The IBM Advantage for Implementing the CSCC Cloud Customer Reference Architecture for Internet of Things (IoT)

Introduction

This paper describes how you can use IBM products and services to support the best practices for architected Internet of Things (IoT) solutions provided in the Cloud Customer

You can use the architectural components described in the CSCC paper to build IoT solutions using cloud computing components. These components can be mapped to a range of IoT devices and distributed IoT systems appropriate to the nature of the physical entities monitored and controlled by the system. We also include recommendations for using IBM products and services to deploy and manage IoT systems that align with the architecture defined in the CSCC paper.

Before we look at the actual architecture for an IoT solution, let’s take a look at some of the factors that are driving the need for more IoT solutions.

Four key technology shifts are driving the need for IoT systems. These shifts are:

• Availability of more, less expensive and widely installable IoT devices
• Advanced analytics, which can derive actionable insights from masses of device data
• Cloud computing as a growth engine for business
• New ways for businesses to engage with customers

Growth opportunities provided by IoT and insight from IoT data across the enterprise can give companies in any industry a competitive advantage. IoT helps companies realize greater innovation, more effective operations, and increased customer and employee engagement.

**Innovation**

• Use direct feedback from products instrumented with IoT sensors to driven innovation in product development
• Apply IBM’s strength in advanced analytics and cognitive insights to drive new business opportunities from real-world measurements
• Be more competitive through better business engagement by incorporating cognitive insights, weather, analytics, security, and data streaming capabilities into solutions
• Gain competitive advantage over competitors by being first to market using techniques such as cognitive personalization of connected products
• Employ remote monitoring of equipment in the field to change service and support from being reactive to being proactive - enabling new business models of selling equipment as a service

**Operations**

• Enhance operations by applying real-time responsiveness to optimize asset productivity and increase operational efficiency
• Use IoT data and cognitive insights to optimize the use of resources (worker, energy, expertise)
• Provide safer work environments; by connecting sensors in work enviroments and on workers to detect and address hazardous conditions
• Transform automotive industries by gathering data from vehicle sensors, combine this data with other data sources for real-time analysis, and provide actionable insights for both drivers and for service and support
• Improve collaboration across operations, maintenance, reliability, and engineering, supporting and contributing to operational excellence
• Enable organizations to get better insights from their assets to ensure performance and improve associated processes to those assets to sense, communicate, and self-diagnose issues of intelligent assets and equipment so they can optimize their
performance and reduce unnecessary downtime

Customer and employee engagement
• Provide equipment manufacturers the opportunity to engage and form relationships with the end user by instrumenting equipment, creating a business model that drives consumer engagement and lowers field service costs
• Allow insurers to provide more value to policyholders by offering proactive protection of their assets with notifications of potential problems, managing risk, and improving customer satisfaction, embracing the future with intelligent home, auto, business
• Digitize the retail store experience to yield detailed information and analytics about customer visits, including demographics and conversions for optimized channel distribution, inventory, and campaigns

Functional Requirements
• Easy-to-use, secure applications
• Ability to add new IoT devices to the system with minimal effort
• User opt-in to share personal information including location
• Smart homes equipped with the network of sensors, interconnected devices, and gateways
• Cloud IoT platform with robust device management, data identity services, and analytics
• Enterprise network, containing existing enterprise applications, services, and data
• System should provide operational alert and notification support for medical devices
• Device registration and user authentication to provide authentication services that directly bind an identity (for example, user, mobile device, vehicle, application) to its digital identity
• Reporting and analytics capabilities to create key performance indicators (for example, dashboards, graphs, and charts to view risk, compliance, and audit metrics by a variety of parameters, analytic tools with issue tracking and reporting functionality with graphical dashboards)
• Real-time alerts for maintenance and security monitoring
• Plug-and-play interoperability between IoT devices with the adoption of open standards
• Straightforward integration with related data feeds such as weather information

Non-functional Requirements
To create an effective IoT system, you must account for non-functional requirements like security and risk management, scalability, RAS, and mobile support.

Security and risk management [2] [3]
• Ensure protection of personal data
• Protection of the environment communicating with the device; networks need to be protected to prevent hackers from finding a way to intercept network
• Support for authentication (device, system/application, and user), authorization, auditing, administration, encryption/decryption, data integrity, and key management, and managing identity and cryptographic key information
• All devices in the environment must be managed and maintained, and devices, gateways, routers, and other infrastructure must be regularly updated to apply all
security patches and fixes
  • Ability to detect, respond, resist, and recover from attacks
  • Transactional integrity for procurement, purchase, and supply-chain processes, including manufacturing and delivery; prevent introduction of incorrect data or program codes; ensure the physical security of the production environment where devices and systems are manufactured

Safety
  • For IoT systems that have actuators operating on real-world things, safety is a primary consideration and systems must be designed to fail safe and to ensure the safety of humans and equipment

Scalability
The number of concurrent devices and users connecting to the IoT platform must be scalable. The solution platform must scale to support the next generation of devices and to anticipate a newer generation of connected devices that will furnish higher resolution of data streams

Reliability, availability, and serviceability (RAS)
High availability and resiliency of cloud IoT infrastructure and enterprise environments

Mobile support
The IoT device must provide a gateway to enable mobile computing devices such as mobile phones and tablets to serve as gateways to the cloud-based analytics platform. Clearly differentiate between a mobile device that’s acting as an IoT or gateway device from one that’s simply a user interface provider (user input/output device) in an IoT solution.

Cloud Customer Reference Architecture for IoT

Figure 1 shows the elements that may be needed for any IoT solution.
IBM offers a solution for most elements shown above, other than devices and IoT gateways in the proximity layer. A range of devices is available, each suited to a particular IoT use case. IBM solutions are able to connect to these devices to build the overall system.

The IBM Watson IoT™ platform delivers a number of the capabilities in the provider cloud, including the device registry, device identity service, and device management. IBM Watson IoT in combination with IBM Bluemix® provides for API management, IoT transformation and connectivity, and necessary services, plus the means of providing application logic. Watson IoT goes further and supplies services for analytics, visualization, and process management. Also available are transformation and connectivity components that connect from the provider’s cloud system to existing enterprise network assets such as enterprise data stores and existing enterprise applications.

Myriad analytics capabilities are available, including Apache Spark, SPSS® predictive analytics, Watson (cognitive) APIs, IBM Watson IoT™ Platform Analytics Real-Time Insights, BigInsights® for Apache Hadoop service, Geospatial Analytics service, and the Streaming Analytics service. The transformation and connectivity components connect from the provider cloud systems to existing enterprise network assets, such as enterprise data stores and existing enterprise applications.

IBM prebuilt SaaS applications can address specific operational or business capabilities used with many types of IoT systems or sensors, including:

- Facilities and real estate optimization (TRIRIGA®)
- Enterprise asset management (Maximo®)
- Predictive maintenance and quality (IBM PMQ)
- Continuous engineering (Rational® suite)
A series of security services are available, including IBM Single Sign-On, IBM Security Access Manager, IBM Security Directory Server (ID and access management), QRadar® (monitoring), IBM Security AppScan (testing), and IBM Secure Key Lifecycle Manager (key management). For the development lifecycle: IBM DevOps Services, IBM continuous engineering (CE), IBM UrbanCode®, and more. IBM also has offerings in terms of peer cloud services that can be used by IoT systems – notably the Weather Channel service, which can provide streams of relevant weather information, often vital when dealing with physical entities.

Computation and storage for IoT can be done in many places – device, gateway, cloud or data center-hosted environments. Typical communications flow is often via device/gateway, to/from cloud and can also occur between peer systems (devices) and/or gateways.

**Components of a Cloud RA for IoT**

Figure 2 shows the capabilities and relationships for supporting IoT using cloud computing.

![Figure 2. Cloud components for IoT](image)

The cloud components of an IoT architecture are positioned within a three-tier architecture pattern comprising edge, platform, and enterprise tiers, as described in the Industrial Internet Consortium Reference Architecture [4].

The edge tier includes proximity and public networks where data is collected from devices and transmitted to devices. Data flows through the IoT gateway or optionally directly to or
from the device then through edge services into the cloud provider via IoT transformation and connectivity.

The platform tier is the provider cloud, which receives, processes, and analyzes data flows both in flight and at rest from the edge tier and provides API management and visualization. It also provides the capability to initiate control commands from the enterprise network to the public network.

The enterprise tier is represented by the enterprise network and includes enterprise data, an enterprise user directory, and enterprise applications. The data flow to and from the enterprise network takes place via a transformation and connectivity component. The data collected from structured and non-structured data sources, including real-time data from stream computing, can be stored in the enterprise data.

IoT systems need application logic and control logic in a hierarchy of locations, depending on the timescales and datasets that are needed to inform decisions. Some code may execute directly in the devices at the very edge of the network or, alternatively, in the IoT gateways close to the devices. Other code executes centrally in the provider cloud services or in the enterprise network.

When code executes in the IoT gateways or the devices, it’s sometimes referred to as “edge computing.” It’s also sometimes referred to as “fog computing” to contrast with centralized “cloud computing.” Sometimes fog computing can contain one or more layers below the cloud that each could potentially provide capabilities for a variety of services like analytics. This design allows for flexibility in how connectivity and services are designed for optimization and resiliency.

IoT governance and security subsystems span all elements of the architecture to ensure controls and policies for all data and applications are defined and enabled across the system. Compliance is tracked to ensure controls are delivering the expected results.

The remainder of this section describes the various components in detail.

**User Layer**
There are two types of users in this layer—the IoT user and the end-user application.

- **IoT User:** The IoT user is a person or an automated system that makes use of end-user applications to achieve a goal. The IoT user is one of the main beneficiaries of the IoT solution.
- **End-user Application:** A domain-specific or device-specific application that an IoT user may use on smartphones, tablets, PCs or on specialized IoT devices, including control panels.

**Proximity Network**
The Proximity Network is made up of the physical entity, device, and IoT Gateway.

**Physical Entity**
The physical entity is the real-world object that is subject to sensor measurements and actuator behavior. It is the “thing” in the Internet of Things. This architecture distinguishes
between the physical entities and the IT devices that sense them or act on them. For example, the thing can be the ocean, and the device observing it is a water temperature thermometer. Another example is a depot shipping parcels: The parcels are the physical entities, and there are devices with sensors that observe and identify each parcel (for example, via RFID tags or via barcode readers). The RFID tag reader is one thing and the parcels are something completely different – the identity of the parcel is the physical entity here.

**Device**

Contains sensor(s) or actuator(s) plus a network connection that enables interaction with the wider IoT system. There are cases where the device is also the physical entity being monitored by the sensors, such as an accelerometer inside a smartphone.

Key capabilities of a device include:

- **Sensor/Actuator** – The sensor and actuator senses and acts on physical entities. A sensor is a component that senses or measures certain characteristics of the real world and converts them into a digital representation. An actuator is a component that accepts a digital command to act on a physical entity in some way.
- **Agent** – Provides remote management capabilities for the device, supporting a device management protocol that can be used by the device management service or IoT management system.
- **Firmware** – Software that provides control, monitoring, and data manipulation of engineered products and systems. The firmware contained in devices such as consumer electronics provides the low-level control program for the devices.
- **Network connection** – Provides the connection from the device to the IoT system. This is often a local network that connects the device with an IoT gateway – low power and low range in many cases to reduce the power demands on the device. However, there are cases where the network connection is direct to the public network and no IoT gateway is required. In IoT systems, a wide range of alternative communication mechanisms are used and include local area networking using low-power, low-range methods, such as Bluetooth, Bluetooth Low Energy (BTLE), and others to reduce the power demands on the device. It may also include local area networking using WiFi, or wide area networking using 2G, 3G, and 4G/LTE.
- **User interface** – Allows users to interact with applications, agents, sensors, and actuators. This component is optional since some devices have no user interface and all interactions take place from remote applications over the network.

**IoT Gateway**

The gateway is a means for connecting one or more devices to the public network (typically the internet). Because the gateway is essentially a decoupling element, other capabilities are also available. Often, devices have limited network connectivity due to a number of reasons, including the limitation of power on the device, which can restrict the device to using a low-power local network. The local network enables devices to communicate with a local IoT gateway, which is then able to communicate with the public network. The IoT gateway often has other capabilities, including the ability to filter and intelligently react to data, the ability to send and receive
data or commands to and from the internet, and the ability to run application or service logic locally (processing data and executing control logic without the need to communicate to a central location). It can also provide operational efficiency by allowing multiple devices to share a common connection.

Key capabilities in this domain include:

- **Application logic** - Provides domain-specific or IoT solution-specific logic that runs on the IoT gateway. For IoT systems with actuators that act on physical entities, a significant capability of the application logic is the provision of control logic, which makes decisions on how the actuators should operate, given input from sensors and data of other kinds, either held locally or held centrally.

- **Analytics** - Provides analytics capability locally rather than in the provider cloud.

- **Agent** - Allows management of the IoT gateway itself and can also enable management of the attached devices by providing a connection to the provider cloud layer’s device management service via the device management protocol.

- **Device data store** - Stores data locally. Devices may generate a large amount of data in real time, so it may need to be stored locally rather than being transmitted to a central location. Data in the device data store can be used by the application logic and analytics capability in the IoT gateway.

**IBM Capabilities for IoT Gateway**

IBM does not build gateway hardware, and partners with gateway manufacturers, such as Cisco Systems Inc., to provide direct device connectivity. IBM Edge Analytics Agent runs on those gateways to provide a variety of capabilities, including connectivity to the Watson IoT cloud platform and the ability to run analytics on the gateways themselves to filter and summarize data, take local actions, and forward events and a subset of the data to the cloud. You can globally configure the agent, update it from a cloud environment, and cache its configuration and analytics at the gateway so that it can continue to provide its functions even when disconnected from the cloud. This is especially important in environments with intermittent connectivity.

**Public Network**

**Peer Cloud**

The peer cloud is a third-party cloud system that provides services to bring data and capabilities to the IoT platform. Peer clouds for IoT may contribute to the data in the IoT system and may also provide some of the capabilities defined in this IoT architecture.

It is likely that larger IoT systems, such as those involved in smart cities, actually involve the combination of a series of smaller IoT systems, each addressing part of the solution. These systems of systems involve connections between multiple peer cloud systems, each of which may have IoT devices and associated applications and services. Connecting these individual systems can enable larger, more comprehensive solutions.

**Edge Services**

Services needed to allow data to flow safely from the Internet into the provider cloud and into the enterprise. Edge services also support end-user applications.
Key capabilities in this domain include:

- **Domain Name System Server** - Resolves the URL for a particular web resource to the IP address of the system or service that can deliver that resource.

- **Content Delivery Networks (CDN)** - Support end-user applications by providing geographically distributed systems of servers deployed to minimize the response time for serving resources to geographically distributed users, ensuring that content is highly available and provided to users with minimum latency. Which servers are engaged will depend on server proximity to the user and where the content is stored or cached.

- **Firewall** - Controls communication access to or from a system, permitting only traffic meeting a set of policies to proceed and blocking any traffic that does not meet the policies. Firewalls can be implemented as separate dedicated hardware, or as a component in other networking hardware such as a load-balancer or router or as integral software to an operating system.

- **Load balancers** - Provides distribution of network or application traffic across many resources (such as computers, processors, storage, or network links) to maximize throughput, minimize response time, increase capacity, and increase reliability of applications. Load balancers can balance loads locally and globally. Load balancers should be highly available without a single point of failure. Load balancers are sometimes integrated as part of the provider cloud analytical system components like stream processing, data integration, and repositories.

*IBM Capabilities for Edge Services*

These capabilities are well documented in [IBM Advantage supporting the Web Application Hosting Reference Architecture](https://www.ibm.com/support/entry/portal/commerce/ibm advantage). IBM Bluemix supports services for DNS, firewalls, load balancing, and CDN. IBM Security Network Protection is a next-generation intrusion prevention system that can be used to monitor network traffic and provide protection from hidden security vulnerabilities. Finally, IBM DataPower® provides load balancing and SSL termination. It helps quickly secure, integrate, control, and optimize access to a range of workloads through a single, extensible, DMZ-ready gateway.

The IBM VPN service provides a secure IP-layer connectivity between your on-premises data center and your Bluemix cloud. It leverages the Internet Protocol Security (IPsec) suite for protecting IP communication between endpoints residing on your private subnets. An IPsec-compatible VPN gateway is required in your on-premises data center for establishing secure connectivity with IBM VPN service.
Provider Cloud

The Provider Cloud provides core IoT applications and associated services, including storage of device data, analytics, process management for the IoT system, data visualizations, and hosting components for device management, including a device registry.

Key capabilities in this domain include:

- IoT transformation and connectivity
- Application logic
- Visualization
- Analytics
- Process management
- Device data store
- API management
- Device management
- Device registry
- Device identity service
- Transformation and connectivity

A cloud-computing environment provides scalability and elasticity to cope with varying data volume, velocity, and related processing requirements. Experimentation and iteration using different cloud service configurations is a good way to evolve the IoT system, without upfront capital investment.

IoT Transformation and Connectivity

This capability enables secure connectivity to and from IoT devices. This component must be able to handle and perhaps transform high volumes of messages and quickly route them to the right components in the IoT solution.

Key capabilities in this domain include:

- **Secure connectivity** - Provides secured connectivity, which authenticates and authorizes access to the provider cloud.
- **Scalable messaging** - Enables messaging from and to IoT devices. Scalability of the messaging component is essential to support high data volume applications and applications with highly variable data rates.
- **Scalable transformation** - provides transformation of device IoT data before it gets to provider cloud layer, to provide a form more suitable for processing and analysis. This may include decoding messages that are encrypted, translating a compressed formatted message, and normalizing messages from varying devices.

IBM Capabilities for IoT Transformation and Connectivity

The IBM Watson IoT Platform can be used to provide IoT transformation and connectivity. This managed service provides secure connectivity for devices, allowing them to connect either directly or through a gateway. Data from the devices can be retrieved and analyzed in real time, and application logic can also use the platform to query the current state of a device.
The IBM Watson IoT platform lets you perform the following tasks:

- Create and manage applications
- Create, connect, and manage devices
- Extend device management with custom actions
- Create and manage gateways
- Retrieve device data

Devices can connect to the IBM Watson IoT platform using either HTTP or the MQTT messaging protocol.

Provided toolkits help you develop device firmware or application software that uses the platform. Watson supports the following language environments:

- Python
- Node.js
- Java™
- C#
- Embedded C
- mBed C++

**Application Logic**

The core application components typically coordinate the handling of IoT device data, the execution of other services and supporting end-user applications. An event-based programming model with trigger, action, and rules is often a good way to write IoT application logic. Application logic can include workflow and may also include control logic, which determines how to use actuators to affect physical entities.

**IBM capabilities for application logic**

Application logic can be written in many languages. In particular, IBM Bluemix provides runtimes for Cloud Foundry applications written in Node.js, Java (WebSphere® Liberty Profile), Swift, Python, and Go. Node-RED is a tool to develop Node.js applications, and OpenWhisk is a runtime environment designed explicitly for event-driven application development. Both are particularly well suited to IoT.

**Visualization**

Visualization enables users to explore and interact with data from the data repositories, actionable insight applications, or enterprise applications. Visualization capabilities include end-user UI, admin UI, and dashboard as sub-components.

Key capabilities in this domain include:

- **End-user UI** – Allows users to interact with enterprise applications, analytics results, and the like. This also includes internal or customer-facing mobile user interfaces.
- **Admin UI** - Enables administrators to access metrics, operational data, and various logs.
- **Dashboard** - Allow users to view various reports. Admin UI and dashboard are internal-facing user interfaces.
IBM capabilities for visualization

The IBM Watson IoT platform provides visualization dashboards. Other products or services, including embedded reporting, SPSS, Cognos, and the like also provide visualization.

Analytics

Analytics is the discovery and communication of meaningful patterns of information found in IoT data, to describe, predict, and improve business performance.

Key capabilities in this domain include:

- **Analytics Data Repository** - Supports legacy, new, and streaming sources, enterprise applications, enterprise data, cleansed data and reference data, as well as output from streaming analytics. Capabilities include exploration and archiving (for storing, exploring and augmenting large data sets using a wide variety of tools); deep analytics and modeling (application of statistical models to yield information from large data sets comprised of unstructured and weakly-structured elements); interactive analysis and reporting (tools to answer business and operations questions over Internet scale datasets); data cataloging (results from discovery and IT data curation create a consolidated view of information reflected in a catalog). See *How IBM Leads in Building Big Data and Analytics Solutions in the Cloud* [5] for more information on big data and analytics reference architectures for using cloud computing.

- **Cognitive** - These capabilities create an intelligent system that learns at scale, reasons with purpose, analyses to predict, prescribe, and discover from massive datasets of interconnected physical, social, enterprise and other entities, and closes the loop with machine-generated advice, assistance, and actions, in a manner that self-learns and adapts, for enabling augmented human intelligence through man/machine collaborations.

- **Actionable Insight** - Insights that ultimately drive actions that may be used by business applications from data collected, processed, and stored in the data repositories. Capabilities include analytics-based and operational decision management, discovery and exploration across a variety of sources to provide business users with new visibility into business performance, predictive analytics (extracts information from existing datasets to determine the current state, identify patterns, and predict future trends), analysis and reporting (reports of operational and warehouse data to business stakeholders and regulators where big data typically increases the scope and depth of available data), content analytics (enables businesses to gain insight and understanding from their structured and unstructured content), planning and forecasting (enables faster and more efficient development of plans, budgets, and forecasts by creating, comparing and evaluating business scenarios).

- **Streaming Computing** - Accepts and processes in real time large volumes of highly dynamic, time-sensitive continuous data streams from a variety of inputs such as sensor-based monitoring devices, messaging systems, and financial market feeds. Capabilities include real-time analytical processing, which applies analytic processing and decision-making to in-motion and transient data with minimal
Latency, and data augmentation, which filters and diverts in-motion data to data warehouses for deeper background analysis).

**IBM capabilities for analytics**
Tools available to develop and run analytics applications include Spark as a Service, Spark Streaming, SPSS Predictive Analytics, Watson APIs, IBM Watson IoT Real-Time Insights, BigInsights for Apache Hadoop service, Geospatial Analytics service, and Streaming Analytics service. For detailed information about how IBM supports other analytics services, see the IBM Advantage for CSCC Cloud Customer RA for Analytics.

**Process Management**
Process management involves planning, developing, deploying, and monitoring the performance of a business process.

**IBM capabilities for process management**
The Maximo Asset Management service supports business processes to manage all types of assets, including plant, production, infrastructure, facilities, transportation, and communications. IBM TRIRIGA provides strategic facilities planning, implementation, and management capabilities.

**Device Data Store**
The device data store stores data from the IoT devices so the data can be integrated with processes and applications that are part of the IoT system. Devices may generate a large amount of data in real time, requiring the device data store to be elastic and scalable.

**IBM capabilities for edge services**
IBM Object Store provides cost-effective storage for large volumes of data produced by IoT applications. If more rapid access to the data is required, solutions can choose between relational databases such as dashDB®, or noSQL data stores such as Cloudant® or MongoDB. For intensive analytics, the BigInsights for Apache Hadoop service includes an embedded HBase database.

**API Management**
API Management capabilities publish catalogs and updates APIs in a variety of deployment environments. Enables developers and end users to rapidly assemble solutions through discovery and reuse of existing data, analytics, and services.

**IBM capabilities for API management**
IBM API Connect provides streamlined control across the API lifecycle and also enables businesses to gain deep insights around API consumption from its built-in analytics.

**Device Management**
Device Management capabilities provide an efficient way to manage devices securely and reliably from the cloud platform. Device management contains device provisioning, remote administration, software updating, remote control of devices, and monitoring devices. Device management may communicate with management agents on devices using management protocols, as well as communicate with management systems for the IoT solutions.
**IBM capabilities for device management**
The IBM Watson IoT platform supports device management and allows for the creation of customized command sets to meet the needs of the specific application.

**Device Registry**
The Device Registry stores information about devices that the IoT system may read, communicate with, control, provision, or manage. Devices may need to be registered before they can connect to and or be managed by the IoT system. IoT deployments may have a large number of devices, so scalability of the registry is important.

**IBM capabilities for device registry**
The IBM Watson IoT platform can be used as the device registry.

**Device Identity Service**
The Device Identity Service ensures that devices are securely identified before being granted access to the IoT systems and applications.

**IBM capabilities for identity services**
In IoT systems, device identification helps address threats from fake servers or fake devices. The IBM Watson IoT platform can be used as the device identity service.

**Transformation and Connectivity**
Transformation and Connectivity services enable secure connections to enterprise systems and the ability to filter, aggregate, or modify data or its format as it moves between cloud and IoT systems components and enterprise systems (typically systems of record). Within the IoT reference architecture, the transformation and connectivity component sits between the cloud provider and enterprise network. However, in a hybrid cloud model these lines might become blurred.

Key capabilities in this domain include:

- **Enterprise Secure Connectivity** - Integrates with enterprise data security systems to authenticate and authorize access to enterprise systems
- **Transformation** - Transforms data going to and from enterprise systems
- **Enterprise data connectivity** - Enables provider cloud components to connect securely to enterprise data. Examples include VPN and gateway tunnels

**IBM Capabilities for Transformation and Connectivity**
The IBM Bluemix Secure Gateway service brings hybrid integration capabilities to your Bluemix environment. The gateway provides secure connectivity from Bluemix to other applications and data sources running on-premises or in other clouds. A remote client is provided to enable secure connectivity.
Enterprise Network
The Enterprise Network hosts a number of business-specific enterprise applications that deliver critical business solutions along with supporting elements like enterprise data. Typically, enterprise applications have sources of data that are extracted and integrated with services provided by the cloud provider. Analysis is performed in the cloud-computing environment, with output consumed by the enterprise applications.

Systems of record data have generally matured over time and are highly trusted. They remain a primary element in reporting and predictive analytics solutions. Systems of record data include transactional data about or from business interactions that adhere to a sequence of related processes (financial or logistical). This data can come from reference data, master data repositories, and application data used by or produced by enterprise applications functionally or operationally. Typically, the data has been improved or augmented to add value and drive insight. Enterprise data may be input into the analysis process through data integration or directly to the data repositories as appropriate.

Enterprise User Directory
Stores user information to support authentication, authorization, or profile data. The security services and edge services use this to control access to the enterprise network, enterprise services, or enterprise specific cloud provider services.

IBM Capabilities for Enterprise User Directory
IBM Directory Server fills this important function.

Enterprise Data
Includes metadata about the data, as well as systems of record for enterprise applications. Enterprise data may flow directly to data integration or the data repositories providing a feedback loop in the analytical system for IoT. IoT systems may store raw, analyzed, or processed data in appropriate enterprise data elements.

Key capabilities in this domain include:

- **Reference data** - Provide context about collected data.
- **Master data repositories** - Can be updated with the output of analytics, to assist with subsequent data transformation, enrichment and correlation. They can support analytics and feed other analytics models when those models execute.
- **Transactional data** - Data about or from business interactions that adhere to a sequence or related processes (financial or logistical). This data can come from reference data, master data repositories, and distributed data storage.
- **Application data** - Data used by or produced by enterprise applications functionally or operationally. Typically, the data has been improved or augmented to add value and drive insight.
- **Log data** - Data aggregated from log files for enterprise applications, systems, infrastructure, security, governance, etc.
- **Enterprise content data** - Data to support any enterprise application.
• **Historical data** - Data from past analytics and enterprise applications and systems.

**IBM capabilities for enterprise data**
IBM products well suited to support the volume of enterprise data generated by IoT include IBM InfoSphere® Master Data Management (MDM), IBM DB2®, HBase, BigInsights, FileNet®, and Big SQL.

**Enterprise Applications**
Enterprise applications consume cloud provider data and analytics to produce results that address business goals and objectives. Enterprise applications can be updated from enterprise data or from IoT applications, or they can provide input and content for enterprise data and IoT applications.

Key capabilities in this domain include:

- **Customer experience** – Customer-facing systems are a primary system of engagement that drives new business and helps service existing clients at lower cost.
- **New business models** – Alternative business models that focus on low cost, fast response, and great interactions are all examples of opportunities driven by cloud solutions.
- **Financial performance** – Financial applications can be made more efficient as data is consolidated and reported faster and more easily.
- **Risk analytics** – Use risk analytics to evaluate threats to the business, such as fraud or hacking. Elastic resource management means more processing power is available in times of heightened threat.
- **IT economics** – Used to streamline IT operations as capital expenditures are reduced while performance and features are improved by cloud deployments.
- **Operations and fraud** – Cloud solutions can provide faster access to more data, allowing for more accurate analytics that flag suspicious activity and offer remediation in a timely manner.

**IBM capabilities for enterprise applications**
IBM offers a range of specific applications suited to enterprise requirements, such as IBM Maximo for asset management, the IBM Fraud and Abuse Management System, IBM Watson Analytics, and IBM risk management solutions.

**Security**
Security in IoT deployments must address IT security as well as operations technology (OT) security elements. The level of attention to security and the topic areas addressed vary depending upon the application environment, business pattern, and risk assessment. A risk assessment takes into account multiple threats and attacks along with an estimate of the potential costs associated with such attacks.

In addition to security considerations, connecting IT systems with physical systems requires you to consider how the IoT system might impact safety. IoT systems must be designed, deployed, and managed in a way where the operators can always bring the system to a safe operating state, even when disconnected from communications with other systems that are part of the deployment. Indeed, disconnecting from communications may be part of the security measures put in place to help secure the IoT deployment.
There are several areas of security to consider:

- Identity and access management
- Data protection
- Security monitoring, analysis, and response
- System, application, and solution lifecycle management

**Identity and Access Management**

As with any computing system, there must be strong identification of all participating entities – users, systems, applications, and, in the case of IoT, devices and IoT gateways – through which those devices communicate with the rest of the system. Device identity and management involves multiple entities, starting with chip and device manufacturers, including IoT platform providers, and also including enterprise users and operators of the devices. In IoT solutions, many of these entities will communicate and address the IoT devices throughout their operational lifetime.

**IBM capabilities for identity and access management**

The IBM Watson IoT Platform provides capabilities for registering IoT devices and gateways, allowing for identification, authentication, and access control of what devices and gateways can perform in a connected environment. In addition, the Watson IoT Platform has functions for identifying applications that may communicate with and use devices and gateways and invoke Watson IoT platform APIs to perform other IoT-related tasks. User authentication is handled through IBM Bluemix and integration with IBM Single Sign-On capabilities. This allows for a wide range of user/human authentication mechanisms, as well as a wide range of user registries ranging from popular public registries on the Internet to client-specific enterprise or customer-centric registries.

**Data Protection**

Data in the device, in flight throughout the public network, provider cloud, and enterprise network, as well as at rest in a variety of locations and formats must be protected from inappropriate access and use. You can use multiple methods, and indeed, in many cases, you can apply multiple methods simultaneously to provide different levels of protection of data against different types of threats or isolation from different entities supporting the system. Protecting communications links may be used in addition to individual data field level encryption or signing done on the device to provide both end-to-end and point-to-point communications protection. Data at rest in different formats may be encrypted at the field, database, and even whole disk/media level to protect against leakage and improper usage. Increased data collection also results in a need to consider potential privacy implications, requiring additional attention to data segregation, redaction, and special handling requirements.

It is important to consider whether the data involved in an IoT system would include not only personally identifiable information (PII) – which implies legal and regulatory obligations – but also data related to individuals in some way. In some cases, devices may be directly associated with individuals, or individuals may be the physical entities that are the subject of sensor data which, while not personally identifiable information (PII), is definitely information that most would expect to be considered personal. Further, with enough of this observed information, the aggregate data could be enough to identify the person it was gathered from. While PII is usually the subject of laws and regulations, these other types of PII should be treated carefully, and the IoT system must be designed to give appropriate protection to these types of data. Protections may involve where and how data can be
stored, the identified owner of the information, and what data usage restrictions need to be enforced.

Data protection considerations can have a range of implications. For example, it may be the case that data collected by the device must be stored in the same vicinity of the collection, either on the device or on an IoT gateway that is close to the device and cannot be transmitted to a central location such as the provider cloud.

**IBM capabilities for data protection**

Data protection is provided in IBM Bluemix services, such as Cloudant, dashDB, MongoDB, and MessageHub, as well as Spark-based processing systems. Appropriate attention to configuration and connection settings is important when constructing the solution.

IBM IoT solutions including IoT for Electronics, IoT for Automotive, and IoT for Insurance employ these services with careful attention to configuration settings so that appropriate data protection is used.

**Security monitoring, analysis, and response**

To detect and react to active attacks or anomalous behaviour, every system must have built-in monitoring of the environment. Because of the scale of IoT systems, both in the number of devices as well as the amount of information being processed, automated responses to known attacks and automatic detection of suspicious behaviour are required. These responses may include temporary isolation, quarantine, or the removal of parts of the IoT system, as well as formal incident response processes for addressing vulnerabilities that are discovered after the systems have been put into service.

Like IT security, there is a need for disclosure of vulnerabilities so that affected parties can appropriately mitigate the risk and make changes and updates in a timely manner. Because attacks can come in a variety of different forms, all attacks must be expected, planned for, and responded to. As just one example, an attack might come in the form of injection of fake, erroneous, or erratic sensor data into the IoT system in an attempt to steer automated decision-making parts of the system to act in a desired (by the attacker) manner. Such attacks must also be expected, planned for, and responded to.

**IBM capabilities for security monitoring, analysis, and response**

IBM offers several mechanisms for monitoring and analyzing data communications traffic from and between computing systems.

IBM products such as QRadar can be used in conjunction with the Watson IoT Platform and services. These offerings, combined, allow for monitoring, analysis, and response to situations that can arise as IoT devices connect with IT systems to construct a solution. The Watson IoT Platform provides devices management, security configuration, and risk-management features for device and gateway-specific monitoring and response.

**System, application, and solution lifecycle management**

Lifecycle management of the IoT system is complex, multi-faceted, and has relationships with identity management, device management, the supply chain, application and software development, through to system operations and change management of deployed and in-service systems.
Attention to security in all of these areas is required to prevent a variety of attacks ranging from malicious code insertion to inappropriate firmware/software deployment, to effective cryptographic key management. Code, key material, and even physical components must be verified as they flow from procurement and creation through to their installation into the devices, IoT gateways, and other systems that make up the IoT system. The IoT system should also provide the capability to update individual components in a secure way, both to address vulnerabilities and also to address functional enhancements over the lifetime of the system.

**IBM Capabilities for System, Application, and Solution Lifecycle Management**

When constructing an IoT solution and designing, building, and deploying IoT devices and gateways as a part of that solution, you must pay careful attention to how software or firmware will be managed on the devices and gateways. Device and gateway manufacturers and users should determine the most appropriate firmware update mechanism to employ, including what firmware-over-the-air (FOTA) vendors to work with. IBM offers interfaces in the Watson IoT Platform to signal to an applications or solutions when a firmware update is necessary and to observe the firmware levels reported by devices and gateways.

Validating that the firmware and software running in devices and gateways is as expected is also important. If code running in these systems has been tampered with or is corrupted in some way while in operation, inappropriate behavior from the device or gateway may occur. Technologies for in-memory scanning for unexpected code modifications, such as those provided by Arxan (www.arxan.com), can help detect and respond to such attacks.

All interfaces in an IoT solution must be tested for potential vulnerabilities. Ongoing penetration testing of devices, gateways, and all other externalized parts of the solution is necessary to detect potential weak spots and take steps to mitigating these. IBM AppScan Source (for static, source-code analysis) and AppScan Enterprise (for dynamic, web/HTTP interface-based testing) can be applied to any software or interfaces, respectively, which are exposed as part of the solution. This can include interfaces exposed by devices and gateways, even if those interfaces are meant for local, isolated administrative connections.

Managing keys in devices and gateways can be challenging. IBM Secure Key Lifecycle Management (SKLM) offers mechanisms for performing key management operations. Device and gateway manufacturers can employ these services to assist with key management operations for keys deployed into devices and gateways.

**IoT Governance**

As described in the IoT Security section, there are many challenges in securing an IoT solution. Oversight and procedures must be used to ensure that when new vulnerabilities and threats are discovered, there is a means and mechanism for addressing these threats in IoT systems.

An important difference in IoT systems differ from traditional IT systems because exploits and failures in IoT systems have the potential to cause serious harm to humans, property, and the environment. Physical devices and equipment are usually in service for much longer periods of time than typical computing systems such as servers, PCs, tablets, and other mobile devices. IoT equipment is often installed in locations where change or replacement is complicated due to great cost or inconvenience. Because of these reason, IoT systems must be designed and deployed with change/update/modification in mind, along with strong governance to ensure that such change is done appropriately, safely, reliably, and
securely. Indeed, IoT system change is likely to be needed long after device warranty periods have expired as it is well known that physical systems are often used for long periods of time.

Strong governance procedures are needed to determine and enforce the appropriate in-service lifespan for devices and to plan non-disruptive, secure change-overs as new systems are introduced into the system. IoT governance complexities are similar to the complexities in hybrid cloud computing. Definition, planning, and oversight might include both technical and operational staff. Visibility into SLAs, change management, and other policy and process areas can be expedited by selecting tools to simplify data collection, reporting, and notifications.

**IBM capabilities for IoT governance**

Managing the software lifecycle of firmware, software, applications, analytics processing, and user interface functions of an IoT solution is a complex task. Often, each piece of software that makes up the solution is built using a specific software development methodology—one IoT solution could be built using a wide range of different software development methodologies that will be followed across this spectrum of software which makes up the solution. IBM DevOps Services, IBM Continuous Engineering (CE), and IBM Continuous Lifecycle Management (CLM) offerings provide a flexible set of tools and function for managing software development and deployment lifecycle. IBM UrbanCode offerings can also assist in managing software deployments across development, test, and production environments, further assisting organizations in maintaining strong IoT governance.

The Provider Cloud components may also be subject to change over time. For example, the analytics components and their associated software may undergo regular enhancements to improve their performance and reliability. Appropriate governance must be in place to ensure that changes to these components are understood ahead of time and that the changes do not have an adverse impact on the overall IoT system.

**The Complete Picture**

Figure 3 provides a detailed view of all of the components, subcomponents, and relationships in a cloud-based IoT solution architecture.
IBM Product Support for IoT Solutions using Cloud Solutions

Now that we've reviewed the component model for an IoT solution using cloud computing, let's look at how IBM products can be used to implement an IoT solution. In previous sections, we highlighted IBM's end-to-end solution for deploying an IoT solution using cloud services.

The figure below shows how IBM capability maps to specific components in the reference architecture.
Scenarios
Now that you understand the architectural components of an IoT solution in the cloud, let’s look at how to use IBM products to implement common scenarios using this architecture.

- Scenario 1. Smart Homes Insurance scenario
- Scenario 2. Connected Care Analytics
- Scenario 3. Smart Home Connected Appliance Scenario
- Scenario 4. Real-time Motor Monitoring
- Scenario 5. Industrie 4.0/Industrial IoT

These scenarios reuse the components that the organization is currently using in traditional data centers, which we depict as part of the enterprise zone of the architecture.

Scenario 1. Smart Homes Insurance Scenario
Figure 5 illustrates the flow of a connected insurance service use case for IoT.
In this example, smart homes with connected devices and sensors provide insurance companies the ability to improve the service to policyholders while gaining insight into risks in the home. Connected devices allow policyholders to receive notification of potential danger to the home and engage with the insurer in a more proactive manner.

By connecting homes, insurers, and other services, the connected insurance service uses key components of the IoT reference architecture. As an example, leak detection sensors and valves can enable the policyholder to monitor water leaks and offers protection from resulting damage. The sensors are purchased from multiple sources and installed in the home, which includes connecting them to the device maker’s cloud services. The policyholder authorizes the insurance cloud service to connect to the device maker’s cloud service granting access to the device data. The device maker is responsible for the life cycle of the devices and the insurance company benefits from access to the data from these devices and provides an improved experience to its policyholders.

Basic information flow:

1. Sensors and actuators are deployed in the home and attached to the device maker’s cloud service. As an example, the sensors can detect water leak detection, water flow, temperature, and the actuators can include automatic water shutoff valves.
2. The homeowner logs into the insurance mobile application and authorizes the insurance service to access the device maker’s (peer) cloud and their device data. The mobile application sends the authorization token and insurance company identifier to the cloud service. This information is used to map the user, devices, and insurance policy within the cloud service. The device cloud service is used because the device makers have already deployed into their own cloud and owns the life...
cycle of the device as well as the user experience with the devices.

3. The insurance service receives authorization, device details, and the insurance ID from the insurance mobile application and processes this in several nodes (application logic, device registry, and device data store). The devices are registered with the device registry, and data mapping is updated in the application logic component. IBM Bluemix Liberty or Node.js can be used for providing the application logic, which can use the IBM IoT for Insurance Service from Bluemix. IBM Cloudant, dashDB, or Object Storage can be used for device data store. IBM Watson IoT Platform can be used for device registry.

4. The insurance service application uses the authorization token to connect to the device maker (peer) and request the data. The application is configured to pull data on a configured interval. In addition to device data, the application can be configured to access other data sources such as a weather data service for use in analysis. IBM Bluemix IoT for Insurance Service is a Bluemix service that collects, manages, and analyzes data from connected policyholders. IoT for Insurance helps provide personalized risk assessment, real-time protection, and policy cost reductions.

5. Data from devices and other sources such as the weather service are continually updated and sent to analytics systems to determine if a potential risk threshold has been exceeded. This data is analyzed to determine if there is a potential for damage to the home (including water damage, freeze potential, etc.). Device data from sensors in connected homes provides insight into potential problems in the home such as water leak or humidity. The Weather Company data service on Bluemix integrates weather data from The Weather Company into Bluemix applications, and it can retrieve weather data for an area specified by geo-coordinates.

6. Once it is determined that there is a problem, using the analysis from Step 5 notifications are sent to the homeowner and to the insurance company. The homeowner can then take action to respond to the notification and determine if damage has occurred, and the insurance company can initiate a claim process.

7. If damage has occurred, the insurance business process of claims management is initiated. The insurance business processes can be accomplished in the cloud service, their enterprise applications, or their mobile applications. This is dependent on how and where the insurance company decides to perform the business logic. IBM Real-time Insights or MessageHub can be used for managing the process flow. Typically, this is done using the insurance company’s existing claims management system.

A cloud architecture makes this type of solution easier to implement and maintain. As demand increases, more resources must be acquired.

Scenario 2. Connected Care Analytics

Figure 6 illustrates the flow of a connected car analytics use case for IoT.
**Background** – There are two people in this scenario. A 75-year-old male driver has a heart condition and wears a Fitbit to monitor biometrics like heart rate. A female driver, 35, has an active lifestyle and wears an Apple watch which she has enabled to share information. Both drivers register for a “Better Driving Behavior Program.”

Both drivers have a known profile, created as an enterprise record, that is based on their hometown location, driving records, daily driving route, speed, current weather, road conditions, and other features. These relate to a set of KPIs that provide metrics on how to measure such features. Because the drivers have opted into the better driving behavior program, we can monitor the devices the drivers have given access to or permissions for. This information is shared between the device and the providers, the drivers, their emergency contacts, and doctor’s office.

When the man and woman drive and interact with their devices and gadgets, the IoT framework picks up all data points. The analytics engine built for this framework evaluates any changes in driving behavior and flags any anomalies that need to be acted upon. The engine also recognizes information that the systems need to learn about as its normal or new behavior that needs to be acted upon in the future.

**Runtime Flow**

1. The user registers and creates a profile in the Enterprise User Directory, and links existing social media accounts to a doctor’s network. IBM Security Directory Server, a lightweight implementation of Lightweight Directory Access Protocol (LDAP) is used for security and identity management. It acts as a foundation for deploying comprehensive identity management applications and advanced software.
architectures. A custom web or mobile application logic is used to build the user profile. The user's record gets updated in the Enterprise User Directory.

2. The user connects his or her vehicle to a device registry service and to a global network of devices for identification and broadcast message. IBM Watson IoT Platform allows registration of the device and device type, and the payload information is configured accordingly in the platform for downstream and upstream consumption. The user's record gets updated with the devices in the device data store. The IBM Watson IoT Platform boilerplate comes configured with Cloudant as a NoSQL Database as a Service that can be used to store telemetry and other sensor information from devices for long-term storage and retrieval.

3. The user updates his or her user preferences like data capture setup, special alerts, thresholds, emergency contacts, and application settings. A suitable web or mobile application built on the IBM MobileFirst platform or a custom enterprise application that leverages services like Node.js or WebSphere Liberty can be used for these user interactions.

4. The device captures motion, telemetry, and geospatial data by monitoring interactions from a fitness tracker, Apple watch, and cell phone usage. The setup in the steps above allows the IBM Watson IoT to capture all this information and interaction the user has with the devices over time. Additional services like streaming analytics ingest, analyze, monitor, and correlate data as it arrives from real-time data sources. View information and events as they unfold.

5. Via edge services, the user application sends data from the Internet, like social media accounts, or weather and road conditions.

6. The IoT transformation and connectivity service enables secure connectivity to the registered IoT devices (like vehicles, Fitbit, Apple watch). The IBM Watson IoT Platform enables this transformation and connectivity.

7. Devices from the male driver record abnormal medical stress and driving pattern. Devices from the female driver record a phone call and an erratic driving pattern. The application correlates information and evaluates the next best action due to the anomalies and persists in corporate data store. Both drivers are sent appropriate alerts, and the application follows the escalation path as defined in preferences. Custom application code and scenario logic is embedded in Node.js or WebSphere Liberty services that allow devices to send information to downstream data stores and application processing engines for correlation and actions.

8. The analytics engine implements machine learning and applies heuristics, statistics, classifiers, dimensional reduction, and collaborative filtering for anomaly detection and remediation. It updates in-memory processors for quick processing real-time transactions. IBM analytical, predictive, and machine learning capabilities provided with Spark as a Service on Bluemix, IBM Data Science Experience, IBM SPSS, and IBM Watson APIs can be used to understand the behavior of these interactions, their trends, anomalies, outliers and for statistical and predictive learning. Offline learning and online scoring machine learning predictive models can be interjected into the data processing pipelines for applying the analytics where needed. In addition, there are other analytics services available on Bluemix that can be used.

These include:
IBM Watson IoT Context Mapping Service enables your application to analyze moving object trajectories by using road network-based geospatial services. It provides real-time query interfaces to access road network data and search services by unique index structure and advanced cache mechanisms.

IBM Watson IoT Driver Behavior Service lets you analyze drivers’ behavior from vehicle probe data and contextual data.

Geospatial Analytics help track when devices enter, leave, or hang out in defined regions.

9. The transformation and connectivity service allows for secure connection to enterprise systems to look up event information. IBM DataPower and IBM Integration Bus services are used for this.

10. The enterprise application maintains business models like customer experience and risk evaluation and is used for lookup or transaction processing or publishing a new event rule, audit processing. This data is loaded in memory for access to the analytics engine. Enterprise applications are typically custom and specific to the enterprise and are outside the scope of a single IBM product. However, there are many capabilities and solutions provided by IBM commerce, travel, and transportation that allow internal and external users to review business outcomes, experiences, trends, health of programs, sales and revenue information, for example. IBM Cognos is one such enterprise application that can be used in such scenarios.

11. This service manages process workflow and coordinates the REST-based services used in your apps. IBM Bluemix Car Diagnostic API, Real-Time Insights and other microservices running on NPM or Node.js can be utilized to deploy process workflows. The IoT Car Diagnostic API can help you to assess the health status of a vehicle, by translating OBD error codes in a human-readable form.

12. The IoT governance maintains policies and terminology of the business applications and rules around accessing that information. IBM security, audit, and governance capabilities in QRadar Security Intelligence platform and SIEM capabilities, along with Guardium® audit, compliance, and vulnerability capabilities, and capabilities in the information governance catalog in the IBM Information Integration Suite, provide a complete policy-based secure and controlled environment.

13. Visualization provides active, descriptive reports and dashboards to the user. IBM supports open technologies like Rave, D3, Angular, and Brunell along with enterprise offerings from IBM Cognos and Watson Analytics.

14. The end-user application provides the engagement model for the user in the form of a mobile or web application. IBM MobileFirst mobile or custom mobile or web-based applications can be used to surface various usage metrics to end users or to provide an interactive environment. IBM API Connect service can surface these services to many users.

Scenario 3. Smart Home Connected Appliance Scenario
A manufacturer and its ecosystem partners can provide end-user remote control and better customer support for connected appliances for smart homes.

Figure 7 illustrates the flow of a connected appliance and smart homes scenario for IoT.
1. A smartphone app used by the appliance owner registers the customer’s ownership and provides the end user with the ability to control the appliance. The IBM IoT for Electronics service (available in Bluemix) provides a sample mobile application. This makes use of the Mobile Cloud Access service that is also available in Bluemix.

2. Customer registration details are recorded in the manufacturer’s systems of record. The Bluemix Secure Gateway service can provide connection to the system of record.

3. The appliance is registered in the cloud provider’s registry, and appropriate security permissions are established. The IoT for Electronics service provides owner registration services, and the Watson IoT Platform provides a registry of the actual devices.

4. While in the house, the end user can use the smart phone app to check the status of the appliance and can send commands to the appliance, for example to adjust a temperature setting. In this case, the app connects directly to the appliance. The IoT for Electronics sample application shows how to do this.

5. Devices embedded in the appliance send data to the app and respond to its commands. The IBM Watson IoT Platform sends data to the app.

6. The app can communicate with the cloud provider to offer the same capabilities when the user is not physically in the house. In this case, the device also communicates with the cloud provider and communications happens via the IBM Watson IoT Platform.
7. Application logic can be used to influence or control the appliance as well, for example a washer/dryer might not start immediately, but might delay to get a better energy rate. Applications could be written using a Bluemix runtime, for example the Node.js Cloud Foundry runtime.

8. Usage and operational data can be collected from the devices in the appliance and stored in a device data store. IBM Watson IoT Platform can store data directly in the Bluemix Cloudant NoSQL DB service, but other storage services can be used instead.

9. This data can be analyzed, either in real time or retrospectively, for example for:
   - Preventive maintenance
   - Understanding what features are used from appliance (for future marketing or cross selling)
   - For rental/lease of the appliance (pay as you go)

   The IBM Predictive Maintenance solution can be used for preventive maintenance. You can gain usage insights by collecting the device data in a big data store and running analytics applications against it.

10. Third-party ecosystem providers can connect in via API management to offer further services, for example selling accessories or consumables (e.g., soap). API management can be provided using IBM API Connect for Bluemix.

**Scenario 4. Real-time Motor Monitoring**

This sample solution monitors a torque motor in real time on a shop floor, and the application notifies the technician automatically in case of any variance from standard operating parameters. This Watson IoT cloud native application is integrated with Maximo asset management system to minimize operational changes. There are also Bluemix APIs used for notification and SMS alerts. The IBM predictive maintenance cloud service enables operations, manufacturing, production, and maintenance personnel in asset-intensive industries to use predictive analytics to improve asset availability, increase throughput, minimize unplanned outages, and reduce maintenance costs.

This architecture offers the capability to develop predictive models to analyze asset performance data in real time, calculate asset health scores, and predict potential asset failure. Please refer to [IBM Cloud Architecture Center](https://www.ibm.com/cloud/ibm-cloud-architecture-center) for additional details on this Watson IoT solution.

Figure 8 illustrates the flow of the real-time motor monitoring scenario for IoT.
1. A user interacts with a machine (physical entity). In this scenario, the interaction is with a servo motor via sensor to monitor its performance attributes to enable preventive maintenance.

2. IBM IoT Gateway receives the data from the IoT-enabled torque motor and is converted into MQTT format.

3. The MQTT data from the IoT gateway is received by edge services that are enabled by Bluemix API management framework, which the real-time insight application uses.

4. API management enables the bi-directional connectivity into the IoT-enabled device from the Bluemix application.

5. The APIs and the IoT devices are authenticated using IoT foundation APIs in Bluemix. These APIs are enabled by API management and device registry to ensure sensor and API authentication.

6. API and device authorization passes the received data to the PMQ application via a device identity service.

7. The application logic checks for exceptions, boundary conditions, and other anomalies in real-time.

8. Via Transformation and Connectivity services, the workflow integrates with Maximo and notifies service representatives for real-time predictive maintenance.

9. The IBM cloud platform and Watson IoT APIs complete the business process automation and operations integration. This enables new business models, which helps to improve operational efficiency.
Scenario 5. Industrie 4.0/Industrial IoT

Industrie 4.0/Industrial IoT focus primarily on business scenarios integrating vertically (from machines to cloud), horizontally (among supply networks), or along the life cycle of the product. Given the focus on integrating the operational technology (OT) layer with the IT layer in a manufacturing context, Industrial IoT represents a special case of the general IoT reference architecture. This is due to the nature of its closed environment with some specific requirements, three layers (edge, plant, cloud/enterprise), as well as the importance of the flexibility of functional deployment among the three layers, which is a strong differentiation of IBM’s Industrie 4.0 approach.

Note that the three-layer approach results from the need for the individual factory (or "plant") to continue operation even if external connections to enterprise and cloud systems should fail – stopping the production lines for an external connection failure is unacceptable. This three-layer approach also occurs in other IoT scenarios such as smart buildings, where a local entity must continue operating smoothly even if connectivity to centralized IT systems fails.

This scenario from automotive manufacturing monitors production equipment and tools for various performance metrics and performs analytics on this data both at the edge (applying the emerging edge analytics architecture) on a Cisco Edge device and at the enterprise layer, as part of the IBM IoT platform. The equipment in this example, which includes robots (used for welding) and handling equipment (conveyors, palletizers), is already instrumented and is being monitored by either Omron or Fanuc programmable controllers.

Other tools and devices are involved in the operation as well. A welder attachment is connected via the Fanuc controller, and image-processing equipment (being used for inspecting welds) is attached via the Cisco edge device. RFID is used in this example for identifying pallets and WIP; RFID is also integrated via the Cisco Edge device.

Figure 9. Industrie 4.0/Industrial IoT architecture
Background of the reference architecture: Devices and production machines associated with production operations are typically managed by existing DCS/SCADA systems, which can be integrated by industry protocols such as Profinet, OPC, MODBUS, etc. Some newer equipment is embedding technology that allows it to communicate with the outside world through IT protocols such as MQTT.

At the edge, gateways are typically used to integrate with the existing systems and equipment and are also becoming more capable of running edge analytics, applying rules, and even storing data locally to support operations at the edge. It is quite possible that the edge will completely handle an interaction with equipment with no involvement of the plant or enterprise layers. In other cases, the information from the edge will flow up through the plant or to the enterprise where plant and enterprise analytics will be performed in a similar way. The edge and plant need to be able to operate as a stand-alone unit from the enterprise, so some capabilities of the platform need to be in both the plant and the enterprise.

Information from intelligent devices and production machines can be communicated up through the layers (with appropriate filtering and aggregation along the way). It is also possible for information from the devices and machines to be communicated directly to the plant or enterprise layers, assuming the devices and machines have that capability (for example, through embedded technology).

The steps involved in this example are as follows:

1. Information is collected from the equipment and tools initially by programmable controllers connected to the equipment through proprietary equipment interfaces. The controllers in this example have an embedded piece of software called deviceWISE (from Telit) that can be configured to pass controller and device data to the upper layers of the architecture via standard IT protocols like MQTT and MQ (or via JDBC writes to a database) periodically, or based on conditions. The information can also be transformed (mediated) as needed before it is passed on. The same component, deviceWISE, is used for the same purpose within the Cisco Edge (IoT Gateway) device.

2. Analytics are performed on the outbound information in the OT/IT hub (in this example is realized by a Cisco Edge device that is embedded with IBM Edge Analytics Agent (part of the IBM IoT Gateway). Depending on the result of the Edge Analytics, command data is sent back down to the equipment. This is the reverse of the flow into the edge and uses the DS broker and deviceWISE to issue the command and transform it into the specific protocol and data needed by the equipment, in the case of the Image Processor or RFID attached equipment, or the controller, in the case of equipment and tools managed by the Fanuc or Omron.

3. The DS Broker component of the Cisco Edge controller forwards events, based on configuration, to the Plant Service Bus, which, in this example, is the IBM Watson IoT Platform running on Bluemix. In some cases, where plant data is not allowed to leave the premises for example, the Plant Service Bus might instead be realized by the IBM Integration Bus (IIB) Manufacturing Pack with the IBM IoT Platform running in the cloud at the enterprise level.

4. Operational data is collected at the plant level (after normalizing and cleansing) to support plant-level analytics as part of the shop floor analytics loop. An information model, based on the ISA-95 industry standard, is used to support the analytics and is
also used for dashboards and reporting as well.

5. Within the Plant Service Bus, analytics and rules determine the required actions for this event. Required actions can include feedback, but can also include triggering actions represented in a workflow. This could be simple analytics such as threshold monitoring or trending, but it could also be model-based analytics, looking at the performance of a production device, a tool, a work cell, or a production process (depending on where we are in the architecture). In this example, the analytics component of the IBM IoT Platform is used for this purpose. In other situations, IBM products/offers such as Predictive Maintenance, Predictive Quality, Plant Performance Analytics, or SPSS might also be used.

6. If warranted (based on configuration and applied analytics/rules), a plant-level workflow is triggered. This workflow is composed to use Plant IT System of Record (SoR) services in combination with platform services. The services here could correspond to a Manufacturing Execution System or Enterprise Asset Management system. They could also be platform-provided services (e.g., Watson). The workflow is implemented, in this example, as an IBM Integration Bus flow.

7. Based on the result of the analytics and rules, or the workflow (if executed), information may flow back to the Edge and Production Equipment, which results in dynamic reconfiguration of the manufacturing process.

**IoT Development**

For many organizations, building an IoT solution is new territory. Frequently, these systems involve mobile devices, multiple external data streams, and third-party APIs. Whether your business has expertise in these areas or is just starting out, IBM Bluemix offers an efficient way to begin building IoT applications—from minimum viable product to full functionality. The combination of composable services, templates for quick start on IoT and mobile development, including cognitive solutions, powerful data management and data science tools support the range of development activities across the IoT architecture.

**Deployment Considerations**

Deciding which elements of an IoT solution belong on a specific cloud service type—hybrid, public, private/dedicated, or on-premises (local)—is an important decision. Clearly defined requirements related to data sovereignty, regulatory compliance, scalability, availability, and usage peaks are imperative to the decision-making process. The sheer amount of data associated with live data streams from manufacturing sensors or consumer devices means that all aspects of messaging, connectivity, and data management are of the utmost importance.

Once an organization has defined functional and non-functional requirements for their solution, they can model their capacity and performance requirements, analyze existing enterprise systems and infrastructure, and review compliance and risk exposure to come up with their workload assessment. IBM offers workload affinity engagements to assist customers in deciding what cloud service type is best suited to their needs. Established business with strict compliance needs frequently choose a hybrid cloud adoption path. The
following section discusses the most important areas to consider when deploying an IoT solution.

Cloud infrastructure and services offer tremendous flexibility because they don't have to focus as heavily on how components are physically connected. Even though scalability and elasticity are inherent in cloud and reduce the need for exact capacity and resource forecasts, advanced planning is still important. This planning gives organizations a reasonable expectation of operating expenses and sets up the necessary monitoring and automation to deliver the best service at the best cost. IBM cloud service offerings include tools and engagements that help decide where to place specific workloads, such as CloudMatrix brokerage, as well as the means to monitor and manage day-to-day operations and billing.

This section offers guidance for how to provision data and computing resources using the IBM cloud platform and cloud services.

IBM offers a variety of APIs, data transformation, and storage options as cloud services. All offerings provide the necessary scalability and elasticity to meet the data throughput and transactional loads associated with IoT. These include:

- IBM Bluemix
- Cleversafe Object Storage
- DataPower

These offerings also function in a hybrid architecture, allowing the enterprise to leverage existing investments and knowledge.

**Common Criteria for Cloud Environments**

While no single cloud environment optimizes all these criteria, defining the most important ones for your customers will go a long way towards ensuring user satisfaction and meeting your budget. Visibility into services is the key to managing satisfaction and cost. IBM Bluemix provides a single interface to manage platform and infrastructure services and billing.

Specific criteria to consider include:

- Scalability and elasticity
- Data bandwidth
- Data sovereignty
- Resilience
- CPU and computation
- Data volume
- Security
- Optimized provisioning

**Scalability and elasticity**
Elasticity is the ability for a cloud solution to provision and de-provision computing resources on demand as workloads change. Public clouds have a distinct advantage since they generally have larger pools of resources available. You also benefit by only paying for what you use. Private clouds and dedicated hardware can make up some of the difference with higher bandwidth data paths.

IBM Bluemix Infrastructure as a Service allows the creation of a dedicated, private cloud that is based on bare metal and can burst into public cloud as needed. This option allows the architect to design an IoT solution that takes advantage of the best functionality of dedicated and public services. IBM Blue Box is another infrastructure option for a managed OpenStack in the cloud.

Data bandwidth

Public and private clouds need to be optimized for big data. Large cloud data sets requiring fast access benefit from processing components with fast and efficient data access. In many cases, this means moving the processing to the data, or vice versa. Cloud systems can effectively hide the physical location of data and processing. Tuning activities can be carried out continuously with minimal impact on deployed applications. The elasticity of APIs and connectivity services is also key. IBM offers a range of solutions for moving and managing data set, particularly unstructured data.

Data sovereignty

The physical location where data is stored may be regulated, with regulations varying from country to country. This is particularly the case for personally identifiable information and for sensitive data such as health data and financial records. The European Union has particularly stringent regulations that apply to the PII of European citizens. As a result, any IoT cloud system must account for data sovereignty rules and store and process data only in those locations permitted by the regulations. This requires the provider cloud to provide the cloud service customer with control over storage and processing locations. IBM Bluemix PaaS and IaaS have data centers in 40 locations, satisfying EU and other data sovereignty regulations.

CPU and computation

The availability of inexpensive commodity processors means that public, private, and hybrid cloud server farms are typically highly scalable. Modern development environments using Hadoop, Spark, and Jupyter (iPython) take advantage of these massively parallel systems. Streams and high-speed analytics are an emerging area where cloud applications use more powerful processor pools to enable real-time, in-motion data solutions.

Dedicated hardware allows for faster development and testing prior to migration towards hybrid and public environments. IBM offers multiple, fully managed and customer managed options in support of big data and analytics.
Data volume
In IoT systems, the data volume can exceed a threshold at which the traditional analytic toolsets and approaches may no longer scale to meet performance requirements. So careful planning to store data in public cloud or private cloud or traditional data center is very important. Data streaming in cases of weather or maps that use GPS may result in huge data set for analysis. Also, all data loses relevance over time. Data retention requires a little experimentation, unless specifically governed by regulatory or other policies. Public clouds offer the flexibility to store varying amounts of data with no advance provisioning. In-house cloud storage solutions can offer long-term storage cost advantages when volume is predicted in advance.

Security
As more data about people, financial transactions, and operational decisions is collected, refined, and stored, the challenges related to information governance and security increase. The data privacy and identity management of devices and individuals is very important for cloud computing. The cloud generally allows for faster deployment of new compliance and monitoring tools that encourage agile policy and compliance frameworks.

Cloud data hubs can be a good option by acting as focal points for data assembly and distribution. Tools that monitor activity and data access can actually make cloud systems more secure than stand-alone systems. Hybrid systems offer unique application governance features: Software can be centrally maintained in a distributed environment with data stored in-house to meet jurisdictional policies.

Optimized provisioning
Optimized cloud provisioning can help you select the right product family for a given set of usage criteria. IBM Cloud Brokerage can help automate provisioning based on automated assessment based on an organizations strategy and policies.

Hybrid cloud and IoT
Similar to data-intensive solutions in e-commerce, the enterprise moving to IoT environments frequently need to combine public cloud, private cloud, and on-premises components to create a hybrid cloud. See the CSCC Practical Guide to Hybrid Cloud Computing [6] for more information about hybrid cloud planning, governance, and operations. The IBM Cloud point of view is to offer choice with consistency, giving you the ability to:
  - Extend an existing investment via a range of cloud services
  - Position environments in public, dedicated, or local spaces as needed to satisfy regulatory or security requirements
  - Gain elasticity by leveraging off-premises systems that are a mirror to on-premises, all while keeping visibility across the entire architecture.

IBM hybrid offerings include:
  - IBM Bluemix PaaS
Businesses implementing hybrid cloud solutions are looking for flexibility and agility in delivering new capabilities. Efficiency in process and data collection are often the drivers of these initiatives. The broad availability of embedded sensors and cellular, WiFi or network connectivity of devices supports the expansion of IoT. Because of the need to combine multiple data sets to serve a variety of user personas, IoT solutions for B2B and B2C are frequently the entry point for hybrid cloud adoption.

The following example illustrates the new business models and approaches possible when adopting hybrid cloud deployment for IoT systems.

**IoT for connected cars** - The IoT solution for connected cars is a real-time event detection and management system designed to securely detect, analyze, and handle events generated by connected cars. Some of the information with historic and maintenance data for car manufacturer will stay in the dedicated private cloud or in their traditional data centers while other generic information and their integration with third-party cloud services may stay in public cloud. Connected cars need real-time information about weather, traffic, and map data which comes from peer cloud services. For the data privacy and sovereignty requirements, data with personal information about customers may reside in on-premises data centers in specific countries. With use of hybrid cloud, only we can handle all these specific needs.

**Summary of Key Considerations**

The architect of a consumer-centric, business focused or industrial IoT solution must navigate a complex set of concerns. Among these concerns, the architect must consider end-to-end security, management of massive amounts of data, and ensuring that the velocity of data transfer and overall connectivity meets business requirements or contractual obligations. IoT solutions, with their combination of multiple device types integrating with multiple system types, also require the kind of adaptive operation supported by continuous deployment methods, cloud resilience, and elasticity.

Architects will be most successful when they keep these considerations in mind:

- Design to meet needs for rapid change and updates in connected devices and sensors
- Build monitoring and adaptive management into the system
- Design with data security and privacy requirements at the fore
- Ensure high performance across all components, with special attention to where the ingestion of real-time data streams occurs
- Plan system interfaces and services for the greatest flexibility
- Ensure future interoperability by choosing open standards-based components wherever possible
- Make data security a focal point across the architecture
Conclusion
This paper offers a deeper understanding of the CSCC Cloud Customer Reference Architecture for IoT and introduces key concepts for creating an efficient, scalable, secure IoT architecture and gives you guidance on how to integrate your on-premises and enterprise systems. To frame your specific solution compared to real-world experiences, this paper also offers practical guidance in the form of deployment options and use-case scenarios based on actual IBM customer implementations. As you can see, IBM products support the key capabilities required to realize and operationalize an IoT architecture. IBM provides first-class product support for IoT and the cloud architecture for customers.

Acknowledgements
Eric Libow, Gopal Indurkhya, Heather Kreger, Tim Hahn, Peter Niblett, Mike Edwards, Thomas S. (Scott) Wallace, Tejinder Luthra, Ramesh Menon, Karolyn Schalk, Elizabeth Koupman, Glenn Daly, Robert Flaherty, David Noller, and Plamen Kiradjiev

References
[2] The Industrial Internet Consortium’s Industrial Internet Security Framework (IISF) paper
[4] The Industrial Internet Consortium’s Industrial Internet Reference Architecture IIAR paper
[5] How IBM leads in building big data analytics solutions in the cloud