Disruptive Analysis

*Don't Assume*

# Examining the role of DPI in device-aware policy & transcoding

A Disruptive Analysis *thought-leadership* paper

***Commissioned by Continuous Computing (now Radisys)***
***November 2010***

Author: Dean Bubley
Contact: information@disruptive-analysis.com

***Summary:*** *Mobile broadband data traffic is growing extremely quickly, especially with the growth of smartphones, tablets and PC modems. While capacity is being added, operators need to deal with real-world constraints and manage network resources more effectively.*

*This will increasingly mean that policies, charging and traffic management will need to be "device-aware". Different devices have different usage patterns and impacts on the network, in terms of outright volumes, signalling load and mobility. Even different versions of operating system or software application can be important. Creating and enforcing device-specific data plans and policies can protect the network – and potentially create better user loyalty and operator revenues. In the future, the network's intervention with devices will extend further still, with client software (such as dashboards connection managers) forming part of the overall holistic traffic management and policy ecosystem.*

# Background

This white paper covers the potential for device-specific policy management, pricing and content transcoding in mobile data networks. It has been written by the independent industry analyst and consulting firm Disruptive Analysis, and sponsored by Continuous Computing (now Radisys), as part of an initiative to promote thought-leadership, differentiation and innovative networking concepts for the mobile broadband and network policy-management marketplace. The opinions expressed are Disruptive Analysis' own, and are not specific endorsements of any vendor's or operator's products or strategy.

# Introduction: devices' impact on networks

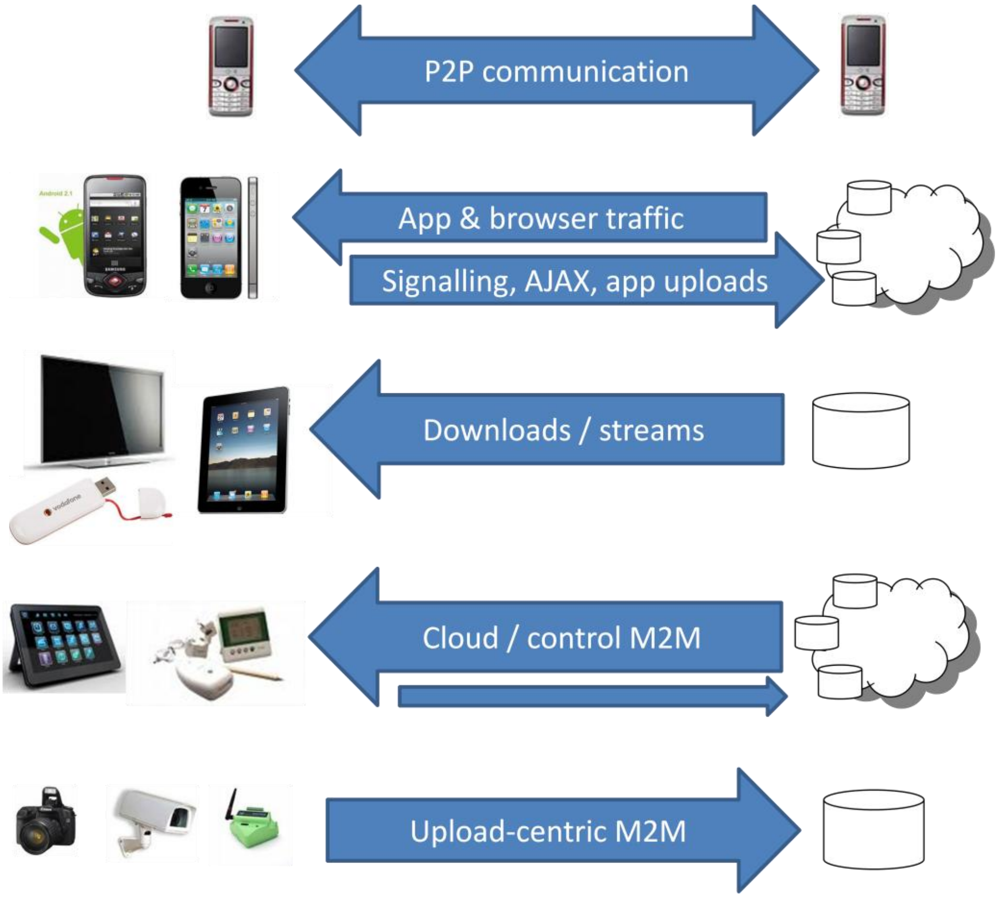
It is no secret that different mobile devices have vastly different impacts on operators' networks – in terms of traffic volumes, signalling load, usage patterns – and also realistic revenues and business models.

Laptops with integral 3G modems or USB dongles / datacards tend to have infrequent, but long, high-volume sustained sessions. Smartphones are usually used more regularly, but for shorter periods – often generating signalling events but only consuming limited amounts of data. New generations of M2M (Machine to Machine) devices will have widely varying usage dynamics – smart meters and health-monitoring devices may have very low volumes of data traffic but need absolute priority and guarantees, for safety reasons. Tablet-style devices like the Apple iPad may start to be used for bulk content downloads or video streaming, mostly in indoor locations. Connected CCTV cameras and sensors will be very upload-centric.

Not all of these devices will be Internet-connected, either – even though they may attach to a 3G/4G radio or fixed broadband line. Not all data over carriers' networks is to and from the Internet – and certain types will have very specific requirements. The recent Google / Verizon proposals on Net Neutrality point to the role of separate "managed services" running in parallel to "real Internet access" on a service provider's network.

Managing this diversity of usage models is going to be a major challenge for mobile operators in the future – as well as their fixed-line peers. Disruptive Analysis believes that network policy-management and enforcement will increasingly become device-specific – not least because it is easier to implement (technically and perhaps legally) than application-based alternatives. Often, operators control the sales channels for devices, and are thus able to specify particular terms and conditions for their use.

**Figure: Mobile data usage models varies strongly by device type**




**Source: Disruptive Analysis**

There will be numerous methods by which devices and policy management / DPI (Deep Packet Inspection) platforms interact – differential plan prices and performances, on-device tools for users to track data usage or be up-sold, varying approaches to offload, assorted algorithms for connection-management and so forth. While some of these approaches require operators and their vendors to deploy software onto the devices, or conduct remote-management and configuration themselves, other approaches are likely to be simpler and wholly network-based.

As the rest of this paper discusses, one of the main network-centric approaches to dealing with devices is around content optimisation and transcoding.

# Device-specific plans, business models & tariffs

It is already very common for operators to have device-specific tariffs and data plans, for example those specifically offered for iPads, MiFi-type personal hotspots, or BlackBerries. These often have their own unique terms, policies and prices.

In Disruptive Analysis' view, this is a good way for operators to manage overall traffic levels passively, but can be more powerful when combined with more deliberate and interventionist approaches to managing high-volume content.

In many ways, this approach to pricing is not new – operators have often applied different data plan structures and costs to 3G modems and handsets, for example. This has enabled operators to pitch 3G dongles against fixed-broadband competitors, reflecting typical higher monthly usage and so on. It has also reflected the lack of a voice plan sold concurrently, lack of (or lower) device subsidy and various other factors. There have also been dedicated BlackBerry plans from many operators, incorporating BES BlackBerry Enterprise Server) / BIS (BlackBerry Internet Service) email connectivity, and reflecting the specific deals cut with RIM. In some places it is now possible for travellers to rent iPhones or MiFi's by the day or week.

But increasingly, data plans are becoming more granular still – a trend likely to continue as we gain new device form factors. iPhone and iPad plans are specific to those products – and easily enforceable (for now) through the use of MicroSIMs which cannot be swapped around. MiFi and other personal hotspot products, which are inherently multi-device tethers, may also be subject to different plans and policies.

As an example, in the UK, the operator 3 has a well-segmented and device-specific set of tariffs: (*prices as at 10 August 2010*)

- Apple iPad MicroSIM only plans: 1GB @ £7.50 / month or 10GB @ £15 / month [1-month rolling contract]

- Laptop SIM-only plan: 5GB @ £15 / month

- Handset SIM-only Internet plan: 1GB @ £5 / month (which also includes circuit-based Skype calling)

With new devices using prepay and non-subscription models, we can expect to see even more granularity – a navigation device might come with a year's free traffic reports including data, while a camera might be bundled with 1000 wireless photo uploads. Add in plans which include or exclude WiFi or femto data, and it gets murkier still.

The interesting thing here is that, in essence, we are getting a sort of blurry form of policy management and mobile traffic management by the back door. Although the correlations are not perfect, typical iPhone usage is different than typical BlackBerry usage, or assorted other products. Less / more video, less / more social networking , less / more web browsing, more / fewer notifications and so on. It's quite easy to skew the prices and tiers to favour the less network-hungry products – or implicitly reward manufacturers for creating "non-aggressive" devices that don't hammer the RNCs (Radio Network Controllers) with signalling traffic so much.

What is less clear is whether prioritising **device types'** traffic is the same in terms of Net Neutrality as prioritising **application types**. Is it fair, reasonable or legal to distinguish between them? Even if they are not dynamically prioritised, it could be possible to rate-limit them – for example peak speeds of 1Mbit/s download vs. 3MBit/s. Under absolute purist views on Net Neutrality, it would probably also fall foul of the strict rule-making, but it is probably easier to justify discrimination between hardware than between websites or online services. It is certainly much easier to distinguish between device types than application types in the network.

But looking to the future, it seems likely that even this style of discrimination will be insufficient to control the volume of rich content, especially over non-offloaded macro 3G networks which are the most congested. It is probable that video traffic, in particular, will need more active management – especially as capabilities are extended to new devices, or even new OS / application refreshes of devices already in service.

The risk for operators here is that new applications can change device usage patterns, through appstore downloads and "viral" adoption. This could hugely increase video consumption or other rich media (for example, "augmented reality" content) on a particular device class, literally overnight. In those cases, the role of DPI might be to spot new and emerging trends rapidly, so that network policy changes, device plan price updates, or direct collaboration with software partners can be pursued without delay.

# The role of content transcoding

Within a given category of device, there is tremendous variation in both "network-friendliness" and typical use cases. RIM's BlackBerry range tends to be used for less video streaming than Apple's iPhone – and has a particular emphasis on data compression. That may change with the new Torch model and OS version 6. Until recently, iPhones did not support multi-tasking, while Android and Symbian devices enabled various applications to perform background tasks that consumed network resources.

It is becoming clear that video downloads – either streamed or downloaded, are fast becoming the source of much of the mobile data traffic occupying 3G (and, in the future, LTE) networks. This is true across device form factors – consumers watching full-screen Hulu and iPlayer on notebooks, YouTube and other clips on smartphones, as well as a likely broad range of new services on dedicated consumer electronics and tablet products.

Future evolution to higher definitions and even 3D seems certain. Left unmanaged, it is likely that networks will suffer congestion in some instances, users will receive sub-par viewing experiences, while content publishers and their advertisers will not obtain the reach and quality they expect.

Although it is possible for operators to perform across-the-board compression or optimisation of video, this is a very blunt instrument which could have negative consequences. The legal implications of transcoding data without the agreement of the user and/or content provider is highly debatable. Dealing with video on a device-specific basis allows any such intervention to be tied back to the relevant service plan – and also perhaps to "upstream" providers of content or applications.

There are various possible usage scenarios for policy in relation to device-specific content optimisation and transcoding:

- Automatically reducing downstream video traffic to a pixel-resolution corresponding to the device's screen size, or potentially reducing frame-rate or colour depth if this is over-specified for the device's capabilities, where all parties agree.

- At times or locations of network congestion or overload, it may be permissible to compress or transcode video to a lower bitrate without explicit agreement from publisher or user – although it would probably be appropriate to notify both parties, and perhaps zero-rate the traffic, or use some other means of compensation. It should be noted that this is likely to be highly controversial, if conducted without sensitivity and cooperation.

- Enabling operator-branded video applications and clients on mobile devices to display video or other content from other sources, for example when embedded in mash-ups. This could also work with the operator's in-house browser/widget framework – or deliver video from the operator's own or preferred CDN (Content Delivery Network).

- Encoding a unique identifier or watermark into the transmitted video, indicating the device to which it was transmitted. This could be offered as a managed-DRM (Digital Rights Management) service to the video publisher – for example embedding a key corresponding to the device's IMEI number in the video stream, to enable any future pirated versions to be traced back to the original receiving device.

- Potential for working collaboratively with the end-user, for example through an on-device client or notification engine to ask, "*Reformat video for better speed and lower data consumption?*"

- Management of uplink traffic – for example, scheduling full-resolution image or video uploads for quiet periods, sending lower-quality thumbnails immediately.

When one considers the range of visual content that extends beyond video, it is likely that many additional use cases will emerge over time. Augmented reality applications, for example, may create significant traffic loads, as might "cloud gaming" services which transfer screen images over the air. Indeed, the whole area of cloud computing and cloud services is likely to drive the need for mobile network transcoding or data-awareness to optimise performance.

# Potential use cases for DPI & policy

While it is possible for operators to enforce device-specific policies simply by routing them to different APNs (Access Point Names), this is itself complex to manage, given the huge proliferation of products. It also fits poorly with SIM-swapping, roamers or the use of generic data SIM cards with "vanilla" 3G-enabled devices bought through non-operator channels. The use of tethering and MiFi-type products also makes it difficult to deal with two-device combinations.

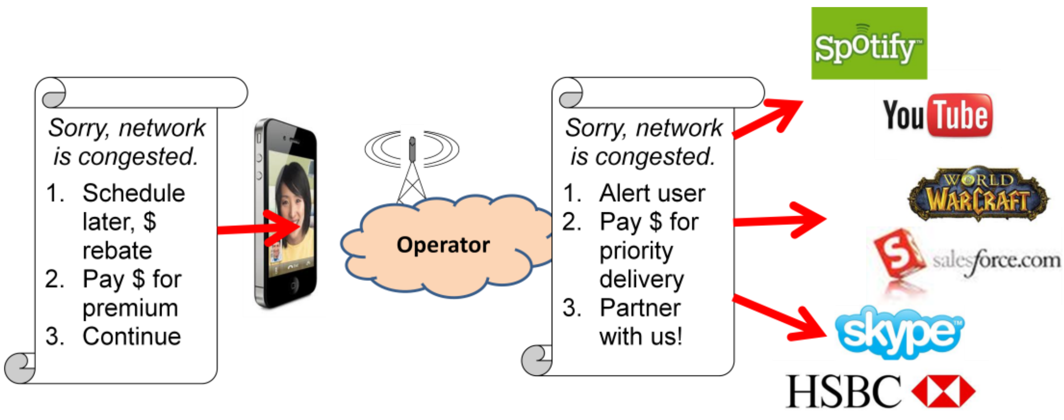
Using DPI resources in the network, it is possible to detect specific brands and models of device, as well as particular browsers or other client software in use. These can then be categorised and grouped in a variety of ways – for example by screen size, by manufacturer, by OS variant and so on. Although basic information can be derived from the IMEI, it is worth noting that different configurations for the same device can have a sizable impact on the network – for example, Apple's major OS updates, which can also change the firmware on the radio baseband chip.

The simplest use case for DPI in this context is in observation and reporting – tracking trends in usage between devices, or mapping device data consumption against specific tariffs and terms / conditions of the plan. This can inform marketing staff of the success of their pricing strategy, highlight particular application / content providers to target for partnerships, or assist operational staff in assessing the impact of new codecs or applications on the network.

The next stage of device-specific DPI is in enforcement and direct intervention in the traffic stream. It is possible, for example, that particular policies might be applied – at network level – for all Android devices, or all made by a specific OEM. There may be interesting business models here, for example working with a video publisher such as YouTube to *upgrade* certain video traffic in exchange for a share of incremental advertising revenues associated with the streams.

Ultimately, DPI could form part of the dynamic use case scenarios outlined above – perhaps detecting in real-time certain flows or application streams, alerting the user / publisher to congestion issues and offering alternatives.

**Figure: Towards a Congestion API for users and application providers**




**Source: Disruptive Analysis**

In the future, we are also likely to see differential treatment of PC-based, smartphone and "new device" mobile data flows. This may lead to more sophisticated, nuanced methods of managing high-volume traffic. For example, during periods of network congestion, it may be considered important to leave images or video for medical devices absolutely untouched, while treating more generic tablets or smartphones more aggressively.

# Pitfalls and concerns

All this said, there are still significant issues around network-based modification of Internet content, without the specific knowledge and agreement of the publisher and end-user, especially if it degrades quality. For example, if the user pays a 3rd-party content provider for HD (High Definition) video, but it is downgraded during transit, there are potential liability consequences.

There is also the possibility that some applications will use widescreen content, with the user able to zoom in on specific portions. It must be remembered that video is not "an application", merely a traffic flow – there are 10,000s of video-based applications, and it is essentially impossible for the network to always guess correctly what the *real* end use of a given flow might be.

Further, an increasing number of handsets are used as tethers, or perhaps with video-out ports, rather than for displaying video on their own screen. Some are even being equipped with projectors. This would mean that re-formatting video for the expected screen resolution could actually worsen QoE inadvertently. Nevertheless, even with these scenarios, there is value in the operator knowing that a mismatch exists between video quality and handset display resolution – perhaps driving suggestions for new service sales, or encouragement for using offload techniques rather than downloading in a busy cell.

It must also be recognised that generic devices are increasingly being used for specialist tasks – either standalone, or with accessories. A tablet might be used to display a children's cartoon – or a critical safety video for aircraft maintenance engineers. Heavy-handed approaches to content transformation must be avoided.

Lastly, it is unclear what risks may occur from a mismatch between network-resident content transcoding, and the application's own rate-adaptive behaviour. It is quite possible that the app itself will monitor perceived network quality, and adjust its own codec, frame-rate or other parameters to optimise performance for the user. There is a chance that this could work "in anti-phase" to whatever the network is trying to enforce itself.

# Conclusions

To sum up, different devices have different network impacts. In many ways, device type is a good proxy for application usage profile, especially where the service provider has dictated specific prices or terms & conditions. But ultimately, price / plan approaches to segmentation and mobile broadband management are quite blunt and inflexible tools – especially where usage of a particular device class changes quickly.

There are therefore definite use cases for network elements both monitoring and transcoding data and content on a device-specific basis. However, it is critically important that this is approached with a full understanding of the context – and ideally the knowledge and support of both end-user and the publisher of the video or application.

Used in parallel with pricing, policy management and monitoring via DPI makes sense, especially on congested networks populated by traffic from increasingly diverse device types.

Over time, Disruptive Analysis expects mobile policy management to involve device-side software as well as network-based elements. Alerts or notifications to the user of network congestion, options or recommendations for content format choice or transcoding, or rate-adaptive applications that work collaboratively with the operator's infrastructure.

Overall, device-specific functions in the DPI and policy domain will become ever more important – perhaps more than those geared towards "application", as specific hardware products are much easier to categorise.

# About Disruptive Analysis

Disruptive Analysis is a technology-focused advisory firm focused on the mobile and wireless industry. Founded by experienced analyst Dean Bubley, it provides critical commentary and consulting support to telecoms / IT vendors, operators, regulators, users, investors and intermediaries. Disruptive Analysis focuses on communications and information technology industry trends, particularly in areas with complex value chains, rapid technical l/ market evolution, or multi-sided business relationships.

Currently, the company is focusing on mobile broadband, network policy-management, operator business models and services, voice and personal communications applications including VoIP and IMS, smartphones, Internet/operator/vendor ecosystems and the role of governments and regulation in next-generation networks.

The company produces research reports and white papers, conducts consulting projects on technology strategy and business models, and provides speakers and moderators for workshops and conferences.

**For more detail on Disruptive Analysis publications, workshops and consulting / advisory services, please contact information@disruptive-analysis.com**

Disruptive Analysis' motto is *"Don't Assume".*

**Website: www.disruptive-analysis.com**
**Blog: disruptivewireless.blogspot.com**
**Twitter: @disruptivedean**

## Intellectual Property Rights / Disclaimer