PROCERA

The Evolution and Virtualization of Big Data:

Structuring Subscriber Data for Broadband Operator Transformation and Profitability

Virtualization is empowering operators to transform their business. Successful transformation requires better data.

EVOLUTION OF THE BROADBAND LANDSCAPE

The broadband landscape is undergoing a dramatic shift as consumers begin to utilize higher bandwidth connections that fixed and mobile operators are making available as well as the growing ubiquity of WiFi connections.

First, the number of devices directly and indirectly connected to the network is accelerating. Many consumers have multiple connections to the Internet, not only a fixed and a mobile, but often a WiFi connection and maybe even multiple mobile devices (for the same person, not just for multiple family members!). The Internet of Everything (IoE) is connecting everything it possibly can to the Internet – from utility meters to vending machines to refrigerators to connected cars. Consumers are embracing wearable technologies like smart watches and fitness trackers that use mobile devices to connect back and update cloud services.
With the dramatic increase in the number of devices connected to the Internet, it follows that consumers are using more bandwidth. A study by Intel estimates that 1.6B TB of data per minute is consumed on the network – ranging from video services like Netflix and YouTube to messaging applications like Skype to social networking and browsing. Consumers often have multiple applications open at the same time – downloading files and email while chatting and streaming audio or video to their devices.

And these applications are increasingly encrypted – which is good for the consumer for privacy and security, but creates a unique challenge for broadband operators who are trying to understand usage trends on their network to deliver a better experience to subscribers. Most of the systems that are deployed in broadband networks today cannot keep pace with the rapid change in the application landscape, as encrypted traffic requires higher levels of sophistication to distinguish one application from another. Almost ½ of Procera's application signatures cover encrypted applications, and we release a signature update every week to keep our accuracy levels high in identifying traffic.

And all while the technology landscape is changing, broadband operators are also undergoing a significant business transformation to take advantage of virtualization – specifically Network Function Virtualization (NFV). NFV is promising to deliver a more cost effective, flexible, and agile network that will allow operators to deliver services faster than they ever could in the past. This transformation will take time to create the orchestration systems required to make the network truly software defined, but operators have already started shifting their buying patterns to include virtualization as a mandatory option from vendors during the buying process. And let's not forget the battle cry from both government regulators as well as consumers – "We want more bandwidth for less money".

Broadband operators need better data to empower business and network transformation. Data-driven organizations will have a competitive advantage in the market

VIRTUALIZATION OF BIG DATA COLLECTION

To better compete in this environment, broadband operators are trying implement Big Data projects that give them new insights into their network and subscribers. Operators have many disparate systems that send bits of information into the Telco's data mart, but the data is fragmented and not very granular. Operators often have probes deployed in their networks, but most focus on signaling information and they lack KPIs on the actual broadband experience delivered to the subscriber – resulting in customer complaints of a slow network and the operator responding that their network is up, so it can't be their problem. These systems are not integrated into the OSS/BSS and do not perform subscriber correlation in real-time, losing opportunities to proactively deliver targeted offers and rectify problems before a subscriber realizes that there is an issue on the network. Operators are also looking to monetize their subscriber usage data externally, but they often lack the granularity and correlation of that data to maximize the value to external audiences like enterprises, retail, or advertising companies.

So what should a broadband operator be looking for in a Big Data intelligence solution? First, in today's environment, the solution needs to be virtualized. Virtualized solutions offer significant cost and operational benefits for operators, not only by leveraging COTS hardware, but also by creating generic deployments that the software hosted on the server can be adjusted or changed without any physical updates to the network. This operational flexibility can translate directly to cost savings for the operator because they can limit the licenses that they buy to exactly what they need for the network. "Exactly what they need" is defined significantly different in a virtualized deployment than a physical deployment. Imagine the ability to spin up a probe instance on your network when you need it (maybe this is due to an upcoming online event, unusual traffic spikes, or a complaint from a customer) and once you are finished with your investigation, re-use that license somewhere else. Another option would be to pay only for the amount of time that the probe is active or exactly how much traffic was processed – rather than paying for the peak usage bandwidth or subscriber count. As orchestration systems mature in the NFV market, this vision of a dynamic network will be achievable.

Once you can get data from any point in the network thanks to virtualization, it is time to focus on what data that you need to assess the subscriber experience. As mentioned previously, relying on signaling information in today's data-centric will not deliver useful root cause analysis intelligence on why a subscriber is having difficulty access Facebook or a specific YouTube Video. The centerpiece of the data collection must be the subscriber – not an IP address. But simply knowing how much traffic that the subscriber is sending or receiving is also not sufficient to determine the subscriber experience. The subscriber experience is a combination of factors – the instantaneous performance of the access network, the device being used, the application(s) and content being consumed, the location of the subscriber, their service plan, or even the peering connection that the content is traversing. This correlation of information provides a fairly comprehensive view of a subscriber's data experience by focusing on the most likely points of failure or degradation.

If the above attributes are tracked for each subscriber active on the network, what data points are critical? Both the US and EU call out three metrics in their network neutrality proposals as critical for measuring the subscriber experience: Throughput, latency, and packet loss. When these metrics are applied to the expectations of the common applications that subscribers use on broadband – web browsing, video streaming, social networking, gaming, voice, and file upload/download – it is possible to get a very good view of the subscriber experience. However, the measurements need to be done for all subscriber traffic, and taken at sub-second intervals to gain real visibility. This can be done anywhere in the network – from the access to the core to the peering point – to enable the network operator to gather the subscriber experience intelligence where it makes the most sense for their deployment.

Real-Time access to structured data can transform an operator's ability to engage subscribers with value-added offerings and enhanced QoE

But in the era of Big Data, there is a lot more valuable data that can be gleaned from the network that can be important to the broadband operator. Understanding the applications and content that is being consumed by subscribers and breaking that down by location, devices, service plans, and the other subscriber attributes are important when doing capacity and service planning. Identifying the source of network issues and determining why a node is congested (due to increasing video usage rather than file sharing) can help determine what action should be taken to alleviate the congestion. Monitoring peering links for quality and application content ensures that you are getting the best ROI when you upgrade the link performance. There is also a lot of valuable subscriber data (aggregated and anonymized) that broadband operators can monetize with enterprises, advertisers, and industry analysts to increase their profitability. The more contextual subscriber information that the operator has, the more opportunity to discover nuggets of gold in their Big Data.

EVOLUTION OF BIG DATA USAGE

So the operator now has a nice collection of subscriber intelligence – both experience as well as operational and business data. How should that data be made available to the different departments inside the operator? There two facets to the use of data inside an operator – let's think of them as Small Data and Big Data.

Small Data is intelligence that can be made available to the operator in real time. Real time data needs to be very specific and actionable to be useful. One example might be if a subscriber initiates a video stream, and the operator wants to offer a turbo boost for the duration of the video to the subscriber so they can watch the video in HD. It could also be that the subscriber has had a bad experience (which the network detected in real time) and the operator sends them a rebate or a credit to keep the customer from churning to another operator.

Big Data is pretty much anything else. Any data that can be collected and stored, and then mined using analytics tools. The better the data that feeds the analytics system, the better conclusions and recommendations that can be made by the analytics systems. Big Data can be used for internal network planning, service planning, budgeting and business planning, or even external monetization opportunities.

So a good system can provide both Small Data and Big Data. The combination is powerful when the capabilities can be delivered by a single system, especially if it is subscriber focused.

INTRODUCING EVOLUTION

eVolution is an all virtual data probe that empowers fixed, cable, mobile, and WiFi operators to run their networks more efficiently and profitably. eVolution delivers a unique approach to providing structured data that can be used in real-time to enhance the subscriber experience for broadband subscribers.

Engineered from NFV-ready PacketLogic Technology

eVolution is based on industry-proven PacketLogic technology that has been engineered to run as individual Virtual Network Function Components (VNFC) as part of an ESTI-defined Virtual Network Function (VNF) environment. eVolution runs on Commercial Off the Shelf (COTS) hardware leveraging Intel processors, and can run on VMware or KVM hypervisors. The eVolution VNF will support NFV orchestration standards as they mature, and includes a VNF Manager and License Manager that support creating new instances of eVolution than can be deployed when and where the operator needs data collection.

eVolution utilizes the Datastream Recognition Definition Language (DRDL) signature database to identify applications, with weekly updates to keep unknown traffic at a minimum.

eVolution is empowered by Procera's industry-leading PacketLogic technology, including the DRDL signature database that is updated weekly

DRDL supports over 2500 signatures (Nov 2015) with over half of those signatures covering encrypted applications, a key requirement from network operators as encryption becomes pervasive on networks, limiting the effectiveness of current data probes. Procera's Perspectives can also be layered onto the eVolution VNF - including Subscriber and Score - to add structure to the data.

The data in eVolution can be accessed through multiple mechanisms depending on the use case for the deployment. LiveViewTM can be used to view the network traffic in real-time, with drill down from all network traffic to a single subscriber flow. LiveViewTM APIs can be accessed by OSS/BSS systems for integration into engineering and customer care systems. Industry standard IPFix can stream flow records to real-time or historical mediation systems decorated with valuable Perspective structure. The data can also be stored in the Insights Storage system where it can be accessed via ODBC for integration into Hadoop or other large scale Big Data solutions. Any or all of these options are available to network operators to leverage the data provided by eVolution.

The use cases for the eVolution data in real-time include Customer Care Alerts, Network Troubleshooting and Forensics, feeding Threat Management Systems, Congestion Management, or Real-Time subscriber offer solutions. Historical use cases include Capacity and Service Planning, Regulatory Compliance with ScoreCard's experience scoring, and targeted data wholesaling of aggregated subscriber data.

The eVolution data collection probe is deployed on COTS hardware in a fully virtualized environment supporting common virtualization and NFV technologies. The collection can take place anywhere in the network, at the edge of the network, in the core, or at the peering point, depending on the goal of the analytics. The performance of the individual systems can range from a few megabits to hundreds of gigabits, meeting the needs of the smallest and the largest networks in world. The solution is managed with Procera's VNF Manager, which can dynamically instantiate instances wherever and whenever needed from a pool of licenses that delivers just-in-time bandwidth and optimizes CAPEX and OPEX spending.

Procera's Perspectives technology differentiates our solution from traditional probe-based solutions by having a providing structured data on the overall subscriber experience that includes traffic (i.e. applications), subscriber, device, topology, routing, score (quality of experience), RAN, and content - all placed in context with each subscriber data flow on the network. Each Perspective is a software license that can be unlocked as needed to provide more visibility into the subscriber experience.

With this structured data, Big Data systems can more easily correlate the subscriber experience for integration into Customer Care, Engineering, Marketing, and Executive-level dashboards and analytics offerings. The use of IPFix to stream intelligence in real-time ensures that the data provided to Big Data systems can be acted upon in real-time, solving problems before subscribers know that they exist. Data can also be stored in the Insights Storage database and access via ODBC.

eVolution can be integrated with existing OSS/BSS systems for real-time use cases and Big Data systems for historical use cases

DRDL supports over 2500 signatures (Nov 2015) with over half of those signatures covering encrypted applications, a key requirement from network operators as encryption becomes pervasive on networks, limiting the effectiveness of current data probes. Procera's Perspectives can also be layered onto the eVolution VNF - including Subscriber and Score - to add structure to the data.

The data in eVolution can be accessed through multiple mechanisms depending on the use case for the deployment. LiveViewTM can be used to view the network traffic in real-time, with drill down from all network traffic to a single subscriber flow. LiveViewTM APIs can be accessed by OSS/BSS systems for integration into engineering and customer care systems. Industry standard IPFix can stream flow records to real-time or historical mediation systems decorated with valuable Perspective structure. The data can also be stored in the Insights Storage system where it can be accessed via ODBC for integration into Hadoop or other large scale Big Data solutions. Any or all of these options are available to network operators to leverage the data provided by eVolution.

The use cases for the eVolution data in real-time include Customer Care Alerts, Network Troubleshooting and Forensics, feeding Threat Management Systems, Congestion Management, or Real-Time subscriber offer solutions. Historical use cases include Capacity and Service Planning, Regulatory Compliance with ScoreCard's experience scoring, and targeted data wholesaling of aggregated subscriber data.

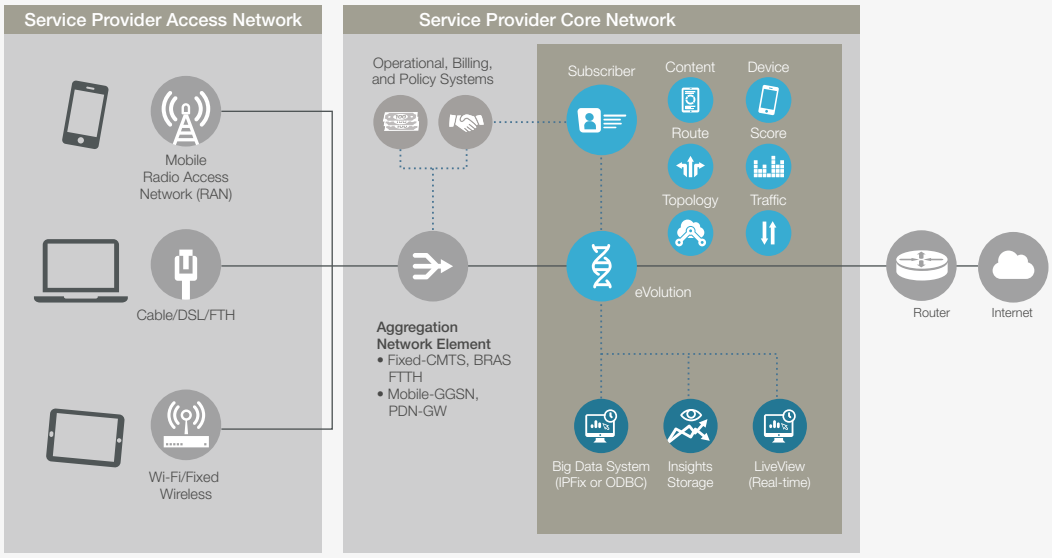
The eVolution data collection probe is deployed on COTS hardware in a fully virtualized environment supporting common virtualization and NFV technologies. The collection can take place anywhere in the network, at the edge of the network, in the core, or at the peering point, depending on the goal of the analytics. The performance of the individual systems can range from a few megabits to hundreds of gigabits, meeting the needs of the smallest and the largest networks in world. The solution is managed with Procera's VNF Manager, which can dynamically instantiate instances wherever and whenever needed from a pool of licenses that delivers just-in-time bandwidth and optimizes CAPEX and OPEX spending.

Procera's Perspectives technology differentiates our solution from traditional probe-based solutions by having a providing structured data on the overall subscriber experience that includes traffic (i.e. applications), subscriber, device, topology, routing, score (quality of experience), RAN, and content - all placed in context with each subscriber data flow on the network. Each Perspective is a software license that can be unlocked as needed to provide more visibility into the subscriber experience.

With this structured data, Big Data systems can more easily correlate the subscriber experience for integration into Customer Care, Engineering, Marketing, and Executive-level dashboards and analytics offerings. The use of IPFix to stream intelligence in real-time ensures that the data provided to Big Data systems can be acted upon in real-time, solving problems before subscribers know that they exist. Data can also be stored in the Insights Storage database and access via ODBC.

Figure 1

FLEXIBLE EVOLUTION DEPLOYMENT



ABOUT PROCERA NETWORKS

Procera Networks, the global Subscriber Experience company, is revolutionizing the way operators and vendors monitor, manage and monetize their network traffic. Elevate your business value and improve customer experience with Procera's sophisticated intelligence solutions.
For more information, visit proceranetworks.com or follow Procera on Twitter at @ProceraNetworks.

PROCERA

CORPORATE OFFICES
Procera Networks, Inc.
47448 Fremont Blvd Fremont,
CA 94538
P. +1 510.230.2777
F. +1 510.656.1355

CORPORATE OFFICES
Procera Networks
Birger Svenssons
Väg 28D 432 40 Varberg, Sweden
P. +46 (0)340.48 38 00
F. +46 (0)340.48 38 28

ASIA/PACIFIC HEADQUARTERS
Unit B-02-11, Gateway Corporate Suite,
Gateway Kiaramas
No. 1, Jalan Desa Kiara,
Mont Kiara 50480 Kuala Lumpur,
Malaysia