

IT CHALLENGES IN THE MOBILE COMMUNITY

Considerations for CTOs when deploying
IT infrastructure for **mobile service** delivery



This white paper is based on research with mobile network operators and with content and service providers operating in the mobile community.

EXECUTIVE SUMMARY

RESEARCH METHODOLOGY

UK-based research agency FirstPartner undertook qualitative research to explore the commercial drivers, connectivity requirements, and hosting and colocation strategies for mobile content and service providers and mobile network operators. In-depth interviews were held with network operators and content and service providers, including those in mobile advertising, payments, content and app delivery, location-based services, messaging services and machine-to-machine (M2M) services.

The growing demand for mobile services across sectors ranging from media and entertainment to retail and financial services raises both commercial opportunities and IT challenges.

Based on research conducted by FirstPartner with mobile-centric service providers and network operators, this white paper reviews the challenges facing CTOs in the mobile community. We examine the importance of latency, and therefore connectivity, for businesses serving a global customer base with real-time services. We look at the different attitudes of early-stage and more mature companies to the use of public cloud services.

We conclude that mobile-centric businesses aren't always fully aware of the impact of their IT infrastructure deployment decisions and that there's significant value in doing more to analyse and understand the value of hybrid infrastructure deployments (either a mix of public and private cloud services or a combination of cloud and traditional on-premise hosting). We also believe that members of the mobile community may be missing a significant opportunity to optimise connectivity, reduce costs and improve flexibility by collocating infrastructure in carrier-neutral data centres.

THE RUSH TO MOBILE

IN THIS SECTION

The rapid growth of mobile channels and services is a game changer:

- Established players face competition from new entrants
- New entrants don't have time to take it slowly
- Everyone must meet increasingly high user expectations
- The mobile service delivery community grows increasingly complex

Smartphones and tablets are now the first screen for many consumers, and Gartner forecasts that by the end of 2013 mobile phones will overtake PCs as the most common web access device worldwide.¹

This rapid growth in the uptake of mobile services provides huge opportunities for new entrants and forces established players to review their strategies and business models. It also makes it difficult, especially for new entrants, to balance the pressures of product development with the need to grow rapidly to meet massively expanding demand.

For consumers and businesses, increased reliance on mobile computing and continually improving hardware drive expectations of high service quality, irrespective of the device or network being used.

The growing complexity of the mobile service delivery community (see figure 2) presents further challenges for a number of players in the mobile space, including those in media delivery, payments and advertising, all of which depend on an interconnected and interdependent web of partners to deliver services effectively.

Figure 1:

GROWTH IN MOBILE USAGE IN WESTERN EUROPE



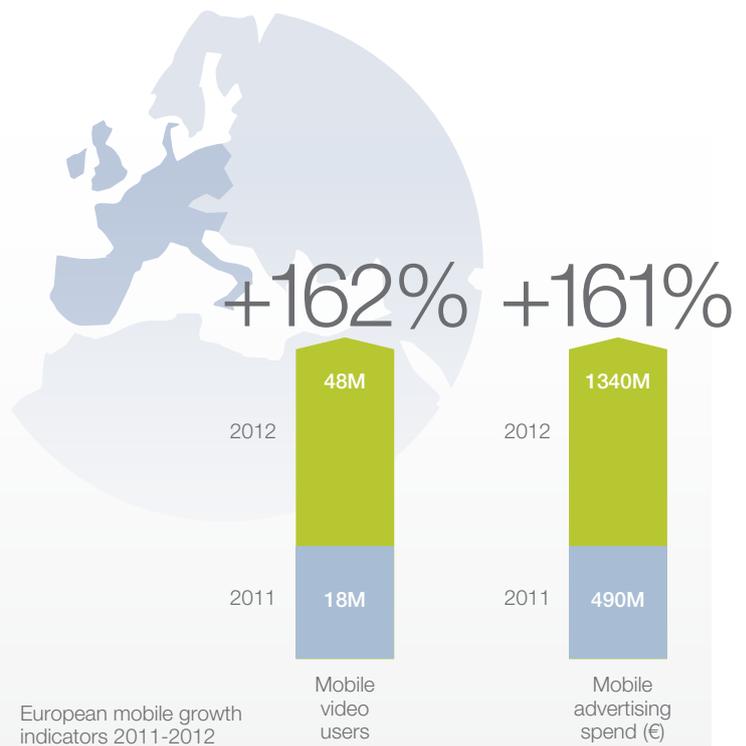
33% Internet users accessing via a smartphone in 2012



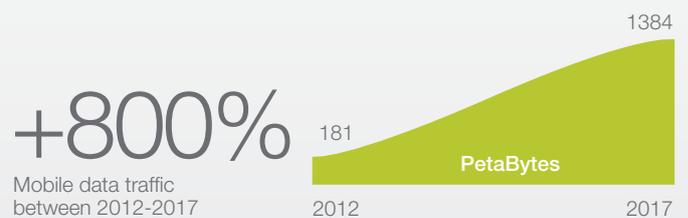
75% Of newly acquired mobile devices in Dec 2012 were smartphones



3.1x The rate of mobile vs. fixed IP traffic growth 2012-2017



European mobile growth indicators 2011-2012



Mobile data traffic between 2012-2017

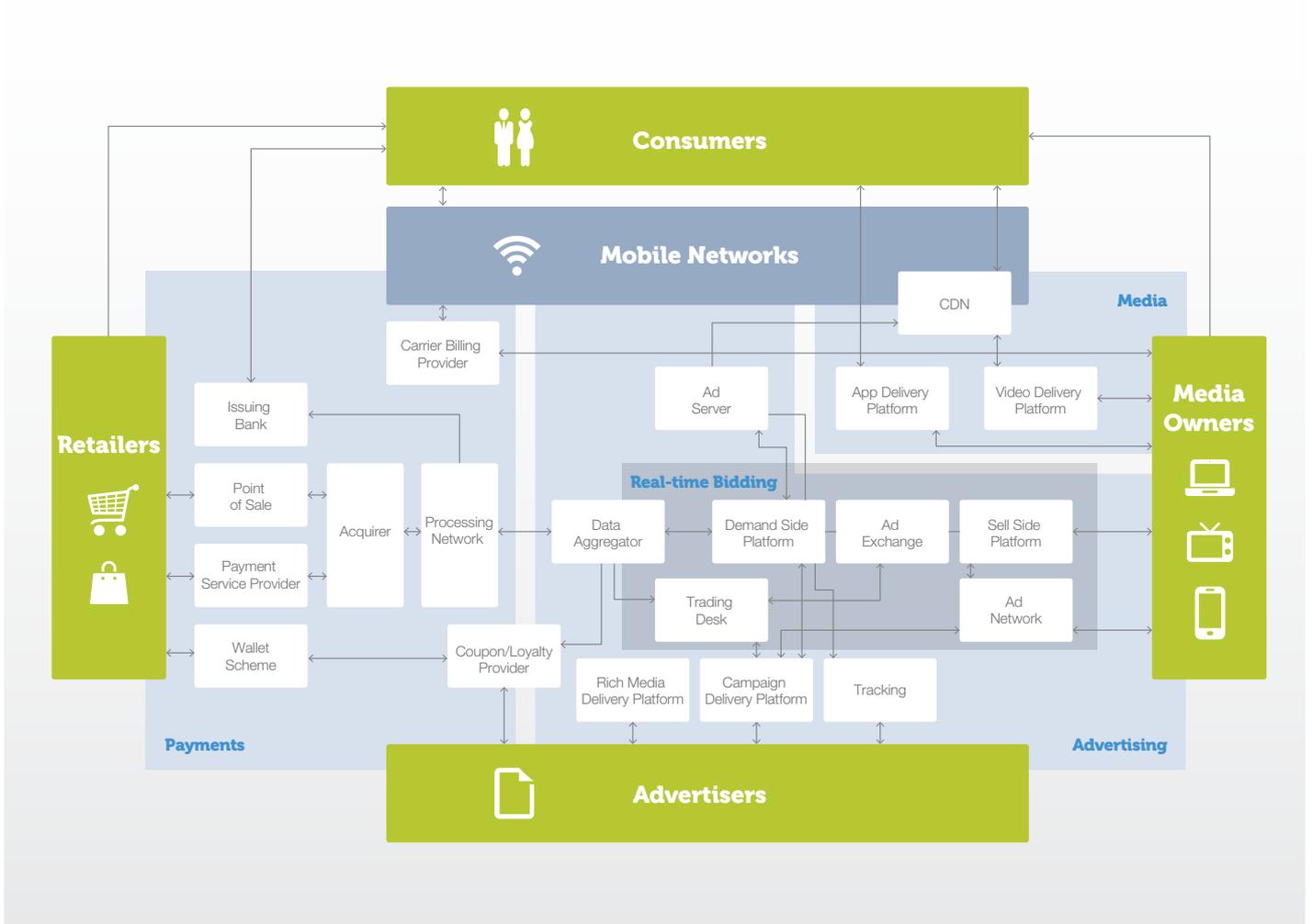
Sources

- Ofcom International Communications Market Review 2012
- comScore MobiLens
- FirstPartner Western Europe Mobile Advertising Spend Forecast 2012-2017
- Cisco Visual Networking Index 2012-2017

¹<http://www.gartner.com/newsroom/id/2429815>

Figure 2:

MOBILE SERVICE DELIVERY COMMUNITY: MEDIA, ADVERTISING & PAYMENTS



THE MULTIPLE CHALLENGES FACING CTOs

IN THIS SECTION

CTOs must consider a range of factors in their approach to deploying IT infrastructure for mobile services:

- Real-time performance, especially in latency-critical services
- International requirements and their impact on the physical location of infrastructure
- Protection of sensitive data through appropriate storage and sharing
- Investing appropriately to deal with volatile demand

REAL-TIME BIDDING

Real-time bidding (RTB)

matches bids for individual ad impressions in real time, optimising advertiser spend and publisher inventory. When a user views a mobile page or app, a **sell-side platform (SSP)** representing the publisher presents the impression, together with pricing and consumer profile information, to one or more **ad exchanges**. **Demand-side platforms (DSPs)** submit bids on behalf of advertisers and agencies interested in the audience and price range. The winning offer/bid combination is matched by the exchange and the ad is served. *All this happens in the fractions of a second between the user requesting the page and the page being served with the ad.*

Within the context of the business challenges posed by rapid growth in mobile services, the research identified a number of essential areas of focus for CTOs charged with delivering mobile services.

DELIVERING REAL-TIME PERFORMANCE

While some mobile services can absorb network delays of a few hundred milliseconds (ms) or more, others can't. Whenever speed really matters, variability in cellular network coverage and data throughput make delivering a consistent experience more challenging than over a fixed broadband network.

One of the most latency-critical services is real-time mobile advertising; with ad requests needing to be bid, matched and served in a round trip time of around 100ms:

"You have such a limited period of time to get the ad in front of the consumer; and if you don't get the ad in front of the consumer you don't get any revenue for it so it's really, really important." Mobile ad exchange.

The processing involved in matching hundreds of thousands of requests with bids is extremely complex, and the critical consideration is reducing latency through connectivity.

"We are squarely in that kind of access-at-scale market, and that means direct interconnects with all the major exchanges in the RTB side." Mobile advertising demand-side platform.

DELIVERING INTERNATIONAL REACH

For most of the providers interviewed, the ability to serve an international customer base is a given, necessary to compete in the market and essential to company growth. Retailers, media companies and app developers in North America and Asia address mobile users in Europe, Africa and South America, and vice versa.

"Our big focus is all about the connected world. It's about connecting to services that are disparate around the world." Mobile messaging technology and services provider.

"We have 500 merchants currently on the platform and around 380 network operators. We are the middle man managing that from a technical and business perspective. My merchant network is mostly in the US and Europe. My carrier network is very strong in Europe and in Asia because that's where carrier billing is used." Carrier billing payment service provider.

It's also about delivering locally appropriate advertising and payment collection services: a Korean games developer needs to collect payments from players in France via card or other payment options preferred in France; German consumers viewing US sites should be served German ads; and so on.

Given the importance of real-time service delivery and the resulting need to reduce latency, a key factor in international service delivery is the physical location of service deployments. Providers need a physical presence in or near target regions to effectively serve customers and partners there. For example, it's becoming increasingly essential for high-growth US and Asian companies to duplicate infrastructure in Europe to serve European markets, and to find the right geographical locations to optimise connectivity for their target regions.



“We’re going to see more and more advertising inventory coming on the market; and when you operate across geographic markets you want to host as locally as you can, to improve or to reduce latency.” Mobile ad server.

Typically mobile service providers will deploy IT infrastructure in one to three data centres per region for service continuity, in some cases integrated with public cloud services to provide additional flexibility and backup.

“Regionally you want to have some redundancy as well, so I’d rather have two European data centres, two US data centres, potentially two Asian data centres in the future; and be able to have a framework that allows continuous uptime in the event there is a major outage.” Mobile advertising demand-side platform.

PROTECTING SENSITIVE DATA

Customer data is a key component of mobile services, used to target advertising and recommend content as well as handle billing and payments. Like all organisations subject to data privacy regulations and other compliance and security requirements, service providers must take responsibility for the integrity of personal and commercially sensitive information and control where such data is stored.

“There are issues around transporting certain types of content and data in and out of country; for example in sectors like tele-health where there are all sorts of issues around data privacy, you can’t backhaul data out of one country to a server in another and bring it out again. It’s a bit of a minefield.” M2M service provider.

With data security a major concern for mobile network operators, some insist that third parties with whom they share sensitive customer data must be located within the operator’s own data centres.

“The policy is to allow third-party content providers to use our data centre facilities but this is on a justified demand basis and may depend upon the service they’re providing. It’s not an automatic provision. Where a service uses our data, is of high value or of a sensitive nature, we do care and it would be allowed accommodation.” Mobile network operator.

DEALING WITH RAPID, VOLATILE GROWTH

The rapid and unpredictable growth in the mobile market makes it difficult for companies to effectively plan their infrastructure investment, introducing significant risks of either exceeding or under-utilising capacity. This is a particular challenge for the many early-stage companies that have entered the mobile space:

“This is a dynamic situation. We’ve been going for six years and looked at predicting growth sectors every year and got it wrong. Everything in M2M is going to grow. We don’t know which sector is going to be the dominant one.” M2M connectivity provider.

As well as coping with rapid underlying growth, significant short-term and seasonal demand peaks are a common issue in some sectors, notably consumer media. In recent years both mobile advertising and m-commerce have experienced strong peaks in the quarter leading up to the end-of-year holiday season, while major events and TV shows always cause a peak in demand for associated mobile apps and streaming services.

COMMON MITIGATION STRATEGIES

IN THIS SECTION

Addressing the challenges of mobile delivery takes various forms:

- Recognising the importance of connectivity to reduce latency
- Utilising public cloud services, a strategy particularly adopted by early-stage companies
- Using hybrid environments as maturity develops and where security and compliance are significant issues

Companies are adopting various strategies to address the challenges of delivering real-time mobile services internationally, without compromising service quality, compliance or security. But in our view, more needs to be done to understand the options and their implications.

OPTIMISING CONNECTIVITY

The underlying need to optimise connectivity is widely accepted, especially in the context of reducing latency through partner and inter-regional connectivity.

“There will be an ever-increasing focus on latency so we’ve got to do everything we can to reduce it. Sometimes it’s very good and sometimes we need to improve on it. Sometimes it’s just because you’ve got a transatlantic hop and that can’t be avoided.” Mobile ad server.

“So how do we best do this interconnection, particularly over the ocean; you lose quite a bit of speed in the transaction. We have a redundant data centre setup in Europe and then we have a hub in North America that connects for merchants, and strong interconnection between the two hubs.” Carrier billing payment service provider.

In an ideal world, key partners within a latency-critical interdependent community would be colocated in common data centres.

“Where we can, we will colocate physically in the same data centre. If you’re putting, say, 20ms each way onto that 100ms round trip then we’ve effectively reduced our decision time – and therefore the complexity of our data and algorithms that can be applied at impression time – by 40%. So you know, the closer that gets to zero, the better.” Mobile advertising demand-side platform.

At present there’s a view among mobile-centric businesses that this ideal scenario may be impractical, given the number of potential integration partners and competitive pressures:

“You’ve got to remember we are just one of maybe 10 sources of supply that these people hook into, so are they going to follow us and get closer to us and maybe further away from the other nine? I don’t know; in terms of getting ads delivered, it would always be better to have everyone sitting on the same environment, but in terms of the realities of it, what you’re asking is the entire industry to make a procurement decision and I just don’t believe you would ever get agreement.” Mobile ad exchange.

This is a common enough view, but other industries have found that what may seem impractical often isn’t, and a community of interest can come together in a data centre hub without any centralised decision-making. For example, it’s become extremely common for the capital market trading community to colocate trading infrastructure in relatively few common data centres because of the significant benefits, including cost and latency reduction. This is the case even though it’s a highly competitive industry and participants need to connect with multiple exchanges and information providers worldwide. Within the video content space, too, a colocation strategy is being pursued for real-time bidding by a major ad exchange platform.

In the absence of colocation, it certainly remains true that well-connected data centres or cloud services are essential to achieving the necessary interconnectivity with key partners wherever they may be.

STRATEGIES TO DEAL WITH VOLATILE GROWTH

The research suggests that strategies for dealing with volatile growth vary by company size and maturity, among other factors.

Many early-stage companies are growing up using public cloud services because they can be deployed immediately, require no up-front capital investment, and deliver real-time flexibility to meet volatile growth and short-term changes in demand.

“A large number of companies in the online real-time bidding and mobile real-time bidding space are hosting their services in Amazon.” Mobile demand-side platform.

“It works really well for our business – you are paying for computing time and bandwidth rather than the physical cost of the hardware. We can be told about extra demand in the morning and by the afternoon our system has automatically scaled itself to support a million concurrent users or more. That’s the advantage with hosting versus the full-on data centre.” Startup mobile interactive TV company.

Figure 3:

ARC OF MATURITY & HYBRID CLOUD INFRASTRUCTURE





But complete reliance on public cloud infrastructures may not be appropriate as a company matures or if it faces specific security or compliance challenges.

Larger and high-growth companies voice uncertainty about the cost benefits of public cloud services for their core mobile applications, once a certain size threshold is reached.

“A lot of start-ups will go into Amazon and not really consider what the consequences are as they start to scale up. Our research at this point is that it’s more expensive to run in a cloud environment like that.” Mobile advertising demand-side platform.

The research indicates that companies haven’t generally analysed the impact of relying entirely on public cloud services. When asked about their understanding of the threshold that might trigger a move to or from a public cloud model on the basis of cost, a typical response was:

“To be honest with you I couldn’t tell you. We haven’t done that analysis. I’ve seen a few articles on the internet where people talk about the tipping point. But we’ve not looked at it internally for a while.” Mobile advertising demand-side platform using a public cloud service.

Clearly more should be done to analyse and understand the implications for high-growth companies of a strategy that they choose when they’re still young and small.

Finally, organisations handling or processing sensitive data expressed concerns over the appropriateness of public cloud services:

“On the payments side of the business we have quite a lot of dedicated hardware all clustered, sitting in our ring-fenced portion of the data centre. When we are sending out low-risk information, catalogue information, product information, typically we put that on a public cloud.” Mobile e-commerce provider.

“We won’t look at public cloud services as they cannot be protected and cannot be isolated. The need is to house the equipment and to have a solid internet networking layer.” M2M connectivity provider.

“We are using some third-party services but we also have an operations team that look after our own environments. We are US government- and US emergency services-certified so we have to have a degree of control and visibility as we are accountable at the end of the day.” M2M service provider.

It seems that while public cloud services are ideal for startup companies and those without a heavy compliance burden, maturity demands a solution that matches flexibility and cost minimisation with high levels of control. Hybrid approaches combine dedicated infrastructure and private cloud services with public cloud services to provide a balanced model.

BALANCING DEMANDS: THE POTENTIAL OF CARRIER-NEUTRAL DATA CENTRES

IN THIS SECTION

Carrier-neutral data centres have great potential to help mobile-centric organisations address their key infrastructure needs:

- Very high levels of connectivity
- Hubs for cross-connected communities to minimise networking costs and latency
- Ideal for balancing flexibility and control through hybrid deployments that combine public cloud services with private infrastructure

We've seen that two key themes characterise the issues posed by mobile services:

- **Optimising connectivity:** to reduce latency and ensure quality of service for a global customer base.
- **Balancing flexibility and control:** to ensure cost-effectiveness and cope with volatile demand, without compromising security or compliance.

For all members of the mobile community, carrier-neutral data centres (CNDCs) can offer a highly effective way of delivering against both requirements.

Because true CNDCs are independent of network, hardware and software providers, they can attract a wide variety of network operators, ISPs, internet exchanges, content distribution networks (CDNs) and cloud service providers to their facilities.

CNDCs can therefore act as 'connectivity hubs' for communities of companies in the mobile space. These hubs can either be physical clusters of interdependent companies colocated within the same data centre, utilising shared resources and low-cost cross connects; or they can be 'virtual hubs': highly connected centres offering cost- and latency-efficient connectivity to geographically dispersed organisations.

CNDCs can also be ideal for hybrid deployments that combine the security and control offered by dedicated or 'shared private' infrastructures, with the flexibility of public cloud services or edge computing. Why? Because as well as offering scalable space and power for dedicated infrastructure deployments, a cloud-neutral CNDC will offer excellent connectivity to multiple public cloud providers, possibly even through cross connects.

CONCLUSION: THE WAY FORWARD FOR MOBILE-CENTRIC ORGANISATIONS

AUTHOR BIOGRAPHIES

Mike Hollands has over 18 years' experience in the communications sector, holding a variety of product management, marketing and business development roles at leading service providers before joining Interxion. As Director of the Connectivity Segment at Interxion he has a neutral and insightful perspective into the challenges facing mobile service providers and network operators across Europe.

Richard Warren is FirstPartner's managing director. With a long background in telecommunications and media product development and marketing, he has led international research and strategy projects for major companies including BT, Vodafone, Oracle, Microsoft, TomTom and Sky.

The criticality of connectivity and the need to balance flexibility and control make infrastructure deployments, and therefore data centre choices, very important for mobile-centric businesses. In such a fast-growing and changing community, we believe they should be considering hybrid configurations and should regularly be asking the following questions:

- How can we accommodate international growth while at least maintaining, if not improving, quality of service?
- How can we meet evolving data storage and privacy requirements?
- How can we control performance and security while maintaining flexibility and minimising cost?
- Are current infrastructure deployments in the optimum location and delivering the connectivity required?
- Are current cloud deployments cost-effective and providing the level of control required?

In considering the implications of these questions, we also believe that all members of the mobile service community should seriously consider the value of a CNDC in terms of connectivity options, location, flexibility and cost efficiency, together with their access to public cloud services.

ABOUT INTERXION

Interxion (NYSE: INXN) is a leading provider of carrier and cloud-neutral colocation data centre services in Europe, serving a wide range of customers through over 35 data centres in 11 European countries. Interxion's uniformly designed, energy efficient data centres offer customers extensive security and uptime for their mission-critical applications. With over 500 connectivity providers, 20 European Internet exchanges, and most leading cloud and digital media platforms across its footprint, Interxion has created connectivity, cloud, content and finance hubs that foster growing customer communities of interest.

For more information, please visit www.interxion.com

interxion™

www.interxion.com

INTERNATIONAL HEADQUARTERS

Main: + 31 208 807 600

Fax: + 31 208 807 601

E-mail: hq.info@interxion.com

EUROPEAN CUSTOMER SERVICE CENTRE (ECSC)

Main: + 44 207 375 7070

Fax: + 44 207 375 7059

E-mail: customer.services@interxion.com