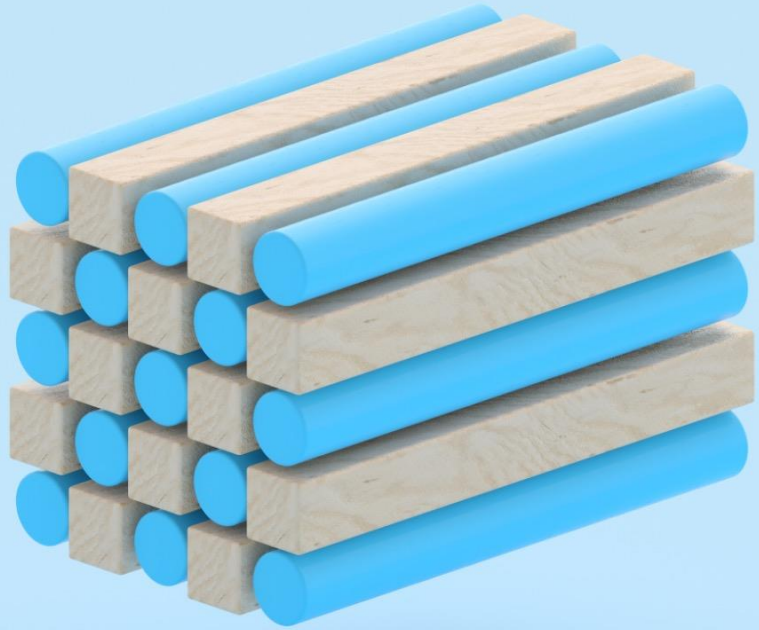
Next Generation Data Management Capabilities

Data Management at the Edge

SDC India

Roopesh Chuggani
Senior Engineer, File System Group, NetApp
5th Dec 2020

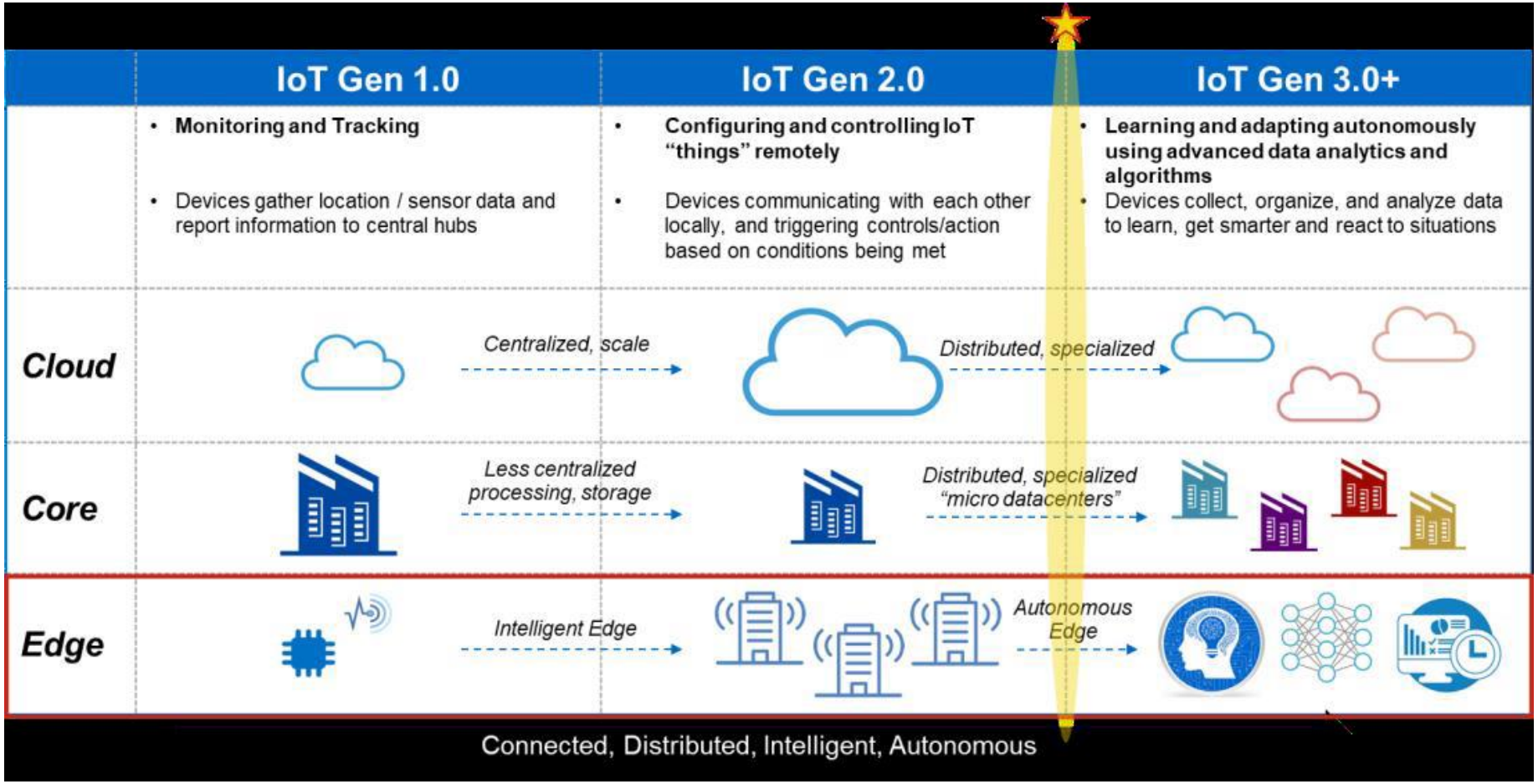


AGENDA

- IoT Landscape
- Edge Ecosystem
- Data Management Capabilities
- Key Takeaways

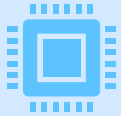
IoT Industry Landscape

Evolving Landscape of IoT



Emerging IoT Verticals

Areas of Opportunity



Industrial IoT

Real-time plant and field condition monitoring
Data analytics to determine enterprise best practices
Data and plant security, backup and recovery

Video Surveillance

Camera with intelligent analytics
Next-generation video management systems (VMS)

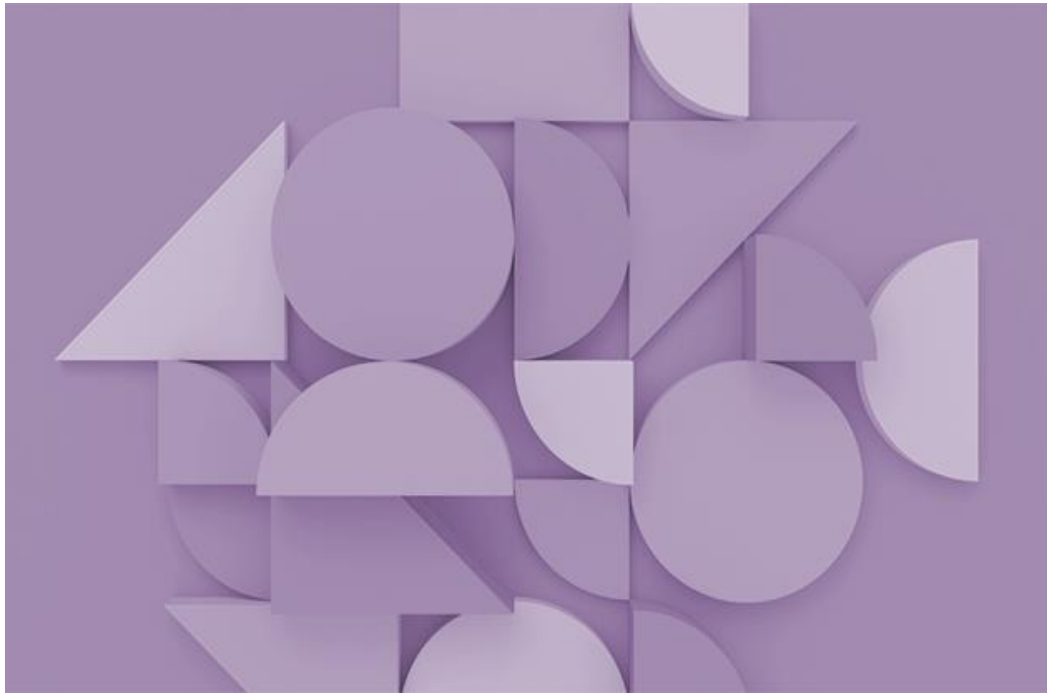
Autonomous Vehicles

Simulation and testing (including data mule)
In-Vehicle Entertainment

Edge Ecosystem

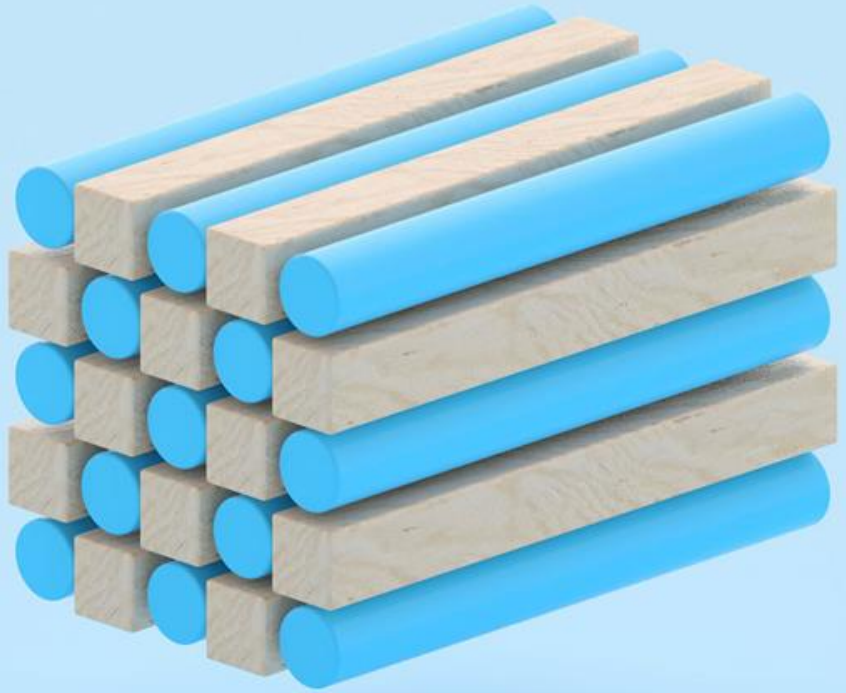
Hardware and Software Components

- Resource constrained environment
  - Few KB to Few GB of RAM
  - Limited Storage
- Power consumption is a key consideration
- Emergence of Container & Kubernetes
- Move towards microservices based architecture
- Standardization still a key challenge
- Increasing usage of open-source tools



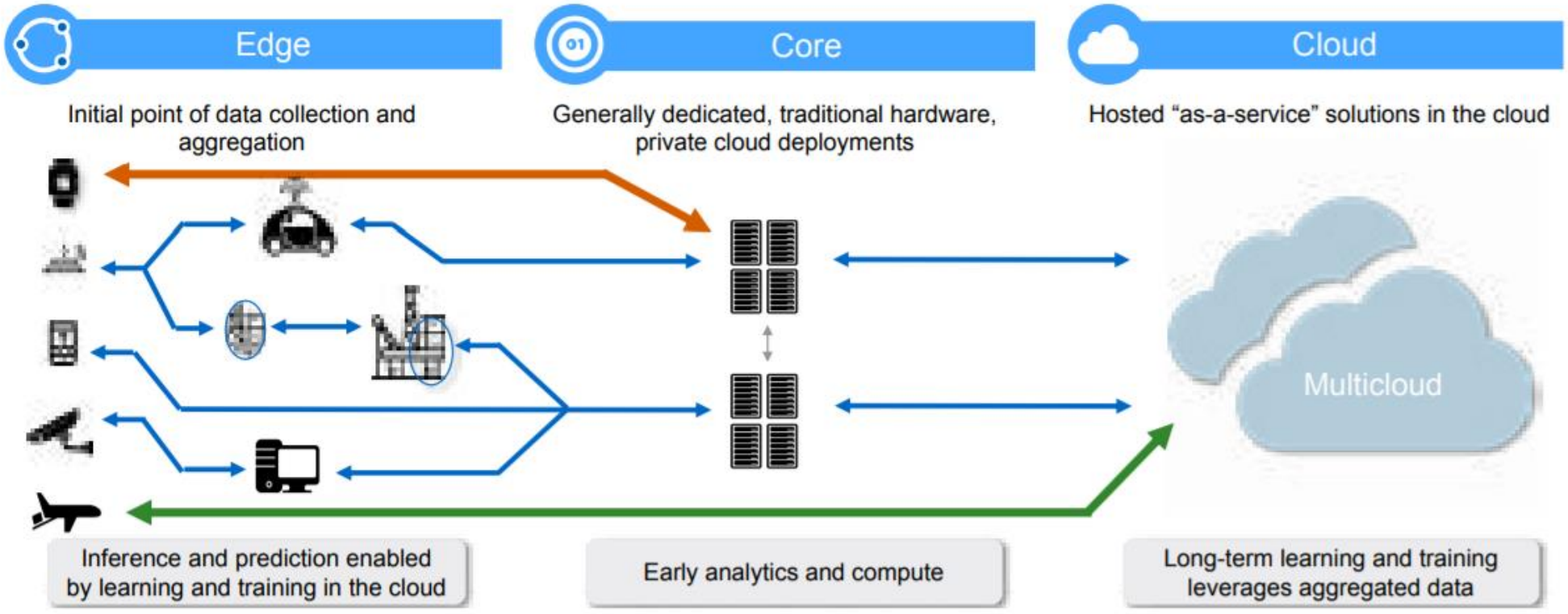
Data Management Capabilities at Edge

- Intelligent Data Mobility
- Data Security
- Storage Efficiency
- Autonomous Management



Intelligent Data Mobility

Edge to Core to Cloud Data Pipeline



Source: PwC research, TechCrunch, Andreseen Horowitz, Gartner, Mary Meeker Tech Trends

Data Characteristics

How are the edge data patterns different?

Ingest
- Normal-compute
- Write-heavy
- Sequential
- Small to large files

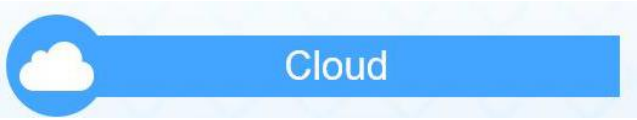
Inference
- Inference GPU
- Mostly in-memory reads
- Sequential reads

Transform
- Normal Compute
- Read/write-heavy
- Random/sequential
- Small to large files

Explore
- Compute-intensive
- Read-heavy
- Random
- Small to large files

Train
- Compute-intensive
- Read/write-heavy
- Random
- Small-files

Inference
- Compute-intensive
- Read-heavy



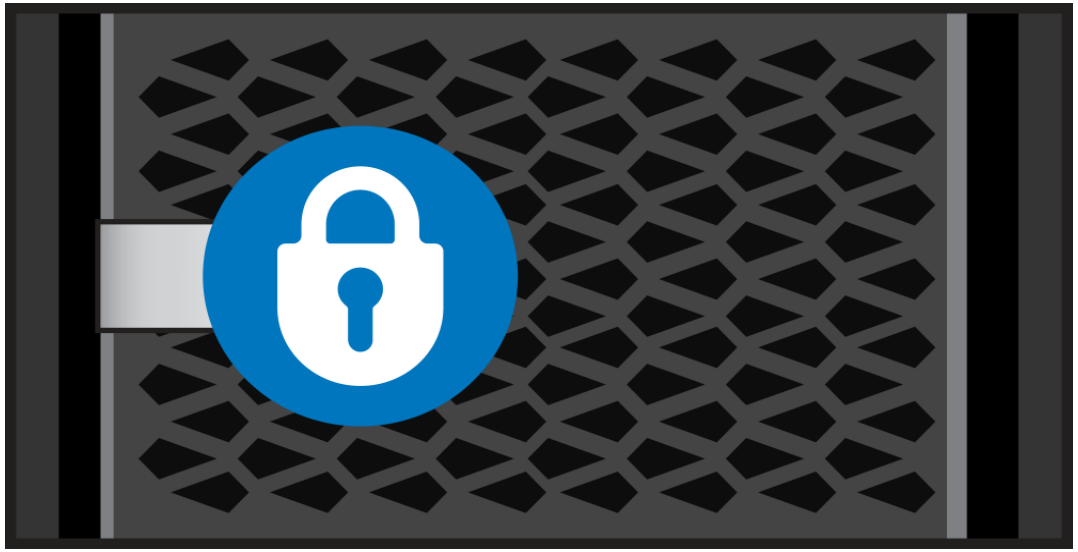
Archive/Backup
- Low-compute
- Write-heavy
- Sequential
- Small to large files

https://www.netapp.com/pdf.html?item=/media/2129-wp-7271pdf.pdf

Data Security

Capabilities & Challenges

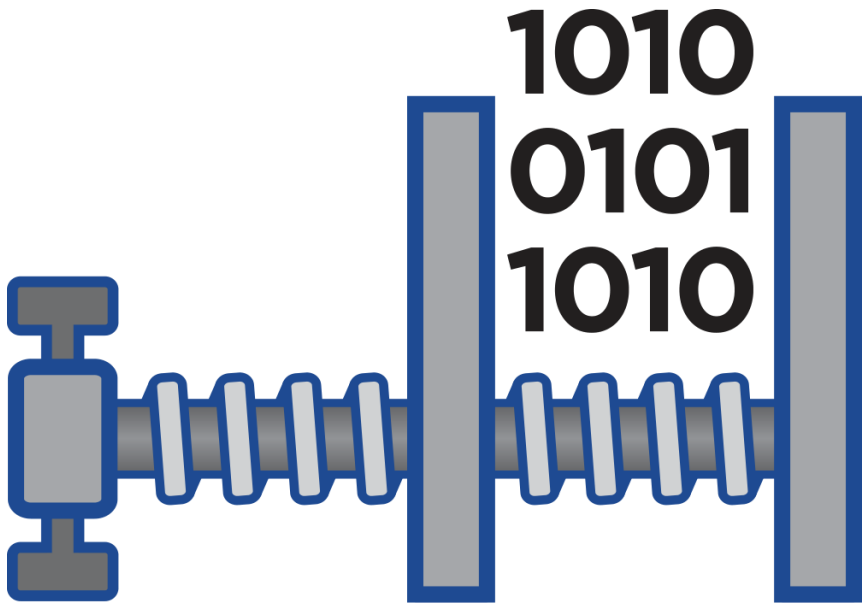
- Capabilities required
  - Data-at-rest Encryption
  - Data-in-motion Encryption
  - Data Governance
- Challenges
  - Power Efficiency
  - Device Identification
  - Key management for large number of devices



Storage Efficiency

Rethinking Needed

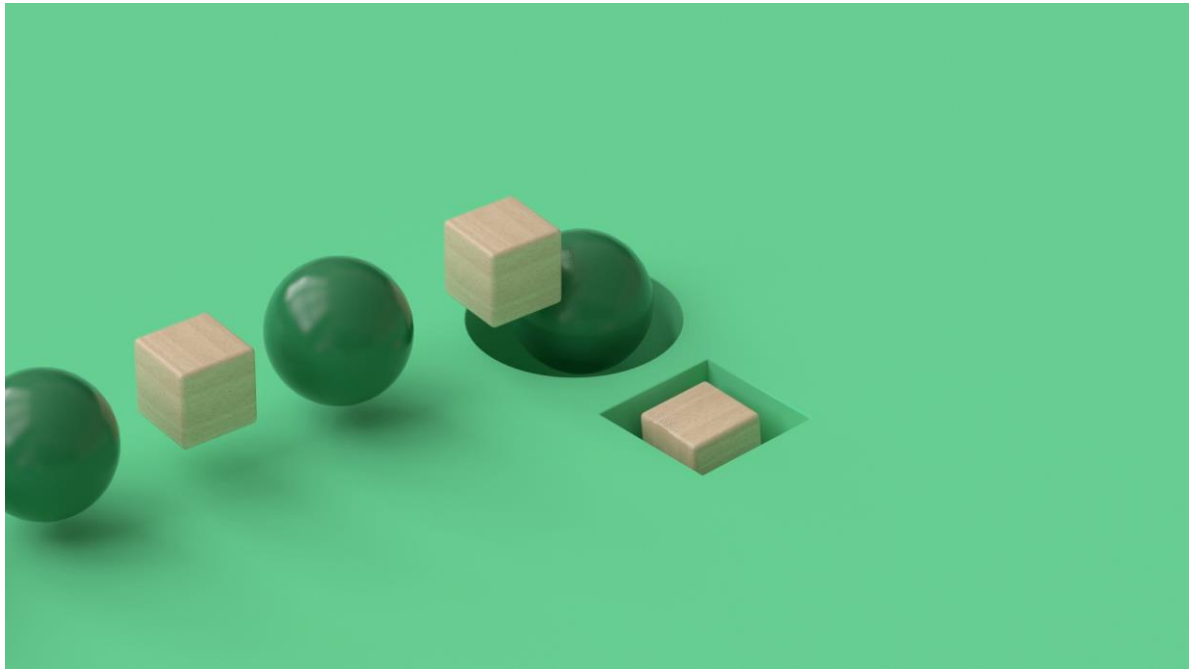
- Traditional compression techniques are not suitable for data reduction at edge
- Edge data has little to no duplication
- Emergence of application aware compression techniques such as LASzip for Lidar data, time series data specific databases



Autonomous Management

No Human intervention required!

- Automated Configuration
- Self Diagnosis/Healing Capabilities
  - Learning from past events
- OTA incremental upgrades
- Container & Kubernetes
- Usage of Pub/Sub Architecture
- Quality of Service



Key Takeaways

- Most data would be generated and consumed at the edge with limited data/results being transmitted to core and cloud
- Next Gen IoT architecture would be driven by distributed compute models & microservices
- Edge ecosystem is very different from a traditional datacenter
- Highly desirable data management capabilities include intelligent data mobility, security, efficiency and autonomous management

Thank You