially be underutilized capacity. As we can see, busy-hour traffic is projected to grow over time, representing 166% of average hour traffic in 2014 and 183% in 2018.

A good example of unpredictable traffic is the success of apps like Periscope and Meerkat, as illustrated during the Mayweather-Pacquiao boxing match[4], where a large number of people in the audience started live-streaming the match. Most of us think of data services in the context of streaming or downloading, but on that night, for these apps users, upstream traffic exceeded downstream.

The implications for network operators are multiple:

- Traffic can grow manifold for events with a large audience;
- Data traffic surges can be downstream and/or upstream;
- Sudden traffic surges can cause congestion if not properly managed.

WHAT WOULD IT TAKE TO HAVE DEMAND-BASED PRICING?

With net neutrality provisions being reviewed in many markets, active traffic management (banning, throttling, prioritization, etc.) is not possible to curb or control these unexpected traffic spikes. The most efficient means to allocate bandwidth justly among users in a given location may become dynamic pricing. Many consumer goods and services are based on demand and supply: airline tickets, Uber rides, hotel room prices vary based on occupation rates. What would it take to have demand-based pricing in wireless?

First, all traffic would need to be aggregated and visualized in real-time per location (or cell). Then, usage has to be aggregated per subscriber, as a single device can have many data connections. Then, the utilization has to be compared to the cell's capacity to determine the utilization threshold. Finally, pricing plans must be created based on various thresholds; for our purposes, we'll call them the green, yellow and red thresholds.

Congestion status	Green	Yellow	Red
Network usage	<60%	>60%, <80%	>80%
Speed	Unlimited	Up to 5 Mpbs	Capped at 2 Mbps
Data usage cost	100%	120$	200%

Table 5: Demand-based Pricing Based on Cell Capacity

In a legacy network, provisioning this type of rule is fairly difficult and rigid. Real-time congestion information is available in the RAN, but not in the core network without introducing costly probes. Virtualized environments allow for a more flexible approach, where intelligence and virtualized network elements (such as the PGW) can be pushed closer to the edge. This implementation allows for a more coherent picture of network utilization in real time, together with location and subscriber data, to enable dynamic pricing models. The green, yellow and red indications can be made available as the user is browsing and can vary over time and location.

Additional benefits can be derived from changing usage costs based on promotions (e.g., browse free today on your birthday or 50% usage on Facebook). These complex use cases can only be enabled with stateful configurations and associated traffic management rules being centrally provisioned, managed and pushed to the appropriate network elements.

THE VALUE OF DYNAMIC PRICING

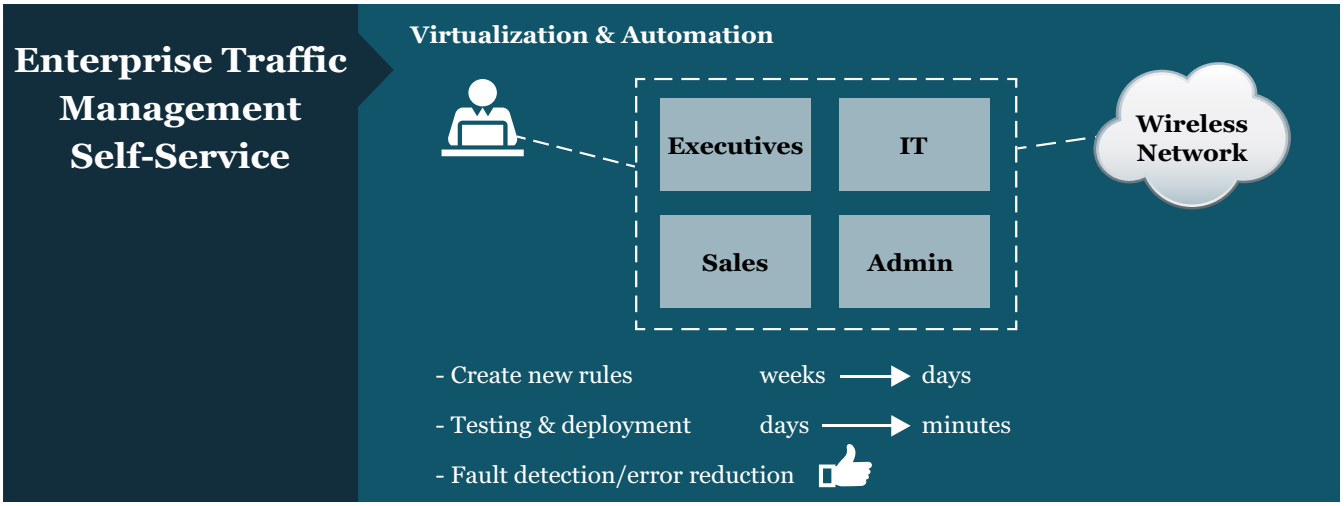
As the price plan becomes dynamic, the trigger mechanisms can also vary over time and location. In our example, elements impacted by the creation of such a service could be the P Gateway, the PCRF, the charging function, the HSS and the DPI. In a legacy network, the level of integration and the necessary real-time collaboration between the RAN and core network would usually prevent such a service from being launched, so we have little in terms of basis for comparison.

The benefit of these pricing plans is twofold. For one thing, we know that providing sliding prices and

speed increments is a better method than limiting traffic speed with a hard ceiling or "best effort" where everyone's traffic is impacted equally for everyone without choice. This method provides subscribers in a congested network with a choice to use less data-intensive apps (stop streaming HD, for instance) or pay a premium to increase/conserve their quality of service. Naturally, as more users elect to reduce their usage to avoid a price increase, this also reduces the overall network congestion level.

As a second benefit, a number of users will likely elect to continue using data-intensive services and may even be prepared to pay a premium for better quality of service, which will increase overall revenue.

2. Enterprise Traffic Management Self Service



It is no secret that customer relationship management and charging is a huge challenge for any consumer-facing organization. On the one hand, physical or live interaction is very expensive (e.g., $35 to $45 for a call to a customer center). On the other hand, automated or online systems offer a limited range of options. Consumers and enterprises increasingly wish to control their online experience. This is reflected in the increase in self-serve applications launched by network operators globally, which show up-to-the-minute data consumption, roaming charges, etc. In many cases, these operator-branded apps provide non real-time data, with accuracy varying from a few hours to a few days.

This limitation is the direct result of network elements' limited configuration options and the length and cost associated with changing these configurations. This problem is even more acute for enterprise customers, as they rely on external providers for their connectivity and have limited capacity to manage data traffic without specialized equipment.

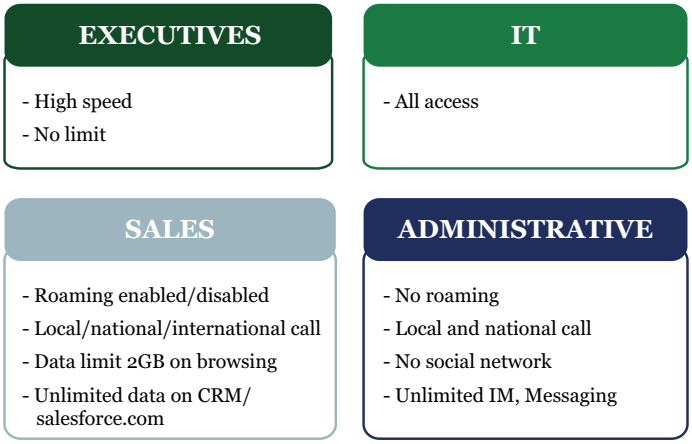
ACMP CUTS DEPLOYMENT TIME FROM DAYS TO MINUTES

We have seen how creating new traffic management and charging rules is a costly and long operation in legacy networks and how this can be reduced drastically with automated service configuration management. This model is probably even more important for operators that wish to extend their

services to SMBs and larger enterprises. These customers have sophisticated traffic management needs that are not easily satisfied as a subset or combination of rules deployed in legacy systems. In such a case, we have seen that the ACMP can reduce the time to create new rules from weeks to days and testing and deployments from days to minutes. More importantly, beyond the gain in time and reduction in costs, automation ensures fault detection, which eliminates the risk of service failure due to technical mistakes.

The agility afforded by the ACMP can be passed on to large enterprise customers through a dedicated interface and sandbox for their IT department. From there, enterprises can directly select and change service configurations with or without the operator's assistance, depending on the service level agreement. SMBs could profit from more service flexibility as well by selecting their plans, which could then be tailored to their needs.

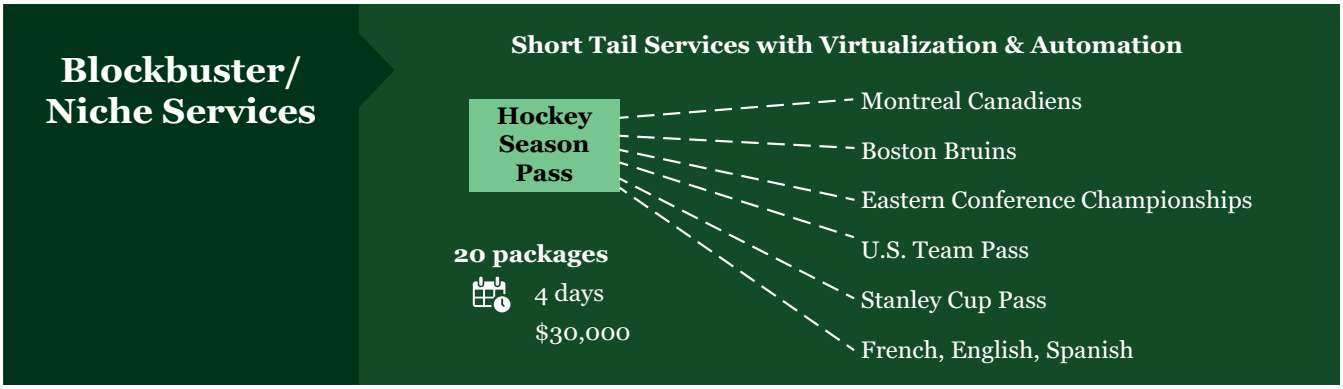
It is not unusual for enterprise network administrators to limit or manage different classes of services for different internal clients:



Setting up these different profiles internally can be cumbersome because the enterprise IT system is not integrated with the operator's IT system. Of necessity, enterprises end up fitting their internal customers into operator-offered packages, which in many cases do not align completely with their needs. Creating specific packages for specific user groups is expensive and requires much time in design, integration and roll out. As a result, these implementations remain limited to large enterprise customers, usually those with hundreds of thousands of employees.

EPC virtualization, together with automated service provisioning management, allows service providers to drastically reduce the costs and the complexity associated with the creation and maintenance of these traffic plans. One could even push the model further, envisioning that large enterprises will, in time, implement their own EPC and provisioning systems to integrate with the operator's, enabling direct and total control of enterprise traffic within the limits imposed by the mobile network operator.

3. Blockbuster or Niche Services?



Network operators overwhelmingly cite service agility as one of the main reasons for deploying virtualized, software-defined architectures. Service agility is not only about launching services faster; it is also about the creation of an ecosystem of microservices.

Until now, operators had no choice but to adopt the "blockbuster" strategy in terms of service creation. The level of preparation, planning, approval, testing and costs of deployments were so high that only services with the widest mass market appeal could conceivably support a positive business case.

The introduction of an automated service provisioning platform changes this paradigm by substantially reducing cost and time-to-market and by removing risks due to human errors with stateful service templates and service provisioning and testing automation.

If we accept that the ACMP, once used for a service, can enable the launch of new services in a matter of days and a few thousand dollars of direct cost, the risk of failure of an individual service from a business-case perspective becomes much lower. At that point, operators might want to invest in creating a portfolio of services using a short tail/long tail strategy. Some services in the long tail are known to be massively attractive and provide most of the revenues, so these will necessitate most of the marketing investment for promotion, while other services (short tail) might be targeted at a much smaller portion of the population but demand much lower investment and thus generate better margins.

3.1. Short Tail Services

Traditional short tail services revolve around mass-market public interests such as sports, politics and global events. Canada, for example, is a country of hockey fans and network operators vie to secure the rights to retransmit the regular season and Stanley cup games. These license rights cost billions of dollars to secure for TV, online and mobile retransmissions. In 2012, Rogers, a leading network operator in Canada, paid $5.2 billion to acquire part of these rights for 12 years[5].

On the mobile front, operators traditionally sell season passes, where a subscriber can buy access for the full season to all games for $199.00 per season or $29.99 per month. Language is region-dependant

and, if you live in Ontario, you will get English commentary by default, even if you'd rather watch your games in French.

Certainly, the high cost of securing these rights qualifies as a blockbuster approach for TV, which is still the predominant vehicle for their consumption. The emerging mobile viewership, however, could do with more choices. For instance, Rogers has launched an additional GamePlus app, which offers multiple camera angles to mobile users and the ability to interact with them. This option was compared to a "business class" offer by Rogers' CEO[6], as a start towards service and consumer segmentation.

20X THE CHOICES FOR $30,000

As we have discussed, the cost of service introduction is high for these type of new services: creating new price packages, ensuring available network capacity at game time, etc. Now, if the service had been provisioned for the initial mainstream offering using an ACMP, the creation of variants for specific targets would have been trivial in terms of incremental effort and cost. At that point, it would have been possible to create packages for the different target demographics such as:


Montreal Canadiens fans


Boston Bruins fans

Eastern Conference pass

US Teams pass

Season pass

Stanley Cup pass


French, English, or Spanish languages, etc.

All these different packages would be possible today, with little additional costs, since the rights are already purchased and the video streams are already produced for TV and mobile. All that is needed is a flexible and cost-effective configuration tool to create and maintain each variation. Each package takes on average two days to create and test and can be launched all at once or in a staggered manner. The overall budget to create 20 variations and packages should be 40 days and close to $30,000. This example, of course, is not specific to hockey or Canada and can be translated to any sport, season, tournament or country.

3.2. Niche Services

Potentially more interesting than the blockbusters or short tail services are long tail services or microservices for operators with the right strategies. In his book "The Long Tail: Why the Future of Business Is Selling Less of More," Chris Anderson, the editor of Wired magazine, makes a convincing argument that, given the ability to afford and find services and goods tailored to their specific tastes, consumers will adopt fewer best-sellers in favor of many customized experiences. We certainly see that dichotomy in mobile with the success of Apple's iPhones and the ever-changing catalogue of top-grossing apps.

The only way to economically produce microservices is with a low-margin cost structure borne by

having a single vendor with an integrated silo. This approach, however, might sacrifice the long-term flexibility of a multivendor approach, where functions and vendors can be replaced in a rapid cycle using automation and virtualization, but where service management, integration and configuration remain abstracted, centralized and stateful.

SPORTS FANS

When you go to a Formula 1 Grand Prix or a NASCAR race, the event usually lasts four days, from trial runs to the race itself. Anyone who has attended these events may have seen people walking around with wireless TV devices. These connected screens stream the event in real time, offer unique camera angles and provide replays and statistics. They're typically rented for the duration of the event for between $50 to $250.

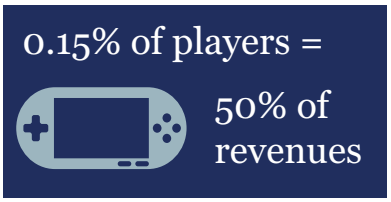
With everyone carrying a smartphone or a tablet in their pocket nowadays, and many network operators sponsoring/acquiring rights for retransmission of these events, there is no reason we couldn't have a better experience directly on our screens without having to rent an additional device. The issue is that most network operators wouldn't even attempt to create a service for a few thousand fans in a limited venue for a few days. Again, the planning, configuration, testing and deployment of such a service would far outweigh prospective revenues. Realistically, though, many fans in these venues rent devices or try to stream TV content while there, so this is really a missed revenue opportunity.

With an automated service configuration tool, creating a service variation such as this one is not only cost effective enough for such a small scale, but it is also easy to take down and archive until the next event or venue. The same templates with different variables can be used for other events, ranging from boxing matches to concerts end even charity events. When the incremental cost to serve a specific event is tens of thousands of dollars, it becomes much easier to address smaller, high-margin demographics than with one-size-fits-all price plans.

MOBILE GAMERS

Gaming as an industry has surpassed Hollywood in terms of revenue over the last five years. We routinely see gaming content consumption (game traffic, downloadable content, game video channels on YouTube) increase on fixed and mobile networks. Steam, a video game platform that connects players, normally accounts for 1 to 2% of fixed internet traffic in North America, but when Grand Theft Auto 5 and Call of Duty were launched, Steam accounted for up to 5% of peak traffic.

In the mobile industry, only 0.15% of gamers account for 50 percent of all in-game revenue[7]. In 2015, these revenues were estimated at close to $30 billion[8] with popular games like Game of War or Big Fish Casino averaging $400 per user in in-app transactions.



This relatively small population of gamers who spend a lot of money, spend it not only on games themselves but on in-game purchases. Most of this revenue flows to the game publishers and the platforms that serve them. Most hard-core gamers also spend money on anything that can give them an edge against their competitors. This translates into specialized desktops, programmable keyboards, expensive joysticks, specialized headsets… you get the idea.

In mobile, most hard-core gamers buy upgrades and add-ons within the game to enhance their experience, but little is done on the network to provide them with a better experience. Many real-time strategy or sports games are very sensitive to lag and latency. We know that mobile networks are prone to longer latency, which can be reduced with techniques such as TCP optimization or by creating dedicated bearers with specific QoS levels.

Guangdong Mobile, part of China Mobile, partnered in 2015 with Tencent, a leading multiplayer online game provider in China, to launch a dedicated offering for mobile gamers, promising up to 6x less latency. The users of the service buy games and optimization packages to enhance their experience.

While China might be the largest mobile gaming market in the world by the sheer virtue of its size, the proportion of mobile gamers is high in Europe and North America (over 50%), while Latin America is the fastest-growing market for mobile gaming. Creating a gaming package that would reduce or guarantee homogenous latency among players could be a desirable service even for casual gamers. Let's assume that the deployment of a virtual EPC is dedicated to this service and integrated with the existing core network elements, with a starting subscription fee of 80 cents per month. If only 3% of a mobile gaming population of 5 million players were to adopt the service, the revenue over a five-year period would exceed $10 million with a break even in less than four months from project start.

This is a perfect example of the creation of a service for a seemingly small demographic with a large disposable income dedicated to buy better mobile network performance. Such a service is difficult to create and justify for a small market segment, but if the capabilities to orchestrate its creation are there and all that is needed is to create a new configuration across the e-node B, P Gateway and DPI, a short project of a couple of weeks and a few tens of thousands of dollars might be sufficient to justify the business case and create a differentiating service for a high-margin group.

About Affirmed Networks

Affirmed Networks pioneered the virtualized Evolved Packet Core function even before Network Function Virtualization was invented. The company has designed its vEPC solution from the ground up to be cloud native. The benefits in scalability and elasticity have been proven commercially with over xx commercial deployments globally. Its vEPC seems like an ideal starting point for my investigation, since it has been publicly commercially deployed at AT&T, one of the most advanced and vocal proponent of service agility through software defined networks and virtualization.

AFFIRMED SERVICE AUTOMATION PLATFORM (ASAP)

ASAP is Affirmed Network's effort at automating and industrializing network element configuration management. Recognizing that network operators' SMEs wanted to rely on a single interface for configuration management while retaining complete control over each individual element configuration table, Affirmed developed ASAP as a vendor-agnostic configuration builder and manager.

The core of the platform is a transaction engine responsible for translating generic intent-based service configuration MoPs into vendor-specific configuration files or instructions, conveyed via connectors to the recipient network element in their native protocol.

A graphical user interface allows the SME to create sets of modular nested instructions that can then be used as building blocks for service creation or change. For instance, the creation of a new MVNO in an operator network is a meta-project composed of many complex discrete operations. These operations rely on detecting and authorizing a subscriber, creating price, services and QoS plans and associating the corresponding traffic control and management policies. Even if an operator does not create MVNOs very often in their network, recording the MoP in an automated system such as ASAP allows them to easily duplicate, re-use or modify any of the modular recipes in order to create a new enterprise client, a new QoS rule, or a new traffic type policy.

The benefits are multiple, ranging from the creation of an effective service configuration catalogue, to the reduction of errors from recurrent operations and the drastic compression of time, resource and costs necessary for service changes and new service launches.

For more information please go to
http://www.affirmednetworks.com/products-solutions/service-automation.

About Red Hat

Red Hat is at the forefront of open source software development for enterprise IT, with a broad portfolio of products and services for commercial markets.

That vision for developing better software is a reality, as CIOs and IT departments around the world rely on Red Hat to deliver solutions that meet their business needs. Trusted and innovative solutions that provide technology leadership, performance, security, and unmatched value to more than 90% of Fortune 500 companies, preparing them for the future of IT.

Red Hat delivers this technology to customers in two ways:

1. Offering commercial products for demanding environments

Deploying solutions from Red Hat and its partners helps customers benefit from the best of both worlds—innovative technology coupled with the services, security, predictability, and quality that their businesses demand.

2. Sponsoring freely available open source projects

Red Hat is an expert at collaboration—developing new technology, solving common problems and sharing the benefits with their customers and partners.

About Mellanox Technologies

Mellanox Technologies (NASDAQ: MLNX) is a leading supplier of end-to-end Ethernet and InfiniBand interconnect solutions and services for servers, storage and hyper-converged infrastructure. Mellanox interconnect solutions increase data center efficiency by providing the highest throughput and lowest latency, delivering data faster to applications and unlocking system performance capability. Mellanox offers a choice of fast interconnect products: adapters, switches, cables, software and silicon that accelerate application runtime and maximize business results for a wide range of markets including Web 2.0, cloud, storage, high performance computing, enterprise data centers and financial services.

Mellanox products and solutions are uniquely designed to address the virtualized infrastructure challenges, delivering best-in-class and highest performance server and storage connectivity to various demanding markets and applications, combining true hardware-based I/O isolation and network convergence with unmatched scalability and efficiency. Mellanox solutions are designed to simplify deployment and maintenance through automated monitoring and provisioning and seamless integration with the major cloud frameworks.

Hardware Based I/O Isolation

Mellanox ConnectX® adapters and Mellanox switches provide a high degree of traffic isolation in hardware, allowing true fabric convergence without compromising service quality and without taking additional CPU cycles for the I/O processing. Mellanox solutions provide end-to-end traffic and congestion isolation for fabric partitions, and granular control of allocated fabric resources.

Accelerating Storage Access

In addition to providing better network performance, ConnectX's RDMA capabilities can be used to accelerate hypervisor traffic such as storage access, VM migration, data, and VM replication. The use of RDMA pushes the task of moving data from node-to-node to the ConnectX hardware, yielding much faster performance, lower latency/access-time, and lower CPU overhead.

About {Core Analysis}

Patrick Lopez
Founder & CEO
{Core Analysis}
www.coreanalysis.ca
Blog (http://coreanalysis1.blogspot.com/)
Twitter: @coreanalysis

Patrick Lopez has nearly 20 years of international progressive experience in product and technology introduction in the United States, Canada, Switzerland, Ireland and France.

Founder and CEO of {Core Analysis}, he provides advisory services to technology vendors, board of directors, carriers and venture capital firms on OTT video. As an analyst, he presents at influential industry forums and conferences and publish an acclaimed blog, industry articles and reports. In 2016, he is chairman / speaker at 5G World Summit, NFV World Congress, and SDN / NFV Summit.

Recent achievements:

- {Core Analysis} was exclusive advisor to Opera Software in its acquisition of Skyfire for $155 millions.
- Presented at NAB (National Association of Broadcasters) on PayTV vs OTT strategies and business models.
- Presented at United Nations' ITU Telecom on convergence of broadband and broadcast.

Patrick has collaborated to various industry reports including Deutsche Bank, JP Morgan, Morgan Stanley Credit Suisse First Boston, IDC, Frost & Sullivan, Yankee group, Ovum, Informa… and has written several articles in collaboration with The Wall Street Journal, TMCNet, Wireless Week, RCR Wireless News, CNN and CNBC Europe.

Patrick holds a MBA in Corporate Management and Bachelor Degree in Marketing Strategy.

[1]Cisco VNI 2016, Ericsson Mobility report 2015
[2]Allot Mobile Trends H1/15
[3]Cisco Visual Networking Index 2014
[4]Pirates Crash Mayweather – Pacquiao fight – The NewYorker –May 4 2015
[5]Swrve survey, October 2015
[6]NHL, Rogers announce landmark 12-year deal
[7]CBC.ca article October 2014
[8]Newzoo Mobile Games Market Growth Above Expectations