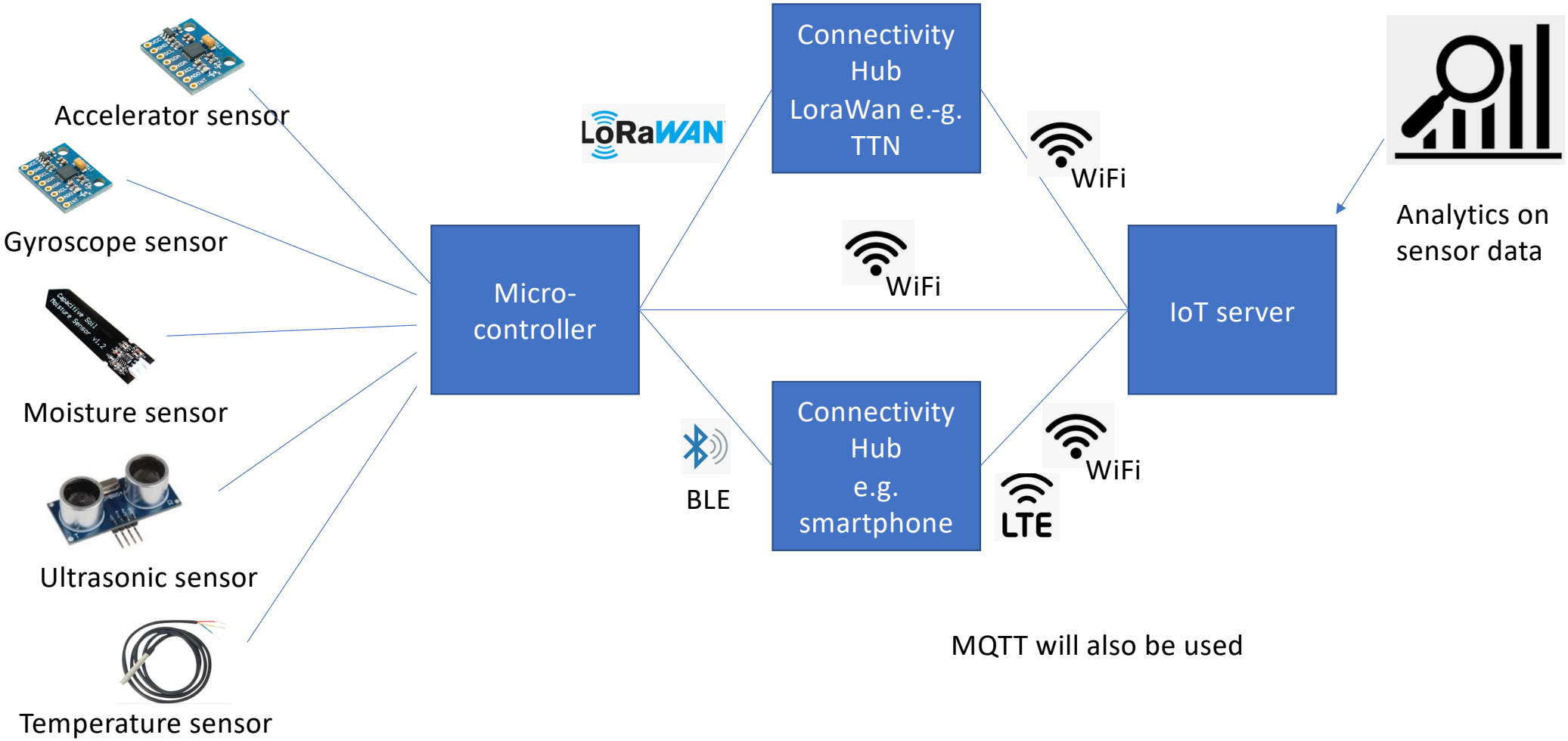
IoT Internet of Things

Hartmut Seitter

Block diagram IoT Lab – the architecture overview diagram we discussed so far



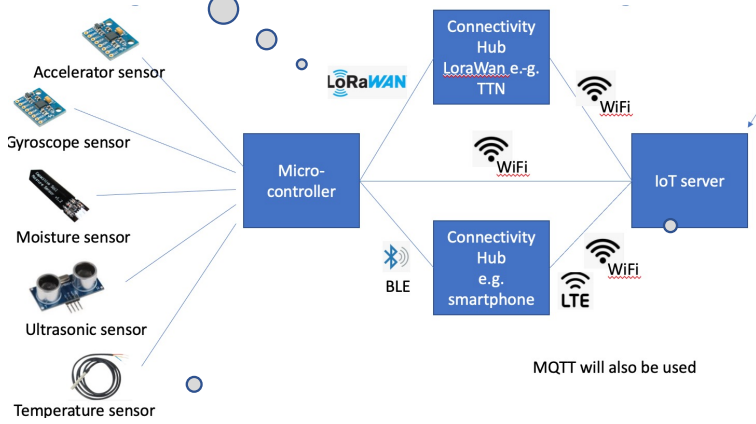
More topics to consider for an ‘End to End’ IoT Solution environment

IOT user

Data flow
Data values

Non functional requirements

Register and Monitor End Nodes



Analytics
AI
Machine learning

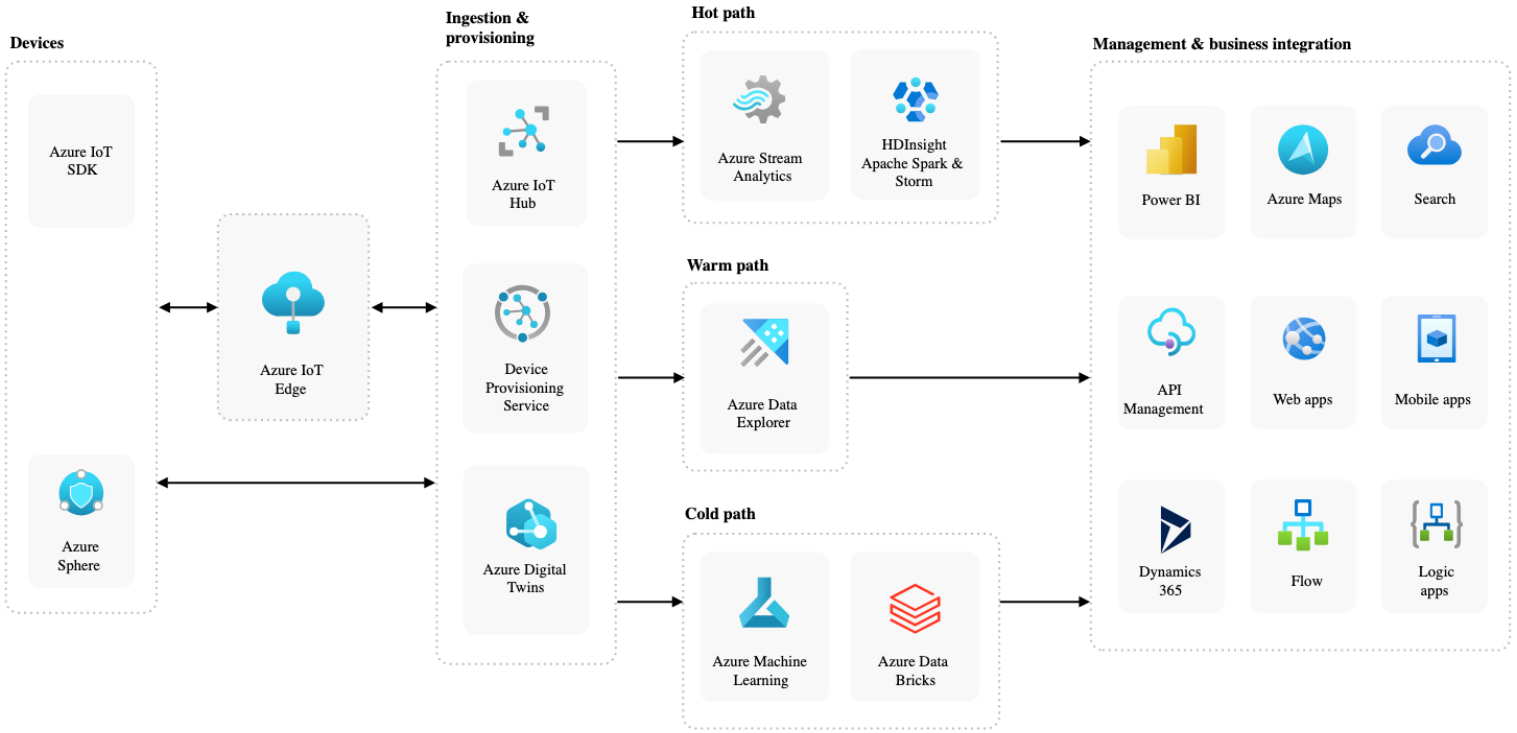
Storing the data for analytics ‘data lake’

Software updates End Nodes

IoT Security and IoT Governance

Different types of data
Timeseries data, Geodata, in motion data, telematic data

Let’s have a look to the Azure IoT Reference Architecture



Microsoft Azure

The ‘device’ side

- Azure IoT Device
  - Azure IoT supports a large range of devices, from microcontrollers running Azure RTOS and Azure Sphere to developer boards like MX Chip and Raspberry Pi. Azure IoT also supports smart server gateways capable of running custom code. Devices might perform some local processing through a service such as Azure IoT Edge, or just connect directly to Azure so that they can send data to and receive data from the IoT solution.

- Azure IoT Hub
  - Azure IoT Hub is a managed service hosted in the cloud that acts as a central message hub for communication between an IoT application and its attached devices. You can connect millions of devices and their backend solutions reliably and securely. Almost any device can be connected to an IoT hub.

- Azure IoT Hub Device Provisioning Service
  - Microsoft Azure provides a rich set of integrated public cloud services for all your IoT solution needs. The IoT Hub Device Provisioning Service (DPS) is a helper service for IoT Hub that enables zero-touch, just-in-time provisioning to the right IoT hub without requiring human intervention. DPS enables the provisioning of millions of devices in a secure and scalable manner.

The ‘application’ side

- Azure IoT hot path
  - The **hot path** analyzes data in near-real-time as it arrives. Hot path telemetry must be processed with very low latency. The hot path typically uses a stream processing engine. Consider using services such as Azure Stream Analytics or Azure HDInsight. The output might trigger an alert, or be written to a structured format that can be queried using analytical tools.

- Azure IoT warm path
  - The **warm path** analyzes data that can accommodate longer delays for more detailed processing. Consider Azure Data Explorer for storing and analyzing large volumes of data.

- Azure IoT cold path
  - The **cold path** performs batch processing at longer intervals, like hourly or daily. The cold path typically operates over large volumes of data, which can be stored in Azure Data Lake Storage. Results don't need to be as timely as in the hot or warm paths. Consider using Azure Machine Learning or Azure Databricks to analyze cold data.

End to End point of view
Walk through – The device itself and the device registration

- **An IoT device** (source: Microsoft azure)
  - Process an unique identity that distinguishes it within the solution
  - Has properties, or state that the application can access
  - Sends events to the IoT platform for the application to act on
  - Receives commands from the application to act on

- **Device Metadata and device registration (these are sample attributes)**
  - Device Type
  - Type of Sensor
  - Settings e.g. which parameter the operator can set
  - Measurement e.g. telemetry the device sends
  - Location Info
  - Rules and Actions e.g. to enable operators to monitor the device
  - Hardware and software version
  - Value range
  - Activating / onboarding the device (e.g. OTAA over the air activation / ABP activation by personalization) – Initial key settings
  - …

➢ **An application component is required to cover this functionality**

End to End point of view
Device monitoring and measurement monitoring

- **Is device monitoring required?**
  - **Keep alive ping**
  - **Watchdog functionality**
  - **Remote reset**
  - **Does the device operate differently from ‘normal operation’**
- **Is measurement monitoring required?**
  - **What to do if no data arrive**
  - **How to handle missing values**
  - **How to handle out of range values**
  - **…..**
- **Alerting**
  - **When something occurs which should be reported**

➢ **An application component is required to cover this functionality**

End to End point of view Walk through – Software updates and ‘OTAA over the air activation’

- **Device activation over the air (remember IoT Security – LoRaWAN example)**
- **Software updates over the air?**
    - **Software updates over the air**
        - **Memory available for new download**
        - **Fall back scenario**

➢ **An application component is required to cover this functionality**

End to End point of view
Walk through – IoT Security – IoT Governance

- **Authentication**
  - Device Authentication - This involves verifying the identity of IoT devices before allowing them to connect to a network or communicate with other devices.

- **Integrity**
  - Data integrity ensures that the information collected by IoT devices is accurate and has not been tampered.
  - Methods; checksums, hashes, and cryptographic signatures

- **Confidentiality**
  - Confidentiality ensures that data collected and transmitted by IoT devices is kept secure and not exposed to unauthorized individuals or entities.
  - Method: Using secure communication protocols (like HTTPS, MQTT with TLS, or CoAP over DTLS)

- **Robustness**
  - How fault tolerant the system will be. Redundency (e.g. backup power supplies)…..

End to End point of view
Walk through – IoT Security – IoT Governance

- **Availability**
  - Availability is often quantified as a percentage representing the time during which devices and services are operational. E.g. 99,1%
- **Reliability**
  - Reliable IoT systems consistently perform their tasks correctly, providing accurate data and responses under normal and varying operating conditions.
- **Interoperability**
  - Interoperability often relies on the establishment of common standards and protocols that enable devices from different manufacturers to communicate effectively.
  - Examples include MQTT, CoAP, and HTTP/REST for messaging,
  - as well as standards like Zigbee, Z-Wave, and Bluetooth for wireless communication and Thread / Matter
- **Accessability**

End to End point of view
Walk through – What to do with the data received

- **Data lake (big data)**
  - **Format of the data**
  - **Type of the data**

    - **Timeseries data**
      - **Characteristic: Sequence of values or events that are recorded in specific time intervals. Each data value is associated with a time stamp.**
      - **Time series data often show**
        - **Trends**
        - **Seasonality**
        - **Cycling Patterns**
        - **Stationarity**
        - **Autocorrelation**
        - **Noise**
        - **Irregular Values**

      - **Very often are timeseries data stored in timeseries databases**

End to End point of view
Walk through – What to do with the data received

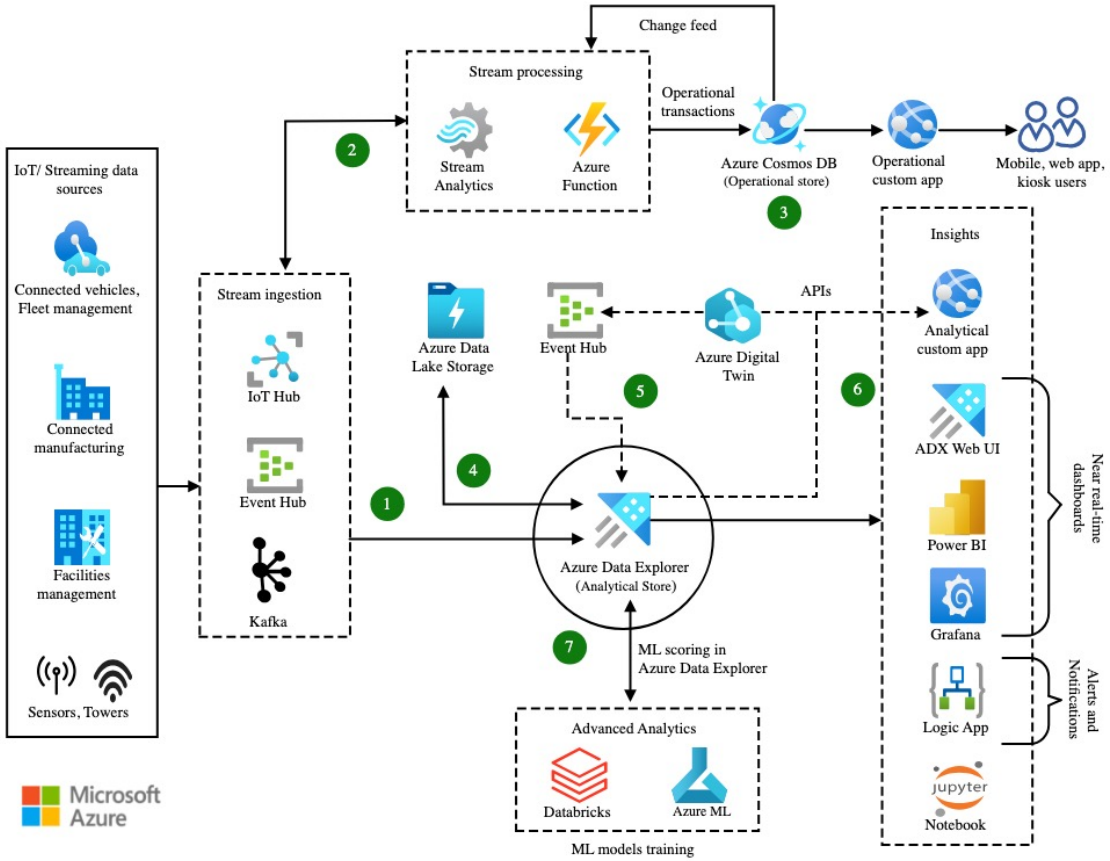
- **Data lake (big data)**

  - **Geolocation data**
    - Geolocation data refers to information that identifies the geographic location of a device, object, or individual.
    - Geolocation data are stored in SQL, NoSQL and Databases specialized for geolocation data such as PostGIS (extension of PostgreSQL), Oracle Spatial, MongoDB with geospatial querying features
  - **Telemetry data**
    - Telematic data refers to information collected from remote devices and transmitted over telecommunications networks. This type of data is commonly associated with vehicles and includes a combination of telecommunication and monitoring technologies
  - **Images, Audio data**
  - **Automation data**
  - **…..**

End to End point of view
Walk through - Analytics

- **Prescriptive analytics**
  - is used to analyse which steps to take for a specific situation. It's often described as being a combination of descriptive and predictive analysis. When used in commercial applications, prescriptive analytics helps decipher large amounts of information to obtain more precise conclusions.
- **Spatial analytics.**
  - This method is used to analyze location-based IoT data and applications. Spatial analytics deciphers various geographic patterns, determining any type of spatial relationship between various physical objects. Parking applications, smart cars, and crop planning are all examples of applications that benefit from spatial analytics.
- **Streaming analytics.**
  - Streaming analytics, sometimes referred to as event stream processing, facilitates the analysis of massive "in-motion" data sets. These real-time data streams can be analyzed to detect emergency or urgent situations, facilitating an immediate response. The types of IoT data that benefit from streaming analytics include those used in traffic analysis, air trafficking, and the tracking of financial transactions.
- **Time series analytics.**
  - Time series analytics is based on time-based data, and data are analyzed to reveal any anomalies, patterns, or trends. Two systems that greatly benefit from time series analytics are health-monitoring and weather-monitoring systems

Sample – IoT Analytics with Azure Data Explorer



Dataflow in IoT Analytics with Azure Data Explorer

- Azure Event Hubs, Azure IoT Hub, or Kafka ingest a wide variety of fast-flowing streaming data such as logs, business events, and user activities.
- Azure Functions or Azure Stream Analytics process the data in near real time.
- Azure Cosmos DB stores streamed messages in JSON format to serve a real-time operational application.
- Azure Data Explorer ingests data for analytics, using its connectors for Azure Event Hubs, Azure IoT Hub, or Kafka for low latency and high throughput.
- Alternatively, you can ingest blobs from your Azure Blob Storage or Azure Data Lake Storage account into Azure Data Explorer by using an Event Grid data connection.
- You can also continuously export data to Azure Storage in compressed, partitioned Apache Parquet format, and seamlessly query the data with Azure Data Explorer. For details, see Continuous data export overview.
- …. and so on – see https://learn.microsoft.com/en-us/azure/architecture/solution-ideas/articles/iot-azure-data-explorer

Dataflow in IoT Analytics with Azure Data Explorer

- Azure Data Explorer provides native advanced analytics for:

  - Time series analysis.
  - Pattern recognition.
  - Anomaly detection and forecasting.
  - Anomaly diagnosis for root analysis.

- **Very often the anaconda framework and Jupiter notebooks are used to analyse data and get insight into the data.**

End to End point of view
Walk through – Non functional requirements

- Performance
  - Function offload to the edge. The system must have a centralized configuration with offload to the edge for rules-based decision-making and streaming analytics.
  - Hybrid cloud. It must be possible to deploy the platform layer on premises, in the cloud, or part on premises and part in the cloud.
  - Modular design. The microservices orientation allows for the modular design of IoT solutions.
  - Unstructured and structured data. The system must handle structured data—both real-time messages and bulk transfer—and unstructured data, including acoustic files, image files, recordings, and text documents.
  - User response time. The system must provide acceptable response times to users regardless of the volume of data that is stored and the analytics that occurs in background.
  - Real-time communications (request/response). Bidirectional, near real-time communications must be supported. This requirement is related to the requirement to support industrial and device protocols at the edge. Responses might need to happen in 1 second or less depending on the use case. This requirement also drives a need for edge-level analytics and decision-making for that class of use cases.
  - Time-series capture granularity. A timestamp to at least the millisecond is required so that even if the data is too late for real-time control, the catch-up analytics can see the sequence of events.
- Scalability
  - Horizontal scalability. The system must handle expanding load and data retention needs that are based on the upscaling of the solution scope, such as extra manufacturing facilities and extra buildings.

End to End point of view
Walk through – Non functional requirements

- Security
  - Data security. All persisted data requires secure access.
  - Device security. Devices must register and communicate securely, such as by using transport layer security (TLS). Unauthorized devices are prohibited.
  - User security. User logon to any device must be secure and validated for each role. This security must be attached to the LDAP or user registry that customers use.
  - Application security. Authorized users of the system who exchange information must be able to do so with appropriate security controls.
- Usability
  - Mobile support. Users must be able to interact in the same roles and on the same tasks on computers and mobile devices where practical, given mobile capabilities.
- Manageability
  - Incident management. The system must include support for alerting, notification, and incident management.
  - Solution management. The system must support centralized solution management so that system's support personnel can quickly determine the root cause of a problem and fix it to avoid downtime.

End to End point of view
Walk through – Non functional requirements

- Availability
  - High availability (HA). Some IoT solutions and domains demand highly available systems for 24x7 operations. That said, this type of system isn't a *critical production* application, which means that operations or production don't go down if the IoT solution is down.

- Maintainability
  - Adaptable and flexible. It must be possible, and relatively easy, to rapidly adapt the system to a change in processes, participants, or information that is exchanged.
  - Maintainability and uptime. It must be possible to perform maintenance on the system without violating service level agreements (SLAs) for uptime. In a 24x7 environment, you can't take down the system completely to perform maintenance.

- Volumetrics
  - Big data. The system must store and analyze volumes of data, both historical and current, on scales that are commonly known as *big data*.
  - Platform speed, capacity, and accessibility. The platform must support high volumes of near real-time data transmissions 24x7. This requirement includes support for remote locations and mobile data sources.