MPLS&SRV6AINETWORLD
25/27MAR25

palaisdescongrès
deparis

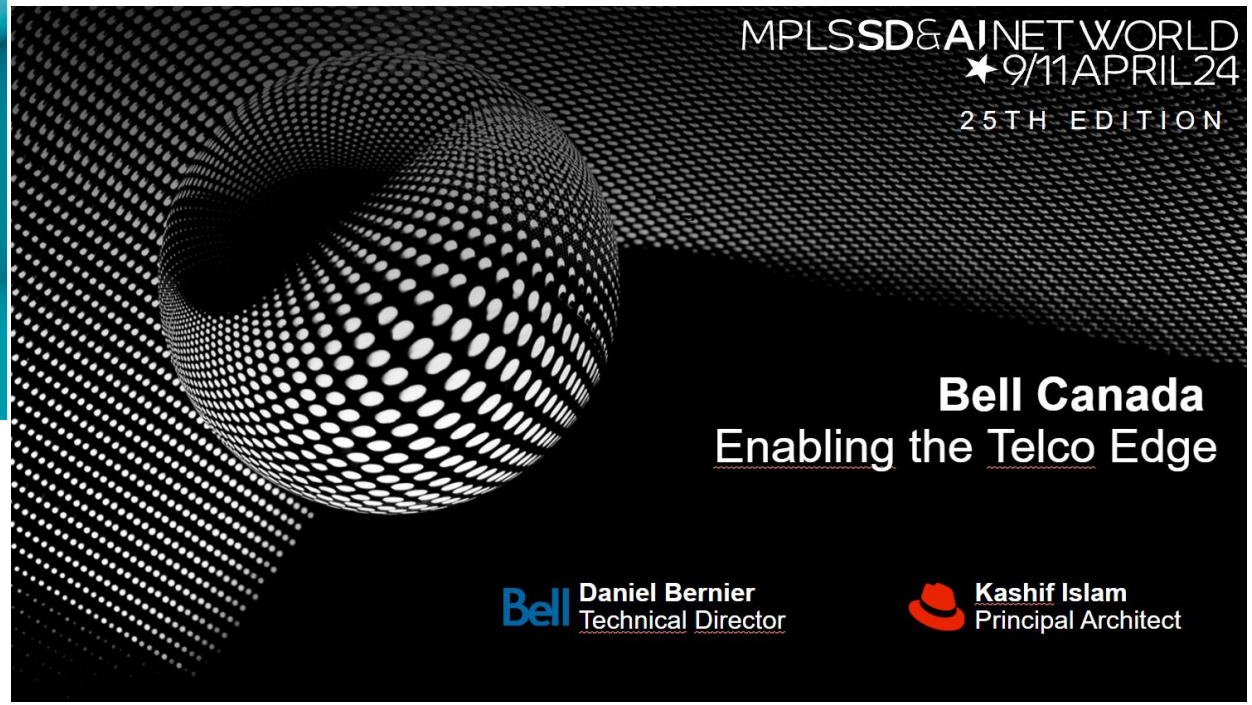
Agentic Edge: An AI Centric Approach to Provider Edge

26TH EDITION

Kashif Islam
Principal Architect

Syed Hassan
Principal Architect

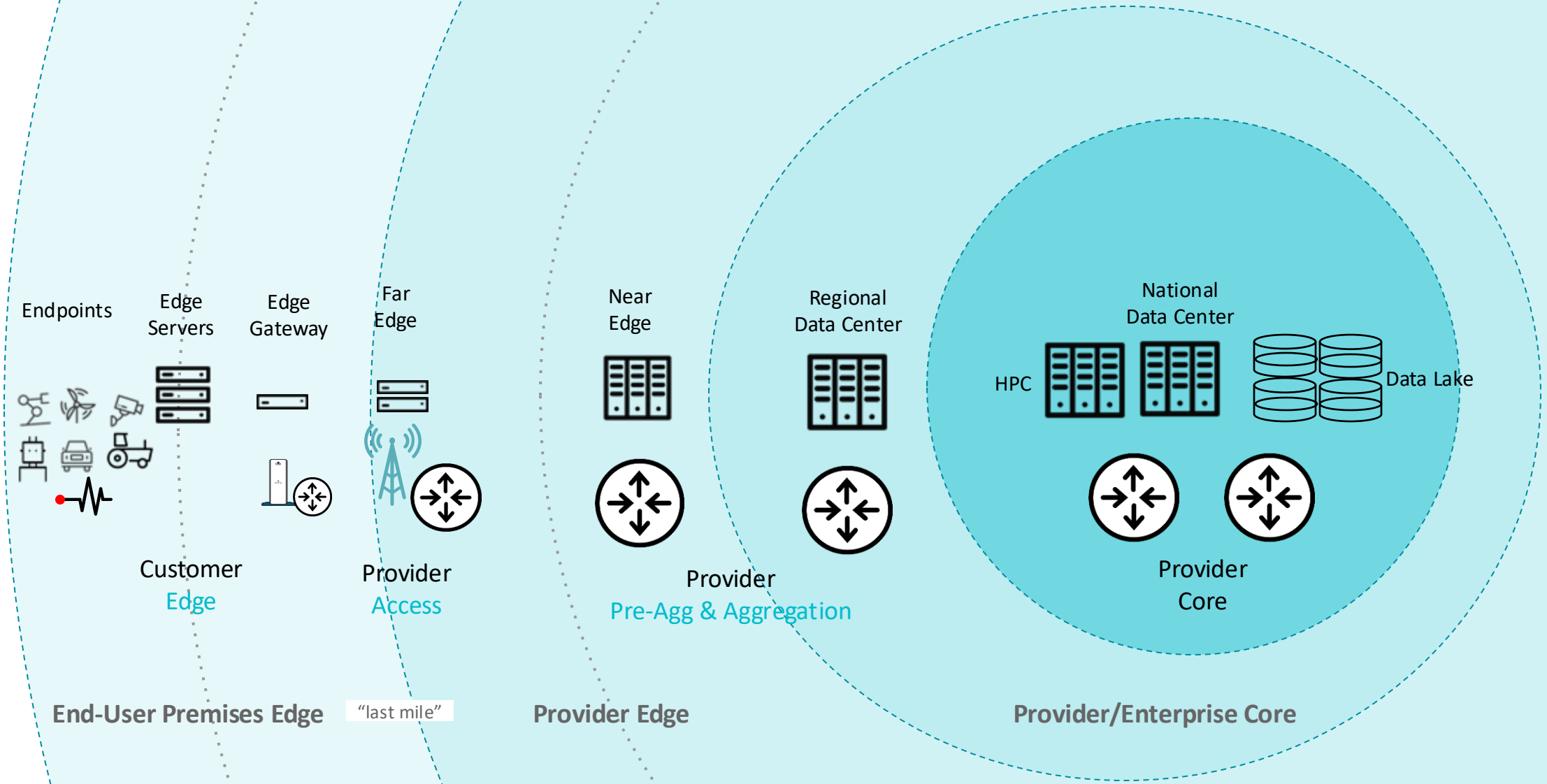
Our Edge Story at MPLS World Congress So Far !



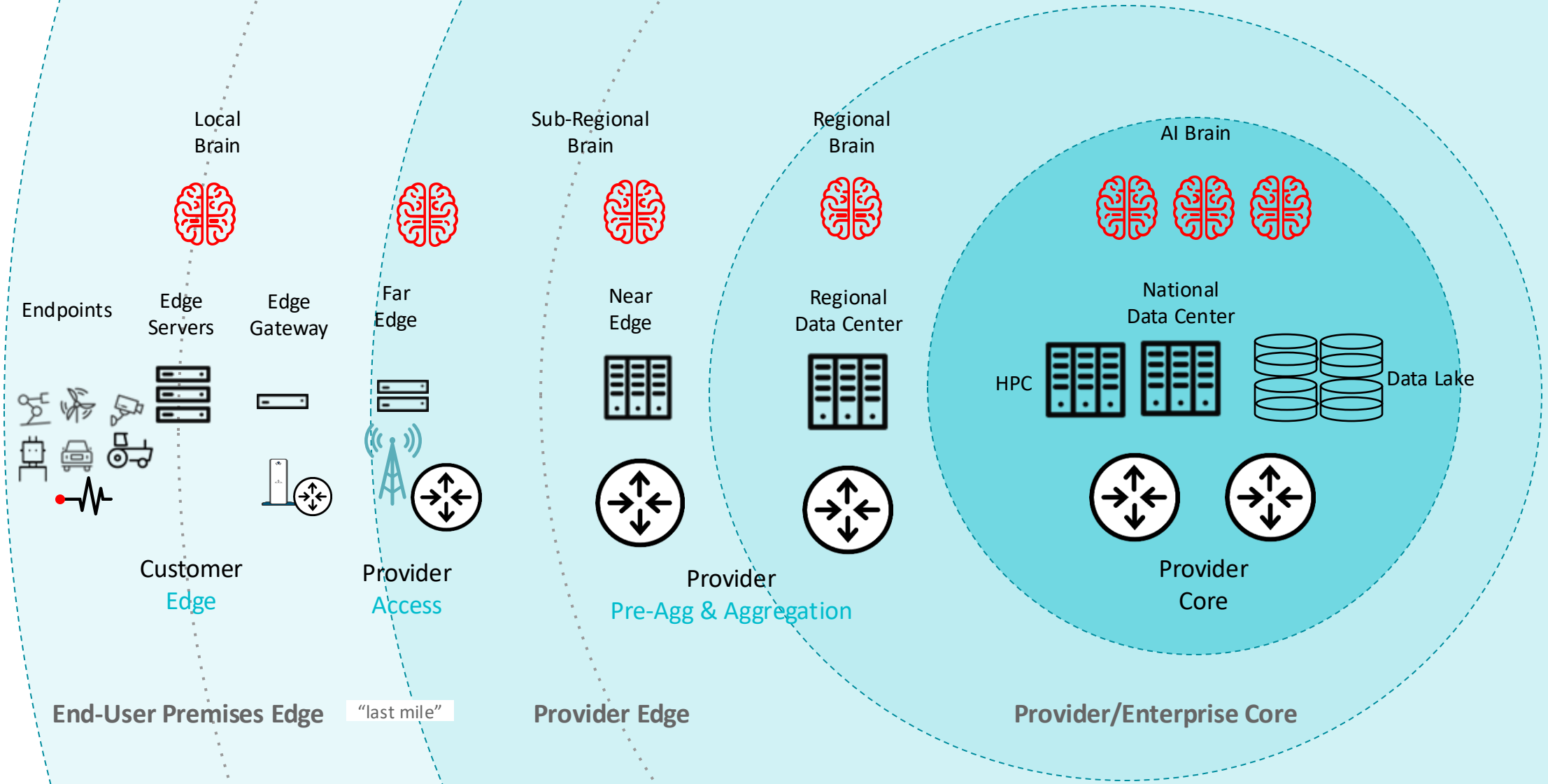
Red Hat Telco: Bringing Telco Cloud, Cloud-Native Apps and Provider Networks together!!

Red Hat

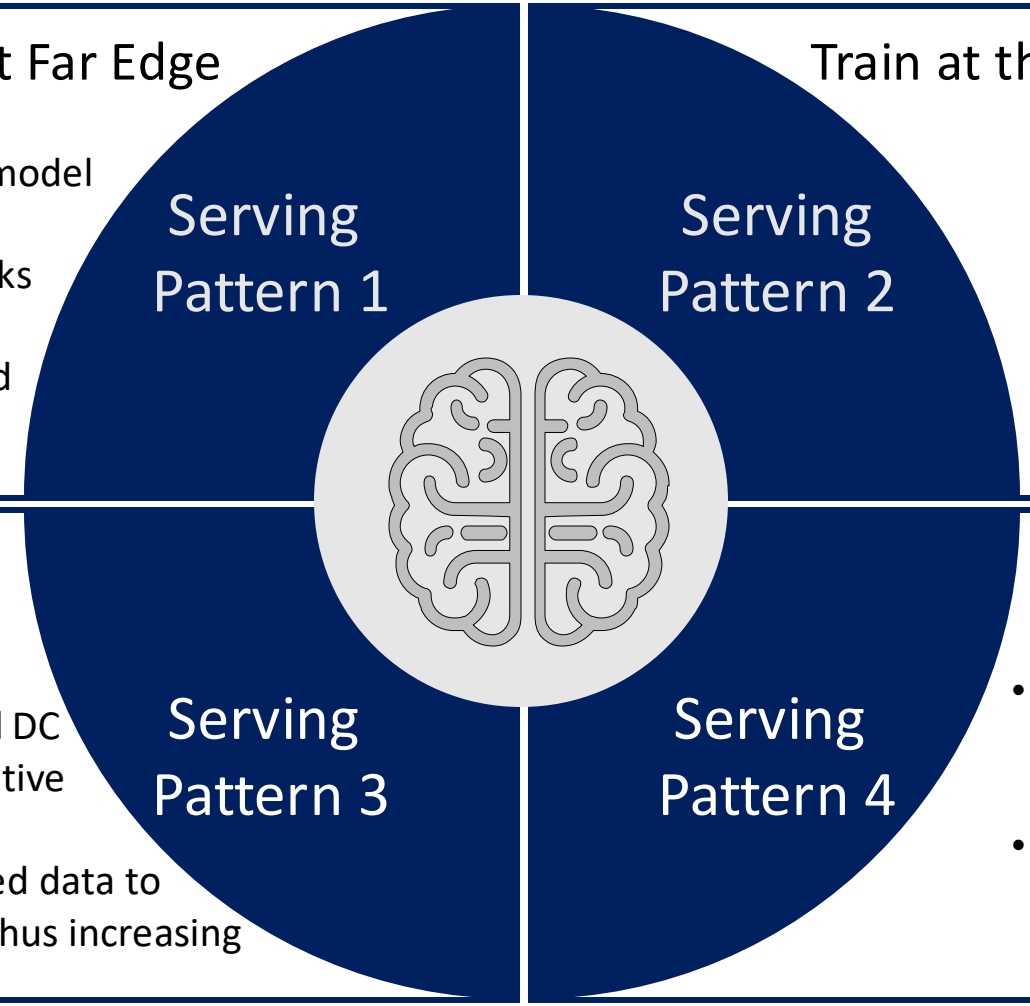
Understanding the Evolving Edge



The Intelligent Edge !!



AI Serving Patterns



Train at the Core, Serve at Far Edge

- Commonly used option
- HPC at the core DC provides model tuning capabilities
- Ideal for latency sensitive tasks
- Avoids backhauling high BW date (e.g. 4K videos processed for facial recognition)

Train at the Core, Server at Near Edge

- Best for slightly more latency tolerant tasks
- Offers higher processing capabilities enabling complex task
- Potentially higher impact

Everything at the Core

- Traditional model
- Cost effective w.r.t Capex as HPC is all collocated in central DC
- Not effective for latency sensitive tasks at the edge
- Backhauling all edge-generated data to core will strain the network, thus increasing OPEX

Train/Tune at Near Edge, Serve at Far Edge

- Uncommon, but effective for Small Language Models or tuning of pre-trained Models
- Cost effective, offers quick turnaround and startup time

Red Hat

Understanding Agentic AI

Predictive AI

- Classifies, predicts, or identifies data patterns
- Easier to integrate models into existing systems and architectures
- Capable of leveraging commodity CPU compute infrastructure for model services, training and tuning
- Runs most of the business today: Netflix movie recommendations, shopping cart recommendations, etc

Generative AI

- Creates new content by learning patterns from data
- Customers import, tune and manage Large (and sometimes Small) Language Models – LLMs and SLMs
- More computationally intensive that predictive AI, due to model size, data parameters and dimensions
- Requires more performance heaving and costly GPU infrastructure for model serving and tuning
- Starting to take hold in enterprises

Agentic AI

- Uses mathematical systems to make decisions and take action based on predictive and/or generative modeling
- Goal oriented behavior that uses multi-step reasoning and adaptive learning through a continuous feedback loop that evaluates impact and adjust accordingly in real time
- Autonomously perceives its environment, plans and carries out an action –or a series of actions–on behalf of the user

Red Hat

Defining an Agentic Edge

Assistants

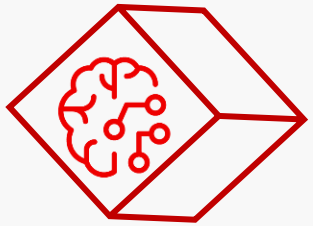
Information Retrieval
Prescriptive Tasks
Single Step Process
Can live on CNF, VNF or PNF
Multiple Assistants may gather information for Agents to provide correlation and perform action

Examples: transport assistant on a router, RAN assistant(s) on CU, DU

Agents

Multi step processing
Correlation of multiple data sources
Autonomous action taking
Self-correcting
Possible communications with other agents on behalf of humans

Example: Sensor fusion in robotic systems or RAN optimization agents in AI-RAN



Models

Problem Solving
Logical Thinking
Pattern Matching
Inferencing Engine
Large Language Models
Small Language Models

Examples: Gemini, OpenAI, Anthrop, Mistral, LLaMa, Granite, etc.

Platform

Consistent Policy and Governance

Red Hat Device Edge | Red Hat OpenShift | Red Hat OpenShift AI | NVIDIA | AMD | intel

Red Hat

Multi-Purpose Agentic Edge

Agentic Edge for Optimization	• Utilize AI Agents at the Edge to **increase network performance** • **RAN optimization** through AI RAN Agents • Traffic classification, network fault isolation and **remediation**
Agentic Edge for Services	• Agentic Edge as an enabler for **Next Gen services** • AI Factory & **AIaaS** • **Agent as a Service** for domain specific Edge AI offerings ○ Smart traffic Management, Sensor Fusion for Robotic assembly lines, Retail agents, and more

Red Hat

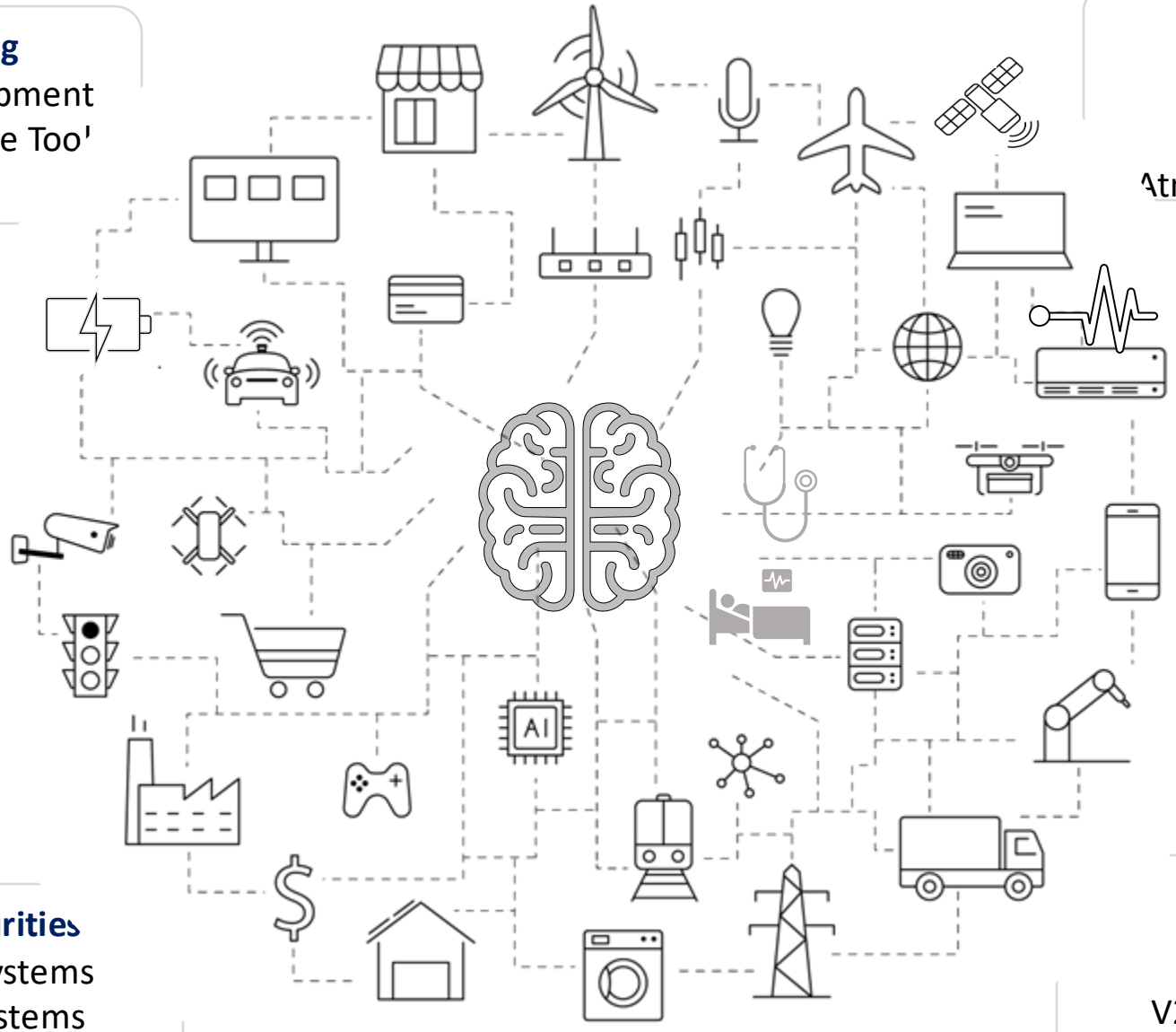
Agentic Edge in Action

Manufacturing
Manufacturing Equipment
Processing / Machine Too'
Industrial Robotics

Energy & Resources
Sustainability Initiatives
Power Control Systems

Security and Logistics
Automated Guided Vehicles
Facial Recognition
Smart Shelves

Finance & Securities
Securities Trading Systems
Order Processing Systems
Price Update Systems



NTN and Aerospace
Aircraft Control Systems
Satellite Edge Compute
Imagery analysis
Atmospheric anomaly detection

Medical & Healthcare
Medical Imaging Systems
Blood Analysis Systems
Telemedicine Systems

RAN Optimization
Smart RF Utilization
Intelligent Beam Forming
Adaptive Smart Handoff

Transportation
Train Control Systems
V2X, FSD, VTOL, Robo Taxi

Agentic Edge in Action

Agentic RAN Edge

- CSR traffic assistants **monitor network traffic**
- RAN assistants monitor **RF traffic patterns**

- Agent at the Far or Near Edge
  - **Correlate** Traffic Patters
  - **Analyze** SLA, Capacity, Power
- **Acts**
  - Turn on/off different sectors/carriers based on goal

Red Hat

Agentic Edge in Action

Agentic RAN Edge

- CSR traffic assistants **monitor network traffic**
- RAN assistant monitor **RF traffic patterns**

- Agent at the Far or Near Edge
  - **Correlate** Traffic Patters
  - **Analyze** SLA, Capacity, Power
- **Acts**
  - Turn on/off different sectors/carriers based on goal

- Agent, with help from local brain, **Correlates**
  - Camera feed
  - Sensory Data
- Agent **Take Actions**
  - Shutdown assembly lines
  - Shutdown access to unsafe area (e.g. mines)
  - Issue evacuation order

Assistants gathers inputs from:

- Cameras for safety violations
- IOT Sensors for Unsafe environment
- QoE sensors for product quality

Agentic Workplace Edge

Red Hat

Agentic Edge in Action

Agentic RAN Edge

- CSR traffic assistants **monitor network traffic**
- RAN assistant monitor **RF traffic patterns**

- Agent at the Far or Near Edge
  - **Correlate** Traffic Patters
  - **Analyze** SLA, Capacity, Power
- **Acts**
  - Turn on/off different sectors/carriers based on goal

- Agent, with help from local brain, **Correlates**
  - Camera feed
  - Sensory Data
- Agent **Take Actions**
  - Shutdown assembly lines
  - Shutdown access to unsafe area (e.g. mines)
  - Issue evacuation order

Assistants gathers inputs from:

- Cameras for safety violations
- IOT Sensors for Unsafe environment
- QoE sensors for product quality

Agentic Workplace Edge

Agentic Non-Terrestrial Edge

- Assistant in VLEO/LEO use imaging and environment **gather data** over a large geographical area

- Agent (NTN or TN based) **correlates** information from NTN agents (and potentially TN agents)
- **Acts**:
  - Notify authorities – public works, firefighting dept
  - Dispatch Firefighting Robo-VTOLs

Red Hat

Agentic Edge in Action

Sub Regional Brain

Agentic RAN Edge

- CSR traffic assistants **monitor network traffic**
- RAN assistant monitor **RF traffic patterns**

- Agent at the Far or Near Edge
  - **Correlate** Traffic Patters
  - **Analyze** SLA, Capacity, Power
- **Acts**
  - Turn on/off different sectors/carriers based on goal

Local Brain

Agentic Workplace Edge

- Agent, with help from local brain, **Correlates**
  - Camera feed
  - Sensory Data
- Agent **Take Actions**
  - Shutdown assembly lines
  - Shutdown access to unsafe area (e.g. mines)
  - Issue evacuation order

Assistants gathers inputs from:

- Cameras for safety violations
- IOT Sensors for Unsafe environment
- QoE sensors for product quality

Hybrid Brain

Agentic Non-Terrestrial Edge

- Assistant in VLEO/LEO use imaging and environment **gather data** over a large geographical area

- Agent (NTN or TN based) **correlates** information from NTN agents (and potentially TN agents)
- **Acts**:
  - Notify authorities – public works, firefighting dept
  - Dispatch Firefighting Robo-VTOLs

Red Hat

Agentic Edge is the logical evolution of Provider Edge that build upon MEC and accelerates the adoption of AI at the Edge

Red Hat

upperside conferences