Big Data and Internet Thinking

Chentao Wu
Associate Professor
Dept. of Computer Science and Engineering
wuct@cs.sjtu.edu.cn

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

Download lectures

- ftp://public.sjtu.edu.cn
- User: wuct
- Password: wuct123456

- http://www.cs.sjtu.edu.cn/~wuct/bdit/

Schedule

- lec1: Introduction on big data, cloud computing & IoT
- lec2: Parallel processing framework (e.g., MapReduce)
- lec3: Advanced parallel processing techniques (e.g., YARN, Spark)
- lec4: Cloud & Fog/Edge Computing
- lec5: Data reliability & data consistency
- lec6: Distributed file system & objected-based storage
- lec7: Metadata management & NoSQL Database
- lec8: Big Data Analytics

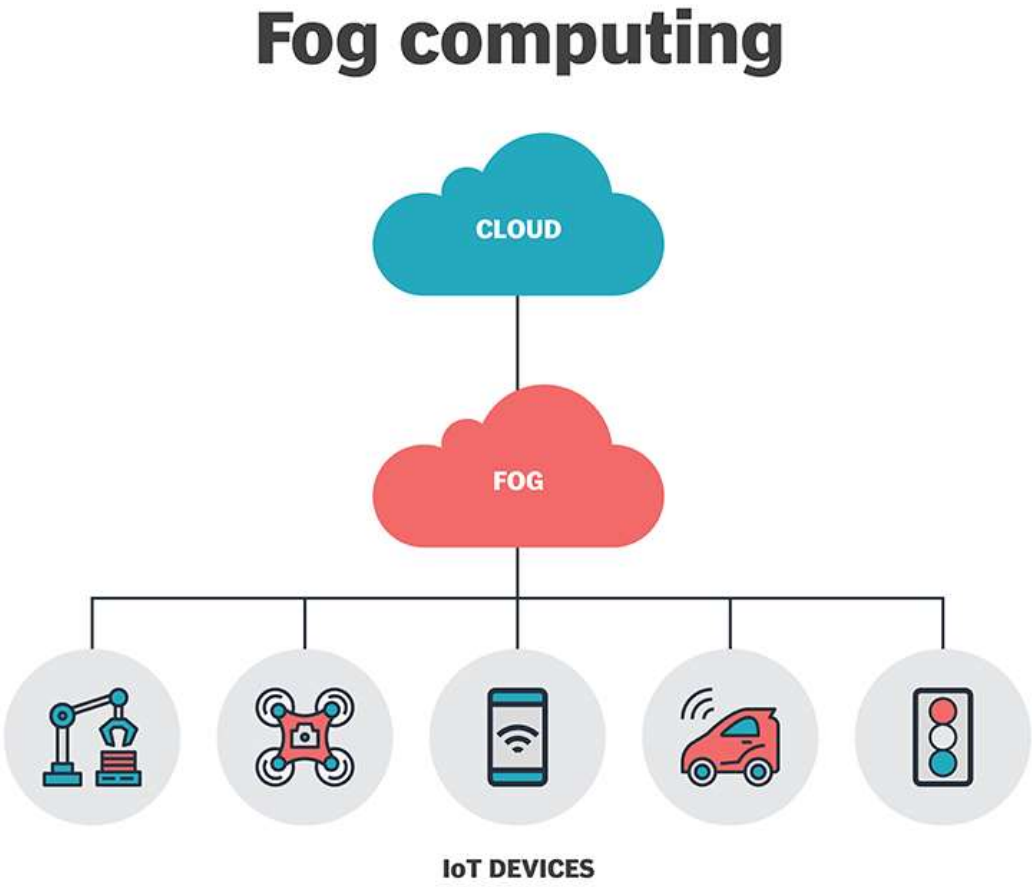
Collaborators

Google 谷歌

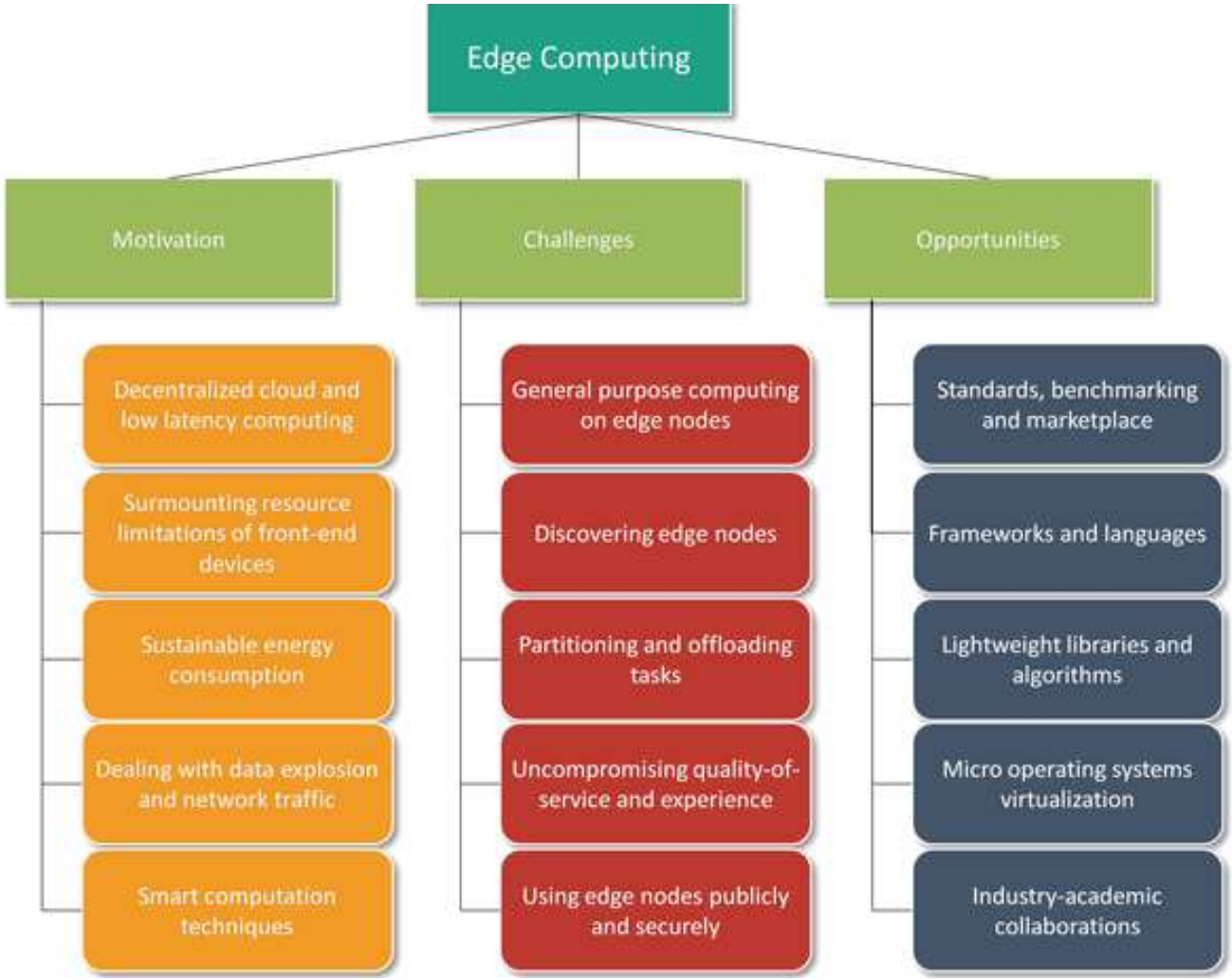
阿里云 aliyun.com

DELL EMC

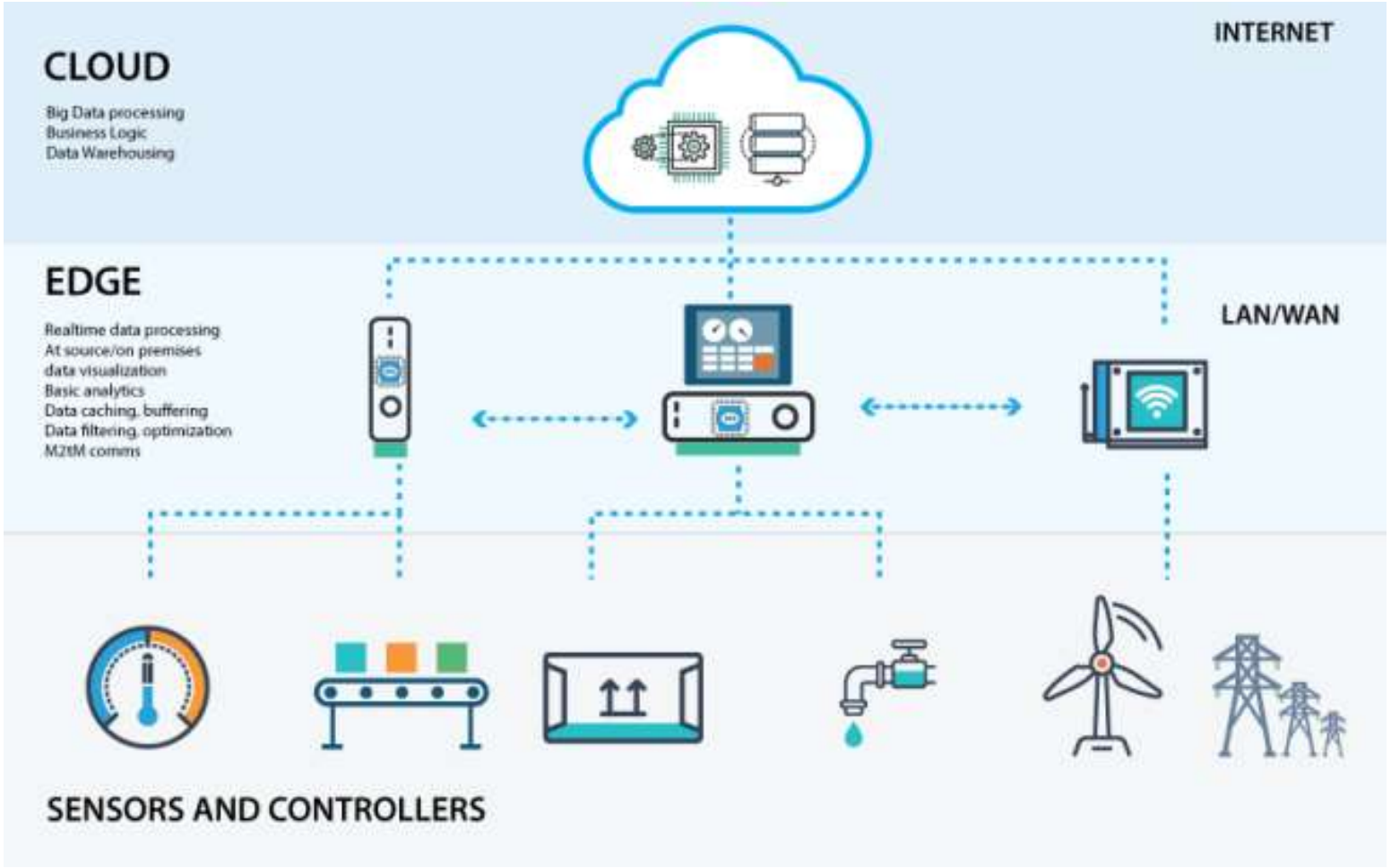
Fog Computing



Challenges

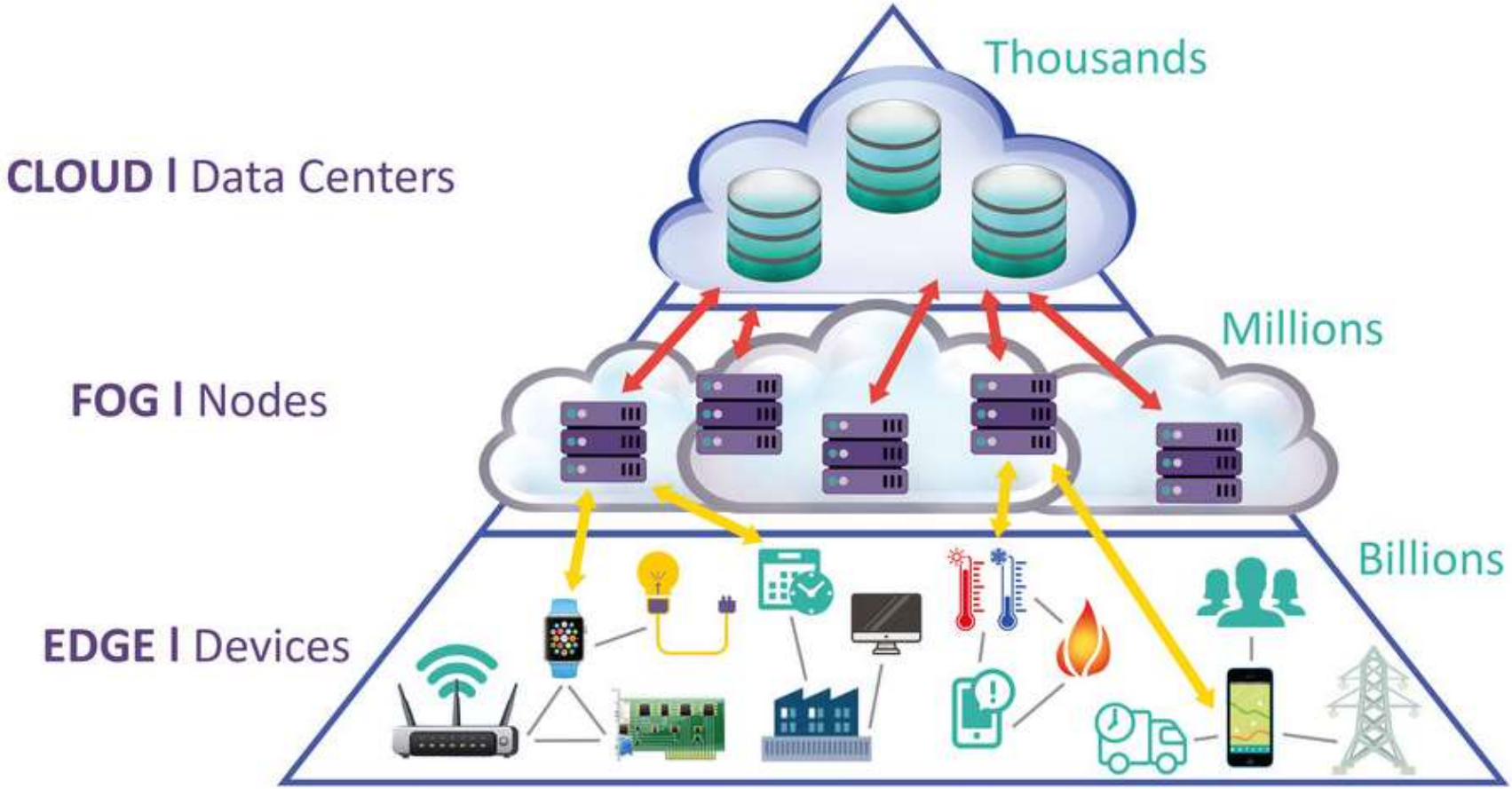

Edge Computing
Motivation
Challenges
Opportunities
Decentralized cloud and low latency computing
Surmounting resource limitations of front-end devices
Sustainable energy consumption
Dealing with data explosion and network traffic
Smart computation techniques
General purpose computing on edge nodes
Discovering edge nodes
Partitioning and offloading tasks
Uncompromising quality-of-service and experience
Using edge nodes publicly and securely
Standards, benchmarking and marketplace
Frameworks and languages
Lightweight libraries and algorithms
Micro operating systems virtualization
Industry-academic collaborations

Cloud-Fog-Edge

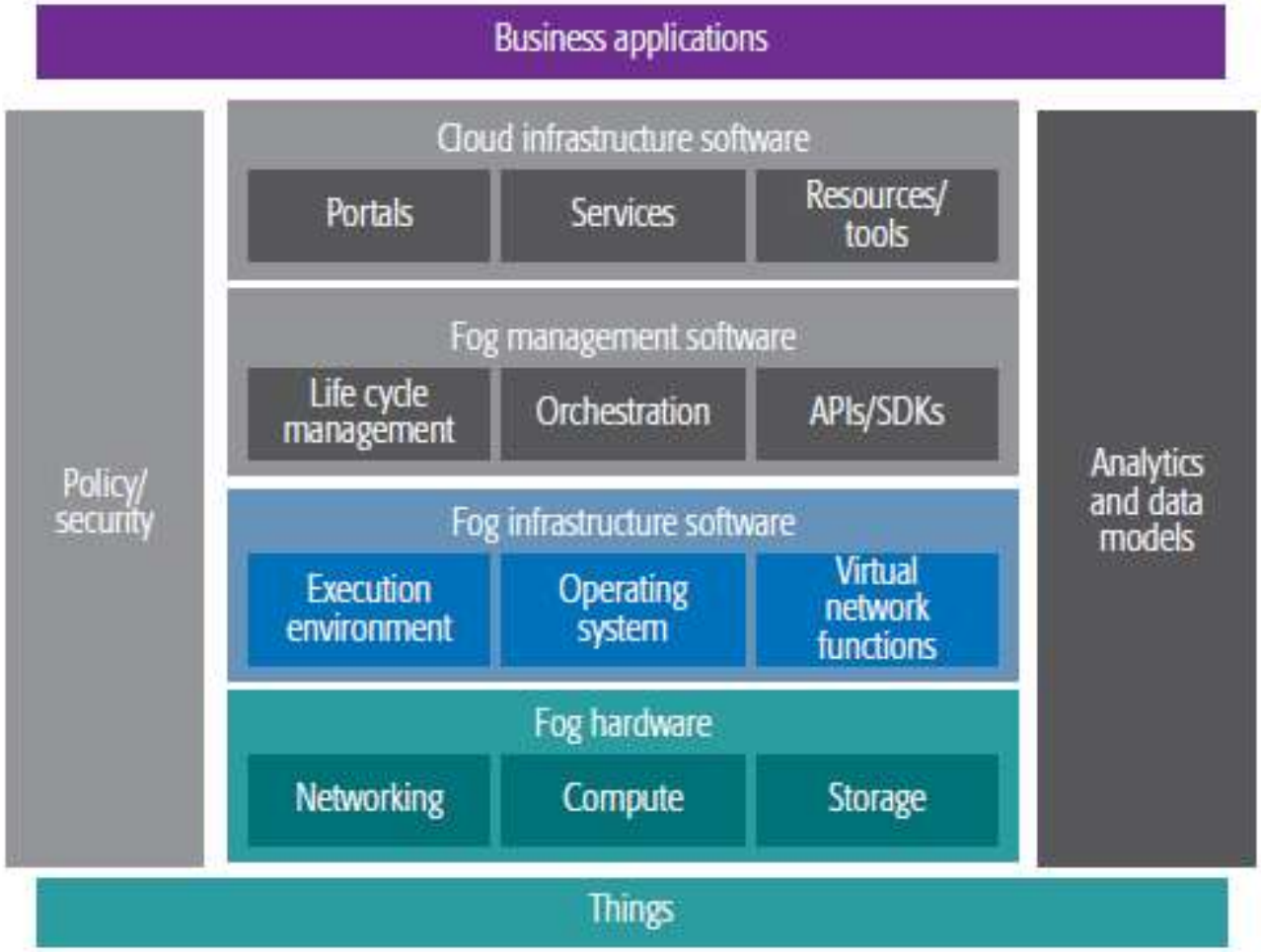


Cloud-Fog-Edge Devices

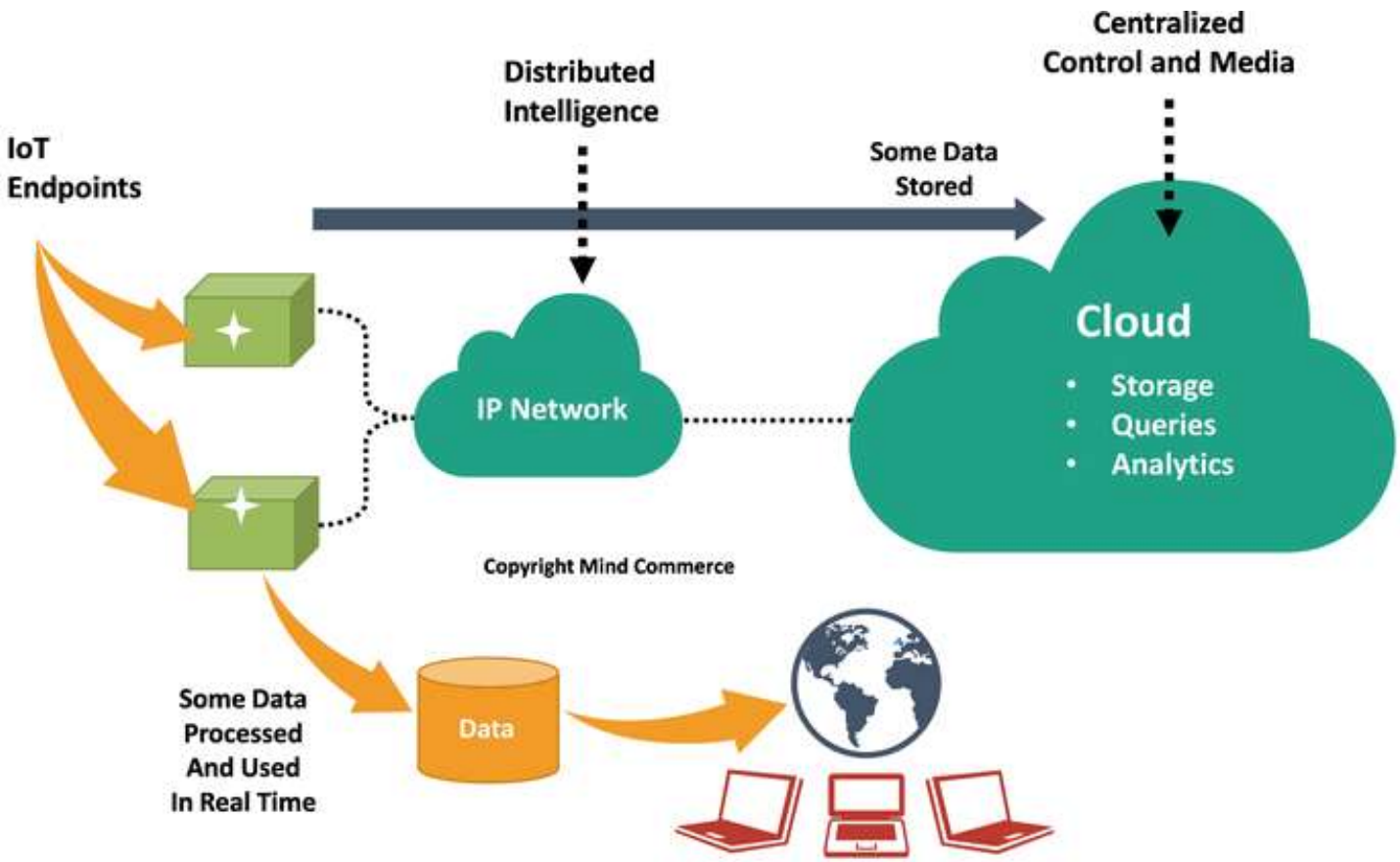
Cloud-Fog-Edge Architecture (1)



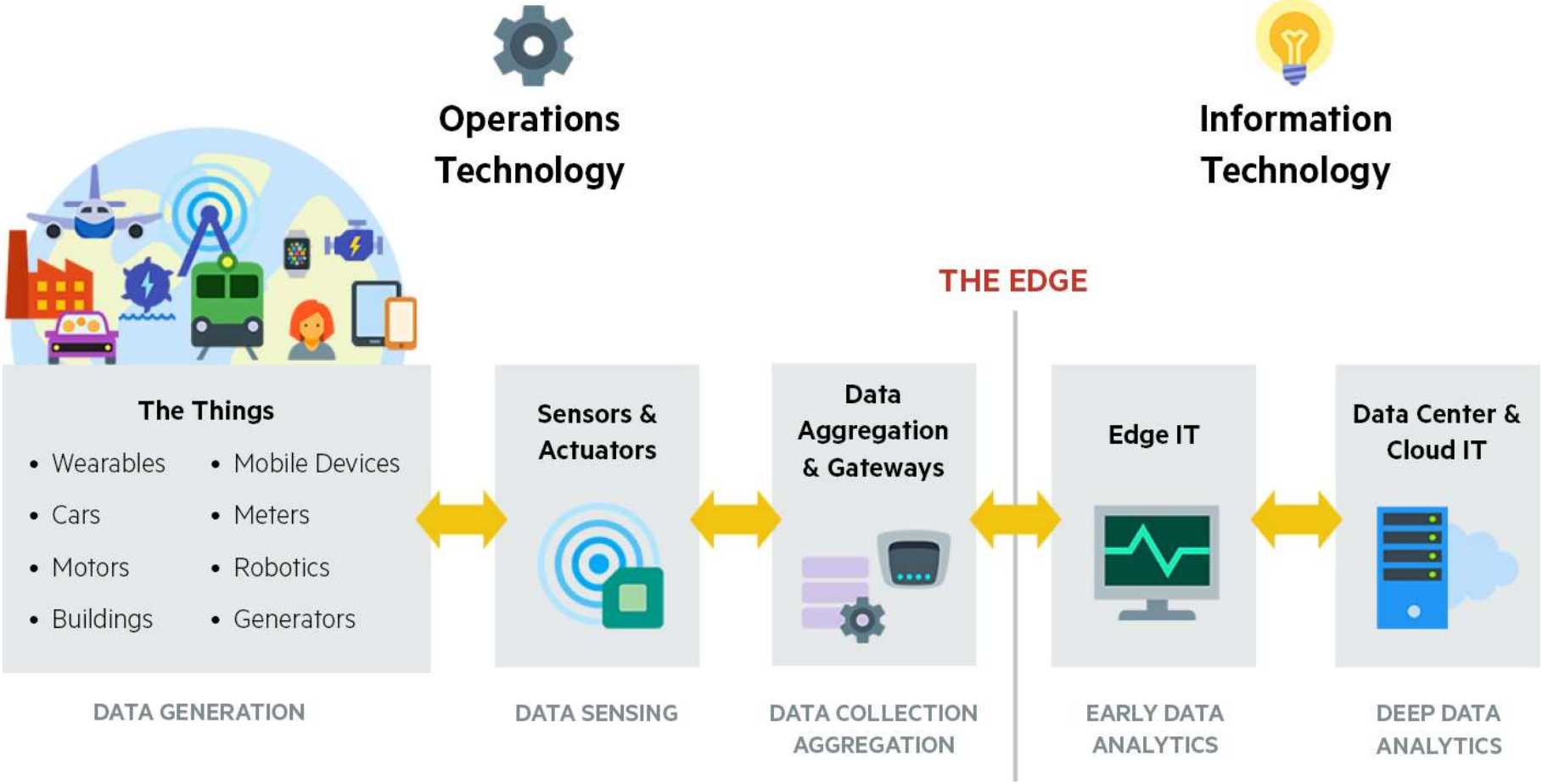
Cloud-Fog-Edge Architecture (2)



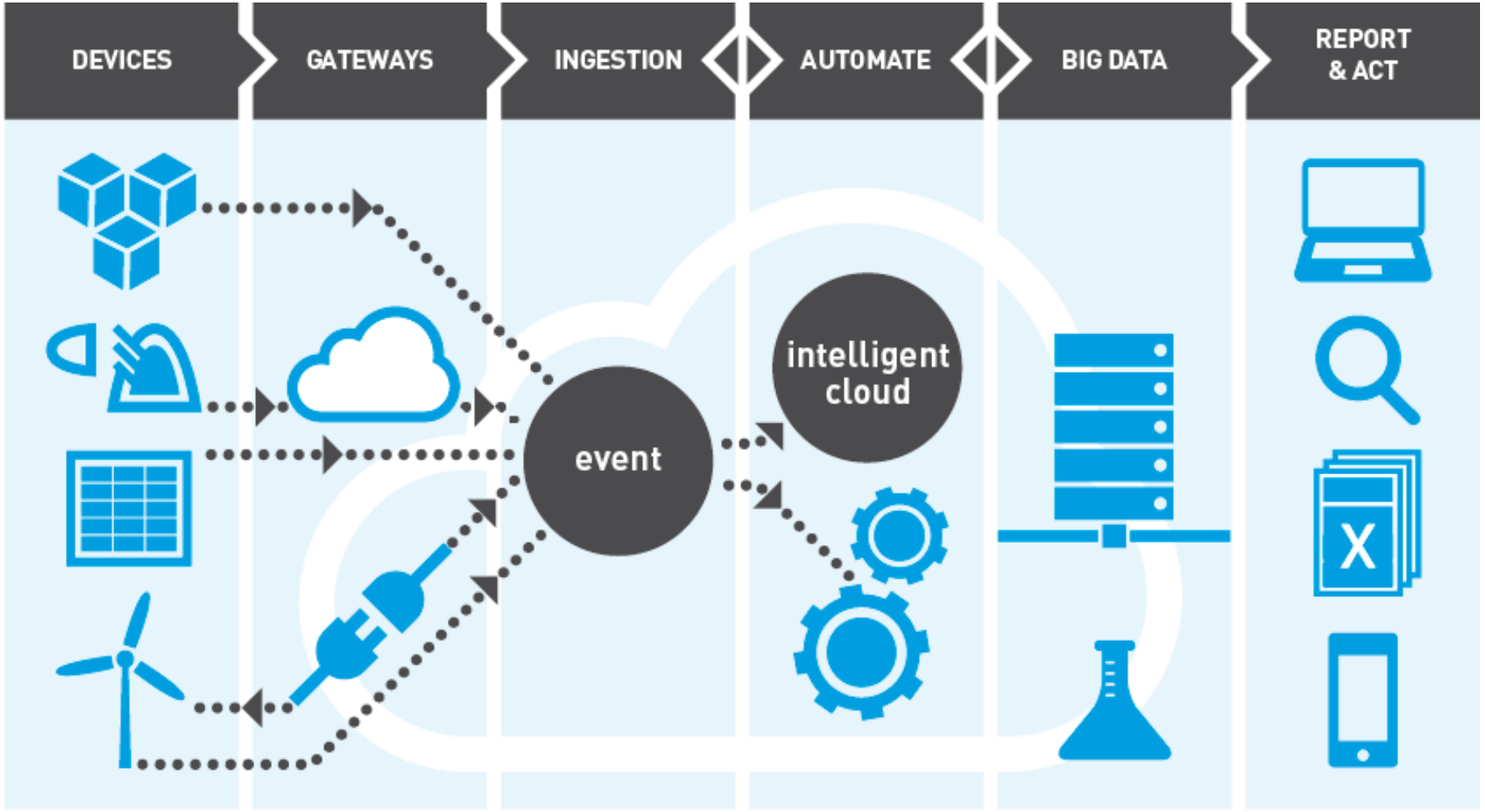
Functions of Cloud-Fog-Edge



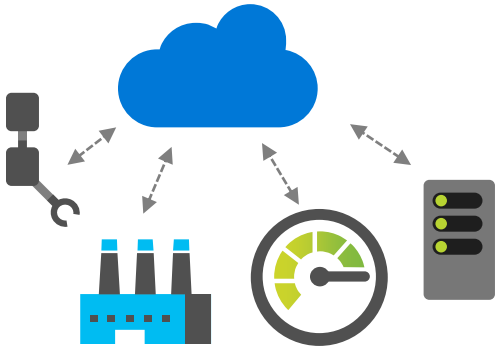
Fog/Edge Computing is the Primary Choice to Handle Real Time Data

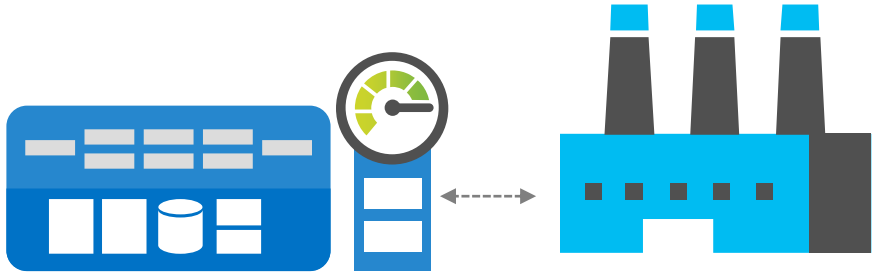


IoT End-to-End Value Chain



IoT in the cloud and on the edge





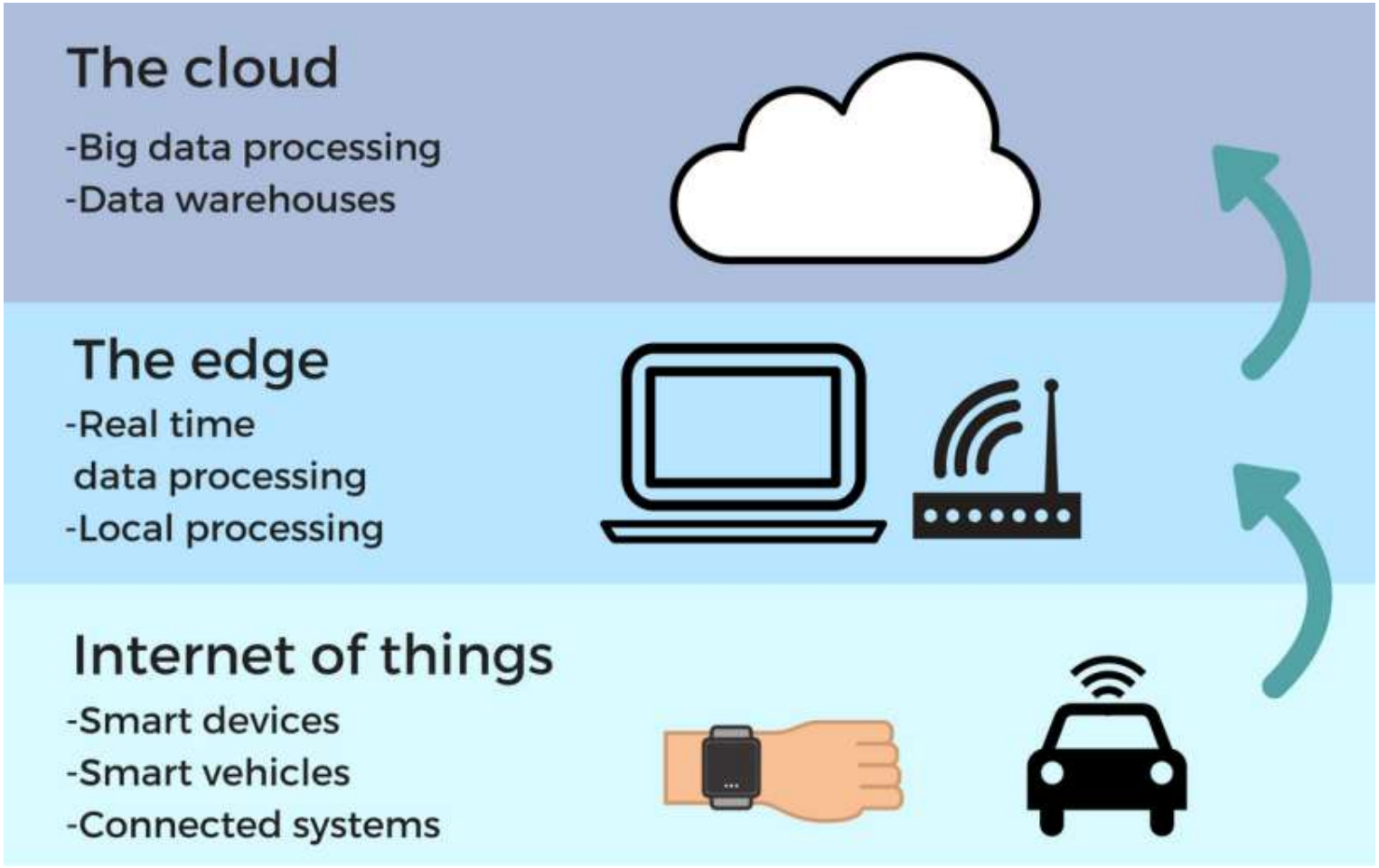
IoT in the Cloud

- Remote monitoring and control
- Merging remote data from across multiple IoT devices
- Near infinite compute and storage to train machine learning and other advanced AI tools

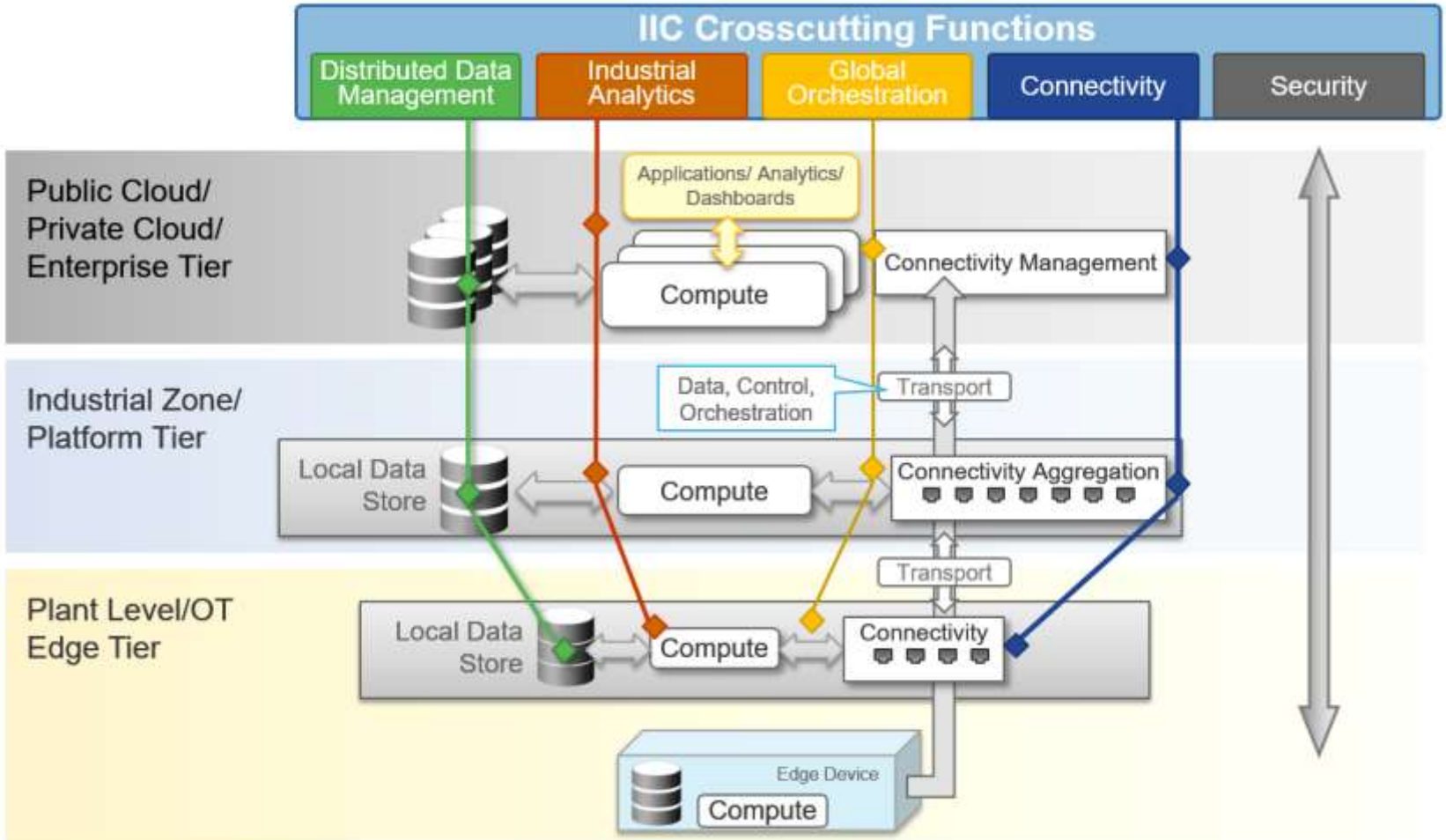
IoT on the Edge

- → Low latency tight control loops require near real-time response
- → Public internet inherently unpredictable
- → Privacy of data and protection of IP

Data Processing in Cloud-Fog-Edge

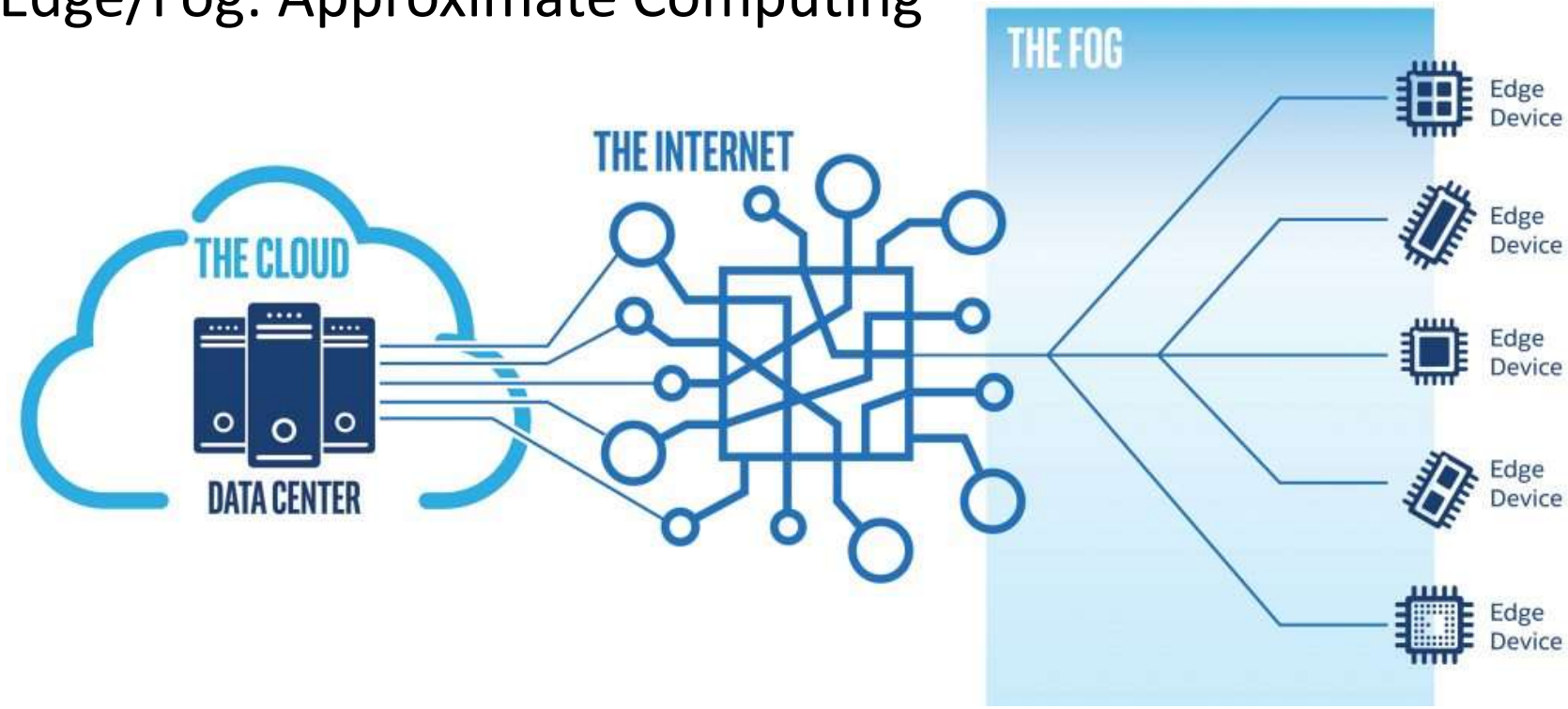


Heterogeneous/Homogeneous Cloud/Fog computing



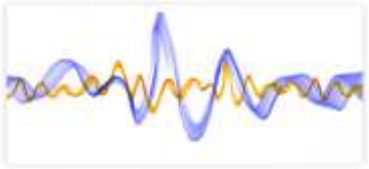
Heterogeneous/Homogeneous Computing Framework

- Cloud: Parallel and Distributed Computing
  - Map-Reduce, Graph Computing, Stream Computing
- Edge/Fog: Approximate Computing



Why Approximate Computing?

image, sound
and video processing



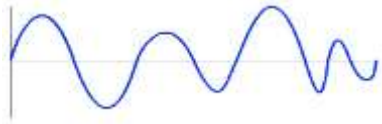
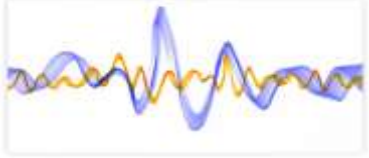
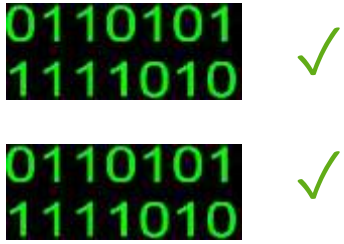


image rendering

sensor data analysis,
computer vision



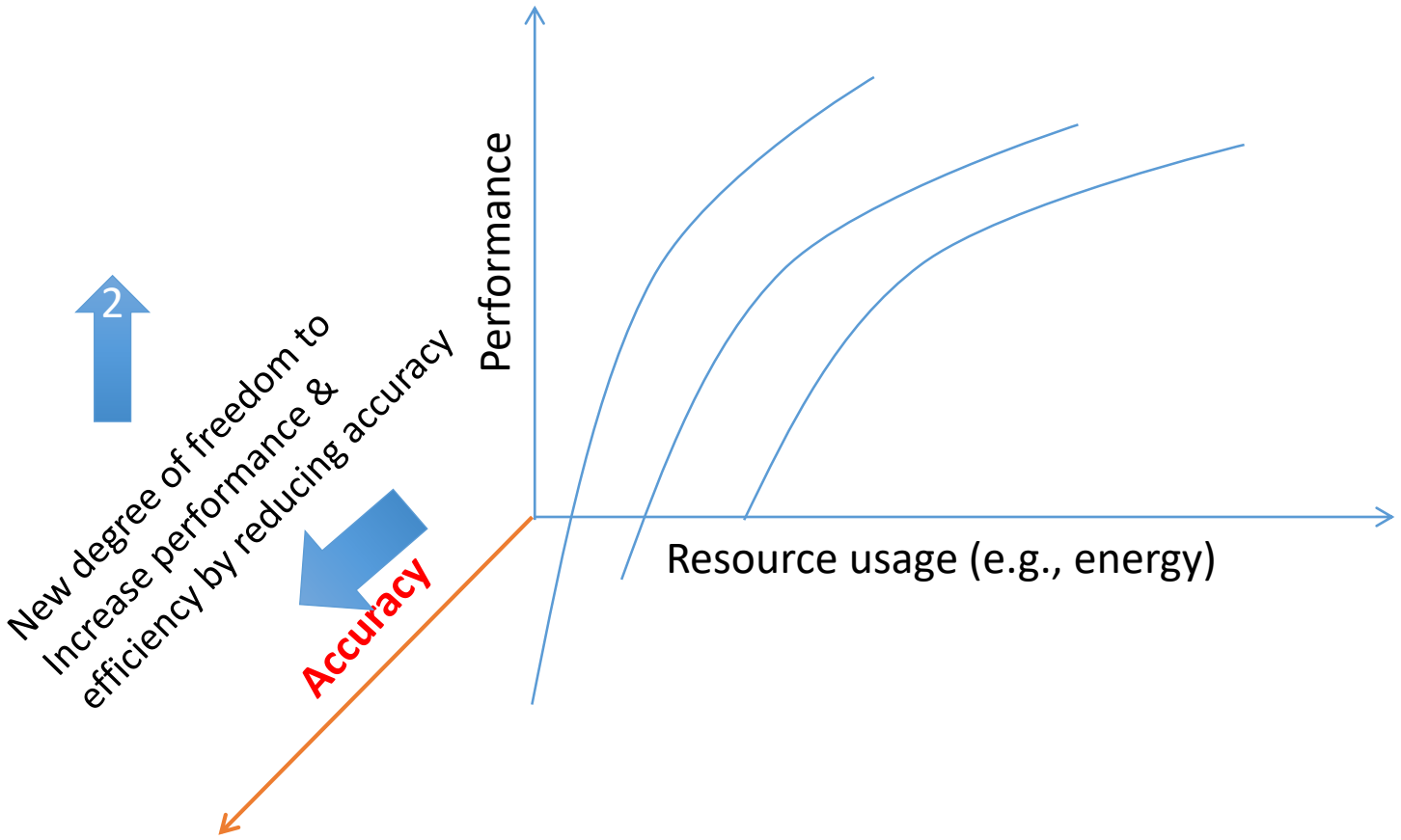
simulations, games,
search, machine learning



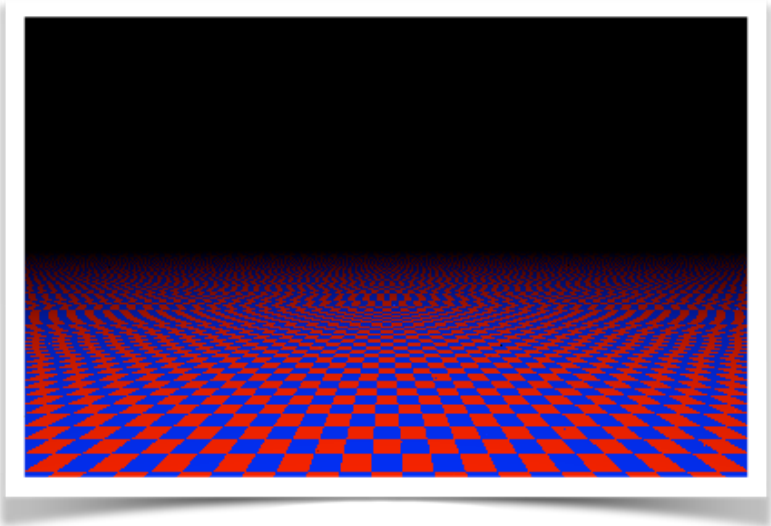
- Inexact/imprecise input data
- Approximate/iterative algorithms
- Loose constraints on output

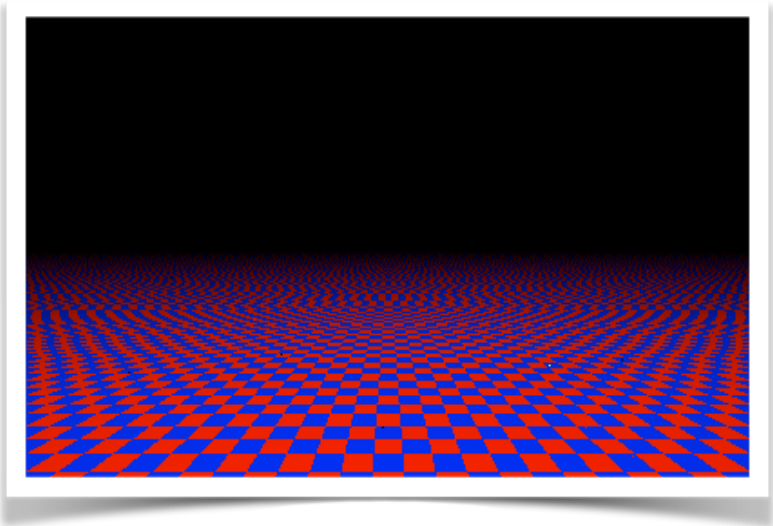
Where a lot of (most?) resources go!

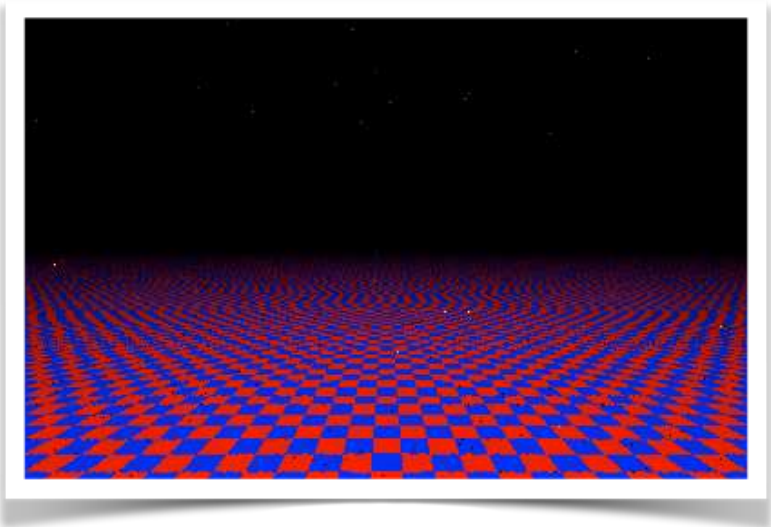
Point of View on Approximate Computing

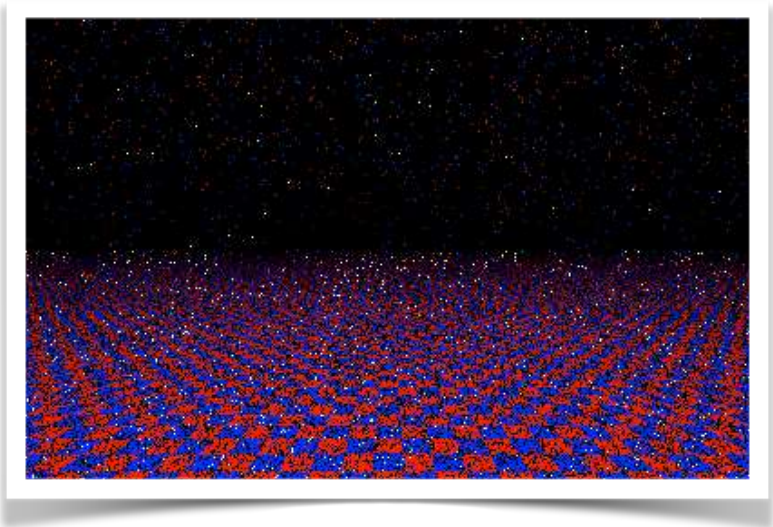


Approximate Computing Example - Images(1)

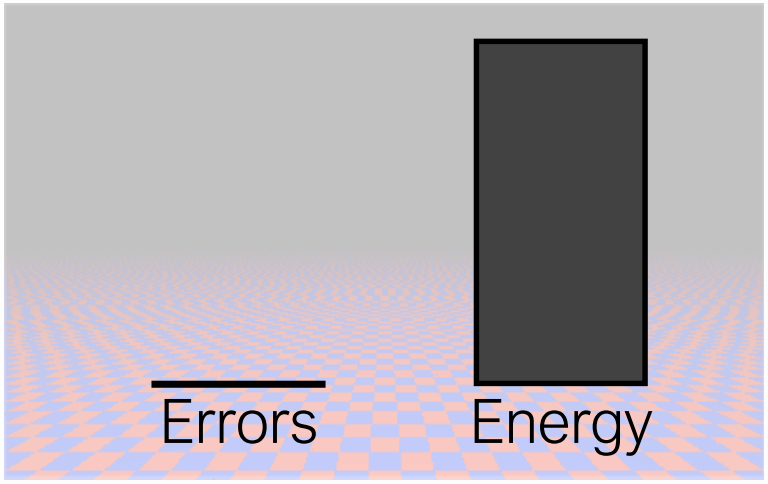


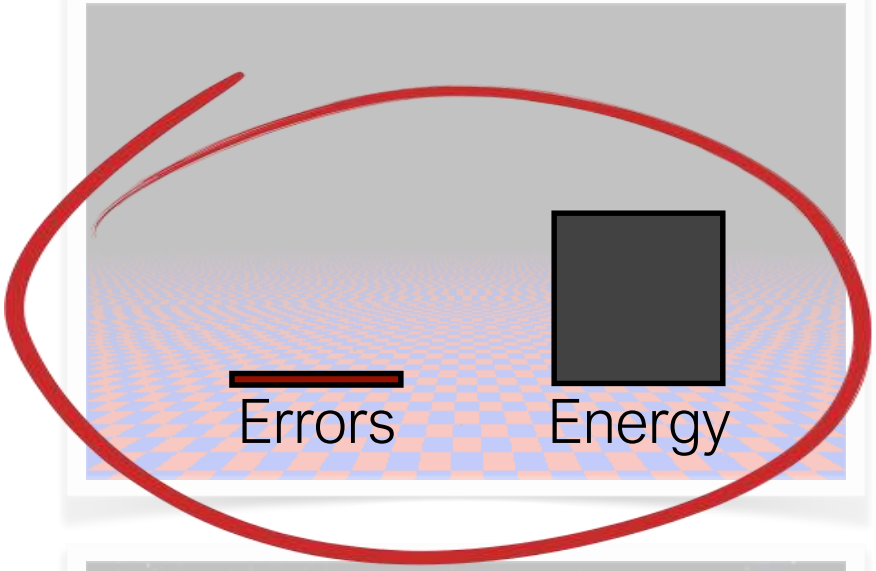


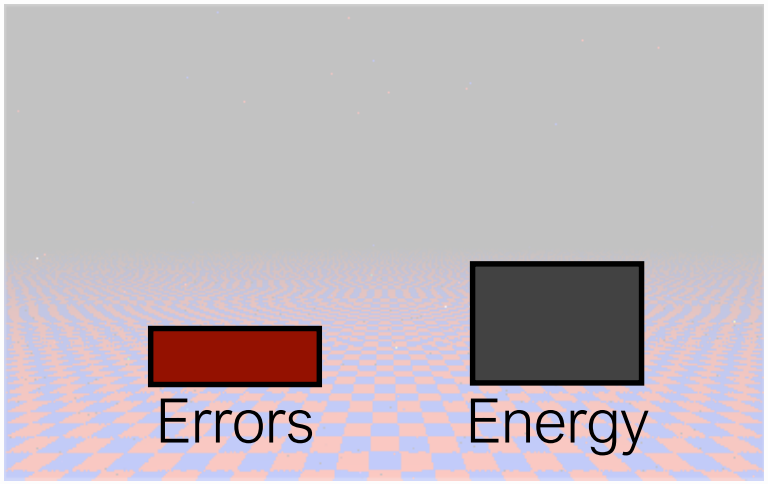


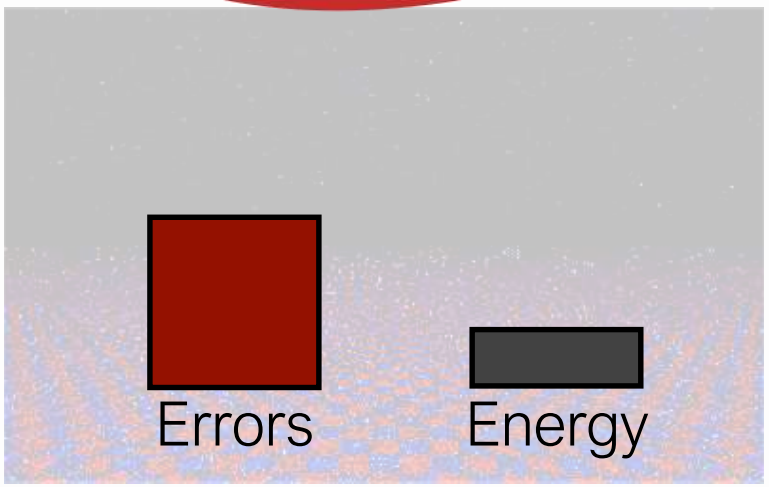


Approximate Computing Example - Images(2)





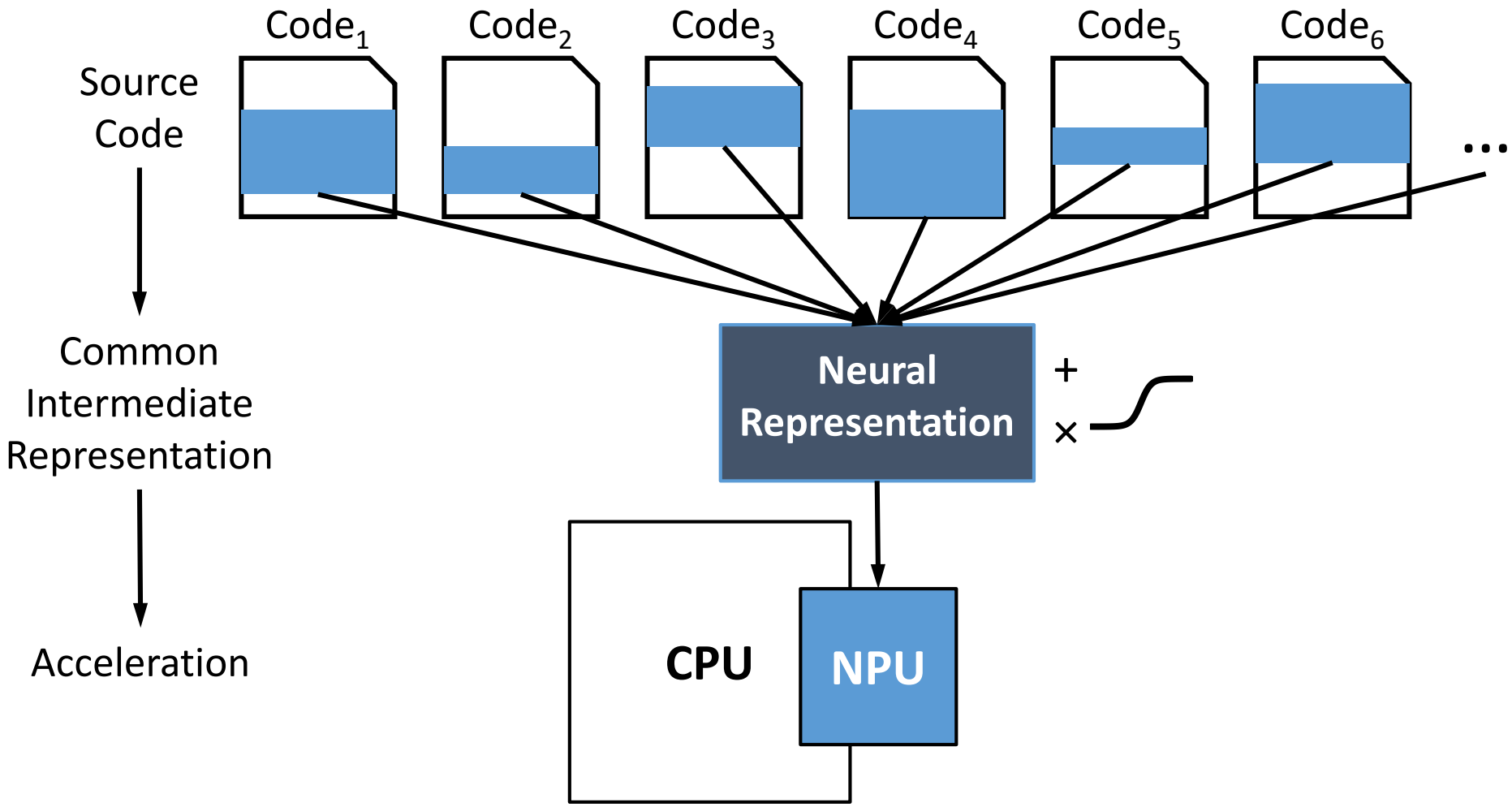




Approximate Computing in Different Areas

PL	EnerJ (UW), Passert (MSR/UW), Rely (MIT), Relax (Wisconsin) Uncertain<T> (MSR), Eon (UMass)
Compiler	Probabilistic transformations (MIT)
Runtime	Green (MSR), PowerDial (MIT), soft error control (UCLA), SAGE & Paraprox (Michigan), Swat (UIUC)
OS/DB	BlinkDB (Berkeley/MIT)
Architecture	ANNs (UW, MSR, INRIA, Wisconsin, Qualcomm) Using Neural Nets for code approximation (GAtech/UW/MSR) Stream Processing (Princeton)
Hardware	Stochastic Processors (UIUC), ERSA (Stanford), Flikker (MSR), QUORA (Purdue), Approximate Storage (MSR, UW) Probabilistic CMOS (Rice), approximate components (Purdue)

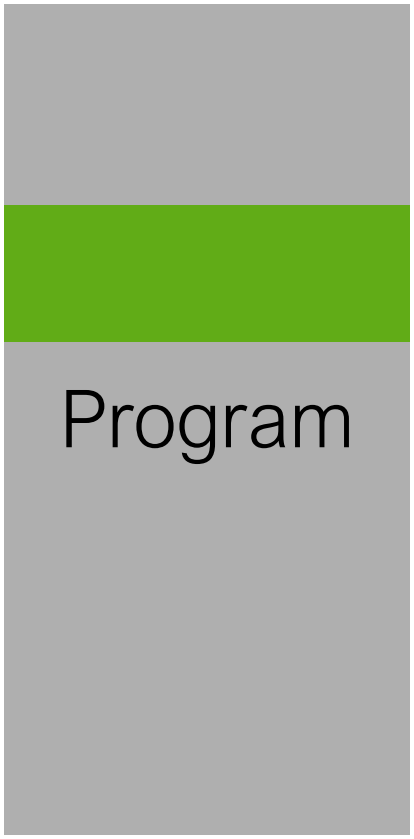
Approximate Computing Using Neutral Networks



Approximate Computing – Program (1)

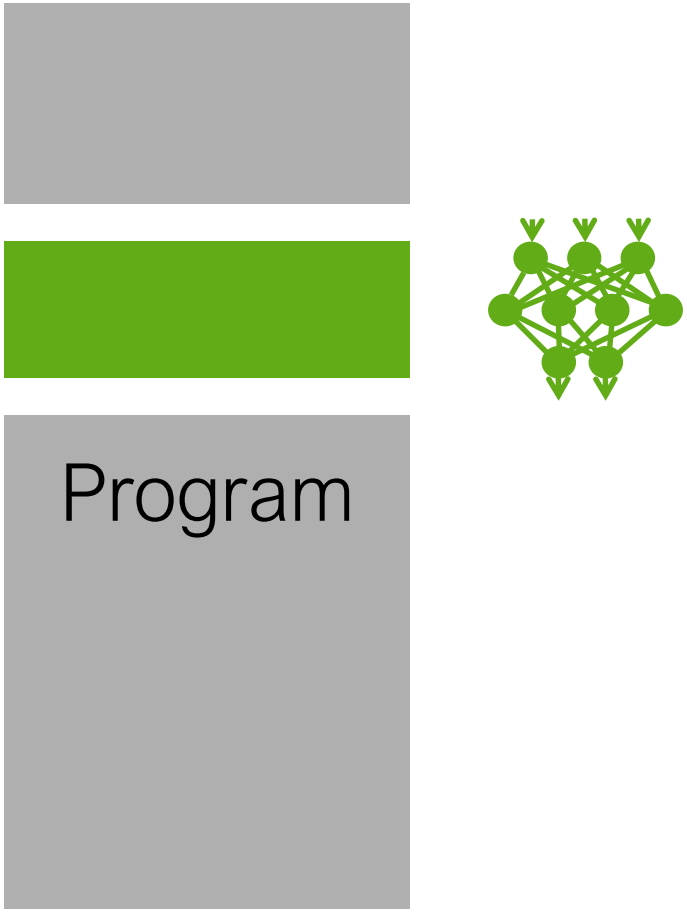
Program

Approximate Computing – Program (2)



Find an approximate program component

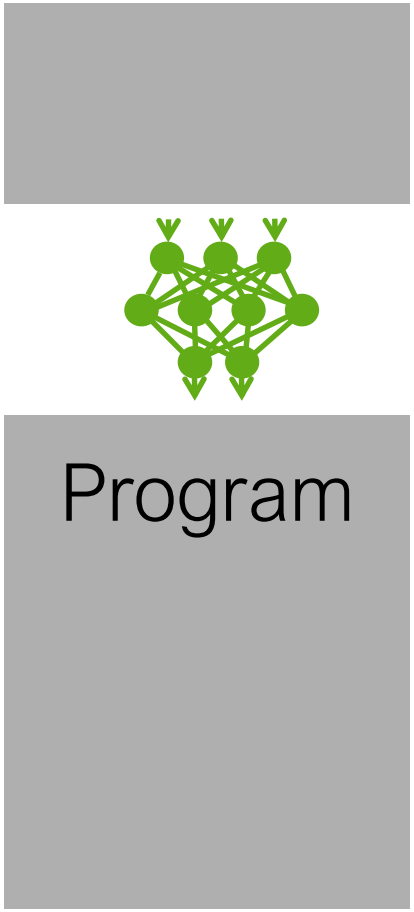
Approximate Computing – Program (3)



Find an approximate program component

Compile the program and train a neural network

Approximate Computing – Program (4)

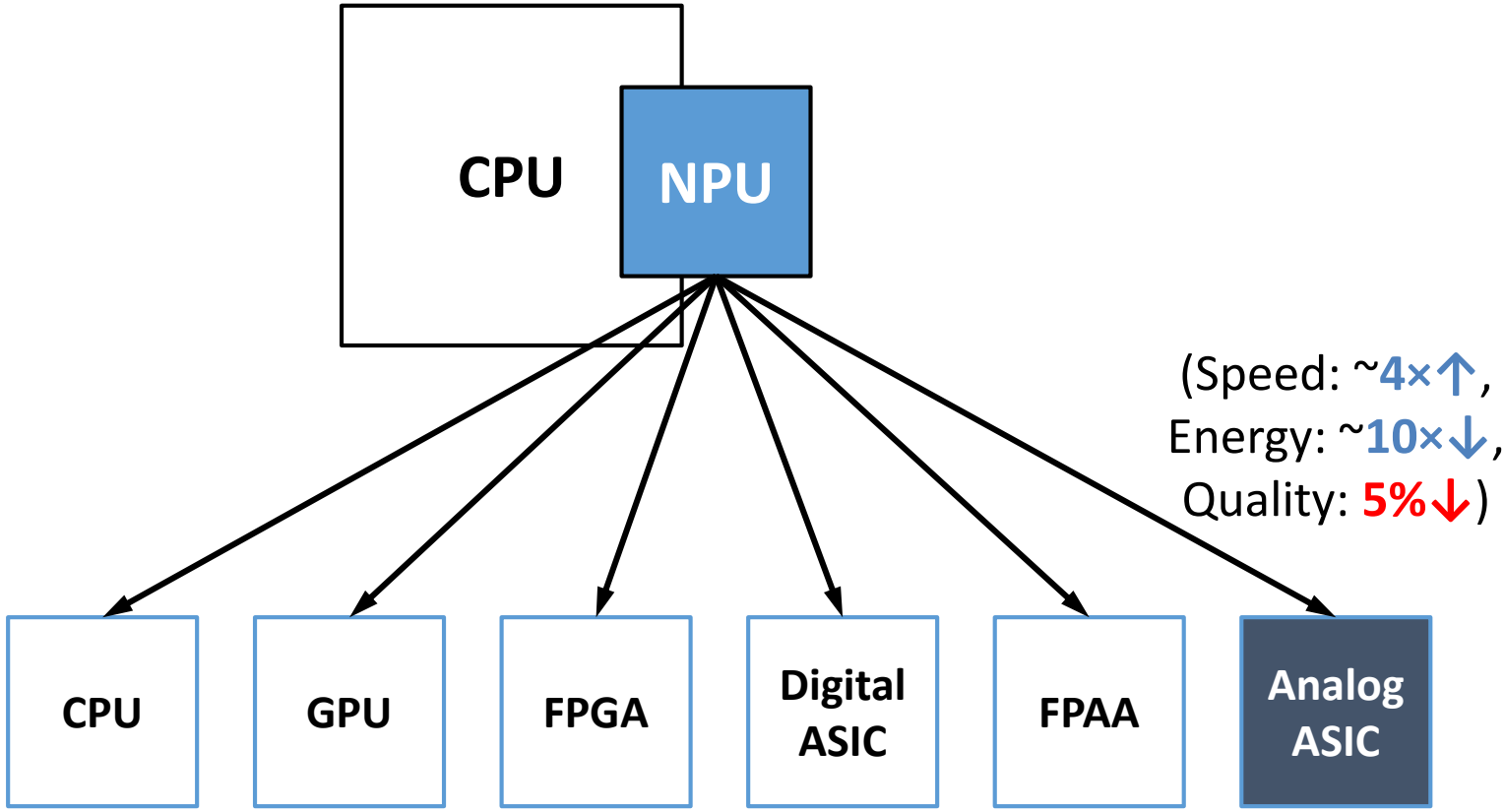


Find an approximate program component

Compile the program and train a neural network

Execute on a fast Neural Processing Unit (NPU)

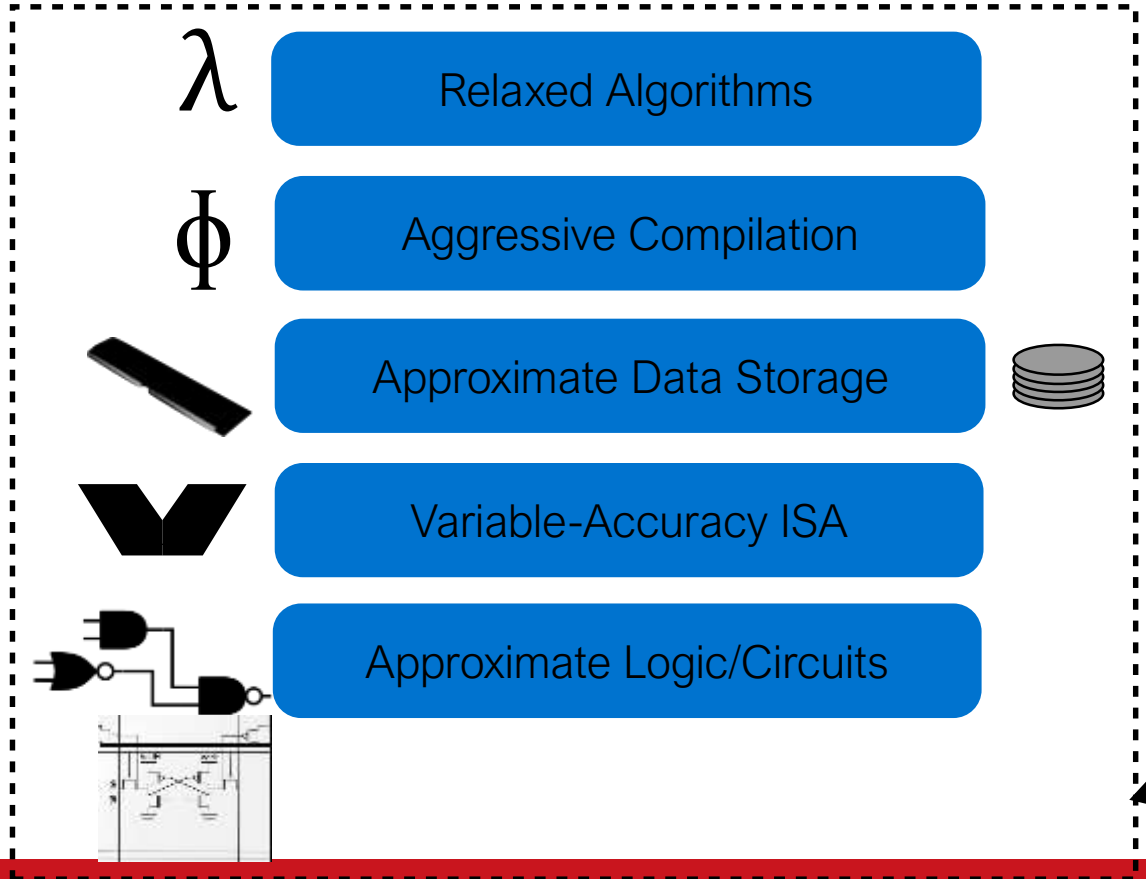
Approximate Computing – Neutral Network Acceleration



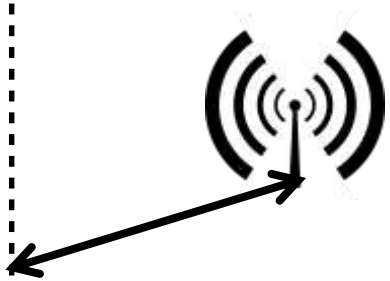
Approximate Codes

Disciplined Approximate Programming
(EnerJ, EnerC,...)

```
int p = 5;
@Approx int a = 7;
for (int x = 0..) {
        a += func(2);
        @Approx int z;
        z = p * 2;
        p += 4;
}
a /= 9;
p += 10;
socket.send(z);
write(file, z);
```

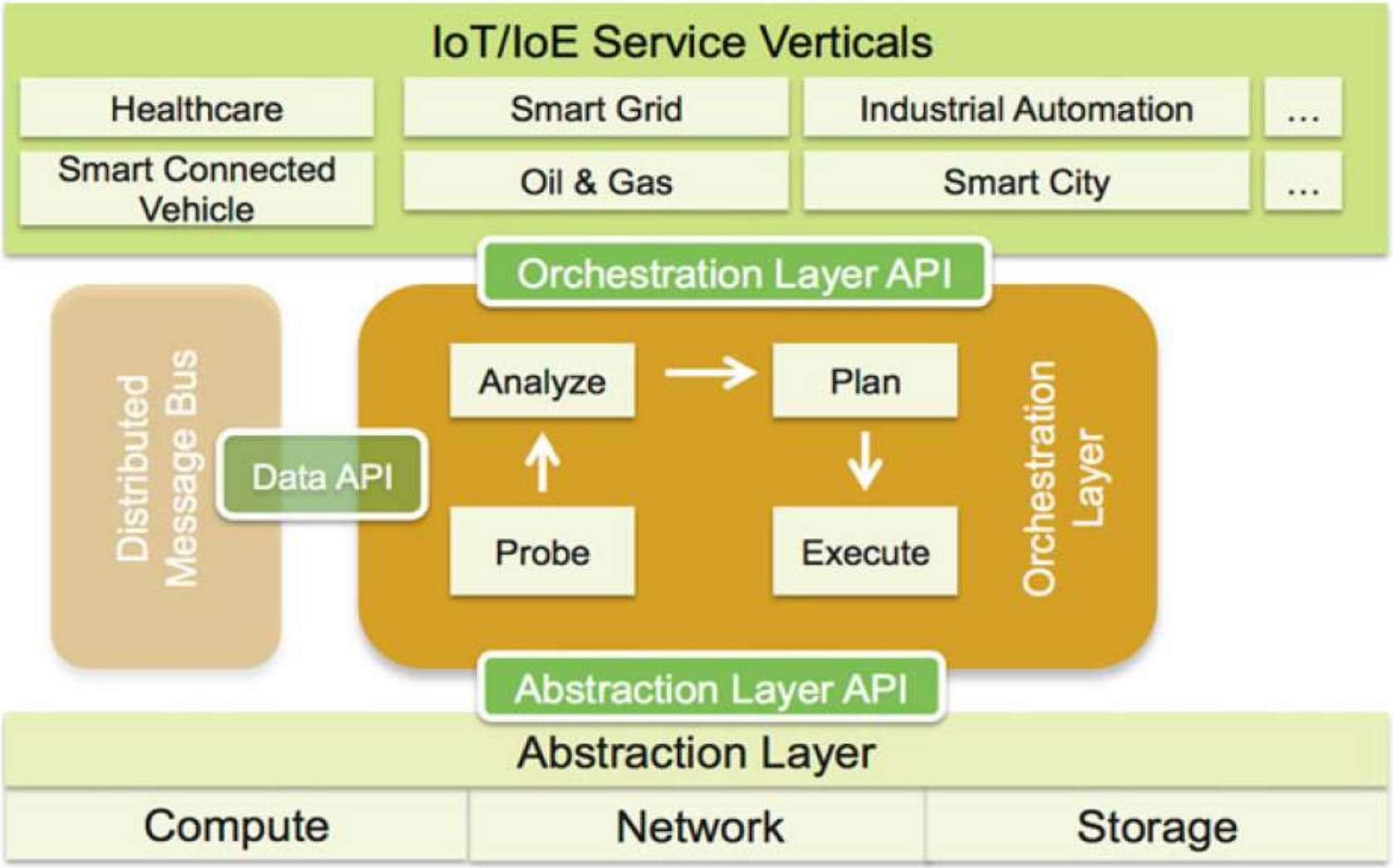


Variable-quality wireless communication

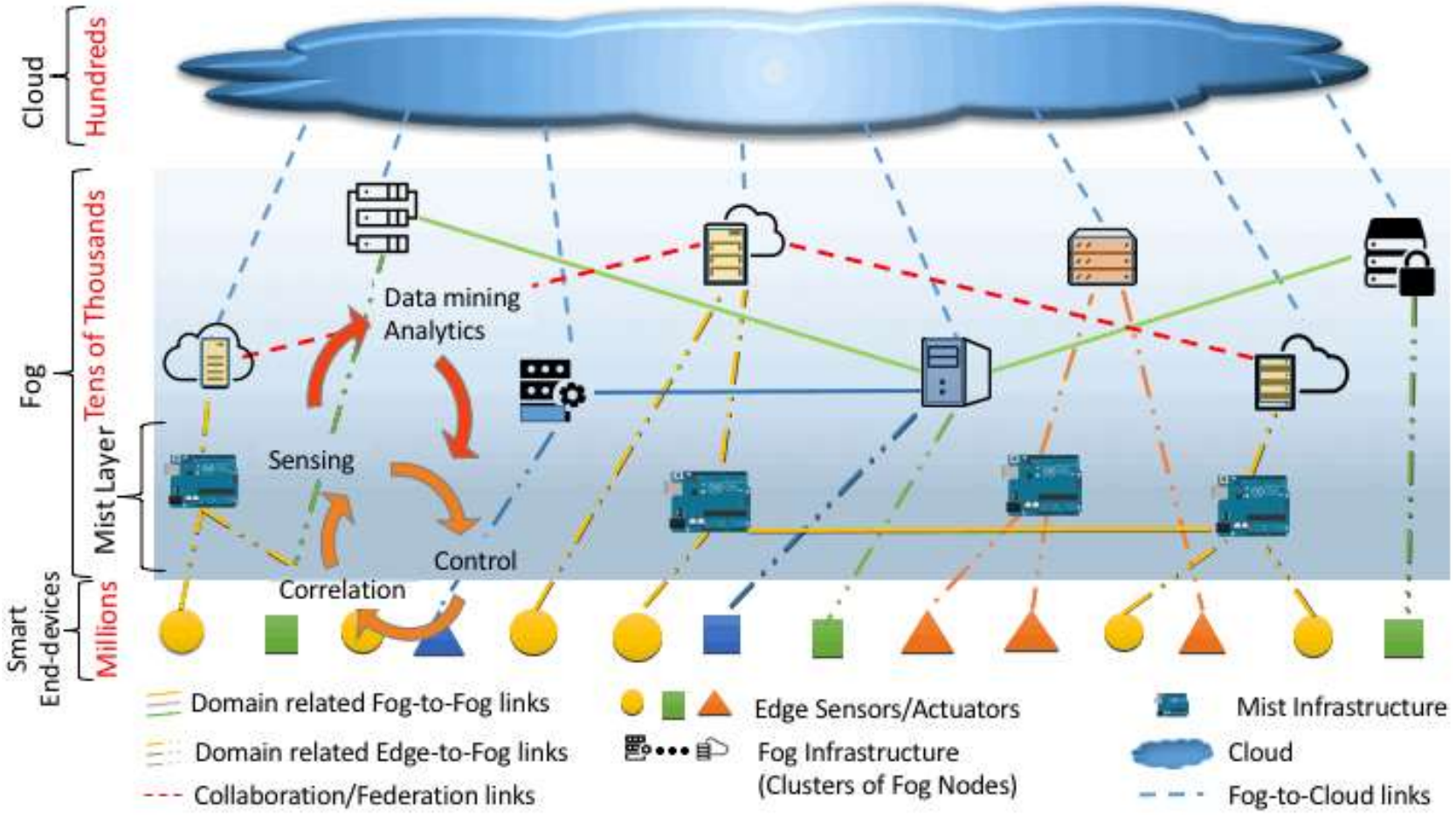


2 Intro. to Fog/Edge Networking

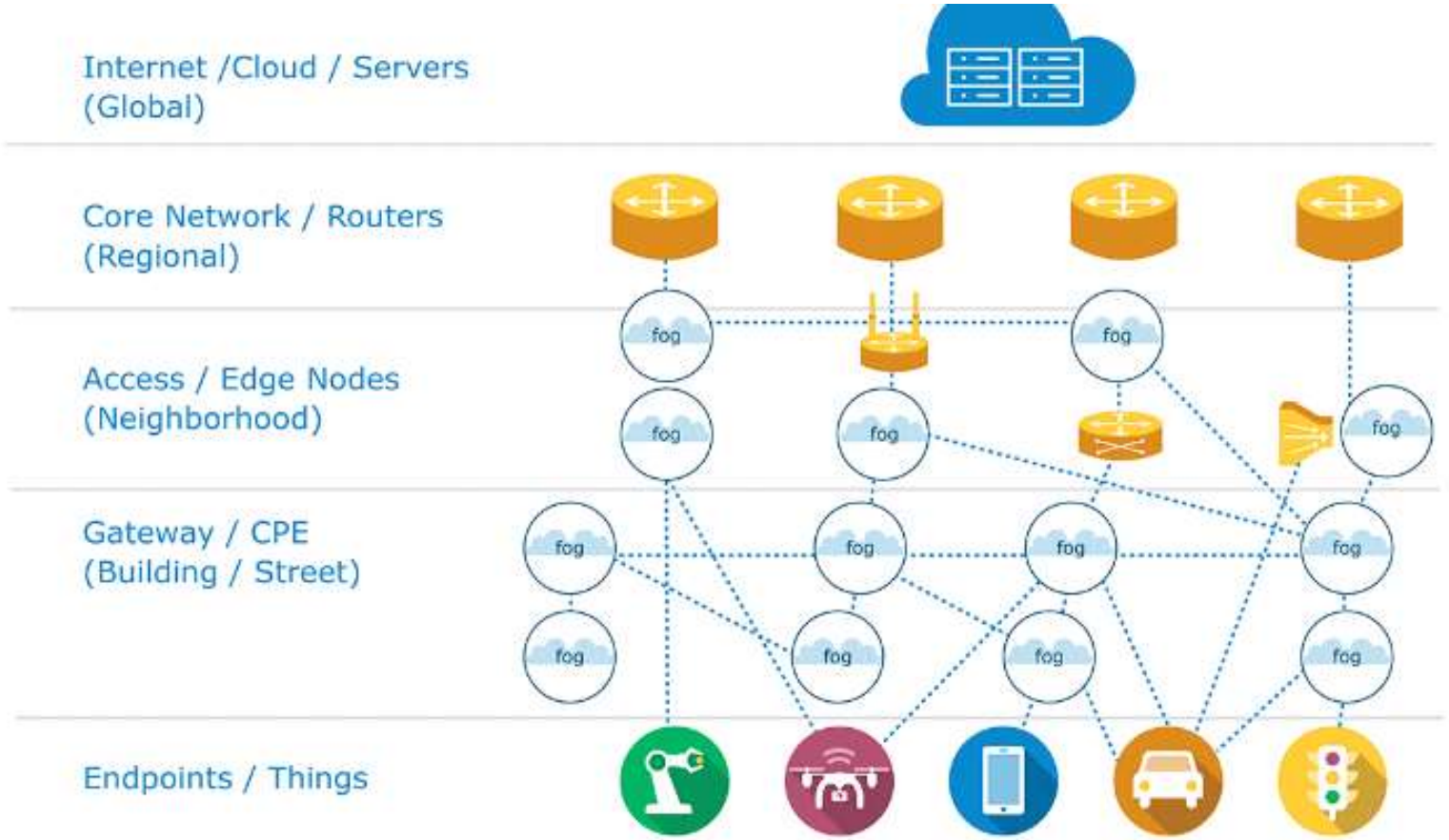
Edge Architecture



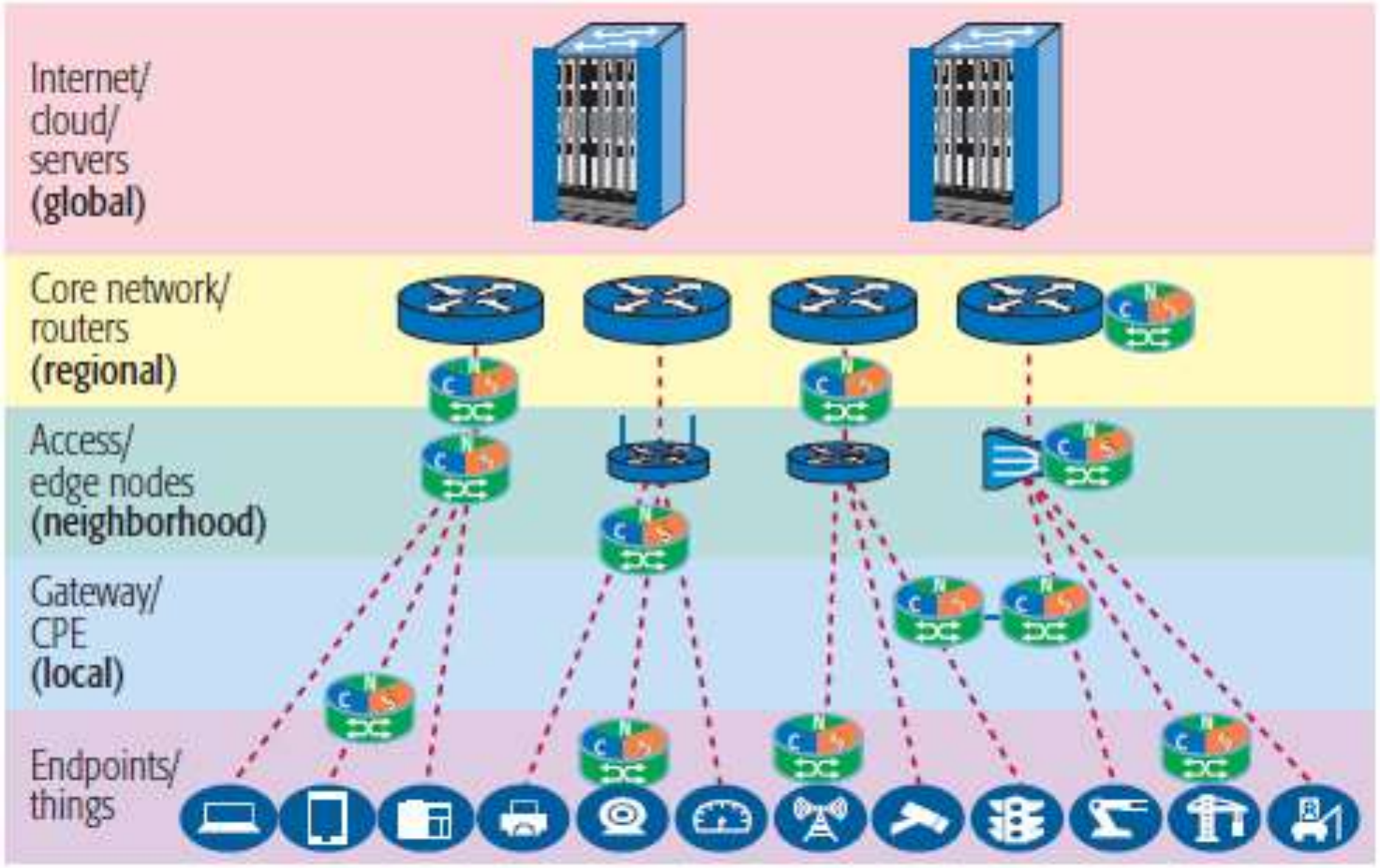
Edge-Fog-Cloud Network



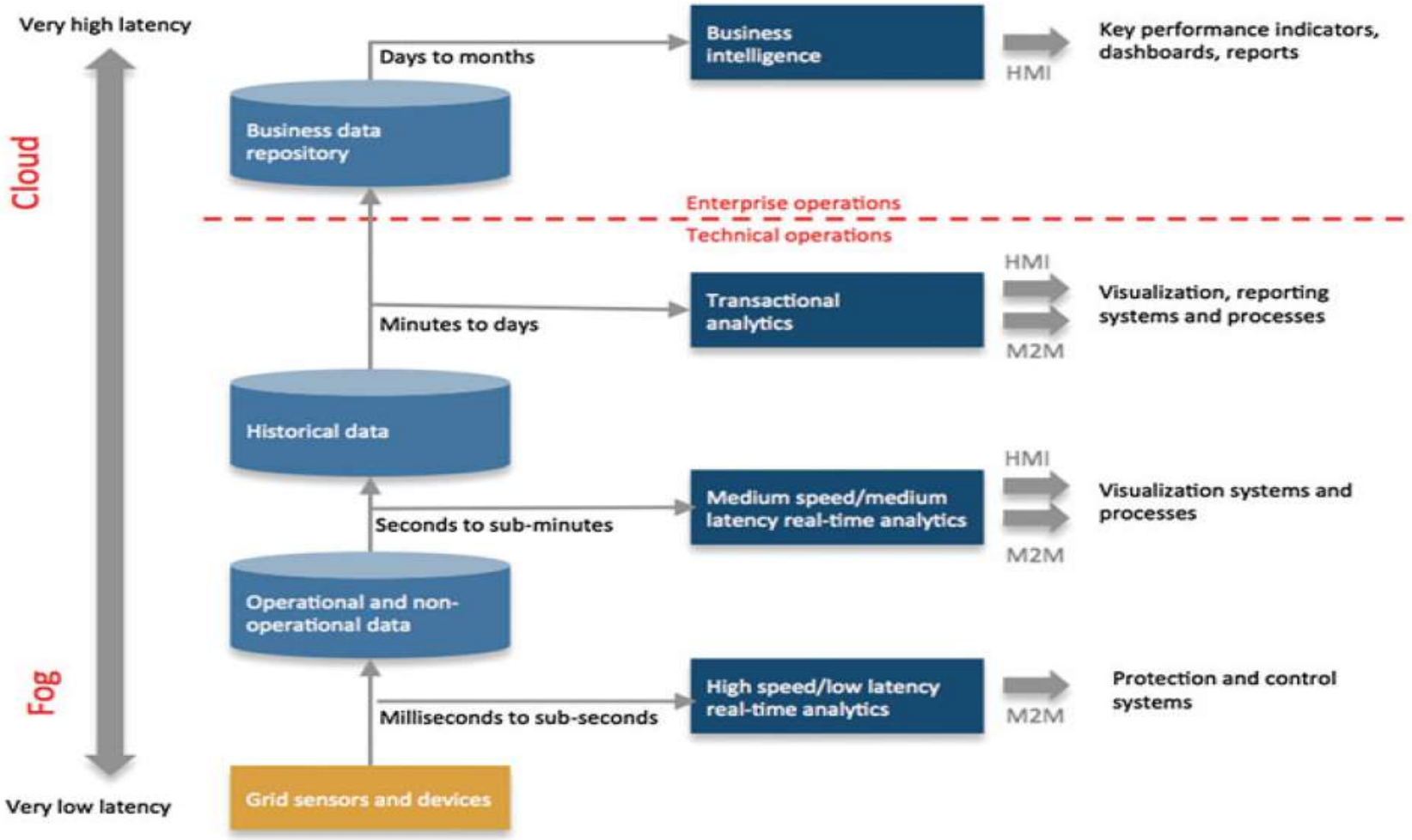
Layered Network (1)



Layered Network (2)

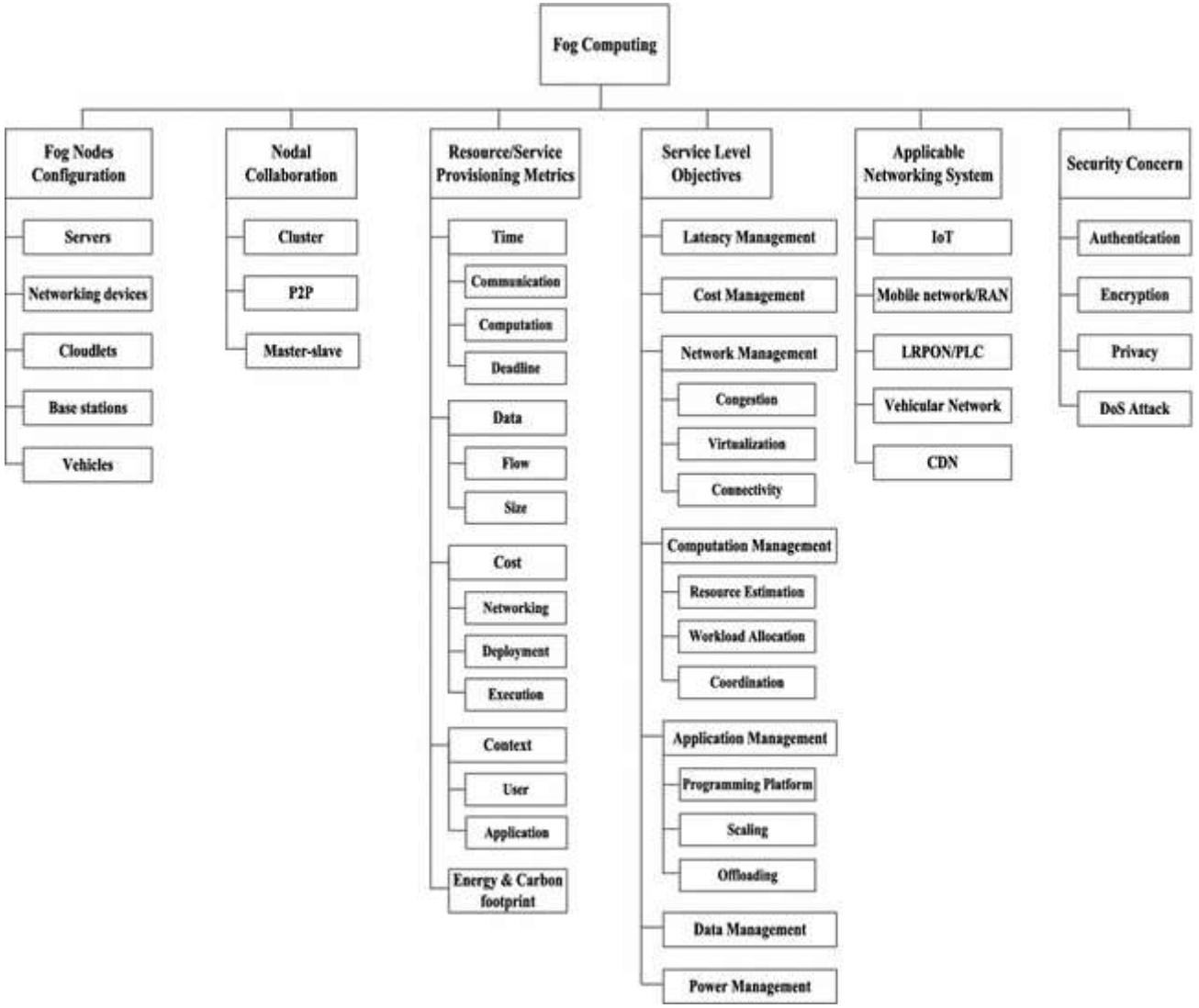


Different Requirements on Latency

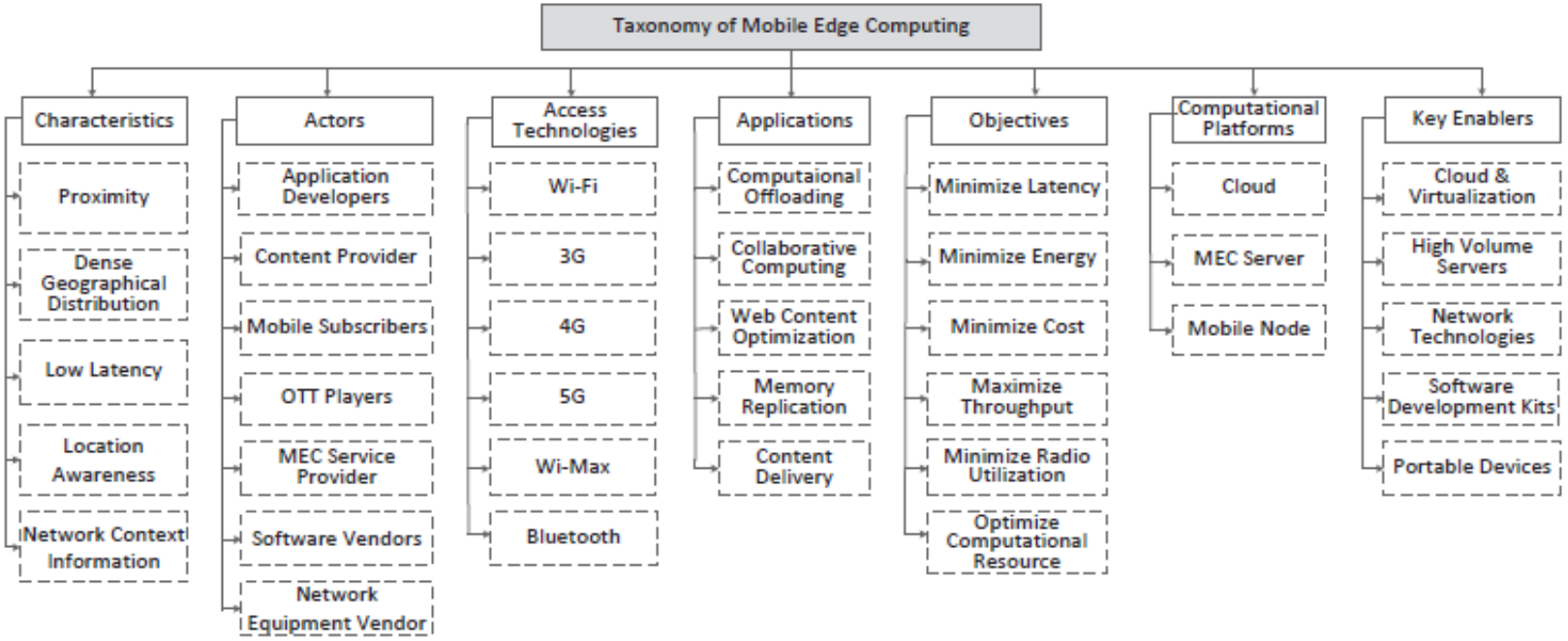


Note: HMI = Human to machine; M2M = Machine to machine

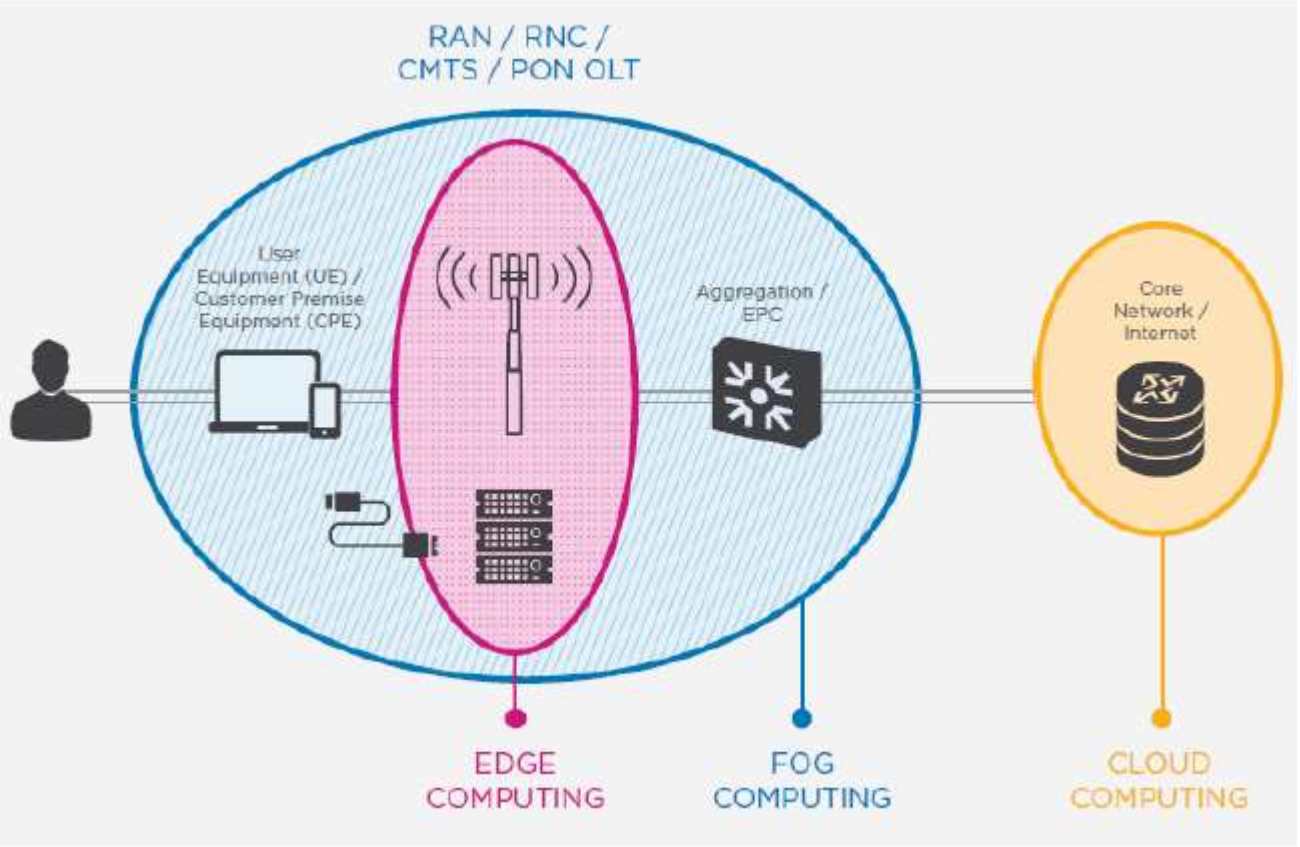
Different Network Topology



Different Network Protocols

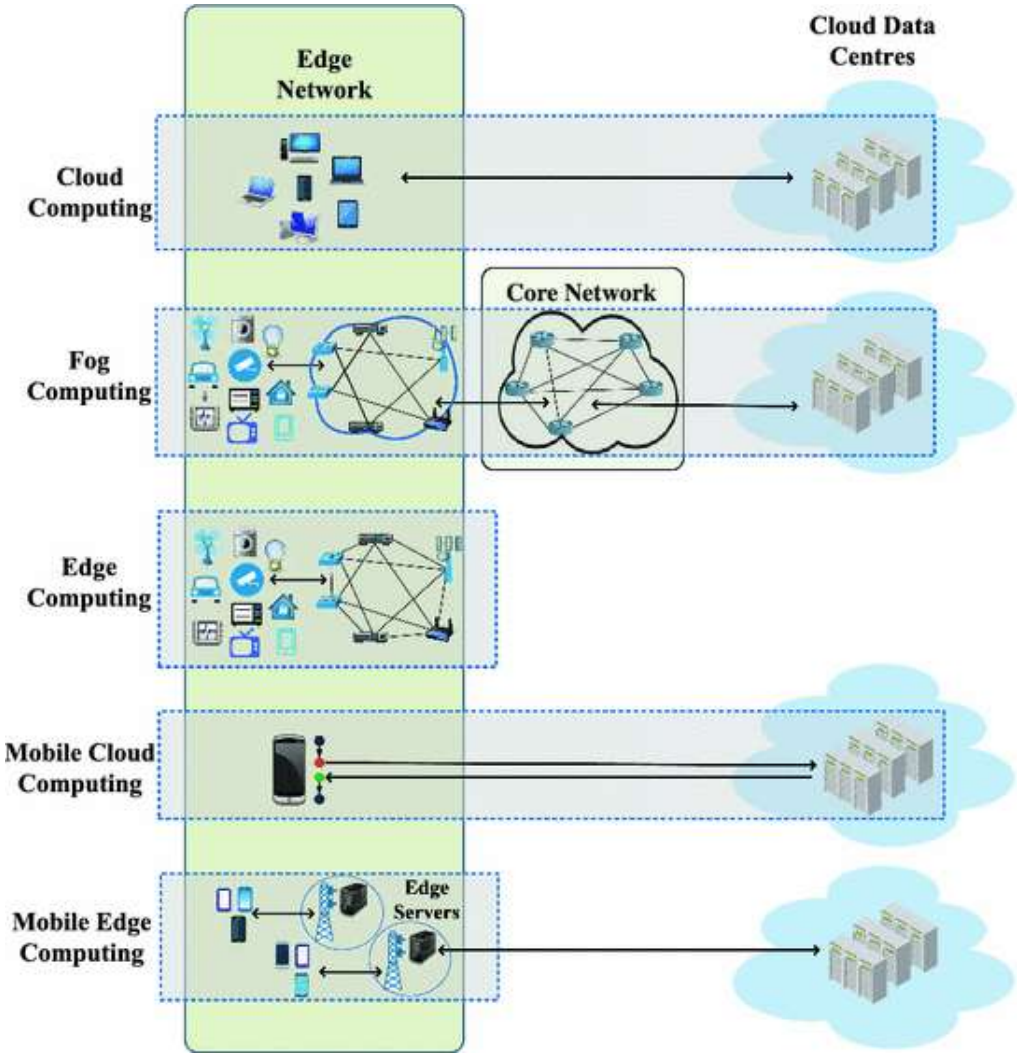


Different Network Accesses

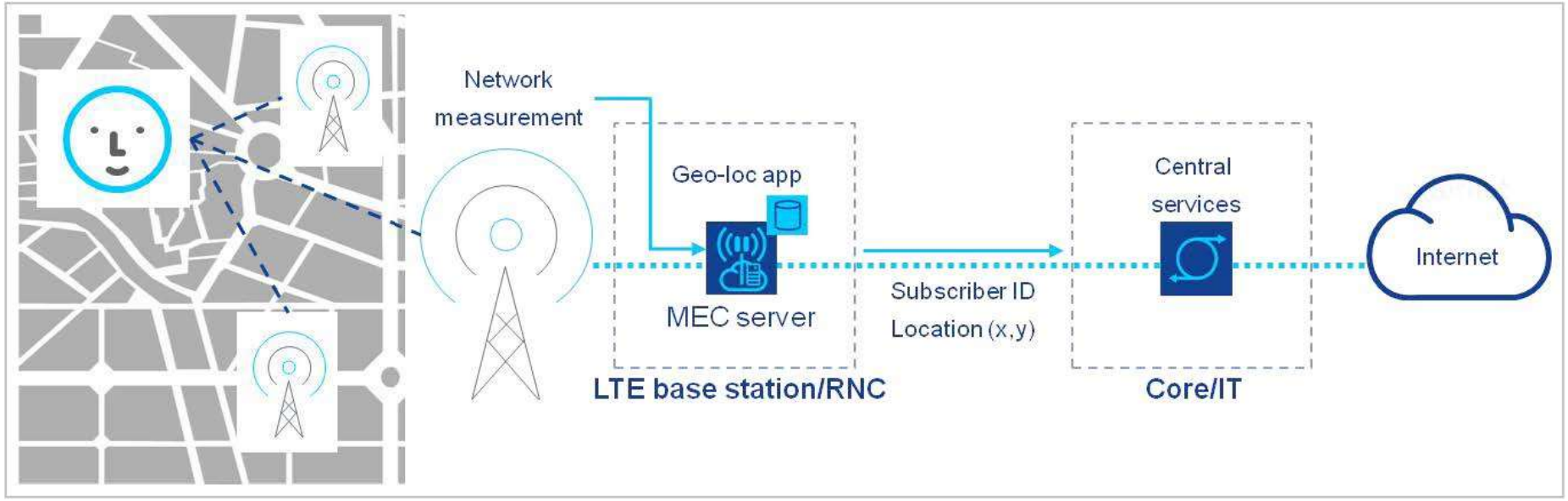


- RAN - Radio Area Network for LTE/5G
- RNC- Radio Network Controller for WiFi
- CMTS- Cable Modem Termination System
- PON OLT for fiber
- EPC – evolved PacketCore

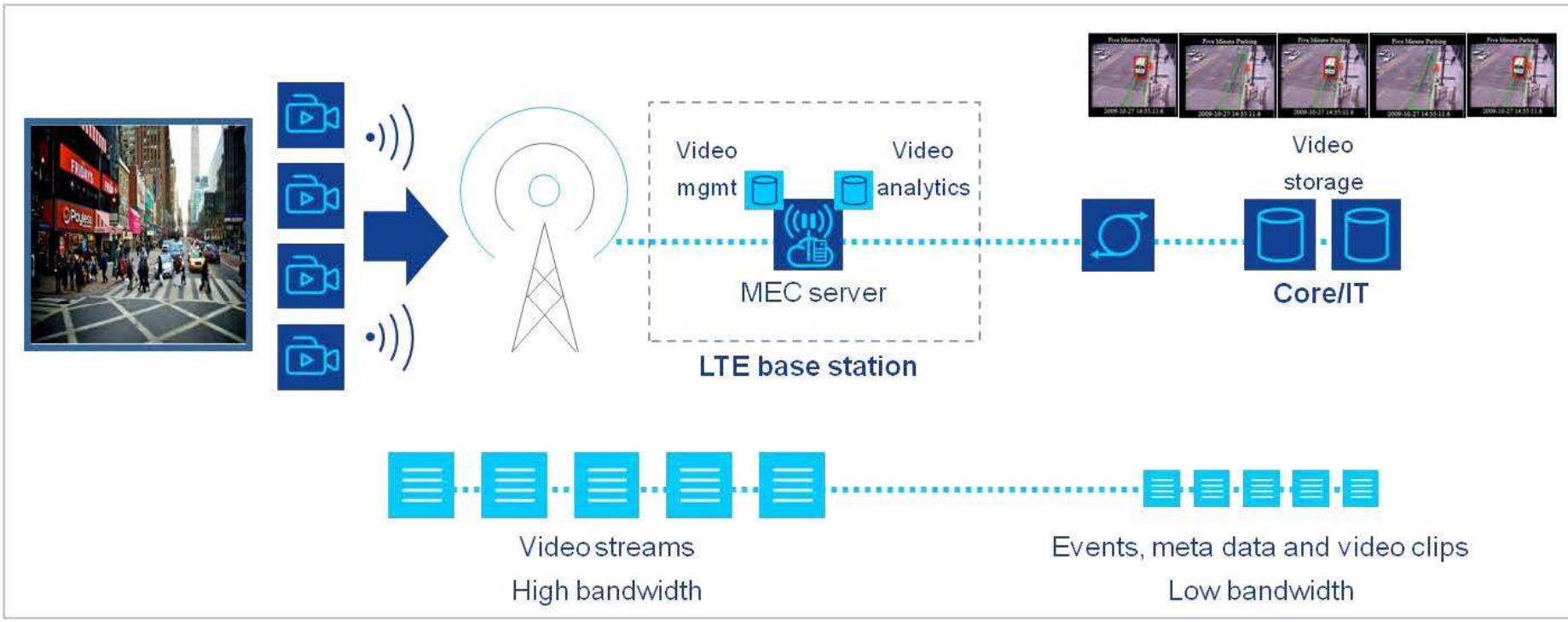
Different Network Connections



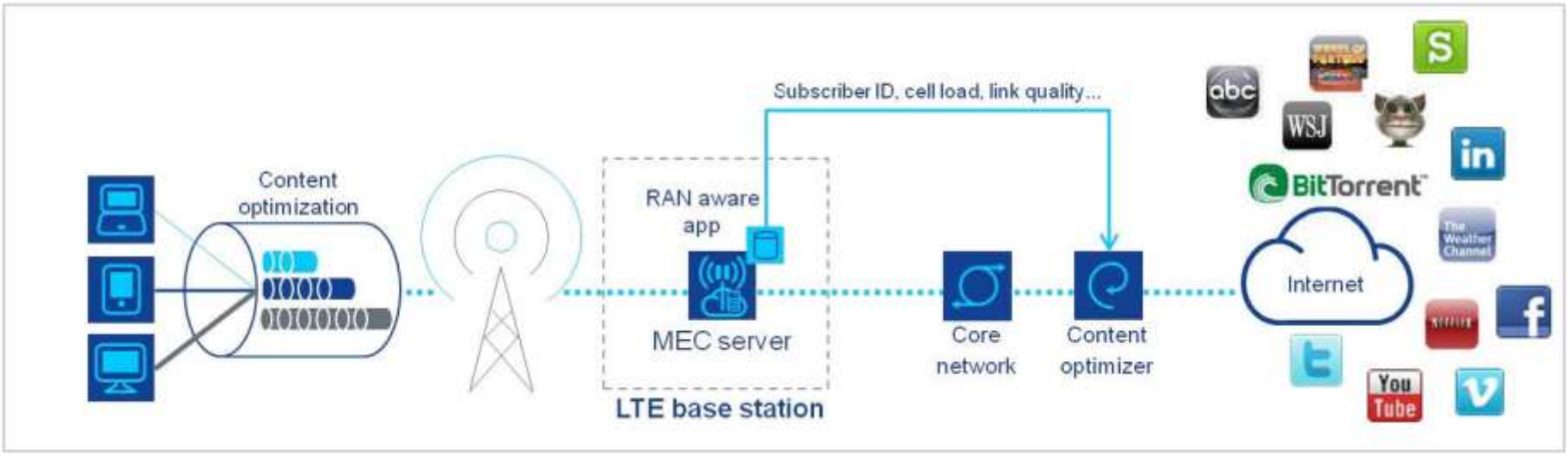
Applications – Device Location



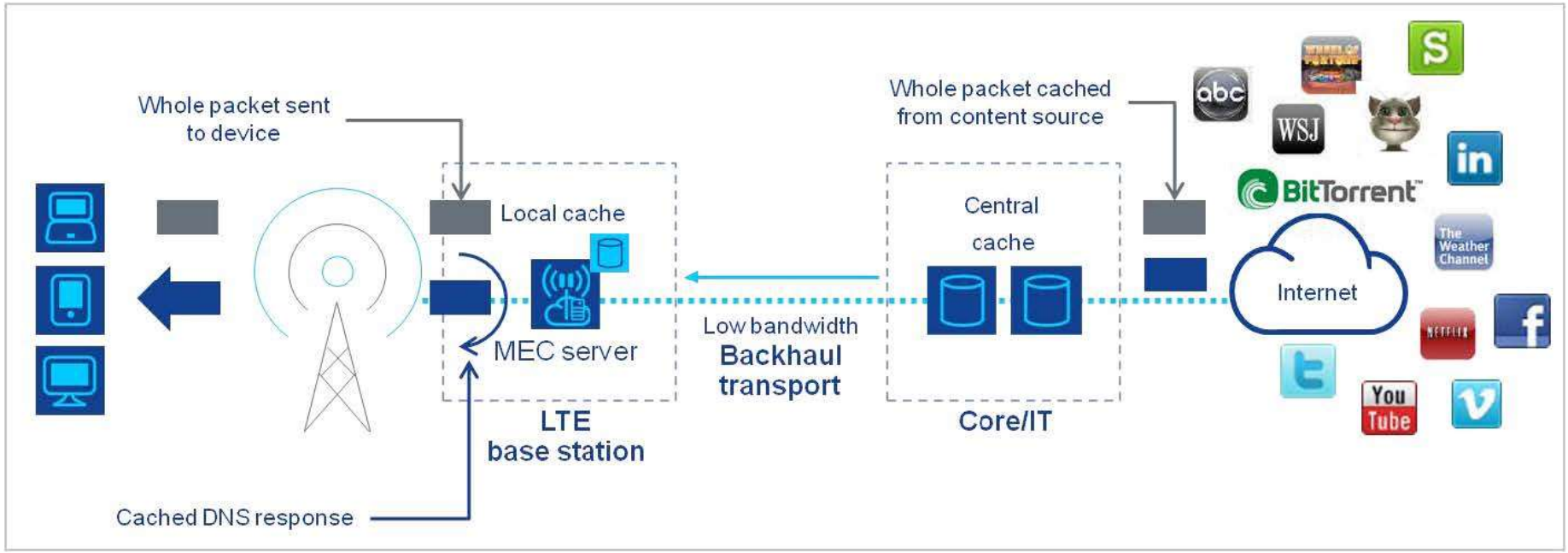
Applications – Video Analytics



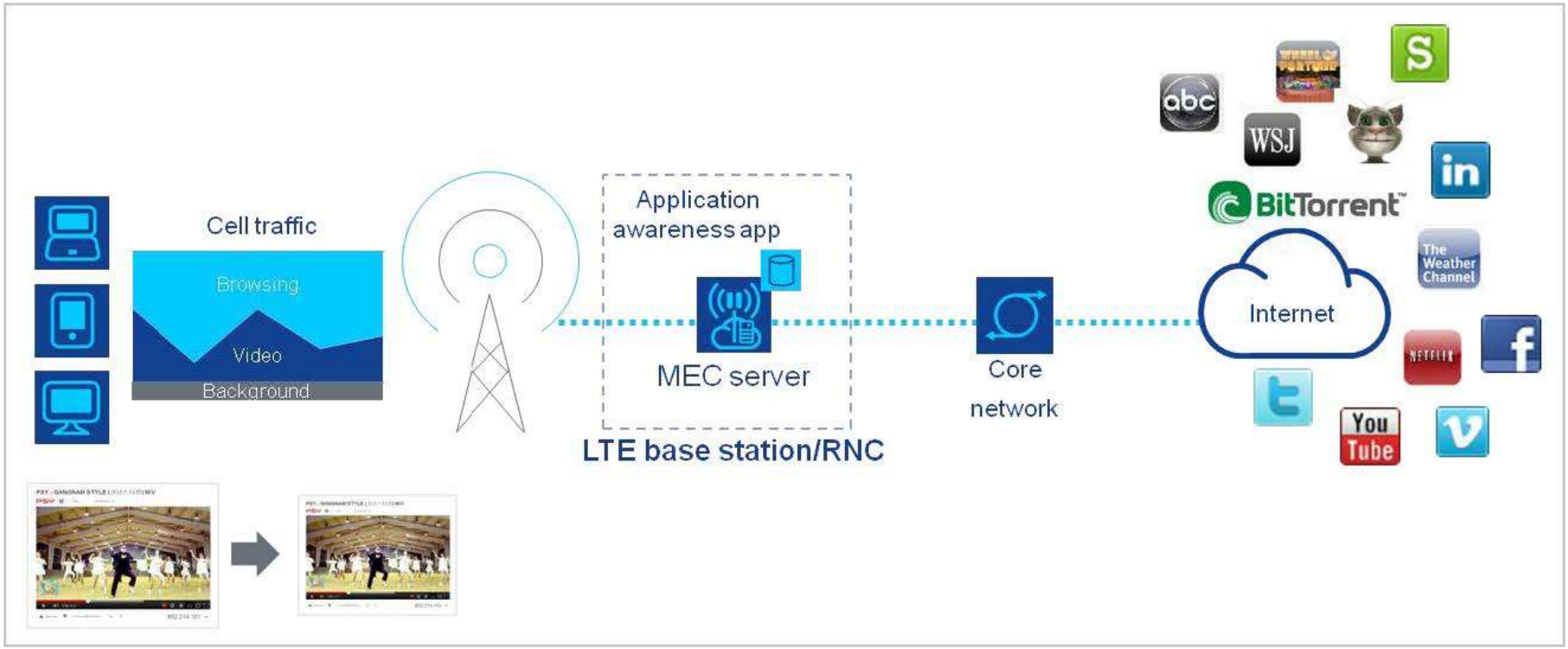
Applications – Content Optimization



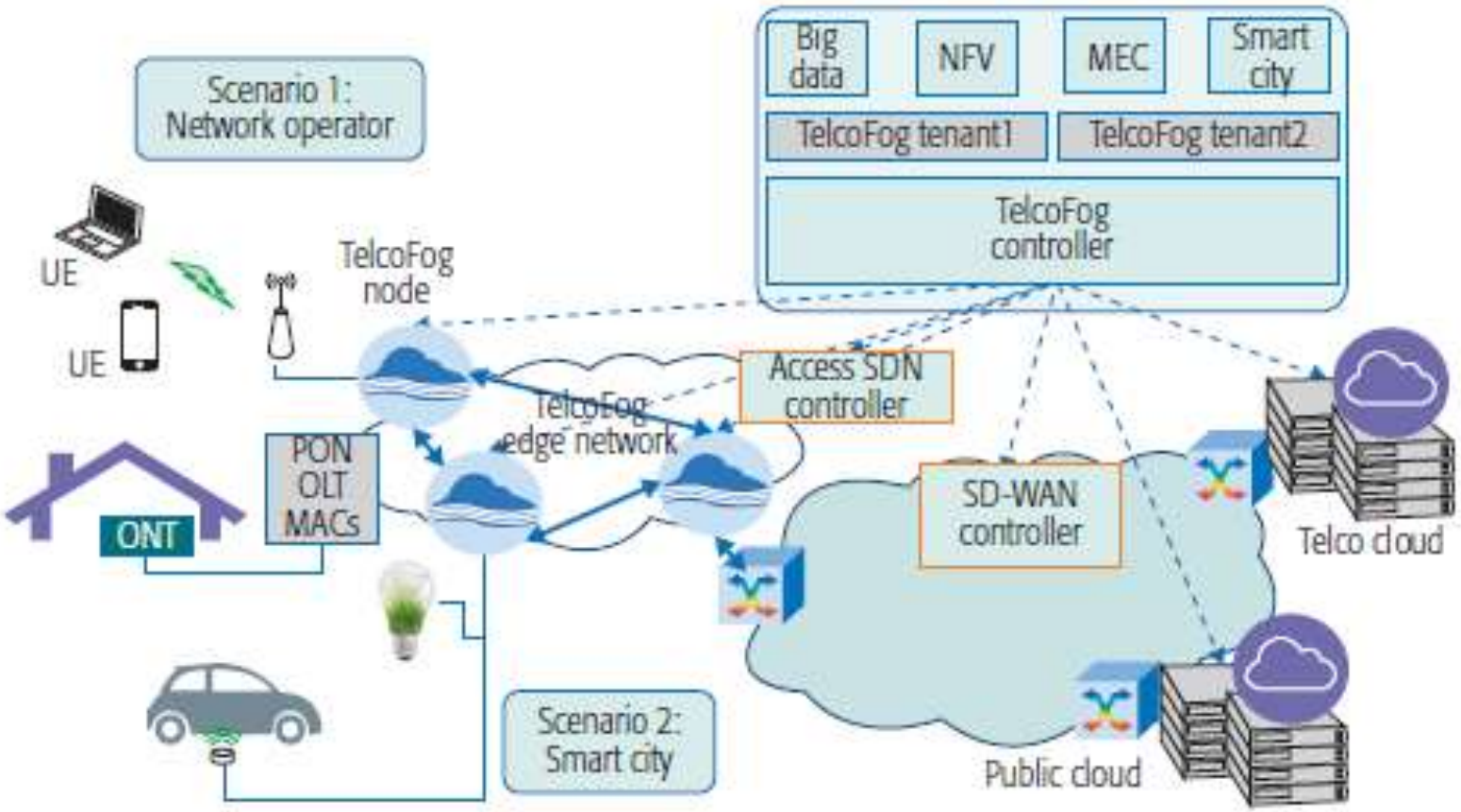
Applications – DNS Caching



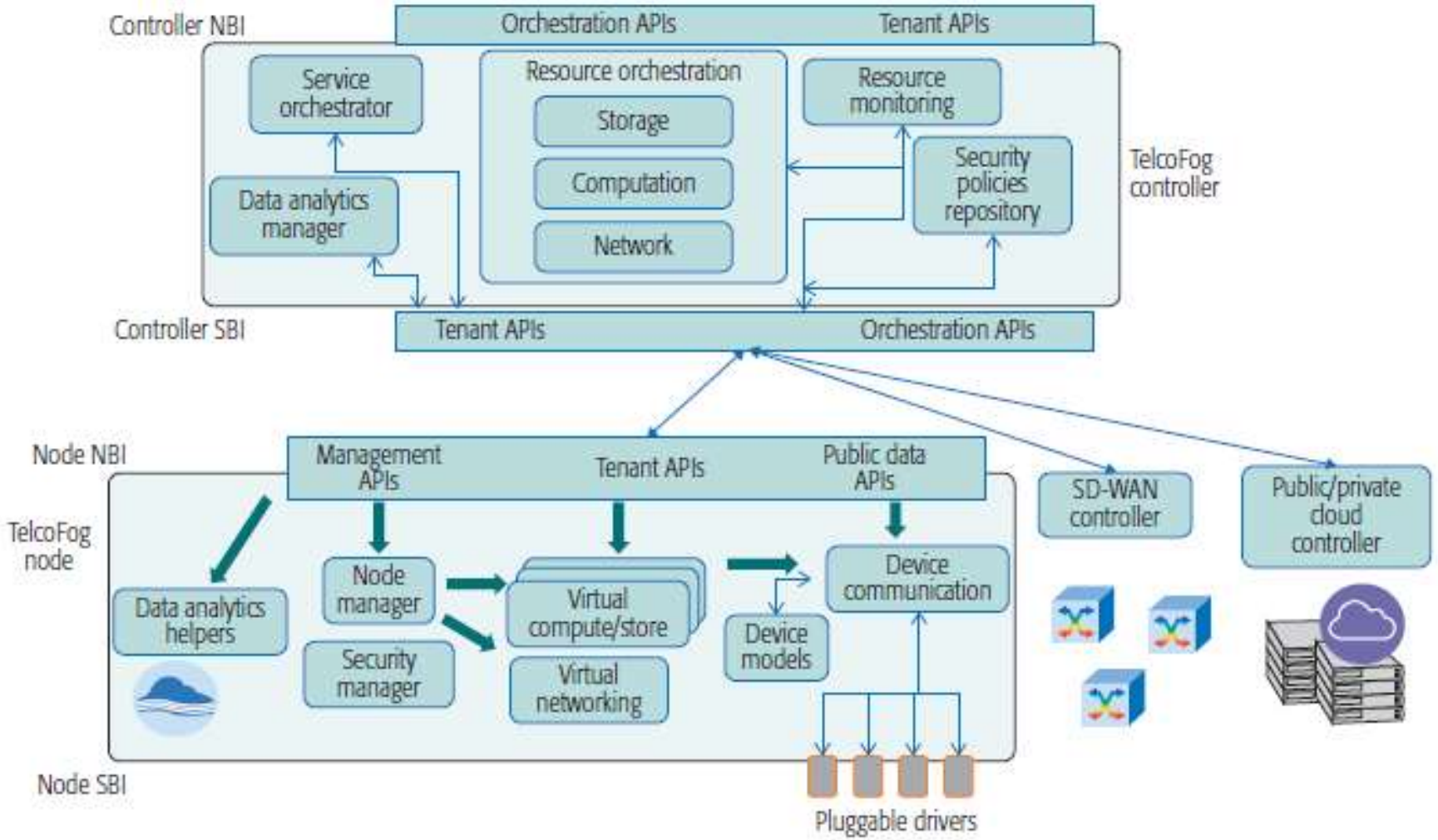
Applications – Application Optimization



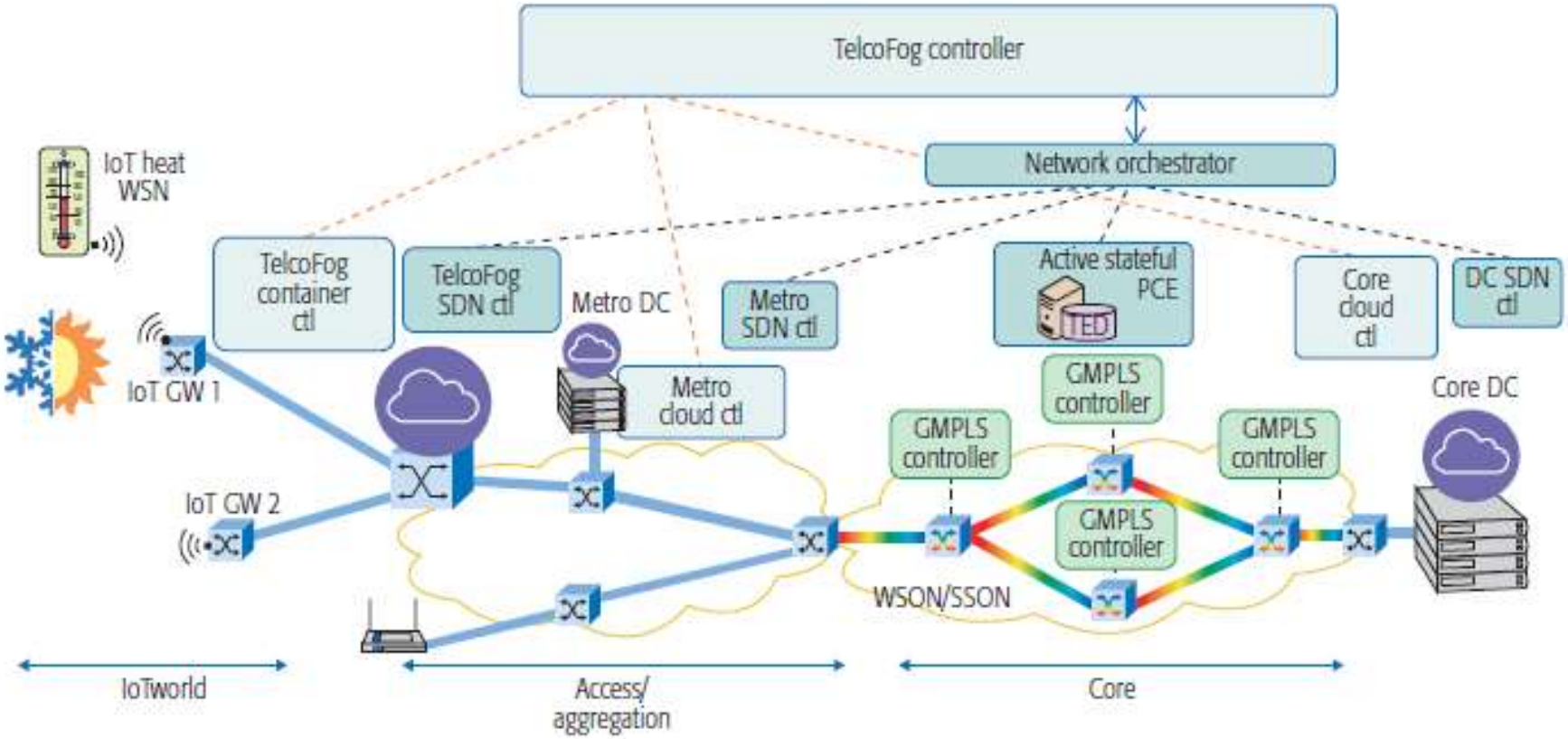
Edge-Fog-Cloud Network Example: TelcoFog (1)



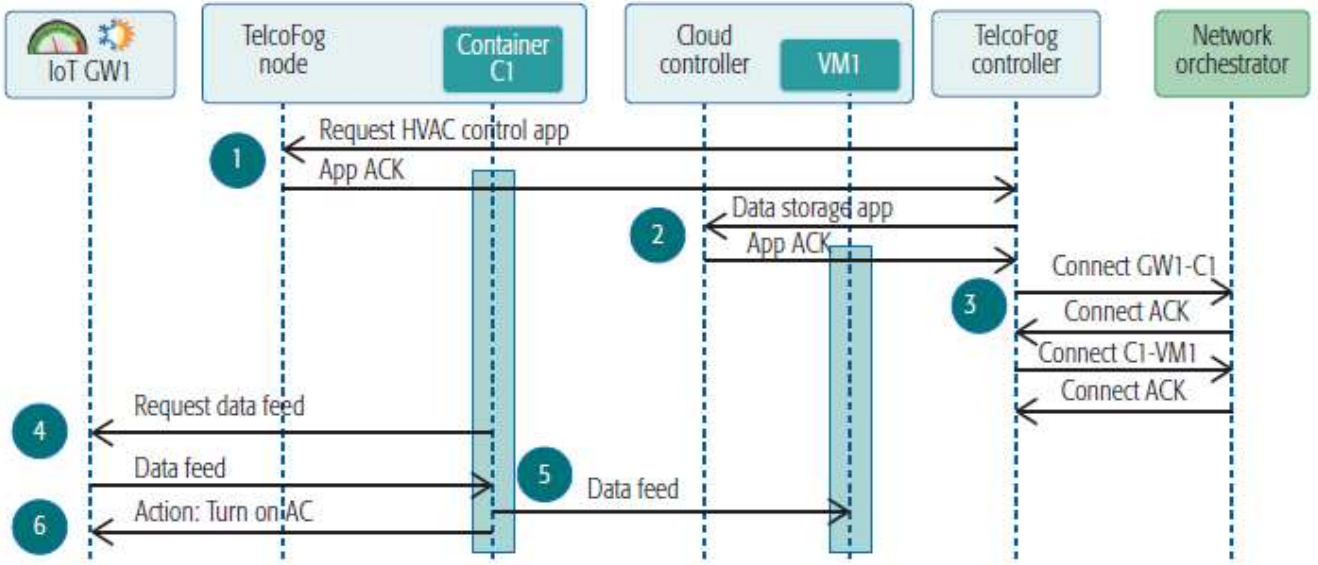
Edge-Fog-Cloud Network Example: TelcoFog (2)



Edge-Fog-Cloud Network Example: TelcoFog (3)



Edge-Fog-Cloud Network Example: TelcoFog (4)



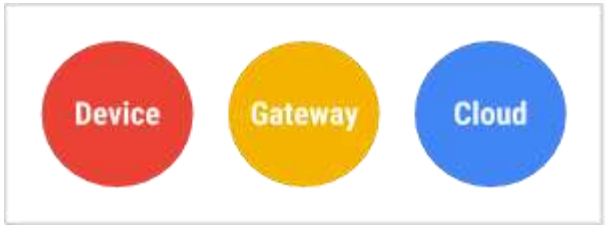
1	TELCOFOG_CTL	FOG_NODE_1	HTTP	POST /restconf/config/container/hvac_container/ HTTP/1.1 (application/json)
	FOG_NODE_1	TELCOFOG_CTL	HTTP	HTTP/1.0 200 OK (application/json)
2	TELCOFOG_CTL	CLOUD_CTL	HTTP	POST /v2.1/e9102671a6004a209f8ccec14cadf2fd/servers HTTP/1.1 (application/json)
	CLOUD_CTL	TELCOFOG_CTL	HTTP	HTTP/1.1 202 Accepted (application/json)
	TELCOFOG_CTL	CLOUD_CTL	HTTP	GET / v2.1/e9102671a6004a209f8ccec14cadf2fd/servers/45244ef2-8a9a-448b-bd38-62c31f8edd2a HTTP/1.1
	CLOUD_CTL	TELCOFOG_CTL	HTTP	HTTP/1.1 200 OK (application/json)
3	TELCOFOG_CTL	NET_ORCH	HTTP	POST /restconf/config/Context/_connectivityService/gw1_c1/ HTTP/1.1 (application/json)
	NET_ORCH	TELCOFOG_CTL	HTTP	HTTP/1.0 200 OK (application/json)
	TELCOFOG_CTL	NET_ORCH	HTTP	POST /restconf/config/Context/_connectivityService/c1_vm1/ HTTP/1.1 (application/json)
	NET_ORCH	TELCOFOG_CTL	HTTP	HTTP/1.0 200 OK (application/json)
4	HVAC_CONTAINER	IOT_GW1	HTTP	GET /restconf/config/HVAC/Temperature/ HTTP/1.1
	IOT_GW1	HVAC_CONTAINER	HTTP	HTTP/1.0 200 OK (application/json)
5	HVAC_CONTAINER	MYSQL_CLOUD_VM	MySQL	Request Query
	MYSQL_CLOUD_VM	HVAC_CONTAINER	MySQL	Response OK
6	HVAC_CONTAINER	IOT_GW1	HTTP	PUT /restconf/config/HVAC/Ac/ HTTP/1.1 (application/json)
	IOT_GW1	HVAC_CONTAINER	HTTP	HTTP/1.0 200 OK (application/json)

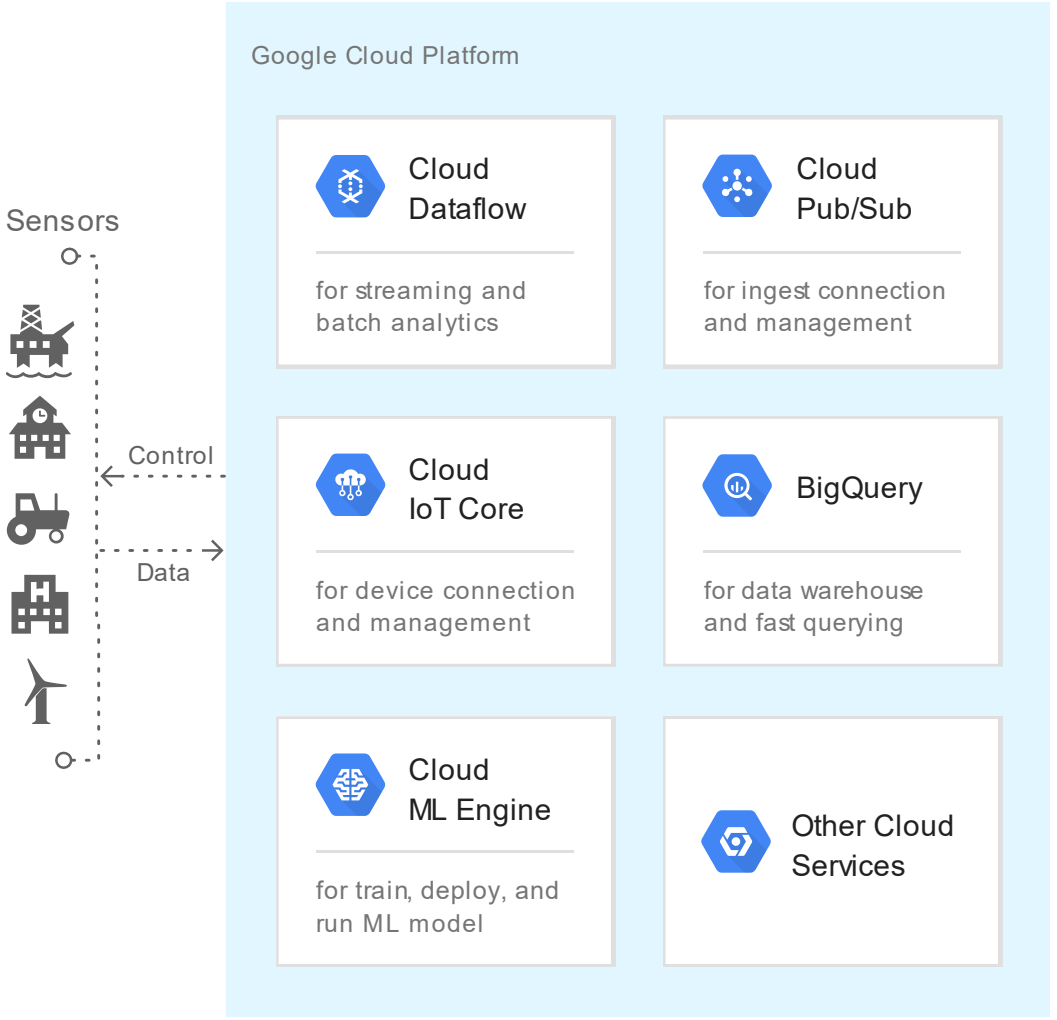
Container deployment delay	399 ms
Connectivity service setup delay (GW1 – C1)	736 ms
HVAC command delay	186 ms
VM deployment delay	16.17 s
Connectivity service setup delay (C1 – VM1)	2.86 s
Edge orchestration delay	1.32 s
Total orchestration delay	20.35 s

3

Industrial Solutions

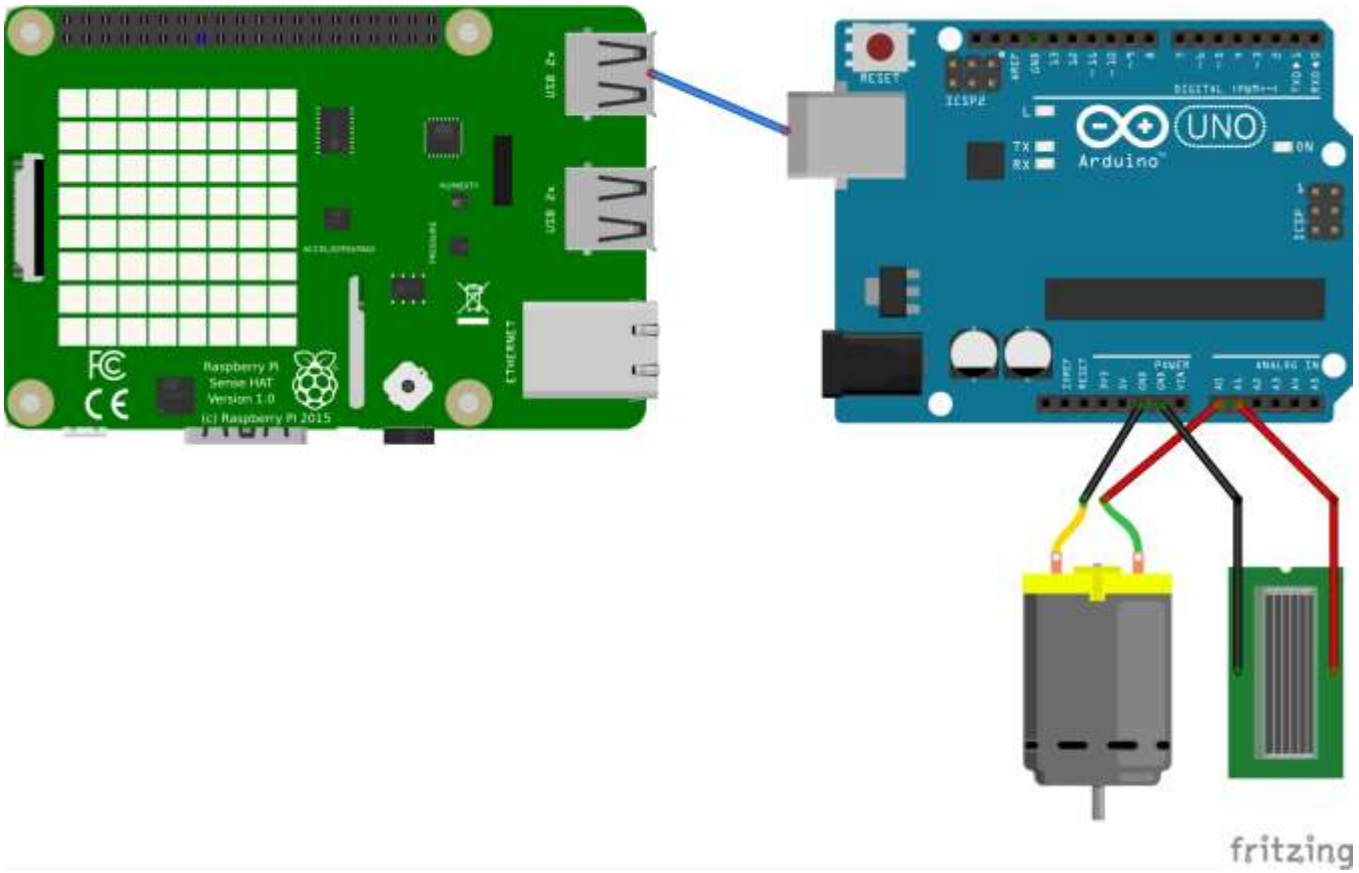
Google IoT Solution



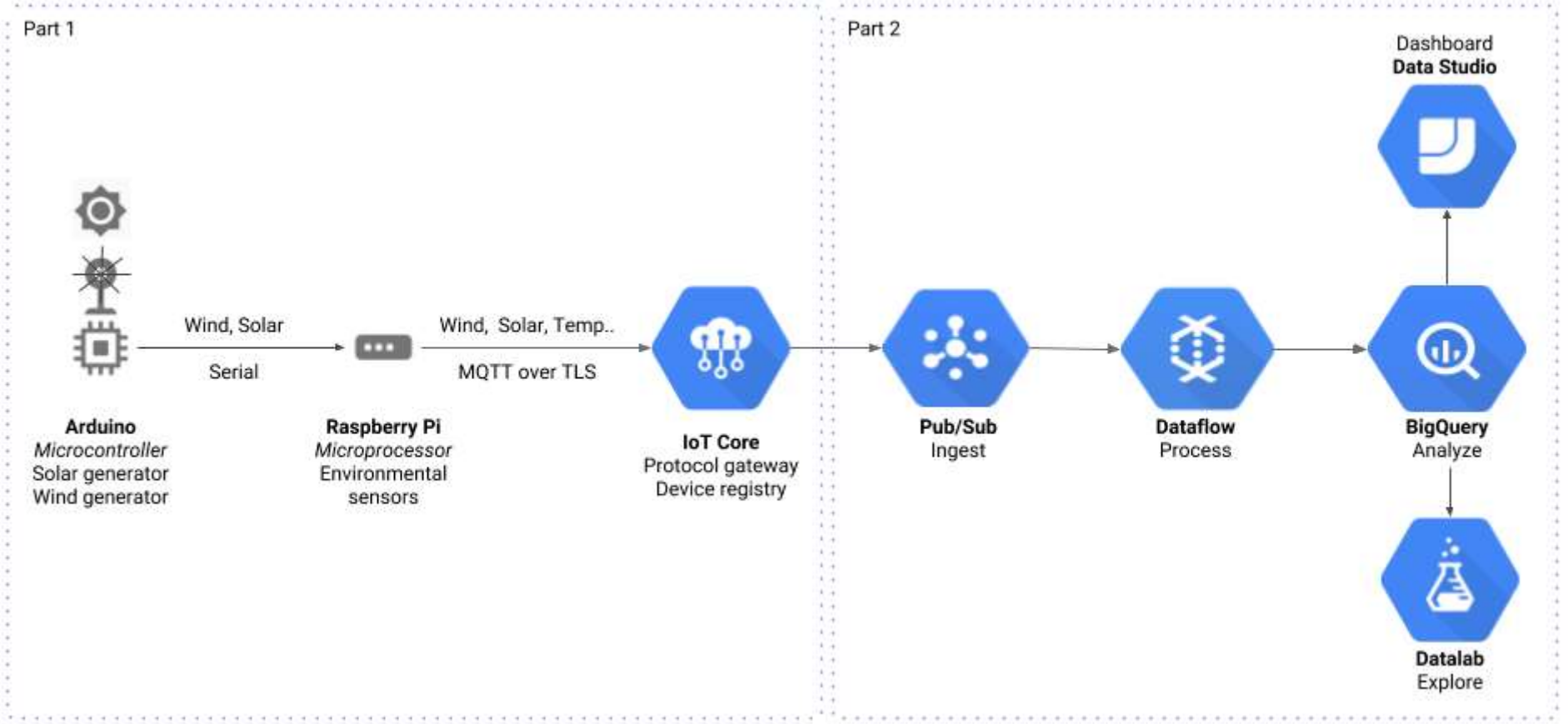


Build & train ML models in the cloud

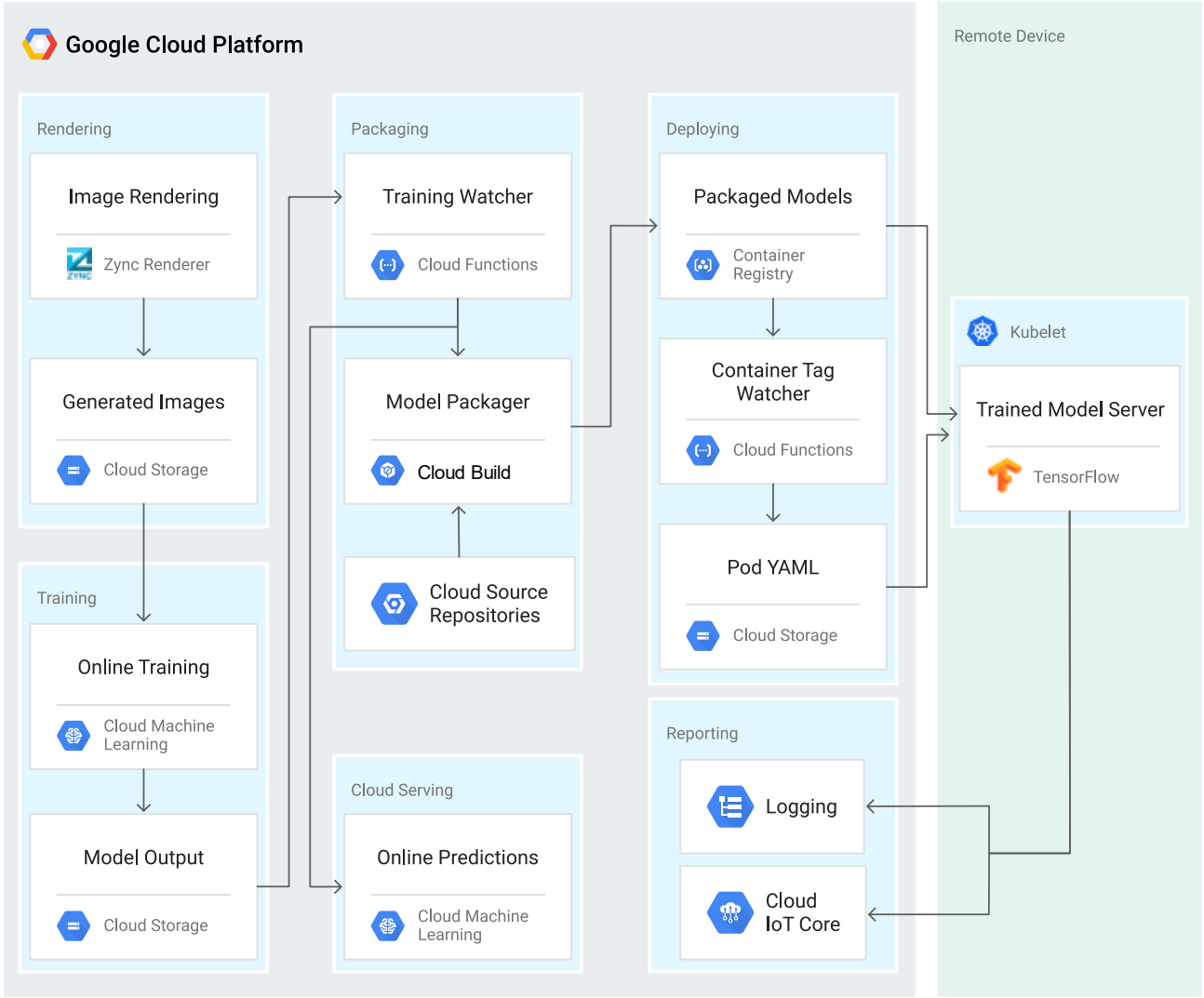
Google Edge Computing (1)



Google Edge Computing (2)

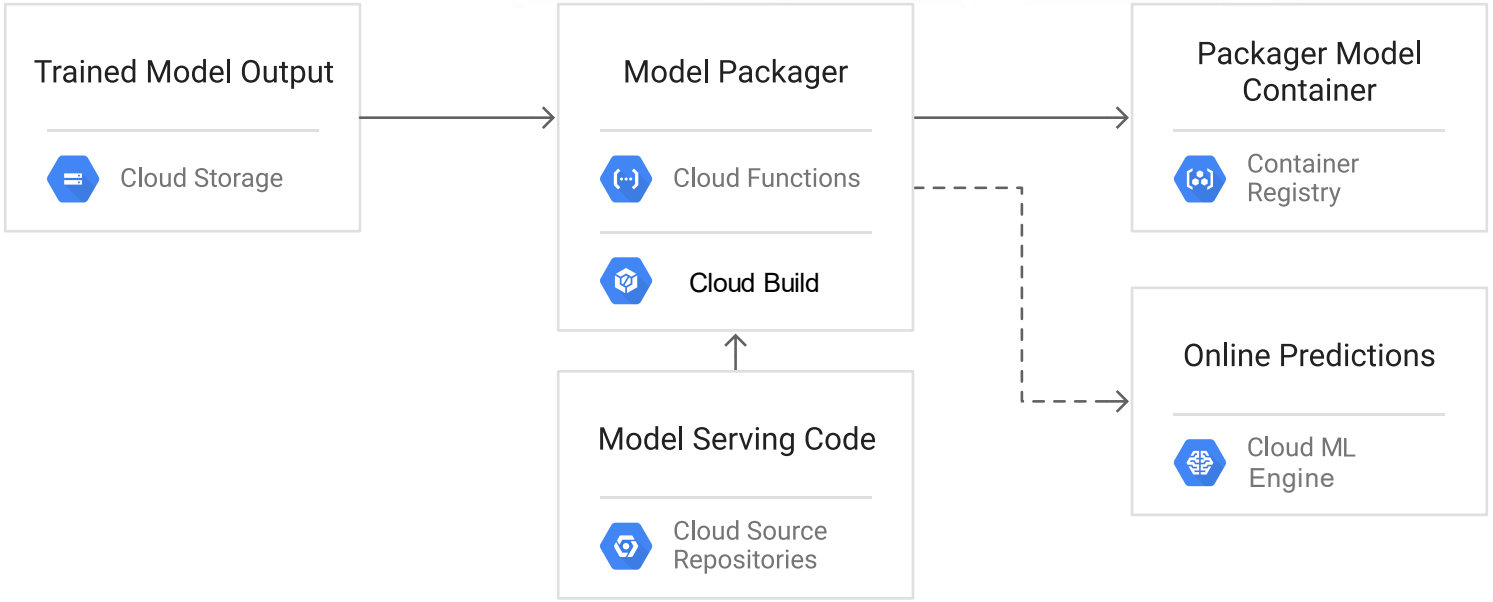


Cloud & Edge Fusion – System Architecture



Cloud & Edge Fusion – Model Training

- Fog: Collecting Data
- Cloud: Rendering & Training Data



Cloud & Edge Fusion – Virtualization (1)

Schema Details Preview

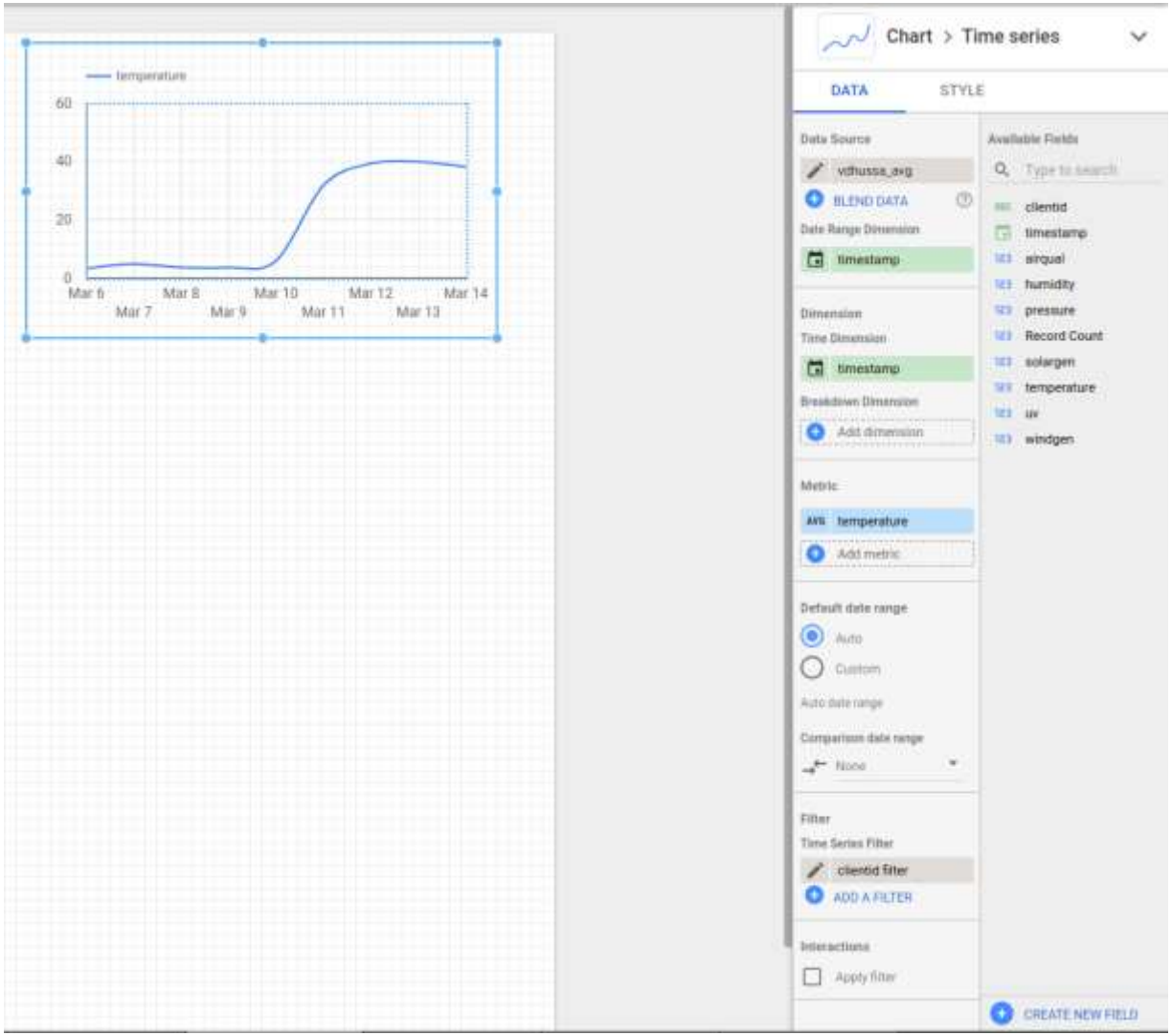
Row	timestamp	clientid	meter	temperature	pressure	lux
1	2018-09-20 00:25:22 UTC	enviro-kit-0123AEA107D0D077F2	0.011679170507812498	26.0	1016.78882812	1639
2	2018-09-26 03:11:28 UTC	enviro-kit-0123AEA107D0D077F2	0.0032490240234375003	27.0	1018.60605468	52
3	2018-09-26 03:11:28 UTC	enviro-kit-0123AEA107D0D077F2	0.0032490240234375003	27.0	1018.60605468	52
4	2018-09-20 00:32:18 UTC	enviro-kit-0123AEA107D0D077F2	0.039799405078125	28.0	1016.87734375	1550
5	2018-09-20 00:32:20 UTC	enviro-kit-0123AEA107D0D077F2	0.0539848353515625	28.0	1016.880625	1553
6	2018-09-20 00:32:31 UTC	enviro-kit-0123AEA107D0D077F2	0.0095739123046875	28.0	1016.88375	1575
7	2018-09-19 00:08:15 UTC	enviro-kit-0123AEA107D0D077F2	0.0393482783203125	28.0	1016.42046875	1601
8	2018-09-20 04:46:56 UTC	enviro-kit-0123AEA107D0D077F2	0.037593896484375	28.0	1014.71710937	1603
9	2018-09-20 04:46:57 UTC	enviro-kit-0123AEA107D0D077F2	0.035588888671875	28.0	1014.71726562	1605
10	2018-09-20 04:47:15 UTC	enviro-kit-0123AEA107D0D077F2	0.035388387890625	28.0	1014.7051562500001	1606
11	2018-09-20 04:47:25 UTC	enviro-kit-0123AEA107D0D077F2	0.035187887109374996	28.0	1014.71382812	1608
12	2018-09-19 00:07:29 UTC	enviro-kit-0123AEA107D0D077F2	0.0335337556640625	28.0	1016.41050781	1610
13	2018-09-20 04:50:51 UTC	enviro-kit-0123AEA107D0D077F2	0.0362405162109375	28.0	1014.69402343	1612

Cloud & Edge Fusion – Virtualization (2)

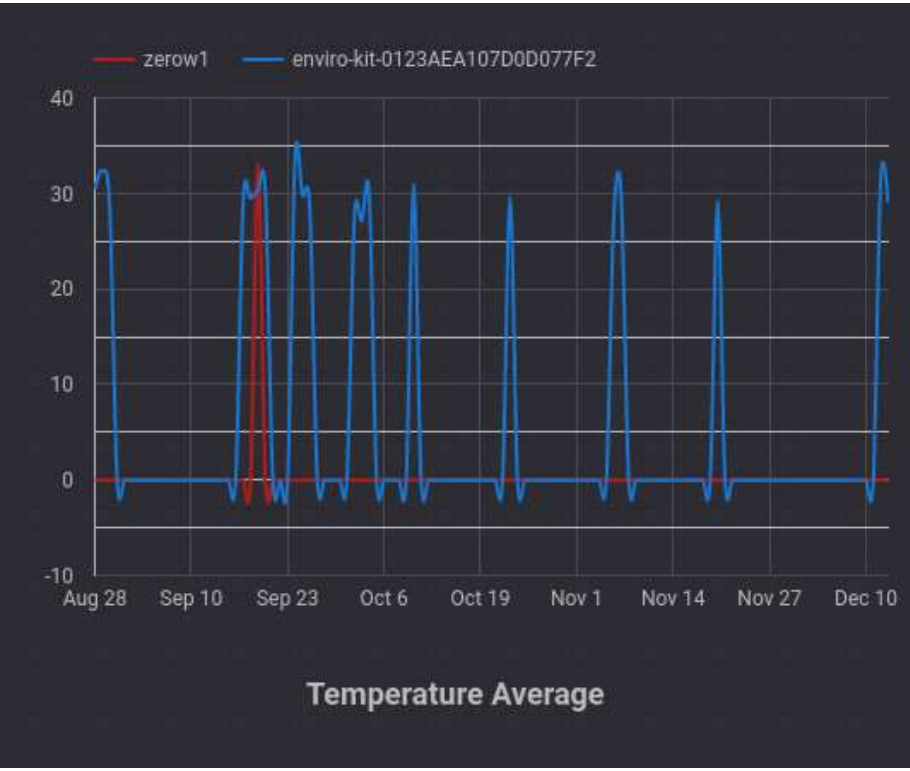
EDIT CONNECTION

Index	Field	Type		Aggregation	Description
1	Record Count	123	Number	Auto	
2	timestamp		Date (YYYYMMDD)	None	
3	clientid	ABC	Text	None	
4	temperature	123	Number	None	
5	pressure	123	Number	None	
6	humidity	123	Number	None	
7	uv	123	Number	None	
8	airqual	123	Number	None	
9	windgen	123	Number	None	
10	solargen	123	Number	None	

Cloud & Edge Fusion – Virtualization (3)

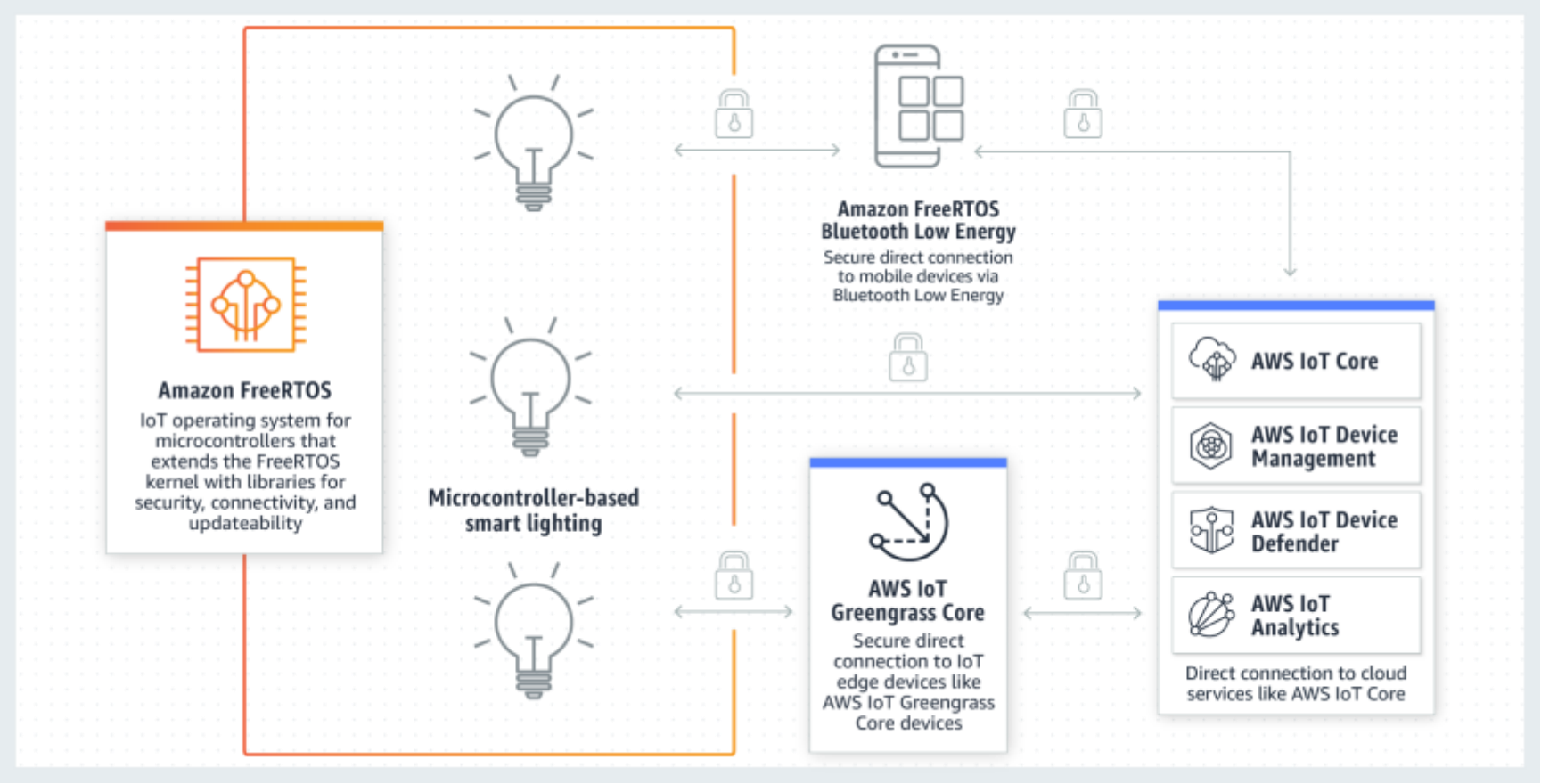


Cloud & Edge Fusion – Virtualization (4)

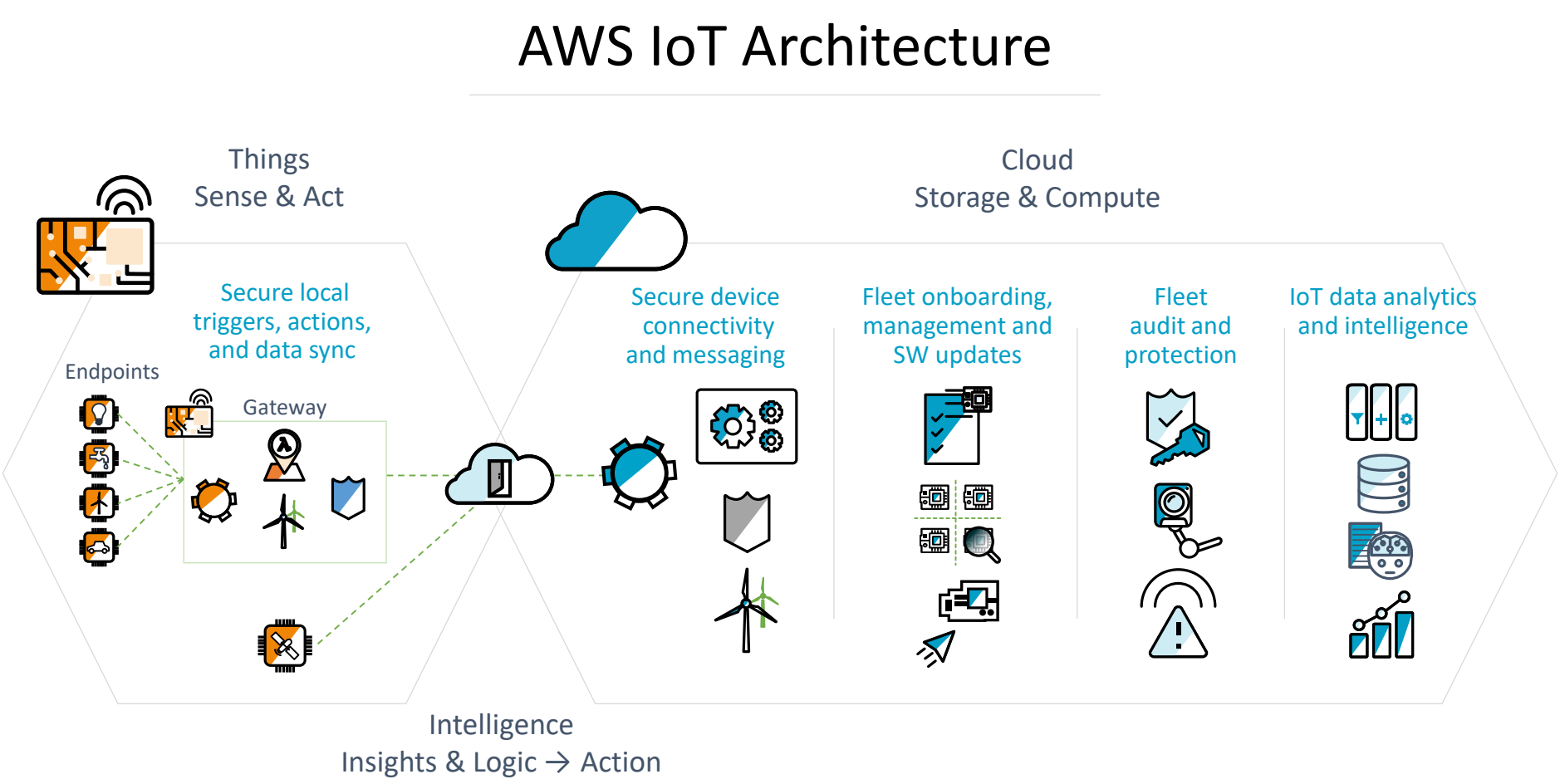


Amazon AWS IoT Solution

- FreeRTOS: IoT operating system
- Greengrass： Seamless expansion to edge devices

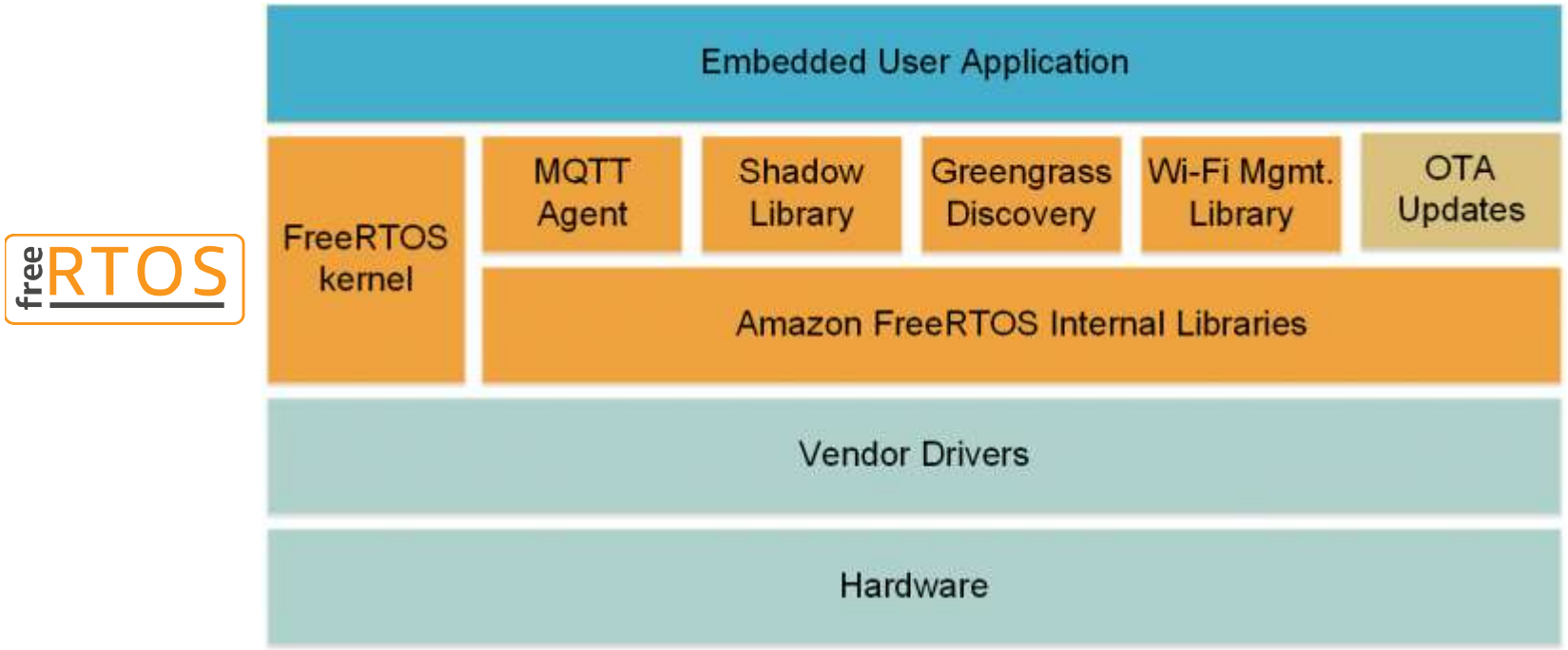


Amazon AWS IoT Architecture



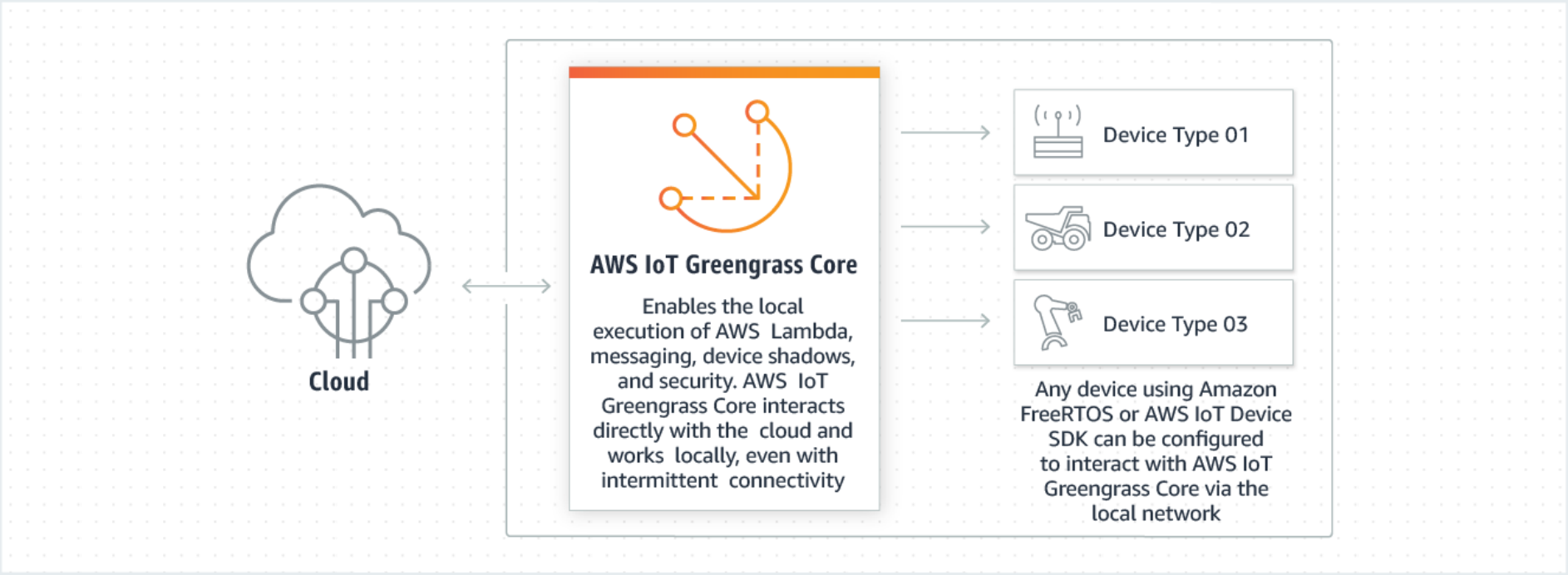
Free RTOS - an open source IoT OS

- FreeRTOS: https://www.freertos.org/



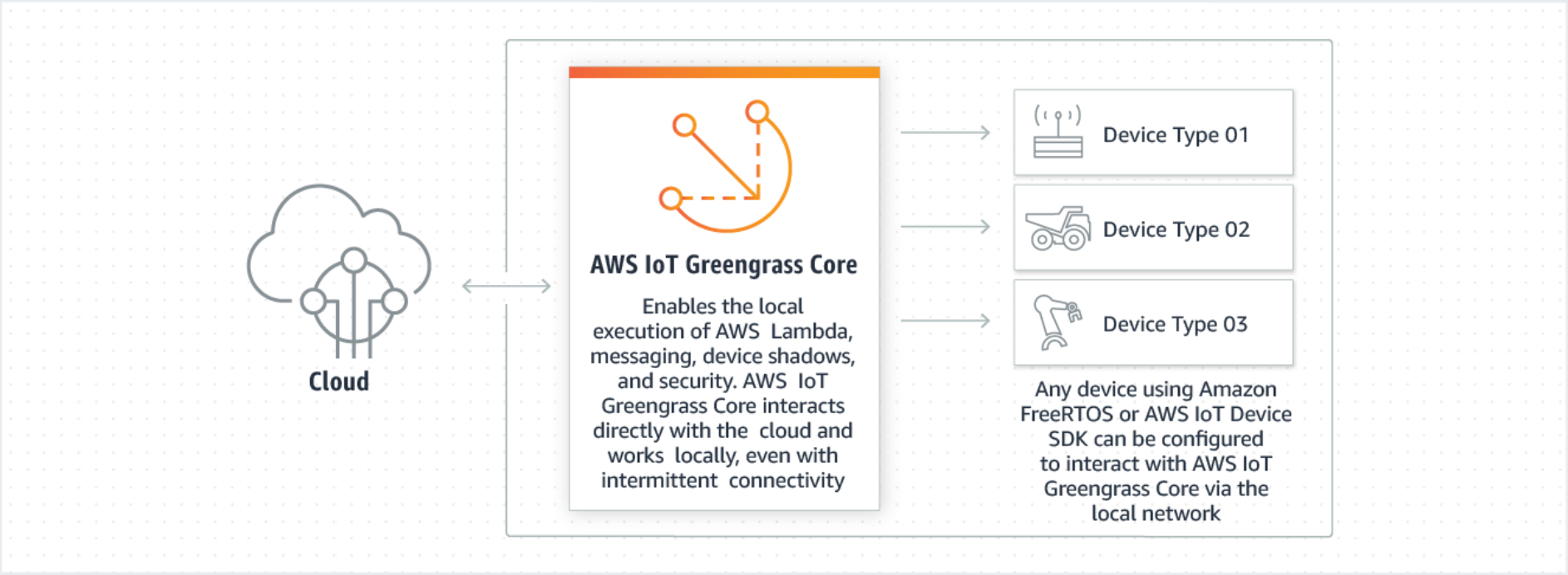
Greengrass – AWS Edge Computing Platform (1)

- Greengrass is an edge/ fog node with certain computing and processing capability in AWS



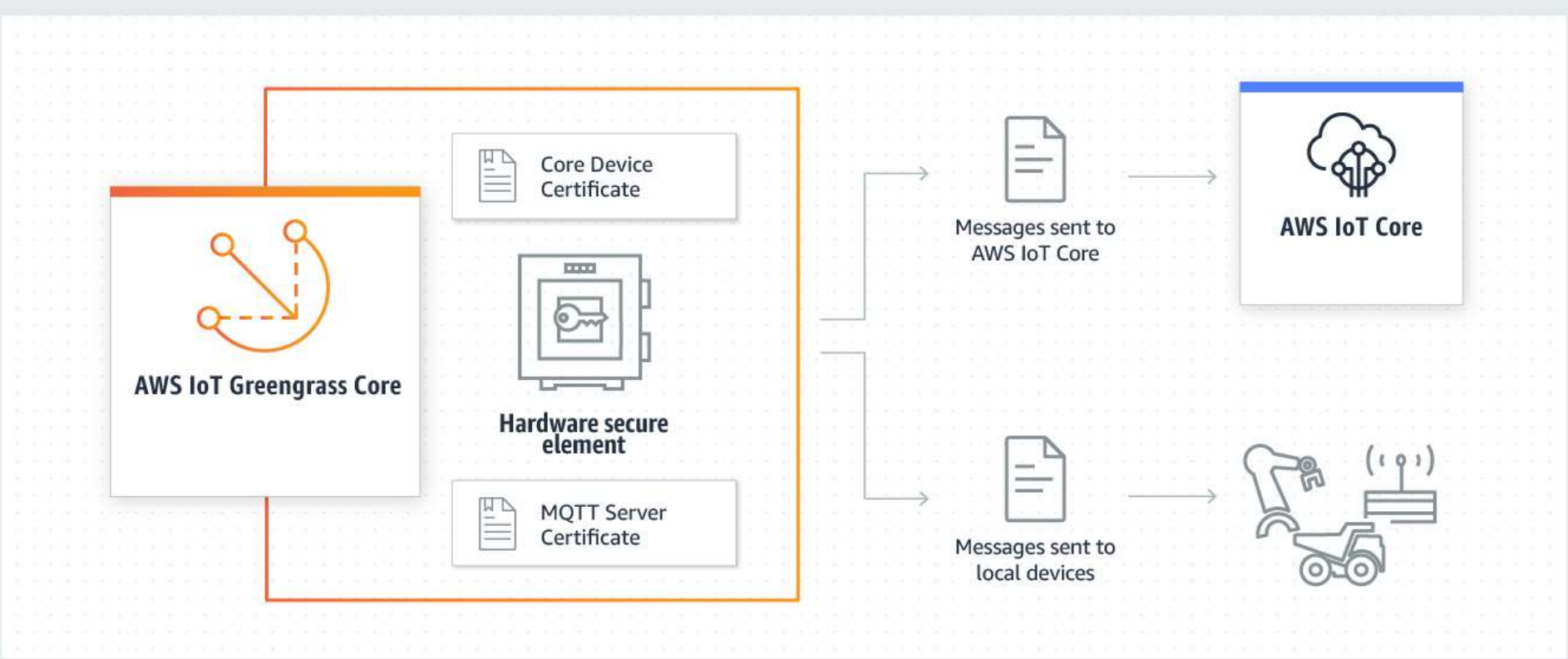
Greengrass – AWS Edge Computing Platform (2)

- Greengrass provides connector, connecting edge-fog-cloud nodes, and realizing adaptive configuration

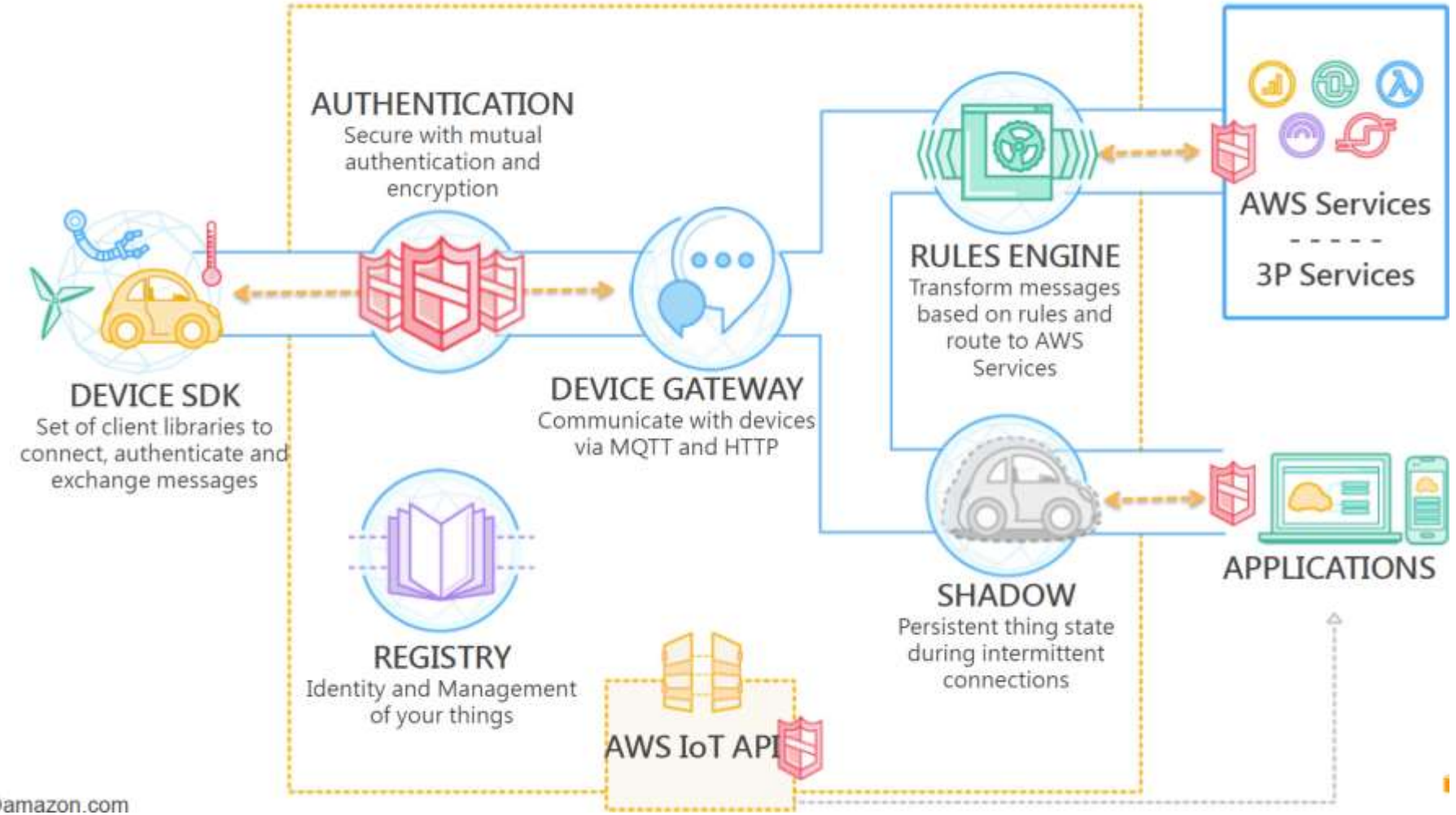


Greengrass – AWS Edge Computing Platform (3)

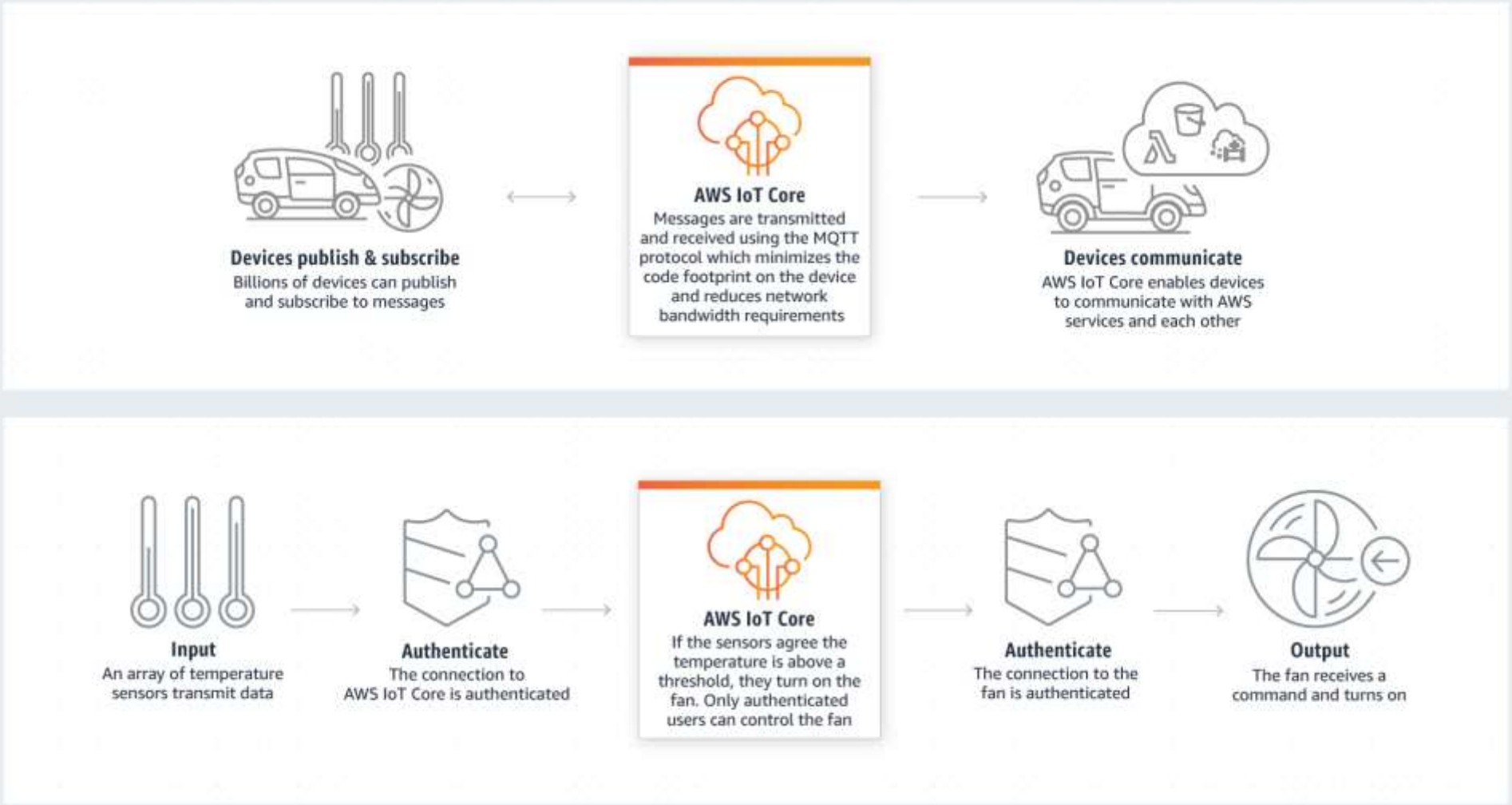
- Greengrass provides good authorization and privacy protection mechanisms



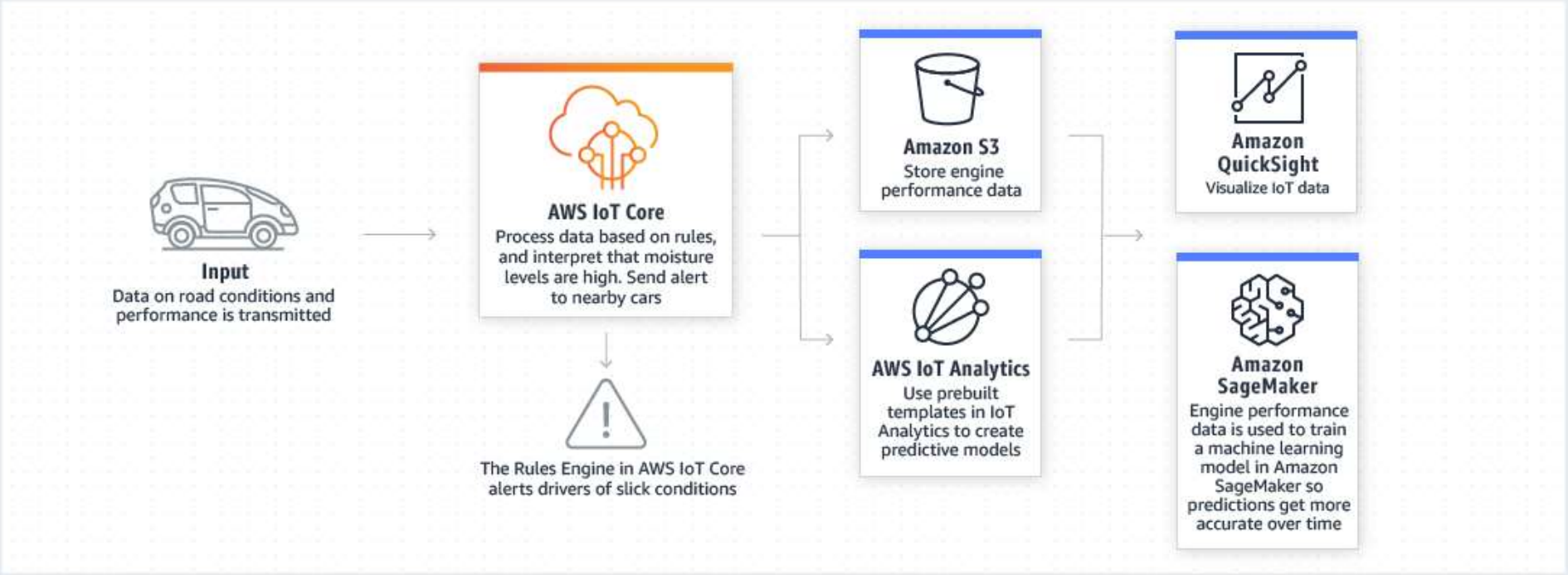
Greengrass – AWS Edge Computing Platform (4)

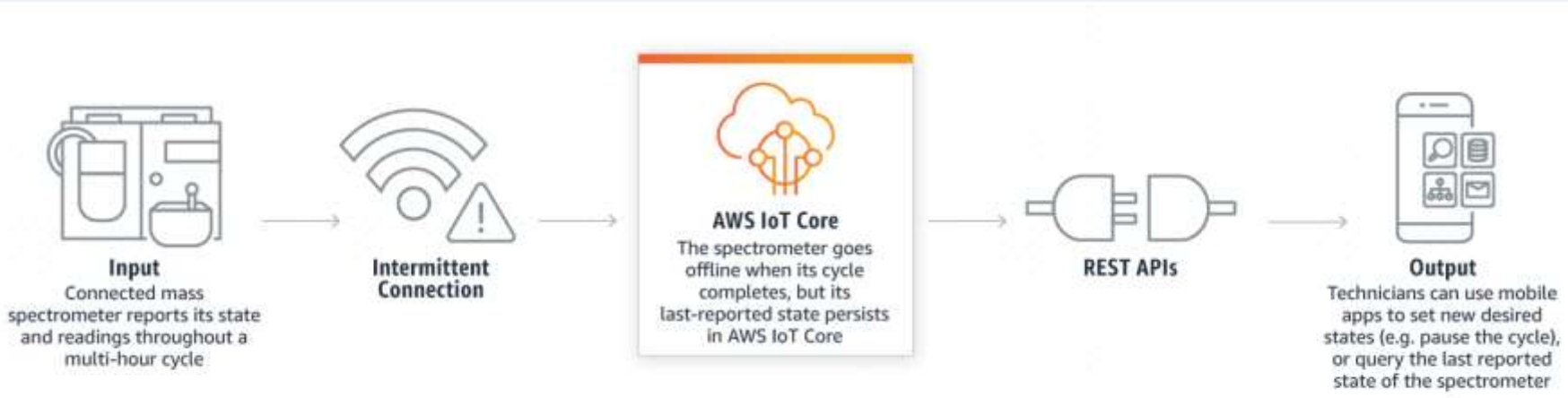


AWS IoT Core – Edge node (1)

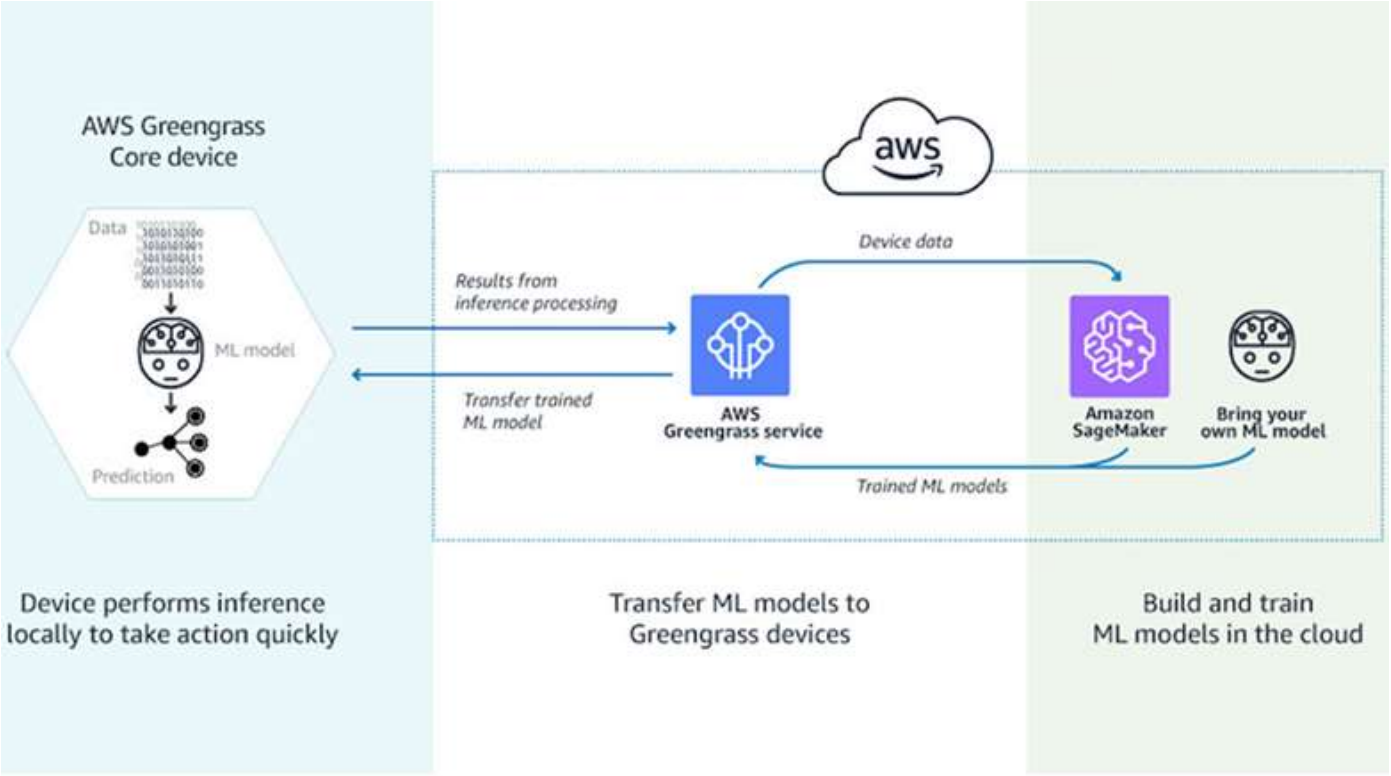


AWS IoT Core – Edge node (2)

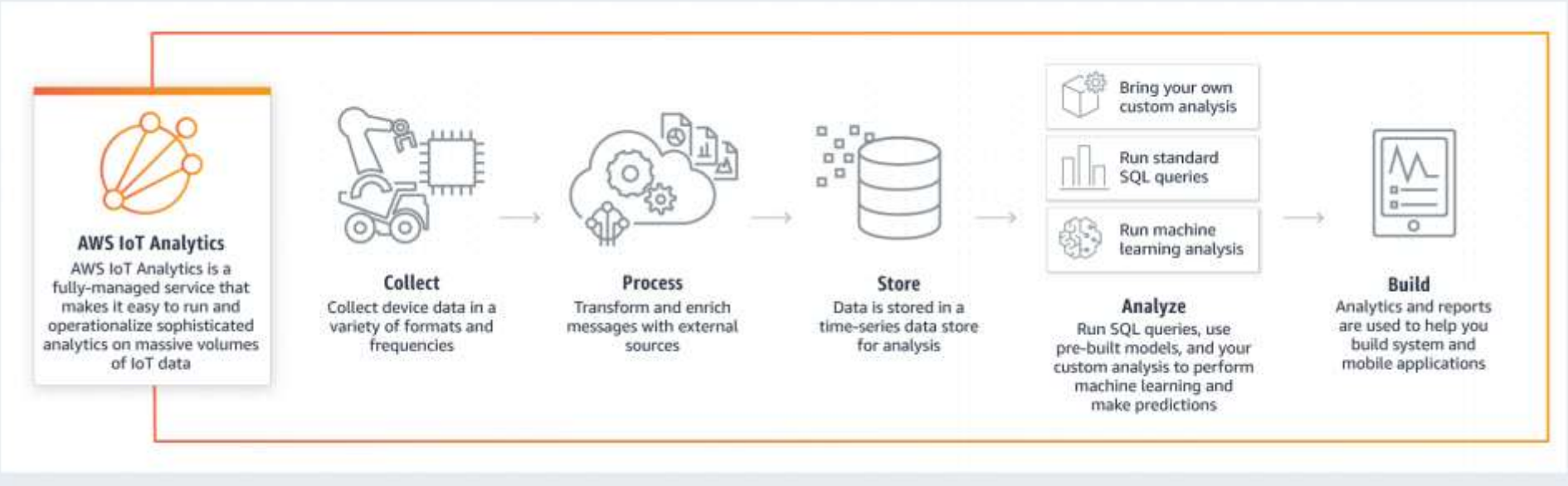




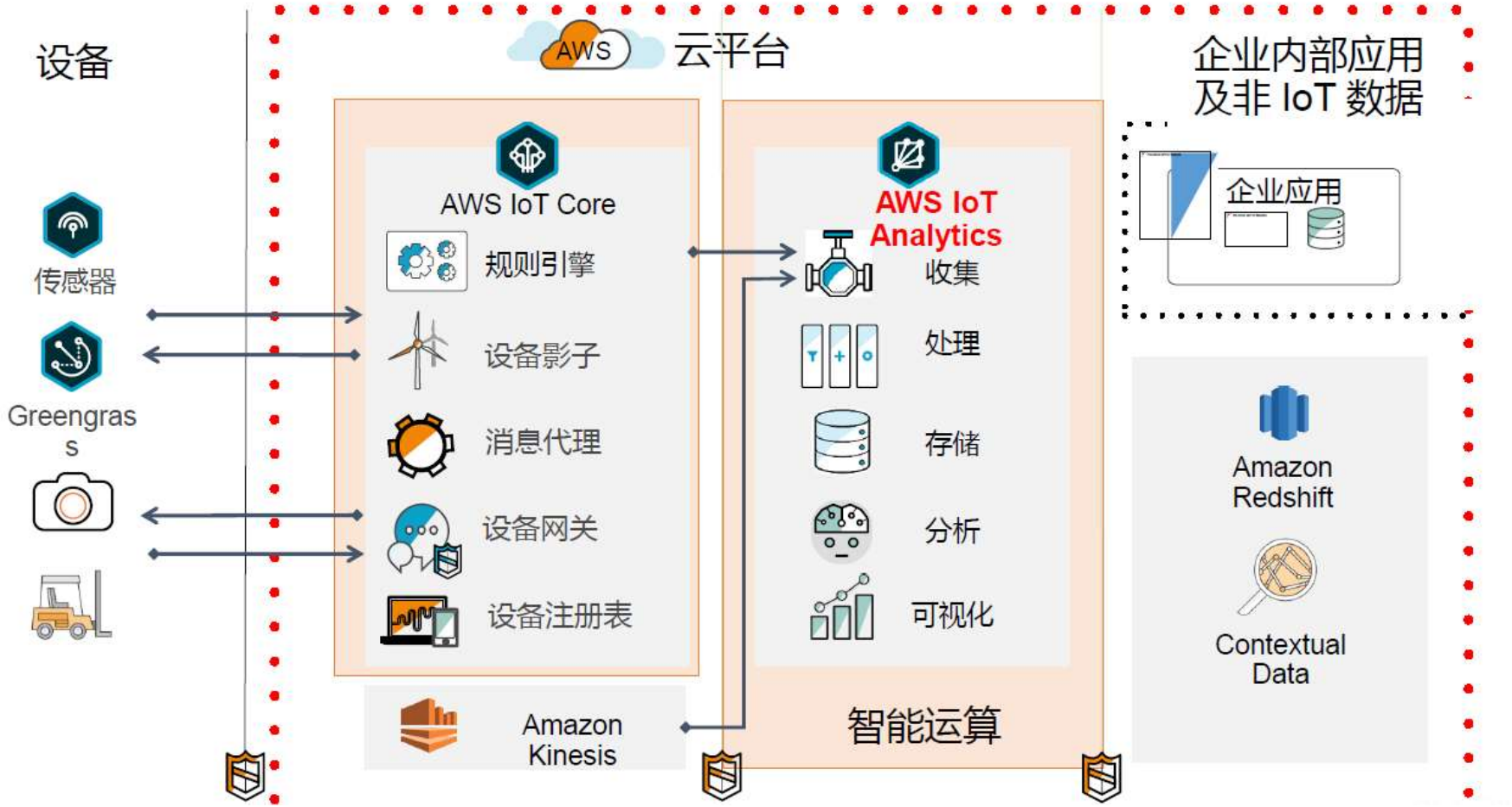
Greengrass → IoT Analytics



AWS IoT Analytics (1)



AWS IoT Analytics (2)

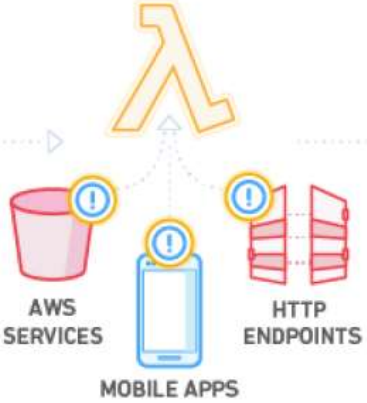


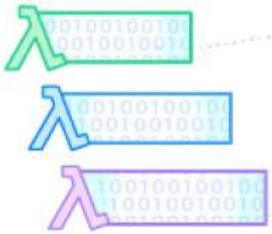
AWS Lambda (1)

- AWS lambda is a fine-grained method for deploying code, managing services, and monitoring the health of lightweight services. similar to Alibaba microservice.

AWS Lambda (2)

- AWS lambda is a new pricing and service model





开发过程中无须担心底层服务器管理

针对一系列的触发事件编写响应的处理程序代码

按需执行代码，持续的扩展性能

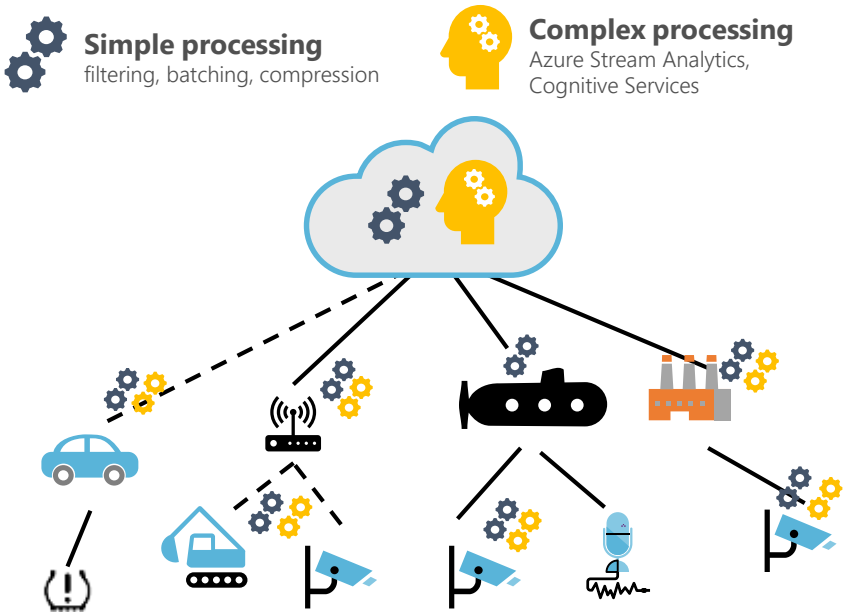
仅仅为调用次数付费

AWS Lambda (3)


Amazon Kinesis
Social media stream is loaded into Kinesis in real-time
Lambda is triggered
AWS Lambda
Lambda runs code that generates hashtag trend data
Amazon DynamoDB
The hashtag trend data is stored in DynamoDB
Social media trend data immediately available for business users to query


Online order is placed
Amazon DynamoDB
Order data is stored in an operational database
Lambda is triggered
AWS Lambda
Lambda runs data transformation code
Amazon Redshift
Lambda loads results into data warehouse
Analytics generated from data

Microsoft IoT Solution



Secure

Provides a secure connection to the Azure IoT Edge, update software/firmware/configuration remotely, collect state and telemetry and monitor security of the device

Cloud managed

Enables rich management of Azure IoT Edge from Azure provide a complete solution instead of just an SDK

Cross-platform

Enables Azure IoT Edge to target the most popular edge operating systems, such as Windows and Linux

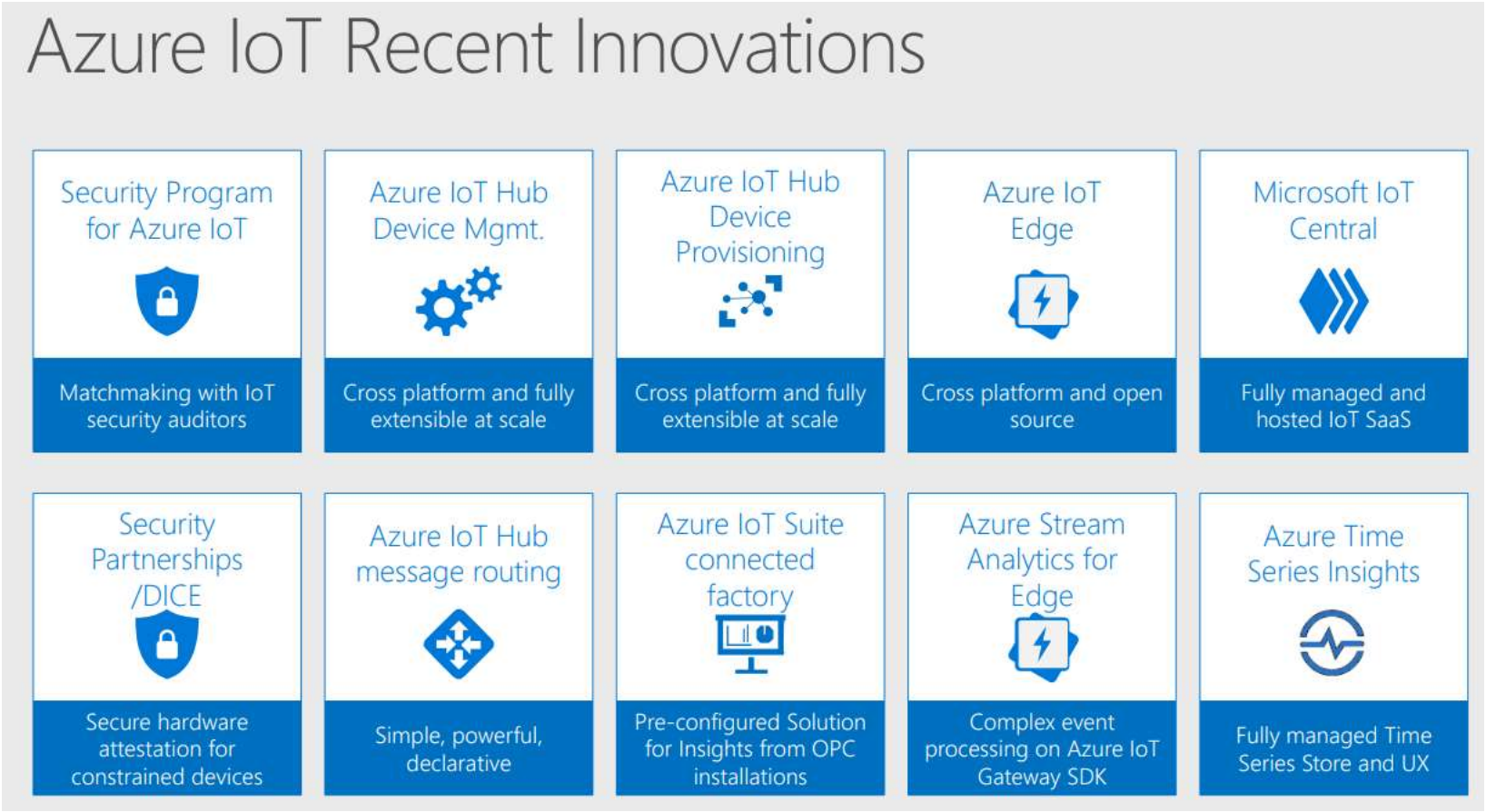
Portable

Enables Dev/Test of edge workloads in the cloud with later deployment to the edge as part of a continuous integration / continuous deployment pipeline

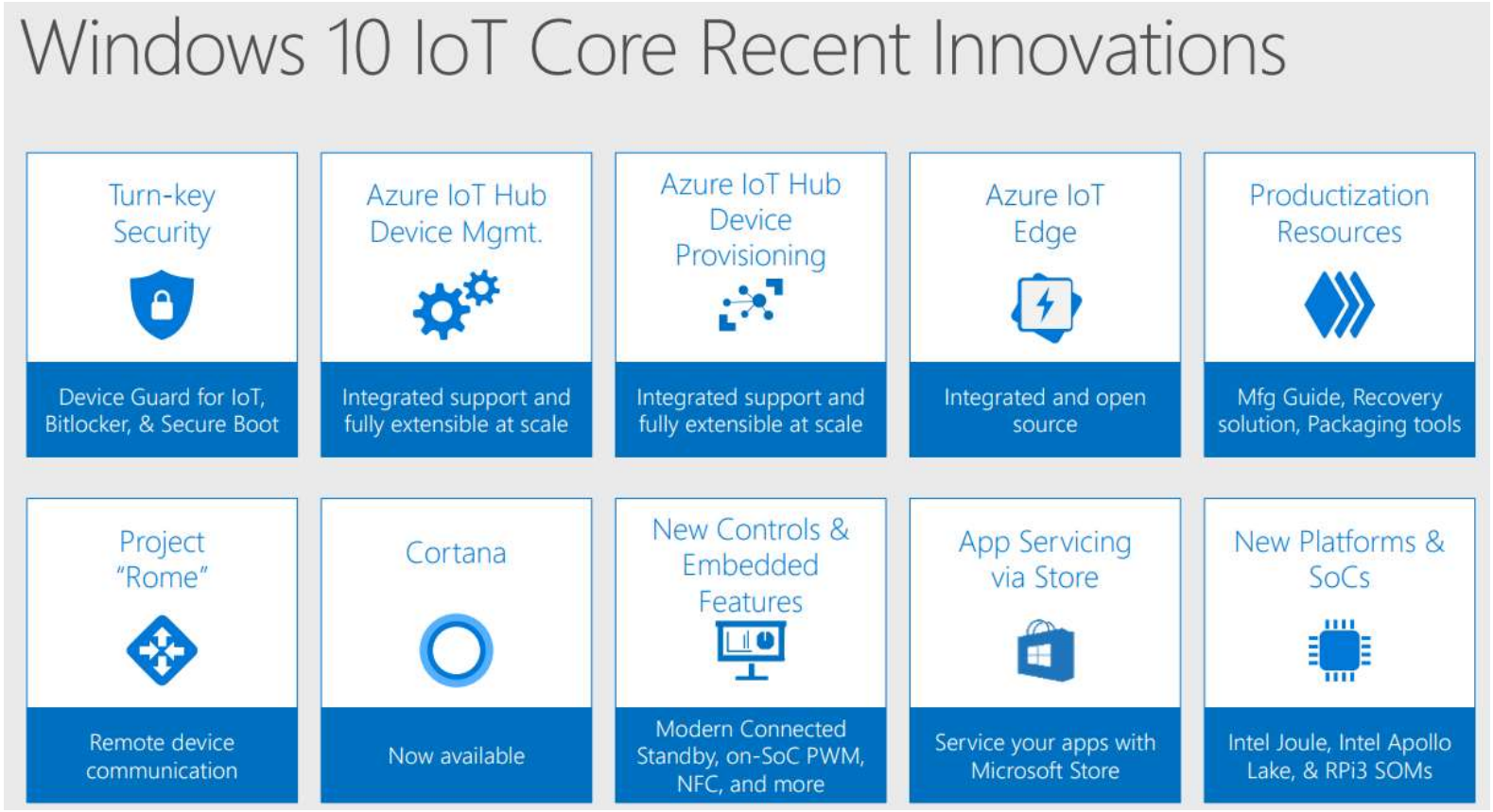
Extensible

Enables seamless deployment of advanced capabilities such as AI from Microsoft, and any third party, today and tomorrow

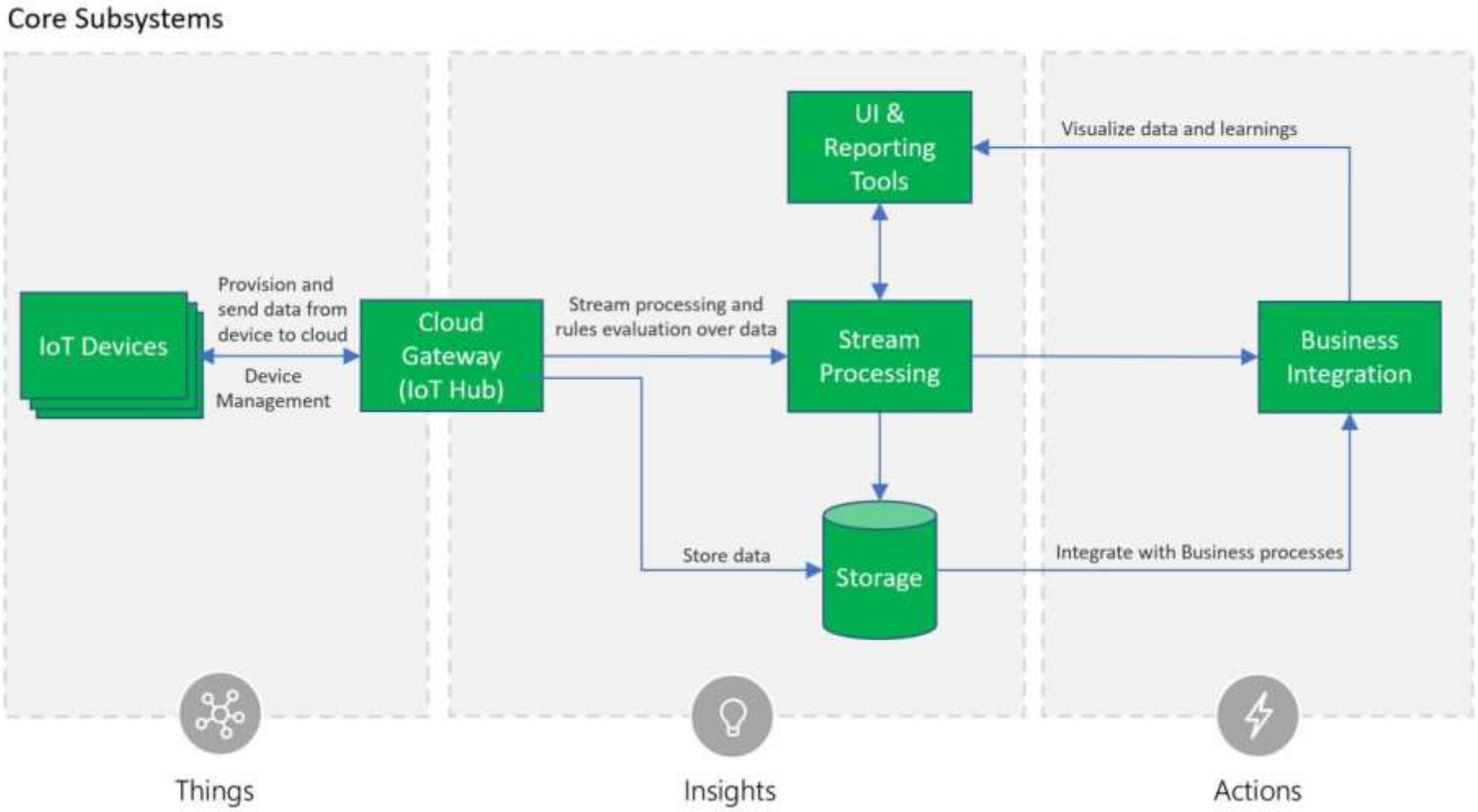
Microsoft IoT Core Innovations (1)



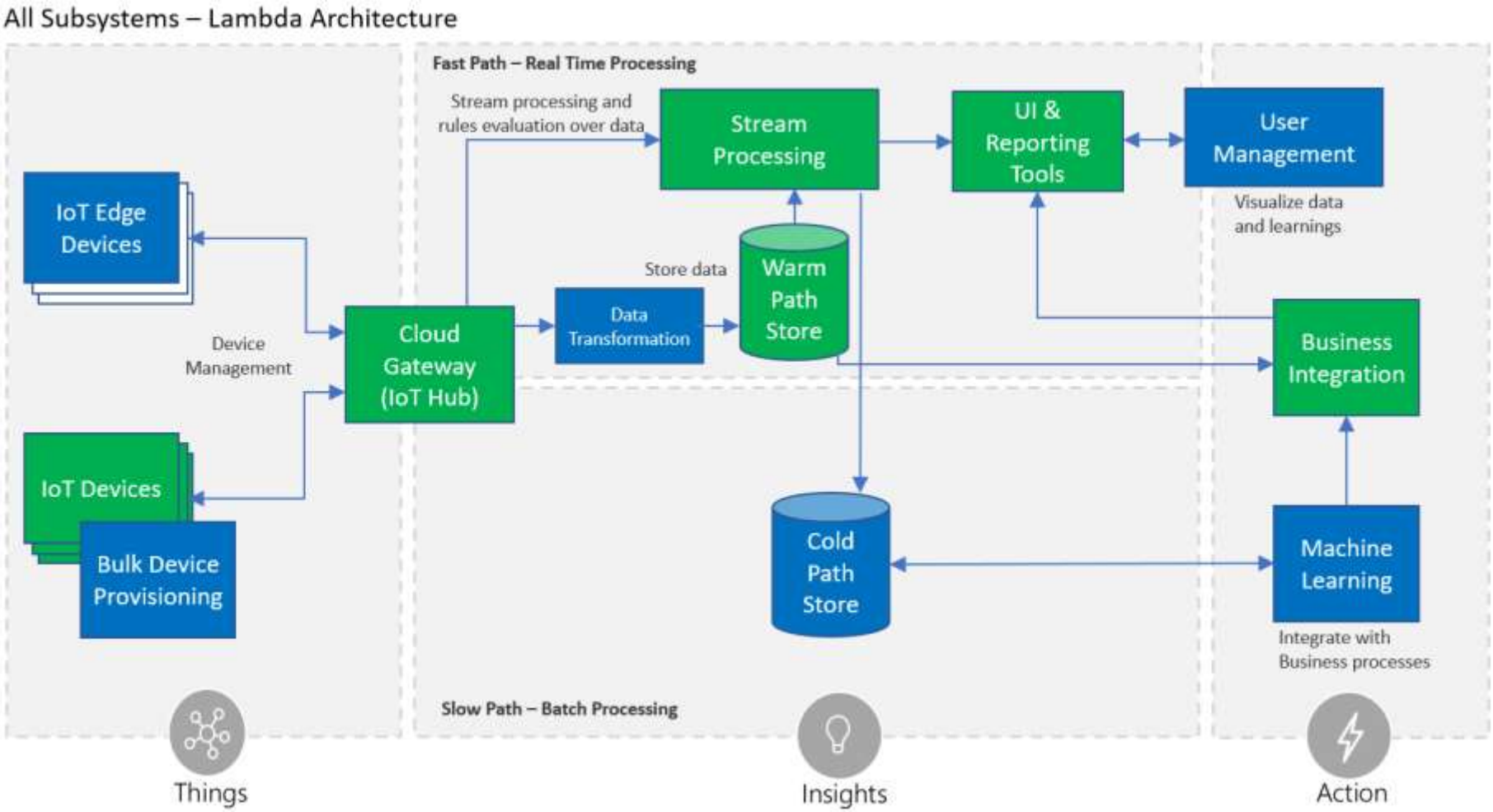
Microsoft IoT Core Innovations (2)



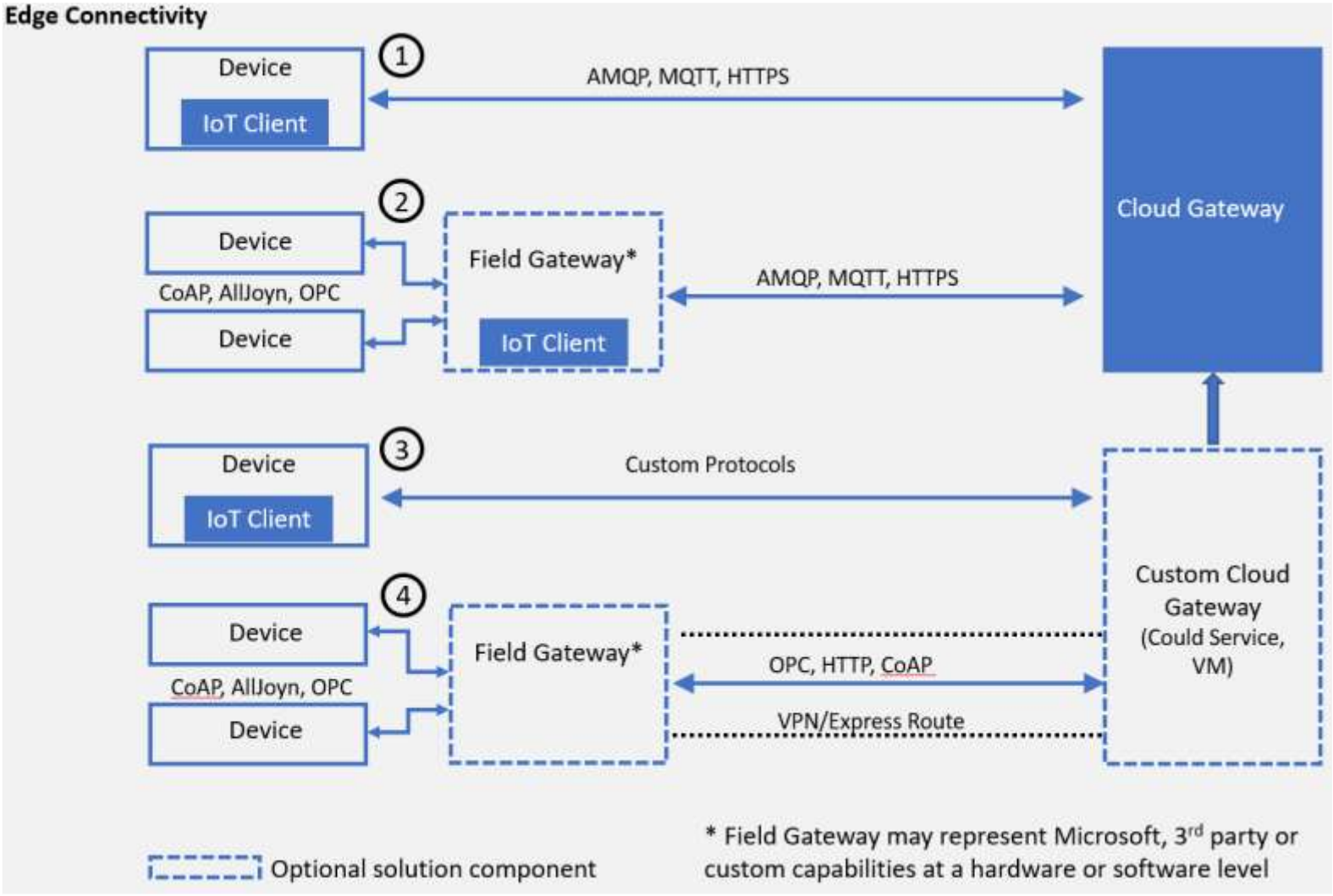
Microsoft IoT Processing Procedure



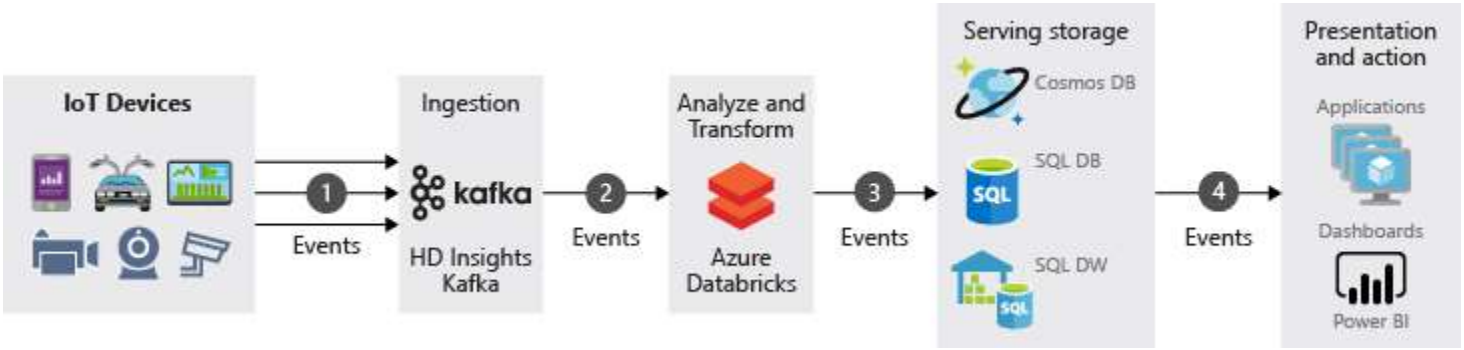
Microsoft IoT Intelligent Processing Lambda Architecture

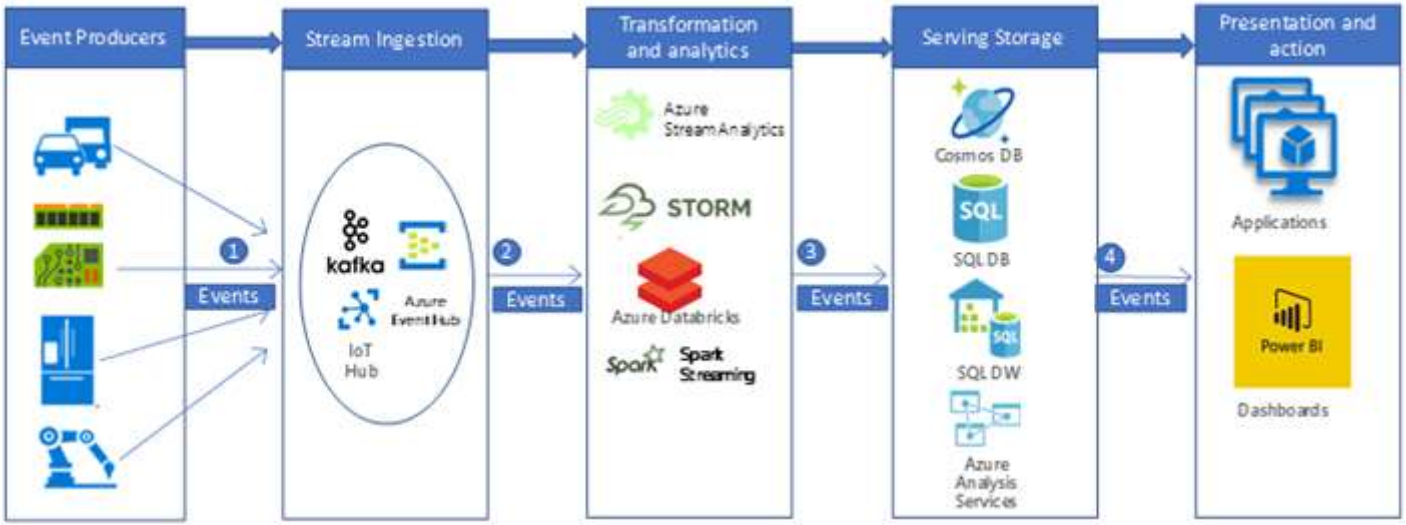


Microsoft IoT Connection Procedure

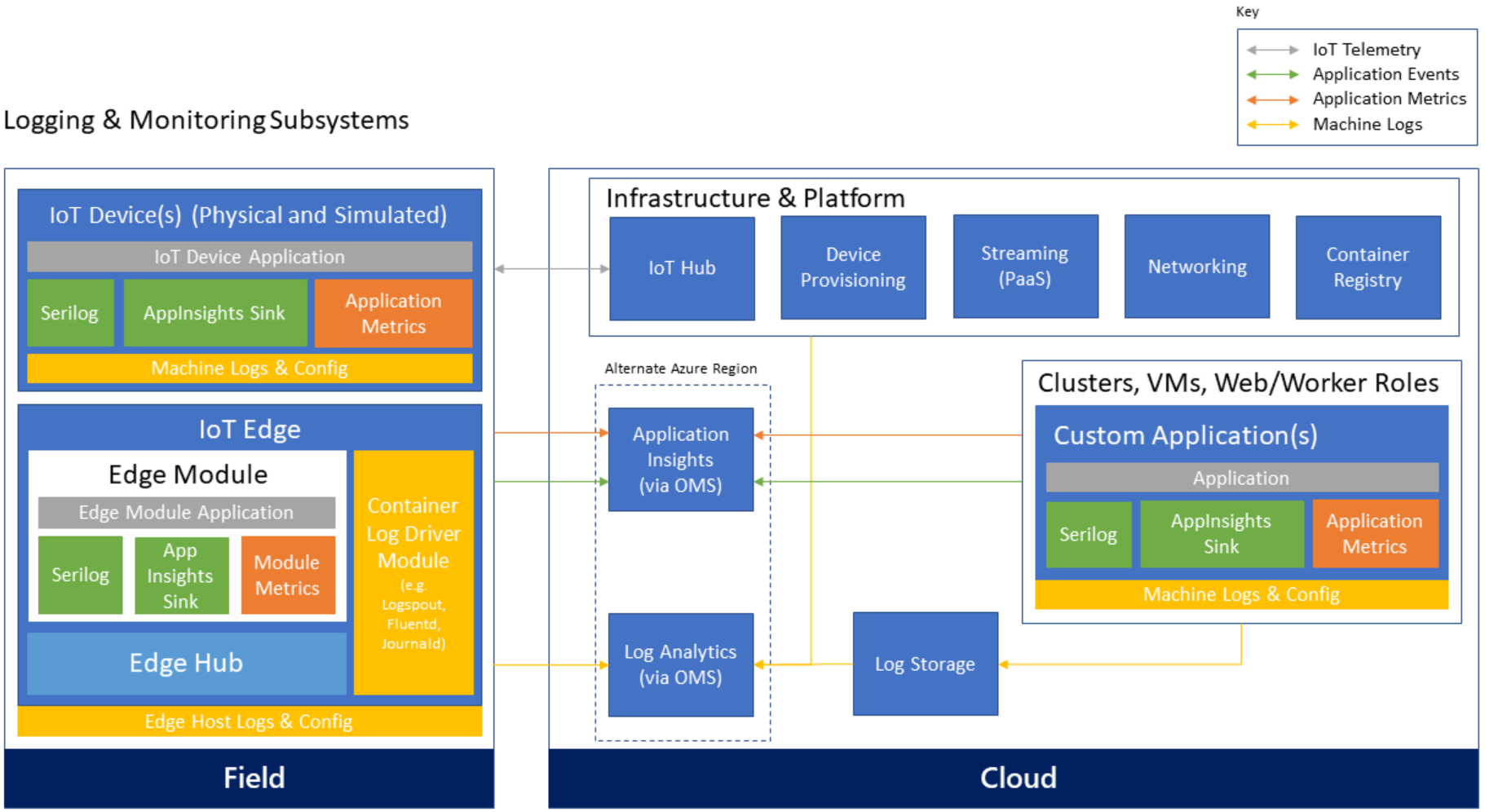


Microsoft IoT Edge-Fog-Cloud Fusion (1)





Microsoft IoT Edge-Fog-Cloud Fusion (2)



Thank you!