UNIT IV: M2M and IoT Technology Fundamentals

Devices and gateways, Local and wide area networking, Data management, Business processes in IoT, Everything as a Service(XaaS), M2M and IoT Analytics, Knowledge Management.

4.1 Devices and gateways

4.1.1 Introduction

- There is a growing market for small-scale embedded processing such as 8-, 16-, and 32-bit microcontrollers with on-chip RAM and flash memory, I/O capabilities, and networking interfaces such as IEEE 802.15.4 that are integrated on tiny System-on-a-Chip (SoC) solutions.
- Such devices enable very constrained devices with a small footprint of a few mm2 and with a very low power consumption in the milli- to micro-Watt range, but which are capable of hosting an entire Transmission Control Protocol/Internet Protocol (TCP/IP) stack, including a small web server.
- A device is a hardware unit that can sense aspects of it's environment and/or actuate, i.e. perform tasks in its environment.
- A device can be characterized as having several properties, including:

  • Microcontroller: 8-, 16-, or 32-bit working memory and storage.

  • Power Source: Fixed, battery, energy harvesting, or hybrid.

  • Sensors and Actuators: Onboard sensors and actuators, or circuitry that allows them to be connected, sampled, conditioned, and controlled.

  • Communication: Cellular, wireless, or wired for LAN and WAN communication.

  • Operating System (OS): Main-loop, event-based, real-time, or full featured OS.

  • Applications: Simple sensor sampling or more advanced applications.

  • User Interface: Display, buttons, or other functions for user interaction.

  • Device Management (DM): Provisioning, firmware, bootstrapping, and monitoring.

  • Execution Environment (EE): Application lifecycle management and Application Programming Interface (API).

4.1.1.1 Device types

- Group devices into two categories

  • **Basic Devices:** Devices that only provide the basic services of sensor readings and/or actuation tasks, and in some cases limited support for user interaction. LAN

communication is supported via wired or wireless technology, thus a gateway is needed to provide the WAN connection.

• **Advanced Devices:** In this case the devices also host the application logic and a WAN connection. They may also feature device management and an execution environment for hosting multiple applications. Gateway devices are most likely to fall into this category.

4.1.1.2 Deployment scenarios for devices

- Example deployment scenarios for basic devices include:

• **Home Alarms:** Such devices typically include motion detectors, magnetic sensors, and smoke detectors. A central unit takes care of the application logic that calls security and sounds an alarm if a sensor is activated when the alarm is armed. The central unit also handles the WAN connection towards the alarm central. These systems are currently often based on proprietary radio protocols.

• **Smart Meters:** The meters are installed in the households and measure consumption of, for example, electricity and gas. A concentrator gateway collects data from the meters, performs aggregation, and periodically transmits the aggregated data to an application server over a cellular connection. By using a capillary network technology it's possible to extend the range of the concentrator gateway by allowing meters in the periphery to use other meters as extenders, and interface with handheld devices on the Home Area Network side.

• **Building Automation Systems (BASs):** Such devices include thermostats, fans, motion detectors, and boilers, which are controlled by local facilities, but can also be remotely operated.

• **Standalone Smart Thermostats:** These use Wi-Fi to communicate with web services. Examples for advanced devices, meanwhile, include:

• **Onboard units** in cars that perform remote monitoring and configuration over a cellular connection.

• **Robots and autonomous vehicles** such as unmanned aerial vehicles that can work both autonomously or by remote control using a cellular connection.

• **Video cameras** for remote monitoring over 3G and LTE.

• **Oil well monitoring** and collection of data points from remote devices.

• **Connected printers** that can be upgraded and serviced remotely.

4.1.2 Basic devices

- These devices are often intended for a single purpose, such as measuring air pressure or closing a valve. I
- In some cases several functions are deployed on the same device, such as monitoring humidity, temperature, and light level.
- The main focus is on keeping the bill of materials (BOM) as low as possible by using inexpensive microcontrollers with built-in memory and storage, often on an SoC-integrated circuit with all main components on one single chip (Figure 5.1).
- Another common goal is to enable battery as a power source, with a lifespan of a year and upwards by using ultra-low energy microcontrollers.

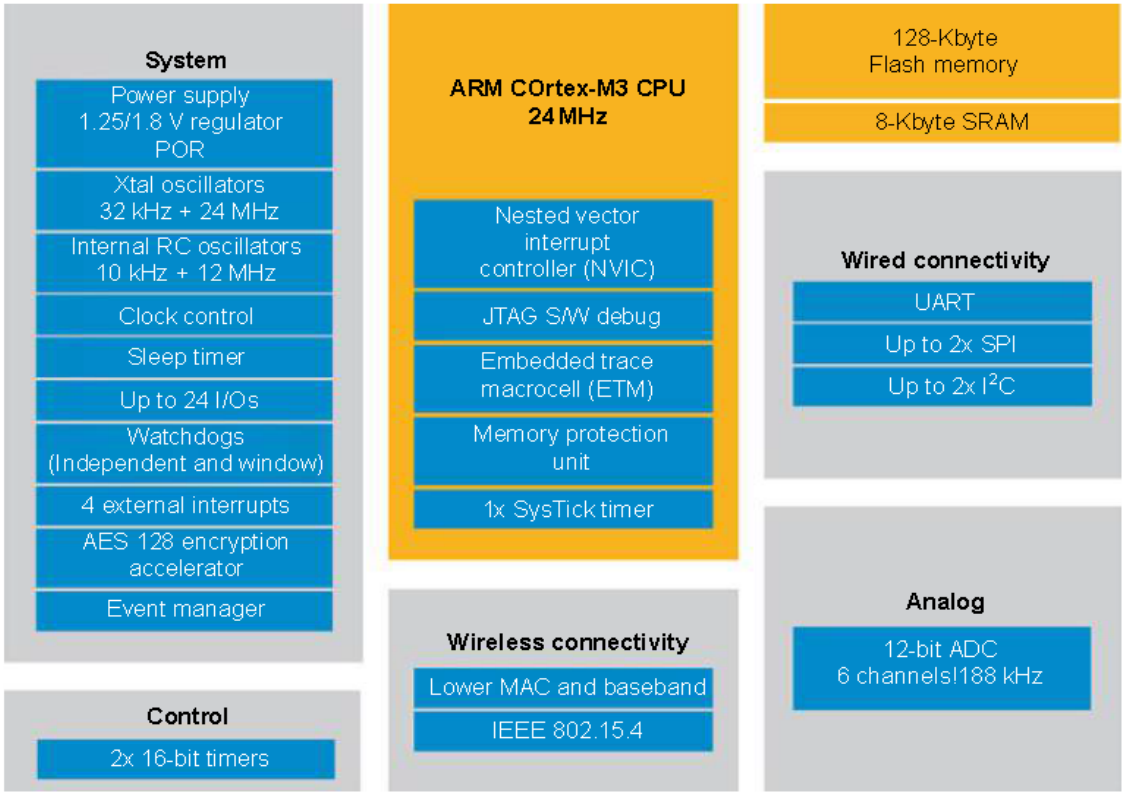

FIGURE 5.1

Example of a microcontroller with integrated STM32W-RFCKIT.

- The microcontroller typically hosts a number of ports that allow integration with sensors and actuators, such as General Purpose I/O (GPIO) and an analog-to-digital converter (ADC) for supporting analog input.
- For certain actuators, such as motors, pulse-width modulation (PWM) can be used.
- As low-power operation is paramount to battery-powered devices, the microcontroller hosts functions that facilitate sleeping, such as interrupts that can wake up the device on external and internal events.

- Some devices even go as far as harvesting energy from their environment, e.g. in the form of solar, thermal, and physical energy.
- To interact with peripherals such as storage or display, it's common to use a serial interface such as SPI, I2C, or UART.
- These interfaces can also be used to communicate with another microcontroller on the device.
- This is common when the there is a need for offloading certain tasks, or when in some cases the entire application logic is put on a separate host processor.
- It's not unusual for the micro controller to also contain a security processor,e.g. to accelerate Advanced Encryption Standard (AES).
- This is necessary to allow encrypted communication over the radio link without the need for a host processor.
- The gateway together with the connected devices form a capillary network.
- The microcontroller contains most of the radio functions needed for communicating with the gateway and other devices in the same capillary network.
- An external antenna is, however, necessary, and preferably a filter that removes unwanted frequencies, e.g. a surface acoustic wave (SAW) filter.
- Due to limited computational resources, these devices commonly do not use a typical OS.
- It may be something as simple as a single-threaded main-loop or a low-end OS such as FreeRTOS, Atomthreads, AVIX-RT, ChibiOS/RT, ERIKA Enterprise, TinyOS, or Thingsquare Mist/Contiki.
- These OSes offer basic functionality, e.g. memory and concurrency model management, (sensor and radio) drivers, threading, TCP/IP, and higher level protocol stacks.
- The actual application logic is located on top of the OS or in the mainloop.
- A typical task for the application logic is to read values from the sensors and to provide these over the LAN interface in a semantically correct manner with the correct units.

4.1.3 Gateways

- A gateway serves as a translator between different protocols, e.g. between IEEE 802.15.4 or IEEE 802.11, to Ethernet or cellular.
- There are many different types of gateways, which can work on different levels in the protocol layers.
- A gateway refers to a device that performs translation of the physical and link layer, but application layer gateways (ALGs) are also common.
- The latter is preferably avoided because it adds complexity and is a common source of error in deployments.
- Some examples of ALGs include the ZigBee Gateway Device which translates from ZigBee to SOAP and IP, or gateways that translate from Constrained Application Protocol (CoAP) to HyperText Transfer Protocol/Representational State Transfer (HTTP/REST).

- Tthe gateway device is also used for many other tasks, such as data management, device management, and local applications.

4.1.3.1 Data management

- Typical functions for data management include performing sensor readings and caching this data, as well as filtering, concentrating, and aggregating the data before transmitting it to back-end servers.

4.1.3.2 Local applications

- Examples of local applications that can be hosted on a gateway include closed loops, home alarm logic, and ventilation control, or the data management function above
- The benefit of hosting this logic on the gateway instead of in the network is to avoid downtime in case of WAN connection failure, minimize usage of costly cellular data, and reduce latency.
- To facilitate efficient management of applications on the gateway, it's necessary to include an execution environment.
- The execution environment is responsible for the lifecycle management of the applications, including installation, pausing, stopping, configuration, and uninstallation of the applications.
- A common example of an execution environment for embedded environments is OSGi, which is based on Java: applications are built as one or more Bundles, which are packaged as Java JAR files and installed using a so-called Management Agent.
- The Management Agent can be controlled from, for example, a terminal shell or via a protocol such as CPE WAN Management Protocol (CWMP).
- Bundle packages can be retrieved from the local file system or over HTTP, for example. OSGi also provides security and versioning for Bundles, which means that communication between Bundles is controlled, and several versions of them can exist.
- The benefit of versioning and the lifecycle management functions is that the OSGi environment never needs to be shut down when upgrading, thus avoiding downtime in the system.
- Also, Linux can be used as an execution environment.

4.1.3.3 Device management

- Device management (DM) is an essential part of the IoT and provides efficient means to perform many of the management tasks for devices:
  • **Provisioning:** Initialization (or activation) of devices in regards to configuration and features to be enabled.
  • **Device Configuration:** Management of device settings and parameters.
  • **Software Upgrades:** Installation of firmware, system software, and applications on the device.

• **Fault Management:** Enables error reporting and access to device status.

- Examples of device management standards include TR-069 and OMA-DM.
- In the simplest deployment, the devices communicate directly with the DM server.
- This is, however, not always optimal or even possible due to network or protocol constraints, e.g. due to a firewall or mismatching protocols.
- In these cases, the gateway functions as mediator between the server and the devices, and can operate in three different ways:
  - If the devices are visible to the DM server, the gateway can simply forward the messages between the device and the server and is not a visible participant in the session.
  - In case the devices are not visible but understand the DM protocol in use, the gateway can act as a proxy, essentially acting as a DM server towards the device and a DM client towards the server.
  - For deployments where the devices use a different DM protocol from the server, the gateway can represent the devices and translate between the different protocols (e.g. TR-069, OMA-DM, or CoAP).
- The devices can be represented either as virtual devices or as part of the gateway

4.1.4 Advanced devices

- An advanced device are the following:
  - A powerful CPU or microcontroller with enough memory and storage to host advanced applications, such as a printer offering functions for copying, faxing, printing, and remote management.
  - A more advanced user interface with, for example, display and advanced user input in the form of a keypad or touch screen.
  - Video or other high bandwidth functions.

4.1.5 Summary and vision

- The most important of these is security, both in terms of physical security as well as software and network security.
- External factors that can affect the operation of the devices, such as rain, wind, chemicals, and electromagnetic influences.
- One of the major effects that the IoT will have on devices is to disrupt the current value chains, where one actor controls everything from device to service.
- This will happen due to standardization and consolidation of technologies, such as protocols, OSes, software and programming languages (e.g. Java for embedded devices), and the business
- New types of actors will be able to enter the market, e.g. specialized device vendors, cloud solution providers, and service providers.

- Standardization will improve interoperability between devices, as well as between devices and services, resulting in commoditization of both.
- Another expected outcome of improved interoperability is the possibility to reuse the same device for multiple services;
- for example, a motion detector can be used both for security purposes as well as for reducing energy consumption by detecting when no one is in the room.
- Thanks to developments in hardware and network technologies, entirely new device classes and features are expected, such as:
  - Battery-powered devices with ultra-low power cellular connections.
  - Devices that harvest energy from their environment.
  - Smart bandwidth management and protocol switching, i.e. using adaptive RF mechanisms to swap between, for example, Bluetooth LE and IEEE 802.15.4.
  - Multi-radio/multi-rate to switch between bands or bit rates
  - Microcontrollers with multicore processors.
  - Novel software architectures for better handling of concurrency.
  - The possibility to automate the design of integrated circuits based on business-level logic and use case.

4.2 Local and wide area networking

4.2.1 The need for networking

- A network is created when two or more computing devices exchange data or information.
- The ability to exchange pieces of information using telecommunications technologies has changed the world
- Devices are known as “nodes” of the network, and they communicate over “links.”
- In modern computing, nodes range from personal computers, servers, and dedicated packet switching hardware, to smart phones, games consoles, television sets and, increasingly, heterogeneous devices that are generally characterized by limited resources and functionalities.
- Limitations typically include computation, energy, memory, communication (range, bandwidth, reliability, etc.) and application specificity (e.g. specific sensors, actuators, tasks), etc. Such devices are typically dedicated to specific tasks, such as sensing, monitoring, and control.
- Network links rely upon a physical medium, such as electrical wires, air, and optical fibers, over which data can be sent from one network node to the next.
- A selected physical medium determines a number of technical and economic considerations.

- Nodes of the network must have an awareness of all nodes in the network with which they can indirectly communicate. This can be a direct connection over one link (edge, the transition or communication between two nodes over a link), or knowledge of a route to the desired (destination) node by communicating through cooperating nodes, over multiple edges.

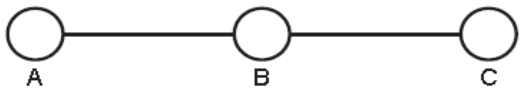

FIGURE 5.2

A network.

- In Figure 5.2 is the simplest form of network that requires knowledge of a route to communicate between nodes that do not have direct physical links.
- if node A wishes to transfer data to node C, it must do so through node B.
- Thus, node B must be capable of the following:
    - Communicating with both node A and node C,
    - advertising to node A and node C that it can act as an intermediary.
- Basic networking requirements have become explicit.
- It is essential to uniquely identify each node in the network, and it is necessary to have cooperating nodes capable of linking nodes between which physical links do not exist.
- In modern computing, this equates to IP addresses and routing tables.
- Consider the differences between streaming video from a surveillance camera, for example, and an intrusion-detection system based on a passive sensor.
- Streaming video requires high bandwidth, whereas transmitting a small amount of information about the detection of an intruder requires a tiny amount of bandwidth, but a higher degree of reliability with respect to both the communications link and the accuracy of the detection.
- Node A is a device that can only communicate over a particular wireless channel of limited range
- Node B is cap able of communicating with node A, but also with an application server with service capabilities (node C, with which it can connect using wired Ethernet, e.g. over a complex link using a standardized protocol and/or web service such as REST at the application layer) over the Internet.
- Node B may be connected to a sub-network (of child nodes, similar to node A) of up to thousands of similarly constrained devices (A1. . .An).
- These thousands of devices may be equipped with sensors, deployed specifically to monitor some physical phenomenon.
- They can only communicate with one another and node B, and may communicate with each other over single or multiple hops.

- Consider that the owner of the WSN wishes to obtain the data from each of the (A1. . .An) devices in the WSN.
- However, the preferred way to read the data is through a web browser, or application on a smartphone/tablet, via node C.
- Therefore, a networking solution is required to transfer all of the WSN data from nodes A1. . .An to node C, through node B.
- This concept maps directly to the M2M Functional Architecture, where nodes A1. . .An are an M2M Area Network, node B is an M2M Gateway, and node C is representative of M2M Service Capabilities and Applications.
- A Local Area Network (LAN) was traditionally distinguishable from a Wide Area Network (WAN) based on the geographic coverage requirements of the network, and the need for third party, or leased, communication infrastructure.
- In the case of the LAN, a smaller geographic region is covered, such as a commercial building, an office block, or a home, and does not require any leased communications infrastructure.
- WANs provide communication links that cover longer distances, such as across metropolitan, regional, or by textbook definition, global geographic areas.
- In practice, WANs are often used to link LANs and Metropolitan Area Networks (MAN)
- LANs tended to cover distances of tens to hundreds of meters, whereas WAN links spanned tens to hundreds of kilometers.
- The most popular wired LAN technology is Ethernet. Wi-Fi is the most prevalent wireless LAN (WLAN) technology.
- Wireless WAN (WWAN), as a descriptor, covers cellular mobile telecommunication networks, a significant departure from WLAN in terms of technology, coverage, network infrastructure, and architecture.
- Difference between LAN and WAN

S.NO	LAN	WAN
1.	LAN stands for Local Area Network.	Whereas WAN stands for Wide Area Network.
2.	LAN’s ownership is private.	But WAN’s ownership can be private or public.
3.	The speed of LAN is	While the speed of WAN is slower

S.NO	LAN	WAN
	high(more than WAN).	than LAN.
4.	The propagation delay is short in LAN.	Whereas the propagation delay in WAN is long(longer than LAN).
5.	There is less congestion in LAN(local area network).	While there is more congestion in WAN(Wide Area Network).
6.	There is more fault tolerance in LAN.	While there is less fault tolerance in WAN.
7.	LAN's design and maintenance is easy.	While it's design and maintenance is difficult than LAN.

- The current generation of WWAN technology includes LTE (or 4G) and WiMAX.
- Acting as a link between LANs and Wireless Personal Area Networks (WPANs), M2M Gateway Devices typically include cellular transceivers, and allow seamless IP-connectivity over heterogeneous physical media.
- In the home, the "wireless router" typically behaves as a link between the Wi-Fi (WLAN, and thus connected laptops, tablets, smartphones, etc. commonly found in the home) and Digital Subscriber Line (DSL) broadband connectivity, traditionally arriving over telephone lines. "DSL" refers to Internet access carried over legacy (wired) telephone networks, and encompasses numerous standards and variants.
- "Broadband" indicates the ability to carry multiple signals over a number of frequencies, with a typical minimum bandwidth of 256 kbps.
- In the office, the Wi-Fi wireless access points are typically connected to the wired corporate (Ethernet) LAN, which is subsequently connected to a wider area network and Internet backbone, typically provided by an Internet Service Provider (ISP).
- The need exists to interconnect devices (generally integrated microsystems) with central data processing and decision support systems, in addition to one another.
- In WLAN technologies, a geographic region can be covered by a network of devices that connect to the Internet via a gateway device, which may use a leased network connection.
- For example, a gateway device can access the IP backbone over a WWAN (e.g. GPRS/UMTS/LTE/WiMAX) link, or over a WLAN link.

- WPANs is the for newer standards that govern low-power, low-rate networks suitable for M2M and IoT applications.
- "IEEE 802.15.4 _ Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR-WPANs).
- This is similar to the evolution of Wi-Fi WLAN technology (e.g. IEEE 802.11, a, b, g, n, etc.).
- Communication ranges for IEEE 802.15.4 technology may range from tens of meters to kilometers.
- Devices in an M2M Area Network connect to the IP backbone, or Network Domain, via an M2M Gateway device.
- Gateway device is equipped with a cellular transceiver that is physically compatible with UMTS or LTE-Advanced, for example, WWAN.
- The same device will also be equipped with the necessary transceiver to communicate on the same physical medium as the M2M Area Network(s) in the M2M Device Domain.
- M2M Area Networks may include a plethora of wired or wireless technologies, including: Bluetooth LE/Smart, IEEE 802.15.4 (LR-WPAN; e.g. ZigBee, IETF 6LoWPAN, RPL, CoAP, ISA100.11a, WirelessHART, etc.),
- The "Internet of Things," as a term, originated from Radio Frequency Identification (RFID) research, wherein the original IoT concept was that any RFID-tagged "thing" could have a virtual presence on the "Internet."
- RFID ,bar codes and QR codes use different technological means to achieve the same result.
- M2M applications become more synonymous with IoT, it is necessary to understand the technologies, limitations, and implications of the networking infrastructure.

4.2.2 Wide area networking

- WANs are typically required to bridge the M2M Device Domain to the backhaul network, thus providing a proxy that allows information (data, commands etc) to traverse heterogeneous networks.
- It is used to provide communications services between the M2M service enablement and the physical deployments of devices in the field.
- WAN is capable of providing the bi-directional communications links between services and devices which is achieved by means of physical and logical proxy.
- The proxy is achieved using an M2M Gateway Device.
- M2M Gateway Device is typically an integrated microsystem with multiple communications interfaces and computational capabilities.
- It is a critical component in the functional architecture, as it must be capable of handling all of the necessary interfacing to the M2M Service Capabilities and Management Functions.

- Example: consider a device that incorporates both an IEEE 802.15.4-compliant transceiver, capable of communicating with a capillary network of similarly equipped devices, and a cellular transceiver that connects to the Internet using the UMTS network.
- Transceivers (sometimes referred to as modems) are typically available as hardware modules with which the central intelligence of the device (gateway or cell phone) interacts by means of standardized AT Commands.
- This device is now capable of acting as a physical proxy between the LR-WPAN, or M2M Device Domain, and the M2M Network Domain.
- The latest ETSI M2M Functional Architecture is illustrated in Figure 5.3.

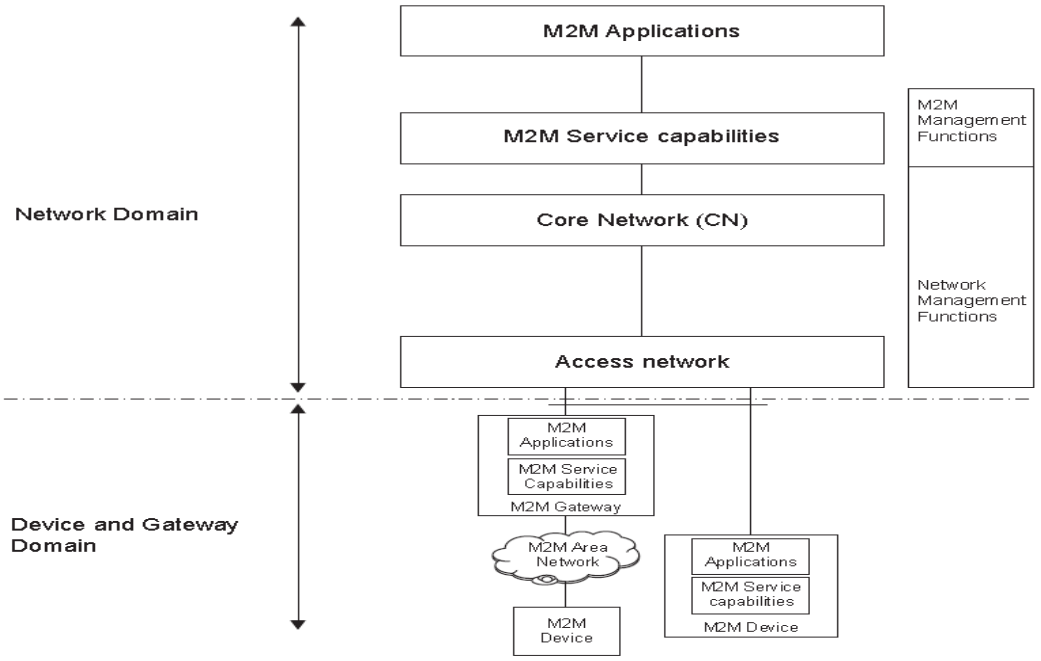

FIGURE 5.3

ETSI M2M Functional Architecture.

- The Access and Core Network in the ETSI M2M Functional Architecture are foreseen to be operated by a Mobile Network Operator (MNO), and can be thought of simply as the "WAN" for the purposes of interconnecting devices and backhaul networks (Internet), thus, M2M Applications, Service Capabilities, Management Functions, and Network Management Functions.
- The WAN covers larger geographic regions using wireless as well as wire-based access.
- WAN technologies include cellular networks (using several generations of technologies), DSL, WiMAX, Wi-Fi, Ethernet, Satellite, and so forth.
- The WAN delivers a packet-based service using IP as default. Circuit-based services can also be used in certain situations.
- important functions of the WAN include:
    - The main function of the WAN is to establish connectivity between capillary

networks, hosting sensors, and actuators, and the M2M service enablement.

- The default connectivity mode is packet-based using the IP family of technologies.
- Many different types of messages can be sent and received. for example, a message sent from a sensor in an M2M Area Network and resulting in an SMS received from the M2M Gateway or Application
- Use of identity management techniques (primarily of M2M devices) in cellular and non-cellular domains to grant right-of-use of the WAN resource.
- The following techniques are used for these purposes:
    - ✓ MCIM (Machine Communications Identity Module) for remote provisioning of SIM targeting M2M devices.
    - ✓ xSIM (x-Subscription Identity Module), like SIM, USIM, ISIM.
    - ✓ Interface identifiers, an example of which is the MAC address of the device, typically stored in hardware.
    - ✓ Authentication/registration type of functions (device focused).
    - ✓ Authentication, Authorization, and Accounting (AAA), such as RADIUS services.
    - ✓ Dynamic Host Configuration Protocol (DHCP), e.g. employing deployment-specific configuration parameters specified by device, user, or application-specific parameters residing in a directory.
    - ✓ Subscription services (device-focused).
    - ✓ Directory services, e.g. containing user profiles and various device (s) parameter(s), setting(s), and combinations thereof.
- M2M-specific considerations include, in particular:
    - ✓ MCIM (cf. 3GPP SA3 work).
    - ✓ User Data Management (e.g. subscription management).
    - ✓ Network optimizations (cf. 3GPP SA2 work).

4.2.2.1 3rd generation partnership project technologies and machine type communications

- Machine Type Communications (MTC) is heavily referred to in the ETSI documentation.
- MTC refers to small amounts of data that are communicated between machines (devices to back-end services and vice versa) without the need for any human intervention. In the 3rd Generation Partnership Project (3GPP), MTC is used to refer to all M2M communication.

4.2.3 Local area networking

- Capillary networks are typically autonomous, self-contained systems of M2M devices that may be connected to the cloud via an appropriate Gateway.

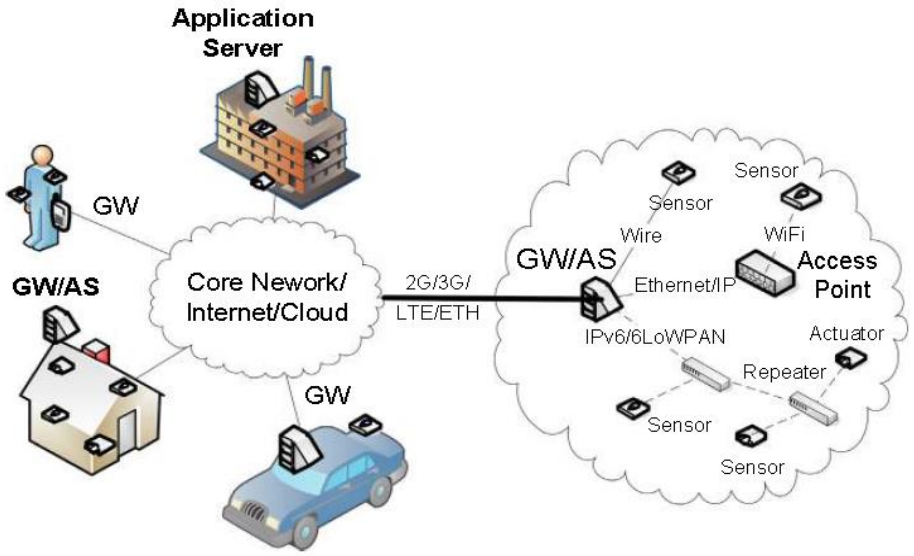

FIGURE 5.4

Capillary networks and their inside view.

- They are often deployed in controlled environments such as vehicles, buildings, apartments, factories, bodies, etc. (Figure 5.4) in order to collect sensor measurements, generate events should sensing thresholds be breached, and sometimes control specific features of interest (e.g. heart rate of a patient, environmental data on a factory floor, car speed, air conditioning appliances, etc.).
- There will exist numerous capillary networks that will employ short-range wired and wireless communication and networking technologies.
- For certain application areas, there is a need for autonomous local operation of the capillary network.
- In the event that application-level logic is enforceable via the cloud, some will still need to be managed locally.
- The complexity of the local application logic varies by application.
- For example, a building automation network may need local control loop functionality for autonomous operation, but can rely on external communication for configuration of control schemas and parameters.
- The M2M devices in a capillary network are typically thought to be low-capability nodes (e.g. battery operated, with limited security capabilities) for cost reasons, and should operate autonomously.
- For this reason, a GW/application server will naturally also be part of the architected solution for capillary networks.
- More and more (currently closed) capillary networks will open up for integration with the enterprise back end systems.
- For capillary networks that expose devices to the cloud/Internet, IP is envisioned to be the common waist.

- IPv6 will be the protocol of choice for M2M devices that operate a 6LoWPAN-based stack.
- IPv4 will still be used for capillary networks operating in non-6LoWPAN IP stacks (e.g. Wi-Fi capillary networks).
- In terms of short-range communication technology convergence, an IPv6 stack with 6LoWPAN running above the physical medium is expected.
- The development of the IEEE 802.15.4g standard, a physical layer amendment to support Smart Utility Networks (SUN) _ smart grid in particular _ designed to operate over much larger geographic distances (wireless links spanning tens of kilometers), and specifically designed for minimal infrastructure, low power, many-device networks.

4.2.3.1 Deployment considerations

- There are increasing numbers of innovative IoT applications (hardware and software) marketed as consumer products.
- These range from intelligent thermostats for effectively managing comfort and energy use in the home, to precision gardening tools (sampling weather conditions, soil moisture, etc.).
- Scaling up for industrial applications and moving from laboratories into the real world creates significant challenges that are not yet fully understood.
- Low-rate, low-power communications technologies are known to be “lossy.” The reasons can relate to environmental factors, which impact upon radio performance, technical factors such as performance trade-offs based on the characteristics of medium access control and routing protocols, and physical limitations of devices (including software architectures, runtime and execution environments, computational capabilities, energy availability, local storage, etc), and practical factors such as maintenance opportunities (scheduled, remote, accessibility, etc.).
- Numerous deployment environments (factories, buildings, roads, vehicles) are expected in addition to wildly varying application scenarios and operational and functional requirements of the systems.
- ETSI describes a set of use cases, namely eHealth, Connected Consumer, Automotive, Smart Grid, and Smart Meter, that only capture some of the breadth of potential deployment scenarios and environments that are possible.
- Assuming that IP connectivity can be the fundamental mechanism to bridge heterogeneous physical and link layer technologies, it stands to reason that fragmentation can continue such that appropriate technologies are available for the breadth of potential application scenarios.

4.2.3.2 Key technologies

- Power Line Communication (PLC) refers to communicating over power (or phone, coax, etc.) lines.

- This amounts to pulsing, with various degrees of power and frequency, the electrical lines used for power distribution.
- PLC comes in numerous flavors. At low frequencies (tens to hundreds of Hertz) it is possible to communicate over kilometers with low bit rates (hundreds of bits per second). Typically, this type of communication was used for remote metering, and was seen as potentially useful for the smart grid.
- Enhancements to allow higher bit rates have led to the possibility of delivering broadband connectivity over power lines.
- There have been a number of attempts to standardize PLC in recent years. NIST recently included IEEE 1901 and ITU-T G.hn as standards for further review for potential use in the smart grid in the United States.
- LAN (and WLAN) continues to be important technology for M2M and IoT applications.
- This is due to the high bandwidth, reliability, and legacy of the technologies. Where power is not a limiting factor, and high bandwidth is required, devices may connect seamlessly to the Internet via Ethernet (IEEE 802.3) or Wi-Fi (IEEE 802.11).
- The IEEE 802.11 (Wi-Fi) standards continue to evolve in various directions to improve certain operational characteristics depending on usage scenario.
- A widely adopted recent release was IEEE 802.11n, which was specifically designed to enhance throughput (typically useful for streaming multimedia).
- Ongoing work such as IEEE 802.11ac is developing an even higher throughput version to replace this, focusing efforts in the 5 GHz band.
- IEEE 802.11ah is allow a number of networked devices to cooperate in the ,1 GHz (ISM) band.
- The idea is to exploit collaboration (relaying, or networking in other words) to extend range, and improve energy efficiency (by cycling the active periods of the radio transceiver).
- Bluetooth Low Energy (BLE; “Bluetooth Smart”) is designed for short-range (,50 m) applications in healthcare, fitness, security, etc., where high data rates (millions of bits per second) are required to enable application functionality.
- It is deliberately low cost and energy efficient by design, and has been integrated into the majority of recent smart phones.
- Low-Rate, Low-Power Networks are another key technology that form the basis of the IoT.
- For example, the IEEE 802.15.4 family of standards was one of the first used in practical research and experimentation in the field of WSNs.
- Low-Rate Wireless Personal Area Networks (LR-WPAN)- It covered the Physical and Medium Access Control layers, specifying use in the ISM bands at frequencies around 433 MHz, 868/915 MHz, and 2.4 GHz. This supported data rates of between 20 kbps up to 256 kbps, depending on selected band, over distances ranging from tens of meters to kilometers.

- Radio duty cycling refers to managing the active periods of the Radio Frequency Integrated Circuit (RFIC) during transmission, and listening to the medium.
- IEEE 802.15.4 defines the PHY layer, and in some instances the MAC layer, upon which a number of low-energy communications specifications have been built. Namely, ZigBee.
- Recent developments, such as the PHY Amendment for Smart Utility Networks (SUN), IEEE 802.15.4g, seek to extend the operational coverage of these networks up to tens of kilometers in order to provide extremely wide geographic coverage with minimal infrastructure.
- 6LoWPAN (IPv6 Over Low Power Wireless Personal Area Networks) was developed initially by the 6LoWPAN Working Group (WG) of the IETF as a mechanism to transport IPv6 over IEEE 802.15.4-2003 networks.
- Specifically, methods to handle fragmentation, reassembly, and header compression were the primary objectives.
- The WG also developed methods to handle address autoconfiguration, the hooks for mesh networking, and network management.
- RPL (IPv6 Routing Protocol for Low Power and Lossy Networks) was developed by the IETF Routing over Low Power and Lossy Networks (RoLL) WG.
- They defined Low Power Lossy Networks as those typically characterized by high data loss rates, low data rates, and general instability.
- No specific physical or medium access control technologies were specified, but typical links considered include PLC, IEEE 802.15.4, and low-power Wi-Fi.
- Typical use cases involve the collection of data from many (for example) sensing points, nodes towards a sink, or alternatively, flooding information from a sink to many nodes in the network.
- Thus, the well-known concept of a Directed Acyclic Graph (DAG) structure was concentrated to a Destination Oriented DAG (DODAG) for the purposes of initial development.
- The group defined a new ICMPv6 message, with three possible types, specific for RPL networks.
- These include a DAG Information Object (DIO), that allows a node to discover an RPL instance, configuration parameters and parents, a DAG Information Solicitation (DIS) to allow requests for DIOs from RPL nodes, and Destination Advertisement Object (DAO), used to propagate destination information upwards (i.e. towards the root) along the DODAG (specific RPL details are available in RFC 6550 and related RFCs).
- The Trickle Algorithm is an important enabler of RPL message exchange.
- CoAP (Constrained Application Protocol) is being developed by the IETF Constrained RESTful Environments (CoRE) WG as a specialized web transfer protocol for use with severe computational and communication constraints typically characteristic of M2M and IoT applications.

4.3 Data management

4.3.1 Introduction

- In the era of M2M, where billions of devices interact and generate data at exponential growth rates, data management is of critical importance as it sets the basis upon which any other processes can rely and operate
- Some of the key characteristics of M2M data include:

  • **Big Data:** Huge amounts of data are generated, capturing detailed aspects of the processes where devices are involved.

  • **Heterogeneous Data:** The data is produced by a huge variety of devices and is itself highly heterogeneous, differing on sampling rate, quality of captured values, etc.

  • **Real-World Data:** The overwhelming majority of the M2M data relates to real-world processes and is dependent on the environment they interact with.

  • **Real-Time Data:** M2M data is generated in real-time and overwhelmingly can be communicated also in a very timely manner.

  • **Temporal Data:** The overwhelming majority of M2M data is of temporal nature, measuring the environment over time.

  • **Spatial Data:** Increasingly, the data generated by M2M interactions are not only captured by mobile devices, but also coupled to interactions in specific locations, and their assessment may dynamically vary depending on the location.

  • **Polymorphic Data:** The data acquired and used by M2M processes may be complex and involve various data, which can also obtain different meanings depending on the semantics applied and the process they participate in.

  • **Proprietary Data:** Up to now, due to monolithic application development, a significant amount of M2M data is stored and captured in proprietary formats. However, increasingly due to the interactions with heterogeneous devices and stakeholders, open approaches for data storage and exchange are used.

  • **Security and Privacy Data Aspects:** Due to the detailed capturing of interactions by M2M, analysis of the obtained data has a high risk of leaking private information and usage patterns, as well as compromising security.\

4.3.2 Managing M2M data

- The data flow from the moment it is sensed (e.g. by a wireless sensor node) up to the moment it reaches the backend system has been processed manifold (and often redundantly), either to adjust its representation in order to be easily integrated by the diverse applications, or to compute on it in order to extract and associate it with respective business intelligence (e.g. business process affected, etc.).

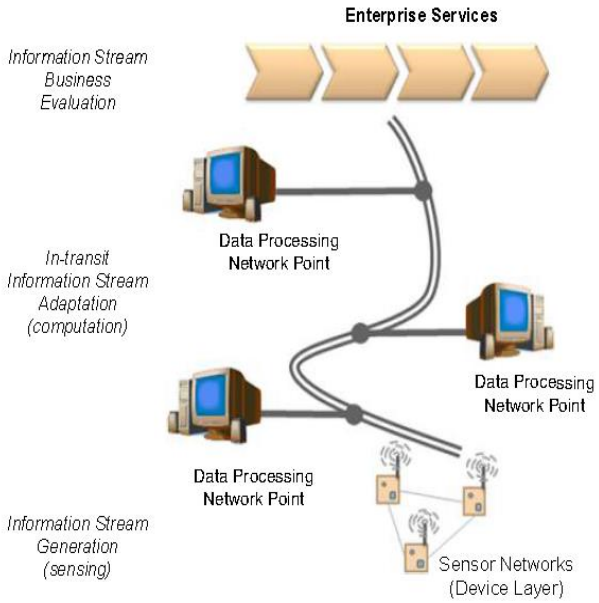

FIGURE 5.5

M2M data from point of generation to business assessment.

- In Figure 5.5, we see a number of data processing network points between the machine and the enterprise that act on the datastream (or simply forwarding it) based on their end-application needs and existing context.
- Dealing with M2M data may be decomposed into several stages.
- Additionally, the degree of focus in each stage heavily depends on the actual usage requirements put upon the data as well as the infrastructure.

4.3.2.1 Data generation

- Data generation is the first stage within which data is generated actively or passively from the device, system, or as a result of its interactions.
- The sampling of data generation depends on the device and its capabilities as well as potentially the application needs.
- Usually default behaviors for data generation exist, which are usually further configurable to strike a good benefit between involved costs, e.g. frequency of data collection vs. energy used in the case of WSNs, etc.

4.3.2.2 Data acquisition

- Data acquisition deals with the collection of data (actively or passively) from the device, system, or as a result of its interactions.
- The data acquisition systems usually communicate with distributed devices over wired or wireless links to acquire the needed data, and need to respect security, protocol, and application requirements.
- The nature of acquisition varies, e.g. it could be continuous monitoring, interval-poll, event-based, etc.
- The frequency of data acquisition overwhelmingly depends on, or is customized by, the application requirements (or their common denominator).
- The data acquired at this stage (for non-closed local control loops) may also differ from the data actually generated.
- In simple scenarios, due to customized filters deployed at the device, a fraction of the generated data may be communicated.
- Data aggregation and even on-device computation of the data may result in communication of key performance indicators of interest to the application.

4.3.2.3 Data validation

- Data acquired must be checked for correctness and meaningfulness within the specific operating context.
- This is usually done based on rules, semantic annotations, or other logic.
- The acquired data may not conform to expectations and data may be intentionally or unintentionally corrupted during transmission, altered, or not make sense in the business context.
- As real-world processes depend on valid data to draw business-relevant decisions
- Several known methods are deployed for consistency and data type checking;
- for example, imposed range limits on the values acquired, logic checks, uniqueness, correct time-stamping, etc.
- In addition, semantics may play an increasing role here, as the same data may have different meanings in various operating contexts, and via semantics one can benefit while attempting to validate them.
- Another part of the validation may deal with fallback actions such as requesting the data again if checks fail, or attempts to “repair” partially failed data.
- Failure to validate may result in security breaches.
- Tampered-with data fed to an application is a well known security risk as its effects may lead to attacks on other services, privilege escalation, denial of service, database corruption, etc.

4.3.2.4 Data storage

- The data generated by M2M interactions is what is commonly referred to as "Big Data."
- Machines generate an incredible amount of information that is captured and needs to be stored for further processing.
- As this is proving challenging due to the size of information, a balance between its business usage vs. storage needs to be considered; that is, only the fraction of the data relevant to a business need may be stored for future reference.
- However, one has to carefully consider what the value of such data is to business not only in current processes, but also potentially other directions that may be followed in the future by the company as different assessments of the same data may provide other, hidden competitive advantages in the future.
- Due to the massive amounts of M2M data, as well as their envisioned processing (e.g. searching), specialized technologies such as massively parallel processing DBs, distributed file systems, cloud computing platforms, etc. are needed.

4.3.2.5 Data processing

- Data processing enables working with the data that is either at rest (already stored) or is in-motion (e.g. stream data).
- The scope of this processing is to operate on the data at a low level and "enhance" them for future needs.
- Typical examples include data adjustment during which it might be necessary to normalize data, introduce an estimate for a value that is missing, re-order incoming data by adjusting timestamps, etc.
- Similarly, aggregation of data or general calculation functions may be operated on two or more data streams and mathematical functions applied on their composition.
- Another example is the transformation of incoming data; for example, a stream can be converted on the fly (e.g. temperature values are converted from _F to _C), or repackaged in another data model, etc. Missing or invalid data that is needed for the specific time-slot may be forecasted and used until, in a future interaction, the actual data comes into the system.

4.3.2.6 Data remanence

- Even if the data is erased or removed, residues may still remain in electronic media, and may be easily recovered by third parties _ often referred to as data remanence.
- Several techniques have been developed to deal with this, such as overwriting, degaussing, encryption, and physical destruction.
- For M2M, not only the DBs where the M2M data is collected, but also the points of action, which generate the data, or the individual nodes in between, which may cache it.

- At the current technology pace, those buffers (e.g. on device) are expected to be less at risk since their limited size means that after a specific time has elapsed, new data will occupy that space; hence, the window of opportunity is rather small.
- In addition, for large-scale infrastructures the cost of potentially acquiring "deleted" data may be large; hence, their hubs or collection end-points, such as the DBs who have such low cost, may be more at risk.

4.3.2.7 Data analysis

- Data available in the repositories can be subjected to analysis with the aim to obtain the information they encapsulate and use it for supporting decision-making processes.
- The analysis of data at this stage heavily depends on the domain and the context of the data.
- For instance, business intelligence tools process the data with a focus on the aggregation and key performance indicator assessment.
- Data mining focuses on discovering knowledge, usually in conjunction with predictive goals.
- Statistics can also be used on the data to assess them quantitatively (descriptive statistics), find their main characteristics (exploratory data analysis), confirm a specific hypothesis (confirmatory data analysis), discover knowledge (data mining), and for machine learning, etc.
- This stage is the basis for any sophisticated applications that take advantage of the information hidden directly or indirectly on the data.

4.3.3 Considerations for M2M data

- The M2M infrastructure in place heavily depends on real-world processes, implying also that a big percentage of data will be generated by machines that interact with the real-world environment, while the rest will be purely virtual data.
- Many of the machines generating this data, which can then be communicated to others (e.g. analytics specialists).
- The end-beneficiaries might acquire information, but do not necessarily need to have access or to process the data by themselves.
- There is a rise of specialists in the various stages of M2M data management that will cooperate with application providers, users, etc. for the common benefit.
- Sharing of data and usage in multiple applications, security and trust are of key importance.
- Security is mandatory for enabling confidentiality, integrity, availability, authenticity, and nonrepudiation of data from the moment of generation to consumption.
- Due to the large-scale IoT infrastructure, heterogeneous devices, and stakeholders involved, this will be challenging.

- In addition, trust will be another major issue, as even if data is securely communicated or verified, the level of trust based on them will impact the decision-making process and risk analysis.
- Managing security and trust in the highly federated M2M-envisioned infrastructures poses a significant challenge, especially for mission critical applications that also exercise control.
- Privacy is also expected to be a significant issue in IoT infrastructures.
- Currently, a lot of emphasis is put on acquiring the data, and no real solutions exist for large-scale systems to share data in a controlled way.
- Once data is shared, the originator has no more control over its lifetime.
- A typical example here constitutes the usage of private citizen data, which could be controllably shared as wished; it should also be possible to (partially) revoke that right at will.
- Data Science in the IoT era is a cross-discipline approach building on mathematics, statistics, high-performance computing, modeling, machine learning, engineering, etc. that will play a key role in understanding the data, assessing their information at large scale, and hopefully enabling the better studying of complex systems of systems and their emergent characteristics.

4.3.4 Conclusions

- Data and its management hold the key to unveiling the true power of M2M and IoT.
- To do so, however, we have to think and develop approaches that go beyond simple data collection, and enable the management of their whole lifecycle at very large scale, while in parallel considering the special needs and the usage requirements posed by specific domains or applications.

4.4 Business processes in IoT

4.4.1 Introduction

- A business process refers to a series of activities, often a collection of interrelated processes in a logical sequence, within an enterprise, leading to a specific result.
- There are several types of business processes such as management, operational, and supporting, all of which aim at achieving a specific mission objective.
- As business processes usually span several systems and may get very complex, several methods and techniques have been developed for their modeling, such as the Business Process Model and Notation (BPMN), which graphically represents business processes in a business process model.

- Several key business processes in modern enterprise systems heavily rely on interaction with real-world processes, largely for monitoring, but also for some control (management), in order to take business-critical decisions and optimize actions across the enterprise.

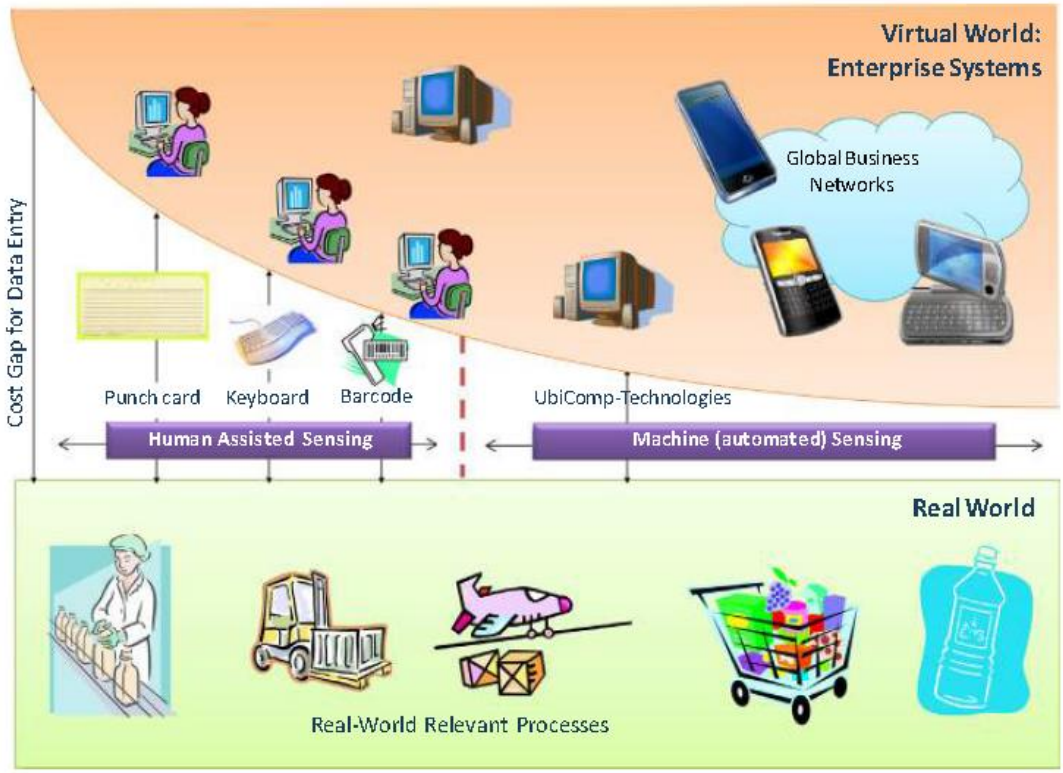

FIGURE 5.6

The decreasing cost of information exchange between the real-world and enterprise systems with the advancement of M2M.

- In Figure 5.6, the dramatic reduction of the data acquisition from the real world
- Initially all these interactions were human-based (e.g. via a keyboard) or human-assisted (e.g. via a barcode scanner); however, with the prevalence of RFID, WSNs, and advanced networked embedded devices, all information exchange between the real-world and enterprise systems can be done automatically without any human intervention and at blazing speeds.
- In the M2M era, connected devices can be clearly identified, and with the help of services, this integration leads to active participation of the devices to the business processes.
- Existing modeling tools are hardly designed to specify aspects of the real world in modeling environments and capture their full characteristics. To this direction, the existence of SOA-ready devices

- (i.e. devices that offer their functionalities as a web service) simplifies the integration and interaction as they can be considered as a traditional web service that runs on a specific device.
- A layered approach for developing, deploying, and managing WSN applications that natively interact with enterprise information systems such as a business process engine and the processes running therein is proposed and assessed.
- M2M and IoT empower business processes to acquire very detailed data about the operations, and be informed about the conditions in the real world in a very timely manner.

4.4.2 IoT integration with enterprise systems

- M2M communication and the vision of the IoT pose a new era where billions of devices will need to interact with each other and exchange information in order to fulfill their purpose.

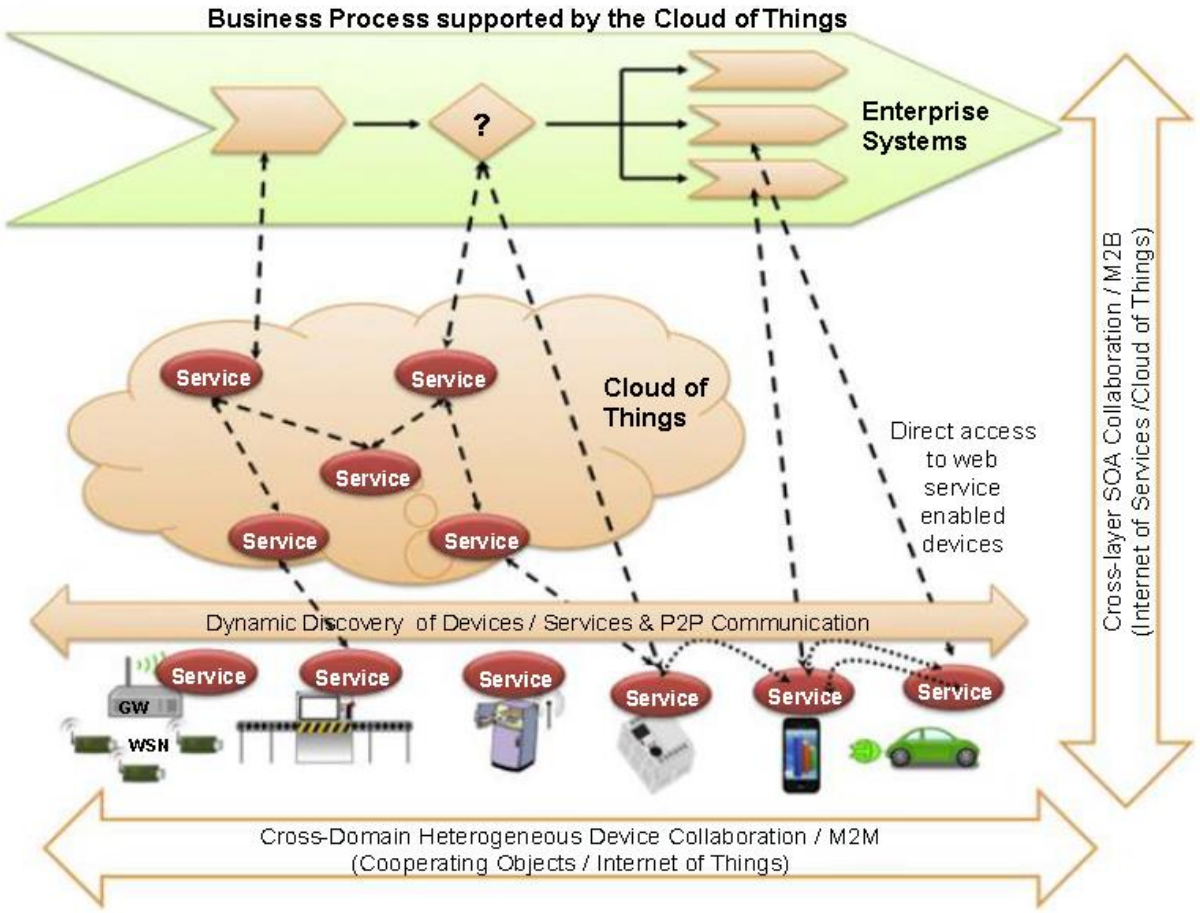

FIGURE 5.7

A collaborative infrastructure driven by M2M and M2B.

- In Figure 5.7, cross-layer interaction and cooperation can be pursued:

  • at the M2M level, where the machines cooperate with each other (machine-focused interactions)

  • at the machine-to-business (M2B) layer, where machines cooperate also with network-based services, business systems (business service focus), and applications.

- Several devices in the lowest layer. These can communicate with each other over short-range protocols (e.g. over ZigBee, Bluetooth), or even longer distances (e.g. over Wi-Fi, etc.).
- Some of them may host services (e.g. REST services), and even have dynamic discovery capabilities based on the communication protocol or other capabilities (e.g. WS-Eventing in DPWS).
- Some of them may be very resource constrained, which means that auxiliary gateways could provide additional support such as mediation of communication, protocol translation, etc.
- Independent of whether the devices are able to discover and interact with other devices and systems directly or via the support of the infrastructure, the M2M interactions enable them to empower several applications and interact with each other in order to fulfill their goals.
- Promising real-world integration is done using a service-oriented approach by interacting directly with the respective physical elements, for example, via web services running on devices (if supported) or via more lightweight approaches such as REST.
- Many of the services that will interact with the devices are expected to be network services available, for example, in the cloud.
- The main motivation for enterprise services is to take advantage of the cloud characteristics such as virtualization, scalability, multi-tenancy, performance, lifecycle
- management, etc.
- A key motivator is the minimization of communication overhead with multiple endpoints by, for example, transmission of data to a single or limited number of points in the network, and letting the cloud do the load balancing and further mediation of communication.
- Content Delivery Network (CDN) can be used in order to get access to the generated data from locations that are far away from the M2M infrastructure (geographically, network-wise, etc.).
- To this end, the data acquired by the device can be offered without overconsumption
- of the device's resources, while in parallel, better control and management can be applied.

4.4.3 Distributed business processes in IoT

- In Figure 5.9, the integration of devices in business processes merely implies the acquisition of data from the device layer, its transportation to the backend systems, its assessment, and once a decision is made, potentially the control (management) of the device, which adjusts its behavior.
- In future, due to the large scale of IoT, as well as the huge data that it will generate, such approaches are not viable.
- Enterprise systems trying to process such a high rate of non- or minor-relevancy data will be overloaded.

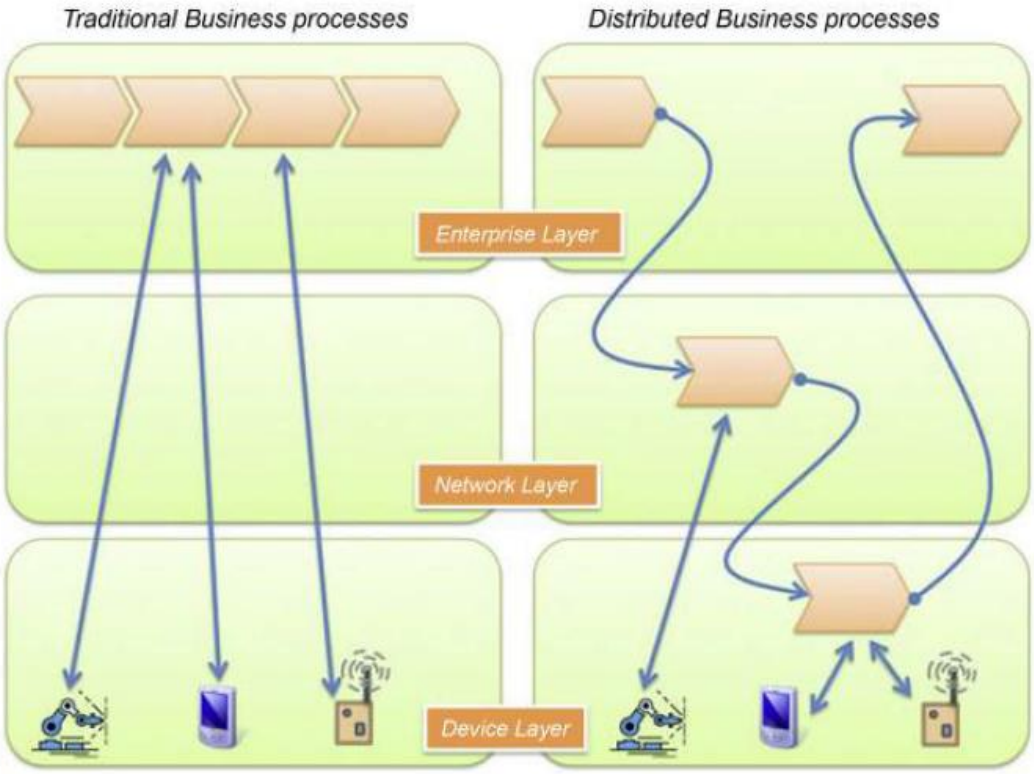

FIGURE 5.9

Distributed Business Processes in M2M era.

- The first step is to minimize communication with enterprise systems to only what is relevant for business. With the increase
- in resources (e.g. computational capabilities) in the network, and especially on the devices themselves (more memory, multi-core CPUs, etc.), it makes sense not to host the intelligence and the computation required for it only on the enterprise side, but actually distribute it on the network, and even on the edge nodes (i.e. the devices themselves), as depicted on the right side of Figure 5.9.

- Partially outsourcing functionality traditionally residing in backend systems to the network itself and the edge nodes means we can realize distributed business processes whose sub-processes may execute outside the enterprise system.
- As devices are capable of computing, they can either realize the task of processing and evaluating business relevant information they generate by themselves or in clusters.
- Business processes can bind during execution of dynamic resources that they discover locally, and integrate them to better achieve their goals.

4.4.4 Considerations

- Existing tools and approaches need to be extended to the make the business processes IoT aware.
- Distributed execution of processes exists (e.g. in BPMN), additional work is needed to be able to select the devices in which such processes execute and consider their characteristics or dynamic resources, etc.
- The dynamic aspect is of key importance in the IoT, as this is mobile and availability is not guaranteed, which means that availability in modeling time does not guarantee availability at runtime and vice-versa.
- Scalability is an aspect that needs to be considered in the business process modeling and execution.
- In addition, event-based interactions among the processes play a key role in IoT, as a business process flow may be influenced by an event, or as its result, trigger a new event.

4.4.5 Conclusions

- Modern enterprises operate on a global scale and depend on complex business processes.
- Efficient information acquisition, evaluation, and interaction with the real world are of key importance.
- The infrastructure envisioned is a heterogeneous one, where millions of devices are interconnected, ready to receive instructions and create event notifications, and where the most advanced ones depict self-behavior (e.g. self-management, self-healing, selfoptimization, etc.) and collaborate.
- Business logic can now be intelligently distributed to several layers such as the network, or even the device layer, creating new opportunities, but also challenges that need to be assessed.
- Future Enterprise systems will be in position to better integrate state and events of the physical world in a timely manner, and hence to lead to more diverse, highly dynamic, and efficient business applications.

4.5 Everything as a service (XaaS)

- Cloud computing is a model for enabling ubiquitous, on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications, and services) that can be provisioned, configured, and made available with minimal management effort or service provider interaction.
- All applications need access to three things: compute, storage, and data processing capacities.
- With cloud computing, a fourth element is added _ distribution services _ i.e. the manner in which the data and computational capacity are linked together and coordinated.

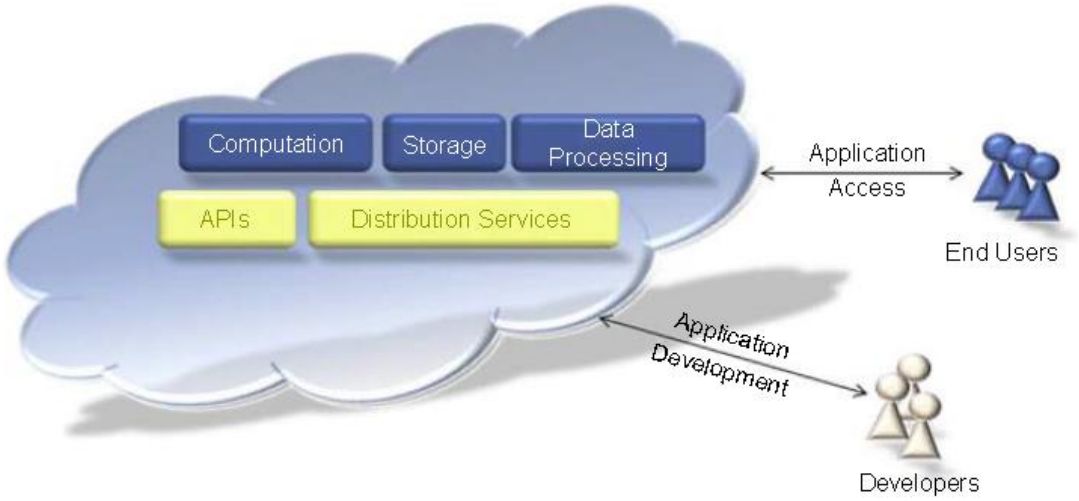

FIGURE 5.11

Conceptual Overview of Cloud Computing.

Characteristics of cloud computing

- On-Demand Self-Service.

  A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed, or automatically, without requiring human interaction with each service provider.
- Broad Network Access.

  Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g. mobile phones, tablets, laptops, and workstations).
- Resource Pooling.

  The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically

assigned and reassigned according to consumer demand. Examples of resources include storage, processing, memory, and network bandwidth.

- Rapid Elasticity.

  Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand.

- Measured Service.
  - ✓ Cloud systems automatically control and optimize resource use by leveraging a metering capability, at some level of abstraction, appropriate to the type of service (e.g. storage, processing, bandwidth, and active user accounts).
  - ✓ Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

For M2M and IoT, these infrastructures provide the following:

1. Storage of the massive amounts of data that sensors, tags, and other "things" will produce.

2. Computational capacity in order to analyze data rapidly and cheaply.

3. Over time, cloud infrastructure will allow enterprises and developers to share datasets, allowing for rapid creation of information value chains.