The Definitive Guide

THE INTERNET OF THINGS FOR BUSINESS

aeris

BY SYED ZAEEM HOSAIN | Chief Technology Officer, Aeris

FOREWORD

In more than 10 years serving the machine-to-machine (M2M) marketplace, I have seen many great ideas and products. First, purpose-driven innovation designed to solve particular problems in manufacturing, transportation, energy, and other industries. Next, technology solutions to regulatory imperatives for everything from the environment to the financial sector. And today, the brand new services offerings and cross-sector, big-data analytics that are transforming M2M into the Internet of Things (IoT).

Analysts are divided on the exact potential of this transformation—but all agree, the impact will be huge. Some predict an economic impact of the IoT equivalent to the US GDP. Others measure the number of connected devices—which will outnumber the people on the planet many times during the next five years. Still others focus on specific societal impacts: lives saved, people fed, reduced hours in transit, or tons of carbon emissions, children educated…

Unfortunately, I have found over the years that a lack of information and understanding about exactly what it takes to bring a connected solution to market has been the number one hindrance to realizing those world-changing ideas. Imagine watching life-saving, people-feeding, environment-saving, child-educating ideas die on the vine for lack of knowledge. Syed Zaeem Hosain wrote this book to address this knowledge gap—to help innovators through the process of bringing their concepts to reality—to the betterment of our world.

Inside, you'll find the first comprehensive primer to outline exactly what it takes to bring your idea to life. It will steer you around common pitfalls and provide the guidance you need to streamline your process. I think, if you're in the business of IoT or looking at starting up a deployment, this book is for you.

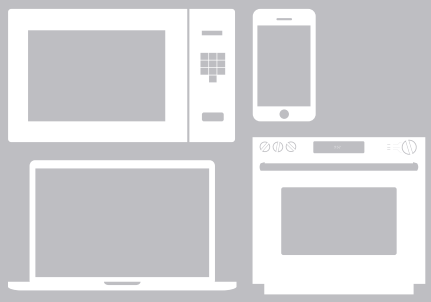
Good reading and good luck.

Stefan Lindvall
CEO, MultiTech Systems
CEO, Connected Development

THE DEFINITIVE GUIDE:
THE INTERNET OF THINGS FOR BUSINESS

CONTENTS

CONTENTS *cont.*



WHAT IS THE INTERNET OF THINGS?

Chapter 1

Back to Table of Contents

1.1 » New Markets, Increased Efficiency

The Internet has become such an important part of daily life that Twitter and Facebook seemed to help topple dictators, and "Google" is a common verb. Instant connectivity has completely changed much of society. Now a new revolution is upon us: The Internet of Things (IoT).

The depth and breadth of IoT connectivity will create new businesses, provide new markets for existing businesses, and improve operational efficiencies. Think tank Machina Research says the IoT market will swell to $4.3 trillion USD by 2020. Gartner predicts that the number of IoT devices will grow to 26 billion units by 2020 in the US.

Vast potential lies within connecting enterprise-scale devices. IoT and machine-to-machine (M2M) communications increase operational efficiency by giving businesses visibility into the details of their operations in ways that could not have been measured before. In **Cisco**'s "Internet of Everything" 2013 report, the highest percentage (27%) of value in future IoT revenue will be in manufacturing. "Smart" IoT/M2M-enabled factories alone could reap $1.95 trillion in profits between now and 2022, thanks to sensors incorporated into machines and processes. Likewise, **Accenture**'s 2014 report, "Driving Unconventional Growth Through the Industrial Internet of Things," finds that manufacturers could boost their efficiency by 30% using IoT. As Accenture notes, manufacturing has already been an early adopter of these technologies.

While these reports point to the growth of IoT in industry and its impact on GDP, their analysis finds that government and industry need to build out the infrastructure and work on "political enabling factors" to push it forward. Accenture further states that 73% of businesses have no IoT plans, yet 87% of business leaders say IoT will be a net creator of jobs. So you could conclude that while many businesses understand the importance and potential of IoT, not all businesses are actively deploying their own projects yet.

1.2 » The Guide to IoT for Business

In this book, we'll focus on how the burgeoning IoT/M2M ecosystem can be used by business. In addition to providing real-time information on devices in the field, IoT works in the other direction too: it will let companies control devices from a central location. This can provide everything from marketing intelligence to improved preventative maintenance. Companies can use IoT for applications as diverse as helping medical professionals care for more patients at the same time or giving retailers the ability to customize advertising to a single individual.

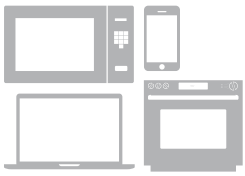
MANUFACTURERS COULD BOOST THEIR EFFICIENCY BY 30% USING IOT.

Back to Table of Contents

To get started with IoT and M2M for your business, you'll need a basic understanding of what makes it all work. You don't need to be an engineer or a data scientist, but it's useful to have a grounding in the concepts of how IoT systems are connected, how they communicate, and how this will impact your organization. We'll present an overview on networking and the Internet and describe the Internet of Things in more detail. In doing so, we'll cover these broad topics:

- The technology that connects the Internet of Things.
- How wireless devices are networked and locate themselves.
- Different types of sensors, how they work, and what they do.
- An overview of security technologies used to protect IoT data.
- How to scale up an IoT project to immense proportions.
- Using Big Data analytics to gain insight from the IoT ecosystem.
- Advice for managing the lifecycle of an IoT deployment.

All of these aspects of the Internet of Things will be addressed from an enterprise point of view for those running small to large businesses. While IoT will make an impact on many everyday consumers' lives, we feel that the end-user world of smartphones, fitness trackers, and connected toasters has been sufficiently discussed elsewhere. We want to look behind the scenes into how these devices are run and managed, where the data they collect goes, and how it's used. If you're in the business of IoT or looking to start up a deployment, this guide is for you.



1.3 » Definition of the Internet of Things

The Internet of Things, also called machine-to-machine communications, envisions a world where both ordinary and exotic devices are connected wirelessly to the Internet and to each other. This means devices that do not already have a network connection may have one added in the future, when it is logical to do so.

One of the most basic uses of the IoT is to connect devices to the Internet so they can report their own status or their local environment. For example, an IoT/M2M device could be a temperature gauge, a location sensor, a device measuring humidity, or an integrated circuit that checks vibration. One or all of these sensors could then be attached to manufacturing machinery, and the data transmitted would help a business track the machine's operations. This data could track required maintenance, improve production efficiencies, reduce downtime, increase safety, and more. Plus, IoT/M2M devices may provide information on the ambient environment of the manufacturing space, such as the temperature, pollution, and other conditions near the machinery, which may be particularly relevant for remote installations.

Most IoT projects are motivated by a need to reduce operating costs or increase revenue. Occasionally, legislation compels companies to deploy M2M applications that support the new law's data needs. Mobility is an obvious factor driving cellular adoption in markets like transportation. Desire for competitive features may inspire IoT applications in consumer high-tech. But whatever the specific purpose, connected IoT/M2M devices can give your business the data needed to streamline workflows, predict necessary maintenance, analyze usage patterns, automate manufacturing, and more.

Back to Table of Contents

1.4 » Examples of IoT in Use Today

As new as the Internet of Things may seem, many network-connected devices are already in use all around us. You've probably heard of connected homes or the smart grid—these are just a few of the IoT/M2M systems aimed at both everyday consumers and large-scale enterprises.

Consumer IoT Applications

While the focus of this book is on business uses for IoT technology, seeing how it applies to consumer devices is relevant for a sense of scale and direct application in today's world. These kinds of IoT devices let individuals control their own network-connected devices from their smartphones and watches or get information about their status from a webpage.

Even seemingly old-school industries like **restaurants** are finding ways to be more efficient and improve the bottom line with networked devices. Restaurants using IoT-connected tablets attached to dining tables are letting guests place orders and pay bills. Early studies have shown these systems can boost appetizer sales by 20% and dessert sales by 30%.

A few of the most popular consumer IoT devices already in use are:

- Nest (now owned by Google) markets a Wi-Fi-connected thermostat that programs itself to adjust to your home habits. The thermostat system can help save energy, and it can be integrated with automated IoT lighting, security, and other tools to create the long-imagined connected home.

- OnStar, by General Motors, is one of the first connected car services launched in the US. Using cellular technology, OnStar provides subscribers with turn-by-turn navigation, hands-free calling, and other remote, in-vehicle services. Since OnStar's creation, most major auto manufacturers around the world have added connected-car features.

- FitBit is one of the best-known Internet-connected fitness trackers on the market, along with the Apple Watch (which has more functions than simply counting steps or heartbeats). These devices are part of a "quantified self" movement that started in the mid-2000s to gain greater personal understanding through data and technology.

Back to Table of Contents

1.4 » Examples of IoT in Use Today

Enterprise IoT Applications

To date, most industrial uses of IoT have been for preventive maintenance. These applications detect when a machine gives off variations in vibration, temperature, speed, or any other metric to signal that they might need maintenance.

But using IoT for preventative maintenance is just a start. This doesn't tap into the ability of network-connected devices to talk with each other, thus letting them work together. For example, a business could use a central monitoring hub or even an engineer with a smartphone to reach out to the machine and make changes on the device or push out new instructions. More and more enterprises are realizing that machine-to-machine communications can create greater efficiencies and reduce production costs far beyond simple maintenance IoT/M2M systems.

IOT DEVICES REDUCED EMERGENCY HOSPITAL ADMISSIONS BY 20% AND EVEN MORTALITY RATES BY 45%.

Currently, the 500,000 long-haul trucks on the Aeris network use customized IoT/M2M **fleet telematics** solutions to enhance delivery performance, increase service radius, reduce payroll, and cut down on idle truck times. In the **utilities** market, "smart" meters delivering automated readings over IoT/M2M networks like Aeris provides could save an estimated $20 billion worldwide by 2020, versus manual meter readings. These automated meter readings are more accurate as well, resulting in fewer billing inquiries and improved customer care.

Some specific IoT/M2M enterprise applications currently deployed in the field include:

- **Novariant** makes an autopilot system called GeoSteer for heavy agricultural machines like tractors and combines. With Aeris' IoT/M2M network providing an Internet connection, farmers can use GeoSteer to keep their machinery on course within inches of accuracy. This is crucial for agribusiness so no farm acreage is wasted during seeding or fertilizing. This application also helps the farm operator access remote portions of large fields.

- **GTX Corp.** provides GPS tracking in personal, wearable devices, powered by the Aeris IoT/M2M network. Its flagship product is a shoe insert equipped with an IoT-enabled GPS chip that can be worn comfortably and discreetly by anyone who might wander out of sight—such as dementia patients. Caregivers for those with Alzheimer's disease have found the GTX shoe device to be a lifesaver for keeping track of people.

- Singapore launched a "**Smart Nation**" program in 2014 with the goal of integrating data and advanced technology to drive economic growth and improve quality of life. Initiatives have encompassed everything from citizens being able to call for a bus from their smartphones to smart homes for the elderly.

- **Telehealth** systems including home monitors, personal alarms, fall detectors, automated medication dispensers, and more, underwent extensive trials in the UK's National Health Service from 2008 to 2013. These IoT devices reduced emergency hospital admissions by 20% and even mortality rates by 45%.

- **Met One Instruments** improves on air-quality monitoring by creating small, effective low-cost IoT/M2M tools that can be used around a neighborhood. Traditionally, the Environmental Protection Agency (EPA) had to send technicians out to service air quality monitors and record data. But Met One's products, using Neo's cellular connectivity by Aeris, are so effective, the EPA bought several units and is considering deploying them throughout the US.

Back to Table of Contents

1.5 » IoT Apps of the Future

Since IoT/M2M applications are growing at such a rapid pace, it's impossible to know exactly what the future holds. But smart grid and connected home technologies are sure to make inroads into people's lives, along with more wearable tech. Perhaps we'll see these businesses take off as well:

- **Fully Autonomous Vehicles:** Google already has driverless vehicles, although they face regulation challenges in most US states. While studies are still being conducted, some suggest that driverless vehicles may be safer than existing human-driven cars. The potential also exists for driverless vehicles to be used for long-haul trucking or automated deliveries and for driverless cars to work seamlessly together with central traffic control systems to relieve traffic congestion.

- **Supply Chain Analytics:** Successful businesses are already seeing value in the data their IoT/M2M systems are collecting about their own manufacturing processes. Real-time analysis of this information along the supply chain is poised to become one of the most effective uses of IoT/M2M technology by streamlining manufacturing, bringing costs down, tracking goods from point to point, and improving customer satisfaction. Organizations that are on top of their supply chains will have a competitive advantage.

- **Micro-Targeted Advertising:** An IoT beacon can deliver a customized restaurant coupon to your smartphone, but maybe next, your favorite musician's song may automatically download to your car's sound system via the in-vehicle connection. Businesses can use Big Data and IoT/M2M technology to target their messages, advertising, and products precisely to the customers who are most likely to want them.

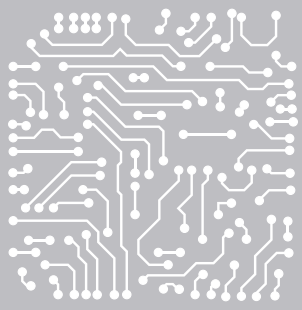
These are just a few examples of where the IoT is heading. According to analyst reports, billions of connected "things" are already in use. More are connecting every day and for various purposes. If your organization has an IoT/M2M deployment, this will generate increasing amounts of data that will need to be crunched. New and better analytics systems will have to be developed, not to mention cloud storage to deal with all this data.

Current devices must be built to last, whether it's a connected car that can receive over-the-air updates or a remote patient monitoring device compatible with the next-generation of cellular networks. Once equipment is installed and in the field, employees and customers will be relying on service for years at different levels of importance. Business needs to plan for the future from the hardware up.

So how are you approaching your IoT/M2M project? How will it reduce your business' operating costs or increase your revenues? Are you planning analytics to take advantage of the Big Data these connected devices will generate? Because, starting now, one thing is clear—the number of connected devices, and the benefits from the data they gather, will both be enormous.

As you begin to consider how your company fits into the IoT's future, let's look under the hood at some of the technology that makes IoT/M2M connections happen.

Back to Table of Contents

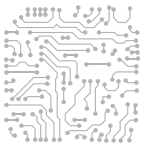


IOT NETWORK TECHNOLOGY

Chapter 2

Back to Table of Contents

To understand how the Internet of Things and machine-to-machine communications work, you'll need a basic overview of the technology that runs the Internet itself. While technology is always evolving, certain principles are common to how networking functions. What changes more frequently are the tools and protocols used to access the network, such as modems, cellular radios, transmitters, and more.



2.1 » Basic Internet Concepts

IoT/M2M connectivity relies upon these essential concepts and devices that make the Internet function:

IP: Traffic on the Internet uses the Internet protocol (IP) to transmit data. This communications protocol has a routing function ideal for Internet connectivity. IP can route data packets across the Internet from a source host to a destination IP address. Everything in such a network will have an IP address, a unique numerical label. The computers and printers in your office generally have private, local-area network IP addresses, while websites such as Aeris.com have public IP addresses.

Packet: Data travels across the Internet in packets. Each packet has both a source and destination IP address, but many packets may be needed to make up one complete "item." For example, one email message can be comprised of many different packets that, when assembled by the recipient's email program, make a finished piece of mail. A webpage retrieved by your browser is also comprised of multiple packets.

To the right is a sample data packet. The innermost layer is shown at the top. This is an Ethernet packet as it is the protocol that carries IP traffic. You can see the source and IP headers. Even though it looks complicated, you can read the actual data in some cases. For example, if it was the webpage of The Washington Post, you would be able to read the text part of page.

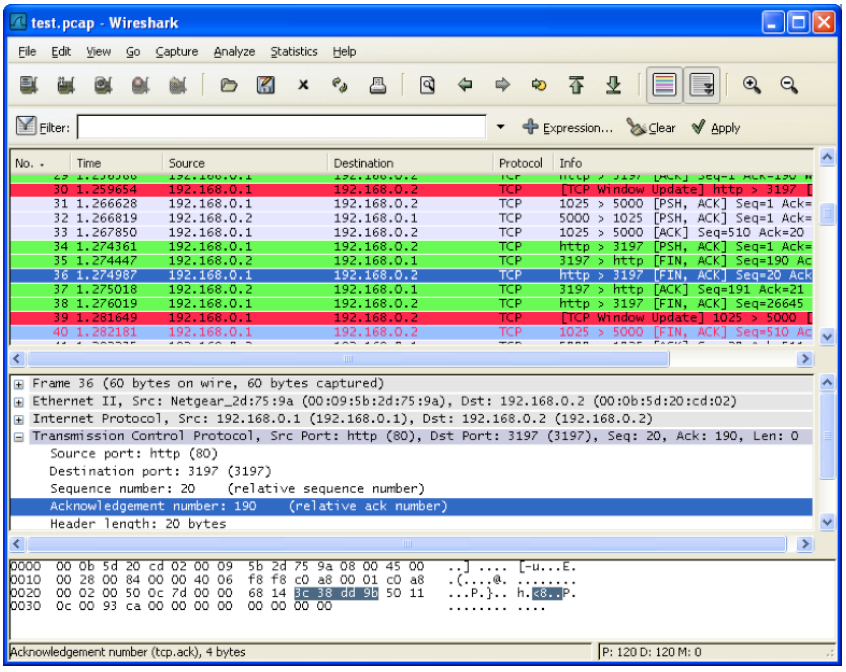


Figure 1. Data Packets

Back to Table of Contents

2.1 » Basic Internet Concepts

Router: A router connects one network to another. For example, your home or office wireless router connects the internal network in your home or office to the Internet. Also, your Internet service provider connects to other providers and broadband Internet backbones using routers.

Modem: A modem is a truncated term meaning "modulator-demodulator," and the device modulates signals to encode digital information and then demodulates the information. For example, a voice-over-IP (VoIP) device connects to a wireless cellular network using a modem. Wireless broadband modems are a popular way for smartphone and laptop users to get Internet connections. Early cellular modems used the 2G standards, but most have moved to the faster 3G technologies. The newest standard is 4G, also called LTE, which is becoming available around the world.

Speed: Internet speed is measured in megabits per second (1,000,000 bits per second). For example, Netflix HD video requires 5.0 megabits per second for viewing, although Netflix will work at speeds as slow as 0.5 Mbps. South Korea leads the world with the fastest average Internet connection speeds of 25.3 Mbps, and Hong Kong ranks second 16.3 Mbps, while the United States has an average best speed of 11.5 Mbps **as of 2014**.

SOUTH KOREA LEADS THE WORLD WITH THE FASTEST AVERAGE INTERNET CONNECTION SPEEDS OF 25.3 MBPS.

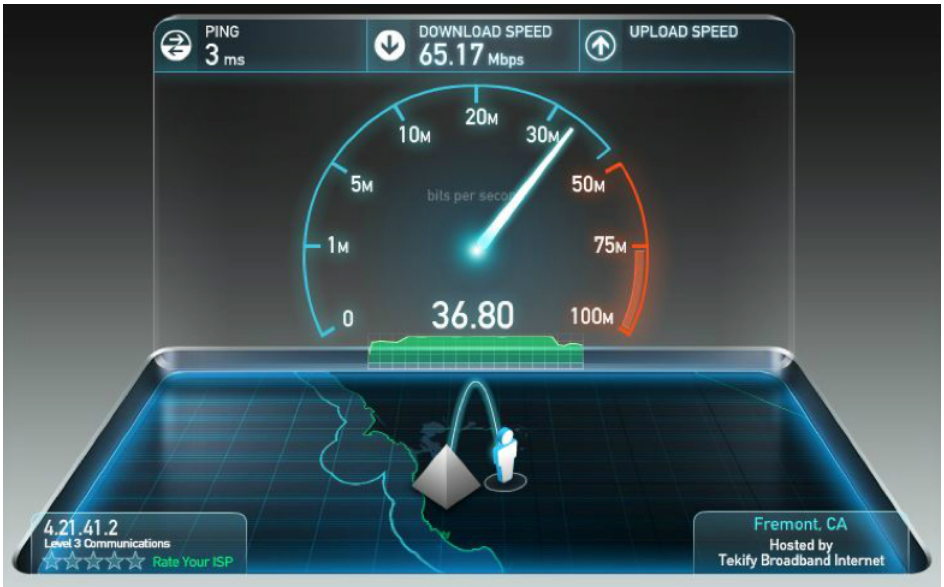
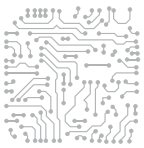


Figure 2. Speed Test

This graphic is from the website SpeedTest.net, which is a common site for checking how fast your Internet connection is operating. In this example, the download speed is 65.17Mbps. The ping is the amount of delay in milliseconds going from your computer to the website where you are going (in this case, SpeedTest.net).



2.2 » Choice of Connectivity

The Internet itself can be accessed in many ways, depending on your device and application. There are pros and cons to each form of connectivity technology, particularly when implementing a large IoT/M2M project.

Internet Service Providers

An Internet service provider (ISP) connects offices and homes to the Internet by taking that network traffic and forwarding it to other networks until it gets to the desired destination. An ISP could be, for example, Telstra in Australia. But it doesn't stop there, because an ISP essentially has another ISP. For example, while Telstra runs the largest Internet network in Australia, it still has to connect to other networks within the country and around the world through both cross-country fiber optic cables and submarine cables. ISPs, such as Telstra, connect to Tier 2 or Tier 3 networks and up to Tier 1 networks that form the Internet backbone. These top-level networks become the principle routes for data transmission around the world.

THE FUTURE HOLDS PROMISE FOR MORE VARIETIES OF WIRELESS DATA TECHNOLOGIES SUCH AS WIDER ADOPTION OF 4G LTE.

Wireless operators like Aeris connect IoT/M2M deployments to the Internet in a similar fashion. A wireless operator has a cellular network that uses fixed transceivers or cell towers instead of wires to transmit signals from the cellular devices into the network. Much like ISPs using other ISPs, wireless operators can also connect to Tier 1, Tier 2, or Tier 3 networks. This is how they deliver traffic on the wireless network when a mobile device requests data.

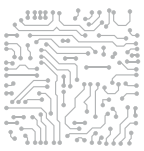
Wired and Wireless IoT Connections

A home, office, or an IoT/M2M-networked device can be connected to the Internet either via a wired or wireless connection. If the connection is wired, it's plugged directly into an Internet router, and the device is needs to remain stationary. A device with a wireless connection can have a cellular modem (described above), a Wi-Fi router, or other connectivity technology, and among other things, this lets the device be mobile.

Wired connectivity was common in the early days of machine-to-machine systems. For example, many factories installed pre-wired systems for supervisory control and data acquisition. For business and residential security systems, alarm panels could use telephone circuits to communicate events—like a burglary or fire—to monitoring stations. However, connectivity was dependent on where an ISP's lines could extend to, and setup could be complicated. These early applications tended to be purpose-built, meaning each industry and company developed its own devices and software systems from scratch.

The 1990s saw a move towards using wireless radio technologies. Ademco Corporation, a leader in intrusion and fire detection systems, began to build out a private radio network to address this need. In 1995, Siemens introduced the first cellular radio module for data transmission applications. Very shortly afterwards, Aeris introduced its MicroBurst® data services using the control channels of the Analog AMPS cellular system, and Ademco became the first major customer to deploy units using this transport.

Back to Table of Contents



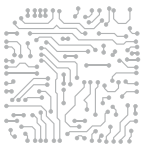
2.2 » Choice of Connectivity

These new technologies further broke machines free from wires, and more IoT/M2M functions were possible in different industries and even for consumer products. OnStar was one of the first connected-car systems in 1995, offering a mix of service and entertainment options. Fleet and container tracking solutions similarly made use of mobile telematics for the trucking and railroad transportation industries. In addition to being mobile, connectivity could extend to more remote locations that wired networks didn't allow.

By 2008, changes in cellular technology introduced digital cellular networks with features such as Short Message Service (SMS) and General Packet Radio Services (GPRS). However, there are two competing types of digital cellular, CDMA and GSM, and different industries sided with each one. The automotive and trucking industries mostly chose CDMA devices, while the alarm and security industries generally picked GSM. By 2017, the largest American 2G GSM operator will sunset its GSM network, so alarm and security systems still using this service have to upgrade or switch cellular systems.

The future holds promise for more varieties of wireless data technologies such as wider adoption of 4G LTE. Also, short-range data transport methods, such as Bluetooth®, ZigBee®, and 6LowPAN, may augment long-range cellular in some applications. More details about IoT cellular connectivity will be discussed in later chapters.

Back to Table of Contents



2.3 » ICANN and IP Addresses

THE WORLD BEGAN TO RUN OUT OF TOP-LEVEL IPV4 ADDRESSES IN 2011.

The Internet Corporation for Assigned Names and Numbers (ICANN) manages top-level domain name assignment and delegates the assignment of lower-level domains so no two domains get assigned the same address. ICANN works with various regional Internet registries—for example, the RIPE Network Coordination Centre is responsible for handing out IP address in Europe, the Middle East, and parts of Asia, while LACNIC is responsible for Latin America. These regional groups assign IP addresses to different countries. Coordination is important because, among other things, the world began to run out of top-level IPv4 addresses in 2011.

Too Many Internet Devices for IPv4

Due to the explosion in the number of websites, mobile devices, and always-on IP connections (the latter of which is crucial to future IoT/M2M deployments), the Internet governing bodies realized that the IPv4 IP address space would not be sufficient over the long term.

Luckily, the shortage noted in 2011 has not had a serious impact on many people yet because of Network Address Translation (NAT). This allows a router to share the same public IP address, or set of public addresses, for all the traffic generated by systems on the internal network. Because of NAT, many systems share a common IP address for external Internet access.

But the long-term solution for accommodating the billions of devices constantly being added to the Internet, especially with IoT/M2M applications, is upgrading IP address space to a larger number range. Currently the vast majority of systems use IPv4 addresses like:

101.10.101.10

This is a 32-bit number comprised of four 8-bit numbers. There are theoretically 255*255*255*255 or approximately 4.2 billion of these numbers available. In actual practice, there are fewer IPv4 addresses because of the groupings into IP address classes. Many addresses have special uses, like 192.nnn.nnn.nnn for internal networks, and some are reserved for the military and other purposes.

The World Is Moving to IPv6

The problem of not having enough IPv4 address numbers will be resolved when the world moves to IPv6. In IPv6, the address space has been expanded to 128 bits (from the 32 bits used in IPv4). This allows 2 to the power 128 (or approximately 3.4 x 10 to the power 38) IPv6 addresses.

Although not yet fully deployed across the Internet, IPv6 networks are already in use by many large corporations and websites. For example, Google and Facebook have provided access to their systems in IPv6 networks.

Ultimately, every device and router will use IPv6 addresses to access the new public Internet. In the interim, gateway systems will provide address translation functions—allowing older IPv4 systems to access future IPv6 networks.

Back to Table of Contents

CELLULAR CONNECTIVITY AND LOCATION

Chapter 3

Back to Table of Contents

In many Internet of Things and machine-to-machine applications, knowledge of the physical location of the remote device as it performs its tasks is an important requirement, particularly if that application's behavior and function depends on the location of the device.

Various mechanisms can provide this physical location with varying degrees of accuracy—the specific accuracy needed depends on the particular application function that uses the location. In applications where the device moves its physical location as part of its normal tasks, cellular technologies are commonly used for data transmissions.

This chapter briefly, and generally, describes the cellular technologies used in IoT/M2M applications and the methods used to determine device location for the applications.

3.1 » Types of Cellular Technologies

This section provides an overview of the cellular technologies available to IoT/M2M devices and applications for long-range data transmissions. These cellular technologies are changing and will continue to change over time. You should assume that new cellular technologies will completely replace existing deployed technologies in time and plan the device and application lifecycles accordingly.

Brief History of Cellular

Cellular service has evolved over time. Often, a fairly major change in the technology rendered a previous technology incompatible and necessitated a replacement of the radios and handset, along with changes in the network to support the new radios.

In the cellular industry, these major changes are loosely termed "generations" to distinguish and summarize their technology, the protocols used, the network changes, and the commercial deployment phases.

Analog Cellular

The first cellular service was an analog cellular system[1] later termed First Generation (1G). In North and South America, this was the Advanced Mobile Phone Systems (AMPS). It was deployed in the US in the early 1980s and was eventually shut down in February 2008.

AMPS used radio frequencies (spectrum) distinct from other wireless services. In particular, the technology used relatively low-power transmissions, which restricted the distance of the radio signals, to reach a tower (also called a base station) where the voice call could be sent into the landline telephone system.

This allowed re-use of the radio channels beyond a particular distance from a tower—each tower received and transmitted only to the cellular radio devices within that range. Grouped into cells (hence, the term "cellular") like a beehive, the tower radio did not communicate with devices outside its cell. Cellular devices communicating in remote cells could use the same radio channels (i.e., hence "re-use" the frequencies) without interfering with calls in the closer cell.

[1]Although analog cellular is no longer used, it is described here for completeness.

Back to Table of Contents

3.1 » Types of Cellular Technologies

In the US, the spectrum used for AMPS was at 850 MHz (termed "cellular") that was grouped into two bands (called "A" and "B")—thus, each market had two cellular telephone service providers that customers could select from.

The A and B bands in each market were subdivided into 30 kHz analog channels. During a voice call, a channel at the tower was dedicated to that call, to transmit and receive from a cellular telephone (also called a "cellphone," "mobile," or "handset"). The voice communication used the entire channel for the call. As can be appreciated, this was a very inefficient use of that available spectrum.

When more and more cellular users began using AMPS, it became clear that the available channels could not support the business requirements of the operators who provided the service. Improvements were needed.

WHEN MORE AND MORE CELLULAR USERS BEGAN USING AMPS, IT BECAME CLEAR THAT THE AVAILABLE CHANNELS COULD NOT SUPPORT THE BUSINESS REQUIREMENTS OF THE OPERATORS.

Thus, radio technologists began to explore ways to use available wireless spectrum more efficiently. The first improvement used digital encoding protocols for the communication rather than analog. Three competing digital systems came into existence: ANSI-136 TDMA, ANSI-95 CDMA, and GSM. Since this was a major change to cellular technology, these new systems (and the additional data transmission protocols—see below) were termed Second Generation (2G) cellular.

Eventually, AMPS and other analog cellular services were shuttered in most parts of the world (in the US, this was the "AMPS Sunset" in February 2008).

ANSI-136 TDMA

To maintain backwards compatibility with AMPS in the early deployments, technologists in the US used a mechanism to slice each AMPS radio channel in time. When speaking into a cellphone, the human voice is converted from the analog electrical signals generated by the microphone into digital bits using an Analog to Digital Converter (ADC). To listen to the received voice from the tower, digital bits are converted into an analog electrical signal using a Digital to Analog Converter (DAC) and then amplified to drive a speaker in the cellphone.

In ANSI-136 TDMA, each voice call was allocated one-third of the time (the "slot") that the channel was active for the transmission of the digitized voice bits. The transmissions were decoded at the towers into multiple voice paths sent into the landline telephone system to their destinations. Hence the general term for the protocol: Time Domain Multiple Access (TDMA). Humans are unaware of the missing "times" when the channel is used for other voice calls, as long as the duration of the missing time is short enough. The TDMA protocol is quite successful at this function.

Back to Table of Contents

3.1 » Types of Cellular Technologies

The standard deployment was called EIA-136 TDMA (eventually ANSI-136 TDMA), and it improved the efficiency of the channel by a factor of three (since each call only used the channel one-third of the time). Essentially, each channel could now support three TDMA voice calls simultaneously rather than one AMPS voice call.[2]

ANSI-95 CDMA

In the 1990s, another new digital protocol was also commercialized. Rather than using TDMA encoding, the digitized human voice bits are combined, or multiplexed, with "codes" using a mathematical algorithm. Thus, this encoding protocol is called Code Division Multiple Access (CDMA).

The combination of voice bits combined with codes allows the data to be transmitted over a single wider channel. In ANSI-95 CDMA, each channel is approximately 1.25 MHz wide. The bits are essentially "spread" across the spectrum width of that channel, and it is thus a "spread-spectrum" communications system.

In general, CDMA protocols are more spectrally efficient than TDMA protocols—this has allowed the deployed CDMA technologies to survive for longer than other protocols. The time slots in TDMA are not necessarily optimal for all use cases and essentially have a limit when every slot in a channel is in actual use for calls.

In CDMA, additional calls are combined (or "spread") using mechanisms that are beyond the scope of this book to describe—suffice it to say that the number of possible calls in a given channel may not be entirely deterministic. A very rough estimate (and this is subject of some heated debate) is that ANSI-95 CDMA was probably 10 to 20 times more efficient than AMPS, while ANSI-136 TDMA was three times more efficient than AMPS.[3]

GSM

In Europe (and eventually most of the world), a different approach was used for the original digital cellular deployments. Although the encoding mechanism is still TDMA, the available spectrum was grouped into 200 kHz channels with eight time slots, rather than 30 kHz channels with three time slots in ANSI-136 TDMA. This system was termed Global System for Mobile Communications (GSM)—a marketing term that described this digital cellular service.

The bandwidth allocations and channel differences in the TDMA transmissions in ANSI-136 TDMA and the TDMA transmissions in GSM are incompatible—a GSM cellphone could not operate in an ANSI-136 TDMA network and an ANSI-136 TDMA cellphone could not operate in a GSM network. Of course, there were other network differences too (such as the messages used in the control channels of the technologies), but the radio technical differences were fundamental.

[2]*The term TDMA is a description of the method and protocol for the data encoding, and "ANSI-136 TDMA" is a specific set of standards implemented for these cellular transmissions. This distinction will become clearer when the GSM technology is described.*

[3]*The term CDMA is a description of the method and protocol for the data encoding, and "ANSI-95 CDMA" is a specific set of standards implemented for these cellular transmissions. This distinction will become clearer when newer technologies are described.*

Back to Table of Contents

3.1 » Types of Cellular Technologies

GSM rapidly became popular in Europe and in other parts of the world—particularly since the early analog cellular systems in many countries were entirely replaced very quickly or were not deployed in the first place in some other countries.

RAPID GROWTH OF GSM NETWORKS AND SERVICES MADE IT A POPULAR CHOICE OUTSIDE THE AMERICAS AND A FEW OTHER COUNTRIES.

This rapid growth of GSM networks and services made it a popular choice outside the Americas and a few other countries. With the far larger deployed base of cellphones, the economies of scale meant that GSM cellphones rapidly became lower in cost than ANSI-136 TDMA cellphones. Thus, the operators in North and South America eventually abandoned ANSI-136 TDMA in favor of GSM to take advantage of this reduced cost.

Data Transmissions

When cellular systems became digitally encoded, it was natural to consider treating the transmitted digital bits as something other than human voice encoded bits—this allowed the deployment of data transmission services for purposes other than human voice. This included communications from mobile radio devices (data handsets) and data cards for mobile computers (laptops) to access the increasingly important Internet and the World Wide Web.

The mechanism for treating the digital bits as application data rather than human voice, was different in the deployed technologies. ANSI-136 departed too quickly for any significant data protocols to be deployed, but both 2G GSM and ANSI-95 CDMA experienced this evolution.

Back to Table of Contents

3.1 » Types of Cellular Technologies

2G GSM Data: GPRS, EDGE

GSM introduced a practical data transmission technology called General Packet Radio Service (GPRS), followed by an improvement called Enhanced Data Rates for GSM Evolution (EDGE) with higher throughput.

These technologies were popular for cellular data communications, although the throughput rates are extremely slow by today's expectations for smartphones that access the Internet. In IoT/M2M applications, however, where the throughput requirements are lower, GPRS is a perfectly good technology for data transmissions.

Thus, GPRS is commonly used around the world for cellular IoT/M2M applications. But it encountered spectral efficiency issues that makes it impractical for use for high-end human smartphone applications. In the US, the largest operator providing 2G GSM announced that it will stop providing GPRS and EDGE data services (and hence, sunset 2G GSM) on January 1, 2017.

Other US operators are likely to follow this example at some time in this decade too, but this is less of a problem for countries where competitive business pressures for wireless spectrum are not as high. For example, GPRS is very likely—albeit not certain—to remain in use in Europe through the middle of the next decade.

GPRS IS COMMONLY USED AROUND THE WORLD FOR CELLULAR IOT/M2M APPLICATIONS. GSM INTRODUCED A PRACTICAL DATA TRANSMISSION TECHNOLOGY CALLED GENERAL PACKET RADIO SERVICE (GPRS), FOLLOWED BY AN IMPROVEMENT CALLED ENHANCED DATA RATES FOR GSM EVOLUTION (EDGE) WITH HIGHER THROUGHPUT.

EDGE was rarely used for IoT/M2M applications, since GPRS (in GSM) and 1xRTT (in CDMA, see below) was sufficient for the vast majority of such uses, and newer data technologies became common for power smartphone users quickly enough.

2G CDMA Data: 1xRTT

Like GPRS in GSM, the CDMA operators in many countries deployed a data transmission technology called 1x Real Time Transmission (1xRTT). This was faster than GPRS in its base throughput rate and has also proven to be very successful for many IoT/M2M applications. Defined into the ANSI-2000 standard, it provided (and continues to provide) a reliable, extensive coverage data network for IoT/M2M applications.

In the US, the wide availability of 1xRTT makes it an easy choice for physically mobile applications, such as the automotive and trucking industry, that need coverage across the continent. The early deployment and expansion of CDMA and 1xRTT (while the other camp was busy with a transition from ANSI-136 to GPRS) led to excellent coverage across the country.

Back to Table of Contents

3.1 » Types of Cellular Technologies

However, the complexity of the CDMA data encoding protocol compared to TDMA also resulted in a higher cost for the radio modules, since chipsets for CDMA radios are more complex. Thus, due to the greater deployment of GSM and economies of scale, 1xRTT modules are more expensive than GPRS radio modules.

3G CDMA (EV-DO)

For smartphone users, the CDMA data standards were substantially improved to enhance their data throughput rates. Technology change cycles added EV-DO Rev. A and EV-DO Rev. B to the portfolio (renaming the first implementation as EV-DO Rev. 0). The changes were added to a new standard called ANSI-2000, which detailed the 1xRTT and EV-DO technologies.

Although used by some IoT/M2M applications, 3G EV-DO has not been extensively used for these kinds of applications, since the higher throughput (compared to 1xRTT) is not strictly required. The excellent coverage and availability of 1xRTT service in the US essentially made it unnecessary to do so, since the radio module costs are higher for EV-DO.

3G UMTS (HSPA/HSPA+)

In the GSM technologies, it became clear over time that the 2G GSM voice and data transports—GPRS and EDGE—that used the TDMA encoding protocol, were not sufficiently spectrum-efficient. The cost of adding new spectrum became much higher, as national governments began auctioning new spectrum for smartphone data uses.

Thus, the standards bodies began defining and deploying a new technology called Universal Mobile Telephone Service (UMTS). They abandoned the TDMA protocols in favor of a new CDMA protocol since CDMA is more spectrum-efficient than TDMA. However, in UMTS, a wider 5 MHz channel differentiates it from the ANSI-95 and ANSI-2000 deployments. The differences are substantial enough that the UMTS protocol is often referred to as using Wide-Band CDMA (W-CDMA) to distinguish it from ANSI-2000 CDMA.

In UMTS, the data technologies have evolved quickly. Early "faster in one direction" transports—such as High-Speed Downlink Packet Access (HSDPA) and High-Speed Uplink Packet Access (HSUPA)—have been mostly replaced by High-Speed Packet Access (HSPA), including variants called HSPA+ that allow for yet faster throughput.

In most IoT/M2M applications, using 3G HSPA is not needed since the performance and throughput of this data technology is high. Indeed, the 5 MHz channel allows it to provide faster overall throughput than EV-DO with its 1.25 MHz channels. However, since 2G GSM data transports (GPRS) have a finite availability in North American markets, there is a need to change, and 3G HSPA is one service that can fill that need.

On the other hand, 3G HSPS is a relatively recent technology and does not have the coverage of 2G GSM GPRS or 2G CDMA 1xRTT. And 4G LTE technologies are also being rapidly deployed. Thus, many 2GSM IoT/M2M applications are either switching to 2G CDMA 1xRTT for an interim solution or leapfrogging 3G to go directly to 4G LTE in the near future. This choice is generally a function of the cost of available radio modules and service coverage.

Back to Table of Contents

3.1 » Types of Cellular Technologies

4G LTE

One limitation of 3G technologies is that they use fixed-width channels. With the ever-increasing number of smartphone data users, the availability of wireless spectrum has created many new bands that are not always optimally usable by 3G technologies. National governments have auctioned a large number of new bands for smartphone users.

To use these new bands, the standards entities developed a new technology for more flexible spectrum use. Since they also had the opportunity to select the encoding protocols to use these new bands, Long Term Evolution (LTE) was designed to use a new protocol called Orthogonal Frequency Domain Multiple Access (OFDMA). Again, the specific mechanism used in OFDMA is beyond the scope of this book, but it has been termed a Fourth Generation (4G) technology, since it is quite different from 3G and also meets some of the original performance requirements set for new cellular implementations under the umbrella of a 4G service.

What is quite important, however, is that LTE is very flexible in terms of the channel widths that can be used, and thus the available spectrum bands can be partitioned into smaller blocks with greater ease. And it also allows existing spectrum to be partitioned into multiple blocks—which can allow an operator to deploy 4G without having to entirely remove older technologies.

The flexibility comes at a price. There are more than 30 bands available for LTE use, and countries have not auctioned or made available the full set of possible bands. Indeed, some bands may be impossible to use for LTE in certain countries because they are dedicated to other uses.

Thus handsets that can be used for LTE everywhere must support a number of different bands, and the addition of each band adds cost, since filters and power-amplifiers inside the radios must support each band. For IoT/M2M applications, this can increase the overall cost of the radio module substantially—perhaps to the point where the LTE device can be financially impractical. Smartphones can absorb the higher cost of multiple band support, since it is a smaller percentage of the overall cost of the phone. This cost issue will eventually drop in impact, because the ever-increasing number of deployed LTE units will cause economies of scale to apply.

FOR IOT/M2M APPLICATIONS, THIS CAN INCREASE THE OVERALL COST OF THE RADIO MODULE SUBSTANTIALLY—PERHAPS TO THE POINT WHERE THE LTE DEVICE CAN BE FINANCIALLY IMPRACTICAL.

In addition, LTE uses the concept of categories (CAT) to define a set of performance metrics that are dependent on other parameters (such as the number of spatial layers, antennas, protocols). Originally defined as CAT 1 through CAT 8, these provided a different range of performance—from 10 Mbits/sec download speeds in CAT 1 through 1200 Mbits/sec downloads in CAT 8.

Back to Table of Contents

3.1 » Types of Cellular Technologies

Most LTE smartphones use CAT 3 and 4 to provide data rates that are sufficient for power users, and CAT 6 smartphones are only just becoming available. For IoT/M2M applications, CAT 1 radios would be sufficient performance, but were not originally developed since the LTE chipsets with CAT 1 support were not deemed adequate for smartphone users.

Recently, the standards bodies also defined CAT 0 radios for LTE that have reduced performance and network requirements, and these are in the process of final definition and ratification. These are expected to be supported in LTE chipsets and within the network (since changes are required in the network deployments too) within the next few years. CAT 0 radios that do not support the higher performance requirements of LTE categories should be less expensive, since the chipsets should be substantially lower in cost too.[4]

5G Cellular Futures

The standards bodies are working on next-generation cellular technologies. These are the Fifth Generation (5G) technologies. At this time, an initial set of requirements and the performance expectations for 5G have been proposed, but formal work on the details and standards have barely begun to be identified. Competing possible implementations have been proposed by academia and technology suppliers, and the standards are not expected to be finalized until 2020.

It is entirely possible that the final standards will be different from the current proposals, but one thing is clear: the explosive growth of IoT/M2M is expected to require a range of capability and support in 5G. Thus, it is very likely that 5G will be extensively used for IoT/M2M, but until it is deployed, we can only speculate on what may be possible.

[4]*The CDMA operators (who had deployed 2G 1xRTT and 3G EV-DO) as well as the GSM operators (who had deployed 2G GSM/GPRS/EDGE and 3G HSPA) are both moving to deploy 4G LTE—this has implications for the types of LTE radios used by these operators.*

Back to Table of Contents

3.2 » Cellular Fall Back

Two Fallback Mechanisms

During the early phases of a new cellular generation deployment, it is often the case that the newer generation is not fully deployed everywhere—the geographical coverage is initially small and expands over time. Thus, the modules must support multiple generations of technologies till coverage is fully complete for the new technology.

Thus, in GSM, all 3G cellular devices—modules, smartphones, and cellphones—are expected to also function in 2G GSM/GPRS and EDGE modes. This allows them to be used in areas where 3G UMTS service may not be available. This increases the cost of the cellular device, but is an acceptable trade-off since it is essential to provide a robust service for all users of the services.

Similarly, in CDMA, the 3G EV-DO modules, smartphones, and cellphones are capable of being used in 2G 1xRTT modes—enabling use in markets where 3G may not be available (this is a relatively rare situation however).

The cellular radios essentially "fall back" from newer generations to older generations when the newer generation service is not available at a particular geographical location. The control of when to fall back (including which technology to fall back to) are incorporated in the Subscriber Identity Module (SIM).

In 4G LTE, there are two technology fallback mechanisms. For the CDMA operators who are deploying LTE, the radio must fall back from LTE to EV-DO and 1xRTT. For the GSM operators deploying LTE, the radio must fall back from LTE to UMTS (HSPA) and then to EDGE or GPRS (since 3G is not available everywhere).

THE CELLULAR RADIOS ESSENTIALLY "FALL BACK" FROM NEWER GENERATIONS TO OLDER GENERATIONS WHEN THE NEWER GENERATION SERVICE IS NOT AVAILABLE AT A PARTICULAR GEOGRAPHICAL LOCATION.

LTE-Only

These fallback mechanisms increase the complexity and cost of the chipsets within the current modules and smartphones. In time, when LTE is commonly available everywhere that cellular services are deployed, it makes sense to use radios that only use LTE services—called LTE-Only modules.

This can, and will, reduce the code of modules substantially—with scale. These LTE-Only devices will approach and even become less expensive than the lowest-cost 2G GPRS radios available today. In a few more years, this should be true for suppliers that provide IoT/M2M modules—customers who want to migrate from 2G to 3G services to 4G may find it worthwhile to wait for this cost reduction in LTE-Only modules to make the transition.

This transition date is dependent on the customer product longevity requirements—clearly 2G GPRS units may stop working in markets (such as the US) soon enough that a transition to a 2G CDMA or a 4G LTE device may be required sooner rather than later.

Back to Table of Contents

3.3 » How to Determine Location

For many IoT/M2M applications, knowledge of the physical location of the devices is important—not only to the device but also to the application servers that process data from the devices.

For example, in consumer automotive, knowledge of the exact location—to a reasonable accuracy—of a vehicle crash is vital so that emergency first responders can be sent to the correct location quickly. Seconds may matter! In truck telematics, a dispatch service may need to know the location of the vehicles in its fleet to optimize the selection of the correct vehicle to handle the specific event—perhaps it is the nearest vehicle to the pickup or one that has the available cargo capacity for the job. In both cases, the knowledge of the device location is important—to a particular degree of accuracy (i.e., the error in the location "fix").

For emergency dispatch, the US Federal Communications Commission (FCC) defined location accuracy requirements that must be made available to Public Safety Access Point (PSAP) personnel. These are often called the "E911" requirements, since the number 911 is used to access emergency services—from landline phones *and* cellphones.

THE FCC DEFINED LOCATION ACCURACY REQUIREMENTS THAT MUST BE MADE AVAILABLE TO PUBLIC SAFETY ACCESS POINT PERSONNEL.

These E911 accuracy requirements are not necessarily sufficient for some IoT/M2M applications—the location error may not allow proper calculation of routes or dispatch, with sufficient optimization. For these applications, alternatives must be used.

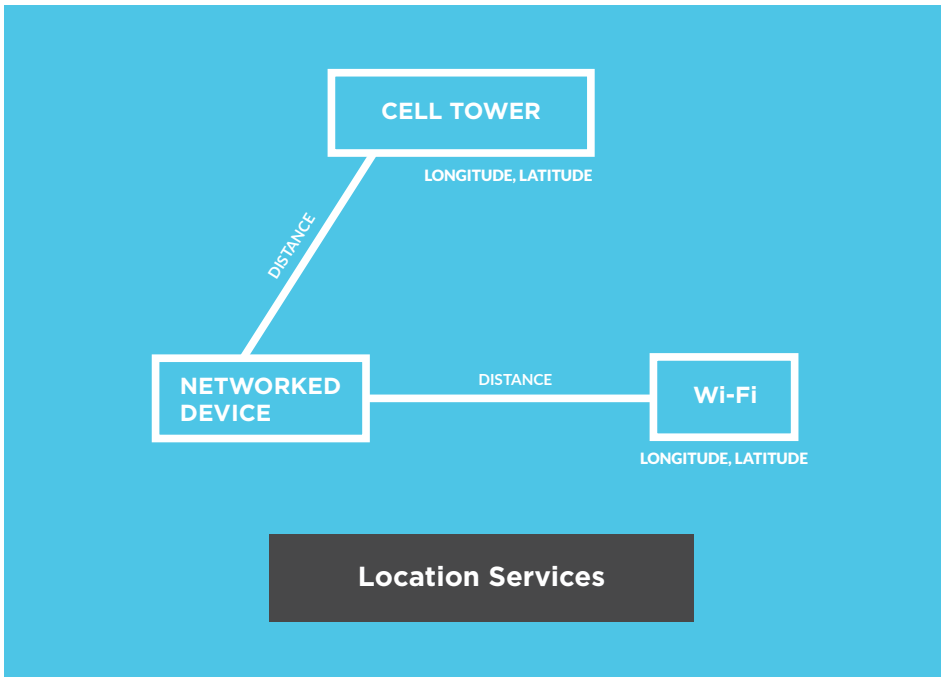


Figure 3. Location Services

Back to Table of Contents

3.3 » How to Determine Location

Location From Cellular Network

Location-Based Services

To support the E911 requirements for physically mobile cellphones used by humans (i.e., which are not fixed at a particular address like a landline phone), cellular operators have implemented various device location mechanisms in their networks. These generally rely on classic radio triangulation techniques that provide the specified degree of accuracy for the E911 requirements.

These network-based location fixes are made available to the PSAP personnel as needed and are also available from operators as Location-Based Service (LBS) information—generally for a fee charged for each location fix of a cellular device. Unfortunately, the cost of these location fixes may be too high for many IoT/M2M uses, and the accuracy may not be sufficient for some uses, and thus, has not proven to be a common technique. Thus, using the GPS (as well as GLONASS, and soon, Galileo) system may well prove to be a superior solution for applications.

Global Positioning System

Many cellphones are now equipped with Global Positioning System (GPS) support that allows the phones to determine their location and provide that information to the cellular network, for E911 and other purposes. Enabling this function is often an available choice in cellphones equipped with GPS.

In IoT/M2M applications, most modules have built-in GPS support (sometimes including support for both systems operated by the US and Russian governments). These can be used by the application firmware in the device when needed for a particular function—such as responding to a location fix request by a dispatch application.

GROUND-BASED REFERENCES CAN BE USED BY CERTAIN RECEIVERS TO GREATLY ENHANCE THE BASIC ACCURACY OF THE GPS SYSTEM FROM 15 METERS TO LESS THAN 10 CENTIMETERS.

What Is the GPS System?

In the latter half of the last century, the US Department of Defense deployed a set of 24 satellites into Earth orbit for a very singular purpose: it allowed a GPS-equipped device to determine its location on the surface of the earth with very good accuracy.

Originally intended for military users, the US government made the system and its information available for civilian use in the 1980s, without any fee or subscription charge. This enabled a large number of new location applications around the world.

For example, the truck telematics industry relies heavily on the GPS system to locate trucks and trailers. Hikers and off-road personnel can use hand-held GPS trackers to avoid becoming lost. High-accuracy augmented GPS systems are used by semi-automated farming equipment since these can often locate the vehicle within a few centimeters on the surface of the Earth! Survey equipment can use GPS to accurately measure location for mapping and thus increase map quality—improving route guidance systems in vehicles.

Back to Table of Contents

3.3 » How to Determine Location

A satellite service similar to GPS, called GLONASS, has been deployed by the Russian government. The European Union is in the process of launching its own system called Galileo (named after the historic astronomer). As of this writing, the Galileo constellation of satellites is not yet operational.

Within time, Galileo will provide a free, low-precision location fix with an accuracy of 1 meter, but high-precision fixes will only be provided for a fee. Since it is a new system, it also has new features that are not available in the older US GPS and Russian GLONASS systems. For example, Galileo has radios that are planned to support a unique relay service for Search-and-Rescue (SAR) distress signals, allowing emergency dispatch around the planet.

In addition to the satellite GPS system transmissions, enhancements are available to dramatically improve the location accuracy. For example, a set of ground-based references can be used by certain receivers to greatly enhance the basic accuracy of the GPS system from 15 meters to less than 10 centimeters. This enhanced system is called Differential GPS and enables farms to use automated equipment that need a very high accuracy location fix.

How Does Basic GPS Work?

In the American GPS system, 24 GPS satellites orbit the Earth twice a day[5] in a very precise manner at an altitude of approximately 20,000 km while transmitting accurate time signals from their on-board atomic clocks to ground GPS receivers.

These GPS receivers take the received time data and use triangulation (more correctly, "trilateration" using points of intersection of circles on a sphere—angles are not measured) techniques to determine the location of the receiver. The receiver essentially compares the time a signal was transmitted by a GPS satellite to the time it was received—this time difference allows the receiver to determine its distance from that satellite.

When this time difference and distance is determined from a number of GPS satellites, the location of the receiver can be determined within about 5 to 10 meters of accuracy on the surface of the Earth. At least three satellites must be used for a latitude-longitude fix on the surface of the Earth, and a fourth satellite can then determine the altitude of the receiver.

It should be emphasized that the above is a very general description of the method used to determine location from the GPS satellite signals. There are a number of other factors that affect the accuracy and are taken into account by sophisticated receivers. For example, the more satellites the receiver listens to, the better the accuracy. Thus, a 10- or 12-channel GPS receiver (which allows it to listen to 10 or 12 GPS satellites simultaneously) will generally provide a more accurate location fix than an older 4- or 6-channel receiver.

Furthermore, since the GPS satellites are in motion and are quite far above the Earth's surface (i.e., operating in reduced gravity), Einstein's Special and General Theories of Relativity must be used to correct the data (since time literally flows at a different rate for the satellite clocks compared to the Earth-bound clocks).

[5] Contrary to popular belief, GPS satellites are *not* in a geo-synchronous orbit above the same spot on earth.

Back to Table of Contents

3.3 » How to Determine Location

Without proper compensation, the relativistic effects of the speed of the satellites combined with their height could create a net error of about 38 microseconds per day at the satellite clock, compared to an identical ground-based clock. This may seem quite inconsequential, but this difference in time can make the location fix inaccurate within a matter of minutes, to beyond the 5 to 10 meter accuracy of the system. Then, accumulated errors could make the location fixes completely unreliable and unusable in a matter of days to weeks since the GPS system requires nanosecond time accuracy. Fortunately, the GPS system uses these Einsteinian Relativity calculations and corrects to ensure that the time and location accuracy is excellent, and remains excellent, under most conditions.

EINSTEIN'S SPECIAL AND GENERAL THEORIES OF RELATIVITY MUST BE USED TO CORRECT THE GPS DATA.

With multiple location fixes spaced in time, these fixes can also be used to determine other information such as speed and direction (i.e., velocity). Sophisticated GPS tracking devices can use the data to display the location and provide route guidance in friendlier ways than a simple latitude-longitude-height-time record—displayed on a graphical moving map, for example.

Limitations of GPS

Location fixes from GPS are not perfect. In "urban canyons" (i.e., within cities with tall buildings), it may be difficult for GPS receivers to lock onto more than a few satellites, since the signals may be blocked by the buildings. This may reduce the accuracy substantially—regardless, it may remain sufficiently capable for many uses of that location data. A higher-performance GPS receiver with many channels may perform better in urban canyons since it has a better chance of listening to satellites that may be "visible" and not blocked by tall buildings. The signal strength is low enough that many GPS receivers cannot listen to the signals from the satellites when inside buildings and underground garages. This limits their use in indoor applications.

Sometimes, heavily overcast days can reduce the strength of the GPS signals enough to prevent the receiver from locking on to the signals, particularly when the receiver has been re-started from a power-off condition. If the internal clock of the receiver is not sufficiently accurate, the measured time may have drifted, and the receiver could be attempting to listen to a set of satellite signals that are not present—those particular satellites may not be visible.

Back to Table of Contents

IOT SENSORS AND DATA COLLECTION

Chapter 4

Back to Table of Contents

When deploying Internet of Things and machine-to-machine application devices, the connected device generally needs to report more than just its physical location. In this chapter, we describe a few of the more common sensors and what they do.

For example, an IoT/M2M device may measure a particular physical parameter at that location—these physical parameter measurements require sensors that are capable of recording the specific value of that parameter for the device application to fulfill its functions.

Sensors are often integrated circuits that are designed for these kinds of IoT/M2M applications, since the small size and low cost of these chips make them appropriate choices. For example, many of the sensors described in this chapter are available in high-end smartphones. These include accelerometers, thermometers, gyroscopes, magnetometers, and heart-rate monitors—just to name a few—but there are other sensors that are unique to a particular industry or market.

In this chapter, we describe a few of the more common sensors, what they do and how to use them.

4.1 » Typical IoT/M2M Sensors

In most of these typical sensors, the specific mechanism used to measure the physical parameter depends on the ranges being measured, the sensitivity and accuracy desired, whether the sensor could be exposed to adverse environmental conditions, the cost target, etc.

Since it is quite impossible to list every possible sensor, its type, and its capabilities, this section focuses on general descriptions of a few types of sensors rather than making specific recommendations.

Accelerometers

Acceleration is a measure of a change in velocity (change of speed or direction). Accelerometers are devices that measure acceleration. The parameter being measured may be a static force, such as gravity exerted on a device. Other sensors make dynamic force measurements to measure motion changes and vibration.

An example of an acceleration sensor is a chip in a moving vehicle that measures changes of speed and uses high acceleration readings (such as during an accident) to trigger an airbag to protect the passengers.

In some industrial applications, the vibrations detected by an acceleration sensor could be an excellent indicator of a potential problem with a moving part—such as a motor with bearings that are worn. Timely transmission of the data from vibration sensors enables early detection of problems where preventative maintenance could avoid catastrophic failures.

VIBRATIONS DETECTED BY AN ACCELERATION SENSOR COULD BE AN EXCELLENT INDICATOR OF A POTENTIAL PROBLEM WITH A MOVING PART.

Back to Table of Contents

Multi-Axis Accelerometers and Sensitivity

In some applications, there is a need to measure the change in speed or vibration in more than one direction (or dimension). Thus, some accelerometers take readings in more than one axis. Typically, a two-axis sensor measures motion changes and vibration in two dimensions, and a third axis on a three-axis sensor can provide information for three-dimensional physical motion detection.

Accelerometers use differing techniques for measuring the motion changes—generally, there is a physical component that changes an electrically measured characteristic (such as capacitance or resistance) in a material when motion is sensed.

Due to the types of accelerations being measured, the sensitivity of the accelerometer is often over a limited range to the required accuracy that is specific to a particular application use. The choice of sensor to use thus depends on the specific range of acceleration values to be measured for that application.

For example, a shock sensor designed to release a vehicle airbag in an accident measures quite a different range compared to a sensor that measures vibration on a high-speed motor to monitor its bearings. The sensitivity and accuracy required for these widely disparate applications is naturally quite different.

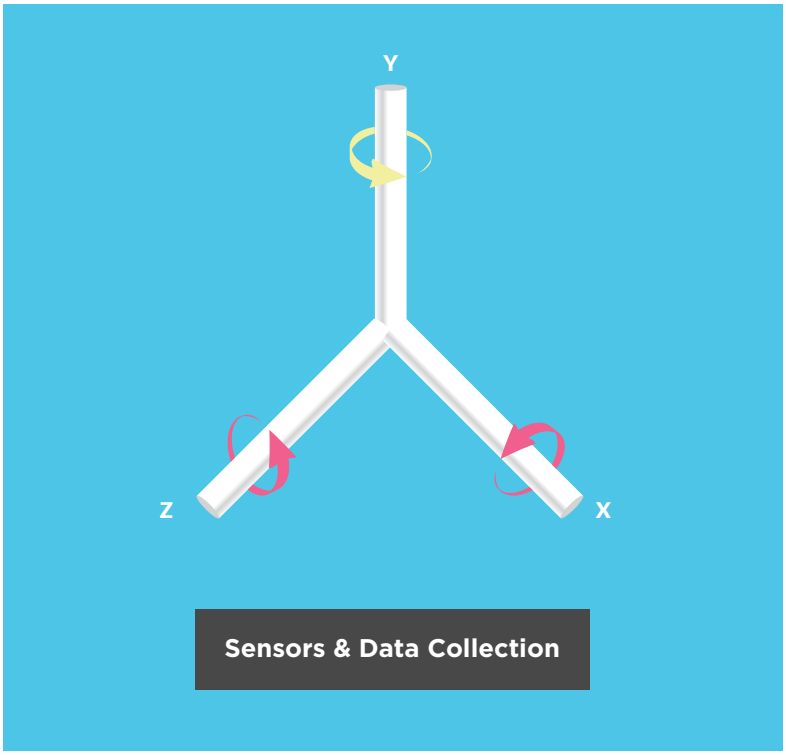


Figure 4. Three-Dimensional Motion Detection

Back to Table of Contents

4.1 » Typical IoT/M2M Sensors

Temperature Sensors

Temperature is a parameter that is often measured and reported, particularly in industrial applications where an accurate temperature reading may be needed for process control. Depending on the desired measurement range, there are various types of available sensors for measuring temperature.

Silicon chip (semiconductor) sensors are easily used in the range from -50 to +150 degrees C. These are quite accurate and linear—to within 1 degree—without the need for extensive calibration. They are as rugged as most integrated circuits (package and metal-can style) and relatively inexpensive.

Thermistor sensors can cover a wider range—from -100 to +450 degrees C for more applications. A thermistor is often more accurate than a silicon chip temperature sensor, albeit at a slightly higher cost per sensor. More importantly, they require a complex correction to achieve good linearity and accuracy over the desired temperature range.

Resistance-Temperature Detectors (RTD) provide yet more range, from -250 to +900 C, but are quite difficult to use since they are more fragile than other types of temperature sensors. They are the most accurate—often a hundred times more accurate than a silicon chip sensor, although this carries the same complex solutions for linearization as thermistors, and some models can be quite expensive.

Finally, for the widest temperature range, particularly for high temperatures, a thermocouple is the correct choice. They are quite rugged and can be used from -250 to +2000 C for many industrial applications, such as chemical process monitors and high-temperature furnaces used in the semiconductor industry.

One important fact about temperature sensors: the response time for the sensor data can be quite slow, since temperature changes are not as "rapid" as other measured physical parameters. The sensors must settle and equalize to the same temperature as is being measured—this must be taken into account when taking readings.

THE RESPONSE TIME FOR TEMPERATURE SENSOR DATA CAN BE QUITE SLOW, SINCE TEMPERATURE CHANGES ARE NOT AS "RAPID" AS OTHER MEASURED PHYSICAL PARAMETERS.

Back to Table of Contents

4.1 » Typical IoT/M2M Sensors

Light Sensors

Light sensors cover a broad range of potential applications—from automated brightness control in cellphones to medical diagnostic equipment. Not surprisingly, there is also a wide range of available light sensors that use different methods for measuring the ambient light.

A very early example of ambient light sensors used in local consumer applications are photocells within lamps that automatically turn the lamps on at dusk and turn them off at sunrise. These are simple light detectors, with equally simple sensitivity controls that are manually adjusted by the owner of the product—the actual value (in lumens) of the ambient light is not measured or reported.

Simple light sensors can also be for proximity detection. Counters in manufacturing systems use the presence or absence of light on photocells to measure products being moved past the counter on conveyors. Garage door systems can reverse direction to avoid hurting children or pets who cross under a closing garage door and temporarily cut the light from a source sending a beam of light across the door opening onto a photocell.

Often, light sensors can be used with light that is not visible to human eyes. Infra-red light sensors can be used as motion sensors in alarm systems or to automatically light a driveway or passage when people and pets come into range. Full-range sensors are used when the light measurements need to correspond to human vision.

As with other types of sensors, the mechanism used to measure ambient light varies depending on the application. Simple Cadmium Sulphide (CdS) or Cadmium Selenide (CdSe) photo-resistors change their resistance as a function of the ambient light. This resistance change can be measured in electronic circuits to provide an indication of a change in the ambient light. It should be noted that these photocell devices can be significantly affected by temperature and are quite unsuitable when accuracy is required.

Common uses of photo-resistors include automated light controls in lamps, dimmers in alarm clocks and audio system displays, control of street lighting systems, etc., where the accuracy of the reading is not a paramount requirement.

Photo-diodes and photo-transistors, with active semiconductor junctions, are used when greater accuracy is required, since the ambient light is converted into a measurable current that can be amplified or converted for a measurement. This measured current can be used to determine the amount of ambient light on the sensor.

Indeed, since semiconductor junctions are affected by ambient light, integrated circuits where this effect is not desired must be enclosed in opaque packages.

Back to Table of Contents

4.1 » Typical IoT/M2M Sensors

MEMS Sensors

In modern, high-end smartphones, integrated chip sensors to measure motion, direction, pressure, magnetic fields, etc., are becoming quite common. These can be used to augment the location information and human motion in the cellphone.

In chip form, these are usually Micro-Electro-Mechanical Systems (MEMS) sensors for many different parameter measurements. The implementation of MEMS uses ultra-miniaturized physical structures—beams, arms, and associated electronics—to measure the motion of the structures when the chips moved. The physical motion is converted to electrical signals that can be measured for the specific function being measured—for example, whether it is rotational motion or air pressure. The device essentially converts a mechanical motion into an electrical signal.

A gyro sensor, for example, senses rotational motion and changes in orientation. These can be used in a variety of applications, such as correcting for hand-held shake in video and still-image cameras and human motion sensing for video games. In smartphone applications, a screen display can be automatically rotated from portrait to landscape display modes when the phone is physically rotated.

MEMS SENSORS CAN BE USED IN A VARIETY OF APPLICATIONS, SUCH AS CORRECTING FOR HAND-HELD SHAKE IN VIDEO CAMERAS.

MEMS sensors are generally manufactured in the same large-scale facilities as semiconductors or chips. This means that the mechanical precision of the devices can be very high and allow for excellent, reliable performance at low cost.

Simple Switch Sensors

At the low end are the simple state or position sensors that provide an “open” or “closed” state. A door or window sensor used in security systems is often a simple magnetic reed relay switch that opens or closes an electrical circuit depending on the position of a small magnet relative to the switch.

These simple magnetic reed relay switches can also be used for sensing when a cabinet—such as a medicine cabinet, oven door, or food storage compartment—has been opened in a senior citizen’s home-monitoring IoT application. A detection of the change of state of such a switch—from open or closed or vice-versa—can be interpreted as evidence that the monitored parent has performed their expected regular daily routine.

Back to Table of Contents

4.2 » Conversion to Digital Data

Because of the wide variety of sensors, the types, the parameter being measured, and what physical phenomenon is converted into a measurable signal, it is difficult to provide implementation details. Thus, this section must necessarily discuss general concepts rather than specifics.

Sensors are often used in local applications, where their signal is processed using circuitry designed for that local application. However, in a sensor that is used for remote data transmission of the measurement, the electrical signal must be converted into a digital value, or number, for the transmission.

The specific electrical signal from different sensors may vary over a wide range of current or voltage or other electrical parameters (such as resistance or capacitance) and often must be converted and amplified into a voltage that can be measured.

If necessary, the signal must be filtered to eliminate noise or to reduce the frequency of the measurement for the requirements of the application. For example, a temperature sensor generally changes its value relatively slowly as the sensor matches its environment. Therefore, a rapid change in reported temperature may be an inaccurate reading, which should be filtered to reduce potential errors.

Input and Output Pins (I/O)

In devices that measure sensors for data transmissions, two input capabilities are generally available:

- A digital input pin that reports an electrical "high" or "low" value in a single digital bit (sometimes grouped into multiple pins and multiple bits).
- An analog input pin that receives a voltage from a sensor and converts that voltage, using an Analog-to-Digital Converter (ADC), to a digital number that represents the sensor value.

These devices may also have output pins where a received value from the network is used to:

- Set a digital output pin to either "high" or "low" state based on an instruction to do so.
- Set an analog output pin to an analog voltage, using a Digital-to-Analog Converter (DAC), representing the received digital number.

Simple Off and On Switches

In simple switch applications, where the state is "open" or "closed," using a digital pin to measure this state and report its value ("0" or "1") is an easy choice. For more complex needs with simple switches, the device may also report *when* the transition from one state (such as "open") to the other state (such as "closed") occurs. That is, it may be equally, or *more*, important to report a change of state rather than the present state of the simple switch.

Back to Table of Contents

4.2 » Conversion to Digital Data

Range of Data Values

In sensors that measure parameters over a range, a single bit is insufficient–the range of the sensor values must be converted into a range of digital values for the application.

However, the specific signal from a sensor may differ widely in its current or voltage or resistance value. This signal–whether it is a current or resistance change–must be "conditioned" or converted to an analog voltage. If the signal from the sensor is a voltage, it may not be in the correct range for an ADC to convert to a digital number and thus may require amplification to a higher or lower range.

SOME SENSORS MAY HAVE BUILT-IN FUNCTIONS FOR CONVERTING THE MEASURED PHYSICAL PARAMETER DIRECTLY TO A NUMBER THAT IS SENT TO THE APPLICATION PROCESSOR.

For example, a commonly available semiconductor temperature sensor provides an electric current of 1 microAmp per degree Kelvin when power is applied. Over a useful range of -50 degrees C (or 223 degrees Kelvin) to +150 degrees C (or 423 degrees Kelvin), this current is approximately 223 microAmps to 423 microAmps.

This current can be used in a circuit with an Operational Amplifier (Op-AMP) and other components (resistors, capacitors, and diodes) to convert to a voltage in the desired operating temperature range being measured. This voltage can then be measured by an ADC and processed by the device taking the temperature measurement for the application function.

Some, usually more complex and expensive, sensors may have built-in functions for converting the measured physical parameter directly to a number that is sent to the application processor or communications module for transmission. For example, a GPS device may report continuous position and time readings on a serial port using common National Maritime Electronics Association (NMEA) formats called NMEA 0183 or NMEA2000®. This signal is already conditioned in a text format that can be used by a device processor to communicate and transmit.

Back to Table of Contents

4.2 » Conversion to Digital Data

ADC and DAC Resolution

When converting the analog voltage signal from a conditioned sensor to a digital value or number, the ADC has a pre-defined resolution. This means that the full range of the measured analog signal varies from a zero value to a maximum high numerical value, with incremental steps defined by the resolution of the ADC.

For example, an 8-bit ADC will convert the signal from a low value of 0 to a high value of 255 (with integer numbers in between) to represent the value of the analog signal in approximately equal steps. This may quite sufficient for many IoT/M2M applications.

In other applications, it may be necessary to use a 12-bit, or even a 16-bit, resolution ADC for the conversion. A 16-bit ADC provides a digital numerical value between 0 and 65535 based on the input analog voltage. With higher resolutions—particularly with low signal levels, the signal conditioning and amplification circuits may need special care to ensure that electrical noise does not result in erroneous readings.

It is important to note that resolution is not the same as accuracy or linearity—it merely states the number of steps between the lowest and the highest value being converted.

A full discussion of these concepts is beyond the scope of this book—interested readers can refer to the data sheets and applications notes from ADC and DAC suppliers for more information.

Modules or External Processors

Quite often, the communications modules or modems—particularly cellular products used in industrial IoT/M2M applications—have multiple input and output (I/O) pins that can provide the conversions. This can be a simple on/off state, using digital input pins, or an analog voltage reading, with an on-board ADC on an analog input pin, converted into a number that is transmitted on the communications network.

Some modules also have digital output pins for setting a state external to the application—for example, to activate a relay to turn on a light, power on an electrical device, disable a vehicle, or perform some similar remote function.

A few modules and modems also have DACs that take a digital number received from the communications network and output an analog voltage on an analog output pin in a specific voltage range. This may be used where an analog voltage is used to control the position of a liquid flow valve or the speed of a motor in an industrial IoT/M2M application.

Back to Table of Contents

4.3 » Calibration and Linearization

As described earlier, a simple application (such as a photocell that controls a lamp to turn on or off based on ambient light) may not need an accurate reading or sensor value. However, when accuracy is important for an application, calibration of the sensor signal may be needed to ensure that the data reading is accurate to the required degree.

For example, the semiconductor temperature sensor mentioned earlier can provide a reading of 1 microAmp per degree Kelvin for the environment it is in. However, does a reading of 273 microAmps actually mean that the temperature is *exactly* 273 degrees Kelvin (0 degrees C)? Or can the reading be erroneous to a certain degree of error? Without calibration, it is difficult to be completely certain, although it is a good estimate of the temperature.

Other temperature sensors are even more problematic. For example, an RTD can be a hundred times more accurate than a semiconductor temperature sensor—but without correction, its readings are quite useless. The RTD requires precise signal conditioning, linearization, and calibration to achieve that accuracy.

These corrections are often applied digitally—the reading from the RTD is first converted to a digital value, and then the correction is applied. Different types of RTDs need different types of corrections. For example, a platinum RTD has two distinct relationships to temperature with different polynomial equations describing its resistance above and below 0 degrees C.

In an RTD, to achieve the best accuracy, the measured signal can be corrected using a variety of techniques: direct math, single linear approximation, or piecewise linear approximation. Each has its advantages and disadvantages.

It is beyond the scope of this book to describe the specifics for correcting the readings from sensors (for example, for correcting an RTD measurement). Suffice it to say that developers designing IoT/M2M applications must take linearization and calibration into account for the specific needs of their application, particularly if the resulting accuracy is important to the function and features.

Back to Table of Contents

4.4 » Specialized Sensors

In industrial and simple applications, many standard sensors have been developed and commercialized over the years. These have evolved and improved over time. The costs and sizes of sensors have been reduced with increasing efficiency and practicality. Recently however, particularly with the start of the IoT/M2M revolution, there has been a great demand for a variety of new parameters to be measured at ever-larger scale and lower cost.

The healthcare industry is among those at the forefront of this revolution. New methods to measure human medical parameters are being researched and commercialized, and this has seen an explosion of new techniques (and sensors) to measure these parameters. In medical monitoring applications, the need for new measurements, reduction in size of devices, and adoption of wearable fitness and medical products, is driving significant research and growth.

Beyond the sensors that are incorporated into hand-held or wearable products (e.g., smart watches, clothing, and bracelets) for reading basic body functions or medical monitoring products (such as continuous blood sugar monitors and insulin dispensers), there is also a need for semi-permanent sensors implanted within the human body. The research into tiny, implantable sensors has been energized by the availability of semiconductor and MEMS solutions—including for mission-critical applications such as cardiac monitoring and vision correction.

THE RESEARCH INTO TINY, IMPLANTABLE SENSORS HAS BEEN ENERGIZED BY THE AVAILABILITY OF SEMICONDUCTOR AND MEMS SOLUTIONS.

For example, medical startups are developing MEMS sensors that are implanted into pulmonary arteries using cardiac catheter procedures similar to angioplasty. These sensors can measure artery pressure and transmit the readings to a nearby wireless device by the patient at home, and the readings can then be wirelessly sent to a database for review by medical practitioners.

Back to Table of Contents

SCHEDULING, ENCODING, AND PROCESSING

Chapter 5

Back to Table of Contents

As mentioned in the previous chapter, data and sensor readings are generally transmitted to IoT/M2M application programs for processing, storage, and business actions.

This may be a relatively short-range transmission—the sensor readings are delivered to a smartphone application using a short-range wireless technology such as Bluetooth, ZigBee, or Wi-Fi, for an action by the owner of the smartphone.

For example, a heart-rate monitor may send heartbeats-per-minute to a smartphone application during exercise, and this can be monitored to modify the specific physical activity. The data can be logged by the application to ensure that the desired fitness goals are being met.

For other IoT/M2M applications, the data is sent over a longer-range transmission to servers and programs, where it is processed for actions or stored for analytics—the data (or patterns in the data) may lead to business actions if appropriate for that specific application.

For example, an airbag deployment notification from a vehicle can be sent to an automotive Telematics Service Provider (TSP) that contacts the driver and connects them to public safety personnel for dispatch of emergency services.

This chapter describes the systems and methods used to encode, transmit, store, and process the data in a server application.

5.1 » Data Transmission Schedules

Devices may transmit their data in real-time, or scheduled rate, or when the device firmware requests a report of an event.

Devices that send their data continuously in real-time or near-real-time are "streaming" applications. The processing of this streamed data requires systems capable of handling the throughput from a large number of devices, particularly if the content is to be analyzed in real-time for specific actions at a remote site.

The cost of transmitting real-time streaming data on "metered" communications networks that charge for "quantity of bytes sent" may be prohibitive for many applications.

Scheduled Transmissions

In some applications, devices transmit on a regular schedule—sometimes sleeping to conserve power until they are woken up by scheduled timers or to report an event.

Devices with accurate time (such as those equipped with GPS capability) must be careful when using regularly scheduled transmissions. In large deployments, if all devices were to wake and transmit at the same time, the simultaneous connection attempts could overwhelm the connectivity paths *and* the server systems that receive and process the data. If possible for an application, randomizing the transmissions can have a major effect on the capacity requirements of the connectivity and the server systems.

Back to Table of Contents

5.1 » Data Transmission Schedules

There are simple ways to achieve this randomization. For example, a device identification number—such as the last four digits of the Mobile Directory Number (MDN) in a CDMA cellular device, modulo 3600—can be used to select the "number of seconds past the hour" when a regular transmission is sent.

Transmit On-Demand

In most IoT/M2M applications, it is typical for the device to transmit "on demand" when an event requires it to do so. For example, a business or residential security system may transmit a signal when a break-in occurs; a car may transmit an accident notification when an airbag deploys or when the driver pushes a concierge button for assistance. These are generally sporadic enough or spaced temporally sufficiently well that they do not create traffic (and server) spikes.

Often, devices that transmit to report sporadic events are also set to transmit a periodic "heartbeat" to report their condition and health—these heartbeat transmissions should also be randomized.

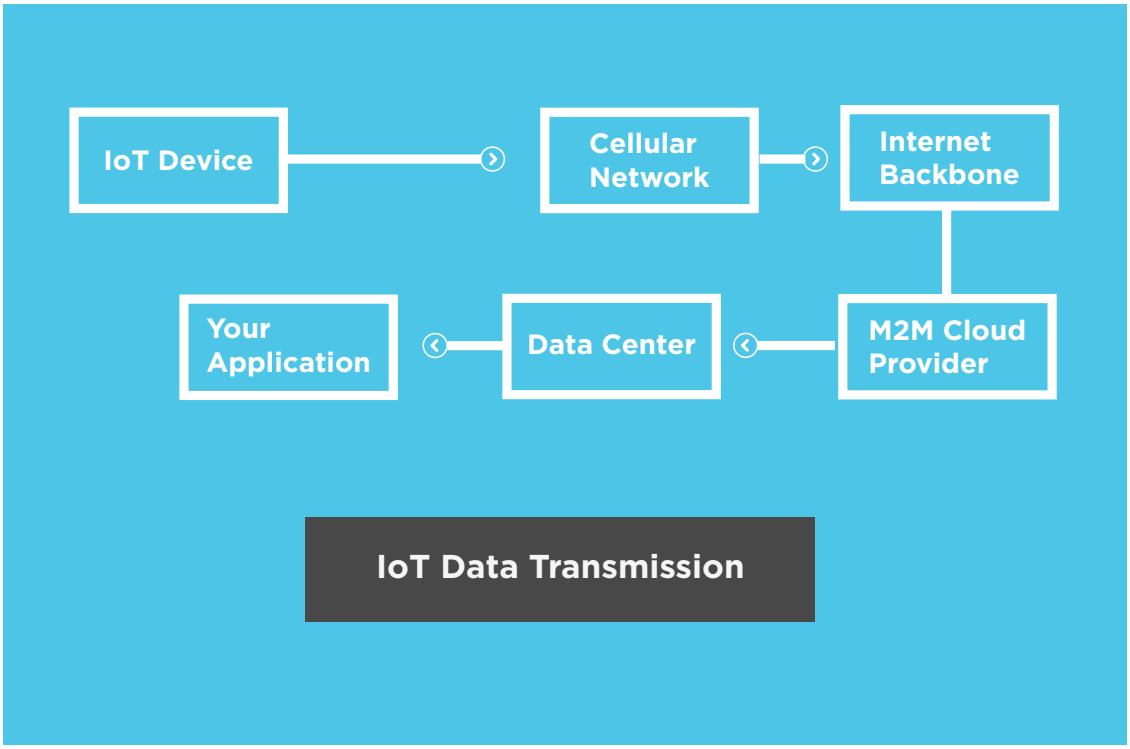


Figure 5. Sample IoT Device Data Transmission

Back to Table of Contents

5.2 » UDP or TCP

We are often asked whether a device should use User Datagram Protocol (UDP) packets or use Transmission Control Protocol (TCP) streaming sessions for the data. The answer, not surprisingly, is: "It depends!"

The Internet Engineering Task Force (IETF) has detailed definitions, but let's briefly describe these two protocols to understand why one may be better than the other for certain IoT/M2M data transmissions. It is important to note that both UDP and TCP are used over an underlying IP connection.

USING THESE TWO PROTOCOLS IS NOT MUTUALLY EXCLUSIVE FOR A GIVEN IOT/M2M APPLICATION.

User Datagram Protocol (UDP)

The UDP format was first defined in an IETF Request For Comment (RFC) specification RFC 768. This protocol provides a procedure for applications programs to send messages to other programs with a minimum of protocol overhead. This protocol is transaction-oriented, and delivery and duplicate protection are not guaranteed.

If an application requires ordered, reliable delivery of streams of data, UDP is not the preferred protocol.

However, the format has lower overhead than TCP—i.e., fewer bytes are sent in the headers of the packets in UDP than TCP.

Transmission Control Protocol (TCP)

The TCP format was first defined in an IETF RFC specification RFC 761. TCP is a connection-oriented, end-to-end reliable protocol and is intended for use as a highly reliable host-to-host protocol between hosts in IP networks and especially in interconnected systems of such networks.

TCP requires that a connection be opened and managed for the duration of the IP data transmission. Within the protocol, transmitted and received packets are acknowledged by the device and the servers.

This format has more overhead than UDP—i.e., more bytes are sent in the headers of the packets in TCP than UDP.

Back to Table of Contents

5.2 » UDP or TCP

Which to Use?

In general, the choice of UDP vs. TCP must take into account:

- The desired balance between the reliability of TCP and the lower cost of UDP, since UDP uses fewer bytes of overhead to transmit the same amount of application data.
- The increased complexity of TCP, where the module must open a data stream to a remote server where programs await connections.
- Careful design of TCP server programs to allow easy scaling as the number of deployed devices increases.
- A requirement for the acknowledgments provided by TCP sessions.

However, it is also important to note that using these two protocols is not mutually exclusive for a given IoT/M2M application.

For some uses, a simple transmission of a UDP packet to a remote host may be quite sufficient–including using independent acknowledgments via UDP. If an acknowledgment is expected, but not received, either side can retry … intelligently (i.e., with limits on number of retries, variable delays between retries, etc.)

For other uses, even in the same application perhaps, a device may open a TCP connection to a server and communicate with the higher reliability of a TCP streaming session to a program that accepts these connections and transmissions.

Often, the amount of data may *require* TCP. For example, if a device needs to transmit a large file (i.e., more than a kilobyte), it is better to use TCP, since the consequences of an error during transmission via UDP could mean that the entire file might require a complete retransmission.

Back to Table of Contents

5.3 » Content Encoding

When a device transmits its data to the servers and receives commands and instructions from the servers, there is a format required for the information sent in both directions. In every application, the devices and servers must agree on the format and information that is sent.

Proprietary Formats

Devices for a particular application and the servers may use a proprietary format for the data encoding. This allows the devices and servers to encode and interpret the content in ways unique to the needs of that application and can often minimize the amount of data sent in that connection session.

Proprietary formats are more difficult to implement initially—since they must be relatively complete for that application to be deployed—as well as difficult to maintain and update later when changes are needed. Most proprietary formats tend not be extensible.

Common Industry Formats for IoT

In addition to proprietary formats and early standardized formats such as Extended Markup Language (XML), there are some industry formats and protocols in use for IoT/M2M data communications for messaging needs, for example:

- JavaScript Object Notation (JSON)
- Constrained Application Protocol (CoAP)
- Message Queuing Telemetry Transport (MQTT)
- Extensible Messaging and Presence Protocol (XMPP)

THESE PROTOCOLS FALL INTO TWO BASIC CATEGORIES: HUMAN-READABLE (JSON, XMPP) AND NON-HUMAN-READABLE (COAP, MQTT).

These protocols fall into two basic categories: human-readable (JSON, XMPP) and non-human-readable (CoAP, MQTT). The human-readable ones are generally much more verbose, but easier to debug during development. The other, non-human-readable, ones are lighter-weight and efficient and can minimize the amount of data sent over the communications path.

Each format and protocol has its pros and cons when used for IoT. The specific choice depends on the needs of the application, the bandwidth of the communications network, the computing power in the sensor or remote device, and other factors.

Back to Table of Contents

5.3 » Content Encoding

JSON

JSON is an open-standard format that sends key-value pairs of information. The "key" is generally the attribute or description of the content sent in the "value." The protocol is described in RFC 7159 from the Internet Engineering Task Force (IETF). See www.ietf.org/rfc/rfc7159.txt for more information.

The JSON format is human-readable and language-independent, and public code for parsing and generating JSON text data is readily available in a variety of programming languages. The format is effectively self-describing, since the definition and value are right next to each other.

For example, the following simplified text illustrates the encoding of a temperature reading of 25 degrees Centigrade from a sensor with a hypothetical sensorID of 123456789:

```
{
        "sensorID" : "123456789" ,
        "temperature" : "25" ,
        "units" : "Centigrade"
}
```

As you can see, the JSON content is verbose and very human-readable. The key-value pairs immediately identify the attribute and its value.

JSON format messages can also be extended readily. For example, the physical location and manufacturer might be added, along with a time-stamp noting the time that temperature was measured.

Of course, the presence of this additional information depends on whether it *should* be transmitted. In the above example, the sensorID could be used to look up the location and manufacturer in a server database (assuming it was accurately stored there at installation of the sensor). However, sending the time-stamp from the device can be more useful since it provides the time when the data was collected (assuming the device knows that time information, of course).

CoAP

As the name implies, CoAP is a format and protocol intended for use in bandwidth limited networks or where minimizing the size of each message transmission is important. The core of the protocol is described in RFC 7252 from the IETF, although extensions to add unique requirements for IoT/M2M are still in progress. See http://www.ietf.org/rfc/rfc7252.txt for more information.

CoAP is a simple protocol that is well-suited for transmissions from small electronic components, such as sensors, and can also be used to control the devices from remote servers. CoAP also includes the concept of "multi-cast" (or "one to many") group communication, where many devices can receive the control information at the same time.

Back to Table of Contents

5.3 » Content Encoding

The protocol provides two types of message: requests and responses using a "type-length-value" (TLV) coding that is different from the JSON format. These CoAP messages are sent using a UDP transport to adhere to the concept of low overhead for the messages.

MQTT

MQTT is another light-weight messaging protocol that is designed for data transmissions from devices operating in bandwidth limited networks. The devices transmit the data to message brokers that are then responsible for sending the content of the messages to clients who are interested in that data and who subscribe to the feed.

This mechanism is the essence of a "publish-subscribe" approach—where data from the devices is published to a broker, and subscribers to that broker can access the data.

Originally developed by IBM, the MQTT protocol was transferred to the OASIS standards body and is now supported by that entity. See www.mqtt.org for more information.

MQTT was originally designed for the IoT/M2M markets for devices transmitting using TCP/IP. To allow simpler electronic devices (such as sensors) to use this protocol, a version called MQTT for Sensor Networks (MQTT-SN) has also been released to extend the protocol beyond TCP/IP.

XMPP

XMPP is an open-standard communications protocol for messages based on XML. It is intended for near-real-time exchange of messages between two (or more) elements on any network. Like XML, it is extensible and can also be used for publish-subscribe message systems.

There are multiple RFCs from the IETF that specify the XMPP standards: the core ones are RFC 3922, 3923, 6120, 6121, and 7622 (see www.ietf.org), although the XMPP Standards Foundation (see www.xmpp.org) is also actively extending XMPP further.

The XMPP protocol evolved from an earlier open-standard protocol called Jabber and was used for Instant Messaging (IM) services as well as Voice over IP (VoIP) control messages. In this last application, XMPP competes with the Session Initiation Protocol (SIP).

When XMPP extensions are used for publish/subscribe services, they are useful for IoT/M2M data applications. However, like JSON, they are human-readable and verbose—perhaps even more verbose than JSON due to the XML roots. This may make it difficult for a small sensor to encode XMPP directly, but a communications device could make the necessary conversion from raw sensor data.

In XMPP, binary files and content can be encoded (using base64 conversion of the binary data to text) and sent using XMPP, but this is likely to be more overhead than is desirable for IoT/M2M applications.

Back to Table of Contents

5.4 » Gateways

In most low-cost sensors—even newer ones that speak IP—it is difficult to provide the data encoding and decoding capability within the sensor. Often, the sensors use short-range communication paths—either wireless or wired—to a device with more computing capacity to actually encode the data and transmit to a remote server.

THE GATEWAY IS A GOOD LOCATION IN THE COMMUNICATIONS PATH TO IMPLEMENT THE DATA ENCODING.

This device may be a unit serving a single sensor and associated application. More often, a gateway is a product with multiple short-range wireless and wired connections to local sensors and a long-range wireless or wired connection to the remote servers.

For example, gateways used in home-automation applications communicate with sensors using Bluetooth, ZigBee, and Wi-Fi, and to the remote servers with cellular or wired-Ethernet connections.

The gateway is a good location in the communications path to implement the data encoding as well as security best practices, with software agents that take the raw information from the sensors and encode the data in the formats described above.

5.5 » Application Servers

The remote data is transmitted to application programs running on the servers that may be dedicated to the task of processing that data—whether it is streaming data or message oriented.

IOT/M2M DEPLOYMENTS SHOULD CONSIDER TAKING ADVANTAGE OF NEWER METHODS LIKE CLOUD COMPUTING.

Typically, these servers are deployed in data centers on the customer premises or in data centers. The programs on the servers receive the data and process them for the specific business action of the IoT/M2M application. This may include storing the data in traditional databases, filtering for erroneous information, alerting when the information is outside pre-determined bounds, displaying the data, reports, etc. The needs vary greatly.

Often, remote devices—even those that are transmitting lightly—cannot tolerate server downtime for any significant duration. Processes and network infrastructure to automatically balance the loads on redundant servers—including at multiple sites—are critical.

For large-scale deployments, the application servers must literally be running continuously with high availability and processing redundancy (including geographic redundancy), particularly for mission-critical applications. With the projected growth of the IoT/M2M market, this will place

Back to Table of Contents

5.5 » Application Servers

an immense burden on servers and data centers. This creates a capital and operation cost of systems, physical site maintenance, power distribution, cooling, etc.

The choice of which server platforms, operating systems, programming languages, etc., is entirely dependent on the entities deploying the IoT application. Traditional IT departments have all the relevant expertise to make these decisions.

However, in most cases where massive growth is expected to occur, IoT/M2M deployments often should consider taking advantage of newer IT methods like Cloud Computing and data traffic reduction methods such as Fog Computing.

Cloud Computing

In recent years, a new phrase called "Cloud Computing" or simply "the Cloud" has been coined to describe the systems that allow processing and storage of information and data in extremely large data centers for a fee. Cloud vendors provide the ability and flexibility to start and stop computing, storage, and networking resources based on the specific needs of the customers and applications using these cloud services.

This has transferred the need for entities and corporations to maintain their own physical hardware, data centers, and data networks, etc., to the cloud providers—eliminating traditional operational burdens of physical site maintenance, electrical power management, environmental conditioning, and system redundancy.

The specific compute, storage, and transport requirements for the cloud customers can then be adjusted fairly dynamically to conform to the needs of the applications being executed. The latest techniques and software for managing the large amount of data can be applied to the data gathered from the devices in the IoT/M2M applications.

These compute elements, storage, and data transport are, of course, provided for a fee—the charges can vary, but can often be high for large-scale applications and large numbers of device deployments.

Fog Computing

The volume of data gathered from a large number of sensors and devices could overwhelm the communications path (transmission and connectivity) or the remote storage capacity and systems that process the data at the customer sites.

ONE SIGNIFICANT ADVANTAGE OF FOG COMPUTING IS THE CONCERN ABOUT SECURITY.

While cloud solutions alleviate this problem, the cost could become very expensive—particularly for streaming applications. Often, a general approach is a "transmit everything and process in the cloud" implementation.

Back to Table of Contents

5.5 » Application Servers

However, if actions based on the data must be processed in real-time or near-real-time, it may be better to process or filter the data remotely—at the device or elsewhere hierarchically in the data flow before it gets to the remote storage. This remote processing and filtering has been termed "Fog Computing."

Fog computing is not without its issues and concerns. If the filtering removes essential information that could be better processed at a central site (such as the cloud) to determine patterns, its use could be a weaker application.

Sometimes, the specific filter used at the remote device may need to evolve or change—thus the devices must be programmable or configurable to the required degree, increasing the complexity of the overall solution.

One significant advantage of fog computing is the concern about security—good security practices can be implemented farther away from the central servers, where a device (or groups of devices) that have been compromised could result in less damage to the overall application deployment.

It also reduces the transport costs of sending a lot of data—much of which may be meaningless, repetitive, or simply not needed—on metered transports where the transport of a large set of data could be expensive.

Back to Table of Contents

SECURITY FOR THE INTERNET OF THINGS

Chapter 6

Back to Table of Contents

In her **keynote speech** at the Consumer Electronics Show in January 2015, the US Federal Trade Commission Chairperson Edith Ramirez noted "any device that is connected to the Internet is at risk of being hijacked." Whether that device is a smartphone, an automobile infotainment system, an automated diabetes monitor, or a GPS-guided farm tractor, specific protections for security of Internet of Things and machine-to-machine devices and applications must be built into the entire solution.

Traditional financial and consumer markets have been targets for misuse of information stored on their systems—including personal credit information, identify theft, misuse of credit cards by unauthorized persons, personal privacy violations, and loss of corporate intellectual property. The financial losses sustained by these security breaches are in the billions of dollars. While attempts have been made to criminalize such nefarious activities, they continue to occur with increasing frequency and are a serious problem for governments, businesses, and individuals.

Business deploying IoT/M2M solutions for their customers and themselves will be held responsible for protecting data and devices, as well as corporate proprietary information. Recent media reports of security compromises in the medical and automotive industries have shown that aspects of such device deployments can be used for purposes other the applications for which they were designed.

This chapter covers basic requirements of security implementations and the different methods commonly used to increase the overall security of IoT/M2M data and applications.

6.1 » Privacy and Security

In the context of IoT/M2M, privacy is concerned with ensuring that data access is limited to the appropriate and authorized parties only. While using tools such as data encryption is an important part of this process, it is just one part of the puzzle, and there are other mechanisms and methods to protect privacy (although not just for IoT/M2M applications):

- Physical access security (for example, secured entrances to data centers).
- Security training (to employees on how to secure computers and devices and to understand data safety).
- Intrusion detection (for systems that process and store the data).
- Software updates (to implement the latest versions of software for security fixes).

U.S. FEDERAL TRADE COMMISSION CHAIRPERSON EDITH RAMIREZ NOTED "ANY DEVICE THAT IS CONNECTED TO THE INTERNET IS AT RISK OF BEING HIJACKED."

Individuals have an expectation of privacy with regard to their personal data, and it is crucial for businesses to consider and implement relevant security methods. In particular, financial and medical industries have specific governmental regulations that govern their products and services in their respective markets. The new IoT/M2M implementations that companies in these industries are deploying may have special testing and certification requirements—particularly in regard to security and privacy issues.

Back to Table of Contents

6.2 » Security Objectives

There are four overall security objectives that must be met for IoT/M2M security implementations:

- Authenticated sender and receiver
- Sender and receiver accessible
- Trust in the data content
- Confidentiality of information

Authenticated Sender and Receiver

In any data connection, it is important for the sender and receiver of information to be authenticated to each other—regardless of whether the device is the sender (for remote data gathering and transmission) or the receiver (for data and control messages from the server). As a security principle when transmitting data, the device must ensure that it is sending its information to the correct server, and when receiving data and control messages, it must ensure that the information is coming from the correct server.

Sender and Receiver Accessible

In any network, the sender and receiver must always be accessible when needed. If the network is not functional, or the server is not executing the correct programs to receive the data, the purpose of the application may be lost. Mission-critical applications, such as automatic crash notification or medical alerts, may fail to work properly if the connection is not reliable. The lack of communication itself means a lack of security.

Trust in the Data Content

The accuracy in the content of the transmitted data is essential—if a device does not encode and transmit data correctly, or the connection is not error-free, the quality and accuracy of the data becomes suspect. Even good data becomes unreliable, and business actions that are taken on the content of the data may not be appropriate.

Mission-critical information is particularly important to keep as error-free as possible. The cost of business actions taken on receipt of incorrect data may be high.

Confidentiality of Information

Finally, the confidentiality of the information must be maintained. Only the correct recipient should have access to the transmitted data, since it may contain proprietary or confidential information. Indeed, privacy laws in many countries require extra care with information regarding individual citizens—for example, in the US, the Health Insurance Portability and Accountability Act (HIPAA) provides specific rules for individually identifiable medical information.

Back to Table of Contents

6.3 » Security Issues for IoT/M2M

Security risks can be recognized and understood, and the implementation of security methods should be incorporated in the IoT/M2M device and software associated with that application. The nature of these new deployments brings new complexities to creating secure solutions.

Most obvious holes in security can be resolved quickly and efficiently. In general, the potential for problems can be managed with confidence in the chosen security methods. However, it is vital to recognize that risks cannot be *completely* eliminated, and there is no single security solution for all possible security requirements for all applications.

Thus, it is critical to assess the level of security implementations that are appropriate for different kinds of data. This assessment must be done early—*during the design of the application, not as an afterthought once many devices have been deployed*!

Before choosing how to secure the application, there are a number of issues to be considered:

- Authenticating presence on multiple transport networks
- Authorization for multiple types of services.
- Scaling to manage the large number of devices in IoT/M2M solutions.
- Automation for application functionality.
- Long lifecycles for deployed devices and applications.
- Implementing security updates in remote devices.

Multiple Networks

Some IoT/M2M devices operate in more than one transport network or technology for redundancy or hybrid solutions. In these devices and solutions, security may be more of a concern in one network compared to the others. For example, a short-range wireless technology such as Wi-Fi can have quite a different security threat vector and potential for breaches compared to a long-range cellular service.

Multiple Types of Services

Applications and devices may be using multiple services, where the required authorizations for allowing a device to access a particular service may differ from one application to another. The authentication mechanisms may also differ, and developers must minimize the risk of a less secure service authentication system from allowing a device to be compromised.

Scaling Growth

In IoT/M2M deployments, there are predictions of explosive growth in the near future—billions of potential devices within the next 5 to 10 years. Thus, in any application where a security problem exists, the overall problem could be greatly magnified by the large numbers of devices that may be affected. This could result in network and data security issues that are difficult to solve, since replacing all the compromised devices could be extremely difficult, perhaps impossible.

Back to Table of Contents

6.3 » Security Issues for IoT/M2M

Automated Functionality

In many IoT/M2M applications, the data is acted upon by automated programs that process the received data and take business actions based on the content. If the transmitted data is compromised, any simplistic responses or automated functions to that compromised data could cascade into difficulty. If some set of devices transmit excessively due to a program error, the servers processing that incoming data could overload and not provide a response to the devices. Simplistic retry algorithms in the devices may create a data storm as a result.

Long Lifecycles

Unlike handsets used by people who change them every few years on average, IoT/M2M devices—particularly in industrial applications—may be deployed for many years and operate relatively continuously over that time. Often, the devices use electrical power rather than batteries (unlike handsets that shut down when battery energy is depleted), and the IoT/M2M devices could continue to use the networks for years. Devices with compromised security could stay operational for lengthy periods.

WHEN A DEVICE SECURITY BREACH IS SUFFICIENTLY CRITICAL THAT THE DEVICE PROGRAMMING MUST BE UPDATED, THE ABILITY TO RE-PROGRAM THE FUNCTIONALITY REMOTELY IS VITAL.

Remote Updates

It is essential to plan and design for device updates over-the-air (OTA). When a device security breach is sufficiently critical that the device programming must be updated, the ability to re-program the functionality remotely is vital. The devices may be in inaccessible locations or a large number of devices must be modified rapidly.

Back to Table of Contents

6.4 » Risk Management and Assessing Impact of Breaches

For some data, the issue of security may not be as critical. For example, if an IoT/M2M application device is collecting temperature information from a residence for monitoring (not control) purposes, the security needs for this data is not as high as a device that collects and transmits credit card information.

Thus, the effort and level of security implementations and methods necessarily differ for these two examples. One may require anonymizing the data source for simplicity and privacy, and the other may require data encryption to prevent unauthorized access to the data.

In all IoT/M2M deployments, it is important to assess the potential for damage caused by a security breach, and implement security solutions accordingly. Ask the following questions (among others):

- If a single device is compromised, can it be used to compromise other devices? The data transport used by that application? The remote application servers? That entire application?
- If an application is compromised and misused, what impact does that security event have? Is it life-threatening to one individual? To more than one individual? An entire population in a region?
- Can a data content breach cause financial harm to an individual? More than one individual? The entire set of people depending on a particular IoT/M2M application to function well?
- How quickly can the specific breach or intrusion be detected? Is it using a well-known target mechanism (such as might exist in a widely used cellular device operating system)?
- Can a compromised device or set of devices be isolated from the application rapidly?

The opportunities for implementing security best practices occur at different points and differing capabilities in the IoT/M2M data chain. As suggested in Figure 6, IoT/M2M developers should assess the opportunity for implementing security best practices (authentication, encryption, etc.) at every point during the design of the application.

For example, the source of data could be a sensor—these are not likely to be compromised easily, since they are so specific to their function, but they still need protection. Because of the simplicity of such sensors, it is often difficult to implement a security solution for them.

However, a compromised sensor could be used to inject false data into the application, where an incorrect action could be taken by a server or human at the remote end of the chain.

A more complex source device, such as a multi-technology gateway connecting to multiple types of sensors, or a cellular modem, offers more opportunity—both for breaches to occur, as well as a location in the chain for implementing a good security solution. For example, a gateway device could have the compute capacity to implement strong encryption algorithms—securing the content further along the chain.

Back to Table of Contents

6.4 » Risk Management and Assessing Impact of Breaches

In general, the "further towards the device" that security best practices can be implemented, the less impact a security breach can have on the overall application.

Each business and its IoT/M2M application implementations will require its own risk assessment to determine the relevant security needs—organizations have to understand the trade-offs they make upfront. It is simply impossible to determine all possible methods by which all such applications could be compromised.

Even if we could determine all possible threat vectors for a particular application, the cost of designing detection and preventative measures to counter every threat might be prohibitively expensive for that application. The best we can do is understand and minimize the risk as best as we can upfront and design the devices and application processes to be as easily updatable as possible.

EVEN IF WE COULD DETERMINE ALL POSSIBLE THREAT VECTORS, THE COST OF DESIGNING PREVENTATIVE MEASURES TO COUNTER EVERY THREAT MIGHT BE PROHIBITIVELY EXPENSIVE.

While server programs and accessible elements of the data chain can be updated more easily, the ability to re-program the devices using OTA updates is key to ensuring that the impacts of security breaches can be contained and repaired.

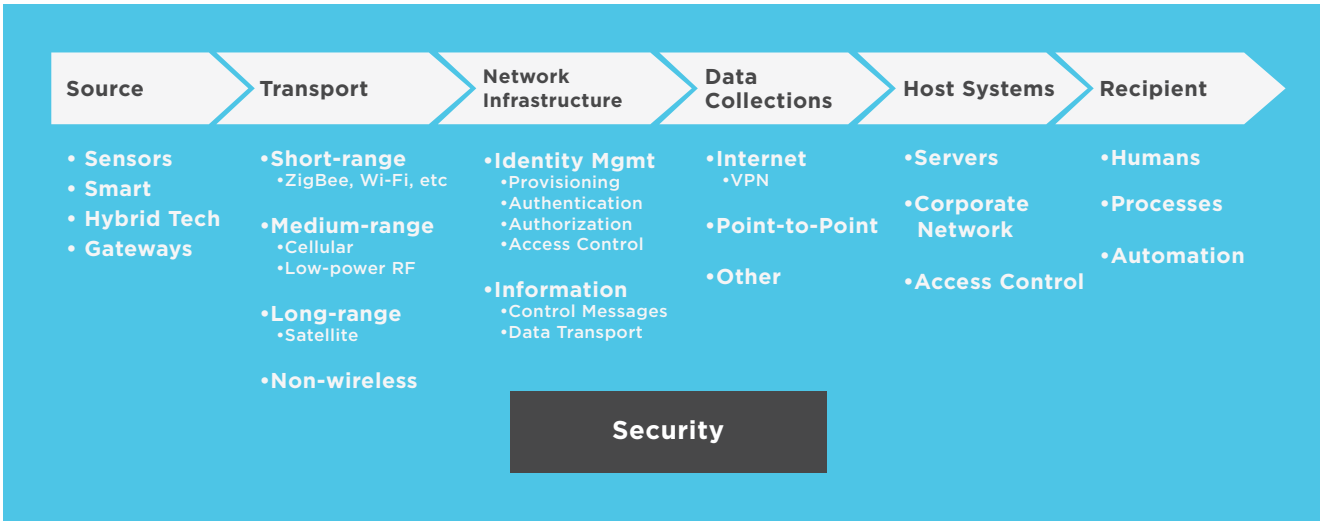


Figure 6. Data Security Flow

Back to Table of Contents

6.5 » Encryption as an IoT Tool

One of the most basic technological tools to secure the content and data in an IoT/M2M deployment is to encode the data so that only the authorized recipient (whether program or human) can decode the data.

After the data is gathered and transmitted by the remote device (or is sent by the server to the device), the content can be encrypted at various points along the network and also when the data is stored.

The basic goals of encryption are to provide:

- **Proof that the sender is valid**—Encryption can make the data's source irrefutable. Techniques such as electronic signatures on a document can be a sign of irrefutability. Proof of who sent data is crucial so that a hacker doesn't steal a session and then pretend to be that user; this is called spoofing.

- **Proof that data was not altered**—Encryption functions can be used to ensure that a change to the data renders the content unusable.

- **Proof that data cannot be read by a third party**—Encryption protects data from being read in transit or upon receipt, except by someone with the correct decryption method.

Data encryption can protect the content in each of these areas to different levels, depending on the need and the specific type of encryption that is used.

Weaknesses in Encryption

No encryption method is perfect—depending on the computing power available at a particular location or the time used by the encryption method, the algorithm may be weak or strong. Strong algorithms may seem impossible to break, but applying enough computing resources to the task could reveal weaknesses that allow the data to be decrypted by unauthorized systems or people.

Indeed, bugs may be discovered in the method itself, or in the particular software implementation. A recent example is the Heartbleed security bug in OpenSSL discovered in 2014. This affected about 17% of the world's web servers and potentially allowed encrypted data to be read. Patches were made to OpenSSL, and a majority of web servers have since been updated.

THE HEARTBLEED SECURITY BUG IN OPENSSL, DISCOVERED IN 2014, AFFECTED ABOUT 17% OF THE WORLD'S WEB SERVERS.

Back to Table of Contents

6.5 » Encryption as an IoT Tool

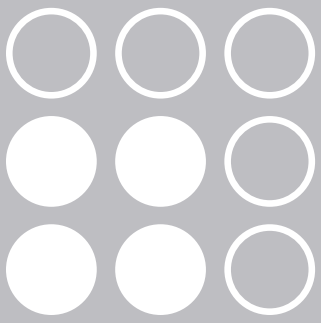
Choice of Encryption Algorithm

It is beyond the scope of this book to describe or recommend a particular encryption algorithm—the specific requirements of the IoT/M2M application or the computing power available at a place in the data chain may drive a preference for a particular algorithm.

Security experts can provide guidance for selecting an approach and should be consulted during the design of the application.

The Information Technology departments at each company may also have specific encryption and security requirements—for example, the use of Virtual Private Networks (VPN) to transport data into its servers for processing and storage.

Back to Table of Contents



IOT SCALABILITY AND ALTERNATIVE TECHNOLOGIES

Chapter 7

Back to Table of Contents

Over the years, the predictions for growth in the Internet of Things and machine-to-machine markets have been staggering:

- 2010, IBM: "A world of 1 trillion connected devices" by 2015.
- 2011, Ericsson's CEO, Hans Vestberg: "50 billion connected devices" by 2020.
- 2013, ABI Research report: "30 billion" by 2020.
- 2013, Morgan Stanley report: "75 billion devices connected to the IoT" by 2020.
- 2014, an Intel infographic: "31 billion devices connected to Internet" by 2020.
- 2014, ABI Research updated report: "41 billion active wireless connected devices" by 2020.
- 2015, Gartner Research: "4.9 billion connected things in use in 2015 … and will reach 25 billion by 2020."

Although the specific predictions and the numbers differ, what is remarkable is that the numbers predicted for 2020 have been consistently extremely high over the years. The markets are experiencing explosive growth around the world, and the numbers are still performing at what Gartner calls the "peak of inflated expectations" in its well-known "Hype Cycle" diagrams. The **Gartner Hype Cycle** showed the Internet of Things had hit the peak of this curve in 2014, so we appear to finally be moving beyond the hype into reality.

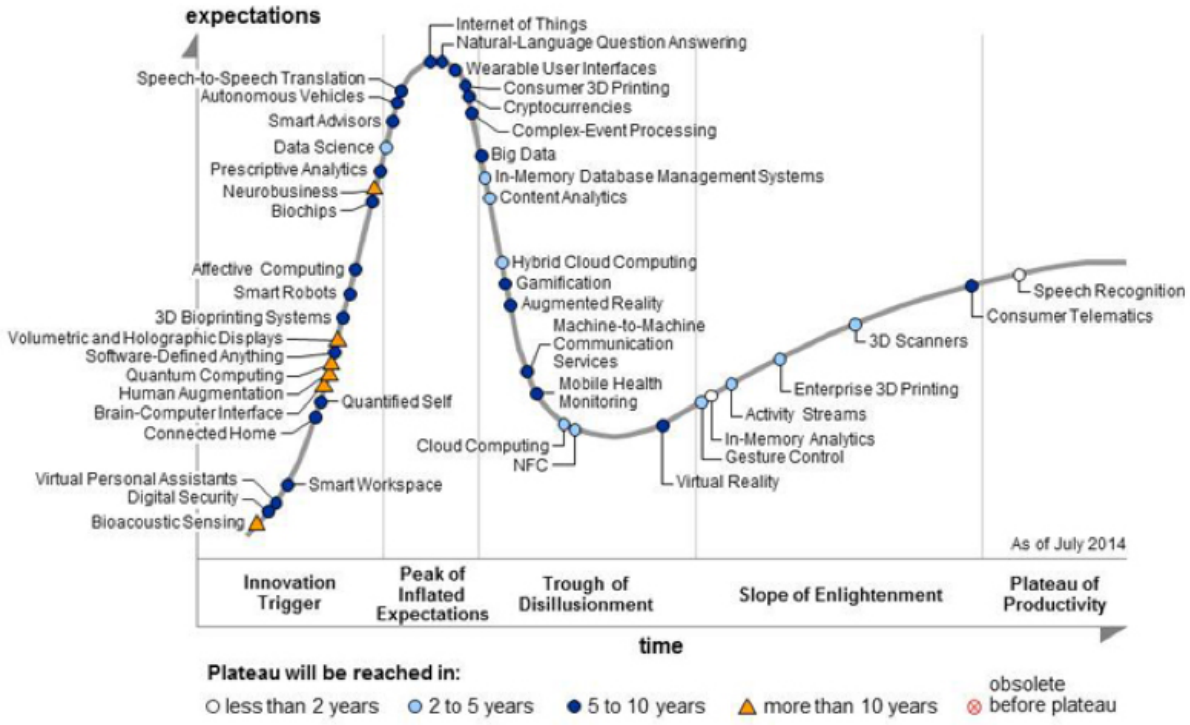


Figure 7. Gartner Hype Cycle, August 2014

Even if the huge numbers forecasted are inaccurate by large percentages, or even a factor of 10, they still represent enormous growth. Indeed, the estimated number of connected devices by 2020 exceeds the projected population of the entire planet by many multiples.

Back to Table of Contents

This explosive growth needs to be managed and planned, if we are indeed going to come close to the predictions for what these markets and industries can do for all of us. Furthermore, cellular is not likely to be the dominant data transport for these large volumes—indeed, it may account for less than 10% of the total devices deployed in IoT/M2M applications.

All of this anticipated growth in the IoT/M2M market will bring new challenges:

- Scaling for the growth in the numbers of devices and applications.
- Providing effective security solutions for the content and solutions (as discussed in the previous chapter).
- Storing the data and providing rapid analysis for action.
- Deploying new wireless and wired connectivity technologies for the increased traffic.
- Managing the connectivity and device "subscriptions" for large numbers of devices.

Therefore, this chapter will also briefly review some of the alternative technologies that are likely to be used for large-scale IoT deployments.

7.1 » What Is Scalability?

In the context of IoT/M2M, scalability is the ability to grow the application, the solution, and platform to keep up with the projected growth in the number of devices, the data traffic from these devices, the applications servers that process and store the received data, the real-time (or near-real-time) streaming data alert systems, the pattern and predictive analytics, etc.

Essentially, this is the ability of the IoT/M2M ecosystem, both for any given application and all such applications in general, to grow at the same rate as the predictions—to make them a reality rather than hype. The demand for IoT/M2M applications, devices. and services will continue to grow exponentially, and companies with connected devices will need to scale their resources accordingly.

Solutions for managing the application must be scalable and designed for growth. For example, most IoT/M2M platforms let the customers rapidly provision the cellular devices for service at volume. Requests are not sent in by humans; rather, automated systems make the requests, and automated systems process the requests.

The Growth Stall

Many companies run into difficulty after deploying their first few hundred or thousand IoT/M2M devices—sometimes, rarely, after tens of thousands of devices. This is not totally surprising, because planning for scalability is difficult and involves many factors, both technological and business-related.

Sometimes, systems and processes simply reach design capacity, and it is time-consuming and costly to change the architecture of the solution or add capacity. Or the cost of operations becomes higher than expected or planned for, which has a deep impact on smaller companies and startups that are resource-constrained. Even seemingly simple tasks such as generating end-user bills and invoices can place unexpected burdens on organizations when scaled up.

7.1 » What Is Scalability?

The key issue for businesses caught in this growth stall is that planning for growth was secondary to getting their products and services launched. This is quite common, but it doesn't have to happen. Successful organizations plan for the entire application lifecycle, from development to operation to scaling to end-of-life.

SUCCESSFUL ORGANIZATIONS PLAN FOR THE ENTIRE APPLICATION LIFECYCLE, FROM DEVELOPMENT TO OPERATION TO SCALING TO END-OF-LIFE.

How Big Can IoT Resource Requirements Grow?

The predictions for deployed numbers of devices are clearly enormous numbers. This has created a need to change some of the resources used for IoT/M2M applications.

Even before the IoT needs became evident, the number of computer systems on the public Internet had increased to the point where the Internet address and numbering method called IPv4 had been exhausted some years ago. The approximately 4 billion possible IPv4 addresses had essentially been used up, as discussed in chapter 2.

And, with the ever-increasing number of IP devices, including cellular smartphones that need an IP address, it is no longer possible to consider using stop-gap measures such as Network Address Translation (NAT), which were introduced for the Internet, for IoT/M2M devices.

Thus, IPv6, which was introduced to increase the number of potential addresses, is a requirement for all future deployments. In theory, this range is large enough that it is unlikely to get exhausted for millennia.

Computing resource can also be scalable—particularly if the device traffic and application processing can be stored and processed. New databases technologies have been deployed that are far more expandable than the traditional databases used in the past three or four decades for data processing.

7.2 » Cloud Computing Revisited

As mentioned in previous chapters, cloud computing technologies have provided a scalable solution for storing and processing the data gathered by IoT/M2M devices.

Since the numbers of devices are growing rapidly, systems to process the data must grow equally quickly. Adding capacity at private data centers is not easy for most companies, since purchasing the physical space, providing for additional power and cooling, increasing the network throughput, installing the computing systems, etc., can take significant effort and time.

Cloud computing suppliers excel at this task–it's their business to provide the compute, network, and general facilities for exactly this growth purpose. Customers using cloud services can "spin up" resources as needed, in step with the IoT/M2M application growth.

7.3 » End-of-Life Management

Yes, many IoT/M2M devices have an end-of-life that must be managed. The service period is generally much longer than the typical period we expect for electronic devices today, particularly for industrial applications. But, once the end-of-life of a device, or all devices within an application, is reached, their removal from service must be managed, to avoid tying up resources.

IT IS IMPORTANT FOR DEVICES TO HAVE AN ABILITY TO BE TURNED "OFF"—PERMANENTLY OR TEMPORARILY—WITH CODE IN THE FIRMWARE AND SOFTWARE OF THE DEVICE.

For example, in cellular networks, devices have a number that identifies them to the network for operational, accounting and authentication purposes. In CDMA, this is the Mobile Identification Number (MIN) or Mobile Directory Number (MDN); in GSM, this may be the International Mobile Subscriber Identity (IMSI) or the Mobile Station ISDN (MSISDN).

These numbers are often from an allocated range, or number pool, and are a resource that must be managed–ideally the numbers are re-used when devices are removed from service permanently.

Devices removed from business service may still have a presence on the networks and impact overall network performance if they are still electronically operational. For example, cellular devices used in automotive applications can be removed from service but could still attempt to "register" on the cellular network every time the vehicle is turned on and off.

Thus, it is important for devices to have an ability to be turned "off"–permanently or temporarily–with code in the firmware and software of the device. This would allow the application servers to effectively remove the device from service, and in the case of permanent removal, allow the device resources (such as numbering) to be re-used for other devices or applications.

7.4 » Selecting Alternative Technologies

When building scalability into an IoT/M2M deployment, selecting appropriate network connectivity is crucial. However, this decision is largely dependent on the type of application. The first question to be resolved is whether the application is fixed or mobile.

Fixed Location Versus Physically Mobile Applications

For simplicity, IoT/M2M applications can be classified into two categories: those that are fixed in one location and those that are physically in motion, while providing the function of the application. These two categories have differing characteristics that affect the specific network selection and implementation for the transport of data from the devices.

In physically mobile applications:

- The devices are installed on moving objects to provide the functionality.
- They physically move from one place to another during the normal operation of the applications.
- During this operation, they often traverse multiple service boundaries (for example, cellular switch boundaries).

In fixed location applications:

- The devices are installed at a single location.
- They generally do not move during the normal day-to-day operation of the applications (although they could be re-installed at some other location during their lifetime).
- During this operation, they are generally in a single service boundary.

Whichever of these two categories the implementation falls into will drive the selection of the network and communications path for the application.

In physically mobile applications:

- Using some form of long-range wireless network is natural and required.
- In this category, using cellular or satellite networks is quite common.
- For some applications that must transmit while traversing service boundaries, the technology must be a Wide Area Network (WAN) with mobility management.

In fixed location applications:

- The devices often use wired networks in deployments where easy wiring solutions are available.
- Wireless networks are also used, however, since network wiring may not be convenient or available.
- The solutions may be hybrid: using short-range wireless to reach a WAN gateway that uses a cellular or wired connection to connect to the servers.

7.5 » Connectivity Options

The range of available data transport technologies for IoT/M2M devices is varied, and new options are becoming available. When planning for scalability, it's important to understand current choices and what's on the horizon.

Fixed location devices are often wired. This could be with a Local Area Network (LAN) such as Ethernet using IP protocols. Older deployments used dial-up telephone lines to reach a remote server directly or connected to the Internet, and cable modem connections are also used where available (also using IP protocols).

For short-range data transmissions, where using a wired solution may not be practical, wireless technologies such as Bluetooth, Wi-Fi, ZigBee, etc., are quite popular. These are common industry standards for which low-cost implementations of the wireless radio and their protocols are available in integrated circuits. The low cost of these short-range wireless technologies enables using them directly within sensors.

FOR SHORT-RANGE DATA TRANSMISSIONS, WHERE USING A WIRED SOLUTION MAY NOT BE PRACTICAL, WIRELESS TECHNOLOGIES SUCH AS BLUETOOTH, WI-FI, ZIGBEE, ETC., ARE QUITE POPULAR.

These short-range wireless technologies are generally quite limited in range—from a few meters to a few hundred meters. If the data needs to go further, the short-range communication is typically sent to a gateway modem that then connects to the servers using cellular, cable, or some other IP network transport.

For medium ranges (for the wireless transport), typical implementations of IoT/M2M solutions use cellular for communication to a nearby tower (generally within a few miles) that then backhauls the data into the Internet or a remote server.

When cellular is not available, such as on ocean-bound ships or remote geographies with low human presence, long-range satellite data services provide a global reach for devices to communicate to a distant server program for that IoT/M2M application.

Wired Data Connections

Wired connections are typically used for fixed location IoT/M2M applications. For reaching a server, the cost of the transport is "shared" with general Internet access. For many device deployments, this is often a very low-cost solution, since the ISP generally does not charge a metered rate—i.e., the fairly low amount of data sent by the devices at a fixed location does not trigger a high cost.

With wired connections, the overall service requires an ISP service or another LAN. The quality of the service and general network availability also depends on the ISP—if it is not able to provide continuous service, some mission-critical applications may experience problems with outages.

7.5 » Connectivity Options

Cellular and Satellite Connectivity

Service coverage and availability for cellular and satellite are generally excellent. Cellular service is available wherever people live and along major highways in most countries. We discuss types of cellular connectivity more thoroughly in chapter 3.

If cellular is not available in a truly remote location, such as an ocean-bound ship or in mountain regions, the coverage from satellite data services is excellent, although some of the satellite services may have relatively higher latencies (the time for a data packet to traverse end-to-end) than other technologies. Coverage inside urban canyons with tall buildings is usually difficult for satellite data services, but this is where cellular services can excel. If required, a hybrid cellular-satellite device, with multiple radios, can provide truly global data access.

A HYBRID CELLULAR-SATELLITE DEVICE, WITH MULTIPLE RADIOS, CAN PROVIDE TRULY GLOBAL DATA ACCESS.

In both cellular and satellite, the cost of the radio can be high relative to the rest of the device, and the radios general consume substantially more electrical energy to transmit—the data range is relatively long.

For example, it would not be practical to equip low-cost sensors or simple IoT application devices with cellular or satellite transports. These would be far better served by the short-range wireless technologies such as Bluetooth or Wi-Fi, etc.

There is one other concern with cellular technologies: the longevity of deployment is driven by smartphone users. Thus, the technologies evolve relatively rapidly and devices using cellular services must be replaced after some period of time—longer than typical smartphone user turnover, but less than older traditional wired technologies.

Short-Range Wireless

In many IoT/M2M applications, short-range wireless data technologies such as Bluetooth, Wi-Fi, ZigBee, etc., are in common use. For certain consumer IoT applications that only transmit to a nearby smartphone belonging to an individual—such as fitness application devices—using Bluetooth and low-power Wi-Fi are common choices. These allow the users to gather data via applications on their smartphone. The need to further transmit the data to central servers for processing is not a paramount requirement but can be done with ease from the smartphone, if necessary.

Short-range wireless is also relatively low-energy, so battery-powered devices are easily designed and deployed. In some low-use applications, the battery may last for months or years before it needs to be replaced. This is a key advantage over cellular and satellite applications that require far more frequent energy replacements (for example, using rechargeable batteries that might last a few days).

7.5 » Connectivity Options

For many home and business applications, a gateway modem that provides one or more short-range wireless technologies for deployed sensors and low-power, low-cost, data transmitters are ideal for a number of IoT/M2M applications. The gateway communicates to the application servers using cellular or wired ISP connections.

Low-Power Wide Area Network (LPWAN)

Recently, the need for low-cost, low-power applications that offer longer transmission ranges (between 2 to 20 miles) has seen the development of a number of new technologies and services competing for the large-scale deployment of consumer and industrial IoT/M2M devices and applications.

These are called LPWAN technologies and include commercial service networks deployed by Sigfox and its licensees in some countries in Europe (and a few cities in the US as of this writing). Similar (but not identical) data transports for IoT/M2M include the technologies developed and deployed by Onramp Wireless and nWave, and the open standards efforts by the LoRa Alliance, that appear to be geared to private network solutions rather than public access data networks.

These technologies currently use unlicensed wireless spectrum at various frequencies. Thus they experience congestion and have technology and data rate limitations that are solved in different ways. For some LPWAN transports, the data rate and message size is low enough that a simple approach to overcome the congestion problems is possible, although the data is mostly one-way (from the device) for low-power use. Others provide more complex data encoding to reach the tower networks, leading to more expensive radio solutions that may work for some IoT/M2M solutions, but not necessarily all.

Finally, the International Telecommunications Union (ITU) is working on a set of standards that extend the 4G LTE technology for use with low-power, and potentially low-cost, radios. This technology is tentatively called “LTE-M” (where M stands for Machine) and could compete strongly with the other LPWAN technologies being deployed today. Since LTE-M is not available today, and not likely to be available for another two to four years, there is an opportunity for these LPWAN technologies to gain a foothold.

THE INTERNATIONAL TELECOMMUNICATIONS UNION IS WORKING ON A SET OF STANDARDS THAT EXTEND THE 4G LTE TECHNOLOGY FOR USE WITH LOW-POWER, AND POTENTIALLY LOW-COST, RADIOS.

7.5 » Connectivity Options

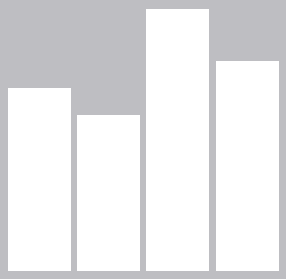
Fifth Generation (5G) Cellular

Although it is perhaps much too early to discuss 5G in any depth, it is important to note that the requirements for the 5G cellular services include a need to accommodate large-scale deployment of IoT/M2M applications and devices.

Overall, the 5G requirements are to provide:

- The transport of 1000x more data volumes that the smartphone users are using today.
- More than 10x to 100x the number of connected devices in use today.
- Dramatically lower latency (for end-to-end data packets) below a few milliseconds.
- Projected 10x longer battery life for low-power devices—up to 10 years.

The ITU has projected that the first set of 5G standards will not be available until 2020, although work is already underway to discover what solutions and technologies—including new radio encoding protocols—will be needed to meet the 5G requirements.



IOT ANALYTICS

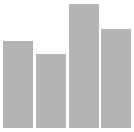
Chapter 8

Back to Table of Contents

This chapter will describe what is meant by analytics in the context of Internet of Things and machine-to-machine applications. Looking at data and finding meaningful patterns in that data is the basis of analytics. These patterns could describe the state of the data, predict an outcome, find correlations between variables, project trends in the data, and more. Analytics is used in many aspects of business, from marketing to risk management. In this chapter, we'll discuss analytics as it relates to IoT/M2M data.

Over time, IoT applications can generate vast amounts of data—this is part of the Big Data revolution that is much hyped in the media. For example, the Aeris IoT/M2M network manages traffic from nearly one billion IoT events each day. The more IoT/M2M data points being transported, the more sophisticated analytics are needed to understand the patterns.

New ways to process and store computing data has made it possible to apply analytics to business problems faster and at a greater scale than ever before. Successful organizations take advantage of these tools and analyze the data their IoT/M2M deployments collect so they can gain insights into everything from how to streamline their manufacturing processes to how satisfied their customers are.



8.1 » IoT Data and Analytics

IoT/M2M devices usually report data in constant streams—these must be processed in real-time or near-real-time. This is a shift from traditional programming frameworks that open a file, read it, process the contents, and then close it. Today, real-time analytics are possible on streaming IoT/M2M data.

Another key to IoT/M2M analytics processing has been the development of open-source distributed storage and distributed processing frameworks such as Hadoop. This allows processing of very large data sets over computer clusters. Hadoop can also be deployed as a cloud-computing service by smaller organizations.

With these scalable technologies capable of managing streams of Big Data, businesses can use various types of analytics to better understand and use the data their IoT/M2M sensors collect.

Back to Table of Contents

8.2 » Types of Analytics

Analytics can be grouped into four broad categories: Descriptive statistics, diagnostic analytics, predictive analytics, and prescriptive analytics.

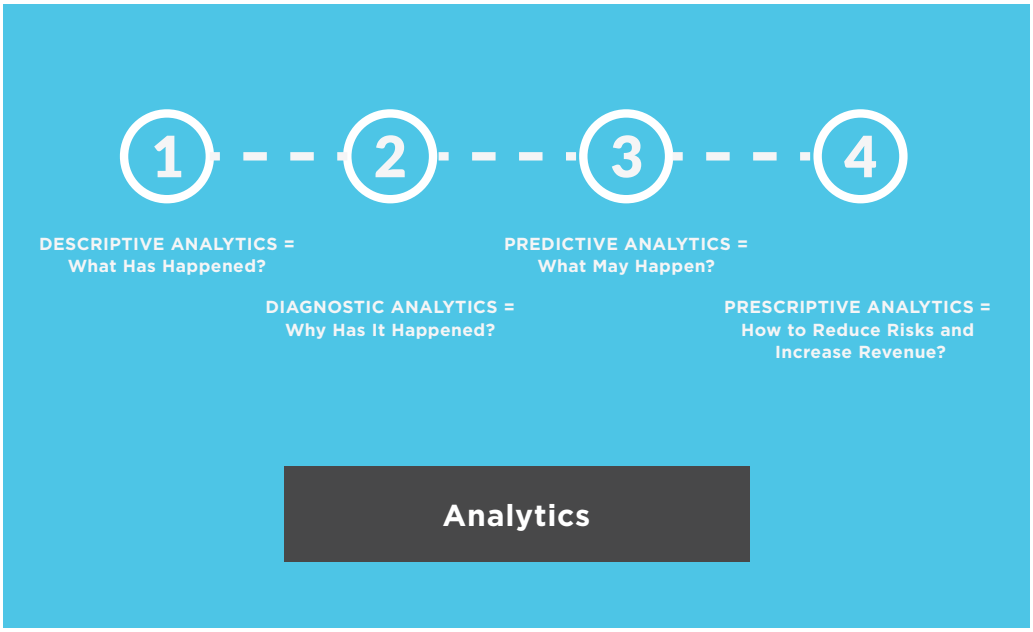


Figure 8. Types of Analytics

Descriptive Analytics

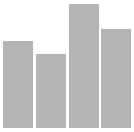
Descriptive analytics, also called descriptive statistics, gives a numerical representation of the data that is on hand right now. It provides a way to express in absolute, unambiguous terms, a quantitative measurement of the current state. This analysis can draw conclusions of data from the past as well.

In a broad sense, descriptive analysis answers these questions:

- What happened?
- How often did it happen?
- How reliable was it?
- How accurate was it?

Knowing the current status of the IoT/M2M data provides a baseline to compare future states against. It is possible to compare basic data from the past to present, tracking progress. Descriptive analytical tools can be as simple as tracking website traffic or more complicated such as cluster analysis in market research.

Back to Table of Contents



8.2 » Types of Analytics

Diagnostic Analytics

This type of analytics is often merged with descriptive analytics, and together they can give data greater interactivity. Where descriptive analytics asks "what happened?", diagnostic analytics asks "why did this happen?" The diagnostic tools can be applied to the data to look for the root causes behind the results observed in the original data.

Usage-based insurance implemented with vehicle telematics is one example of descriptive and diagnostic analytics in action. This type of car insurance bases the driver's insurance premium rates on behavior that is tracked via a GPS-enabled cellular transmitter in the car. The distance a person drives, when, and where are tracked, and this data is used to calculate the insurance cost.

Predictive Analytics

Prediction is one of the main reasons that businesses use analytics in the first place: predictive analytics provides a means of projecting what will happen next, based on what has happened in the past. By finding patterns and trends in the data, it may be possible to predict future results.

While assuming future behavior will be the same as past behavior isn't always the case with, for example, the stock market or consumer purchasing habits, machine behavior is generally highly predictable. In a factory, vibration and temperature data broadcast from an IoT/M2M-connected device can indicate, with a high degree of accuracy, when a machine needs preventive maintenance. Businesses can use predictive analysis in their own IoT/M2M deployments as part of supply chain management and manufacturing processes to increase efficiencies.

Predicating Brake Maintenance in Truck Fleets

Brake balancing in trucking fleets is a complex and expensive maintenance issue. Without regular maintenance, the risk of a truck jackknifing is high. It takes a highly trained technician a great deal of time to check for the combination of brake temperature and pressure on a truck fleet to know when to make an adjustment to the vehicle's brakes.

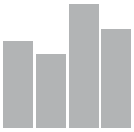
But in **Michael Lawrence-Smith's study**, "Cooperating Artificial Neural and Knowledge-Based Systems in a Truck Fleet Brake-Balance Application," he describes how machine-learning techniques used predictive analysis to improve brake maintenance. These computer-aided systems have a 90% success rate at predicting when to replace brakes, resulting in an annual savings of at least $100,000 for larger trucking companies.

Prescriptive Analytics

Prescriptive analytics is the logical next step from predictive analytics. It asks what a business should do based on the data collected and analyzed. Prescriptive analytics uses models to both recommend actions and forecast outcomes to reduce risk.

Much as descriptive and diagnostic analytics work well together, predictive and prescriptive analytics work hand-in-hand. As past data is used to calculate future results, prescriptive analytics can be used to make better choices and take advantage of opportunities.

Back to Table of Contents



8.2 » Types of Analytics

Google's self-driving cars, for example, use prescriptive analytics to make countless driving decisions, according to **Data Informed**. The cars communicate with the cloud using IoT/M2M systems to obtain data on traffic and weather, which becomes part of their driving computations. The vehicle's on-board computers apply machine learning to the problem of what a car should do based upon predictions of future outcomes. For example, the car's computer may predict traffic based on the time of day and then determine what route to take and what speed to travel safely.

Connectivity or Communication Analytics

A business operating an IoT/M2M service for its customers must keep track of data and manage billing and usage data. Connectivity analytics can provide this data to an organization for administering an IoT/M2M deployment.

For example, a company has to provision IoT/M2M devices on the network, as well as remove or replace any devices that are operating incorrectly. A problematic device could be trying to authenticate on the IoT/M2M network multiple times per minute, which uses up resources, so it will need to be fixed quickly. Connectivity analytics can immediately pinpoint the problem device so it can be removed from the network or repaired remotely.

AerVoyance for Connectivity Analytics

For greater understanding of connectivity analysis, a business could use Aeris' **AerVoyance**, which provides visibility and insight into its IoT/M2M deployments. Focusing on devices, connectivity, and billing information, this tool helps companies effectively manage their IoT/M2M systems through an intuitive, visual presentation.

AerVoyance identifies devices with connectivity issues and allows zooming in on a specific device, time, or usage activity including data, SMS, voice, sessions, cellular registrations, etc. This allows customers to find specific devices that are behaving abnormally. In addition, there is a summary view of the device portfolio including billing, usage, and status. This type of connectivity analytics tool helps businesses understand their aggregate and average connectivity usage over time and better plan deployments.

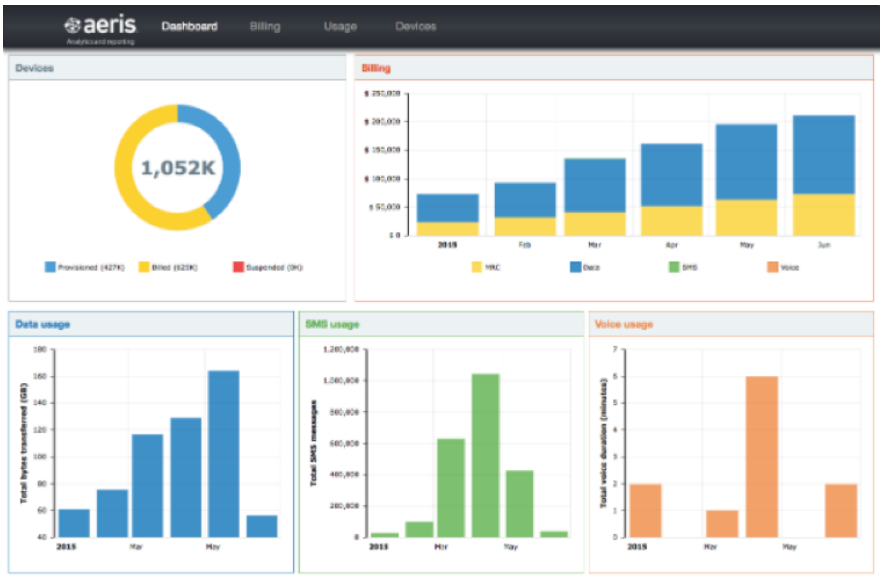
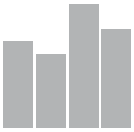


Figure 9. AerVoyance Dashboard

Back to Table of Contents



8.3 » Analytics Tools and Languages

While many different statistical models, tools, techniques, and programming languages can be used in all of these types of analytics, below are few specific ones used to provide analytics functions for IoT/M2M applications.

R Language

R is a programming language for statisticians. Common in academic research, R has also become quite popular with data scientists in the corporate world as business needs for analytics have grown dramatically.

This language can be used to draw graphs and return the numeric results of algorithms run against the data. R has functions to let the programmers run different analytics tools against the data. This makes it easier because it contains packages or functions contributed by different people around the world as open-source software.

MapReduce and Hadoop

MapReduce is a programming model that can be used to query very large data sets, such as those in Hadoop databases. What MapReduce does is take a query and run it across hundreds or thousands of computers to derive an answer to an analytics question. It does this in two steps: map and reduce. The **map** step collects data on each node in the Hadoop Distributed File System. **Reduce** eliminates duplicates and produces the result set.

People can write MapReduce in programming languages like Java. But there also exist specific languages for Hadoop to make the job of creating MapReduce programs easier.

These are just some of the analytical methods and tools available for IoT/M2M solutions. Planning in advance for analysis of the data that will be collected and stored is a crucial part of any IoT/M2M project.

Back to Table of Contents

IMPLEMENTING AN IOT SOLUTION

Chapter 9

Back to Table of Contents

An IoT/M2M deployment has to either increase business revenue or reduce business costs (or both), otherwise there's no reason for a company to pursue it. Either of these objectives can provide a return on investment–it's up to the project manager to determine the specific goals and measurements of this ROI.

Related drivers for IoT/M2M projects can be new regulations and industry requirements, greater efficiencies, more consistent control over processes, predictive visibility into patterns or opportunities, and gaining competitive features that can meet customer requirements. As you build your IoT business model, these factors will weigh differently depending on the product needs and industry.

9.1 » Supply Chain Management

Whether or not the company's business is manufacturing, you will have some level of supply chain management. This is simply planning for the flow of materials and services into and out of the business, managing all the goods required to make your IoT/M2M deployment happen. If your company is building its own IoT/M2M devices from scratch, you'll have a great many materials, parts, and suppliers to account for. If you're assembling devices from readymade components, you need to deal with fewer suppliers. Even if you buy a complete, off-the-shelf device, that still requires sourcing, testing, and managing supply and demand.

9.2 » Cellular Operator Selection

The service provider must be able to deliver several essential requirements for the IoT/M2M project, including low costs, reliable network connectivity, robust service agreements, effective application integration, and flexible rate plans. If they can't deliver on these prerequisites, they are not going to be the right partner.

To help you select the ultimate service provider with the capacity to manage a successful deployment, you may want to ask these questions of any cellular operator during the selection process:

- What are the costs for the entire lifecycle, not just per Kb rates? Make sure you won't be hit with hidden costs from your cellular operator that drive up your IoT/M2M service bill.

- Are they operator-agnostic? Can the service provider expand cellular coverage beyond its own cell towers? Traditional operators only optimize their cellular coverage based on their cost of delivery, and they always prefer to use their own towers, even if the coverage they provide is weak or intermittent. An operator-agnostic provider, like Aeris, can expand coverage where needed, and will offer the strongest signal, regardless of operator, with no interruption in service.

Back to Table of Contents

9.2 » Cellular Operator Selection

- Do they offer remote troubleshooting as well as hands-on support? Cellular operators with remote, real-time troubleshooting capabilities can save you significant costs. Also, an operator with a network operations center support team that deals only with IoT/M2M-related issues is going to be more knowledgeable about your devices and connectivity issues.

- Do they offer a dedicated IoT/M2M network? A network dedicated solely to IoT/M2M traffic won't experience the delays caused by crowds of consumer handsets. The lower latency of an IoT/M2M-dedicated network means you'll be able to rely on mission-critical transmissions to get through the first time.

A NETWORK DEDICATED SOLELY TO IOT/M2M TRAFFIC WON'T EXPERIENCE THE DELAYS CAUSED BY CROWDS OF CONSUMER HANDSETS.

- Do they have APIs for easy integration with your existing systems? Can the cellular operator provide a full suite of free APIs that let you extend the capabilities of your customer-facing applications and back-office solutions, leveraging business applications such as those from SAP, Oracle, and PeopleSoft? These applications are integral elements of enterprise resource planning-based supply chains and are linked to back-office systems with APIs.

- Do they offer pay-per-use as well as per-device billing plans? Can the cellular operator offer rate plans that are flexible enough to meet your needs? When managing IoT/M2M services, it often makes more sense to go with a pay-per-use plan than with a per-device or fixed-data plan. Pay-per-use is most cost-effective for lower-usage device profiles. If your devices have higher-usage levels—10 MB or more—a per-device data plan is your best option.

These are some of the top-level concerns your company should consider when choosing a cellular operator. You'll want to partner with a service provider that suits your business needs and can support your IoT/M2M project over the long term.

Back to Table of Contents

9.2 » Cellular Operator Selection

Operator Support Service Level Agreement

The Service Level Agreement (SLA) you negotiate with the operator defines the scope of your contract with the operator. This is where your business defines its relationship with its network provider, so it's important to specify what will keep your IoT/M2M deployment running.

Things to consider in your SLA include:

- What are the expectations for connectivity? How reliable is the operator's network historically?
- What are the geographic restrictions of the operator's network, if any? Some operators may not guarantee service at all towers or all sections of particular metro areas.
- How long will it take the operator's customer support to acknowledge and then take care of a problem? For example, Aeris' Infinity Support guarantees a five-minute response time, with proactive monitoring and issue identification by IoT/M2M experts, five days a week.

When entering into an SLA, make sure the agreement is realistic, actionable, measurable, calculated, well-defined, mutually exclusive, and completely exhaustive in covering all aspects of the concerned networking services.

Device Certification

Devices must be approved or certified to run on the operator's network. For this certification, the focus is generally on testing the cellular behavior of the device. One example of this might be the behavior of the retry algorithm used by the device if it fails to connect to the application server in your data center. A continuous retry by thousands of your devices at the same time could overload the operator's network. Implementing a random back-off algorithm and testing it prior to certification is a better idea.

Operator certification is also an opportunity to use the application host server software for additional tests that stress the interaction between the device and the server. Unusual scenarios, such as delayed responses from the server (that might be observed during congestion or server scaling), can be used to see if a device handles them gracefully.

Additional certification may be required by standards organizations, regulatory agencies (such as the Federal Communications Commission in the US), or even your customers, particularly if there is end-user integration.

ADDITIONAL CERTIFICATION MAY BE REQUIRED BY STANDARDS ORGANIZATIONS, REGULATORY AGENCIES, OR EVEN YOUR CUSTOMERS.

Back to Table of Contents

9.3 » Normal Operation Considerations

Here are a few of the concerns to be dealt with when IoT/M2M devices are deployed:

- What is the definition of "normal"? What are the baseline transmission patterns and server performance measurements?
- What happens if the IoT/M2M device can't connect to the cloud platform? In addition to having a random back-off retry algorithm, what will the device do with its data? Remember that stale data would be inaccurate when transmitted too late. The device needs to know when to generate an alarm.
- What should a mobile IoT/M2M device do if it loses cell signal? The device needs to know when it is appropriate to hold the data in its queue and retry later.

The range of normal operations will vary for each IoT/M2M deployment, so you'll need to set initial parameters for all aspects of the program. Then you can track performance against this baseline moving forward.

Application Communications Call Flow

This is where the details of the IoT/M2M transmission are agreed upon. Some design issues are:

- Should the device assume there will be a connection when needed or should it be able to queue data for delivery later?
- Will the application "fire and forget" data or will it wait for an acknowledgment? At the network layer, "fire and forget" means to use the UDP protocol for transmission.
- The general call flow is to establish a connection, transmit data, wait for acknowledgment, then disconnect. This is generally a TCP protocol implementation.
- Does the data need to be encrypted? That can increase the amount of data being sent.

Your developers will need to outline each aspect of the IoT/M2M application's call flow, accounting for both standard, predictable behaviors and for outliers.

Back to Table of Contents

9.4 » Customer Support Process

Support staff will need to be trained on the product features and how to operate them, but it's also very important for the support team (especially tier-2 support) to receive training on identifying connectivity issues. This is where a rich set of diagnostic tools from the operator, if available, becomes a huge benefit.

ADVANCE PLANNING PAYS OFF IN A SCALABLE PRODUCT WITH A HIGHER RETURN ON INVESTMENT.

If your tier-2 engineer can log into a portal and see if the device in question has registered on the cellular network and started a data session, then the engineer can observe the recent behavior and can immediately focus the investigation on the root of the problem. Using this observation, the engineer can provide quick feedback to the customer. If these tools are not available, then support sessions can be much slower.

Implementing an IoT/M2M network project requires a great deal of forethought. But this advance planning pays off in a scalable product with a higher return on investment. The next chapter gives an overview of how to manage your IoT/M2M project's lifecycle once you've implemented the solution itself.

Back to Table of Contents

IOT LIFECYCLE MANAGEMENT

Chapter 10

Back to Table of Contents

To plan and deploy a successful IoT/M2M project, you first have to analyze your needs. The IoT/M2M application has to make a strong return on investment, either by increasing revenue or reducing costs. Once you have built a case for making the deployment, then you can outline a full plan that accounts for all the processes, methods, and the design principles for this deployment. After identifying the business needs, you should then evaluate the best practices for your IoT/M2M implementation. Key considerations can include regulatory requirements, operational costs, and customer requirements.

IoT/M2M deployments are long-term investments. Companies offering connected devices and applications must ensure that their solutions are planned with everything from growth through end-of-life in mind. IoT/M2M deployments need reliable, secure, and high-performance connectivity to support device maintenance and troubleshooting, data collection and analysis, and consideration of future upgrades in hardware, software, and connectivity standards.

10.1 » Planning Checklist

An IoT/M2M project manager needs to define the deployment's scope from a business point of view as well as integrating the technical considerations. While you plan for the intensity of the initial deployment activities, you should make sure your company invests in a robust connectivity management platform.

In addition, application and data analytics must be part of your IoT/M2M program. Key stakeholders should understand the fixed and recurring costs, while the IT organization needs to be supportive of the IoT/M2M deployment.

KEY INTERFACE AREAS MAY INCLUDE MANUFACTURING AND SUPPLY CHAIN, IT SYSTEMS, OPERATIONS, SUPPORT, FINANCE, AND REGULATORY APPROVALS.

The project manager must oversee the following considerations:

- Go-to-market business model: How your company will market and sell to customers.
- Supply-chain management: Consideration of device manufacture/assembly, device provisioning on operator's network, testing the communications link at the end of manufacture, and also channel-to-market goals.
- Application communication call flow: How the device will interact with the network.
- Definition of expected usage patterns: Normal operating patterns, over-the-air updates, troubleshooting, dealing with rogue devices.
- Cellular operator selection: Consideration of technologies, geographic coverage, bandwidth, uptime, support, load-balancing capabilities, security, scalability, etc.
- Device operator certification: Are the devices guaranteed to work on the operator's network?
- Operator integration and API integration: How the operator's platform will support your IoT/M2M applications for device management and data collection.
- Operator support Service Level Agreement (SLA) understood and in place: Making sure the service level agreement satisfies all parties involved.

Back to Table of Contents

10.1 » Planning Checklist

- End-to-end application test plan: Testing of the communication links across different RF conditions.
- Customer support process definition and support team training: How you will support your end-customers once the IoT/M2M deployment is up and running.

These are just the basics of what you must plan for as you begin the IoT/M2M lifecycle process. You'll want to work with all the related areas of your business and make sure they are onboard and prepared for their part of the project as well.

The project manager should be aware of how the IoT/M2M deployment impacts other key stakeholders so they can provide input to planning and execution as well as ensure that their own requirements are met. The key interface areas may include manufacturing and supply chain, IT systems, operations, support, finance, and regulatory approvals.

10.2 » Lifecycle Management Phases

A guide to managing the lifecycle of an IoT/M2M deployment could be an entire book in itself. As a brief overview, we'll discuss the phases of product design, scaling for growth, product testing, installation, operations, and end-of-life.

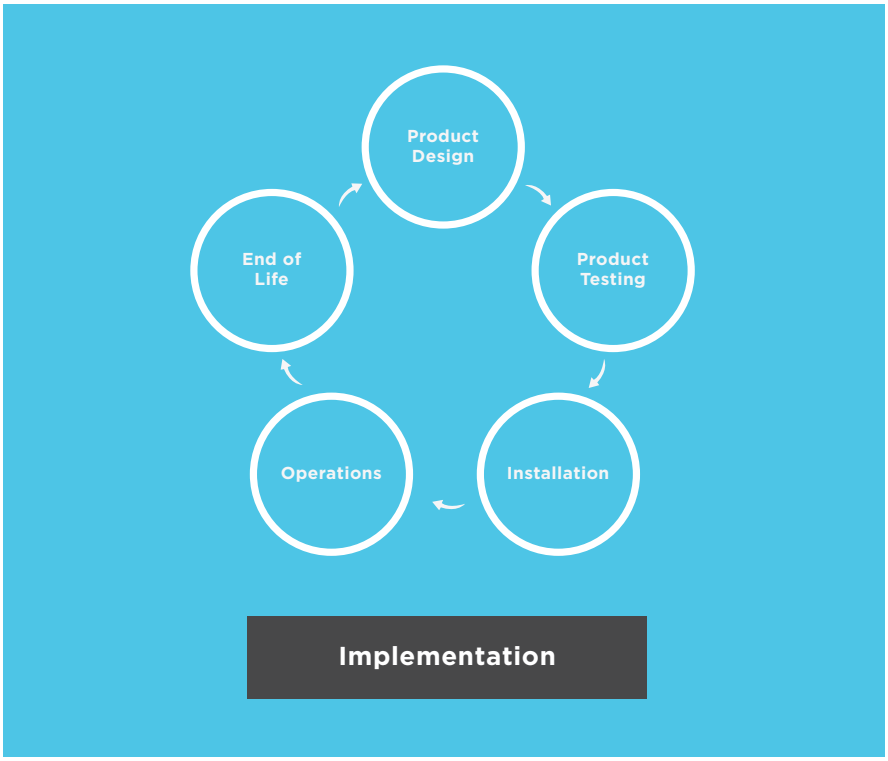


Figure 10. IoT/M2M Product Lifecycle

Back to Table of Contents

10.2 » Lifecycle Management Phases

Product Design

A crucial decision to begin with is whether your company will be manufacturing the IoT/M2M device itself, assembling it from parts, or purchasing a complete device from a supplier. You must understand your business needs and whether a customized or off-the-shelf solution is the best fit.

Whichever route you choose, you'll have to plan for the whole lifecycle of the IoT/M2M product. Consider the longevity of your product–many IoT/M2M devices are deployed for long periods, operating 24/7, and possibly for more than 10 years. These devices may also be in rough or remote locations, so it's important to think about the physical durability of the design. Consider how the device will function under whatever variety of temperatures, humidity, and vibration specifications it may be subjected to in the field.

WHETHER YOU'RE BUILDING AN IOT/M2M SOLUTION FROM SCRATCH OR FROM MODULES OR BUYING A COMPLETE DEVICE, IT ALL COMES DOWN TO HAVING THE RIGHT SENSORS

Finally, what data do you need to collect and transmit? Different sensors are relevant for different needs. Some can be used in combination to provide unique results. Whether you're building an IoT/M2M solution from scratch or from modules or buying a complete device, it all comes down to having the right sensors. Refer to chapter 4 for an overview of commonly used sensor categories.

Scaling for Growth

If you want to deploy large numbers of devices, your manufacturing or assembly process and server and network capacity must all be designed for the expected growth. Without planning for growth, the volume of data from devices may eventually overwhelm your host server systems–for example, for databases used for information storage, as well as any real-time applications that analyze the transmitted content for processing.

Businesses that don't plan for scalability at the beginning may get caught in a "growth stall" after deploying their first hundred or thousand devices. Successful organizations include growth in their lifecycle management. For a discussion of scaling up IoT/M2M deployments, see chapter 7.

Product Testing

Comprehensive testing is not only important to test the core functionality, it's also essential to ensure the IoT/M2M solution operates correctly over your selected cellular network. This usually begins with lab testing in a controlled environment. In the lab, engineers typically ensure that the device has a strong signal level that will provide a high-quality cellular connection and achieve consistent data rates across the air interface.

This is fine for basic functional testing, but negative test cases should be introduced to force poor signal quality conditions that are sometimes present in the field and lead to low bandwidth and

Back to Table of Contents

long latency conditions. Applications requiring near real-time communication may not operate correctly under these conditions, so it is critical to consider this when designing the application, selecting the technology (for example 2G, 3G, and LTE each have different latency characteristics), and when testing the application.

Field-testing a small device population can provide an indication of how the devices will behave after launch. Providing you select an appropriate sample size and deployment environments, if you see a problem with 1% of your devices then, within a margin of error, you are likely to see the same problem in 1% of your devices after launch. If you can address the problem and get the problem device count within an acceptable margin, then there should be a high level of confidence after launch that the percentage of devices with problems will remain manageable.

Installation

Placing the IoT/M2M devices in the field, connecting them to the network, and provisioning them are key steps to installing a deployment. Each of these aspects has its own requirements. Field installation, for example, may be done by third parties who aren't necessarily experienced with your product, so your developers will need to write detailed installation procedures. Automated post-installation verification methods and clear troubleshooting procedures for installation failures can also be very useful.

"PROVISIONING" MEANS REGISTERING THE DEVICE ON THE IOT/M2M PLATFORM, SUCH AS THROUGH A CLOUD-BASED APPLICATION.

Connecting the devices to the network and provisioning them must be done correctly before data-collection can begin. "Provisioning" means registering the device on the IoT/M2M platform, such as through a cloud-based application. Once devices are provisioned and on the network, the IoT/M2M application can begin functioning as designed.

Operations

Once an IoT/M2M deployment is installed, the devices, their applications, and the data generated have to be managed. For the devices, this includes managing device traffic, analyzing device performance, troubleshooting problem devices, managing rogue devices, and updating devices remotely, as well as billing, suspending, and canceling devices. For applications, your operations plan must account for secure and reliable collection, storage, analysis, and publishing of the IoT/M2M data produced by all of your devices.

During initial deployment–after a set of devices are installed and used by end-users–it is a good idea to verify that the device behavior (transmission patterns, retry algorithms, etc.) is as expected. Early detection of problems is important, so that serious problems can be fixed before deploying larger numbers of devices, thus reducing the cost of repair or replacement.

You may also want to track device transmission patterns and server performance measurements to establish a baseline for future monitoring. Later, any data transmission pattern change from

Back to Table of Contents

the baseline can be an early indicator of problems. These early measurements also provide guidance for scaling–the server performance load can be recorded for future expansion planning. During normal operation, the device transmission patterns should continue to be monitored and measured against the patterns recorded during development and initial deployment.

Over-the-Air Updates

IoT/M2M devices will generally need at least one or more application firmware updates during their lifetimes. Planning for this download is critical to reduce costs. Frequently, devices are deployed in remote or hard to reach locations, where sending out a technician would be expensive.

OTA CAPABILITIES CAN ADD COST TO THE DEVICE BUT WILL BE EASILY JUSTIFIED THE VERY FIRST TIME AN IOT/M2M DEVICE NEEDS UPDATING.

Businesses should add over-the-air (OTA) firmware update capabilities to their IoT/M2M devices. This may require extra hardware components; for example, sufficient memory to hold at least two "images" of the firmware–one active and one new–as well as code for the download and verification of the images, restoration of old images, etc. This can add cost to the device but will be easily justified the very first time an IoT/M2M device needs updating.

Developers can also deploy an "incremental" update capability–using software that allows sending smaller chunks of firmware. This uses less memory in the hardware, reduces the update transmission time, etc., but it can increase the coding support and complexity required in the device.

Device Removal and End-of-Life

Products may have a limited lifecycle, whether due to technology, demand, or other business factors. The entire deployment or specific devices may need to be disabled. Businesses need to plan how to effectively remove devices, transition to new technology, or shut down the deployment altogether. Billing issues, customer support, and logistics all must be accounted for.

Back to Table of Contents

10.3 » Pitfalls to Avoid

In general, addressing a problem early, as soon as it arises, is less expensive, particularly if a device recall is required or a large number of devices must be physically serviced by a technician for a replacement or firmware update. Some major problems can be avoided if troubleshooting capabilities are designed into the device, the servers, and the complete IoT/M2M application and product. This can avoid major cost issues after deployment.

IoT/M2M deployments are not immune to disasters and factors beyond human control. When a service disruption is inevitable, the IoT/M2M enterprise must guarantee immediate fixes to the problems affecting its customer base. IoT/M2M solutions can experience both front-end and back-end service failures at times, so you'll need plans in place to deal with the worst. Your network operator should be able to provide assistance and share its expertise to minimize downtime, while your customer support staff addresses end-user concerns directly.

Back to Table of Contents

AN ANATOMY OF AN IOT DEVICE

Personal Fitness Example

3rd Party Services

Examples:
Aggregator services
Analysis services

Cloud Sync Service

aeris
www.aeris.com

Smart Phone App

Automotive Example

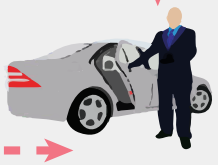
3rd Party Service Providers

Examples:
Charging spots
Usage-based insurance
Navigation
Internet radio

Public Safety

Cloud Automotive Service Delivery Platform

Call Center



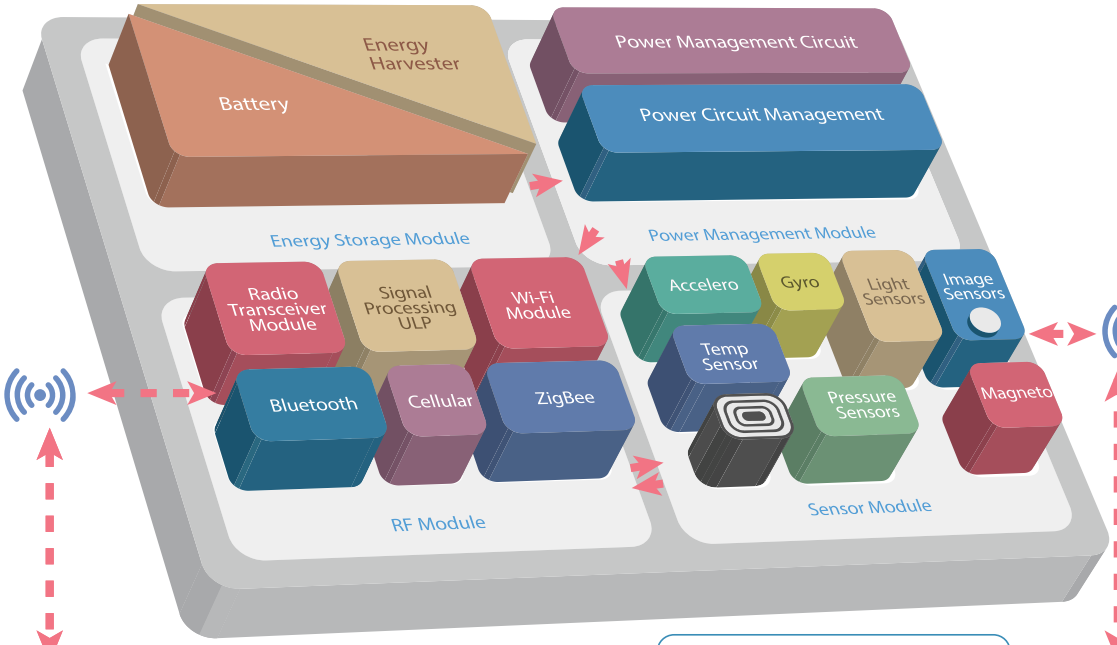
Smart Phone App

ENERGY STORAGE MODULE

- Li-Ion Battery
- AAA/AA Batteries
- Energy Harvesting

POWER MANAGEMENT MODULE

- Ultra-capacitors
- Microbatteries
- PMU



RF MODULE

- Signal Processing Unit
- Wi-Fi
- ZigBee
- Bluetooth LE
- Radio Transceiver
- Duplexer
- BAW
- Cellular (GSM, CDMA, LTE)

SENSOR MODULE (wireless or wired)

- Accelerometers
- Magnetometers
- Gyroscopes
- Acoustic Sensors
- Pressure Sensors
- Humidity Sensors
- Temperature Sensors
- Proximity Sensors
- Image Sensors
- Light Sensors
- Gas RFID Sensors
- Micro Flow Sensors

Back to Table of Contents

DIRECTORY OF IOT/M2M INDUSTRY TERMS

This glossary includes key terms of the Internet of Things and machine-to-machine communications industry, including wireless and cellular technologies spanning many different markets. It is updated to the most recent terminology and usage.

Back to Table of Contents

Directory of IoT/M2M Terms

1xEV-DO	1 times Evolution Data Optimized (used in ANSI-2000 CDMA).
1xRTT	1 times Radio Transmission Technology (used in ANSI-2000 CDMA).
2.4 GHz	A short-range wireless band commonly used in wireless technologies such as Wi-Fi, Bluetooth, and ZigBee.
2G	The second generation of GSM cellular technology that improved performance by adding to the cellular radio spectrum to help solve coverage issues and drops in signal due to urban obstacles. It was also the turning point in moving from analogue transmission methods to digital, adding digital encryption and paving the way for cellular data usage.
3G	The third generation of GSM cellular technology, offering substantially improved data transfer rates over its predecessor, 2G. While the original release of 3G used the UMTS method, improvements have been made to increase capacity and data speeds with additional protocols including HSPA.
3GPP	Third Generation Partnership Project (GSM family of technologies).
3GPP2	Third Generation Partnership Project 2 (CDMA family of technologies).
4G	The fourth generation of GSM cellular technology and the latest upgrade to the GSM network, providing greater data transfer speeds. 4G is also referred to as LTE.
6LoWPAN	A communication protocol that compresses Ipv6 packages for small, low power-devices to let them communicate within the Internet of Things.
802.11ah	New Wi-Fi protocol that uses sub 1 GHz license-exempt bands as opposed to conventional Wi-Fi that operates in the 2.4 GHz and 5 GHz bands.
868 MHz	License-free RF band commonly used for short-range applications such as thermostats, burglar alarms, and industrial uses.
92 MHz	License-free RF band commonly used for short-range applications. The low frequency allows for better penetration through walls and obstacles; however it has a low data transfer rate.
AAA	Authentication, Authorization, and Accounting (see also RADIUS).
ACaaS	Access Control as a Service.
Acceleration Sensing	A MEMS concept referring to the increase in movement of an object from one point to another along a straight line or axis. Typical applications include remote control, pointing devices, gesture recognition, fitness monitoring equipment, etc.

Back to Directory

Directory of IoT/M2M Terms

Accelerometer	A tool that measures changes in gravitational acceleration in the unit it may be installed in. Accelerometers are used to measure acceleration, tilt, and vibration in many devices.
Access Control	A system that determines who, when, and where people are allowed to enter or exit a facility or area. The traditional form of access control is the use of door locks, but modern access control may include electronic systems and wireless locks. Access control may also apply to cybersecurity.
Access Control as a Service (ACaaS)	A recurring fee-based system where a facility manager outsources electronic access control to a third party. Each facility need not maintain a dedicated server.
Access Point	A Wi-Fi node that allows users entry to a network, typically a LAN.
Active Sensor	A sensing device that requires an external source of power to operate.
Actuator	A device that introduces motion by converting electrical energy into mechanical energy in an electromechanical system. (An actuator may also stop motion by clamping or locking.) A dynamo is an example of an actuator.
ADAS	Advanced Driver Assistance Systems.
Additive Manufacturing	The industry-specific term for 3D printing, involving building products by adding layers rather than the traditional technique of removing material via milling.
Addressability	The capacity for an entity to be targeted and found. To be addressable, an entity must be uniquely identifiable, meaning that it must be associated with something–typically an alphanumeric string–that is not associated with anything else that exists within that system.
Advanced Driver Assistance Systems (ADAS)	Digital features incorporated into vehicles to enhance driver safety and performance. ADAS functionality includes digital vision for lane departure warnings, blind spot detection, radar for collision avoidance, and V2V communication for multiple vehicles operating near each other. The data and connectivity integral to ADAS transforms vehicles into IoT devices.
Advanced Encryption Standard (AES)	The specification for encryption of electronic data established in 2001. Operates on a public/private key system, and planning for key management is an important aspect when implementing AES encryption.
Advanced Message Queuing Protocol (AMQP)	An open-source standard for business message communication. Main features include message orientation, queuing, routing, reliability, and security.
Amazon Web Services (AWS)	The name given to a collection of remote computing services, offered by Amazon.com, that combine to make a cloud computing platform.

Back to Directory

Directory of IoT/M2M Terms

Ambient Assisted Living (AAL)	Intelligent systems to assist the elderly and others with daily care activities, often through IoT technology. Application fields are security (for example, observation), functionality (such as automated light switches), and even entertainment.
Ambient Intelligence (AmI)	Sensor-filled environments that interpret and react to human events and activity, and learning to adapt over time, the environment's operation and services change based on that activity.
AMPS	Advanced Mobile Phone System, an analog cellular mobile system using FDMA.
AMQP	Advance Message Queuing Protocol.
Android Wear	An open-source platform that extends the Android system to wearables. The SDK includes an emulator.
Anomaly Detection	A statistical technique that determines what patterns are normal and then identifies items that do not conform to those patterns. Unlike simple classification where the classes are known in advance, in anomaly detection the users don't know what they are looking for in the data.
ANSI-136	American National Standards Institute Standard 41, for TDMA cellular.
ANSI-2000	American National Standards Institute Standard 41, for CDMA2000 cellular.
ANSI-41	American National Standards Institute Standard 41, for control signal messaging on SS7.
ANSI-95	American National Standards Institute Standard 41, for CDMA cellular.
AP	(Wireless) Access Point.
API	Application Programming Interface.
Application Programming Interface (API)	A collection of commands and protocols used to interact with an operating system, device, or specific software component. In IoT, an API lets the developer access the functionality of a device or sensor, such as a thermometer's readings. APIs can be public or restricted to authorized users only.
Application Software	Programs that enable specific, end-user actions. This means the software uses the given potential provided by computers to form an application. Examples include Microsoft Word (text editing), Adobe Photoshop (image editing), and many other programs.
Application Specific Sensor Nodes (ASSN)	Integrating sensors and sensor fusion in a single device, ASSNs have a built-in intelligence to cope with the complexity of applying multiple sensors to a specific problem such as augmented reality, navigation, positioning, and more. Bosch Sensortec's BNO055 is an example of an ASSN.

Directory of IoT/M2M Terms

Arduino	A single-board microcontroller used for prototyping without having to deal with breadboards or soldering. The software to operate an Arduino is free and open source.
ARP	Address Resolution Protocol. A communication protocol used to convert an IP address into a physical address. This way, computers can communicate with each other, despite only knowing each other's IP addresses, by sending an ARP request that informs them about the other computer's MAC address.
AT Commands	Attention commands, developed by Dennis Hayes, that are used to set data connections. The set of short string commands allow developers to set up calls with a modem, as well as perform far more complex tasks. For an example of an AT command set, take a look at Telit's 3G module, the HE910, AT command directory.
Audio Profile	Hardware profile used with Bluetooth applications that include custom AT commands and functionality dedicated to wireless streaming of audio. Examples include A2DP, which allows for streaming of audio to devices such as speakers, where as an audio gateway profile allows for two-way audio communication used in devices such as headsets.
Augmented Entity	A physical entity is represented by a virtual entity on the digital level. An augmented entity combines the two and stands for any combination of the two entities.
Automated Identification and Mobility (AIM) Technologies	A group of technologies that are used to identify, store, and communicate data. An example would be a barcode, though there are many technologies in this area that are used for different services and are often used in combination.
AWS	Amazon Web Services.
BAN	Body Area Network.
Band	A range of frequencies used by a technology for communication purposes. For example, the 2.4 MHz band is used for Wi-Fi and Bluetooth communication.
Beacons	Low-cost devices that communicate with smartphone apps indoors, without the need for GPS. Beacons use BLE and are key enablers for the smart retail category, triggering messages as consumers pass through locations or near products.

Back to Directory

Directory of IoT/M2M Terms

Big Data	Data sets so large that they cannot be used with traditional database tools. Big data often requires massively parallel computing resources to access, curate and analyze. Big data analysis techniques are crucial to such disciplines as spotting business trends and simulation.
BLE	Bluetooth Low Energy.
Bluetooth	Short-range wireless technology standard which operates on the 2.4 MHz band. Bluetooth can be used for sending both data and audio, with popular uses including wireless headsets and cordless keyboards. Bluetooth devices can be set up with different hardware profiles to help perform specific tasks, for example audio adapter, audio headset, serial, and keyboard profiles.
Bluetooth 4.0 (BLE)	The latest iteration of Bluetooth, also called Bluetooth Low Energy (BLE). It offers lower power use for portable devices and new profiles including Bluetooth Mesh, a Bluetooth topology that allows devices to be connected together, sending/repeating commands from the hub to any connected device. Apple's iBeacon is an example of a BLE application, and BLE as many potential uses for IoT devices.
Bluetooth LE (BLE)	Bluetooth Low Energy.
Bluetooth Low Energy (BLE)	The latest iteration of Bluetooth, also called Bluetooth 4.0. It offers lower power use for portable devices and new profiles including Bluetooth Mesh, a Bluetooth topology that allows devices to be connected together, sending/repeating commands from the hub to any connected device. Apple's iBeacon is an example of a BLE application, and BLE as many potential uses for IoT devices.
Body Area Network (BAN)	A wireless network of wearable computing devices and physiological sensors, which may even be embedded inside the body. A BAN may also be referred to as a WBAN (wireless body area network) or a BSN (body sensor network). A key use case for BANs is e-Health applications.
Brick	Slang term for accidentally rendering a device inoperable by changing its configuration or shorting one of its circuits. Used as a verb, as in "what do I do if I brick my Raspberry Pi?" The inert device sits there like a brick.
Bring Your Own Device (BYOD)	Enterprise term recognizing that people are bringing their own Wi-Fi enabled devices into the corporate network.
Brownfield	Brownfield describes the problem and the process of having to consider already existing systems when implementing new software systems.
BS	Base Station. The radios and other equipment at the cell sites that are used to communicate with the cellular devices.

Back to Directory

Directory of IoT/M2M Terms

BSC	Base Station Controller. The equipment that consolidates and controls multiple BS sites (usually, more than one BS is attached to a BSC).
BTS	Base Transceiver Station. This is a machine that enables wireless communication between user equipment, for example a mobile phone or a computer, and networks like the GSM network. The data is received through an antenna and is then processed and transmitted by the BTS to create a wireless connection.
Business Logic	Used to describe processes that are necessary to enable or execute communication between an end user and a database/server. These processes decide how data is transmitted, transformed, or calculated. This does not include the display of data or task-specific commands. It serves as a basis, consisting of algorithms, code, etc.
BYOD	Bring Your Own Device.
CAN	Controller Area Network.
CAN Bus	A message-based, multi-master serial protocol for transmitting and receiving vehicle data within a Controller Area Network (CAN). Sometimes written as "CANbus," the CAN Bus connects multiple Electronic Control Units (ECUs) also known as nodes. Designed initially for automotive applications in 1983, the CAN Bus can be adapted to aerospace, commercial vehicles, industrial automation, and medical equipment.
Card Not Present (CNP)	The type of credit transaction where the merchant never sees the actual card. CNP has the obvious potential for fraud but is vital for newer services such as contactless mobile payments.
CDMA	Code Division Multiple Access. Digital cellular phone service method that separates multiple transmissions over a finite frequency allocation using Spread Spectrum techniques (concept invented and patented by Hedy Lamar).
Cellular Modem	Allows a device to receive Internet access over the cellular mobile networks. Devices can also be configured to remotely connect to a server or device to enable off site access and data collection.
Cellular Router	Allows connected devices to access servers and devices by making an IP connection through the cellular mobile network. Routers allow for multiple devices to be connected and controlled, while built in Open VPN, IPSEC, PPTP, and L2TP, and offer extra device and data transfer security to keep your information safe.

Directory of IoT/M2M Terms

Chief IoT Officer (CIoT)	One of the CxO class of corporate officers, the CIoT coordinates the integration of IoT into the enterprise. Successful CIoTs will break down silos between disciplines such as big data, data analytics, security, communications protocols, etc.
CIoT	Chief IoT Officer.
Class 1 Bluetooth	Offers a greater wireless data transfer distance (over 100m, up to 1km) through using greater power consumption (100mW).
Class 2 Bluetooth	Short-range wireless data transmission (10-20m) which has low power consumption of around 2.5mW.
Cloud	Or the Cloud, meaning cloud computing. The name "cloud" comes from the fluffy cloud typically used in Visio-style network diagrams to represent a connection to the Internet.
Cloud Communications	Communication services being provided by third parties that can be accessed and used through the Internet. The program Skype is one well-known cloud communications application.
Cloud Computing	An approach where information technology capacities (such as storage or applications) are separated from the individual computer and are supplied through the Internet (or an Intranet-based service) at the user's demand. The "as-a-Service" moniker is sometimes used for cloud computing services, such as Software-as-a-Service, Platform-as-a-Service, and Infrastructure-as-a-Service. The backend for many IoT devices may be delivered via the cloud.
Cloud Orchestration	The automated management of a cloud. This includes all services and systems that are part of the cloud as well as the flow of information.
CNP	Card Not Present.
CoAP	Constrained Application Protocol. This software protocol is used in small electronics devices and serves as the interactive communication between those devices.
COBie	Construction Operations Building Information Exchange.
Cognitive Vehicles	A term coined by IBM to describe vehicles that will learn from the behaviors of drivers, occupants, and vehicles around them, plus be aware of the vehicle's own condition and the state of the surrounding environment. A cognitive vehicle will thus be capable of configuring itself to a specific driver, other occupants, and various conditions.

Back to Directory

Directory of IoT/M2M Terms

Communication Model	Communication models try to capture, explain, simplify, and then model communication. One of the oldest and most famous models, the Shannon and Weaver Model, was created in 1949.
Companion Device	In wearables, a companion device requires a parent device, such as a smartphone, to fully operate. The opposite would be a standalone device that can do everything on its own. A companion wearable will typically use Bluetooth to communicate with the parent.
Connected Home	If the devices in a house work interactively and information relevant to residents is accessed via high-speed broadband, it could be called a connected home. This may mean that the refrigerator reports the almost empty milk or that the TV reminds you of your doctor's appointment because it automatically gets this information from the doctor's computer. Related to Smart Home.
Construction Operations Building Information Exchange (COBie)	The COBie approach simplifies the capture and recording of building project handover data, basically by entering things like serial numbers as the project progresses. COBie breaks down the design into Facility, Floor, Space and Zone elements. COBie can be displayed in several interoperable formats.
Controller Area Network (CAN)	In automobiles, a CAN connects Electronic Control Units (ECUs) using a multi-master serial bus (the CAN bus) to control actuators or receive feedback from sensors. ECUs can be subsystems such as airbags, transmission, antilock brakes, or most importantly, engine control. The standard consists of ISO 11898-1 and ISO 11898-2.
COPE	Corporate-Owned, Personally Enabled.
Corporate Owned, Personally Enabled (COPE)	A compromise around pure BYOD, COPE devices allow the user to control much of the data on the device, but the enterprise controls the security model.
Cortex-A	Cortex-A refers to a series of processors from ARM that are equipped with ARMv7 and ARMv8 command sets. They are used for applications that require a lot of processing power, mainly in the areas of mobile handset (smartphones), computing, digital home, automotive, enterprise, and wireless infrastructure.
Cortex-M	Cortex-M is a family of microprocessors developed by ARM which is mainly used in microcontrollers. They range from the cheapest M0 processor up to the Cortex-M4, which is used for effective digital signal control. Applications are found in automotive, gaming, and intelligent consumer products.
CPS	Cyber-Physical Systems.

Directory of IoT/M2M Terms

CR2032	A battery rated at 3.0 volts commonly used in watches, wireless doorbells, and other small devices. Sometimes referred to as a "button cell" or "lithium coin," the battery is shaped like a coin with dimensions of 20mm diameter x 3.2mm height (from which the "2032" is derived). The CR2032 is twice as thick as the CR2016.
Crowdfunding	A novel system for inventors and entrepreneurs to bypass traditional funding methods such as venture capital by raising small amounts from a large group of individual backers. Made popular by sites such as Kickstarter, crowdfunding can act as a pre-ordering system, allowing the project's creator to reduce risk by gaging consumer popularity before production even begins.
Cyber-Physical Systems (CPS)	Systems that combine computer-related and mechanical aspects. A smartphone, for example, combines software, hardware, etc., with a physical device. In general, many mobile or embedded technologies or devices can be called Cyber-Physical Systems, thus applications are manifold. The systems often include some form of sensor which can transfer attributes from the real world to the digital sphere.
Dashboard	A user interface that presents key information in a summarized form, often as graphs or other widgets. Derived from the classic automobile dashboard, the design of the interface depends on what information needs to be monitored or measured.
Data Center	A collective term for the physical site, network elements, systems, etc., that supports computing and network services.
Data Janitor	A subtask of data science concerned with the cleaning up of dirty or duplicative data. Oftentimes the janitor must get data into the correct columns and sort it.
Data Lake	Coined by Pentaho CTO James Dixon, a data lake is a massive data repository, designed to hold raw data until it's needed and to retain data attributes so as not to preclude any future uses or analysis. The data lake is stored on relatively inexpensive hardware, and Hadoop can be used to manage the data, replacing OLAP as a means to answer specific questions. Sometimes referred to as an "enterprise data hub," the data lake and its retention of native formats sits in contrast to the traditional data warehouse concept.
Data Scientist	A job that combines statistics and programming, using languages such as R, to make sense of massive data sets. IoT sensors, for example, create mountains of data, and the data scientist's role is to extract valuable information and detect anomalies.

Back to Directory

Directory of IoT/M2M Terms

Data-Driven Decision Management (DDDM)	An approach to business governance valuing decisions that can be backed up with verifiable data.
Datakinesis	A termed coined by Marc Blackmer, datakinesis occurs when an action taken in cyberspace has a result in the physical world. Industrial Control Systems, for example, are vulnerable to datakinetic attacks where physical equipment such as valves and sensors are compromised and damaged by hackers. Stuxnet is one such example.
DDS	Digital Data Storage. This format is used to store computer data on audio tape. It was developed by HP and Sony in 1989 and is based on the digital audio tape (DAT) format and was a widely used technology in the 1990s.
Degrees of Freedom (DoF)	An engineering concept used in MEMS that describes the directions in which an object can move and generally the number of independent variables in a dynamic system.
De-identification	The stripping away of personally identifiable information from data prior to its use. The process must include the removal of both direct identifiers (name, email address, etc.) and the proper handling of quasi-identifiers (sex, marital status, profession, postal code, etc.).
Demand Response (DR)	The voluntary reduction of electricity use by end users in response to high-demand pricing. Demand response can reduce electrical price volatility during peak demand periods and help avoid system emergencies. An example of DR would be a utility paying Nest to have thermostats turn down air conditioners in empty homes on a hot day.
Device Attack	An exploit that takes advantage of a vulnerable device to gain access to a network.
DG	Distributed Generation.
DIN Rail	A metal rail used for mounting electrical equipment and racks.
Distributed Generation (DG)	Decentralized, modular, and flexible power generation located close to the serviced loads. Distributed microgrids can control smaller areas of demand with distributed generation and storage capacity.
DIY	Do It Yourself. Enthusiasts generally tinker with gadgets or software to improve the functionality or do custom-install projects in their homes.
DNP3 Protocol	An open, standards-based protocol for the electric utility industry with interoperability between substation computers, remote terminal units, intelligent electronic devices), and master stations. Groups of enabled things are organized into namespaces.

Back to Directory

Directory of IoT/M2M Terms

DoF	Degrees of Freedom.
Domain Model	A model that contains all areas and terms related to a certain field of interest. It includes attributes, relations, constrains, acts, etc., that are relevant for a certain task.
Domotics	A combination of "domestic" and "robotics." Also a composite of the Latin domus and informatics, domotics includes home automation systems, home robots, whole house audio/visual systems, and security systems. Domotic devices have the ability to communicate with each other.
Downlink	Abbreviated as DL or D/L, downlink is the process of downloading data onto an end node from a server/target address. In a cellular network this would be seen as data being sent from a cellular base station to a mobile handset.
DR	Demand Response.
DWG	A format for different computer-aided design programs, including AutoCAD. It is used to store two and three dimensional design data and meta data.
EAN	European Article Number. This is used to mark and identify products. Since 2009, it is also called GTIN (Global Trade Item Number). The number is usually found beneath barcodes and consists of up to 13 digits (EAN 13 barcode).
ECU	Electronic Control Unit.
EDGE	Enhanced Data rates for GSM Evolution.
E-Health	Or eHealth, telehealth, telemedicine, and related to mHealth. This is the support of medical processes and applications through information and computer technologies. It can include the gathering and communication of data as well as automated responses of certain devices and processes.
Electronic Control Unit (ECU)	Also known as a node, an Electronic Control Unit is a device, such as a sensor or actuator, that is connected to other devices via a CAN Bus. A vehicle can contain dozens of ECUs for functions such as mirror adjustment, window power, airbags, cruise control, entertainment, and, most significantly, engine control. To form a CAN, two or more ECUs are needed.
Embedded Device Hacking	The exploiting of vulnerabilities in embedded software to gain control of the device.
Embedded Firmware	The flash memory chip that stores specialized software running in a chip in an embedded device to control its functions.

Directory of IoT/M2M Terms

Embedded Software	Specialized programming in a chip or on firmware in an embedded device to control its functions.
Embedded System Security	The reduction of vulnerabilities and protection against threats in software running on embedded devices.
Embrace, Extend, and Extinguish	A strategy associated with Microsoft to defeat open standards with proprietary extensions. Many IoT projects are open source, so this strategy would be anathema to open development.
EMD	Enterprise Mobile Duress.
EMI Protocol	An extension to the UCP (Universal Computer Protocol). It's used to connect to Short Message Service Centers which store, transform, and send short messages.
Energy-Harvesting Technologies	Technologies which use small amounts of energy from close proximity to power small wireless devices. Applications can be found in wireless sensor networks or wearable tech. Energy sources are, among others, sun, wind, or kinetic energy.
Enhanced Data rates for GSM Evolution (EDGE)	This is an enhancement made to 2G GSM networks to improve data transfer speeds and provides downlink speeds of up to 1 Mbit/s and uplink speeds of up to 400KB/s. It builds on available GSM or GPRS standards and is thus easily integrated into the existing network.
Enterprise Mobile Duress (EMD)	Systems designed to detect personnel emergencies within large facilities, such as hospitals or campuses, where determining the physical location of persons in distress is a critical issue. EMD systems are a robust extension of a Personal Emergency Response System (PERS) into the enterprise, focusing on the protection of people from emergency incidents such as violence.
EPCglobal	A nonprofit organization founded by GS1 (former EAN International) and GS1 US (former UCC). It serves to spread, improve, and standardize the Radio Frequency Identification (RFDI) technology and to support communication of gathered data through the Internet.
ESD	Electrostatic Discharge. This discharge can occur if two electrical objects with different electrical charge come in contact with each other. The difference in charge is often due to friction. Sometimes, the short process is accompanied by sparks, as can be seen with lightning. ESD can lead to severe damage to electrical devices (such as generators).
ESN	Electronic Serial Number (in CDMA). Replaced by the MEID.

Directory of IoT/M2M Terms

EtherCAT	A fieldbus system developed by Beckhoff, which allows for real-time Ethernet. It helps to achieve short data update times, accurate synchronization, and low hardware costs, so it can be used specifically for automated or control systems. CAT stands for Controller and Automation Technology.
EV-DO	Enhanced Voice-Data Only (also Enhanced Voice-Data Optimized).
Facility	A structure designed and constructed for a particular purpose, such as a medical facility.
FAKRA	Fachnormenausschuss Kraftfahrzeugindustrie. This is a type of SMB connector used in the automotive industry for connecting coaxial RF connectors which uses snap on connectors.
FDMA	Frequency Division Multiple Access.
Firmware	Programming that's written to the read-only memory (ROM) of a computing device. Firmware, which is added at the time of manufacturing, is used to run user programs on the device.
Firmware Over-the-Air (FOTA)	The process of updating a mobile phone's operating system and software over the network, rather than having the consumer come into a service center for updates.
Fitness Band	A type of activity tracker worn on the wrist, with sensors specifically related to exercise and activity measuring. In contrast to a smartwatch that may include fitness/activity tracking features, a "fitness band" is primarily dedicated to fitness.
Fleet Management (FM)	A broad term referencing a range of solutions for vehicle-related applications. An FM solution is typically a vehicle-based system that incorporates data logging, satellite positioning, and data communication to a back-office application.
FM	Fleet Management.
Fog Computing or Fogging	Also known as fogging, this is a distributed computing infrastructure in which some application services are handled at the network edge in a smart device and some application services are handled in a remote data center–in the cloud.
Form Factor	The physical size, pin-out, and configuration of a component. A family range of module, for example, may include 2G, 3G, and 4G variants to allow PCB designers to design in one module but allow for future upgrades through the product family's road map.
FOTA	Firmware Over-the-Air.

Back to Directory

Directory of IoT/M2M Terms

Galileo	Developed by the European Union and Space Agency, Galileo is a global positioning constellation of satellites which is still in development and will be made up of 30 satellites (27 operations and thee active spares).
Gateway	A link between two computer systems or programs. This way they can share information with each other. The router for your home Internet is one type of gateway.
General Packet Radio Service (GPRS)	A wireless communications standard on 2G and 3G cellular networks which supports a number of bandwidths and provides data rates of 56-114KB/s. As cellular companies move to more advanced networks, GPRS networks may be more cost-effective for IoT networks.
Geofence	A virtual border applied to a physical space. For example, geofencing might be defined around a nursery, and when a mobile device crosses the nursery boundary, an alert is generated. Geofences may be dynamically created and in a telematics application can encompass entire neighborhoods or cities.
Geographic Information System (GIS)	The combination of hardware, software, and data that captures, manages, analyzes, and presents many kinds of geographic data. GIS and location intelligence applications can be the foundation for location-enabled services.
GeoJSON	A dialect of JSON that describes physical places. Features modeled by GeoJSON are points, line strings, polygons, and multipart groups of these types (MultiPoint, MultiLineString, MultiPolygon). Numerous mapping and GIS software packages employ GeoJSON.
Geotagging	The process of tagging a photo, video, or other types of media with coordinates, thus marking it with a location.
GGSN	Gateway GPRS Support Node (see also SGSN).
GIS	Geographic Information System.
Global System for Mobile communication (GSM)	This is the most widely used digital cellular network and the basis for mobile communication such as phone calls and the short message service (SMS).
GLONASS	The Russian global navigation satellite system with a constellation made of 24 satellites orbiting Earth. These multi-constellation GPS modules allow users to access multiple satellite networks, and accessing extra satellites allows for faster and more accurate positioning as well as offering greater resilience when satellites are obscured in areas such as cities.
GNSS	Global Navigation Satellite System. Used when talking about different constellations of satellite navigation systems.
GPRS	General Packet Radio Service.

Directory of IoT/M2M Terms

GPS	Global Positioning System. A system of satellites and radio transmissions that can be used to locate GPS‑enabled hardware anywhere on the planet to a very good accuracy.
Greenfield	In contradiction to brownfield, a greenfield project is a one where no consideration of previous systems is needed, thus already existing standards can be ignored.
GSM	Global System for Mobile communication.
GSM MAP	GSM Mobile Application Part, for control signal messaging on SS7.
HaaS	Hadoop as a Service.
Hadoop	A Java-based, distributed programming framework for processing large data sets. An application can be broken down into numerous small parts, called fragments or blocks, that can be run on any node in the cluster. Hadoop is free and part of the Apache project, sponsored by the Apache Software Foundation.
Hadoop as a Service (Haas)	The running of Hadoop in the Cloud, requiring no local hardware or IT infrastructure. The service is typically elastic, allowing the adding or removal of nodes depending on user needs.
Hadoop Distributed File System (HDFS)	The primary distributed storage used by Hadoop applications. A HDFS cluster has a NameNode that manages the file system metadata and DataNodes to store the actual data.
Haptic Technology or Haptics	Also referred to as Haptics or "touch feedback," haptic technology applies tactile sensations to human interactions with machines. The simplest example is the actuator that vibrates a cell phone, but more advanced haptics can detect the pressure applied to a sensor, affecting the response.
HDFS	Hadoop Distributed File System.
Heating, Ventilation, and Air Conditioning (HVAC)	Sometimes grouped with refrigeration as HVACR, these systems cover both vehicular and indoor building comfort control.
HEM	Home Energy Management.
HEMS	Home Energy Management System.
Heterogeneous Network (HetNet)	Small cell networks using both macro and small cells. HetNets allow mobile operators to better utilize their data networks' capacity.
HetNet	Heterogeneous Network.

Back to Directory

Directory of IoT/M2M Terms

High Speed Downlink Packet Access (HSDPA)	This increases the capacity of UMTS/3G bandwidth to allow for faster download speeds for connected devices.
High Speed Packet Access	An improvement made to transfer speeds over 3G technology through the addition of two new protocols; HSDPA and HSUPA. It offers potential downlink speeds of 14 Mbit/s and downlink of 5.76 Mbit/s.
High Speed Uplink Packet Access (HSUPA)	An improvement made to UMTS to enable faster uploading of data from devices, increasing capacity and throughput while reducing delay.
HLR	Home Location Register.
Home Automation	The automation of certain activities within a household. This can include automated control of lights, doors, and air conditioning, for example.
Home Energy Management	The process of increasing energy efficiency and savings for residential customers, ideally within a larger smart-grid environment.
Home Energy Management System (HEMS)	Equipment and services that optimize energy use in a residential setting while maintaining comfort. A HEMS includes smart appliances, home gateways, smart meters, and information exchange with local utilities via a smart grid.
Host	Computers that provide (or host) certain services or resources within a network that other participants within the network can then access and use. Hosts are the hardware basis for servers, as servers are run on hosts. Often times, they are the central point in a company's data processing process.
HSDPA	High-Speed Downlink Packet Access.
HSPA	High-Speed Packet Access.
HSPA+	Enhanced or Evolved High-Speed Packet Access.
HSUPA	High-Speed Uplink Packet Access.
HVAC	Heating, Ventilation, and Air Conditioning systems.
Hybrid Cloud	A mix of public and private cloud. The distribution of services through private or public channels is decided upon by the users.
IaaS	Infrastructure as a Service.
iBeacon	A technology introduced by Apple that uses sensors to locate iOS or Android devices indoors and can send them notifications via Bluetooth Low Energy (BLE). This can be also used in stores or museums to give further information about item nearby.

Back to Directory

Directory of IoT/M2M Terms

ICCID	Integrated Circuit Chip Identifier.
Identifier	Also just ID, this marks objects for clear identification. Identifiers are usually letters, words, symbols, or numbers that can also be used to create a code that reveals a definite identity after it is decoded.
Identity	Recognizable attributes that are linked to an object, a person, etc. Those attributes expose the entity and allow for clear identification. If two things have the exact same attributes, they usually have the same identity, and they can't be distinguished from each other.
Identity of Things (IDoT)	An area that involves assigning unique identifiers with associated metadata to devices and objects (things), enabling them to connect and communicate effectively with other entities over the Internet.
IEEE 802.11	The family of specifications developed by the IEEE for wireless LAN (WLAN) communications, first adopted in 1997. The addition of a letter, such as 802.11b, indicates a particular specification.
IEEE 802.11ac	Approved in January 2014, this is a wireless standard for high-throughput wireless local area networks (WLANs) on the 5 GHz band. In contrast to the four MIMO spatial streams in 802.11n, the 802.11ac standard supports eight.
IEEE 802.11n	Builds on previous 802.11 standards to use multiple antennas to increase data rates, adding MIMO to the physical layer. The full specification name is 802.11n-2009, which is an amendment to IEEE 802.11-2007.
IEEE 802.11p	Amends wireless access in vehicular environments (WAVE) to the IEEE 802.11 Wi-Fi standard. The amendment defines a way to wirelessly exchange data without the need to establish a basic service set (BSS), since links to roadside infrastructure may be available only for a limited amount of time. 802.11p uses channels of 10 MHz bandwidth in the 5.9 GHz band.
IGES	Initial Graphics Exchange Specification. This is a vendor-neutral, standardized file format used to transfer information between computer-aided design programs. The standard was developed to create a uniform method for exchanging graphical data between the programs.
IGMP	Internet Group Management Protocol. This communication protocol is based on the IP protocol and is used to support group communication. IGMP allows for IP-multicasting that enables the transmission of IP packages to many receivers with one transmission, and this is a requirement for technologies such as Internet television.
IIoT	Industrial Internet of Things.
IMEI	International Mobile Equipment Identifier (used in GSM).

Back to Directory

Directory of IoT/M2M Terms

IMSI	International Mobile Subscriber Identifier (used in GSM and CDMA).
IMU	Inertial Measurement Unit.
Industrial Control System (ICS)	Computer hardware and software that monitor and control industrial processes that exist in the physical world, where operator-driven supervisory commands can be pushed to remote station devices. Industries such as electrical, water, oil, and gas are typical ICS users.
Industrial Internet of Things (IIoT)	A subdiscipline of IoT, encompassing connected large-scale machinery and industrial systems such as factory-floor monitoring, HVAC, smart lighting, and security. This is M2M communication where, for example, equipment can send real-time information to an application so operators can better understand how efficiently that equipment is running. Also referred to as Industry 4.0, Industrie 4.0, and Industrial IoT.
Industrial, Scientific, and Medical (ISM) Bands	An unlicensed part of the RF spectrum used for general purpose data communications. In the US, the ISM bands are 915MHz, 2.4 GHz, and 5.5 GHz, whereas 2.4 GHz is the global unlicensed frequency and has increasing amounts of interference.
Industrie 4.0	Invoking a fourth Industrial Revolution, Industrie 4.0 creates intelligent manufacturing networks where decentralized smart factories can communicate and react to each other autonomously. For example, in an Industrie 4.0 factory, self-predictive systems would trigger maintenance processes autonomously and automatically adapt logistics to the resulting changes in production. The term, also known as Industry 4.0, was first used at the Hannover Messe in 2011.
Industry 4.0	Industry 4.0 is a project introduced by the federal government of Germany and refers to the fourth Industrial Revolution. It is a strategy which aims to make better use of current and future IT-capacities in traditional industries. Also see Industrie 4.0.
Inertial Measurement Unit (IMU)	A MEMS module which measures angular velocity and linear acceleration using an accelerometer triad and an angular rate sensor triad. Other IMU sensors may include magnetometers and pressure sensors.
Information and Communication Technologies (ICT)	The ICT Industry provides access to information through telecommunications. The communications technologies can be things like the Internet, VOIP, wireless networks, or mobile phones.
Infrastructure as a Service (IaaS)	An on-demand business model for IT-capacities. Instead of owning IT-infrastructure or server space, you rent it and pay for it on a per-use basis. Those capacities are usually owned, maintained, and provided by a cloud service.

Back to Directory

Directory of IoT/M2M Terms

Insurance Telematics	Vehicular tracking devices used by automobile insurance companies to alter rates based on driver behavior. Currently, Progressive (Snapshot), Allstate, and others typically track braking and mileage. An excessive number of hard-brakes may indicate risky driving habits, for example.
Intelligent Device	Any type of equipment, instrument, or machine that has its own computing capability. As computing technology becomes more advanced and less expensive, it can be built into an increasing number of devices of all kinds. In addition to personal and handheld computers, the almost infinite list of possible intelligent devices includes cars, medical instruments, geological equipment, and home appliances.
Intelligent Transportation System (ITS)	The application of advanced information and communications technology to surface transportation for enhanced safety and mobility while reducing environmental impacts. EU Directive 2010/40/EU defines ITS in the context of road transport.
Inter-Integrated Circuit (I2C)	I2C, pronounced I-squared-C, is a serial bus that provides communication between sensors and microcontrollers such as the Arduino. In contrast to the full-duplex SPI specification, I2C has a slower data rate, and data can only travel in one direction at a time. Arduino uses 7-bit values to reference I2C addresses, and devices using I2C must use a common ground to communicate.
Internet of Everything (IoE)	A term being promoted by Cisco as a variation or extension of IoT. IoE subtly distinguishes itself by emphasizing the connection of people to things.
Internet of Things (IoT)	Internet-connected physical devices, in many cases everyday objects (things), that can communicate their status, respond to events, or even act autonomously. This enables communication among those things, closing the gap between the real and the virtual world and creating smarter processes and structures that can support us without needing our attention. IoT has evolved from the convergence of wireless technologies, micro-electromechanical systems (MEMS), and the Internet.
Internet of Things Privacy (IoT Privacy)	The special considerations required to protect the information of individuals from exposure in the IoT environment, where almost any physical or logical entity or object can be given a unique identifier and the ability to communicate autonomously over the Internet or similar network.
In-Vehicle Infotainment (IVI)	Systems integrated into automobiles that deliver both entertainment and information content. Typical IVI applications include managing audio, listening to or sending SMS, making voice calls, navigating, and using rear-seat entertainment such as games, movies, games, and social networking. IVI also includes interfacing with smartphone-enabled content such as traffic conditions, sports scores, and weather forecasts.

Directory of IoT/M2M Terms

IoE	Internet of Everything.
IoT	Internet of Things.
IoT Botnet	Internet of Things botnet. A group of hacked computers, smart appliances, and Internet-connected devices that have been co-opted for illicit purposes.
IoT Healthcare	Also called "connected health," this encompasses all advancements in the medical industry that relate to M2M communication and remote sensing.
IoT Security	Internet of Things security. The area concerned with safeguarding connected devices and networks in the Internet of things).
IP Devices	All devices within a network which are labeled with an IP address.
IPSEC	Internet Protocol Security. A set of protocols that provide authentication and encryption to Internet Protocol (IP) packets, adding an extra layer of security on IP communications.
IPv6	IP addresses serve to identify devices on the Internet. IPv6 is the newest Internet protocol, which provides more addresses than the IPv4 protocol.
IPv6 Address	A 128-bit alphanumeric string that identifies an endpoint device in the Internet Protocol Version 6 (IPv6) addressing scheme.
IRIDIUM	A satellite communication constellation that provides global voice and data coverage through its satellite network, operating on the 1618.85 to 1626.5 MHz frequencies.
IS-136	Interim Standard 136 (standard for TDMA Cellular).
IS-95	Interim Standard 95 (standard for CDMA Cellular).
ISM Bands	Industrial, Scientific, and Medical Bands.
ITS	Intelligent Transportation System.
ITU	International Telecommunications Union.
IVI	In-Vehicle Infotainment.
JavaScript Object Notation (JSON)	Used as a lightweight alternative to XML for organizing data, JSON is text-based and human-readable. The format uses "name : object" pairs to organize the data.
JSON	JavaScript Object Notation.

Directory of IoT/M2M Terms

Kevin Ashton	The man who first coined the phrase "Internet of Things" in 1999. Mr. Ashton cofounded MIT's Auto-ID Center which created a global standard system for RFID.
L2TP	Layer 2 Tunneling Protocol. This is a tunneling protocol used to support virtual private networks (VPNs) or as part of the delivery of services by ISPs. It does not provide any encryption or confidentiality by itself, relying on an encryption protocol that it passes within the tunnel to provide privacy.
LED	Light-Emitting Diode.
Light-Emitting Diode (LED)	A semiconductor that generates light via electroluminescence. Infrared LEDs can be used for the remote control units of many consumer electronics.
Link Budget	An accounting of all of the losses in a wireless communication system. In order to "close the link," enough RF energy has to make it from the transmitter to the receiver. (Losses include antennas, structural attenuation, propagation loss, etc.)
LLN	Low power and Lossy Networks.
Local Area Network (LAN)	A network of devices in relatively close proximity, prior to the point of transmission over leased telecommunication lines. The two most common communications technologies used in LANs are Ethernet and Wi-Fi.
Long Term Evolution (LTE) / 4G	LTE, often referred to as 4G, is the latest cellular network type, offering superior data transfer speeds than its predecessor, 3G, and it's part of the GSM upgrade path. Portable devices can now access data at high-speed broadband speeds through LTE. Depending on where in the world you are, LTE may be implemented using different frequency bands. For example, European LTE uses 700/800/900/1800/2600 MHz bands, where North America uses 700/750/800/850/1900/1700/2100(AWS)/2500/2600 MHz.
Low power and Lossy Networks (LLN)	These networks are comprised of embedded devices with limited power, memory, and processing resources. LLNs are typically optimized for energy efficiency, may use IEEE 802.15.4, and can be applied to industrial monitoring, building automation, connected homes, healthcare, environmental monitoring, urban sensor networks, asset tracking, and more.
Low Power Wide Area (LPWA)	These networks are built specifically for M2M communications and offer long-range, low-power consumption. They solve cost and battery-life issues that cellular technology cannot, and LPWA networks solve range issues that technologies like Bluetooth or BLE struggle with.

Back to Directory

Directory of IoT/M2M Terms

Low Power Wireless Sensor Network	A group of spatially distributed, independent devices that collect data by measuring physical or environmental conditions with minimal power consumption.
LTE	Long Term Evolution.
M2M	Machine-to-Machine.
M2P	Machine-to-Person.
Machine Authentication	The authorization of an automated human-to-machine or machine-to-machine (M2M) communication through verification of a digital certificate or digital credentials. Unlike user authentication, the process does not involve any action on the part of a human.
Machine Data	Also known as machine-generated data, this is digital information created by the activity of computers, mobile phones, embedded systems, and other networked devices.
Machine Type Communications (MTC)	A 3rd Generation Partnership Project (3GPP) standard describing machine-to-machine communications. With a wide range of potential applications, MTC communications is gaining interest among mobile network operators, equipment vendors, specialist companies, and research bodies.
Machine-to-Machine (M2M)	A broad term describing technology that allows for one connected device to communicate and exchange information with another connected device, without the assistance of a human.
Machine-to-Person (M2P)	Describes the analytics for big data in a human readable form (e.g., dashboards).
MapReduce	A parallel processing model for handling extremely large data sets. First, a Map process runs to reduce a data set to key value pairs (in tuples), and then a second Reduce process combines those pairs into a smaller set of tuples. First introduced by Google, MapReduce is a concept central to Hadoop.
MCU	MicroController Unit.
MDN	Mobile Directory Number (used in CDMA—conceptually similar to the MSISDN in GSM).
Mechatronics	A combination of the words "mechanical" and "electronics," mechatronics brings together electrical engineering, control engineering, computer engineering, and mechanical engineering disciplines. A warehouse inventory robot would be a mechatronic device, whereas an IoT-enabled sensor device, such as a weather station, could be better classified as a Cyber-Physical System (CPS).

Directory of IoT/M2M Terms

Media Access Control (MAC)	The "layer 2" in a network that allows the physical medium (radio waves or wire signals) to be organized to pass data back and forth. For low-rate data wireless applications, the MAC has many implications on performance.
MEID	Mobile Equipment Identifier (used in CDMA).
MEMS	Micro-Electro-Mechanical Systems.
Mesh Networking or Mesh Network Topology	An ad-hoc, local area network infrastructure where the nodes communicate directly with each other without the need to pass through a central structure such as an ISP. The only way to shut down a mesh network is to eliminate every node. One of the most dramatic demonstrations of mesh technology was during the Hong Kong protests in October 2014 where the direct communication between protestors' devices confounded the government's ability to block communication. The adaptivity of mesh networks makes them ideal for IoT applications.
Message Broker	A middleware program that translates a message from the messaging protocol of the sender into the messaging protocol of the receiver. This way a message broker makes it easier for two applications to communicate.
Message Queue Telemetry Transport (MQTT)	An open, lightweight M2M communications protocol for the transfer of telemetry messages.
Message-Oriented Middleware (MOM)	Middleware that allows for synchronous as well as asynchronous (queue) messaging between distributed systems.
mHealth	Mobile Health. This is the practice of medicine using mobile devices, particularly physiological sensors. Sensors may be enabled to communicate with a user's mobile phone in a Body Area Network configuration. Related to e-Health.
MicroController Unit (MCU)	A full computer on a single chip. The chip contains a CPU, a clock, non-volatile memory for the program (ROM or flash), volatile memory for input and output (RAM), and an I/O control unit.
Micro-Electro-Mechanical Systems (MEMS)	Miniaturized mechanical and electro-mechanical elements, typically used for measurements, such as accelerometers and gyroscopes. Systems-on-a-chip (SoC) technology is used to embed mechanical devices such as fluid sensors, mirrors, actuators, pressure and temperature sensors, and vibration sensors on to semiconductor chips.
MIMO	Multiple Input, Multiple Output (in the context of antennas).
MNO	Mobile Network Operator.
Mobile Network Operator (MNO)	Companies that operate traditional mobile communications networks.

Back to Directory

Directory of IoT/M2M Terms

Mobile Virtual Network Operator (MVNO)	A wireless communications provider that leases the infrastructure over which it proves services.
Modbus	A communication protocol mainly used to connect electronic devices. The Modbus Master (for example, a computer) requests information from the Modbus Slaves (for example, electronic thermometers). Up to 247 Slaves can transmit their information to one Master.
Mote	Short for Remote. A mote is a wireless transceiver that also acts as a remote sensor.
MQTT	Message Queue Telemetry Transport.
MS	Mobile Station (cellular radio handset or cellular M2M device).
MSC	Mobile Switching Center.
MSISDN	Mobile Station ISDN (used in GSM).
MTC	Machine Type Communications.
Multiple DoF Sensing	A MEMS concept referring to the detection of the combined input along multiple axes using multiple sensing types, such as acceleration and rotation. Typical applications include antenna stabilization, robotics and dead-reckoning.
Multiple-Input and Multiple-Output (MIMO)	A radio technology using multiple antennas at both the transmitter and receiver to improve communication performance. MIMO is an important part of wireless communication standards such as IEEE 802.11n (Wi-Fi).
Nagios	Software that monitors IT infrastructures. It includes, for example, immediate problem detection.
Near Field Communication (NFC)	Short-range wireless communication between devices, used in applications such as contactless mobile payments, transport ticketing, and phone-as-key. Using NFC, consumers can pay for retail items simply by bringing their mobile phones into the range of the sensor and confirming the transaction. NFC has been overshadowed in IoT applications by other protocols such as BLE.
Nearables	Coined for the similarity to "wearables," this describes items with nearby tracking devices, or beacons, attached to them. Nearables can communicate with smart devices, such as smartphones, to let the user interact with objects in their vicinity.
NFC	Near Field Communication.

Back to Directory

Directory of IoT/M2M Terms

OBE	On-Board Equipment.
On-Board Equipment (OBE)	Components of a Vehicle-to-Infrastructure (V2I) implementation located in a moving vehicle, communicating wirelessly with roadside equipment (RSE). OBE applications may interface with other vehicle systems via the CAN Bus.
Open Source	A type of software where the source code is available and can be modified and freely redistributed. Open source is the opposite of closed, proprietary systems. Many developers insist that IoT must have open standards to reach its full potential.
Open VPN	An open-source software application that implements virtual private network (VPN) techniques for creating secure point-to-point or site-to-site connections in routed or bridged configurations and remote access facilities. This is a security method which can be implemented on devices such as cellular routers.
Operational Technology (OT)	As opposed to Information Technology (IT), this refers to technologies associated with control and automation. If IT helps run business processes, OT helps execute the physical interactions that control value creation.
PAN	Personal Area Network.
Part 90 Bands	Small parts of the RF spectrum that are made available in small areas to businesses for data or voice communications. Many smart grid providers use part 90 licenses for their wireless data.
Passive Sensor	A device that detects and responds to some type of input from the physical environment.
PCB	Printed Circuit Board.
PDR	Pedestrian Dead Reckoning.
PDU	Power Distribution Unit.
Pedestrian Dead Reckoning (PDR)	A method of indoor positioning that uses a last known waypoint, distance, and direction of travel to calculate the current location of a moving person. PDR may be used to supplement other positioning methods such as GPS. Dead reckoning is subject to cumulative errors.
Pen Testing or Pentest	Penetration Testing.

Back to Directory

Directory of IoT/M2M Terms

Penetration Testing	A method of evaluating the security of a network or system from internal or external threats. Also called pentests, this is part of a full security audit and typically exploits a combination of weaknesses to gain access and then evaluates the capability of the network's defenders to detect and respond to the penetration.
PERS	Personal Emergency Response System.
Personal Area Network (PAN)	Interconnected devices operating in the range of a single person, typically 10 meters. PANS are mostly or exclusively wireless, making the term basically indistinguishable from Wireless PANs (WPAN). WPAN is based on the IEEE 802.15 standard and does not necessarily require an uplink to the Internet. The PAN concept was first developed by Thomas Zimmerman and others at the M.I.T. Media Lab.
Personal Emergency Response System (PERS)	A mobile duress panic alarm component of a monitoring system, typically for the residential market. Modern PERS devices go beyond their origins as a mere push button to include MEMS and various other sensors.
Personal Protection Drone (PPD)	A type of drone, or drone swarm, dedicated to an individual's security. PPDs are non-lethal and may be primarily used to record an encounter or raise an alarm. A use case for a swarm of PPDs may be to hinder the approach of an assailant long enough to facilitate the protected person's withdrawal.
Pervasive Computing	Another term for ubiquitous computing.
Photoplethysmogram (PPG)	A optically obtained plethysmogram using an LED that measures the output volume of an organ, such as the heart. A photodiode measures the amount of light reflected from the LED, which in a heart monitoring application can be translated into a waveform. Respiration can induce variations in the amplitude of the PPG waveform.
Physical Web, The	Google's open standard to allow IoT devices to communicate via web addresses. By using HTTP, users can walk up and access any smart device (such as parking meters and vending machines) without the overhead of dedicated mobile apps.
Platform as a Service (PaaS)	A platform that provides web developers with all the infrastructure they need to develop and run an application.
PoE	Power over Ethernet.
Power Distribution Unit (PDU)	A physical device with multiple outlets that connects electrical power to recipient devices. PDUs can be simple, such as a mounted power strip, or more complex by having power filtering, UPS, load balancing, or intelligent monitoring incorporated in the device.

Directory of IoT/M2M Terms

Power over Ethernet (PoE)	The capability to deliver enough power to operate a device over an Ethernet connection. PoE is useful in certain low-voltage applications, such as passive IP cameras.
PPD	Personal Protection Drone.
PPG	Photoplethysmogram.
PPTP	Point-to-Point Tunneling Protocol. This is a method for implementing virtual private networks (VPNs).
Preboot Execution Environment (PXE)	The ability to manage power over a network connection. A PXE-enabled device can be shut down or restarted via a network connection, allowing for power-hungry devices to be managed remotely.
Private Cloud	A private cloud provides services with cloud characteristics but only within a single organization, for example, one company.
Public Cloud	In a public cloud, cloud services are public and made available for everyone.
Pulse Oximeter	A sensor that measures oxygen saturation in the blood. Saturation of peripheral oxygen, or SpO2, is a measure of hemoglobin saturation and can be measured non-invasively, for example, with a clip on the finger or ear. The sensor typically employs a pair of small LEDs facing a photodiode that measures the amount of light passing through the skin.
PXE	Preboot Execution Environment.
Python	A widely used open-source programming language that can be implemented in variety of ways, including in embedded applications. There is a large library base which can be used by Python applications, helping minimize code and speed up development time.
Python Script Interpreter	A tool that lets you run Python code, something which is now being seen embedded directly into devices such as cellular modules.
Quality of Service (QoS)	Different services that regulate data transfer priorities to identify and control the quality with which a service can be accessed by users. This is especially important if a certain quality (for example, bandwidth) has to be guaranteed to ensure the functionality of a service.
Quantified Self	A movement that started in 2007 that uses modern technical advances to gain more insight into one's own life by collecting data relating to, among other things, health and emotions. This data is then used to improve a person's lifestyle and state of mind.

Back to Directory

Directory of IoT/M2M Terms

Quantum Sensor	A sensor that takes advantage of quantum correlations to produce measurements beyond what are possible with traditional sensors. Taking advantage of unique behavior of systems at the atomic scale, quantum sensors use wonder materials such as graphene and quantum dots.
Radio Frequency (RF)	Radio waves. This term generally means "wireless communication" when referred to in IoT discussions.
Radio Frequency Identification (RFID)	Generally speaking, this is the use of strong radio waves to "excite" enough current in a small tag to send a radio transmission back. It works over short range and only for small amounts of data.
RADIUS	Remote Authentication Dial-In User Service. RADIUS servers are responsible for receiving user connection requests, authenticating the user, and returning all configuration information necessary for the client to deliver service.
Remote Monitoring and Control	The increasingly automated monitoring and control of devices, technologies, or processes. Wireless devices which send information gathered directly to control centers are often used to achieve this.
Remote Sensing	The use of various technologies to make observations and measurements at a target that is usually at a distance or on a scale beyond those observable to the naked eye.
REST	Representational State Transfer. An architecture for web standards, especially for the HTTP protocol. It is supposed to simplify the design of network applications compared to, for example, SOAP.
RESTful Web Services	Web services that are realized within the REST architecture are called RESTful Web Services. Also see REST.
RF Geolocation	A general term that applies to "finding" a radio transceiver with another. GPS is a good example. A good rule to remember is that to do RF geolocation well, you need a large RF bandwidth.
RF Sensitivity	The minimum magnitude of input signal you need based on a specified signal-to-noise ratio to achieve at minimum error rate.
RFID	Radio Frequency Identification.
SaaS	Software as a Service.
SBC	Single Board Computer.
SCADA	Supervisory Control and Data Acquisition.
SDN	Software-Defined Networking.

Back to Directory

Directory of IoT/M2M Terms

SDN	An approach where the control plane and the data plane of a computer network are separated. The control plane is the one making the decision about where traffic is sent, the data plane the one to forward that traffic. The abstraction of such lower function levels should simplify networking.
SDO	Standards Development Organization.
Sensor	A device used to measure a specific characteristic of the surrounding environment, such as temperature. The use of sensors and actuators to connect things to the physical world is a key component of IoT. A properly implemented sensor ideally should be sensitive only to the characteristic being measured and should not interfere with what's being measured nor be influenced by other characteristics.
Sensor Analytics	Statistical analysis of data that is created by wired or wireless sensors.
Sensor Fusion	The process of combining and processing the raw data coming out of multiple sensors to generate usable information. For example, because of the quantity of sensors, a NASA un-crewed vehicle on Mars requires sensor fusion to detect if there has been a failure.
Sensor Hub	A technology that connects sensor data and processes them. This way the hub does part of a processors data-processing job.
Serial Peripheral Interface (SPI)	A specification developed by Motorola for use in short distance communication between sensors and microcontrollers such as Arduino. In contrast to the I2C specification, the full-duplex SPI runs at a higher data rate and is appropriate for applications such as Ethernet and memory cards.
Serial Port Profile (SPP)	A hardware profile used with Bluetooth applications that includes custom AT commands and functionality dedicated to wireless data connections and serial cable replacement.
SGSN	Serving GPRS Support Node (see also GGSN).
Shock Sensing	A MEMS concept referring to the detection of sudden impacts at a predetermined level. Typical applications include shut-off sensing, condition monitoring, and tap detection for data entry.
SIGFOX	A low-bandwidth, wireless protocol that offers excellent range and obstacle penetration for short messages, giving a new low-powered and cost-effective wireless transmission medium for IoT and M2M technologies.
Signal Phase and Timing (SPaT)	Refers to communications associated with the operations of signalized intersections. The major components associated with a SPaT application are roadside equipment (RSE) and onboard equipment (OBE). A SPaT-formatted message can be used to convey the current status of a signal at an intersection.

Back to Directory

Directory of IoT/M2M Terms

SIM	Subscriber Identity Module. A piece of hardware (the "smart card") containing account information for a user on a GSM network. The SIM is inserted into a SIM holder in GSM cellular devices.
Single Board Computer (SBC)	A complete, functioning computer with all functions (I/O, processor, memory) located on one board. Popularized by the Raspberry Pi system, SBCs are constructed in direct contrast to traditional motherboards with plug-in cards for functions like graphics and Ethernet.
SMA	SubMiniature version A. A type of connector commonly used with antenna, giving you male and female coaxial cable connectors that connect with a screw head.
Smart Buildings	Buildings that try to minimize costs and environmental impact. This is achieved by connected systems and efficient use of energy through new, automated technology that intelligently responds to certain circumstances (available solar energy, temperature inside the building, etc.).
Smart Car	An automobile that uses technology to support the driver and create a safer traffic environment. Different systems (inside and outside of the car) are connected and communicate with each other to allow intelligent intervention in dangerous situations and more fluid traffic.
Smart Cities	A concept that tries to create a more intelligent city infrastructure by using modern information and communication technologies. Smart cities propose a more flexible adaptation to certain circumstances, more efficient use of resources, higher quality of life, more fluid transportation, and more. This may be achieved through networking and integrated information exchange between humans and things.
Smart Grid	A general term referring to the application of networking capabilities and computer systems to the electric grid. A smart grid would include smart meters at the point of delivery, allowing for real time monitoring of usage and the adjustment of power settings on some appliances.
Smart Home	The networking of household devices and systems through information and communication technology. This way, processes within a home can be monitored and controlled automatically to optimize quality of life, costs, security, and environmental impact. Related to Connected Home.
Smart Meter	An electronic device that measures and displays resource consumption (of water, gas, electricity, etc.) and communicates this information to the resource distributors and managers (such as utilities and municipalities) and even to consumers. This allows for a more efficient distribution, usage, pricing, and control of these resources.

Back to Directory

Directory of IoT/M2M Terms

Smartwatch	A wristwatch, generally with a display, that interacts with the wearer and can communicate with a network wirelessly (the device may have a USB connection for charging and other functions). Many smartwatches have MEMS and physiological sensors, such as ECG and skin temperature thermometers.
SMC	Short Message Center.
SMS	Short Message Service.
SMSC	Short Message Service Center.
SOAP	Simple Object Access Protocol.
SoC	System on a Chip.
Social Web of Things	The socialization of the Internet of Things. This is the integration of connected things into our social life. An example would be a TV that not only informs you that your favorite TV show is on in an hour, but also lets you know which of your friends like the show too so you can meet up and watch together.
Software as a Service (Saas)	A subscription-based model where a monthly fee is charged for using software, rather than an upfront purchase. SaaS (also spelled SAAS) and cloud computing can give cash-strapped enterprises and startups access to applications such as email and lead management that might otherwise be too expensive to purchase outright.
Software-Defined Network (SDN)	An approach to networking that decouples control of information flow from the hardware and gives it to a software controller. This allows for less data to travel wirelessly, making it a potential strategy for IoT networks.
Spaced Repetition	A quantified self-concept designed to increase the brain's retention of knowledge. Available via apps and cloud-based technologies, spaced repetition operates on the theory that there is an optimum time between memorization drills to maximize retention.
SPaT	Signal Phase and Timing.
SPI	Serial Peripheral Interface.
SS7	Signaling System 7.
Steel Collar	Things in the workplace that replace or augment human labor. A "steel-collar workforce" is capable of tirelessly and efficiently performing repetitive tasks or monitoring. Playing off of the terms "blue collar" and "white collar," the phrase was first coined in the early 1980s referring to a robotic threat to US manufacturing jobs.

Back to Directory

Directory of IoT/M2M Terms

STOMP	Simple (or Streaming) Text Oriented Message Protocol. It's similar to HTTP and allows STOMP clients to communicate with most of the message brokers making it language-agnostic.
Structure Attenuation	The loss in intensity of radio waves through a medium (like radio waves through a brick wall).
Subscriber Identity Module (SIM)	Provided by the Mobile Network Operator, a SIM contains the International Mobile Subscriber Identity (IMSI) and the security parameters to authenticate access to the network.
Supervisory Control and Data Acquisition (SCADA)	An industrial control system typically used for geographically dispersed assets, often scattered over large distances. SCADA is often applied to electrical utilities to monitor substations, transformers, and other electrical assets.
SX1272	First-generation long-range wireless transceiver from Semtech, which introduced a new type of PHY layer modulation. This technology dramatically increases the range of sub-GHz RF communications.
SX1276	Follow on to SX1272 from Semtech, and this part includes frequency coverage for more unlicensed bands worldwide and several modes that increase receive sensitivity.
System on a Chip (SoC)	A single integrated-circuit technology that contains all the necessary circuits and parts for a complete system. A single microchip in a wearable device, for example, could contain an analog-to-digital converter, memory, logic control, I/O, etc.
TaaS	Things as a Service.
TDMA	Time Division Multiple Access.
Telematics	An IT concept regarding the long-distance transmission of data. In vehicles on the move, telematics refers to the integrated use of telecommunications and informatics, such as dashboard screens that show the vehicles current position on a map or in centralized tracking applications.
TETRA	Terrestrial Trunked Radio. This operates as a two-way transceiver and is popularly used by the emergency services as well as on transport such as rail and on marine vessels. It operates on low frequencies split over 4 channels (ranging between 380 and 400 MHz for emergency services and higher for civilian use). The use of low frequencies allows for far greater transmission distances but lower data transfer rates.
Thing, in the Internet of Things	An entity or physical object that has a unique identifier, an embedded system, and the ability to transfer data over a network.

Back to Directory

Directory of IoT/M2M Terms

Thingbot	Something with an embedded system and an Internet connection that has been co-opted by a hacker to become part of a botnet of networked things.
ThingManager	A platform for managing real-world Things and their digital representations.
Things as a Service (TaaS)	The concept of delivering IoT functionality without the end user having to operate or maintain extensive hardware. For example, services such as Hadoop can be delivered in the cloud to receive and process the data generated by IoT-enabled sensor networks.
THNGMNGR	ThingManager.
Thread	A simplified IPv6-based mesh networking protocol geared to the smart home vertical. Developed on low-cost 802.15.4 chipsets, Thread is designed for extremely low power consumption.
Tilt Sensing	A MEMS concept referring to the measurement of the inclination or angle of change with respect to gravity. Typical applications include industrial equipment platform stabilization and landscape/portrait detection on handheld devices.
Transceiver	Short for transmitter-receiver. A transceiver both transmits and receives analog or digital signals. A transceiver is normally built into a network interface card.
Transmission Control Protocol/Internet Protocol (TCP/IP)	The core standard protocol for Internet-based communications. Some wireless systems "break" TCP/IP in order to lower the overhead of the on-air signals.
Transparent Computing	A characteristic of ubiquitous computing where smart devices respond to users' needs in the background. The devices are invisible ("transparent") in the sense that they operate without the conscious thought or interaction of the user who is benefiting from the object or Thing.
TV Whitespace	A new FCC program that makes unused TV station bands available for temporary and controlled use in a small geographic area. This is used mostly by rural Internet service providers and wireless microphone providers.
UART	Universal Asynchronous Receiver/Transmitter.
UBI	Usage-Based Insurance.
Ubiquitous Computing	The concept of embedding microprocessors in everyday things so they can communicate information continuously. Ubiquitous devices are expected to be constantly connected. Utility smart meters are an example of ubiquitous computing, replacing manual meter-readers with devices that can report usage and modify power settings on ubiquitous appliances.

Directory of IoT/M2M Terms

Ultra-Wide Band	A "spark gap" transmitter that emits a very weak, very wide (in frequency) pulse of RF energy. This signal is used mostly for localizing signals. Wide signal bandwidths are good for measuring distance.
UMTS	Universal Mobile Telecommunications System.
Uniform Resource Identifier (URI)	The unique identifier that makes content addressable on the Internet by uniquely targeting items, such as text, video, images, and applications.
Uniform Resource Locator (URL)	A particular type of URI that targets webpages so that when a browser requests them, they can be found and served to users.
Universal Asynchronous Receiver/Transmitter (UART)	A microchip controlling a computer's interface to serial devices, converting the bytes it receives from the computer along parallel circuits into a single serial bit stream. A 16550 UART has a 16-byte buffer.
Universal Authentication	A network identity-verification method that allows users to move from site to site securely without having to enter identifying information multiple times.
Universal Mobile Telecommunications System (UMTS)	Also referred to as 3G cellular technology, this is the third iteration of the GSM. It achieves improved data transfer speeds over 2G by adding additional higher frequency bands (2100MHz).
Uplink	Notified as UL or U/L, this is the process of sending data from your device/computer to a server or target address. In a cellular network, this would be seen as data being sent from a mobile handset to a cellular base station.
Usage-Based Insurance (UBI)	Also called Pay as You Drive (PAYD), UBI bases the insurance rate on pre-defined variables including distance, behavior, time, and place. The data gathering and telematics can be provided by a "black box" in the vehicle, a dongle-type device, or even a smartphone.
V2I	Vehicle-to-Infrastructure.
V2V	Vehicle-to-Vehicle.
V2X	A shorthand to combine Vehicle-to-Vehicle (V2V), Vehicle-to-Infrastructure (V2I), and Vehicle-to-Anything.
Vehicle-to-Infrastructure (V2I)	The communication of smart cars and commercial vehicles with surrounding sensors, such as signal phase and timing (SPaT) information.
Vehicle-to-Vehicle (V2V)	Using a region of the 5.9 GHz band, V2V systems allow vehicles to communicate with each other and with roadside stations. Networks of vehicles can help avoid congestion, find better routes, and aid law enforcement.

Back to Directory

Directory of IoT/M2M Terms

Vehicle-to-Vehicle Communication (V2V Communication)	The wireless transmission of data between motor vehicles.
Vibration Sensing	A MEMS concept referring to the detection of periodic acceleration and deceleration. Typical applications include structural health monitoring, acoustic event triggering, and seismic equipment.
Video Motion Detection (VMD)	A technology that analyzes image data and the differences in a series of images. VMD makes event-driven video surveillance possible, but the potential for false positives creates challenges in storage and alarm verification.
Video Surveillance as a Service (VSaaS)	A managed data service that transfers the monitoring and storage of video to the cloud. VCaaS streamlines security operations by centralizing IT and requires no capital investment in servers but has heavy bandwidth requirements.
Virtual Power Plant (VPP)	In a virtual power plant, different, decentralized power generating plants are connected and are monitored and controlled from a single control center. This way, virtual power plants can integrate smaller energy providers–for example solar or wind parks–into the energy infrastructure. VPPs are also able to flexibly react to changes in demand.
Virtual Sensor	These sensors use data to gather information that would not be measurable by a single device. This way they can attain information that can't be measured directly.
VLR	Visited Location Register.
VMD	Video Motion Detection.
VSaaS	Video Surveillance as a Service.
WAN	Wide Area Network.
WAVE	Wireless Access in Vehicular Environments.
Wearables or Wearable Technology	These are technologies or computers integrated into articles of clothing or accessories that can be worn. Often, the wearble tech is used to quantify a physical process (such as heartbeat monitoring) or to augment human capabilities. Wearables may also be used to control external things, for example, with gestures. Because of the impracticality of wires to transmit sensor data, wearables are almost universally wireless, using a variety of communication protocols such as BLE. Examples include smartwatches, fitness bands, and Google Glasses.

Directory of IoT/M2M Terms

Wi-Fi	Wireless Fidelity. This is a common form of local area network which operates on the 2.4 GHz band. Its popularity has led to a wide variety of devices to become Wi-Fi enabled, including smartphones, cameras, vehicles, and household appliances. Wi-Fi can be embedded into a device through designing in a Wi-Fi module.
Wireless Access in Vehicular Environments (WAVE)	The IEEE 802.11p standard required to support Intelligent Transportation Systems (ITS) applications. ITS applications include data exchange between moving vehicles and between vehicle and ITS-enabled roadside infrastructure.
Wristop	A contraction of wristband and desktop, a wristop computer refers to a wearable that goes on the wrist, such as a smartwatch.
ZigBee	Small range wireless networking protocol that primarily operates on the 2.4 GHz frequency spectrum. ZigBee devices connect in a mesh topology, forwarding messages from controlling nodes to slaves, which repeat commands to other connected nodes. Due to its low power consumption and low data rate, ZigBee has been used in applications such as traffic management, wireless light switches, and industrial device monitoring.
Z-Wave	Wireless communication technology used in security systems and also business and home automation.

The Definitive Guide to the Internet of Things for Business
By Syed Zaeem Hosain, CTO, Aeris
With a foreword by Stefan Lindvall, CEO, MultiTech

First Edition: November 2015

Book Design: Shay Lari-Hosain & Anthony Estes. Covers: ArcherDog. Editor: Trystan L. Bass.

For further information about this book, contact Aeris, 2350 Mission College Blvd, Suite 600, Santa Clara, CA 95054-1574, or **www.aeris.com**.

ABOUT THE AUTHOR

Syed Zaeem Hosain is responsible for the architecture and future direction of Aeris' networks, development programs, and technology strategy. He joined Aeris in 1996 as Vice President, Engineering, and is a member of the founding executive team of Aeris.

He has over 35 years experience in the semiconductor, computer, and telecommunications industries, including product development, architecture design, and technical management. Prior to joining Aeris, he held senior engineering and management positions at Analog Devices, Cypress Semiconductor, CAD National, and ESS Technology. He has presented papers and participated on panel sessions at conferences on Internet of Things (IoT) applications, machine-to-machine (M2M) networks, and cellular technologies for IoT/M2M.

Mr. Hosain is Chairman of the International Forum on ANSI-41 Standards Technology (IFAST), and is on the Board of the Mobility Development Group (MDG). He holds a Bachelor of Science degree in Computer Science and Engineering from the Massachusetts Institute of Technology, Cambridge, MA.

ABOUT AERIS

Aeris is a pioneer and leader in the market of the Internet of Things–as an operator of end-to-end IoT and M2M services and as a technology provider enabling other operators to build profitable IoT businesses. Among our customers are the most demanding users of IoT services today, including Hyundai, Acura, Rand McNally, Leica, and Sprint. Through our technology platform and dedicated IoT and M2M services, we strive to fundamentally improve their businesses–by dramatically reducing costs, improving operational efficiency, reducing time-to-market, and enabling new revenue streams. Visit www.aeris.com or follow us on Twitter @AerisM2M to learn how we can inspire you to create new business models and to participate in the revolution of the Internet of Things.

aeris