



Challenges in IoT Data Management

Vipin Shankar, Calsoft

Speaker

Nitin Singhvi

Vipin Shankar

Sponsor

Calsoft Pvt. Ltd.

Today's workload !!!



Where are we (and the Market) headed?

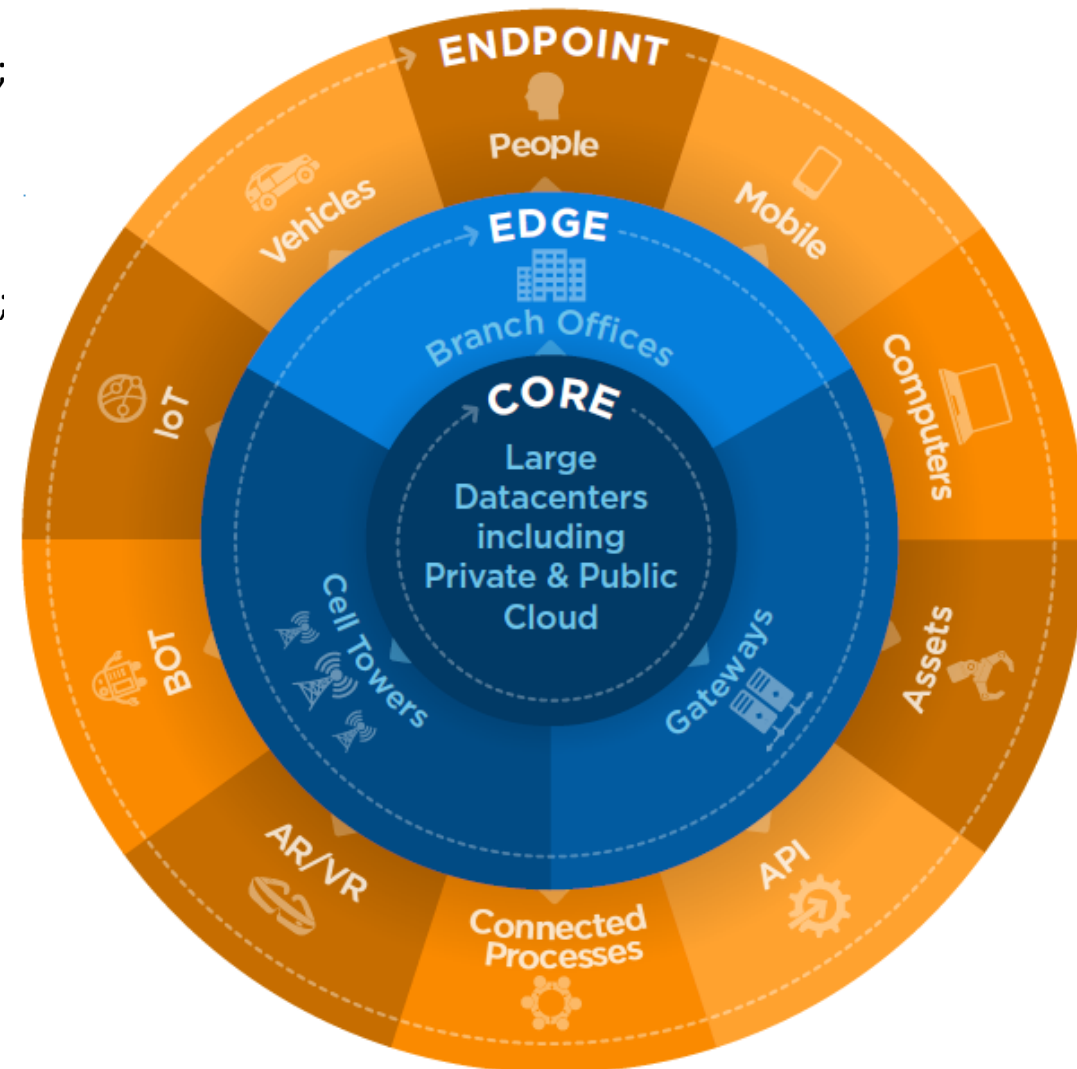


Global DataSphere to Hit 175 Zettabytes by 2025, IDC Says

- The storage industry will ship 42ZB of capacity over the next seven years;
- 90 ZB of data will be created on IoT devices by 2025;
- 49% of data will be stored in public cloud environments by 2025;
- Nearly 30% of the data generated will be consumed in real-time by 2025;

Endpoint Data Sources: IoT Application areas

- Smart Refrigerators/ Thermostats / IP Cameras
- Factory shop floors
- Connected cars and Drones
- Agriculture / Irrigation systems
- Drilling (pressure of gas, water flow, gas flow, sampling, rotational sensors, etc.)
- Smart street lights



Simply put ... Mostly Time Series Data !

- **Sensor Data** - Data about devices / processes being monitored. Structured and unstructured.
- **Platform Data** – Data of the sensors itself [log data] for forensics & troubleshooting.



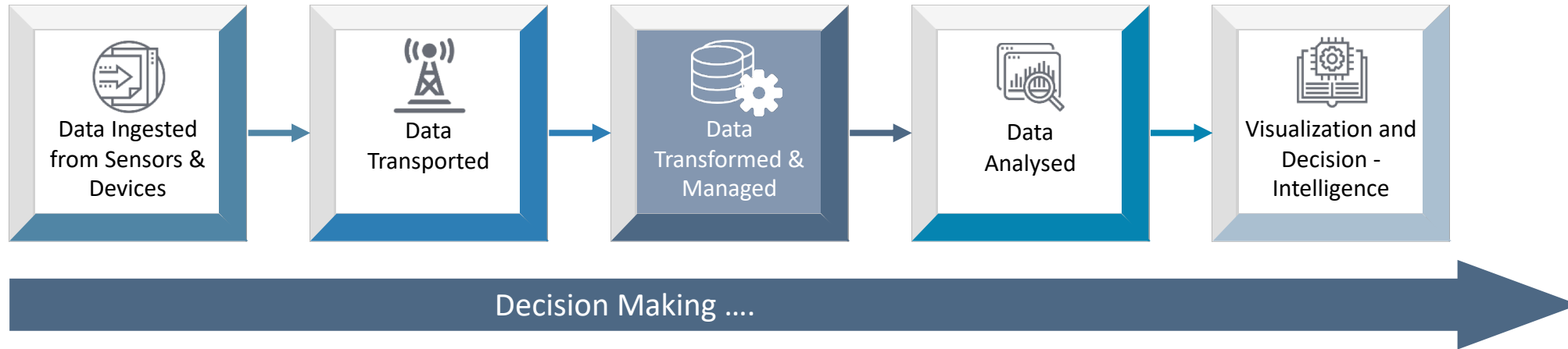
What is IoT Data ?

IoT Data must be:

- **Timely** – Rate of ingestion and query efficiency are key.
- **Accurate** – Data integrity is important
- **Actionable** – Data visualization, Anomaly detection and alerting are essential

Data Management and Analytics process

Data is collected, stored and analysed to gauge, prevent and enhance all events related to the connected device/systems.

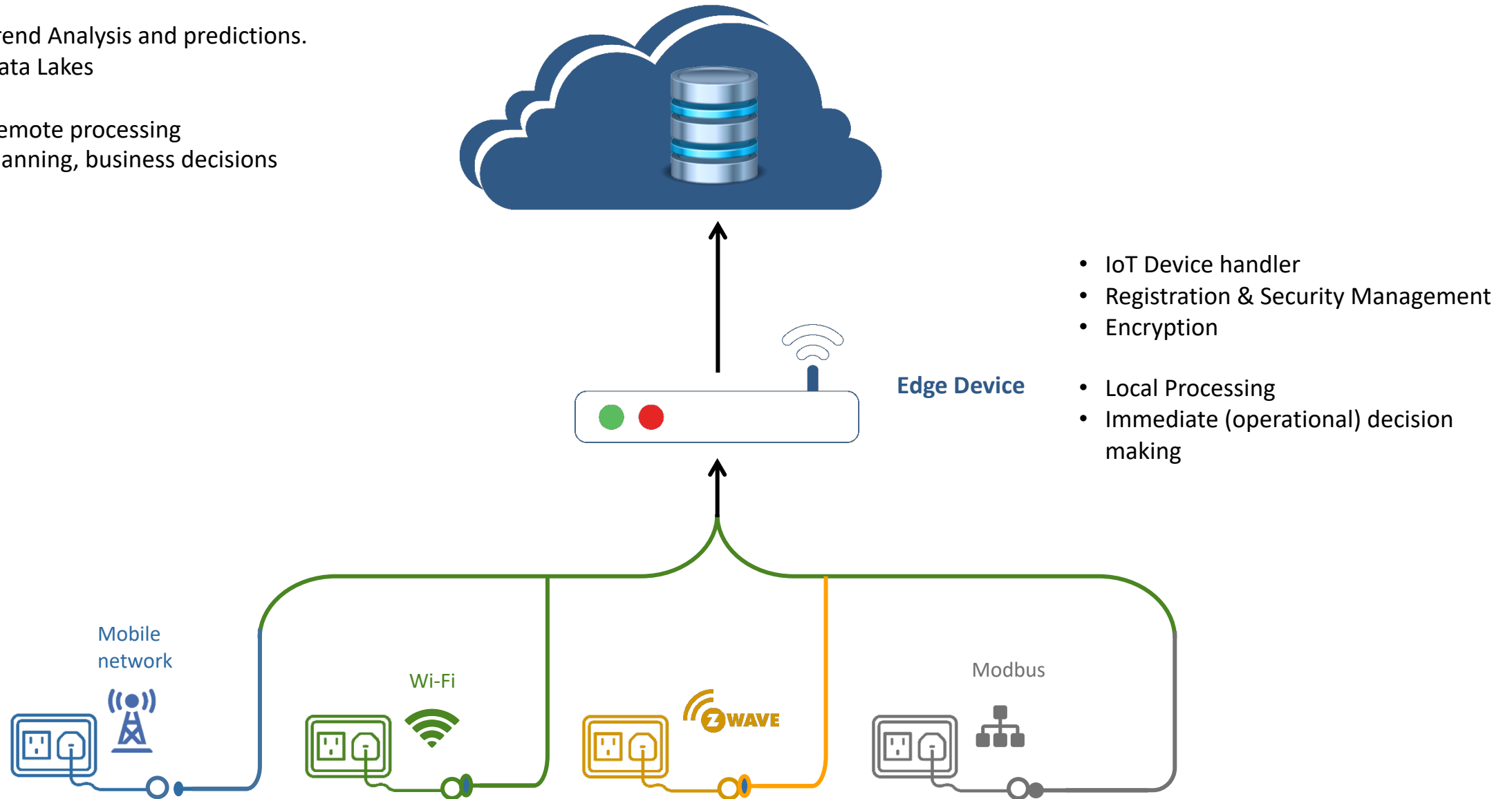


Not all data is equal.

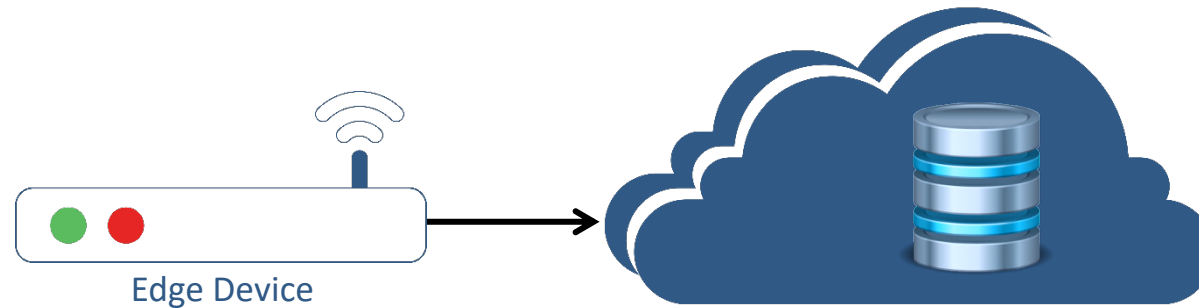
Data of high granularity will be needed at the Edge level.
Typically data averaged-out is send to Cloud.

Data Journey

- Trend Analysis and predictions.
- Data Lakes
- Remote processing
- Planning, business decisions



So where is data stored ?



Local Processing
Immediate (operational) decision making
Write intensive

Fast Storage / SSD
Low Latency

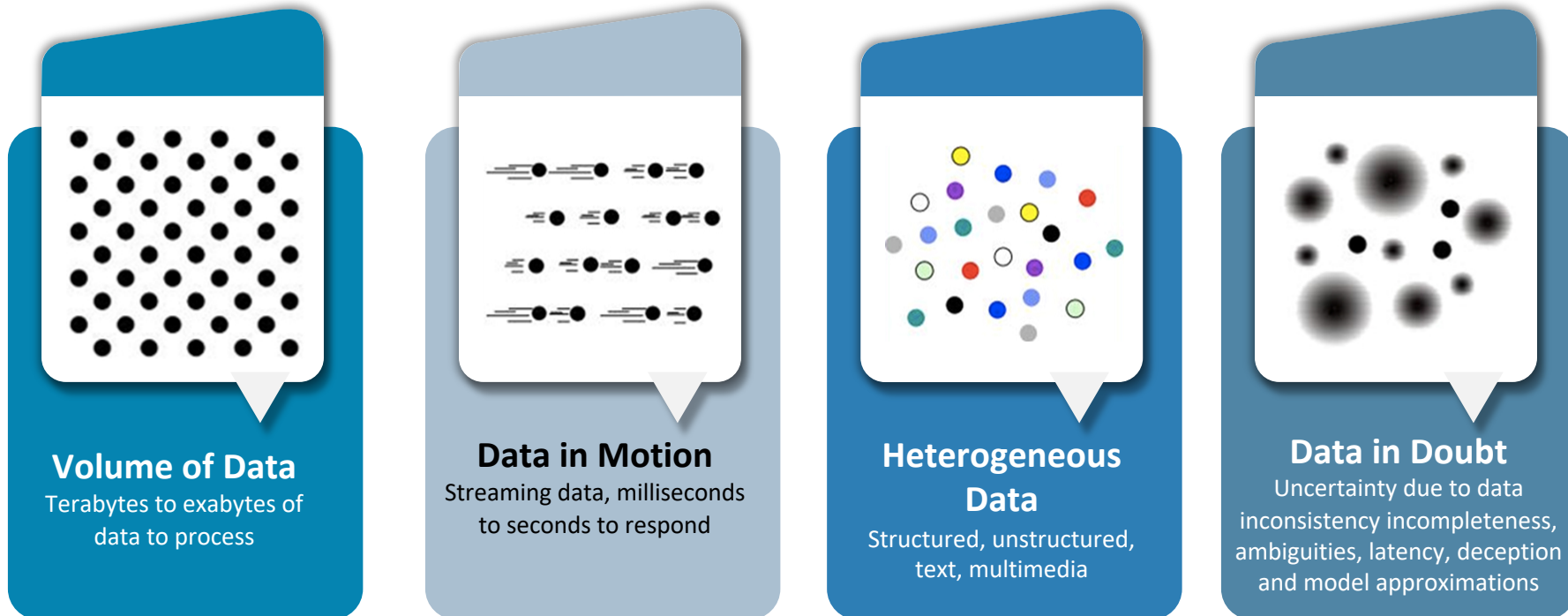
Trend Analysis and predictions.
Data Lakes
Remote processing
Data retention
Write intensive

More Capacity
Cost – Pay-per-user preferred

Challenge#1

Nature of Data in IoT – Multi modal and heterogeneous

- Heterogeneity
- Data collected is multi-modal, diverse, voluminous and often supplied at high speed
- IoT data management imposes heavy challenges on information systems.



Flexibility and Agility, with lower cost

- Rapid provisioning of storage - Enabling free flow of data to and fro through sensors without clogging
- Quick access – Local processing and analytics.
- Transient and Long-Term storage – Raw, Compressed, Aggregated data – Policy driven
- Distributed storage solutions required in order to manage huge data generated (Fog computing)

Real-time distribution of storage and compute for processing

- Pooled compute of the devices connected to the gateway can exceed what is available at the edge gateway
- Dynamic analytics for IoT, architecture allowing analytics to move across IoT compute spectrum

With billions of sensors being deployed, it's imperative to assert data privacy across the IoT ecosystem.

Data Security

Different types of security testing requirements:

- Data protection
- Data encryption (Snooping sensor and analytics data)
- Storage data security at edge and in the cloud
- Mechanism to flush the stale yet critical data

Determining how and where information can travel while keeping all devices up-to-date with business' security policy is crucial.

Endpoint and Network security

- Impersonation of devices
- Infra (Data Plane, N/W, etc.) monitoring and visibility
- Vulnerability and patch management
- Regulatory Compliances

Key opportunities created by the IoT Data Churn

