



## White Paper

# Performance Test Strategies for Next-Generation Networks



Prepared by

Jim Hodges  
Senior Analyst, *Heavy Reading*  
[www.heavyreading.com](http://www.heavyreading.com)

on behalf of

 **Mu** Dynamics®

[www.mudynamics.com](http://www.mudynamics.com)

September 2011

## The Impact of Innovation

There is little (if any) debate within the telecom industry that service innovation is the most essential ingredient of a next-gen all-IP strategy. However, the model of what constitutes service innovation is drastically, fundamentally changing, due to the ongoing migration of subscribers to mobile broadband services leveraging IP end devices such as smartphones, which enable a much more uniform social networking experience than was possible with feature phones and 2G networks.

As a result, many of the guiding principles network operators have relied on in the past are less and less relevant for meeting these new demands. This is especially true of the tools and techniques network operators leverage to test services prior to introduction. Accordingly, this white paper examines and quantifies the challenges specifically associated with testing and introducing new IP-based services.

These challenges, while tightly intertwined, are broken down into three distinct categories for the purposes of evaluation: 1) quantifying the network and business impact of running a diverse set of applications; 2) defining the network impact of IP devices such as smartphones; and 3) the requirement to define best practices for testing these networks using real applications.

Before discussing each of these topics in detail, it's worth defining at a macro level conceptually how next-generation networks (NGNs) and tools differ from legacy networks and tool sets. As shown in **Figure 1**, the differences are considerable.

**Figure 1: Legacy vs. Next-Gen Networks**

CHARACTERISTIC	LEGACY	NEXT-GEN NETWORKS
Service innovation	Driven by telco	Driven by telco and third parties
Device intelligence	Limited and network specific	High and network and Web specific
Application development	Limited reliance on third parties	Heavy reliance on third parties
Application time to market	Medium to long timeline	Extremely short timeline
Application impact	Straightforward to model	Extremely complex to model
Application test cases	<b>Static</b> – application and traffic pattern well defined. Packet inspection not required. Test cases used internally by network operator.	<b>Dynamic</b> – application and traffic patterns volatile. Inspecting packets is required. Test cases may be shared with third parties.
Test tool capability	<b>Limited</b> – focus on testing impact of networks through <i>simulation</i> of data patterns.	<b>Extensive</b> – focus on testing impact of applications and devices based on <i>re-creation</i> of actual usage patterns.
Monetizing service innovation	Relatively straightforward; few if any free applications	Extremely complex given mix of pay by app and free applications

Source: Heavy Reading

In consideration of the above, a key thrust of this white paper is that NGN operators must holistically evolve and adopt a suite of best practices and test tools which possess the flexibility and scalability to exploit device intelligence, dynamic application development and business monetization models.

## Real-Time Growth of Applications

This section of the white paper documents the impact and challenges associated with supporting a diverse set of applications and smartphones.

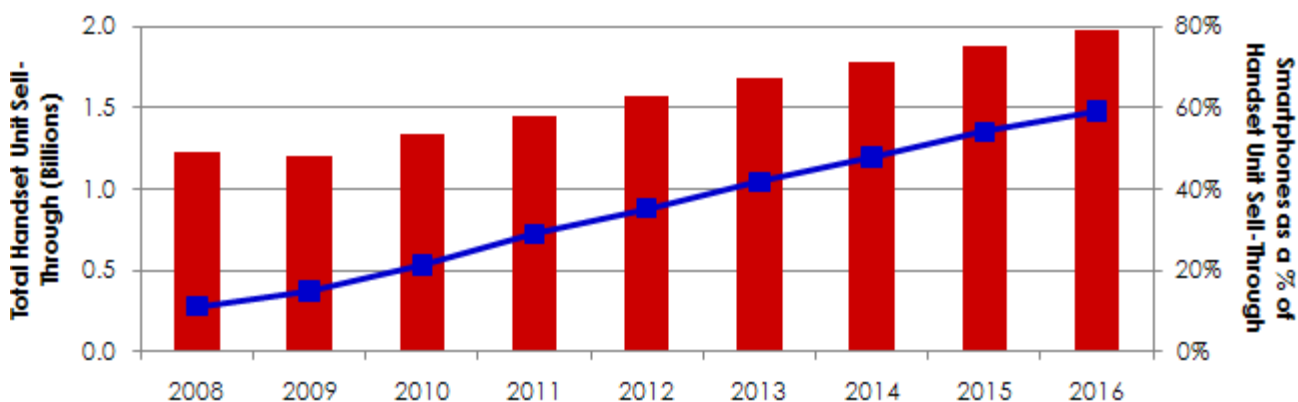
### Applications Get a Facelift

Although there was no precise date when "social network"-based services were introduced, Facebook garnered considerable market scrutiny as early as 2004, when it reached nearly 1 million subscribers. As of 2011, Facebook has grown to a staggering 750 million users. And in many respects, the explosive growth of this application mirrors general growth trends of other real-time applications including YouTube which provided the platform for sharing of personalized video content.

While applications such as Facebook and the pent-up demand they addressed would have been successful on some level, the availability of 3G mobile broadband services from 2005 and the ramp up of smartphones from 2007 were critical factors in redefining the scope of applications and ensuring mass adoption. One key landmark in this evolution was the debut of the iPhone in 2007, which enabled subscribers to download Apple or third-party applications, and paved the way for the integration of real-time applications. Apple's App Store now houses more than 350,000 applications, while Android supports some 245,000 – up from 75,000 less than a year ago and growing by about 27,000 applications per month. Including an estimated 600,000 Facebook applications, there are more than 1 million active applications now available to subscribers across these three platforms alone.

Given the direct link between smartphones and applications, it's not surprising that smartphone penetration is also showing strong growth. While the conversion from feature phones to smartphones is a more gradual process due to contractual obligations and handset subsidization, as price points continue to decline, this will become less and less of an issue. As shown in **Figure 2**, we estimate that by 2016, on a base of nearly 2 billion handset sales, 60 percent (1.2 billion) will be smartphones. Further, over this period it's anticipated that the video capabilities of these devices will continue to improve to deliver enhanced services such as 3D video. One recent example is the Skype video call application for Android devices.

Figure 2: Handset Sales & Smartphones as a Percentage



Source: Pyramid Research

Naturally, as smartphone sales continue to expand in all markets worldwide, so will the number of mobile users who will utilize the video-centric applications noted above. To put this growth in a regional context, as illustrated in **Figure 3**, the number of mobile video subscribers – which includes video downloads, video calling and streaming – will more than double from 261 million to 597 million users over the next five years.

**Figure 3: Mobile Video Subscribers**

REGION	2011	2015	GROWTH IN SUBSCRIBERS
North America	40.4 million	125.7 million	85.3 million
Eastern Europe	9.3 million	22.8 million	13.5 million
Western Europe	49.4 million	88.9 million	39.5 million
Asia Pacific	120.9 million	259.2 million	138.3 million
Latin America	21.8 million	51.1 million	29.3 million
Middle East Africa	19.0 million	48.8 million	29.8 million
Total	260.8 million	596.5 million	335.7 million

Source: Heavy Reading / Pyramid Research

## Application & Network Impacts

Although the growth of real-time video applications unquestionably represents a tremendous business opportunity for network operators, it also presents serious network and application related challenges as a large portion of these applications incorporate a video streaming or sharing component. The most formidable of these are ensuring network scalability and connectivity.

## Network Scalability & Connectivity

Application-wise, perhaps the most serious challenge that network operators confront for video-centric and social networking applications is ensuring that networks can provide subscribers with the services to which they are entitled. Unlike the days when legacy and voice-centric services were the norm, today's networks are much more dynamic and therefore, highly vulnerable to the explosive growth driven by volatile social-driven Internet trends.

In order to fully understand these challenges, it's worth first examining the diverse nature of these applications and device interactions in terms of signaling and session impacts.

## Signaling & Session Impact

In early 2009, as part of a primary research project, *Heavy Reading* interviewed a number of major mobile operators based in North America, and Western and Eastern Europe. One of the topics discussed was the impact of smartphones on network performance.

At this time, none of the network operators believed that smartphones would cause additional congestion on their network. However, shortly following these discussions, a number of operators did in fact start to experience serious signaling and session congestion issues.

This was due to several issues. First, in fairness to operators, little at this time was known of smartphone resource consumption patterns and over the years it has emerged that these devices and tablets have considerably different data usage patterns from the well understood dongle-equipped laptops which still consume the largest portion of mobile data services today.

For instance, Akamai, which benchmarks Web usage patterns, has identified that smartphone users typically utilize much shorter and more numerous sessions than laptops in part given they run a different mix of applications, and many use smartphones for tasks including social network updates which require much shorter sessions. (Source: State of the Internet Q1 2011)

As a result, networks must not only contend with more sessions running concurrently on their network, but also the signaling to set up and tear down these sessions. Compounding the impact, some network operators also learned the hard way that in some cases smartphones may support software features to extend battery life that terminate sessions when data was received even though the session may be still active. This connection-centric approach was a significant factor in overloading network signaling capacity even if networks had been correctly provisioned and engineered to handle the pure data component.

While the adoption of 3GPP features such as fast dormancy will help alleviate signaling issues going forward by providing parameters and procedures defining a switch between active and idle model, network operators will need to remain vigilant as new applications come to market. Moreover, it's also important to recognize that while subscriber usage patterns differ, the profiles of the applications themselves may differ significantly. In this regard, recent history has shown that applications are quite diverse from a session and signaling perspective and may even perform differently on different devices.

Generally speaking, while video applications such as Netflix can have the most impact from a pure data perspective; there may also be significant signaling and session impacts. For example, while Netflix may appear a straightforward streaming service, by design the application has multiple sessions running in parallel which can result in a heavier signaling and bearer load, especially when HD video capabilities are utilized. As an order of magnitude, deep packet inspection (DPI) vendor Sandvine estimates that Netflix traffic represents up to 20 percent of total web traffic, so using a multi session approach can have serious network impact.

In response to this, we have seen over a number of years an increased interest by network operators to implement policy control and DPI to ensure that they manage network resources and applications in the most efficient manner. Essentially by using DPI to identify session and signaling traffic, network operators can invoke policy control and throttling for traffic management and advanced network security. These capabilities are further discussed in **Section IV**.

It's also imperative to note that network performance may be impacted by the devices on which the applications run. One recent example is the excessive signaling performance publicly noted for the Angry Birds game when running on an Android device that utilized an integrated mobile advertising application.

Another example is a large Asian carrier that recently suffered a network outage where a third-party application took the voice-call success rate down to a mere 10 percent, because the signaling traffic generated by the application overloaded its network.

However, this phenomenon is not simply germane to a single game or device use case, but rather a concrete example of performance variability that can result when third-party applications run on different mobile devices across a carrier network. In this regard, we believe it is best to consider all combinations of applications and devices to be potentially vulnerable given the right circumstances, rather than assuming that these are single isolated instances caused by software deficiencies.

Finally, and in an attempt to make applications more network-friendly and to help ensure the best user experience, the industry is starting to see collaboration between application developers, device manufacturers, and operators.

# Network Testing With Apps: The Demise of the "Static Quo"

As we have seen, given the number of variables, estimating, and testing the network performance impact of new and existing applications can be very complex. Therefore, in this section of the white paper we examine the principles for defining best practices for testing networks using these applications.

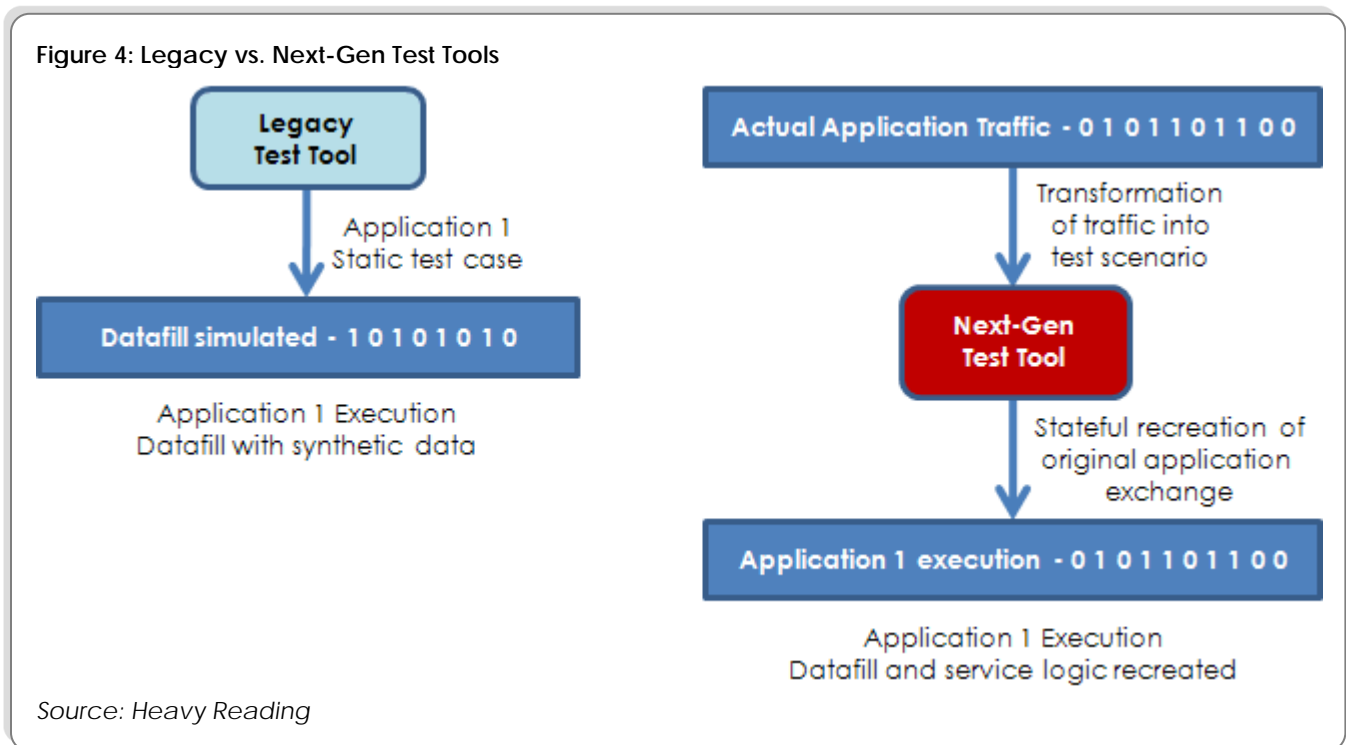
## Legacy vs. Next-Generation Test Tools

Legacy test tools are at a crossroads. While they have evolved and have met operator requirements, utilizing these tools for next-generation applications is problematic. Accordingly, the first step in defining best practices is to isolate the critical capabilities that next-generation tools must support that their legacy counterparts do not. These can be categorized into two areas:

1. Real-time modeling capabilities of today's (and tomorrow's) applications
2. Enhanced library store and third-party testing support.

## Real-Time Modeling of Applications

Modeling the applications is a fundamental component of any suite of test tools. However, as shown in **Figure 4**, one crucial distinction is that legacy tools *artificially simulate* traffic using *static* data, rather than recreating it from actual *dynamic* traffic patterns.



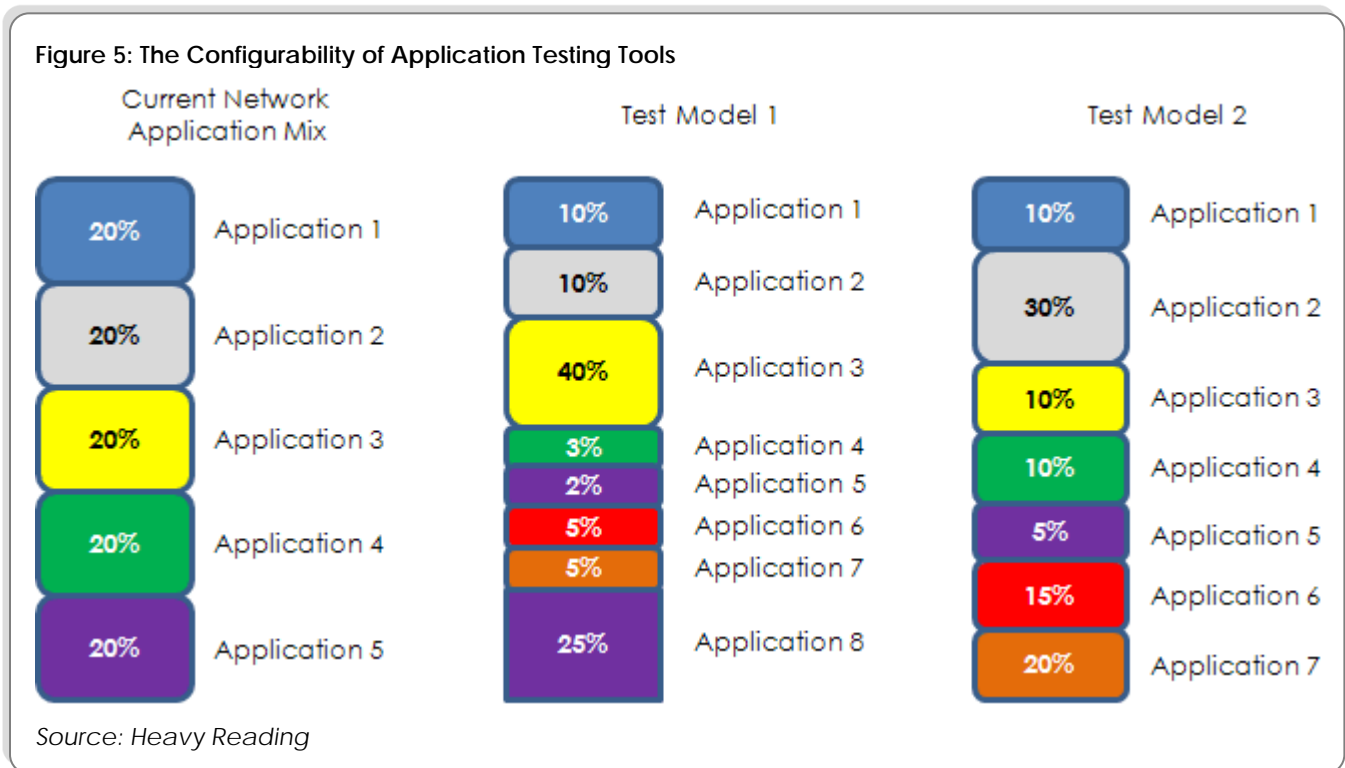
This legacy approach is due to the fact that these tools were primarily designed to load test Layer 2 to Layer 4 network and transport behavior. While this approach was adequate when the signaling and session impact of data applications was relatively low, simply replaying this traffic however will fail to recreate the applications as well as the user behavior and device properties.

By leveraging the actual network traffic captured from a real device running a given application with user operations, a next-generation tool is able to transform the original exchange into a test scenario. This test can then be replayed to faithfully reconstruct the original exchange and truly mimic the application, the device, and the user behavior. By then running these tests concurrently, a high scale environment can be reproduced ensuring the network perceives production-like users and applications.

Without the packet capture of a real traffic profile, measuring the effects of "fuzz test" techniques which methodically inject corrupt data into the traffic fields to assess protocol robustness are much less relevant. This is more and more relevant, given that smartphone malware attacks are on the rise.

A final consideration is application mix. As we have seen, the mix of applications and length of sessions can differ significantly depending on device type and user demographics.

Consequently, as illustrated in **Figure 5**, modern application testing tools must be configurable so that operators can quickly change parameters to support dynamic modeling. These capabilities, which legacy tools lack, allow operators to test the impact of a change in application mix and enable the creation of security threat scenarios.



## Enhanced Test Content Store

In order to support enhanced modeling capabilities, software testing tools also need an expanded set of capabilities for creating and storing an extensive number of profile test cases in a centralized library store. This is essential for ensuring that when new applications or new versions of existing applications are released, the operator can quickly model these in the test environment to determine the network impact and performance. The results can then be used to devise new network policy strategies.

This requirement is largely due to the fact that these applications are consumer- and subscriber-driven instead of telco-driven. This is a mixed blessing, in that while using archived test cases can shorten time to introduce new services and hence drive revenue growth, it also means that an operator may not get a chance to fully test the impact of the application until it is live in the network, which is not ideal from a security and scalability perspective.

Moreover, this approach also enables network operators to share this test profile data with trusted third parties. This is valuable for ensuring these third parties can correctly model service behaviors when developing new applications – thereby shortening application introduction cycles while minimizing the potential for catastrophic adverse effects. Given the volatile nature of application development, we believe these types of shared "social testing" library store initiatives will continue to expand both in scope and relevance going forward.

In response, vendors such as next-gen test tool developer Mu Dynamics with its Mu TestCloud store has developed solutions which support these capabilities as online libraries capable of supporting ready-to-run tests for thousands of applications.

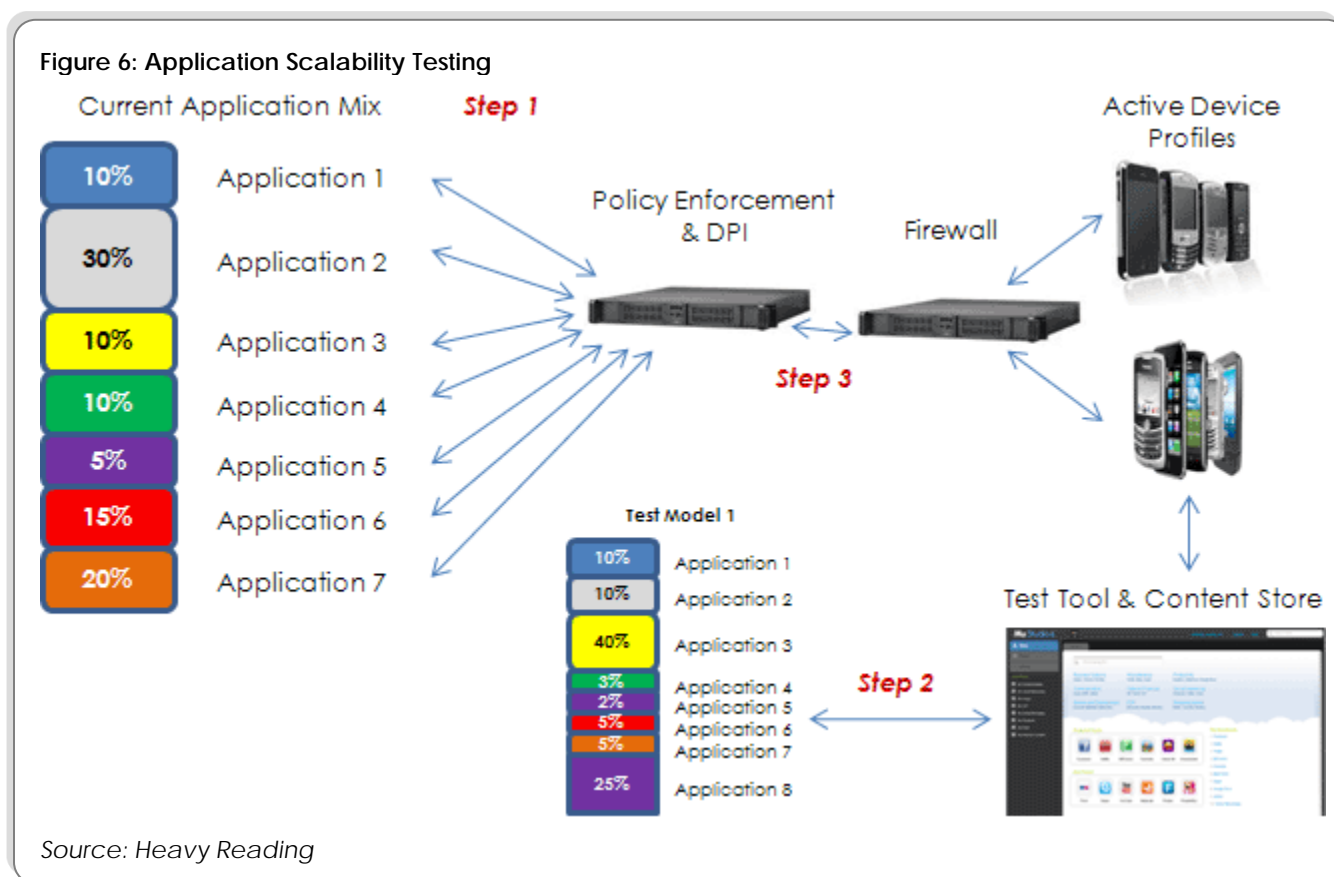
## Use Case Scenarios

In this section of the white paper we illustrate how the testing concepts presented in previous sections can be implemented. To accomplish this, we utilize three specific use cases:

- Application Scalability Testing
- Application Enablement With Policy Control & QoS
- Application & Device Security Testing

### Use Case 1: Application Scalability Testing

As we have seen, the mix of applications running on a network at any one time can vary significantly. Therefore, as illustrated in Figure 6 and detailed in the bullets below, exact replication of the application mix by next-generation test tools is crucial for ensuring that telcos fully comprehend the impact of a change in application as quickly as possible.



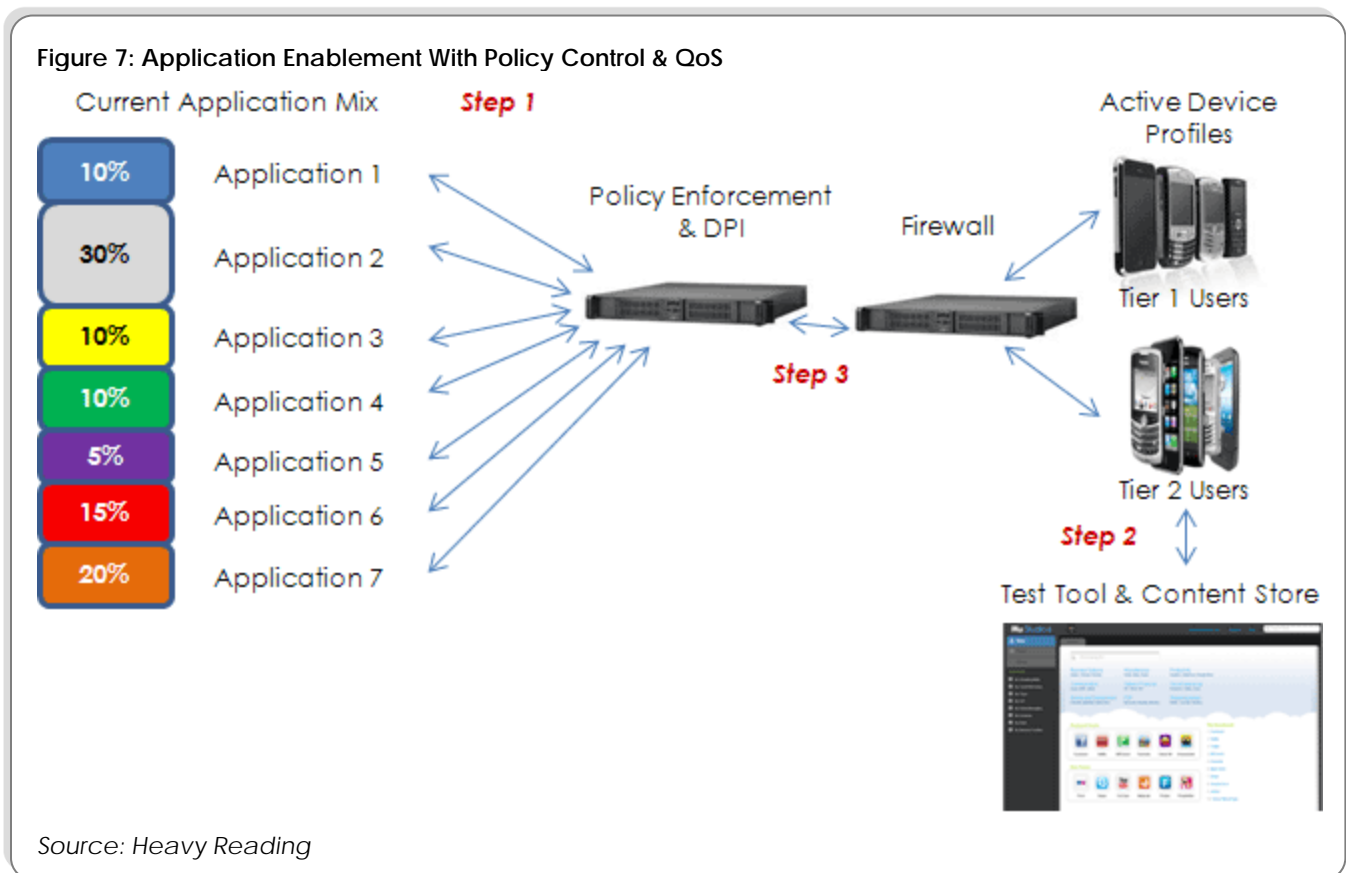
The sequence for Use Case 1 is as follows:

- **Step 1 – Identify Application Traffic Levels:** Network data and application characteristics from DPI and policy control appliances are fed into the test tool for modeling.

- **Step 2 – Test Tool Modeling:** The test tool replicates the exact mix and performance characteristics observed in Step 1. In this case the tool is used to model a change in the mix of applications (2 to 8), including signaling and session impacts. This step may also replicate device-specific impacts for each application based on known device profile attributes. The result of this is the identification of the combination of device and application traffic patterns that represent the greatest network performance threat.
- **Step 3 – Adjust Network Settings:** Based on the above, policy control, firewall and DPI settings may be adjusted to factor worst-case scenarios.

## Use Case 2: Application Enablement With Policy Control & QoS

In this second use case, next-generation test tool capabilities are utilized in conjunction with policy control and QoS to enforce network parameters to validate performance levels.



The sequence for Use Case 2 is as follows:

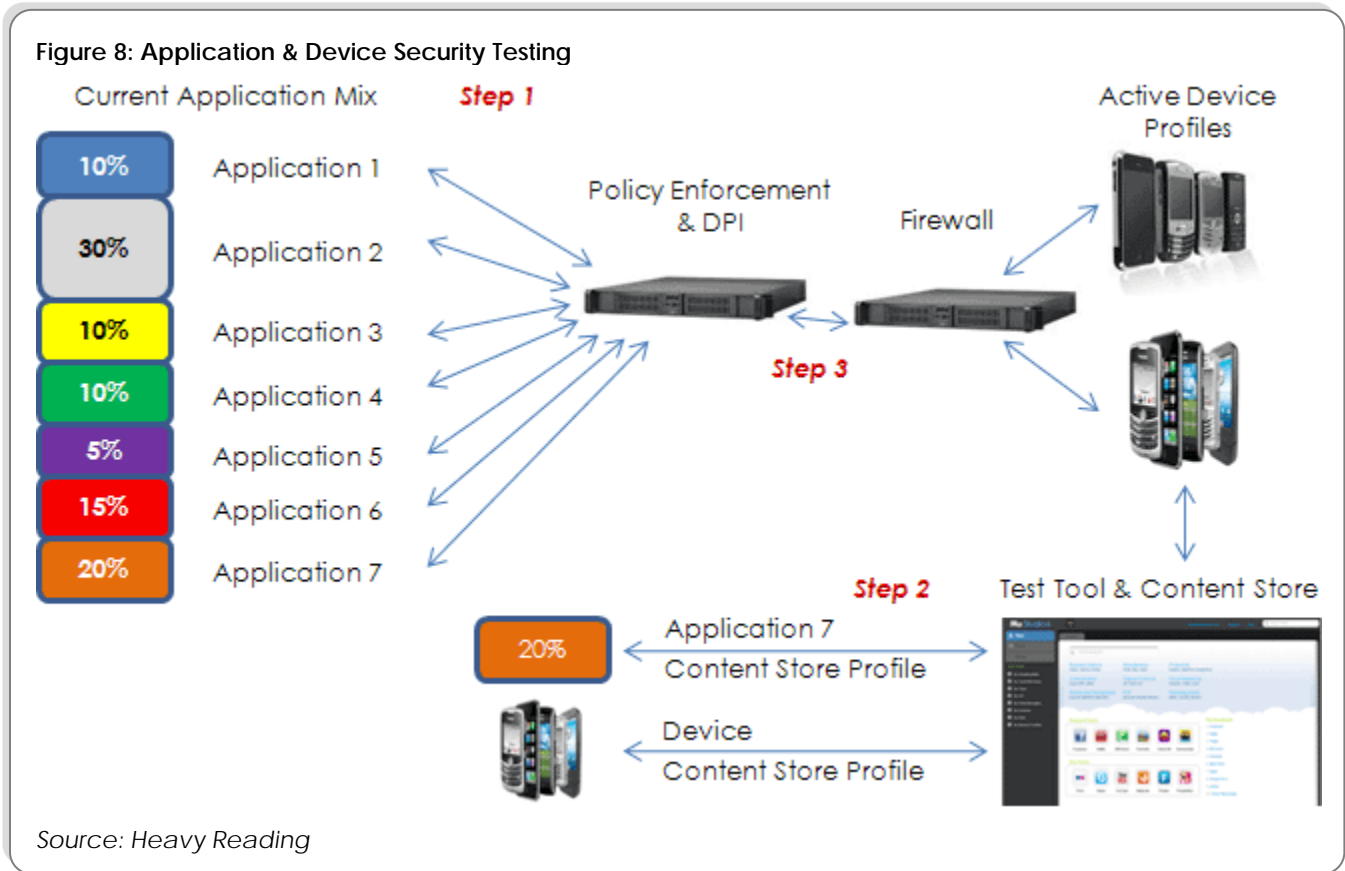
- **Step 1 – Identify Application Traffic Levels:** Network data and application characteristics from DPI and policy control appliances are fed into the test tool for replication.
- **Step 2 – Test Tool Application and Device Analysis:** The test tool replicates the traffic patterns and performance levels of application flows to desired

network performance settings. Key metrics are observed to validate that policies are effective. Policies related to billing and charging based on applications and user tiers (e.g., pre- vs. post-paid) or QoS are measured and modified as necessary to optimize the customer experience.

- **Step 3 – Invoke Policy Control:** In situations where congestion is noted or when an application mix based on the modeling from Use Case 1 is encountered, policy control may be invoked. A number of options are available. The first approach may be to throttle or block certain high-bandwidth applications for all users to maintain network-level QoS. The second may support general or specific application throttling for lower-tier users based on SLA benchmarks and device profiles.

### Use Case 3: Application & Device Security Testing

In this final use case, test tools are utilized to validate application and device profiles to ensure network security is not compromised.



The sequence for Use Case 3 is as follows:

- **Step 1 – Identify Application Traffic Levels:** Network data and application characteristics from DPI and policy control appliances are fed into the test tool for modeling.

- **Step 2 – Test Tool Application and Device Analysis:** The test tool is utilized to replicate valid application and device traffic as well as security attacks that occur on the production network. These attacks' impact on security and the performance of valid applications are measured. Security attack profiles from the content store are used to validate that the network is protected. Valid and invalid attack flows are sent through security systems to test policies such as intrusion protection or intrusion detection and white-listing or black-listing of applications.
- **Step 3 – Adjust Network Settings:** Based on the above, firewall and security policies may be adjusted to address identified vulnerabilities.

## Conclusion & Summary

Despite the explosive growth of applications over the past few years, there is little evidence to suggest that the current pace of innovation at the application layer will subside, especially as IP-enabled devices become the norm. As such, operators are now deploying more intelligent application-aware networks to control these applications, provide increased visibility, and improve quality and security.

In a world with such a dynamic mix of applications, we believe a more feature-rich suite of testing tools from vendors such as Mu Dynamics that can model these applications to determine their impact on the network will become a crucial component of the telco tool box.