metaswitch

Disruptive Analysis

Don't Assume

Integration of IoT & Communications.

Potential opportunities for service providers & cloud-hosted communications services platforms.

A Disruptive Analysis thought-leadership paper.

By Dean Bubley | January 2017

FOREWORD

This paper explores emerging intersections between IoT and real-time communications. It frames and segments the use cases, and examines the opportunity for telecom-type services, especially where enabled by NFV and cloud-based IoT services platforms. The reduced cost and greater flexibility of such platforms may transform telcos' M2M business units, and enable the emergence of other service providers focused on IoT. We may see new MVNO and PaaS players emerge, taking advantage of more flexible infrastructure. IMS may itself move beyond its SIP roots, to embrace a wider diversity of IP "sessions" and M2M messaging functionality.

INTRODUCTION

Everyone is familiar with the projections for the Internet of Things (IoT) – billions of new devices being connected to networks and IT systems, with the potential to transform consumers' lives, businesses' productivity and governments' effectiveness. The communications industry sees IoT as a new source of connectivity and platform revenue, especially welcome as consumer-oriented telephony and SMS decline.

But what is little discussed is how a proportion of IoT solutions will, in fact, link into and drive more usage of traditional communications – increasingly meaning IMS/VoLTE where a 4G connection is involved, for example. There will be many ways that new devices integrate voice and messaging functions (plus video and maybe AR/VR in time).

In some cases, humans will call directly from non-phone devices such as cars or smart watches with integrated voice capabilities. In other applications, an IoT event will trigger a call or message elsewhere, such as a temperature sensor crossing a threshold leading to a call from a supervisor to an engineer.

Another possibility is an IoT solution used to improve a normal call or collaboration – for example, smart headphones providing background noise cancellation, or proximity awareness. And we may simply see the conventional phone network act as a fall-back option if an IoT-based Internet VoIP or video session fails. We will also see combinations of messages together with "calls."

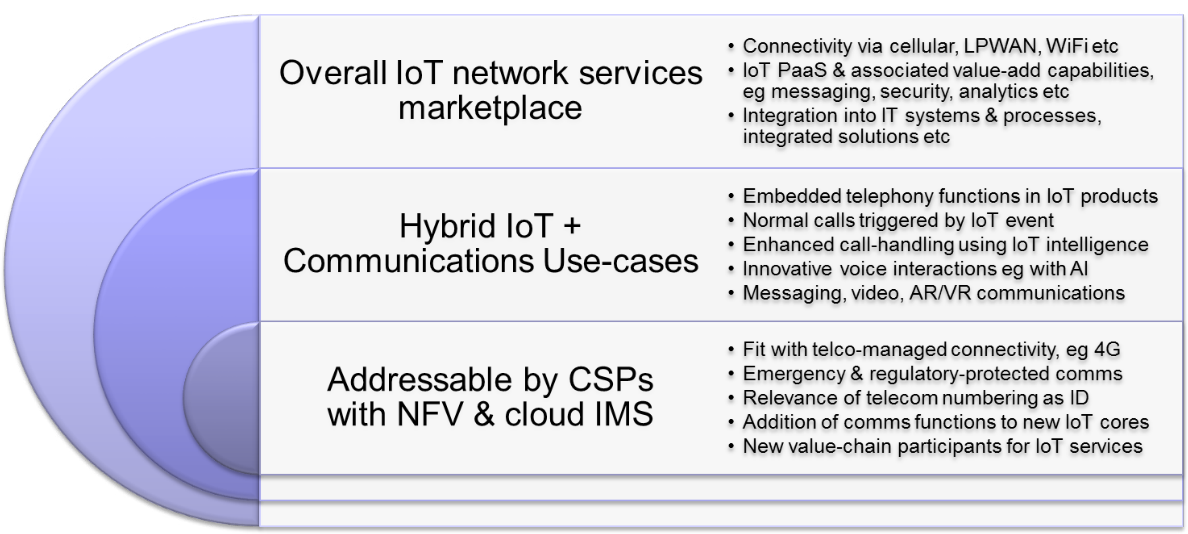
This does not mean all instances of audio/video/messaging in IoT are open to telcos and CSPs (communications service providers). Many video terminals are based on "walled garden" WebRTC platforms. Gunshot detection for emergency services uses a dedicated application and hardware. Most IoT-originated messaging and video will not be linked via operators' platforms optimized for user-to-user communications either (Disruptive Analysis remains deeply sceptical of ViLTE and RCS).

But this still leaves a sizeable overlap between the worlds of devices and carrier-grade telephony, SMS or other simpler and more open messaging applications such as MQTT and CoAP, or REST APIs driven by standards like OneM2M. There may also be future roles for IMS-type infrastructure in managing other network elements such as policy control for QoS, message and media storage or mobile edge-computing nodes, or functions such as security and device provisioning.

TOUCH POINTS BETWEEN IoT & VOICE/VIDEO

Only a fraction of IoT operations will involve voice or video connections – but in a world of tens of billions of devices, even an occasional input or action will drive a sizeable addition to the market for human communications. At this stage of evolution, it is less clear exactly how and where the traditional telecom industry can benefit from this, besides wide-area connectivity, given the fragmentation and complex value chains involved. However, lowering the bars of cost and accessibility/scalability through NFV, open-source and cloud-based infrastructure should help – as should easier-to-use developer toolkits, open APIs and PaaS providers, and the ability to create/prototype concepts more rapidly.

Figure 1: A subset of IoT solutions will involve communications & CSP roles



Source: Disruptive Analysis

Before digging in to the specific use cases for real-time communications around IoT, it is worth first noting the sheer diversity of areas the new industry covers. To help frame the discussion, consider the following (partial) list of major sectors and systems:

- **Vehicles**
  - Connected cars
  - Buses, trucks and& other commercial vehicles
  - Boats and planes
- **Energy & utilities**
  - Smart meters
  - Smart grid control
  - Water and gas distribution systems
- **Smart buildings & cities**
  - Security and& access control
  - HVAC (heating, ventilation, air-conditioning)
  - Parking and traffic systems
  - Infrastructure monitoring
  - Lighting
  - Transport networks
  - Elevators
  - Advertising and vending machines
- **Agriculture & environment**
  - Irrigation
  - Animal tracking
  - Environmental sensors, e.g. pollution measurement
  - Disaster management, e.g. flooding, earthquake sensors/controller
- **Personal & wearable**
  - AR/VR headset
  - Smart watches and& fitness bands
  - Medical monitors and& devices
  - Child / elder trackers and panic -buttons
  - Voice-based virtual assistants
- **Business & industrial**
  - Office machinery (e.g. photocopiers, printers)
  - Robots and& drones with cameras
  - Process control
  - Vertical-specific (anything from hotel minibar, to oilfield pump)
  - AI/voice/chatbot-driven customer service

One of the difficulties with IoT market analysis is this huge fragmentation. The dimensions for analysis span solution cost/complexity, connectivity type (e.g. cellular, Wi-Fi, fibre); criticality/reliability needs; power restrictions; presence of display and mic/speakers; industry sector; and much more. However, it also means that the risk of commoditization across the whole sector is low. There are too many possibilities for any one company or platform to address all of it.

To help make sense of this, from the perspective of voice, M2M messaging and cloud-based platforms for service providers, Disruptive Analysis believes that there are three broad ways in which IoT can intersect with communications apps and services -- although these can all be broken out in much finer grain:

- **An IoT device itself supports communications.** For example, a connected car has a microphone and speaker, used for contact between driver and the emergency services, or perhaps a mechanic for remote diagnostics. This could be applied as well for "call an engineer" buttons, especially on high-value items with usage-based revenue streams, such as photocopiers or "capsule" coffee-machines. For video, there are numerous use cases involving CCTV, drones, or telepresence robots. It may be that the device acts as a telecom end-point, routed across various access networks from an IMS network, or it could communicate via some form of gateway or cloud platform.
- **An IoT event triggers a person-to-person call.** For example, a sensor in a machine detects a fault, and the platform initiates a call from supervisor to a local engineer to check out the problem. In this case, the IoT management and control platform (typically cloud-based) can either directly support telephony/messaging capabilities, or it can use the APIs of a separate service provider. Complex hybrids are possible as well – for example, a sensor highlighting an anomaly with a machine, and alerting a supervisor with a message, whilst simultaneously starting to stream and record an audio or video feed of the apparatus in question.
- **An IoT device enhances a call flow, application, utility or behavior.** For instance, a "smart conference room" adapts audio or video media to better fit the acoustics of the environment, or the seating pattern of the participants. A wearable could recognize a person is running for a flight – or on a treadmill at the gym – and send all but the most important calls to voicemail.

In addition to these main categories, various other possibilities exist. For instance, on an IoT device with no screen, it might be possible to use a voice call or message to a phone for security purposes such as two-factor authentication. Significant growth is also occurring in voice control of devices (like Alexa, Siri or Cortana), although strictly speaking that is not really "communication."

Figure 2: Which device types will embed voice communications capability?

	Narrowband Data	Broadband Data
Integrated Voice Capability Useful	Door Entry System Child Tracker Elder Tracker Medical Kit eg AED PoS Terminal Elevator Luxury Coffee Machine	Car / Truck CCTV AR/VR headset Smart watch Fire Control Panel Digital Advert Displays
Integrated Voice Capability Not Useful	Environmental sensor HVAC control Animal / Pet Tracker Shipping Container Utility Meter Street Lights Waste Container	Drone Industrial Robot Smart Grid Digital Signboards

Source: Disruptive Analysis

There will also likely be multiple solutions for any given problem – a declined transaction at a restaurant could trigger a call directly to a portable PoS terminal, or discreetly connect to a separate phone at the maître d's podium. For each use case, there will be various considerations relating to network architecture, application/interface design, platform capabilities and regulatory requirements.

In addition, not just the device but also the associated network needs to be able to support communication – an IoT object connected directly with LTE or Wi-Fi can support voice on non-constrained devices, but if it uses a low-power WAN like SigFox or NB-IoT for power-constrained devices, or perhaps ZigBee for short-range connections, it cannot. This is considered in a later section of this document.

There is also another important dimension to consider here – there is a distinction to make between single IoT devices and large "fleets" such as sensor grids. It is implausible that every pollution-monitoring unit in a smart city will have its own voice connection and associated number. It will also not be feasible to stream and store "raw" video over cellular connections from millions of CCTV cameras twenty-four-seven – some pre-processing in the device is likely, or a trigger mechanism.

HOW MUCH OVERLAP WITH TELECOM-TYPE SERVICES?

A key question here is the extent of the addressable market for CSPs in mediating IoT/communications interactions -- and by extension, what that means for the platforms enabling them. Some are obvious, such as normal telephony calls from a car. But other instances highlighted need analysis from the perspective of the application designer.

So while it is possible to construct many scenarios for communications involving IoT, not all of these will be appropriate for telecom-style services such as telephony. It seems unlikely that a karaoke machine – even if networked – will be connected to a cloud-based IMS infrastructure, for instance, while an elevator may well be. Similarly, most integral messaging in industrial environments, between sensors and control systems/central IT, will be based on protocols such as MQTT, CoAP and OneM2M, while an alert for a remote-working supervisor might use SMS, which will not be appropriate for normal operations.

The best indicators of "telecom-friendliness" of IoT voice/messaging/video include:

- Similar usage and interaction models to normal communications – for example, a "call an engineer" button works in a similar way to a normal phone call.
- Both ends – or all participants in multi-way interactions – support similar capabilities natively on their devices/platforms.
- The presence of regulatory mandates – e.g. need for lawful interception, support of emergency calls and so forth.
- Relevance of E164 phone numbers for identity or other reasons.
- The need for ultimate levels of reliability and predictability – for instance, if one device or person may not have access to a broadband connection capable of supporting Internet VoIP reliably.

- A need to fall back to circuit-based telephony at least occasionally (e.g. a vehicle or forestry worker outside of 4G network coverage).
- Use of a SIM or eSIM as an authentication mechanism for an interaction.

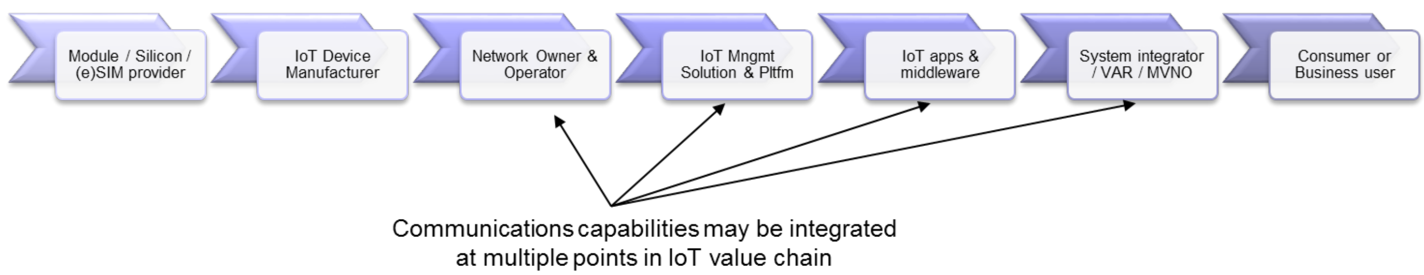
An interesting possibility is that telecom infrastructure, such as IMS, goes beyond its traditional role of setting communications sessions (today, typically based on SIP – session initiation protocol) to extend to more granular Web- and M2M-centric services such as MQTT messaging, or initiating functions such as storage or authentication. Perhaps a more appropriate metaphor could be "interactions" rather than sessions.

THE EXTENDING IoT VALUE CHAIN

An interesting factor here is the diversity of participants in the IoT ecosystem. The range of product types and use cases has led to a proliferation of platform providers, middleware solutions, application vendors and systems integrators. Some of these act in similar fashion to MVNOs, while others are more like conventional IT firms. The touch points for voice/video/messaging are likely to be vertical-specific.

There has been some expansion and consolidation across sectors here as well – for example, with Cisco acquiring IoT connectivity provider Jasper. There have also been increasing links between IT and telecom-type enablers – communications PaaS player Twilio has partnered with T-Mobile US for enabling SIM provisioning for IoT devices. IBM has shown integrations of IoT endpoints with its Watson AI platform to interpret speech or video images.

Figure 3: Lengthening IoT value chain yields many options for integration



Source: Disruptive Analysis

This value-chain extension is both an opportunity and a risk. Product and solution developers will likely have access to multiple different sets of APIs and capabilities provided as a service. They are unlikely to be loyal to any one architecture or set of standards, unless particular criteria such as regulation hold sway. They may swap from third-party platforms to in-house alternatives if costs scale too fast, or if greater control and customization is needed.

SELECTED KEY VERTICALS / USE CASES

The following sections consider some of the most interesting verticals for IoT/communications hybrids. They cover:

- Connected cars and other vehicles
- Wearables and personal/medical care devices
- Smart buildings and cities
- Public safety

CONNECTED CARS & OTHER VEHICLES

Increasingly, cars and other vehicles (buses, trucks, etc.) are being equipped with cellular radios. There are various use cases, including provision of entertainment and Wi-Fi hotspot access to passengers, collecting telemetry data, download of vehicle software updates and more. In future, there will likely be additional capabilities for autonomous/remote driving control, live streaming of sensors or "dash cams" (maybe bot-cams?) or vehicle-to-vehicle interaction.

But as well as data connectivity, there is also a desire for voice and messaging applications. For many years, GM's OnStar system has provided in-vehicle assistance and emergency calls in the U.S. and other markets, while in Europe the new eCall directive mandates that vehicles need to be able to connect automatically to 112 emergency services. There are also likely to be other applications, such as direct voice connections to roadside repair, or engineering assistance/customer service for the vehicle manufacturer. While some communications functions will remain anchored in the driver's or passengers' own phones (and identities), others will relate more to the vehicle itself.

There will likely be a mix of circuit-switched communications and VoLTE used, depending on the use case, and the cellular coverage available in a given location. For now, the European eCall law actually stipulates the use of GSM calls, with a limited amount of data provided in-band as well as voice. While this may transition to VoLTE over time (or whatever Vo5G becomes), it is likely to be a decade or more before the circuit capability is switched off. Other regions may move to all-IP sooner depending on local regulations, or re-farming of 2G/3G spectrum.

One interesting angle here is the emergence of various new players in automotive connectivity. A number of IoT/MVNO specialists that focus on this area are emerging, especially as vehicles often drive across borders and thus roaming needs to be managed. The vehicle manufacturers themselves are also heavily involved in the provision of cellular services, especially as they need to pre-integrate radios and SIMs – and are also looking to areas such as autonomous driving and value-added services. The automotive sector is also one of the lead adopters for embedded eSIMs, as there is a desire to be able to provision (or switch) network providers remotely.

All this means that there will be new market participants requiring platforms for voice capability, for both emergency and non-emergency use. In some cases this may be provided on a wholesale basis by the underlying network operator, but there is also an opportunity for the new tier of service providers to build/operate their own. Cloud-based communications platforms allow for greater flexibility and scalability in such cases.

WEARABLES & PERSONAL/MEDICAL CARE

An important area for communications-centric IoT is the overlap of devices with humans' own space – either where products are worn, or where they monitor health and welfare. While almost everyone has a smartphone with advanced voice and video capabilities, this does not mean other objects cannot also be used for interactions.

The most obvious examples are probably smart-watches and various types of headset. It is already possible to use Samsung's Gear and Apple's Watch for spoken applications, either linked via a phone, or standalone (the Gear S3 Frontier version supports VoLTE). Many people in the technology industry anticipate a huge surge not just in watch/fitness bands, but also AR/VR headsets in coming years. Some form of in-built communications is inevitable in most.

But personal IoT extends beyond these into other niche domains. There is increased interest in tracking devices for children, or perhaps dementia sufferers and others. These could well incorporate both panic/emergency features, and just the ability to reach a relative or helper. Other medical devices could also embed communications functions – for example, the automatic defibrillator machines used to resuscitate heart-attack sufferers could have in-built connections to first-responders to offer assistance, as well as providing data on the patient via the same connection.

The inverse is also possible -- normal "human" communications sessions can be enhanced by incorporating IoT connectivity. For example, a doctor's telemedicine consulting session with a remote patient may blend in streamed data, perhaps from a blood-pressure sensor or an ultrasound detector, connected to the same PC or phone the patient is using to speak.

Many of these areas are ideal for telcos and CSPs to pursue, or at least to enable via MVNOs or platform specialists. The telephony model is likely to prevail, and there may well be a need to link to emergency call systems. Some applications will probably be non-telco-oriented (e.g., in-game communication between players using VR headsets), but that does not invalidate the more general opportunity. There will also be broad requirements to support backup options – traditional calls and SMS especially – as well as integrate with other systems, for example for alerting relatives.

SMART BUILDINGS & CITIES

Another key area for IoT is the creation of more-intelligent networked buildings and cities. This is a huge area in its own right, spanning everything from transport and utility infrastructure, to security monitoring, power management and the delivery of services to workers, visitors and citizens. It also intersects with the smart home marketplace.

While there are unlikely to be direct voice/video/messaging integrations for pollution sensors or public transit timing signs, there are other touch points that do offer opportunities for network operators and other telecom-style service providers. In some cases, these will be IP-era updates to existing circuit-telephony applications, such as elevator emergency calls, or building door-entry systems that dial a resident's phone. But we may also see the addition of voice challenge/response functions to CCTV cameras, the emergence of new public information points or kiosks, and so on.

Smart buildings and cities will also have multiple sources of IoT "triggers" for normal human-to-human communications. Failures of lighting systems could prompt maintenance action, while waste collection will be triggered when bins are mostly full. Home alarms could send message alerts, and simultaneously stream surveillance video to the cloud for recording. While many interactions will be automated and message-driven, others may require voice communication among multiple parties. There may well be associated requirements for security involved too here – authentication of devices and users, secure storage and, potentially, "data integrity" management to ensure that no tampering with the data can occur.

Some of these applications can be easily done with non-telecom VoIP or video communications (e.g. fixed information kiosks with video displays), but others will be stronger opportunities for traditional telephony or enhanced telecom-style services. Anything requiring emergency-call access or interoperability with legacy telephony systems, or which is subject to tight regulations, will probably require a more conventionally managed communications mechanism, rather than one reliant on an Internet connection.

PUBLIC SAFETY

There are many ongoing projects and standardization efforts to improve communications for public safety and emergency services, often linked to IoT innovations as well as IP/broadband upgrades.

This includes both evolving emergency 911/999/112 calling, and updating the services' operational command-and-control infrastructures. In some countries, there is an intention to adapt LTE networks to serve as the basis for future connectivity and communications, typically using the PS-LTE standard being developed by 3GPP. Many new applications require forms of IP communication, such as mission-critical push-to-talk (MCPTT), or broadcast messaging alerts.

There will be a requirement to integrate with numerous non-traditional devices that either have in-built communication functions or trigger calls: for example buildings' fire-alarms and control panels, medical equipment such as resuscitation devices, and the in-vehicle systems already discussed. There will be strong requirements for such devices to support high-quality voice, and for the back-end systems to store communications securely, perhaps for medical records or legal evidence.

There are also parallel efforts to improve the IP connectivity of the answering points (PSAPs), and improve their ability to handle data from the public (e.g., messages and videos from the scene of an accident), from vehicles and from emergency workers themselves (e.g., wearable cameras or drones). There are also many interconnections with other relevant organizations such as hospitals, courts, lawyers and sources of "situational awareness" such as buildings' blueprints or facilities managers. Systems will also likely need to interoperate with older TETRA networks, or circuit-based emergency calling and PSAPs, for many years – it is probable that older infrastructure will be kept for backup and fail-over capability.

Given the varying requirements, policies and budgets of different countries, network architectures will vary substantially. An emergency service may build a completely separate network, even using private frequency bands. Or they may contract or host capabilities in a telcos' network. At present, IMS is not being mandated in PS-LTE, but it appears that some deployments will use it for VoLTE or MCPTT at least (e.g., the new UK Emergency Services Network). And while many of the requirements and projects will be driven by normal emergency calling or radio communications, an increasing intersection with various IoT elements will also occur over time.

REALISTIC TARGET IoT/IMS SERVICES FOR SPs

For CSPs (and their suppliers) to benefit from the IoT/communications opportunity, it will be important to be both realistic and opportunistic. The space is evolving so fast, with so much competition from adjacent sectors, that attempts to create over-ambitious architectures will create delay and lead to failure. Trying to use the trend to shoe-horn unloved solutions like ViLTE into the market is risky.

By the same token, while 5G networks are hugely promising in scope and potential, they are likely several years away from deployment for IoT use cases. By that time, it is likely that some IoT/communications solutions will already have momentum which is difficult to counter. It will also be important to consider integrating 3GPP and non-3GPP technologies pragmatically – the IoT will be a world where a variety of other standards (e.g. LoRA and Bluetooth) coexist with cellular and assorted proprietary networking technologies. Integration and migration can occur at a later date, if needed.

Upfront, it is important to think about different communications media and formats:

- **Telephony:** The mainstay of the telecom and IMS marketplace, this encompasses both the normal "person A calls person B" phone interaction and various other models that can be "hacked" using APIs or voice PaaS options. Generally, it is regulated, subject to lawful intercept and linked to normal E164 phone numbers. **This should be the focus of CSPs' efforts to bridge IoT and IMS.**
- **Conferencing and collaboration:** Business-grade platforms for voice, web and video conferencing. These blend telephony for "dial-in" with an array of controls for meeting management, recording, appointment setting and so forth. They can span from room-sized devices to smartphones, and increasingly incorporate features for file sharing and messaging. There are some links to IMS and a large role for cloud components like SBCs, although the majority is not delivered from carrier NFV platforms.
- **Non-telephony voice and audio:** This covers a broad array of applications and possibilities, including everything from push-to-talk, contact center IVRs, gamer chat, personal AI-type assistants like Siri and Alexa, one-way "whisper mode" chat, emotional analysis, karaoke and even medical diagnosis of breathing sounds. Often, it couples the transport of voice media traffic with various types of processing and computation, frequently cloud-based. It sometimes requires particular device capabilities (e.g. stereo playback, or particular microphone characteristics).
- **Conversational Video:** This covers "video chat and calls," exemplified by applications such as Skype, FaceTime and many others. Increasingly conducted via WebRTC in both browsers and mobile apps, it is becoming much more important, but also highly competitive and requiring a very nuanced understanding of user behavior and psychology, as well as network quality.
- **Non-Conversational Video:** This spans an ever-broadening array of applications, from CCTV camera footage, to video contact-centres for customer service, live-broadcast apps, and increasingly a new world of VR/AR/360 video and "enhancements." It could even include infrared, eye-tracking, game-streaming and a variety of other forms of "moving image." Increasingly, video is being blended into other applications (e.g. social, games, e-commerce, dating), either as embedded video-chat or in other more-complex formats.
- **SMS / MMS messaging:** Along with email, SMS is the most widely used form of text-based message, plus image/media attachments with MMS. It is increasingly used for application-to-person use (A2P) as well as inter-personal (P2P), especially for two-factor authentication and also various transaction notifications. Ubiquitously available for all mobile users, it links to E164 phone numbers and has an important role for all telcos. It is also both regulated and used by other regulated services (e.g. notification of charge for roaming). Although it stems from circuit-based heritage, much modern SMS infrastructure is IP-based, or delivered via cloud-based PaaS or aggregators.
- **IoT M2M messaging:** Many IoT systems themselves use messaging and notifications internally, or for interoperability with other domains. In particular, "publish and subscribe" mechanisms are being widely deployed. MQTT is an old but still-evolving protocol, which is starting to be seen as a de-facto option for many IoT use cases. It also seems to align well with another hot IT trend – so-called "serverless" computing. Numerous other IoT-related messaging and interoperability standards and communities are also emerging, such as CoAP, OneM2M, OIC and AllSeen.
- **Other messaging:** The "other" category for messaging is extremely broad. It implicitly includes email and push notifications for mobile apps, as well as a huge variety of IM and social apps such as WhatsApp, Facebook Messenger, WeChat and many others for both consumers and business. Increasingly these also include voice, video and "bot" capabilities as well as basic IM. There is a broad variety of customization models for white-listing, presence/absence, ephemerality, stickers and emoji, "filters" and "lenses" to modify images, read/write indication and so forth. Many other applications embed their own messaging features, for example for live customer-support chat, or unified communications in enterprise. The telecom industry has long tried to enter this extended messaging market, first with federated IM, and then various iterations of so-called "rich" communications RCS, or proprietary apps, with minimal results. (Disruptive Analysis is a vociferous skeptic of RCS.)

From a cloud-based IMS point of view, the most easily addressed communications types are telephony, SMS and non-telephony uses that can still use telephony as a "raw ingredient." These fit well with the broader trends around cPaaS and UCaaS, exposing communications capabilities to developers, or using them in more sophisticated "private" use cases within service providers' own initiatives. It is also likely that more IoT devices will have microphones than cameras, while displaying text is easier than images.

There is significant scope to reuse components of the traditional telephony stack in an IoT context. Numbering, conditional call-control, caller ID, voicemail, DTMF, ringing and other functions could all be repurposed, if exposed to innovative developers through suitable platforms and APIs. We may also see other capabilities exposed, such as authentication and notifications.

Non-telephony audio applications will probably be much harder for CSPs to offer, and may be a poor fit with conventional IMS or VoLTE applications. However, there may still be a need to transit a gateway from an Internet/WebRTC domain to a telecom one, especially if there is a potential need to fall back to PSTN-type models.

An interesting area that is being discussed in some quarters is the use of IMS as a more generalized platform for cloud services, not just communications sessions. It is possible that it could evolve to become a message-broking platform, especially where there are hybrid message/real-time-comms use cases. Clearly there are various approaches to creating more general IoT interaction and control platforms, but a subset of communications-related instances could fit here.

In theory, conversational IoT-related video is possible with ViLTE, but it has had almost zero deployment, nor support from device vendors or application developers. It would likely need careful work on user experience and interface design to gain any traction (who wants an unexpected video call from an unknown number?), although could conceivably be useful in regulated sectors such as emergency calling from elevators. Both WebRTC and proprietary platforms such as Vidyo are already well entrenched for video applications -- integration into native mobile apps using third-party APIs or open-source components seems commonplace. A similar argument is true for messaging, where IMS RCS competes not just against consumer mobile apps, but also enterprise- and industrial-grade messaging platforms. Given that an increasing proportion of video and message content is blended into apps and browsers, rather than "native" to the device itself, these trends seem hard to counter, unless user behavior changes substantially.

Perhaps the next frontier for IMS-based communications with IoT will emerge with virtual and augmented reality. The complexities of rendering blended video/overlay images, especially with 3D or 360-degree context seen through a headset, may need significant network smarts. For example, QoS control for ultra-low latency guarantees may be more essential, to avoid not just "glitches" in communications, but physical queasiness if perceived reality is not quite "real time" enough to fool the user's inner ear. Some researchers are proposing CSP infrastructures with mobile edge-computing (MEC) functions much closer than a distant data-center. Virtual network functions running in an MEC node a few milliseconds from the end user could be an interesting use case to watch here.

IMS & IoT INFRASTRUCTURE TRENDS: TOWARD NFV & THE CLOUD

This document is not the venue for a full analysis of adoption and architectures for NFV and other types of virtualization. However, it is certainly worth noting that a huge variety of approaches exists, with different carriers' strategies ranging from hugely ambitious and wide-reaching projects such as AT&T's eCOMP, through to many more "tactical" deployments of specific functions from the cloud, or use of third-party hosted capabilities.

IoT is right in the middle of this tectonic shift. Some SPs want to build new, scalable core networks, specifically oriented toward massive M2M-type deployment and provisioning. Others are participating in new value chains oriented around platform-as-a-service (PaaS) concepts that are "cloud-native." An array of participants want to act as MVNOs, managed service providers, eSIM brokers, IoT security specialists and much more – all of these may require network-related functions, even if they don't own the access or radio infrastructure.

In parallel with this, virtualization is also gaining traction for voice, video and messaging. Obviously Internet-based apps and services are inherently "cloudified," but enterprise-oriented and carrier-grade infrastructure (usually involving SBCs and -- for telcos -- IMS) is also moving in that direction. The industry is seeing rising deployment of VoLTE in 4G mobile networks, and various forms of cloud-based UCaaS and collaboration platforms for businesses. Communications-specific cPaaS offers for developers abound, linked to both telecom infrastructure and the Internet.

A likely scenario is that some IoT specialists will find that they need voice/messaging/video capabilities for a proportion of their customers or device fleets. They may have a series of potential value-adds for connectivity, such as security, social or ecommerce integration, with communications being added to the list. The ability to add (say) emergency calling, or VoLTE calling support could prove important – as could non-IMS communications integration with WebRTC or vendor-proprietary systems such as Slack or Microsoft's.

ALIGNMENT WITH eSIM & CELLULAR IoT NETWORKS

IoT devices connect over a wide range of network types, both directly or via some form of hub or gateway. There are many trade-offs here, notably around bandwidth requirements, cost, mobility and power consumption. Historically, many M2M products such as alarms or utility meters have used narrowband GSM connections, which have been available at relatively low cost for some years. Inside a building, most devices tend to use Wi-Fi, Bluetooth or ZigBee for short-range connection to fixed broadband. Conversely, vehicles need full mobility and high-speed connections – LTE is perfect. Numerous niches abound, like fiber connections for industrial machinery in locations with lots of wireless interference, or satellite links for remote sites.

More recently, there has been a push to extend wide-area connections to lower-end devices, possibly with battery life lasting for years. So-called LPWAN (narrowband low-power wide-area networks) span both cellular and non-cellular versions. Variants of 4G called NB-IoT and LTE-M are emerging to compete with proprietary SigFox, open-sourced LoRAWAN and various other technologies to connect sensors, wearables and other "constrained" devices.

The various connection types map to communications services opportunities in different ways:

- Older 2G GSM connections generally support SMS, but only a proportion of modules support (circuit) telephony.
- LTE and Wi-Fi connections can support voice and video, especially where the end device is equipped with SIM / eSIM for identity and authentication.
- LPWAN connections cannot support voice or video, and generally not SMS either. However, they could still provide data "triggers" for calls emanating from a central platform, perhaps with MQTT- or OneM2M-based platforms. LTE-M will likely be better, as it's more of a "mid-power"/mid-bandwidth technology than the others.
- Bluetooth can be used for voice calls or SMS, where the device is being used as a companion to another cellular product such as smartphone.

It is also worth considering the role of eSIM here. The ability to embed tiny soldered-in SIMs into IoT devices, and remotely provision them with a service provider's credentials, should take some friction out of getting cellular devices online. Users will not have to source and install a traditional SIM card, while manufacturers will not have to worry about dust/vibration-prone holders.

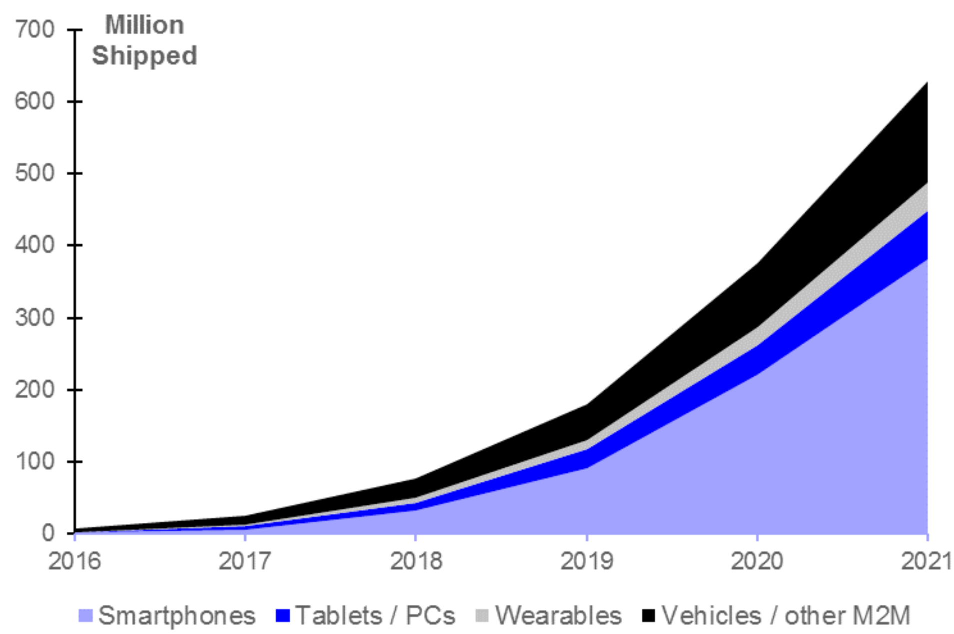
Owners of eSIM-equipped device "fleets" should be able to switch operators for devices already in the field. This should help various cellular IoT technologies diffuse faster into some product categories such as wearables or machinery, particularly for higher-end devices that can bear the extra costs of the radio and eSIM. The technology should also lead to new players in the IoT network-services space, with system integrators and manufacturers acting almost like MVNOs – and, perhaps, running their own communications capabilities as well.

However, the practicalities of eSIM are maturing only slowly. Only a handful of products have shipped so far, and there are still many wrinkles in logistics and user experience to iron out. Disruptive Analysis has recently published a full analysis of the market, and predicts that the eSIM market will show small volumes shipped in 2016 and 2017, with 25 million forecast sales in 2017, especially including connected cars.

From 2018, as experience grows, back-end system adoption becomes broader, and value-chain issues will get fixed. Driven by high-value/high-margin devices, there should be 180m eSIM sales by 2019. Growth then really kicks in around 2020/2021, coinciding with 5G launches, low-power IoT uptake and scale/experience effects pushing eSIM down to mid-range devices. Total eSIM shipments in 2021 are forecast at 629M – maybe about 10 percent of overall SIM sales (which were 5.3B in 2015).

Overall, eSIM coupled with LTE-M (and maybe NB-IoT) should help increase the number of voice-capable IoT endpoints substantially in the medium term.

Figure 4: eSIM shipments, by device category, 2016-2021



Source: Disruptive Analysis

CONCLUSIONS & RECOMMENDATIONS

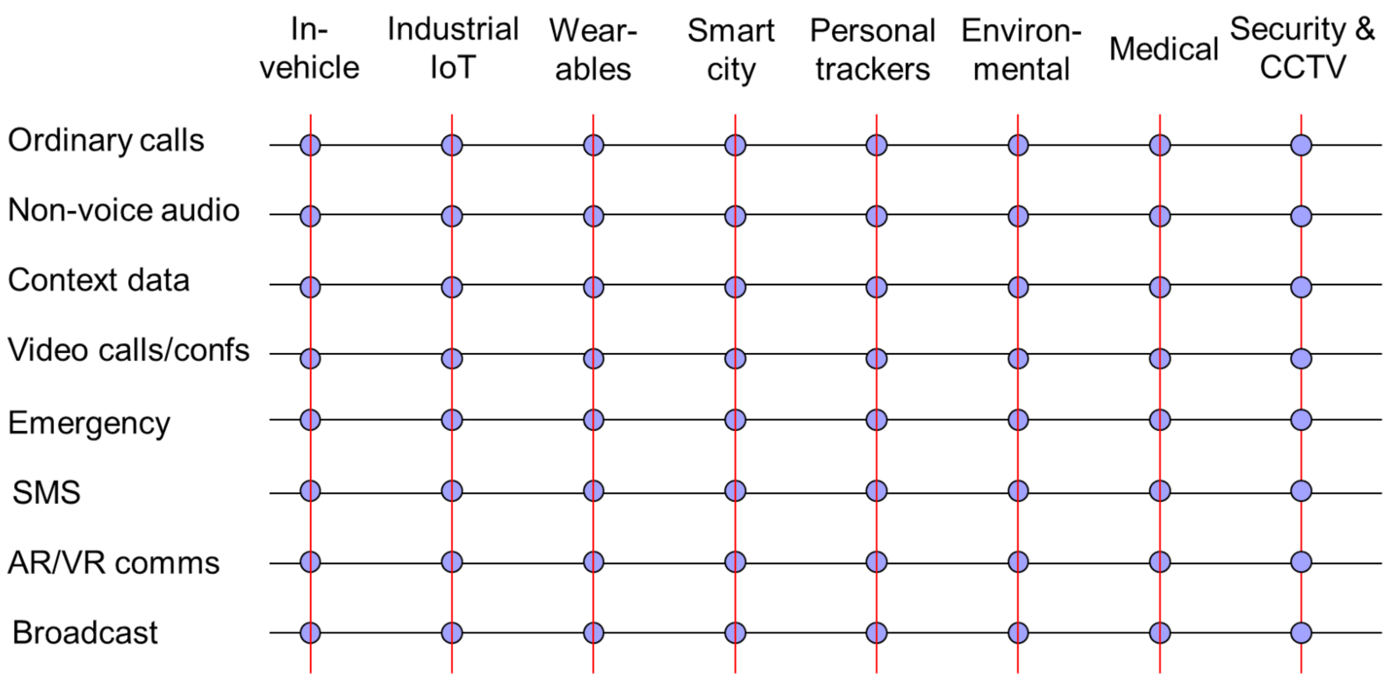
Many IoT implementations will involve communication with one or more humans, directly or indirectly. Some instances will be relatively mundane – for example, devices sending notifications or text-based messages. But a subset will involve real-time communications; that is, voice and video. This could just be an IoT device with an in-built dialler, or an event triggering a traditional phone call, but more interesting and unconventional variants are appearing as well. We will also see hybrids emerge – for instance, message-based notifications (M2M, SIP or HTTP), combined with voice/video calls or streaming from a machine or robot.

A proportion of such communication events are likely to be delivered by CSPs, either end-to-end in the context of a managed IoT service or platform, or through interoperability platforms for origination, termination or some form of media processing. Given the overall trend toward IP communication, this means a role for IMS-based calls in the IoT universe, at least some of the time.

Furthermore, the parallel trend to NFV and virtualization suggests that cloud-based IMS has a role to play here – especially for "greenfield" IoT players such as specialized CSPs, MVNOs or PaaS providers. Historically, IMS has been extremely complex and costly – and thus often a reluctant "forced purchase" for CFOs. Cloud-based, more-scalable variants can improve the economics in terms of both upfront costs and usage-based fees, which broadens the scope. The addition of external APIs and additional non-session capabilities can also potentially move IMS from being just a "Telephony v1.1" platform to something more versatile.

Complex, use-case-specific value chains for IoT will bring in a much wider array of new "service providers" than in the past. Smaller operators without ongoing large-scale NFV transformation projects, or separate IoT divisions of telcos deploying dedicated core networks for M2M, also appear to be potential users of cloud-based infrastructure.

Figure 5: Myriad potential intersections between IoT & communications



Source: Disruptive Analysis

In Disruptive Analysis's view, certain IoT/communications hybrids are more relevant here than others. Variations on telephony and traditional "calls" seem most addressable, as they fit better with readily available IMS VoIP/VoLTE application servers. That is particularly true where such calls may fit with E164 numbering, or need emergency call and lawful-intercept capabilities, or other enablers such as SIM/eSIM authentication – for example, on medical wearables.

There is a broad diversity of applications for enhanced, embedded and contextual telephony, ranging from simple "call an engineer" buttons on complex machines, to in-vehicle voice, or integration with wearables. Using a flexible platform (cloud-based or otherwise) will enable better call-handling, UX and other tweaks – for example, the user behavior of answering a call on a smart-watch could be very different to that on a phone. Combinations of SMS and voice/video can also make sense for such applications.

Perhaps the “killer use case” is around preventative maintenance. If malfunctioning systems can be pre-emptively diagnosed and fixed before they break, rather than afterwards, there is a huge benefit. Such interventions will almost always need human communications – promptly, and with detailed contextual information. Health care is another promising sector, linking doctors, patients and the networked equipment in rapid and secure fashion.

OUTLOOK: TOWARD A BROADER TELCOFUTURISM UNIVERSE

Over time, we should expect to see IoT/communications services become more sophisticated and linked with other technology trends as well.

Clearly, the IoT itself is going to become much more diverse in terms of devices, with more scope for communications innovation. Robots, drones and autonomous vehicles are among the more interesting areas (especially healthcare robots), although any visit to an IT tradeshow these days will see a bewildering array of other innovations from fascinating to absurd. (Maybe “talking to one’s plants” will take on a whole new meaning, if the Internet-connected pots seen at CES2016 embrace voice in future.)

But beyond the devices, there are other developments that play a role here – AI and machine learning are particularly interesting, given the rise of voice-based products such as Amazon’s Echo, or back-end tools such as IBM’s Watson for interpretation. While today these are mostly Internet-based, Disruptive Analysis hopes that the telecom industry will also play a part. Broader use of cognitive and contextual communications should increase where human voice/video functions are linked to analytical observations, sensor input or machine-learned preferences.

Other “telco-futurism” crossovers could also arise here, for instance with communications events stored securely and unchangeably on private blockchains, for multi-party lookups or validation. There are also interesting opportunities around 3D/VR/AR communications. Depending on how images and overlays are transmitted and reconstructed (e.g. using 3D wireframes, eye tracking and processed 2D images and overlays), there may be a need for QoS and local edge-computing resources that go beyond normal connectivity, and which could be a use case for IMS-type infrastructure and service providers.

Overall, the number and depth of intersections between IoT and voice/video/messaging are likely to increase sharply in the next two to five years. Telecom operators and the wider CSP community will be assessing many different approaches and platforms to serve this need. If virtualized and cloud-based IMS continues to become more cost-effective, flexible and accessible to developers, then it can play various roles in the coming hybrid IoT/communications opportunity.

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 telecom/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/market evolution or labyrinthine business relationships. Currently, the company is focusing on mobile broadband, operator business models, the Future of Voice, smartphones, Internet/operator ecosystems and the role of governments in next-generation networks.

Disruptive Analysis attempts to predict -- and validate -- the future direction and profit potential of technology markets, based on consideration of many more "angles" than is typical among industry analysts. It takes into account new products and technologies, changing distribution channels, customer trends, investor sentiment and macroeconomic status. Where appropriate, it takes a contrarian stance rather than support consensus or industry momentum. Disruptive Analysis's motto is "Don't Assume."

For more detail on Disruptive Analysis publications and consulting/advisory services, please contact information@disruptive-analysis.com. For details about Future of Voice & Contextual Communications workshops and publications, please see www.deanbubley.com.

Website: www.disruptive-analysis.com
Twitter: @disruptivedean

Blog: disruptivewireless.blogspot.com
LinkedIn: https://www.linkedin.com/in/deanbubley

Disclaimer: The document has been written by Disruptive Analysis, an independent research and futurist company with long heritage in analysing networks, "the future of voice" and service provider business models. It has been commissioned by Metaswitch, a vendor of software-based infrastructure for CSPs. It is aimed at C-level executives, strategy, planning and technical staff at CSPs and IoT solution providers.

Mentions of companies in this document are intended as illustrations of market evolution and are not intended as endorsements or product/service recommendations.

Every reasonable effort has been made to verify research undertaken during the work on this document. Findings, conclusions and recommendations are based on information gathered in good faith from both primary and secondary sources, whose accuracy it is not always possible to guarantee. Disruptive Analysis Ltd. disclaims all warranties as to the accuracy, completeness or adequacy of such information. As such no liability whatever can be accepted for actions taken based on any information that may subsequently prove to be incorrect. The opinions expressed here are subject to change without notice.