Telecommunication Infrastructure in the AI Era
Efforts in Network Operations and AI Data Centers

Why Do We Need AI?

Users' demands on the network continue to evolve

Increasing Network Complexity

- Diversification of devices and services
- Growth in data volume
- Introduction of new technologies (such as 5G / Beyond 5G and IoT)

Challenges Arising from Complexity

- Difficulty of troubleshooting
- Need for efficient resource management
- Importance of real-time performance monitoring

Benefits of AI

- Greater automation and efficiency
- Enabling faster problem resolution
- Reducing human error
- Lowering operational costs
- Enhancing service quality

Tomorrow, Together
KDDI

Symbiotic Evolution of AI and Networks

The evolution of networks

Network for AI
Networks flexibly responding to AI traffic

The evolution of AI

AI for Network
Harnessing AI to enhance networks

Tomorrow, Together
KDDI

Symbiotic Evolution of AI and Networks

The evolution of networks

Network for AI

Networks flexibly responding to AI traffic

The evolution of AI

AI for Network

Harnessing AI to enhance networks

Tomorrow, Together
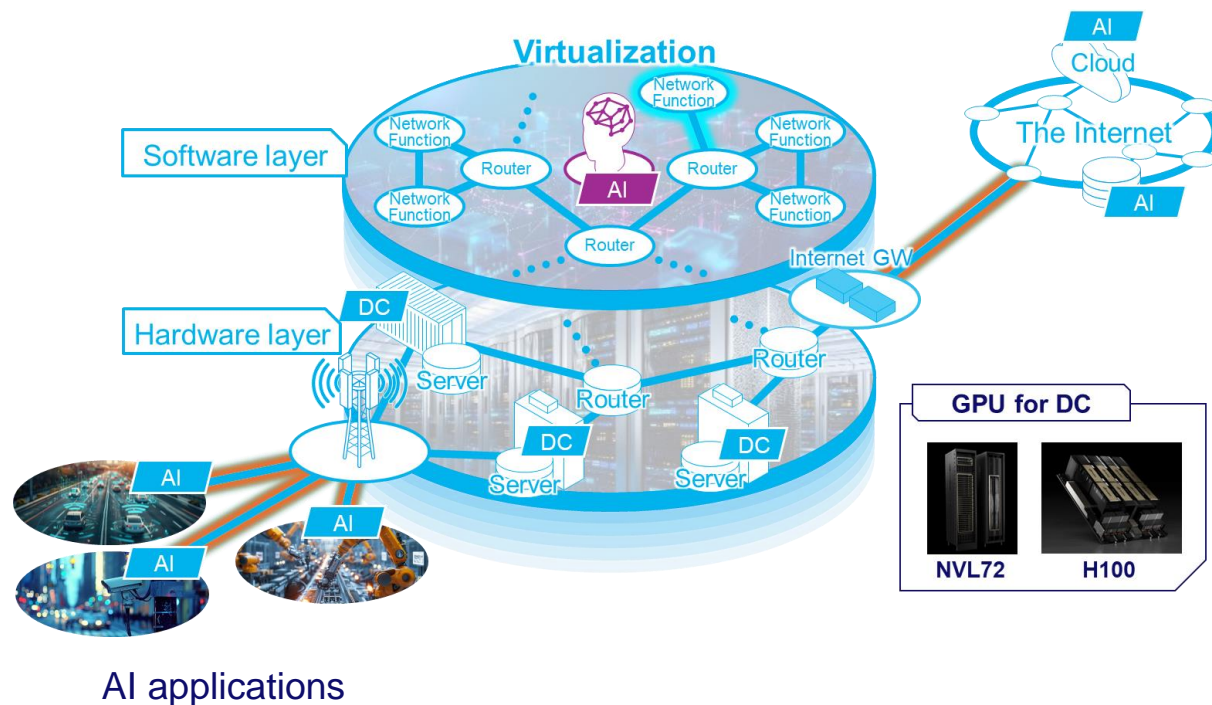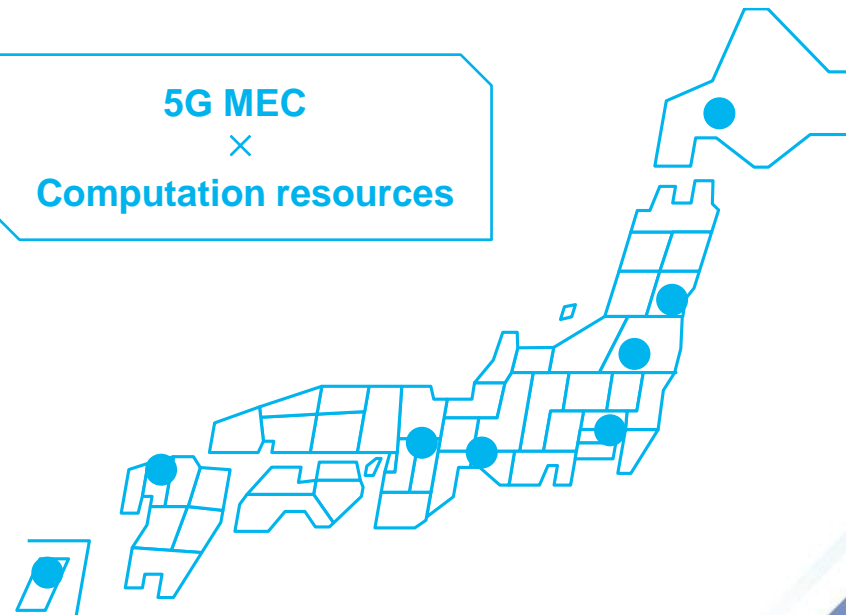
KDDI

Why Do We Build AI Data Centers?

KDDI's virtualization technology and 5G MEC enables low-latency access to AI service.

KDDI's flexible and resilient Networks

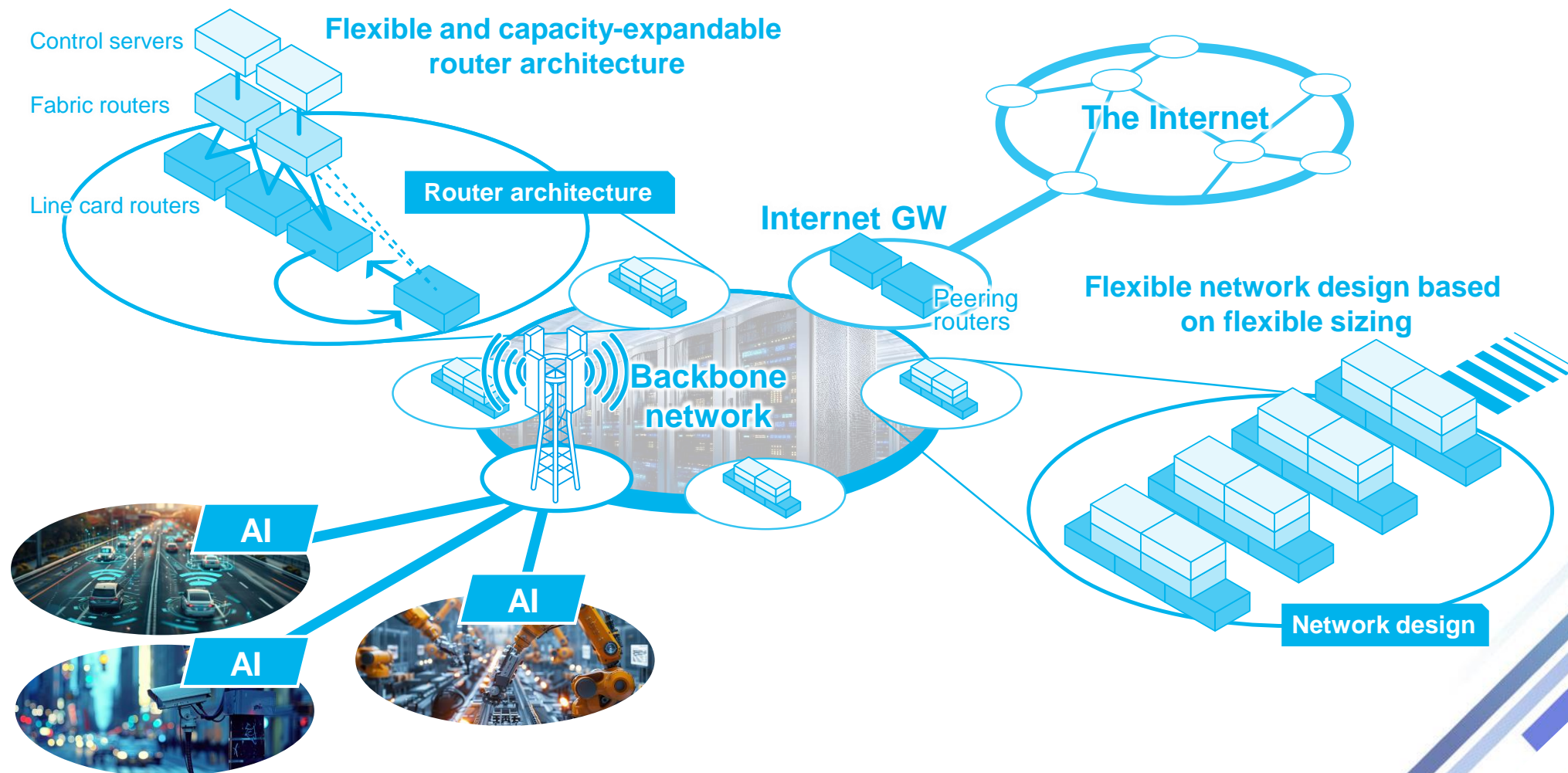**Low-latency, distributed computing infrastructure**

AI applications

Edge computers in each data centers (DC)
delivers seamless connectivity to AI.

Tomorrow, Together

Scale-out Network

Control servers

Fabric routers

Flexible and capacity-expandable router architecture

Line card routers

Router architecture

Internet GW

The Internet

Peering routers

Backbone network

Flexible network design based on flexible sizing

AI

AI

AI

Network design

Tomorrow, Together

Challenges to Build DC in the AI Era

1st: Investiment for DC construction

Integrated GPUs to support large-volume computation especially for AI learning

Mid- to long-term CAPEX
about 640 million euro
(about 65 million euro in grants)

Resource sharing and collaboration

Accelerate business and service co-creation using generative AI
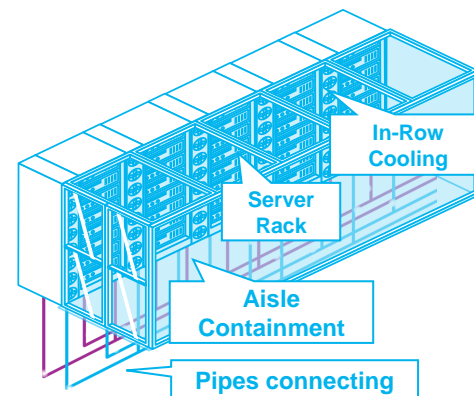
In-house

Industry partners

KDDI BUSINESS

Startups

MUGENLABO
生成AI活用支援プログラム

NICT
国立研究開発法人
情報通信研究機構

Research Institutes
& Municipalities

2nd: Development of liquid cooling systems

In-Row Cooling

Server Rack

Aisle Containment

Pipes connecting

GPU for DC

NVL72
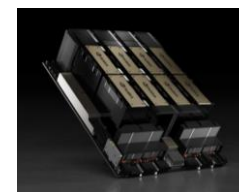
H100

3rd: Efficient construction of liquid cooling ready DC

Server

Building

Capping

LCD Factory

Bus-Duct

Electric Power

Water Power

Piping

Tomorrow, Together
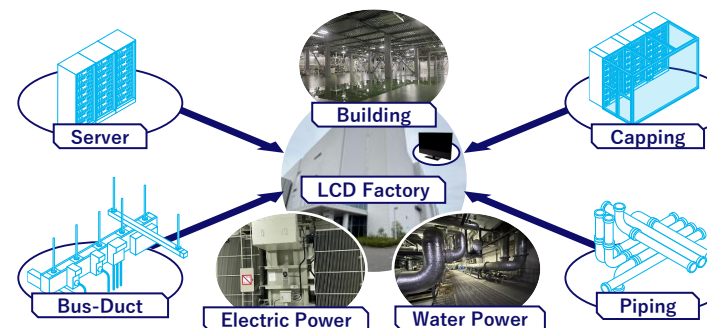
KDDI

Evaluation of Liquid Cooling for GPU Server

KDDI leads the development of liquid cooling and power supply for GPU server at the TELEHOUSE TOKYO Shibuya in collaboration with the partners.
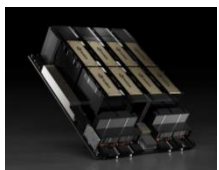
Tomorrow, Together
KDDI ✕ **SUPERMICRO** **Nidec** **SHINOHARA**

GPU for DC

NVL72

H100

In-Row Cooling

Server Rack

Aisle Containment

Pipes connecting

TELEHOUSE TOKYO Shibuya

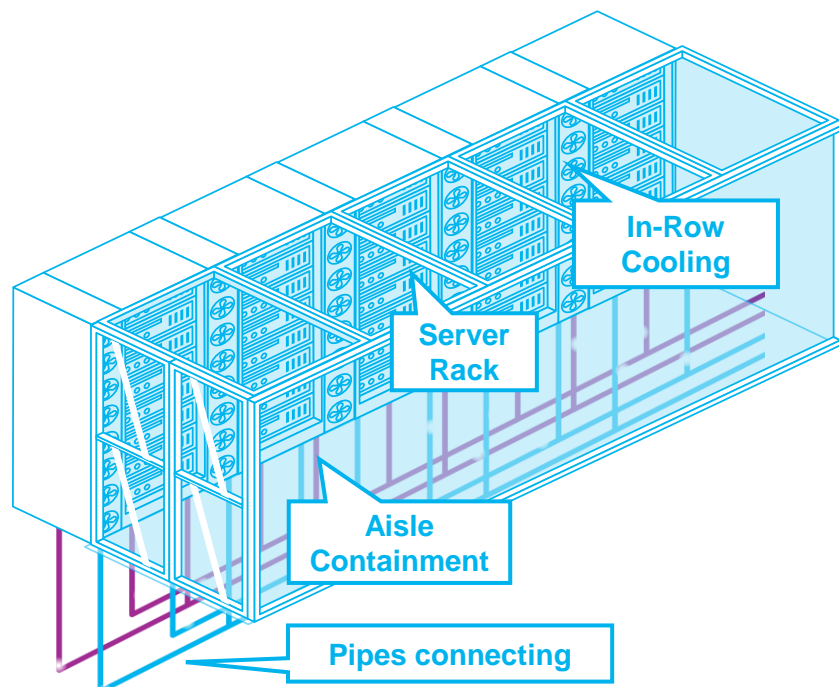*Tomorrow, Together*
KDDI

From Hyper Scale DC to DC in the AI Era

We plan to build liquid cooling ready AI DC as well as installing liquid cooling
to the current Hyper Scaler DC and Container DC.

AI DC	Hyper Scale DC	Container DC
Liquid Cooling and High-Capacity Power	High SLA (Carrier-Grade Quality and Operations)	Flexible and Rapid Deployment
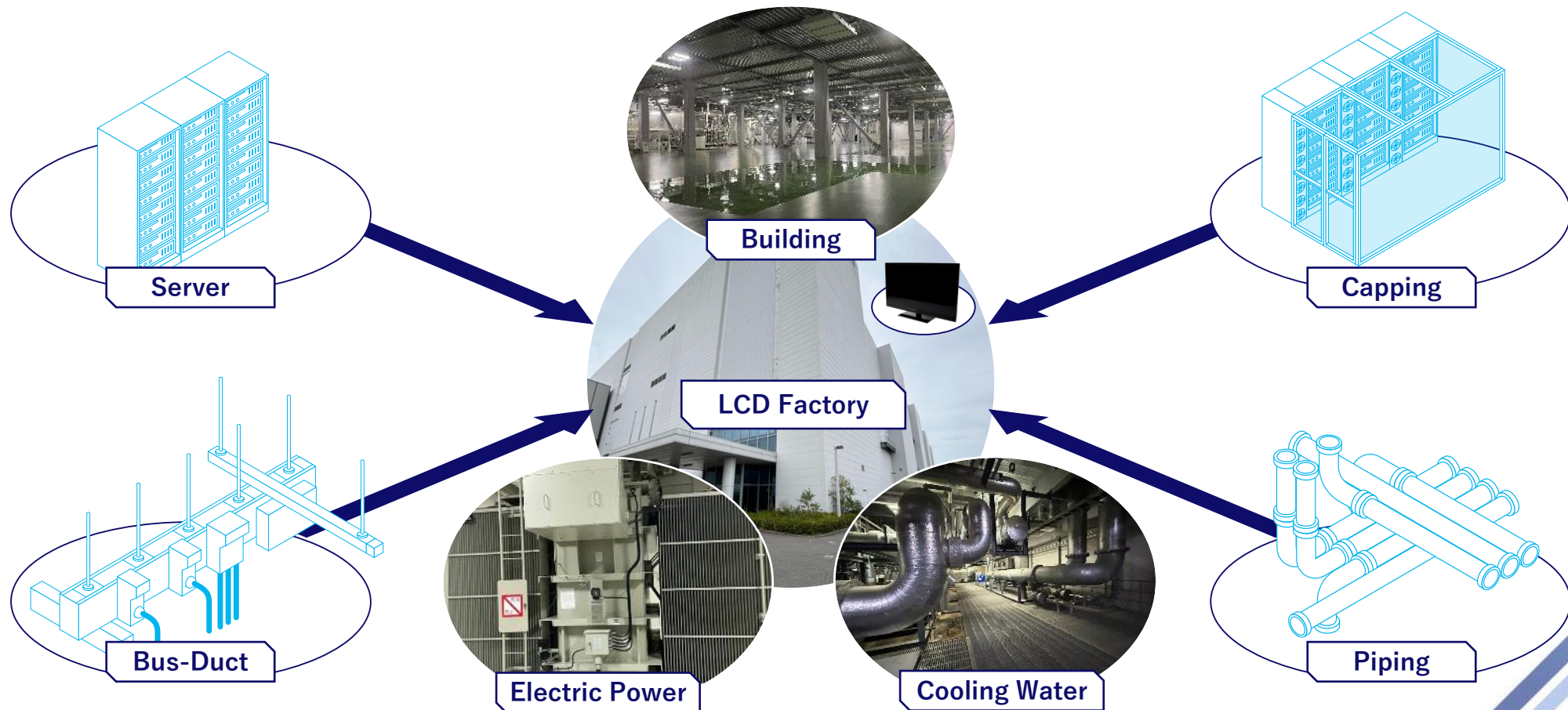
Symbiotic Evolution of AI and Networks

The evolution of networks

Network for AI
Networks flexibly responding to AI traffic

AI for Network

Harnessing AI to enhance networks

The evolution of AI

Tomorrow, Together
KDDI

Zero-Touch Operation

Automate Network Recovery

- Automatically isolating systems by monitoring availability, performance, and equipment logs
- Eliminating manual operations for network operators enables the fastest possible service recovery

Incapable of Handling Irregular Cases

- In exceptional cases where zero-touch processing is challenging, operators must rely on their expertise to manually identify the root cause of the issues

Zero-Touch Operation

Communication services are restored by isolating systems in a zero-touch manner.

Dashboard

Service Monitoring Platform — Failure — Monitoring Performance

Automation Platform — Failure Recovery — Incident Management

Isolate Suspected System

Data（LOG/TRAP）

Mobile Network — IT System
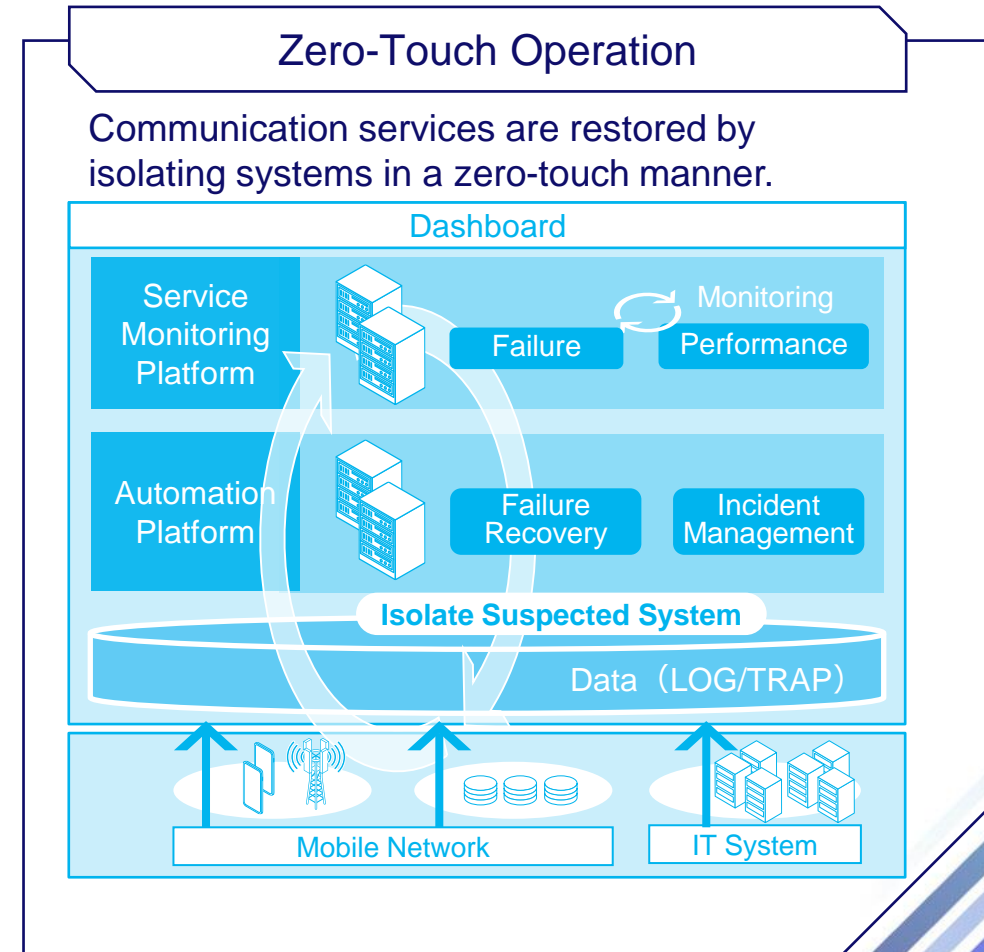
Incident Recovery with Generative AI

Multimodal Generative AI unifies all knowledge

- Multimodal Generative AI excels at interpreting and processing natural language and unstructured data, integrating diverse forms of knowledge to enable more intuitive human-computer interactions

Incident Recovery Support

Even less experienced operator can identify root cause efficiently.

Handling traditional human-dependent tasks

Interacting with operator

Identify Root Cause

Recovery Procedure Exploration

Incident Recovery Support AI (Multi-Agent)

Telco AI
LLM

Operation Manual   Log   Topology

LOG

Dataset

Knowledge Data Extraction

Transform unstructured data into structured data

Unstructured Data

Image Data (Topology)   Text Data

Structuring data with multimodal AI

Telco AI
LLM

Network topology, etc.

Structured Data

Tomorrow, Together

KDDI

Intent-Driven Networking (IDN)

Automate Desired Outcomes

- IDN allows operators to specify desired outcomes like performance optimization without defining technical implementations, as the network configures itself

SLA: Bandwidth utilization must be 50% or lower.

Operator

Example: Traffic Tuning on Transport Network

Input

Feedback

Design

Abstraction Level

high

low

NW Requirement

Business Intent

Service Intent

Resource Intent

Network Monitoring
（Monitoring/Analyzing）

Apply

Orchestrator / Controller

Configuration

Information Retrieval

Overloaded

Offload

Underloaded

Tomorrow, Together

KDDI

Simplifying Intent with Generative AI

Intent is not easy for non-networking backgrounds

- KDDI aims to enhance convenience by leveraging the power of Generative AI to derive operators' intents effectively from user input.
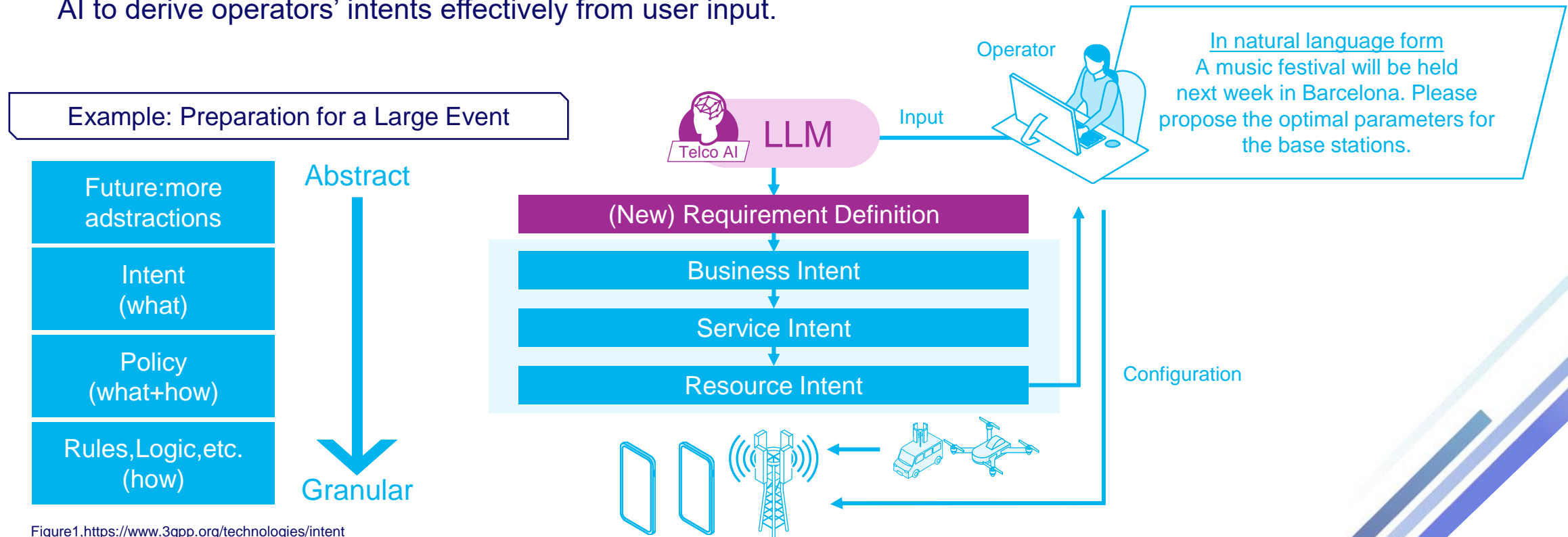
Example: Preparation for a Large Event

Future:more adstractions
Intent (what)
Policy (what+how)
Rules,Logic,etc. (how)

Abstract

Granular

Figure1,https://www.3gpp.org/technologies/intent

Operator

In natural language form
A music festival will be held next week in Barcelona. Please propose the optimal parameters for the base stations.

Telco AI LLM

Input

(New) Requirement Definition

Business Intent

Service Intent

Resource Intent

Configuration

Tomorrow, Together
KDDI

Thank you for your attention