CIS373 - Pervasive Computing Edge/Fog Computing

Erik Fredericks // frederer@gvsu.edu

Adapted from materials provided by Xiang Cao

BUT FIRST

https://futurism.com/the-byte/amazon-abandons-ai-stores

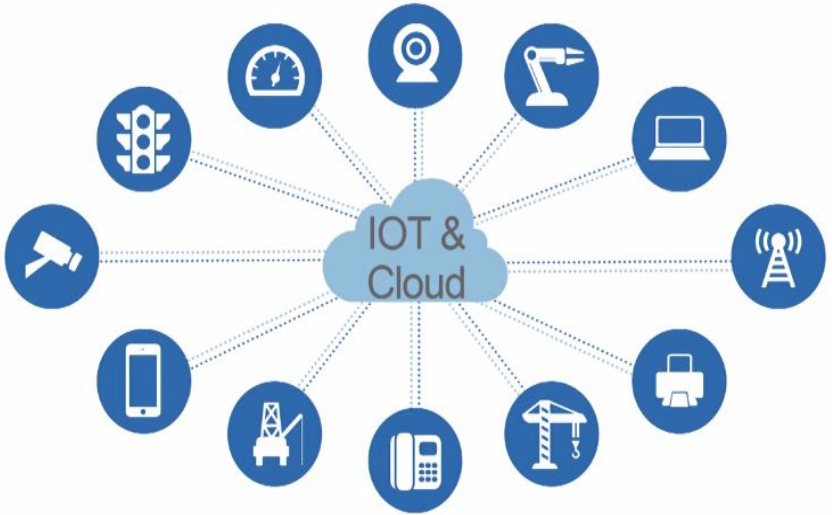
https://gizmodo.com/amazon-reportedly-ditches-just-walk-out-grocery-stores-1851381116

Cloud Computing

Many years ago, the IoT started its communication with the centralized server in the cloud like this, as shown in the figure.

Basically, it is making use of Centralized servers hosted in the core internet rather than using a local server or personal system for huge processing/computation or storage of data.

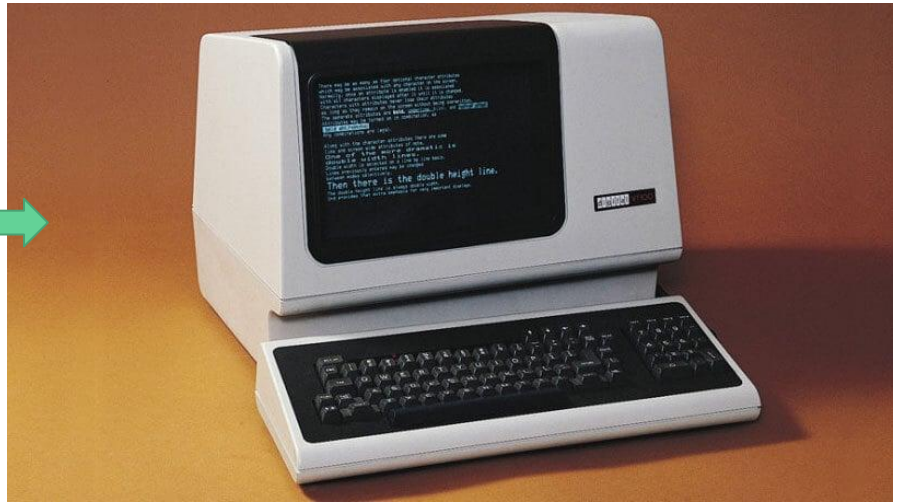
In the present times almost all the organizations uses cloud.



Interestingly...

mainframe

dumb terminal



Cloud computing examples

Spinning up a virtual machine

- Remote machine available for use

Serverless functions

- Remove function available for use

Cloud computing examples

BigQuery - Google Cloud Big Data → GitHub queries

```
#standardSQL
SELECT SUM(copies) FROM `bigquery-public-data.github_repos.sample_contents`
WHERE NOT binary AND content LIKE '%This should never happen%'
```

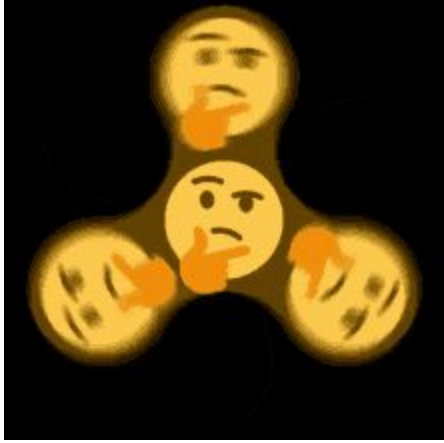
```
#standardSQL
SELECT SUM(copies) FROM `bigquery-public-data.github_repos.sample_contents`
WHERE NOT binary AND (content LIKE '%This should never happen%' OR content
LIKE '%FIXME%' OR content LIKE '%TODO%')
```

Discussion

Need for **edge/fog** computing

- Why can't we just use the cloud for everything?

- Why can’t we just use edge/end systems for everything?



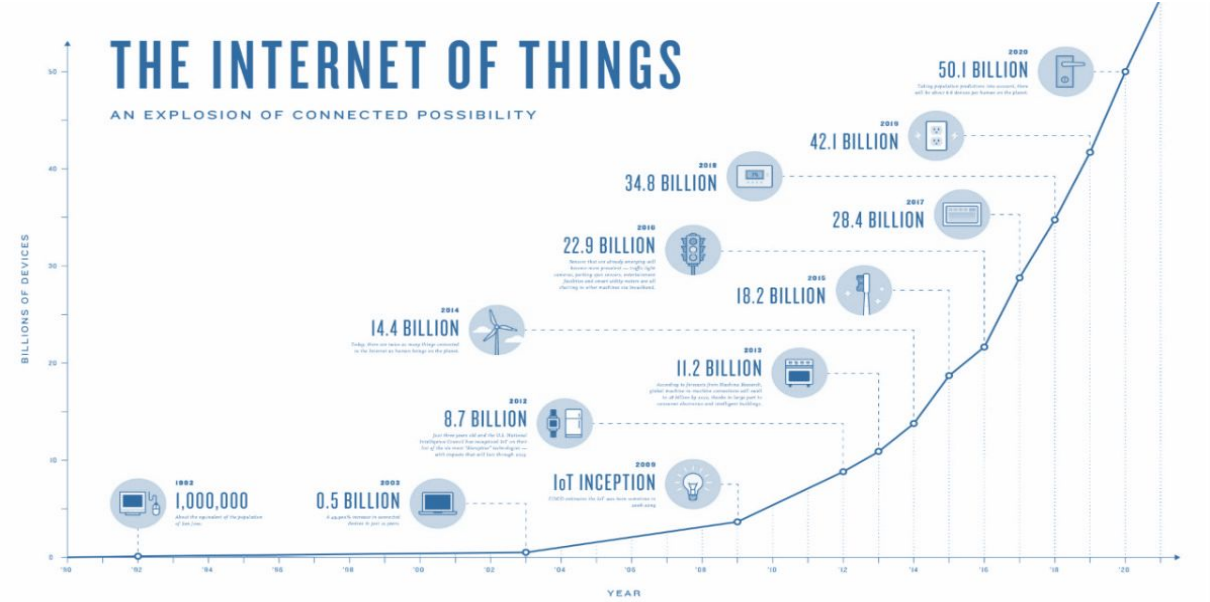
Need for fog computing

Why can't we just use the cloud for everything?

- Cloud computing frees the enterprise and the end user from many details.
- This becomes a problem for latency-sensitive applications.
- **$$$**

Why can’t we just use edge/end systems for everything?

- Physical constraints: Energy, space, etc.,
- Limited computing and storage capacity.



Number of end devices that are connected to internet rose above 50+ billion in 2020.

Cloud computing architectures won't be able to handle the demand of IoT. So only cloud is **not the optimal solution** to handle this massive explosion.

Fog is needed in between to optimize – need for an **intermediary** between cloud and end systems.

Why is **off-premise** cloud not suitable for certain IoT applications?

- Cloud infrastructure is centralized, federated, consolidated, shared, automated, and programmable

- **Latency and response time** is often a critical part, especially when you deal with human life or emergency procedure.

- **Bandwidth cost and capacity** is very often underestimated.
  - If you want to use many smart devices requiring each one to communicate with cloud, then you can quickly reach huge bandwidth requirements at a gateway level.

- **Security and privacy** - transmitting device data over any open and public network is risky

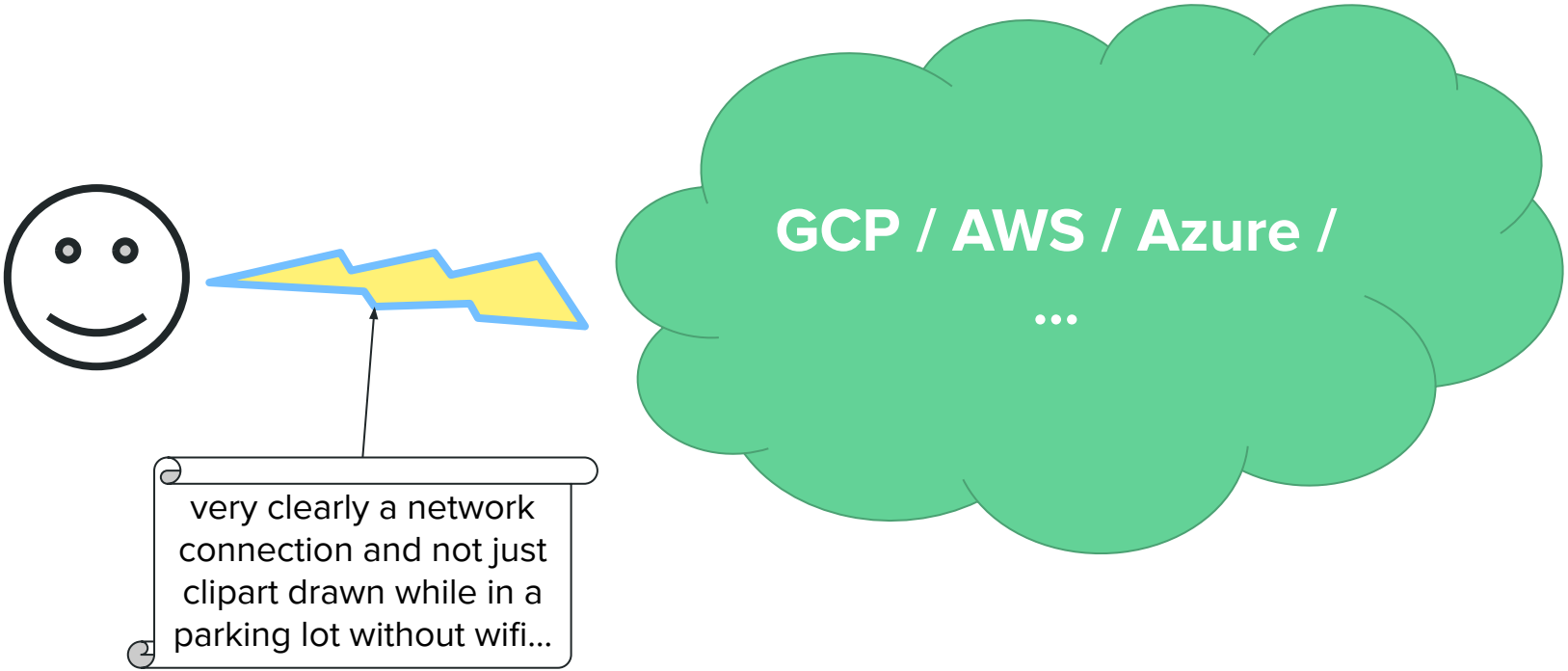
Why is **off-premise** cloud not suitable for certain IoT applications?

- **Power consumption**
  - Cloud computing is **energy-hungry** and that it is a concern for a low-carbon economy.

- **Data amount**
  - In a traditional cloud approach, huge amounts of untreated data are pumped blindly into the cloud.
  - This vision is really not the best efficient and it is much wiser to pre-treat data at a local level and to limit the cloud processes (i.e., cleaning/aggregation)

- **Offline** usages versus **online-only** usages
  - Pure cloud services do not allow offline usages.
  - It is a major shortcoming since smart cities and industry applications require a dual offline/online paradigm.

Ok, so an off-prem cloud doesn't always make sense..



Why do some IoT applications have to be real-time and at edge?

- **Volume and Velocity**
  - Ingesting, processing and storing such huge amounts of data which is gathered in real-time.

- **Security**
  - Devices can be located in sensitive environments, control vital systems or send private data.
  - With the number of devices and the fact they are not humans who can simply type a password, new paradigms and strict authentication and access control must be implemented.

- **Bandwidth**
  - If devices constantly send the sensor and video data, it will hog the internet and cost a fortune.
  - Edge approaches must be deployed to achieve scale and lower response time.

Why do some IoT applications have to be real-time and at edge?

- Real-time data capture, storage, processing, analytics, knowledge discovery, decision-making and actuation
  - **Less Latency and faster response**

- **Combining real-time data with historical state**
  - There are analytics solutions which handle batch quite well and some tools that can process streams without historical context

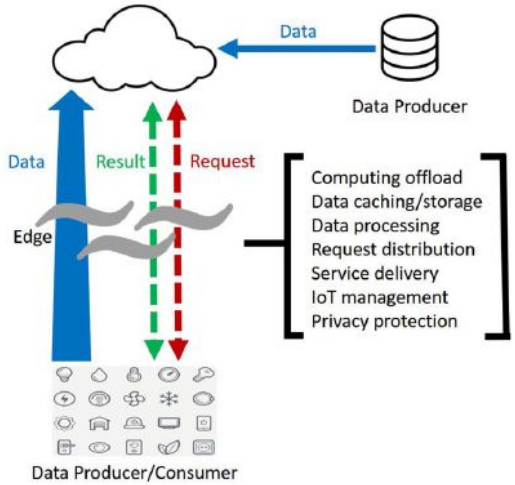
  - HOWEVER: It is quite challenging to analyze streams and combine them with historical data in **real-time**

What is edge/fog computing anyway?

We define “edge” as any computing and network resources along the **path between data sources and cloud data centers**.

Edge computing is typically interchangeable with fog computing.

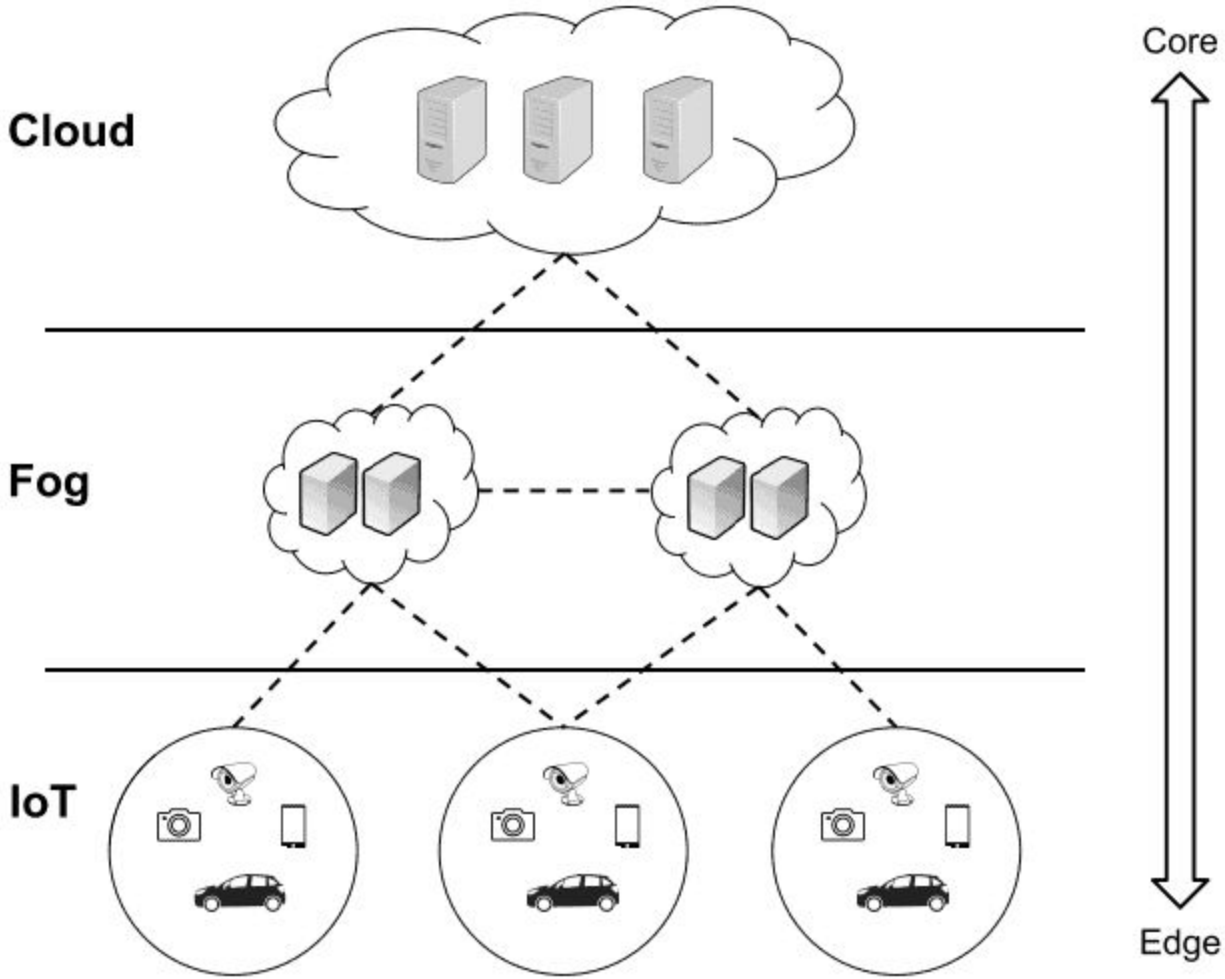
- Delineations between leaf-level nodes and intermediary nodes though!

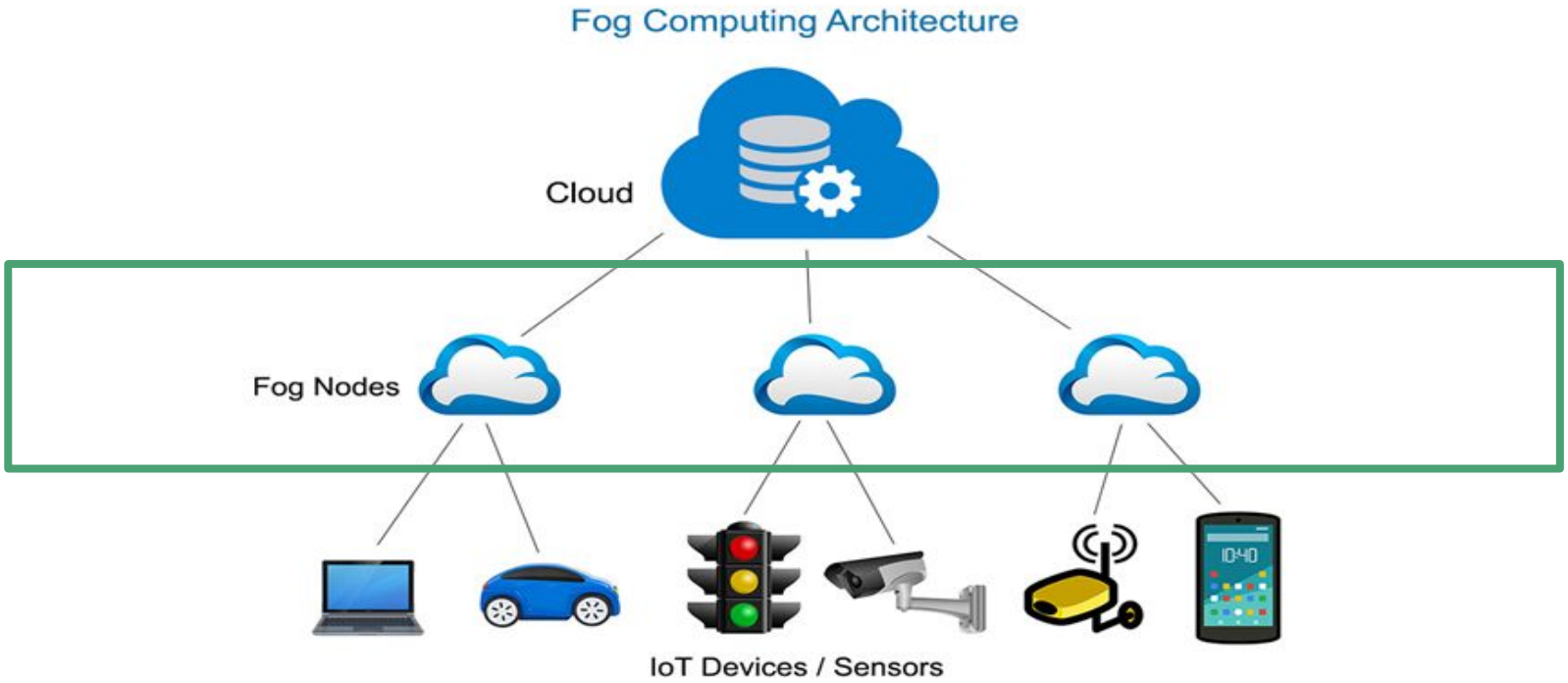


Edge Computing

Edge computing is a method of optimizing cloud computing systems by performing data processing at the edge of the network, near the source of the data. -- Wikipedia: Edge Computing

- Reduce data transfer
- Low latency
- Local feedback loop
- Privacy
- Security


Cloud
Fog
IoT
Core
Edge



Fog nodes can be deployed anywhere with a network connection: on a factory floor, on top of a power pole, alongside a railway track, in a vehicle, or on an oil rig. Any device with computing, storage, and network connectivity can be a fog node. Examples include industrial controllers, switches, routers, embedded servers, and video surveillance cameras.

Fog Computing

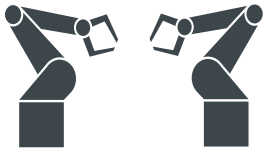
Fog computing is making use of **decentralized servers** in between network core and network edge for data processing and to serve the immediate requirements of the end systems.

Does this replace the cloud?
No, fog computing is non-trivial extension of cloud computing paradigm to the edge of the network.

- The cloud has better computing and storage power than the edge servers. Many computing or storage intensive applications should still run in the cloud.

Device categories

Edge and Cloud



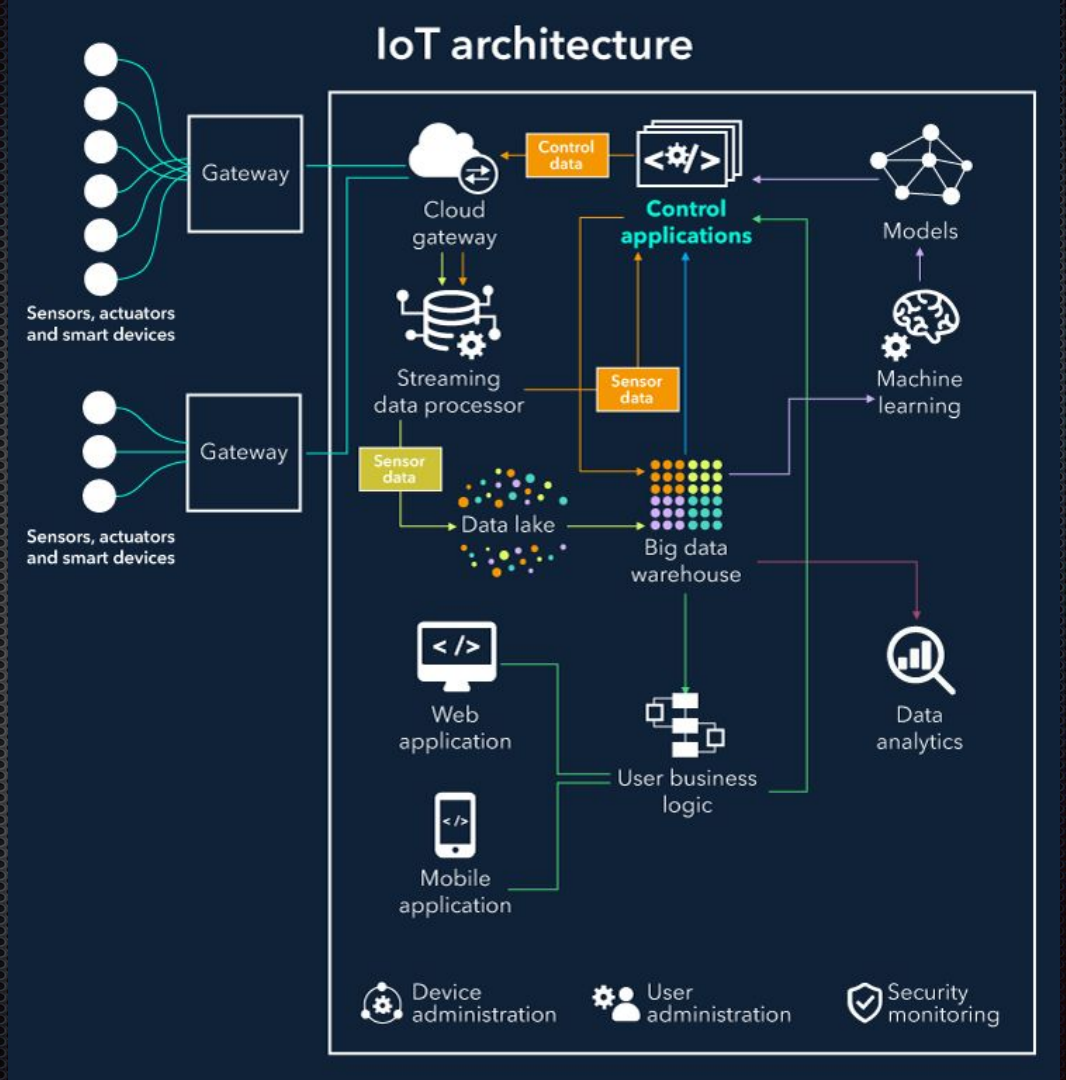
Cloud	Edge
Off site	On site
Limited control	Full control
Virtually unlimited resources	Limited resources
Easy scaling	No or limited scaling
High latency	Low Latency
Limited privacy	Full privacy

(Next slides % Leading Edge Data / Analytics - Dr. Shakhlo Ergashev + Dr. Schymik)

Internet of Things…

CIS641F2019

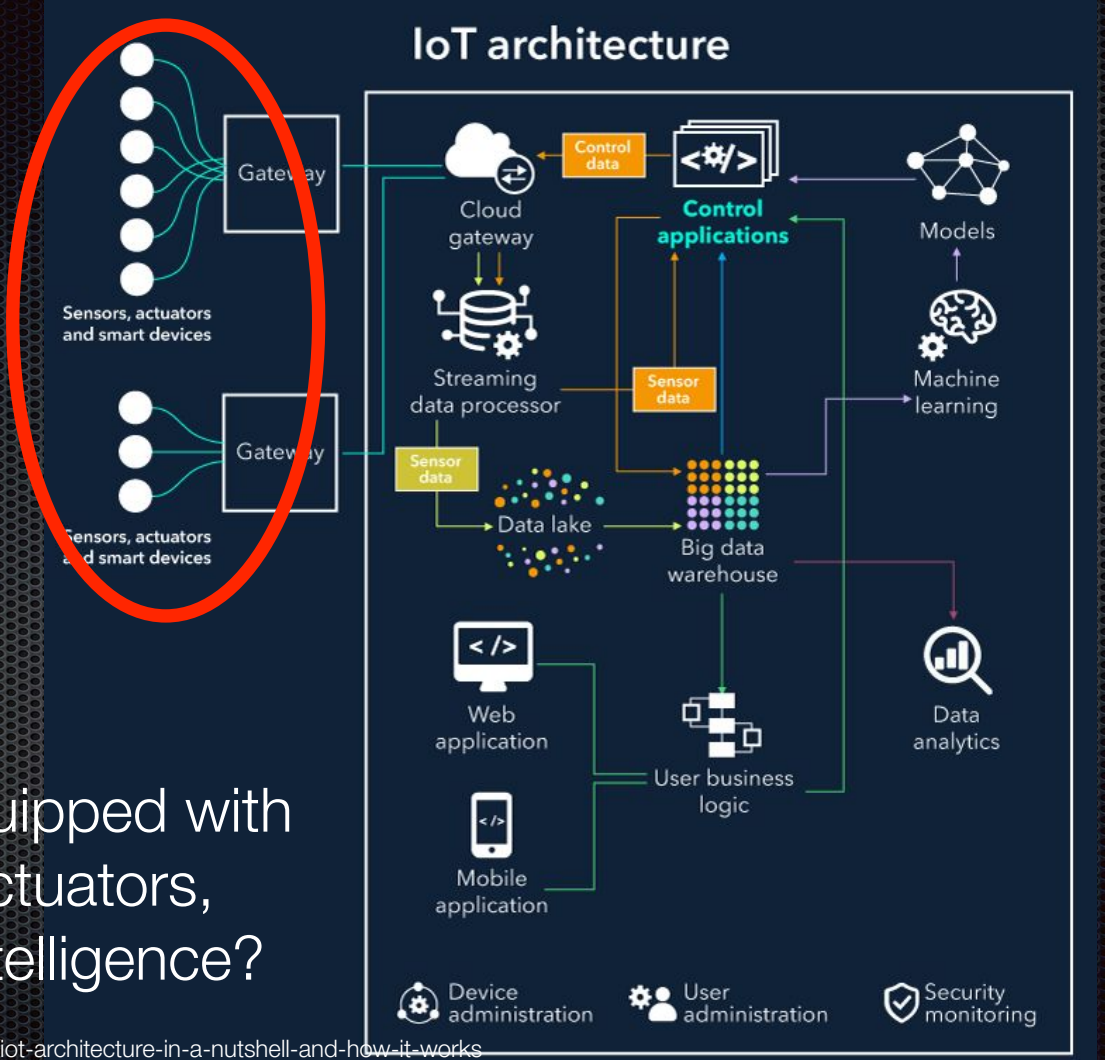
Last Lecture…



.. computational objects and things have been equipped with communication and interactive capabilities of embedded intelligence.

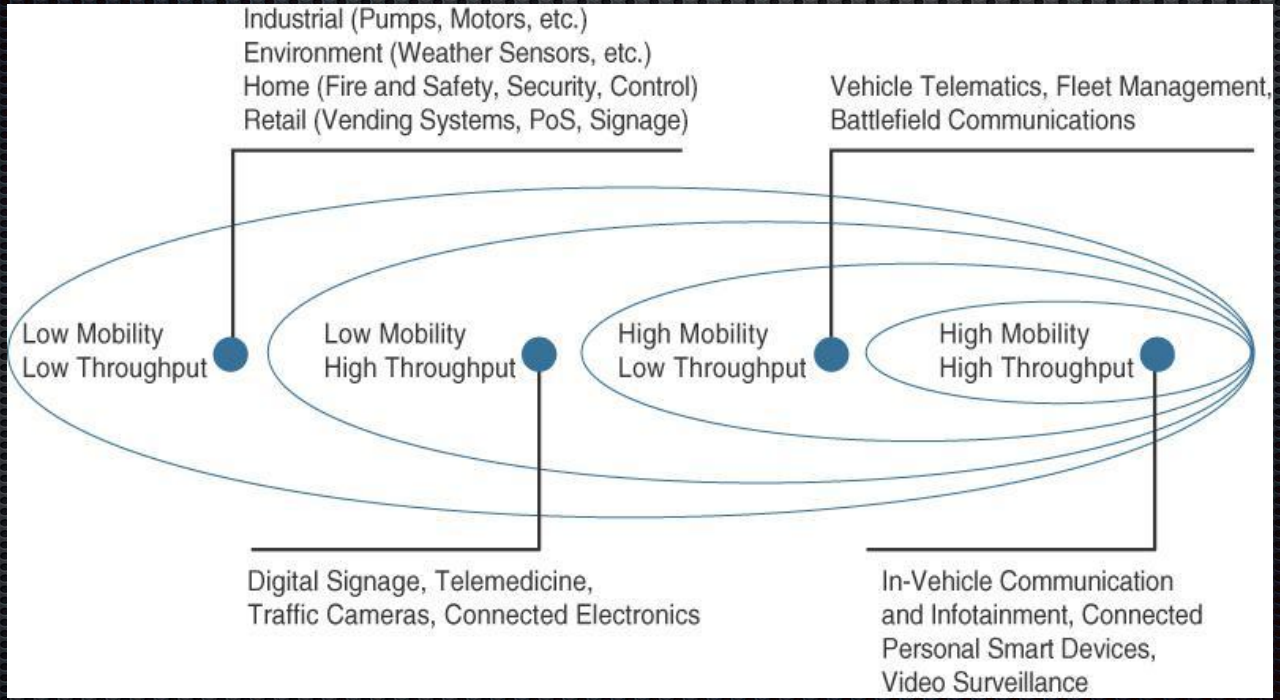
Things

- objects equipped with sensors, actuators, comms, intelligence?



https://www.scnsoft.com/blog/iot-architecture-in-a-nutshell-and-how-it-works

Things…



Hanes, et al, 2017 IoT Fundamentals…

Things classification

Battery or connected?

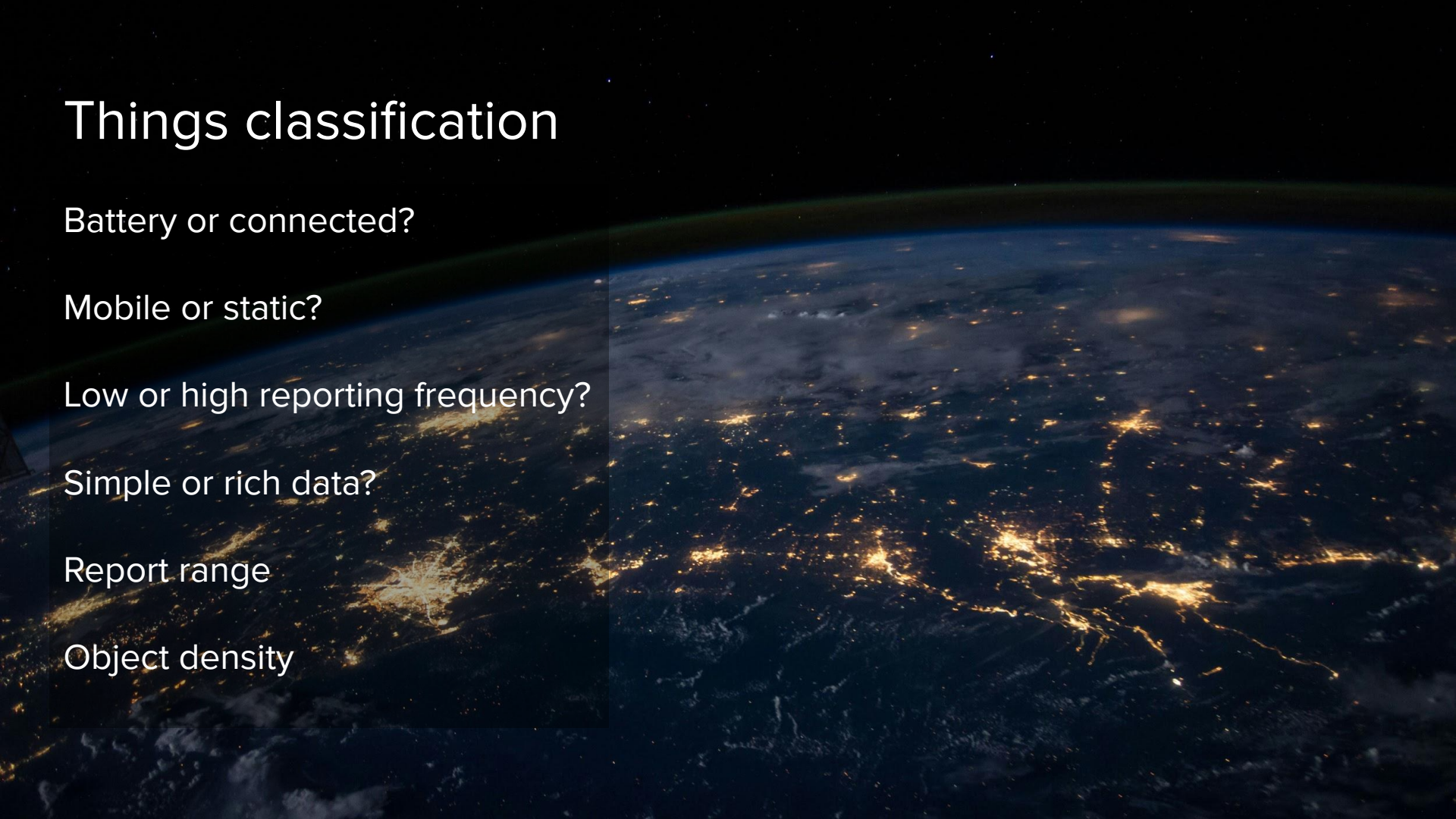
Mobile or static?

Low or high reporting frequency?

Simple or rich data?

Report range

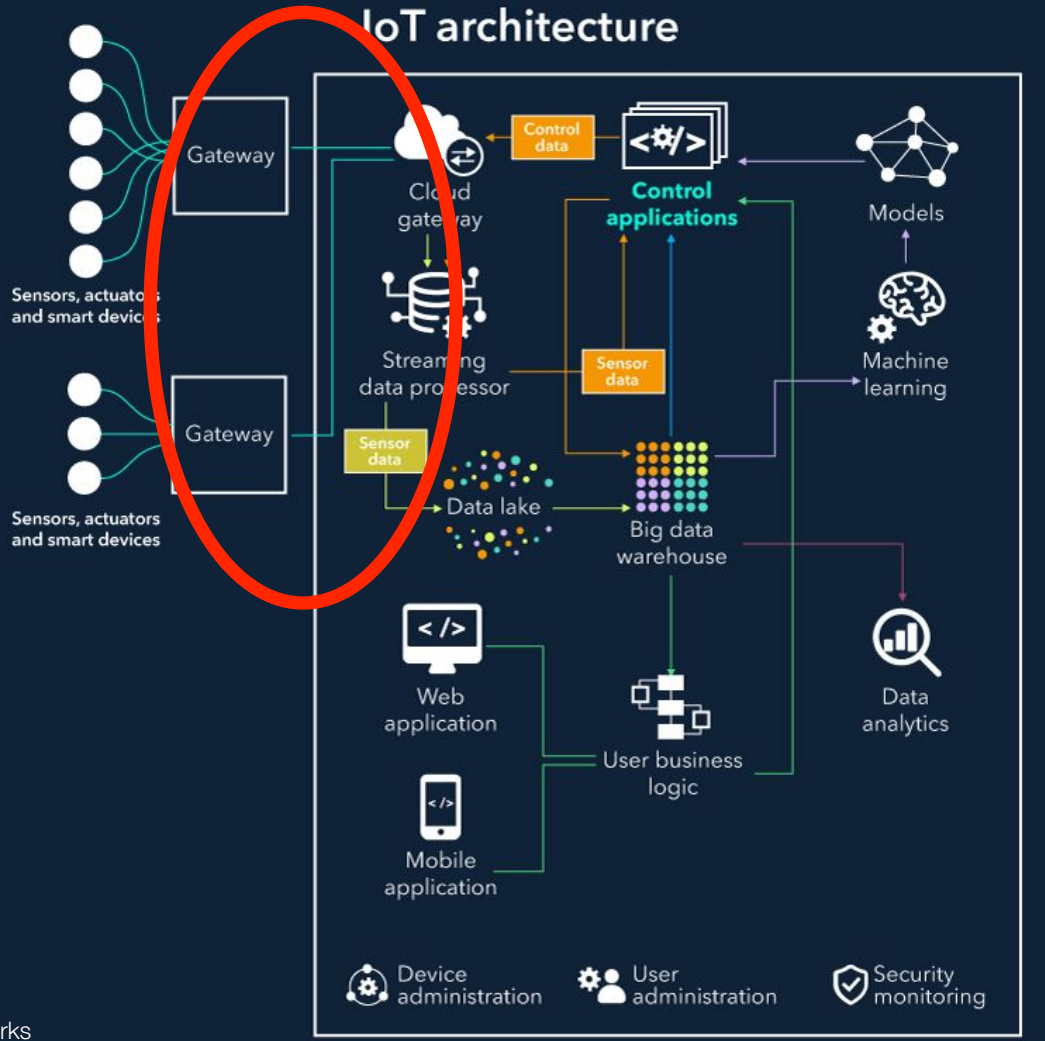
Object density



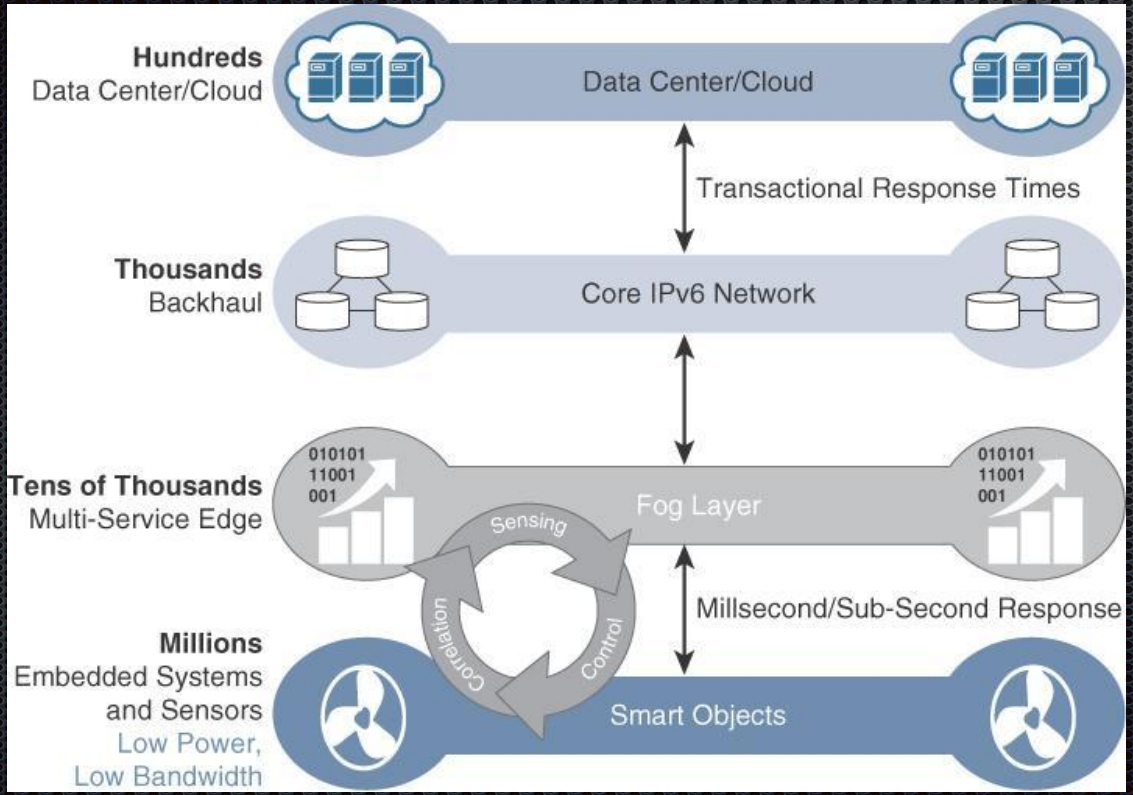
Gateways

- provide connectivity between things and the cloud
- some intelligence - filtering data processing (“fog” nodes)

https://www.scnsoft.com/blog/iot-architecture-in-a-nutshell-and-how-it-works



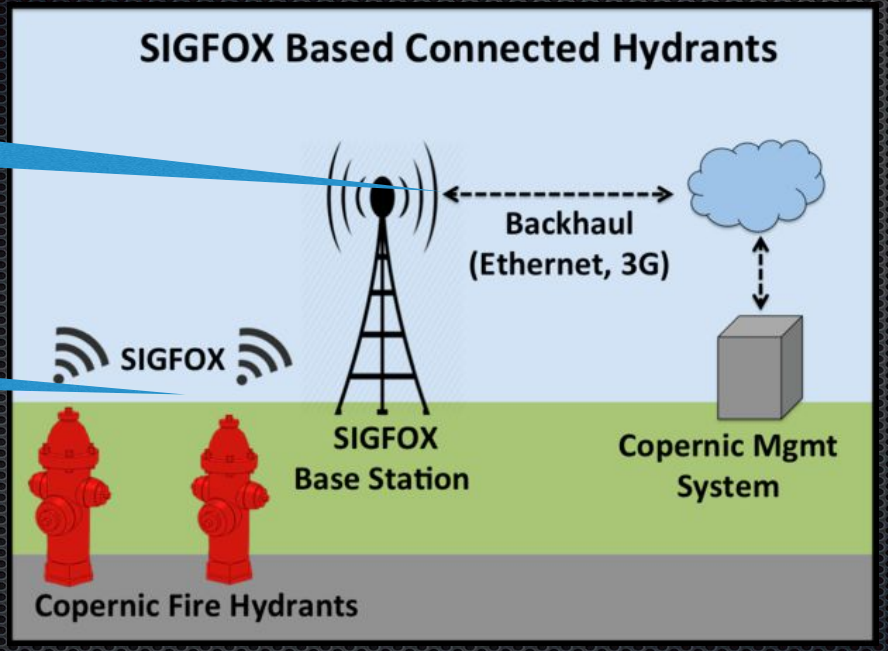
Gateways - Fog Layer



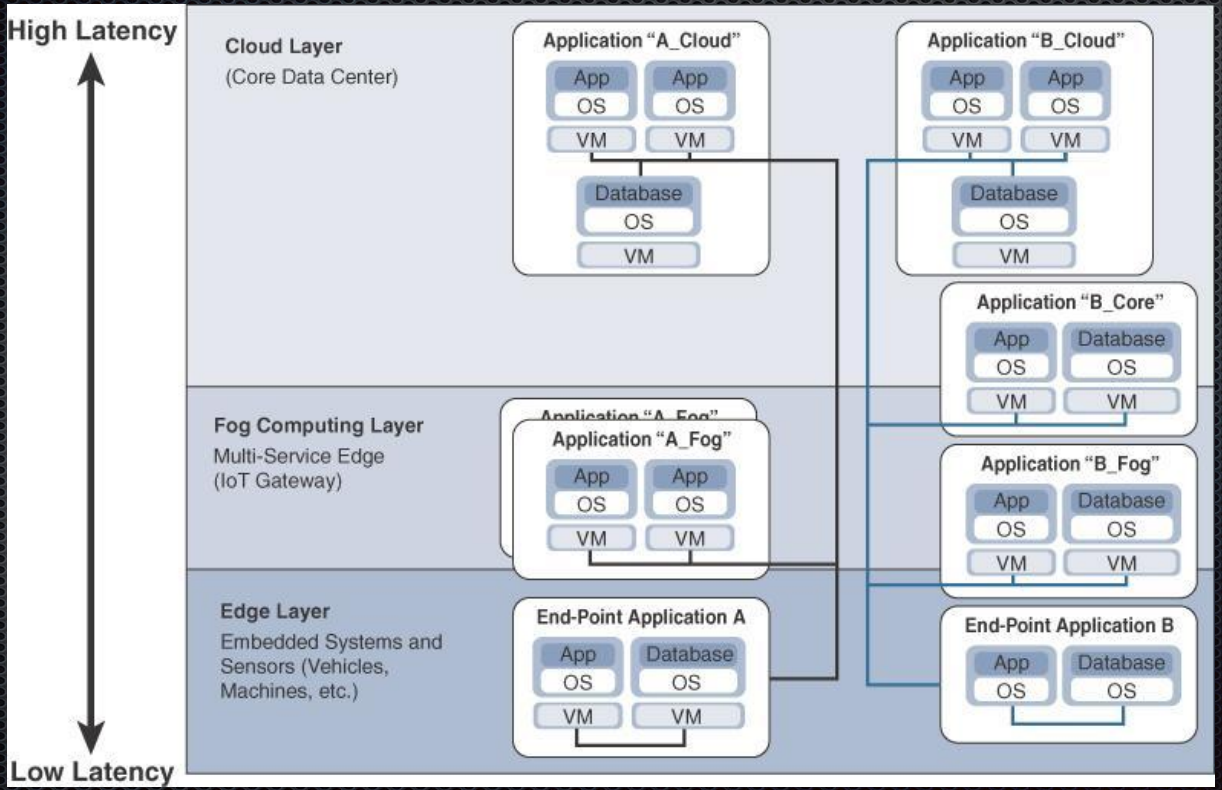
Edge Computing

Fog

Edge

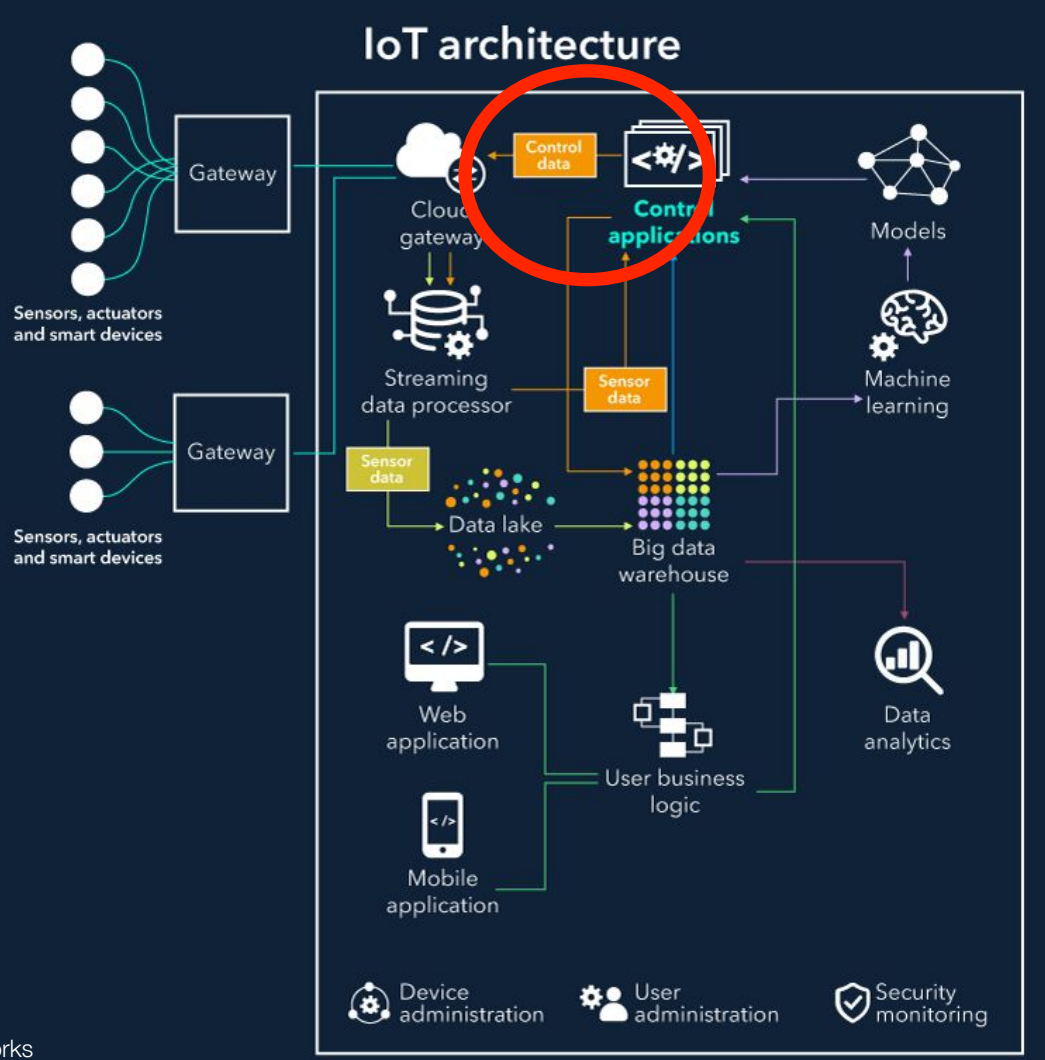


Gateway Hierarchy



Cloud

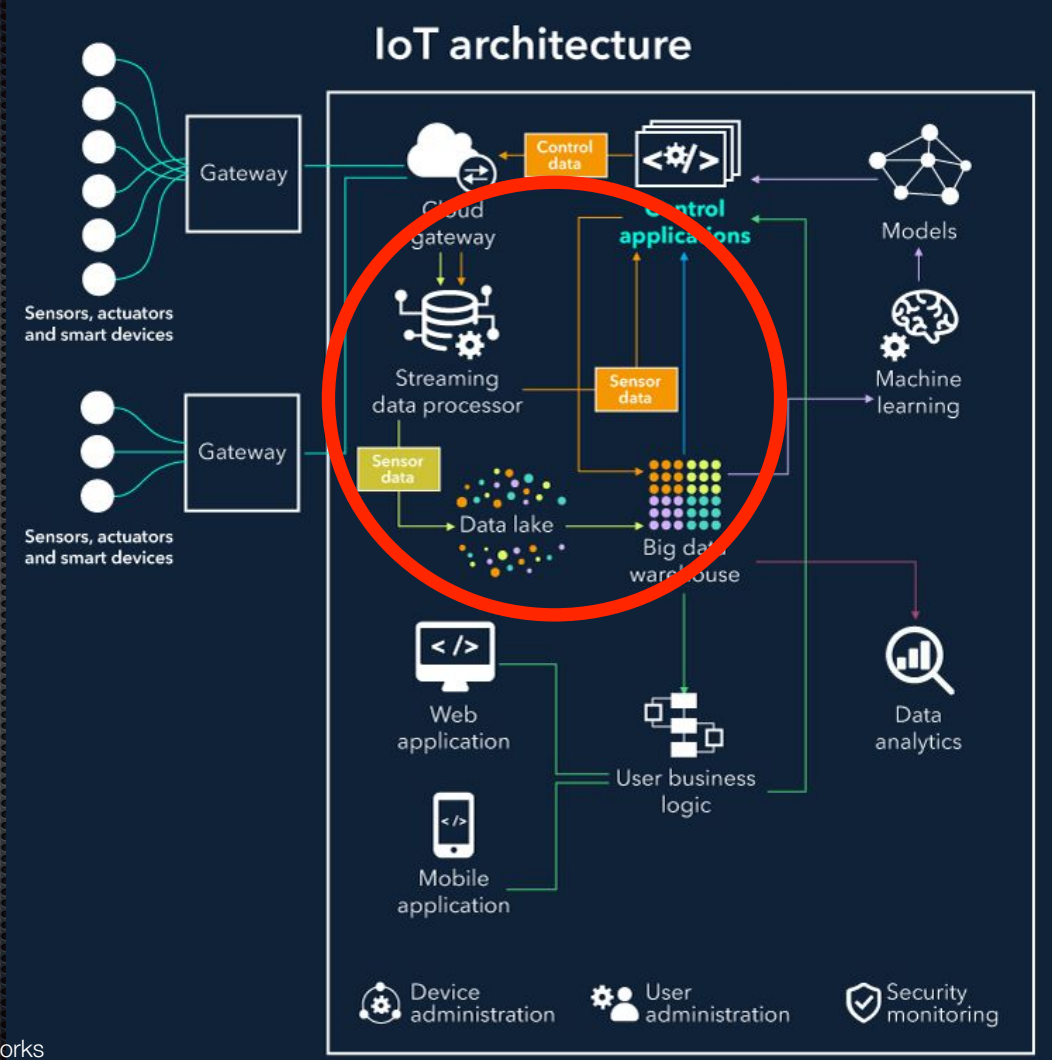
- facilitates transferring data from the IoT gateways to the cloud servers
- data compression
- protocol compatibility



https://www.scnsoft.com/blog/iot-architecture-in-a-nutshell-and-how-it-works

Data Management

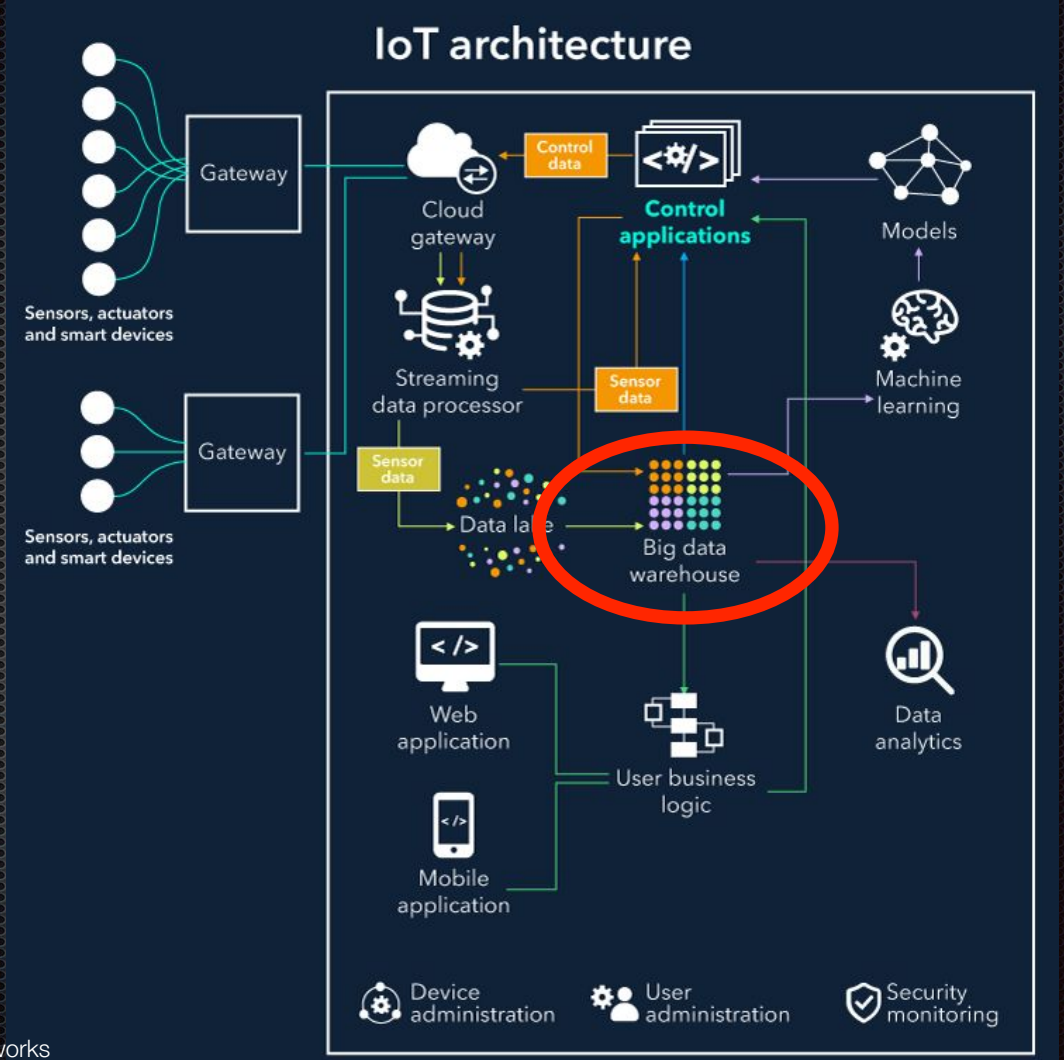
- streaming data processor
  - ETL to data lake



https://www.scnsoft.com/blog/iot-architecture-in-a-nutshell-and-how-it-works

Data Warehouse

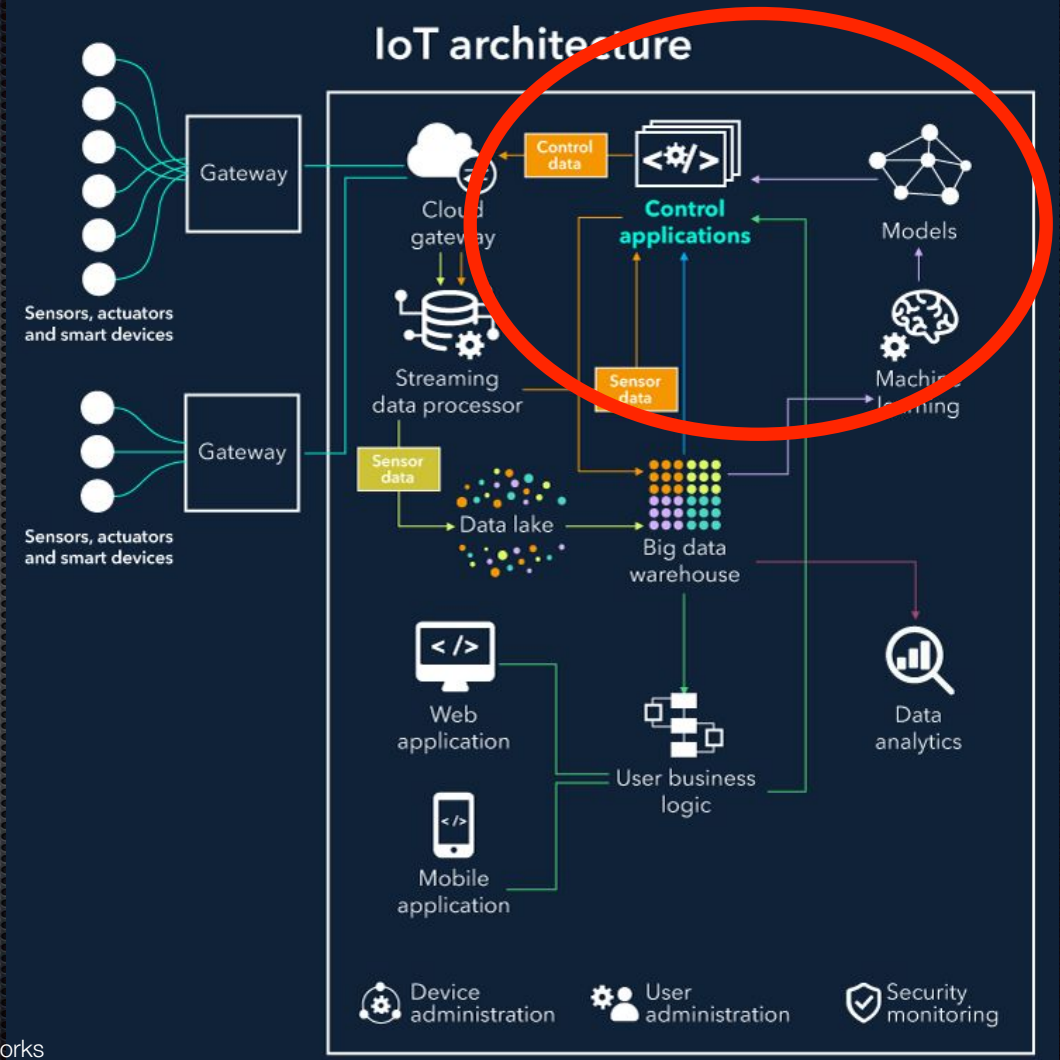
- ETL for specific processing needs



https://www.scnsoft.com/blog/iot-architecture-in-a-nutshell-and-how-it-works

Real-Time Control

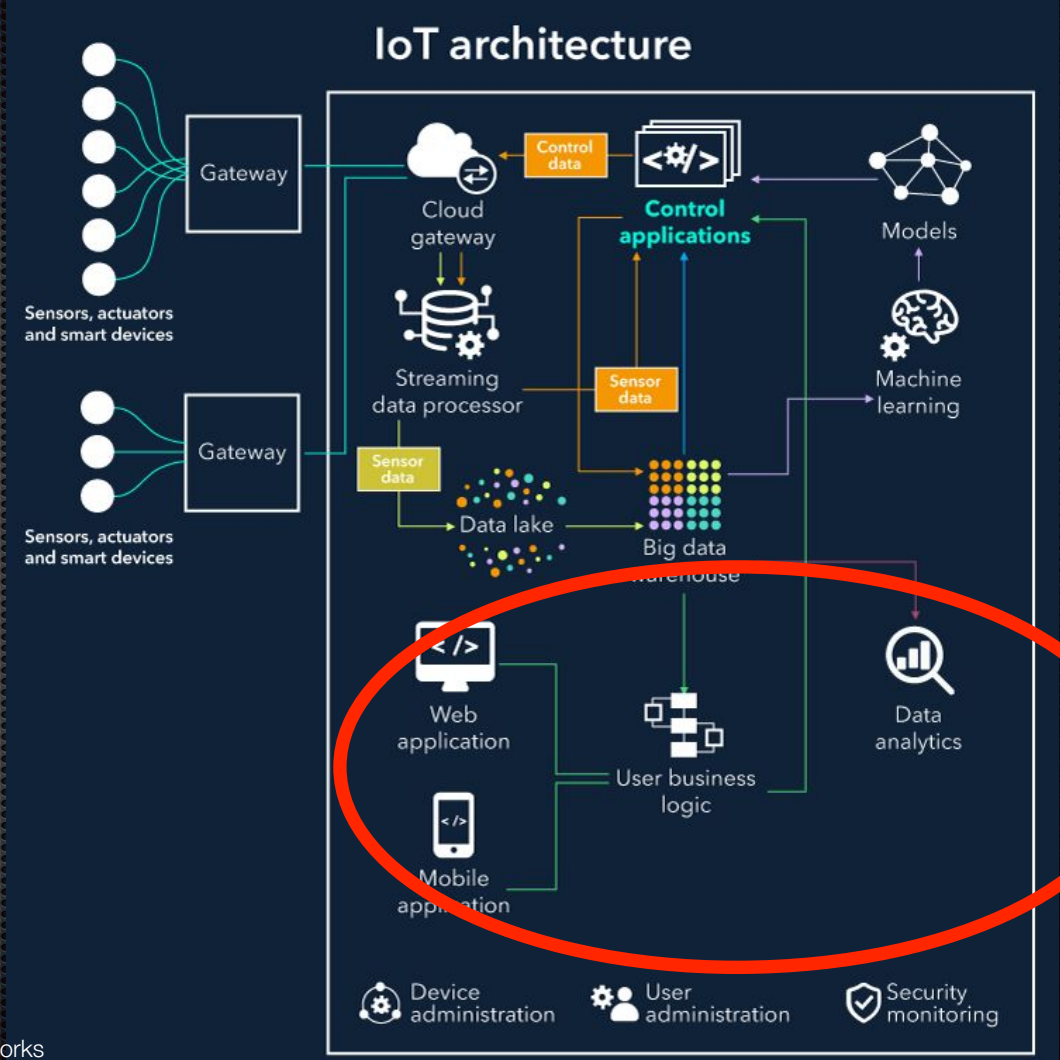
- ML/AI and other real-time modeling feeds Control Apps



https://www.scnsoft.com/blog/iot-architecture-in-a-nutshell-and-how-it-works

Business Logic

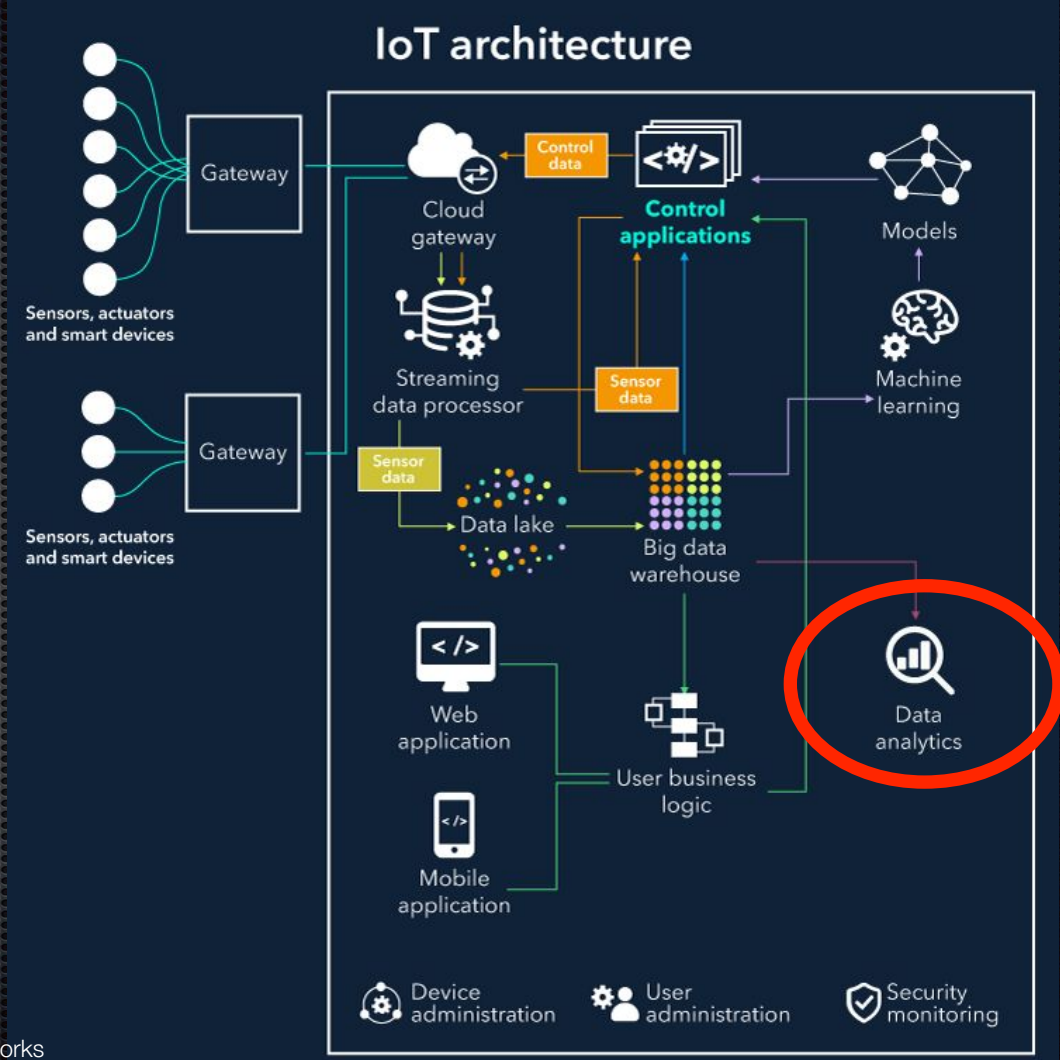
- connects the users to the ecosystem
  - remove monitoring and control
  - configuration, etc.
  - receive warnings/notifications



https://www.scnsoft.com/blog/iot-architecture-in-a-nutshell-and-how-it-works

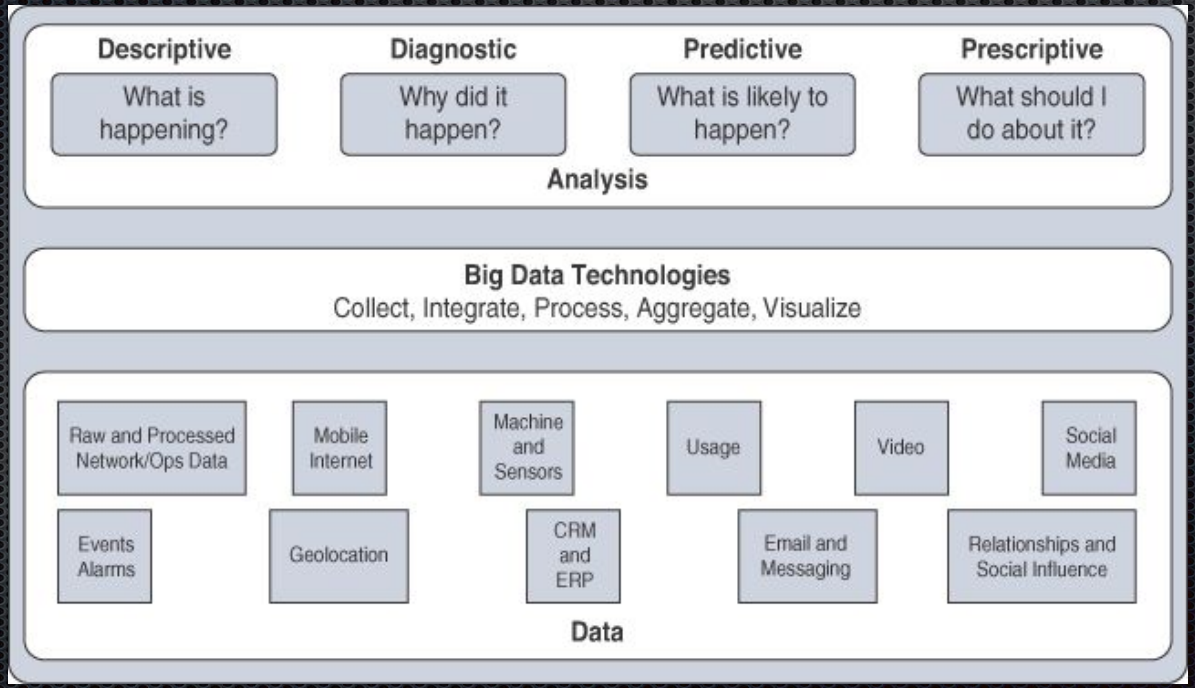
Analytics

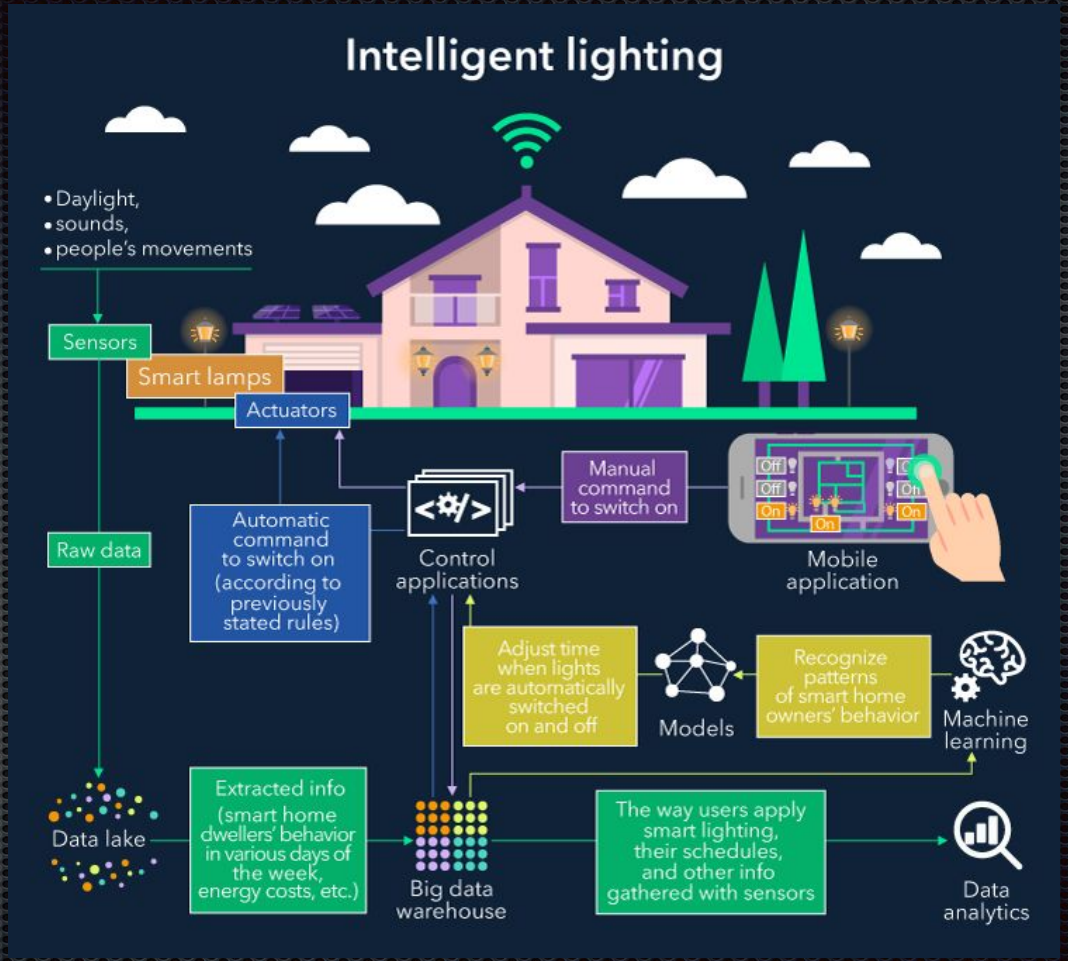
- insights from the data?
- data is $$$



https://www.scnsoft.com/blog/iot-architecture-in-a-nutshell-and-how-it-works

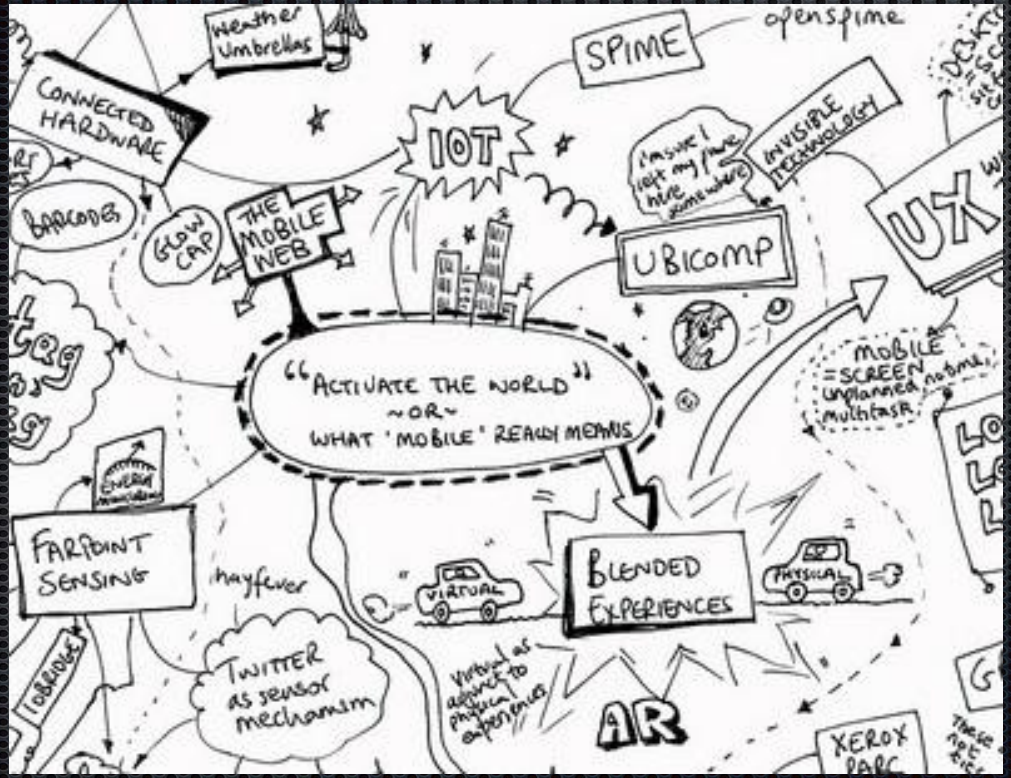
Analytics - a thing in and of itself?





https://www.scnsoft.com/blog/iot-architecture-in-a-nutshell-and-how-it-works

Data Challenges

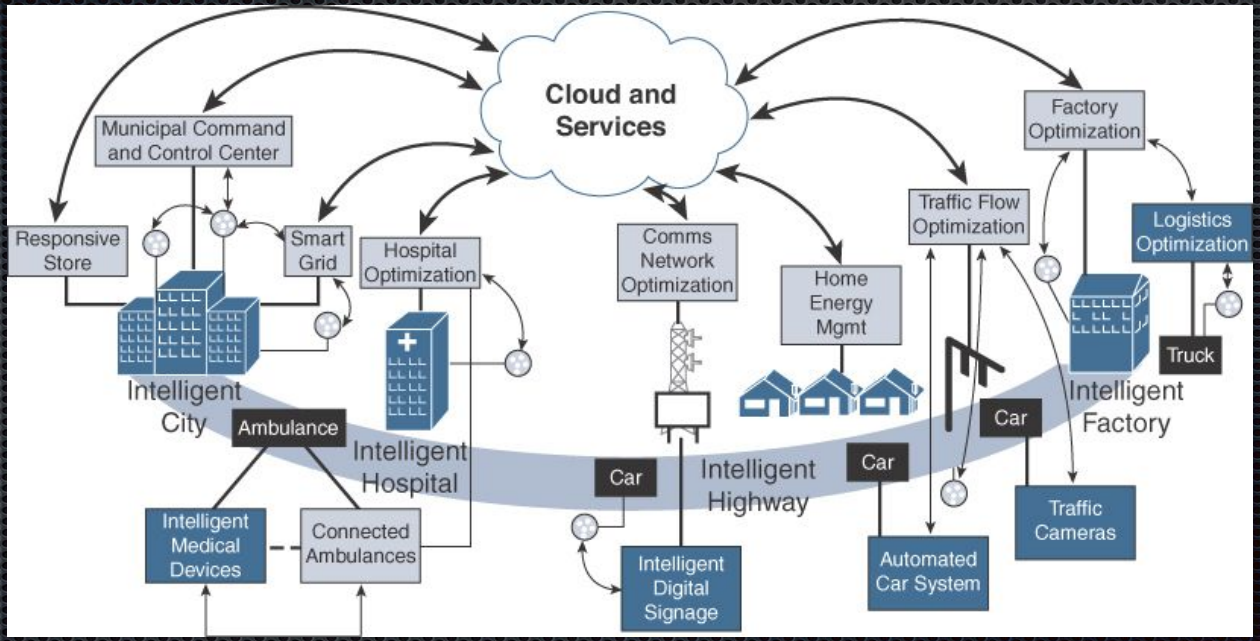


https://www.dataversity.net/reasons-to-consider-a-nosql-solution-for-internet-of-things-data/#

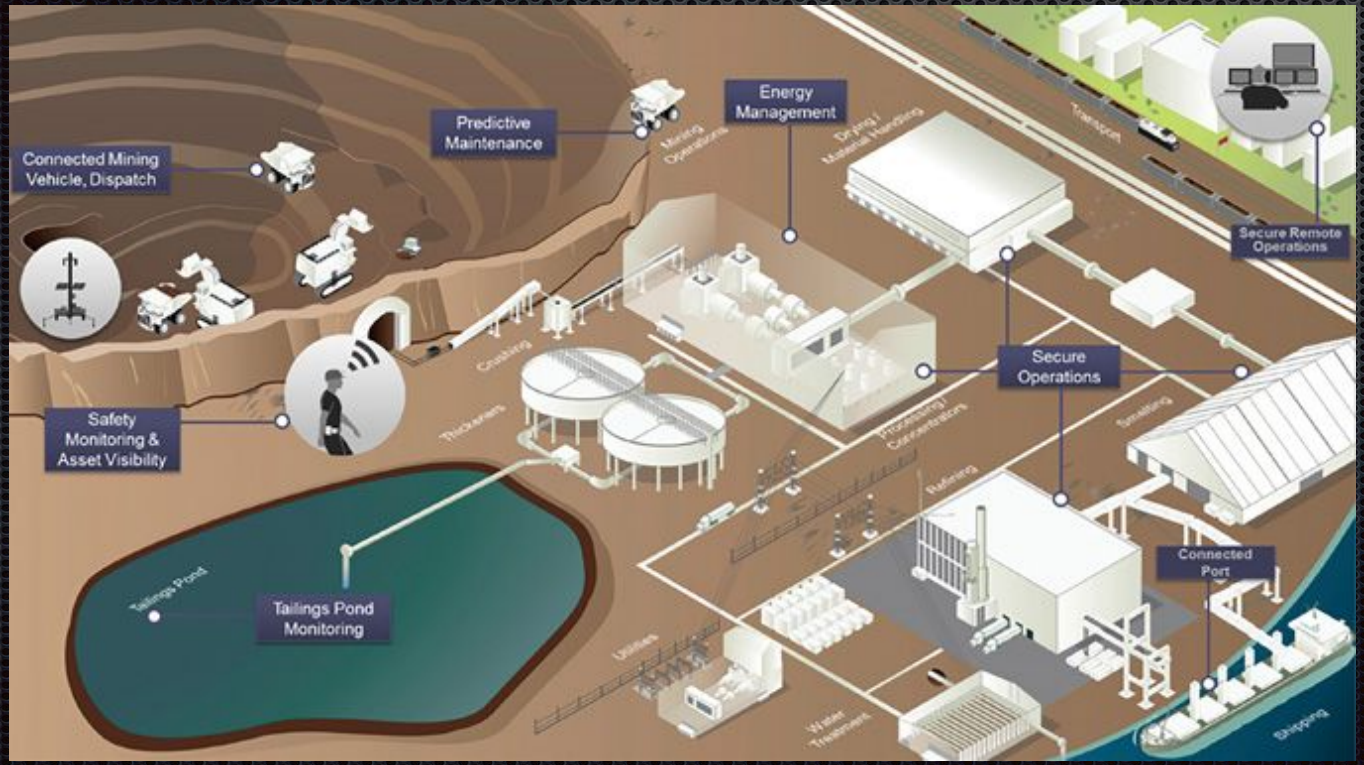
Security concerns...

- we're familiar with classic IT concerns...but wait...there's more
- "Operational Technology"
  - long-lived
  - initial design/maintenance

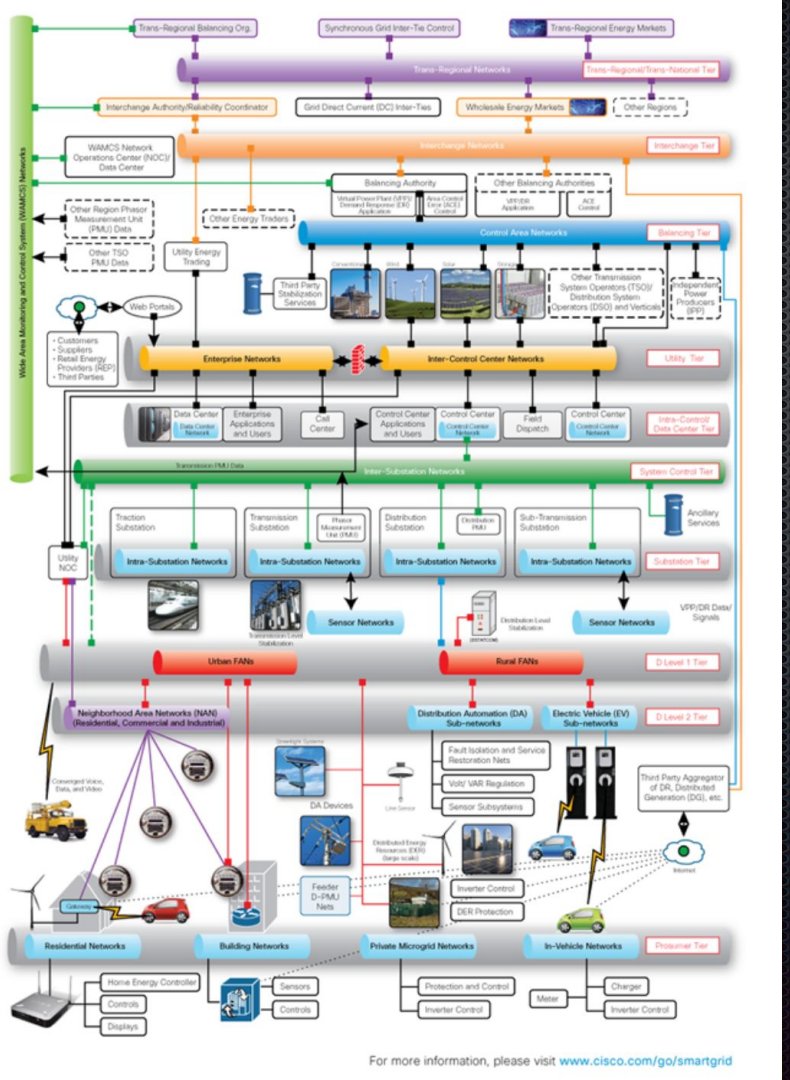
Smart City Example

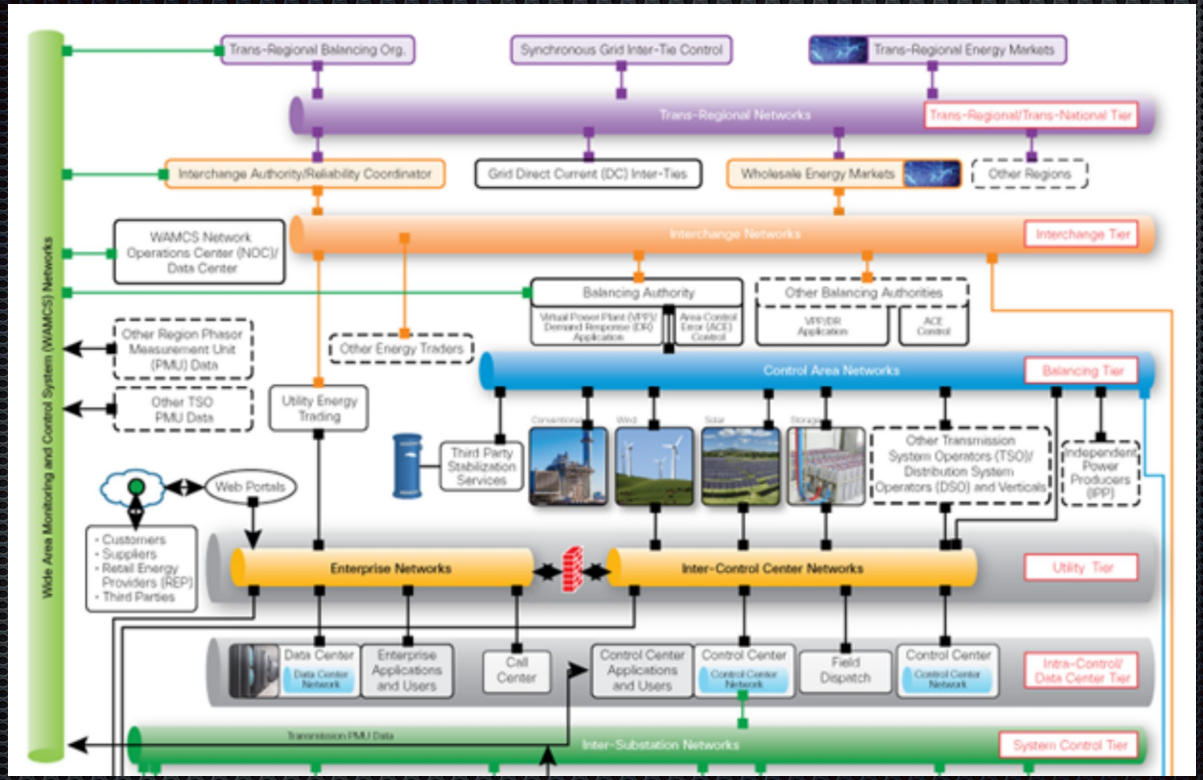


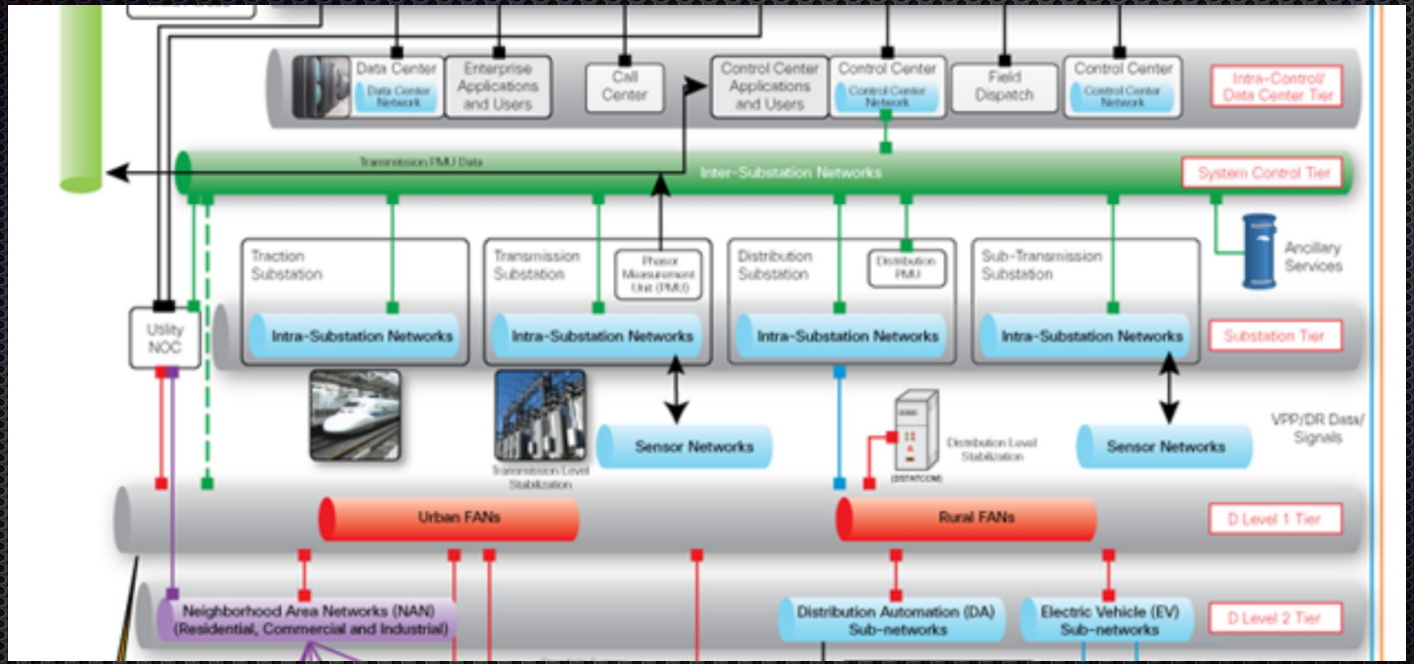
Mining Example

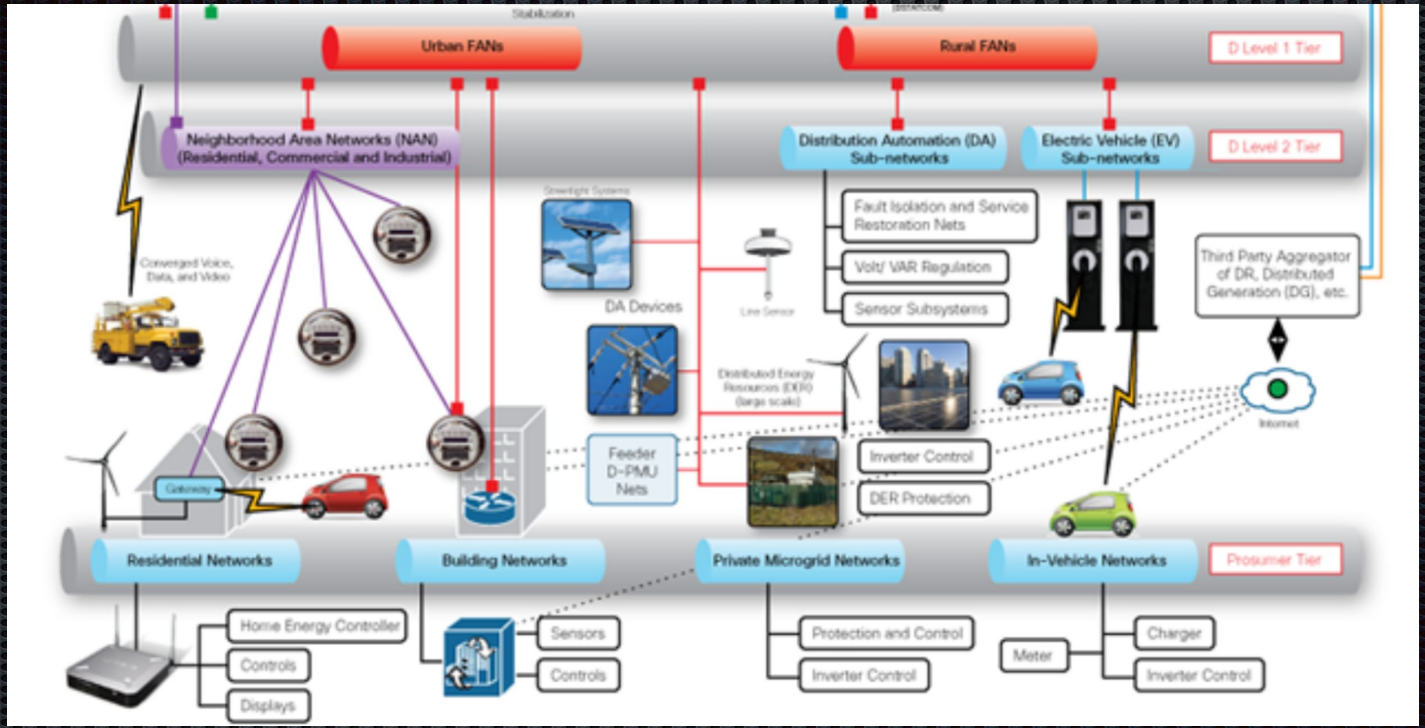


SmartGrid Example

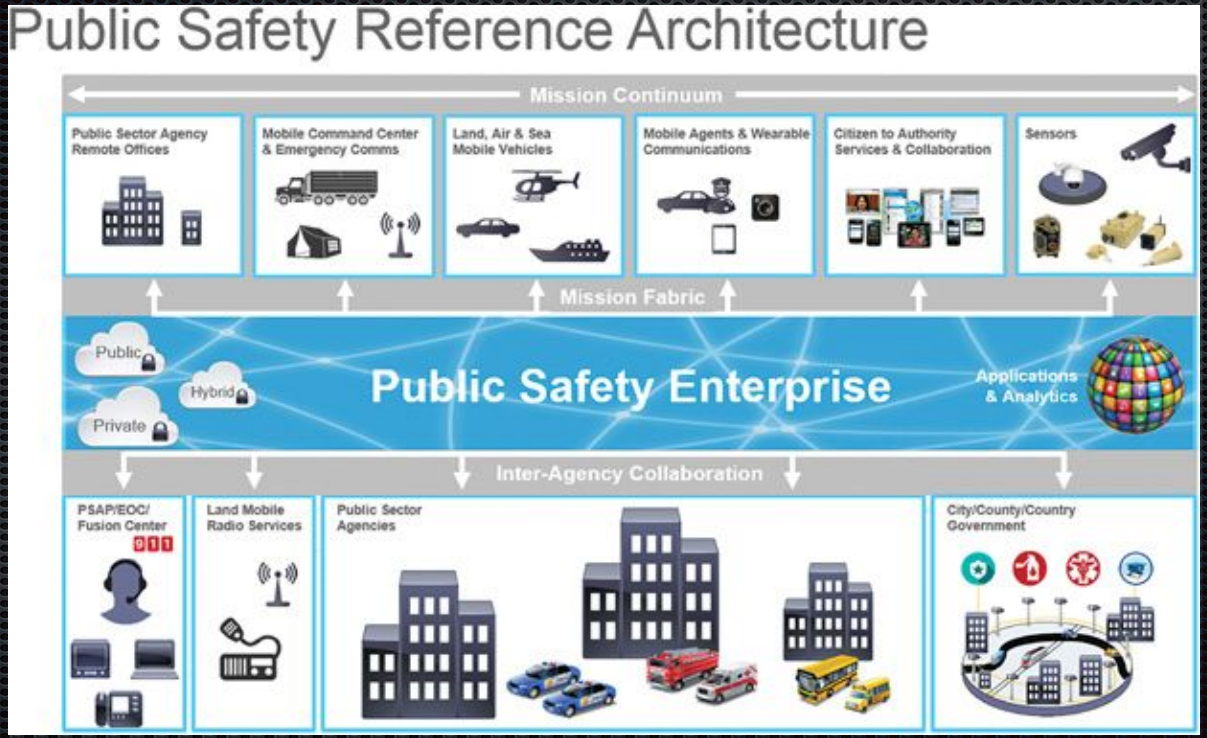



Wide Area Monitoring and Control System (WAMCS) Networks
Trans-Regional Balancing Org.
Synchronous Grid Inter-Tie Control
Trans-Regional Energy Markets
Trans-Regional Networks
Trans-Regional/Trans-National Tier
Interchange Authority/Reliability Coordinator
Grid Direct Current (DC) Inter-Ties
Wholesale Energy Markets
Other Regions
WAMCS Network Operations Center (NOC)/ Data Center
Interchange Networks
Interchange Tier
Balancing Authority
Other Balancing Authorities
Virtual Power Plant (VPP)/ Demand Response (DR) Application
Area Control Error (ACE) Control
VPP/DR Application
ACE Control
Other Region Phasor Measurement Unit (PMU) Data
Other Energy Traders
Other TSO PMU Data
Utility Energy Trading
Control Area Networks
Balancing Tier
Conventional
Wind
Solar
Storage
Third Party Stabilization Services
Other Transmission System Operators (TSO)/ Distribution System Operators (DSO) and Verticals
Independent Power Producers (IPP)
Web Portals
• Customers
• Suppliers
• Retail Energy Providers (REP)
• Third Parties
Enterprise Networks
Inter-Control Center Networks
Utility Tier
Data Center
Data Center Network
Enterprise Applications and Users
Call Center
Control Center Applications and Users
Control Center
Control Center Network
Field Dispatch
Control Center
Control Center Network
Intra-Control/ Data Center Tier
Transmission PMU Data
Inter-Substation Networks
System Control Tier


Data Center
Data Center Network
Enterprise Applications and Users
Call Center
Control Center Applications and Users
Control Center
Control Center Network
Field Dispatch
Control Center
Control Center Network
Intra-Control/ Data Center Tier
Transmission PMU Data
Inter-Substation Networks
System Control Tier
Traction Substation
Transmission Substation
Phasor Measurement Unit (PMU)
Distribution Substation
Distribution PMU
Sub-Transmission Substation
Ancillary Services
Utility NOC
Intra-Substation Networks
Intra-Substation Networks
Intra-Substation Networks
Intra-Substation Networks
Substation Tier
Transmission Level Stabilization
Sensor Networks
Distribution Level Stabilization
(DSTATCOM)
Sensor Networks
VPP/DR Data/ Signals
Urban FANs
Rural FANs
D Level 1 Tier
Neighborhood Area Networks (NAN) (Residential, Commercial and Industrial)
Distribution Automation (DA) Sub-networks
Electric Vehicle (EV) Sub-networks
D Level 2 Tier


Urban FANs
Rural FANs
D Level 1 Tier
Neighborhood Area Networks (NAN)
(Residential, Commercial and Industrial)
Distribution Automation (DA)
Sub-networks
Electric Vehicle (EV)
Sub-networks
D Level 2 Tier
Fault Isolation and Service
Restoration Nets
Volt/ VAR Regulation
Sensor Subsystems
Converged Voice,
Data, and Video
DA Devices
Line Sensor
Third Party Aggregator
of DR, Distributed
Generation (DG), etc.
Distributed Energy
Resources (DER)
(large scale)
Internet
Feeder
D-PMU
Nets
Inverter Control
DER Protection
Gateway
Residential Networks
Building Networks
Private Microgrid Networks
In-Vehicle Networks
Prosumer Tier
Home Energy Controller
Controls
Displays
Sensors
Controls
Protection and Control
Inverter Control
Meter
Charger
Inverter Control

Public Safety Example



Healthcare IoT - Simplified

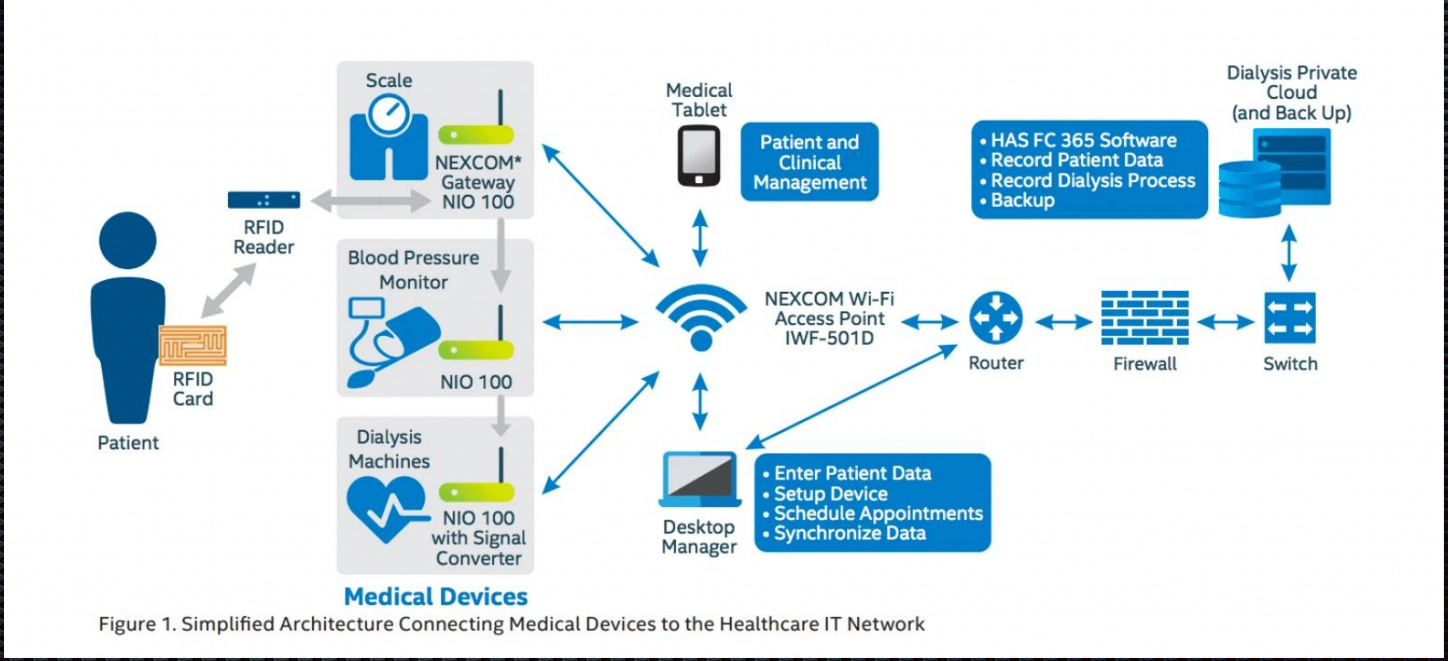
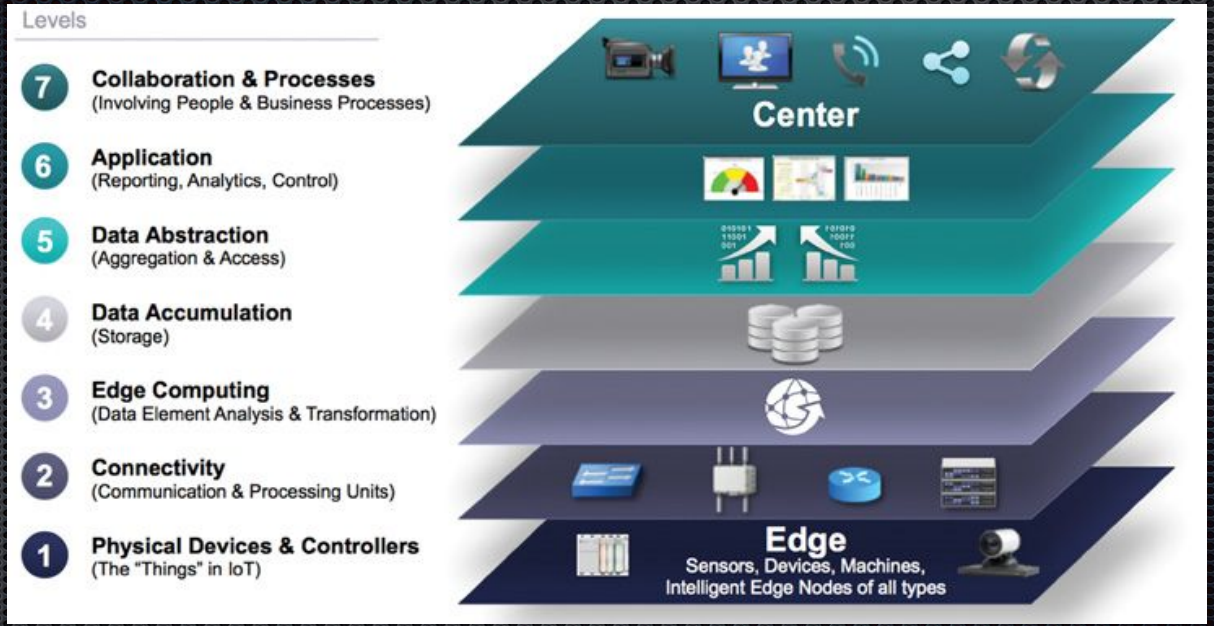


Figure 1. Simplified Architecture Connecting Medical Devices to the Healthcare IT Network

In summary...



Thoughts?

The Industry Use Cases of Fog/Edge Computing

Manufacturing - From creating semiconductors to the assembly of giant industrial machines, edge intelligence enhances manufacturing yields and efficiency using real-time monitoring and diagnostics, machine learning, and operations optimization. The immediacy of edge intelligence enables automated feedback loops in the manufacturing process as well as predictive maintenance for maximizing the uptime and lifespan of equipment and assembly lines.

Oil and gas extraction are high-stakes technology-driven operations that depend on real-time onsite intelligence to provide proactive monitoring and protection against equipment failure and environmental damage. Because these operations are very remote and lack reliable high speed access to centralized data centers, edge intelligence provides onsite delivery of advanced analytics and enables real-time responses required to ensure maximum production and safety.

The Industry Use Cases of Fog/Edge Computing

Mining faces extreme environmental conditions in very remote locations with little or no access to the Internet. As a result mining operations are relying more and more on edge intelligence for real-time, onsite monitoring and diagnostics, alarm management, and predictive maintenance to maximize safety, operational efficiency, and to minimize costs and downtime.

The Industry Use Cases of Fog/Edge Computing

Transportation - As part of the rise in the Industrial Internet, trains and tracks, buses, aircraft, and ships are being equipped with a new generation of instruments and sensors generating petabytes of data that will require additional intelligence for analysis and real-time response. Edge intelligence can process this data locally to enable real-time asset monitoring and management to minimize operational risk and downtime. It can also be used to monitor and control engine idle times to reduce emissions, conserve fuel and maximize profits.

Power and Water - The unexpected failure of an electrical power plant can create substantial disruption to the downstream power grid. The same holds true when water distribution equipment and pumps fail without warning. To avoid this, edge intelligence enables the proactive benefits of predictive maintenance and real-time responsiveness. It also enables ingestion and analysis of sensor data closer to the source rather than the cloud to reduce latency and bandwidth costs.

The Industry Use Cases of Fog/Edge Computing

Renewable Energy - New solar, wind, and hydro are very promising sources of clean energy. However constantly changing weather conditions present major challenges for both predicting and delivering a reliable supply of electricity to the power grid. Edge intelligence enables real-time adjustments to maximize power generation as well as advanced analytics for accurate energy forecasting and delivery.

The Industry Use Cases of Fog/Edge Computing

Healthcare - In the healthcare industry, new diagnostic equipment, patient monitoring tools, and operational technologies are delivering unprecedented levels of patient care but also huge amounts highly sensitive patient data. By processing and analyzing more data at the source, medical facilities can optimize supply chain operations and enhance patient services and privacy at a much lower cost.

Retail - To compete with online shopping, retailers must lower costs while creating enhanced customer experiences and levels of service that online stores cannot provide. Edge intelligence can enrich the user experience by delivering real-time channel personalization and supply chain optimization. It also enables newer technologies such as facial recognition to deliver even higher levels of personalization and security.

The Industry Use Cases of Fog/Edge Computing

Smart Buildings - Among the many benefits of smart building technology are lower energy consumption, better security, increased occupant comfort and safety, and better utilization of building assets and services. Rather than sending massive amounts of building data to the cloud for analysis, smart buildings can use edge intelligence for more responsive automation while reducing bandwidth costs and latency.

Drones/Flying Robots/Unmanned Aerial Vehicles (UAVs) for surveillance and instant delivery – Edge computing facilitates the monitoring, measurement and management of drones.

The Industry Use Cases of Fog/Edge Computing

Connected Vehicles - Connected vehicle technology adds an entirely new dimension to transportation by extending vehicle operations and controls beyond the driver to include external networks and systems. Edge intelligence and fog computing will enable distributed roadside services such as traffic regulation, vehicle speed management, toll collection, parking assistance, and more.

In **automobile manufacturing**, real-time analytics can be performed on sensor streams from the engine and other parts, alerting the driver to potential imminent failure or to the need for preventive maintenance. Such information can also be transmitted to the cloud or EDW for integration into a database maintained by the vehicle manufacturer. Fine-grain analysis of such anomaly data might reveal vehicle model-specific defects that can be corrected in a timely manner.

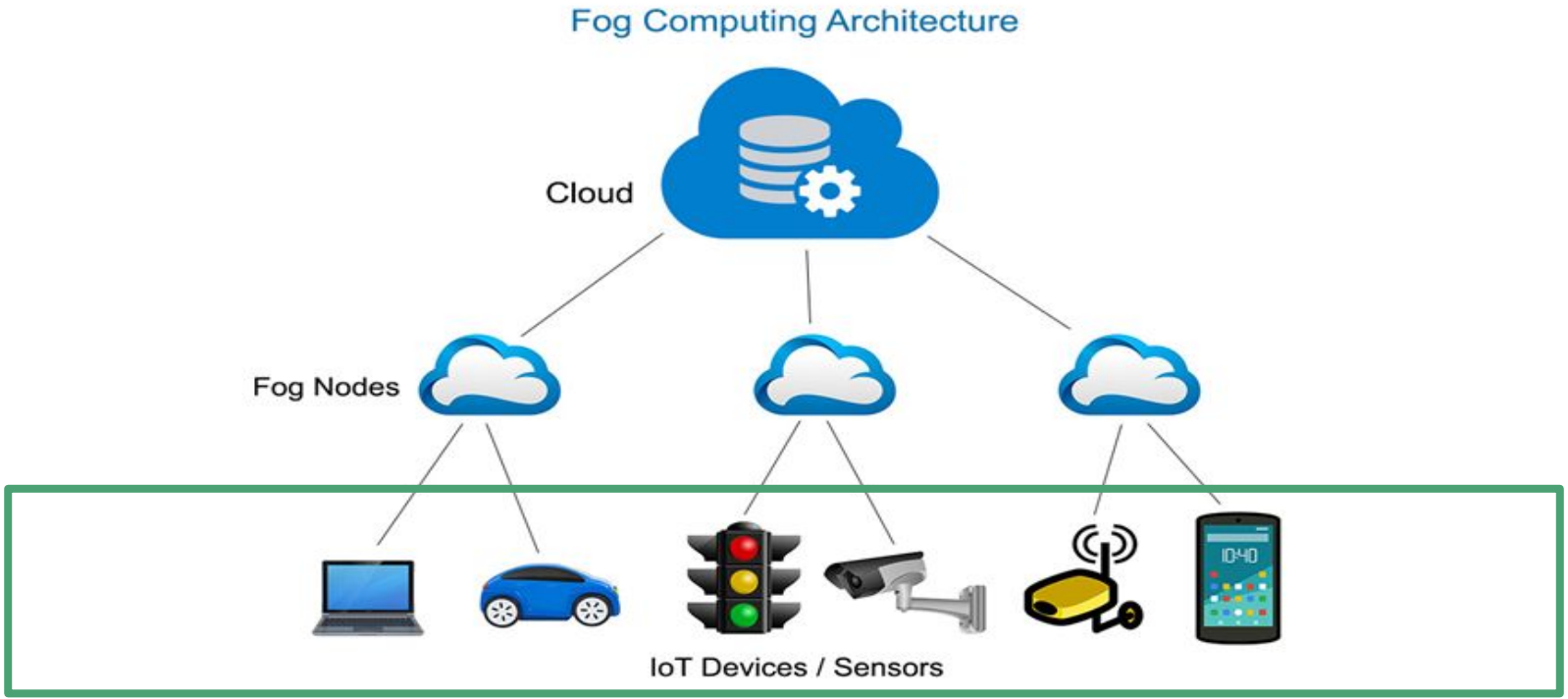
The Industry Use Cases of Fog/Edge Computing

In the **aerospace industry**, the sensors in various parts of the airplane generate huge amount of data on the order of 1 terabyte per 24 hours. Intelligent devices (compared to connected devices) would be of great, and sometimes lifesaving, help as immediate proactive actions based on sensor readings could prevent crucial failures.

The industrial Internet is going to transform the industry by making industrial machines more intelligent and enabling services using real-time data coming from sensors and machines. The intelligent devices will be able to take actions (to optimize processes, improve efficiencies, reduce costs, etc.) based on insights generated from real-time data and analytics.

The Industry Use Cases of Fog/Edge Computing

Smart Cities - Integrating data from a diverse collection of municipal systems (e.g. Street lighting, traffic information, parking, public safety, etc.) for interactive management and community access is a common vision for smart city initiatives. However the sheer amount of data generated requires too much bandwidth and processing for cloud-based systems. Edge intelligence provides a more effective solution that distributes data processing and analytics to the edges where sensors and data sources are located.


Fog Computing Architecture
Cloud
Fog Nodes
IoT Devices / Sensors

Key attributes of edge computing

- Applications that require very low and predictable latency
- Geo-distributed applications
- Fast mobile applications
- Large-scale distributed control systems
- IoT also brings big data *with a twist*: rather than high volume, the number of data sources distributed geographically
  - (and high volume, depending on amount of data sent)

Discussion

Do you think edge/fog computing can completely replace the cloud?

- **Why or why not?**


Networking!

MQTT - Message Queueing Telemetry Transport

https://blog.schneider-electric.com/industrial-software/2016/06/03/basics-mqtt-messaging-protocol-internet-things/

Basics of MQTT

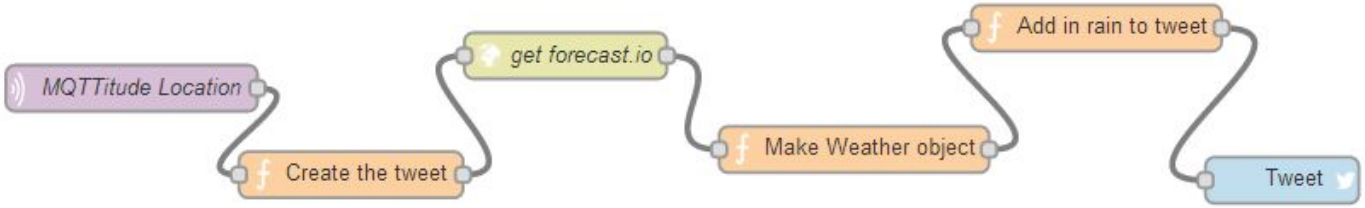
MQTT: "the IoT protocol"

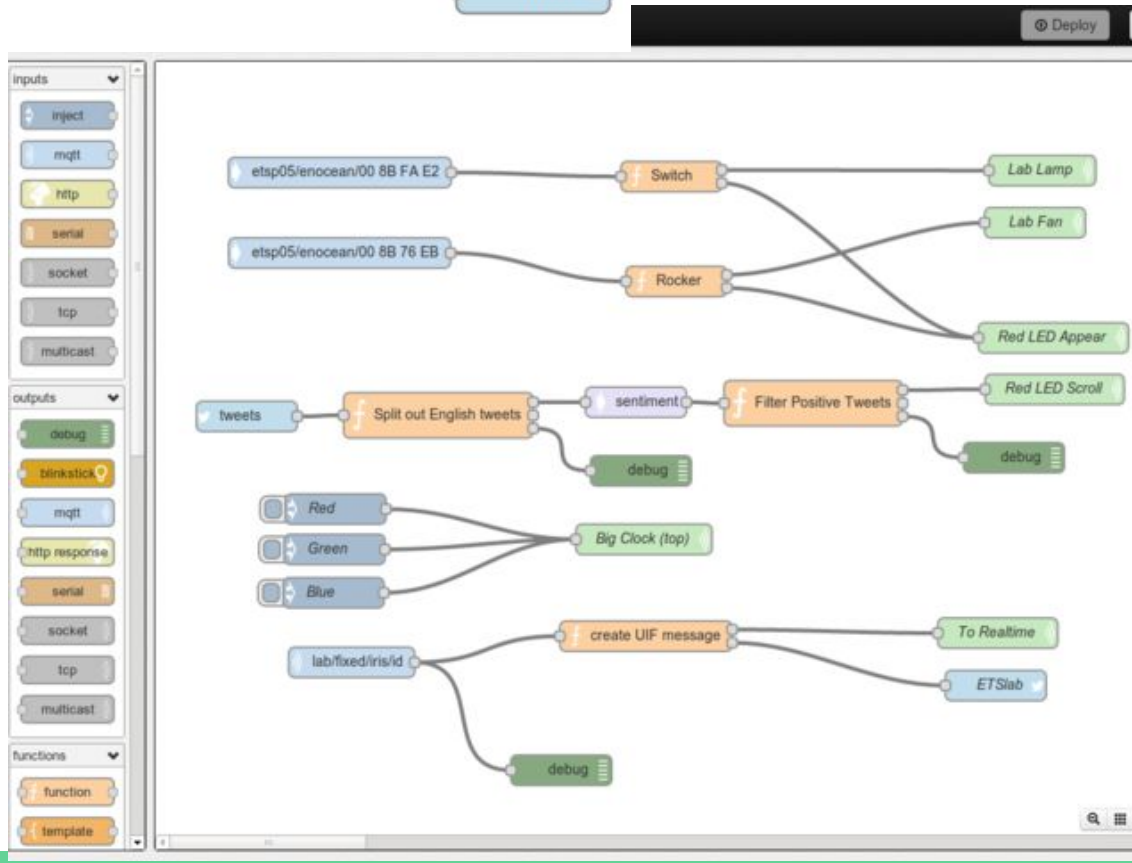
- Open standardized protocol: http://mqtt.org

Easy to use and setup with a wide array of devices and services

- Cloud providers (GCP, AWS, etc.)
- Python bindings
- High-level frameworks (Node Red)

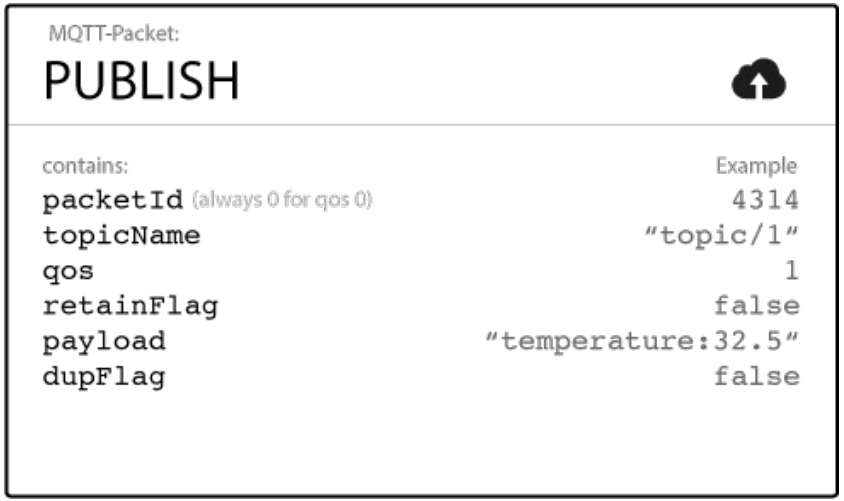
https://mosquitto.org/blog/2013/12/paho-mqtt-python-client/
http://www.steves-internet-guide.com/into-mqtt-python-client/


MQTTitude Location
Create the tweet
get forecast.io
Make Weather object
Add in rain to tweet
Tweet


Deploy
inputs
inject
mqtt
http
serial
socket
tcp
multicast
outputs
debug
blinkstick
mqtt
http response
serial
socket
tcp
multicast
functions
function
template
etsp05/enocean/00 8B FA E2
Switch
Lab Lamp
Lab Fan
etsp05/enocean/00 8B 76 EB
Rocker
Red LED Appear
tweets
Split out English tweets
sentiment
Filter Positive Tweets
Red LED Scroll
debug
debug
Red
Green
Blue
Big Clock (top)
lab/fixed/iris/id
create UIF message
To Realtime
ETSlab
debug

Basics of MQTT

Lightweight
Bandwidth efficient (smaller packet sizes)
Uses publisher / subscriber (pub/sub) model (design pattern)



MQTT-Packet:

SUBSCRIBE

contains:		Example
packetId		4312
qos1	} (list of topic + qos)	1
topic1		"topic/1"
qos2	}	0
topic2		"topic/2"
...		...

MQTT-Packet:

SUBACK

contains:		Example
packetId		4313
returnCode 1	(one returnCode for each topic from SUBSCRIBE, in the same order)	2
returnCode 2		0
...		...

QoS:
0 - best effort (no guarantee)
1 - at least once (possible to be retransmitted)
2 - exactly once (highest reliability - no duplicate messages)

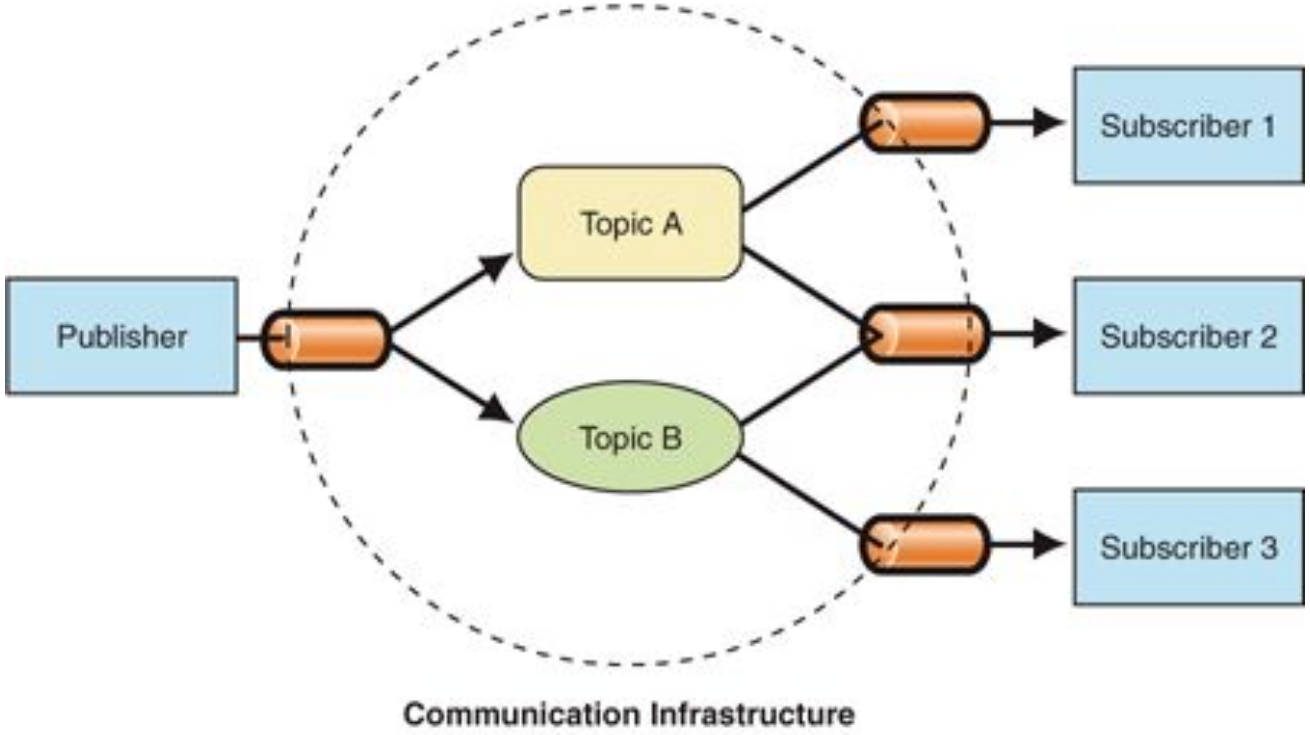
MQTT v.s. HTTP

IBM

Action	HTTP	MQTT
Get single piece of data	302 bytes	69 bytes (<4 times)
Send single piece of data	320 bytes	47 bytes (<6 times)
Get 100 pieces of data	12600 bytes	2445 bytes (<5 times)
Send 100 pieces of data	14100 bytes	2126 bytes (<6 times)

Characteristics	HTTP	MQTT
Style	Document-centric, request/response	Data-centric, publish/subscribe
Verbs	GET/POST/POST/DELETE, complex spec	Pub/Sub/Unsub, simple protocol, easy to learn
Message size	Large message, lots of data in headers	2 bytes in minimum header
Quality of Service	None, requires custom coding in application	3 levels – best-effort, at-least-once, exactly once
Data distribution	No distribution mechanism (1-to-1 only)	Fully supported. 1-to-none, 1-to-1, 1-to-n.

Pub/sub


Publisher
Topic A
Topic B
Subscriber 1
Subscriber 2
Subscriber 3
Communication Infrastructure

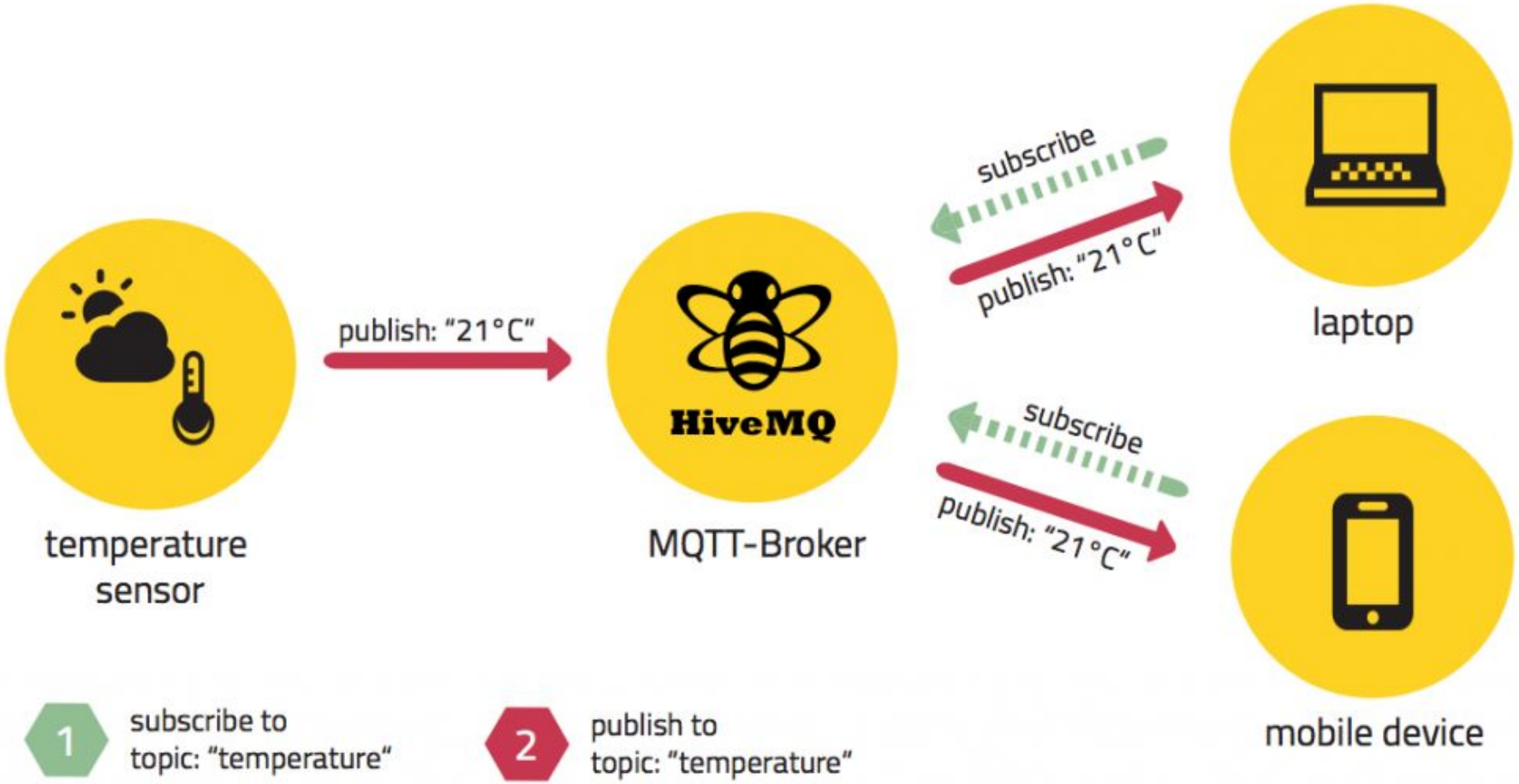
Clients and brokers

Client:

- Publisher or subscriber that connects to a **broker**
- Persistent (maintains connection) or transient (not tracked)

Broker (**central hub**):

- Receiving and filtering messages
- Understanding which clients are 'interested' in data
- Sending messages to subscribed clients
- Authenticating/authorizing clients


temperature sensor
publish: "21°C"
HiveMQ
MQTT-Broker
subscribe
publish: "21°C"
laptop
subscribe
publish: "21°C"
mobile device
1
subscribe to topic: "temperature"
2
publish to topic: "temperature"

Topics

Hierarchical string that filters messages for clients

https://www.youtube.com/watch?v=f5o4tlz2Zzc