ITU Workshop on "Machine Learning for 5G and beyond"

End-to-end network operation automation in IMT-2020 and beyond systems

Ved P. Kafle (NICT), Tatsushi Miyamoto (KDDI),
Takayuki Kuroda (NEC), Taro Ogawa (Hitachi)

This work was conducted as part of the project entitled "Research and development for innovative AI network integrated infrastructure technologies" supported by the Ministry of Internal Affairs and Communications, Japan.

2019 June 17

Outline

- Research motivation
  - 5G service scenarios
  - On-demand network slicing
- Network service design
  - Diverse requirements analysis
- Network design
  - Parameters setting
  - NW construction & deployment
- Resource adaptation
  - Agile control, QoS guarantee
- Failure detection and recovery
  - Dependable service
  - Fast recovery

Research motivation

Networks getting complex; diverse services coexisting

Diverse requirements

IoT


Transport


Healthcare

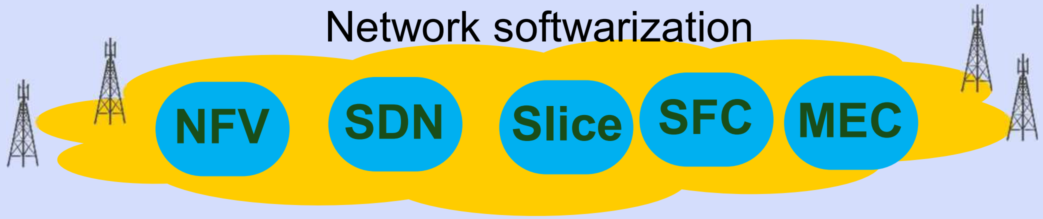

Agriculture


Manufacturing

…

5G

Network softwarization



Challenges

Large scale network, complex operation, demanding high technical skill

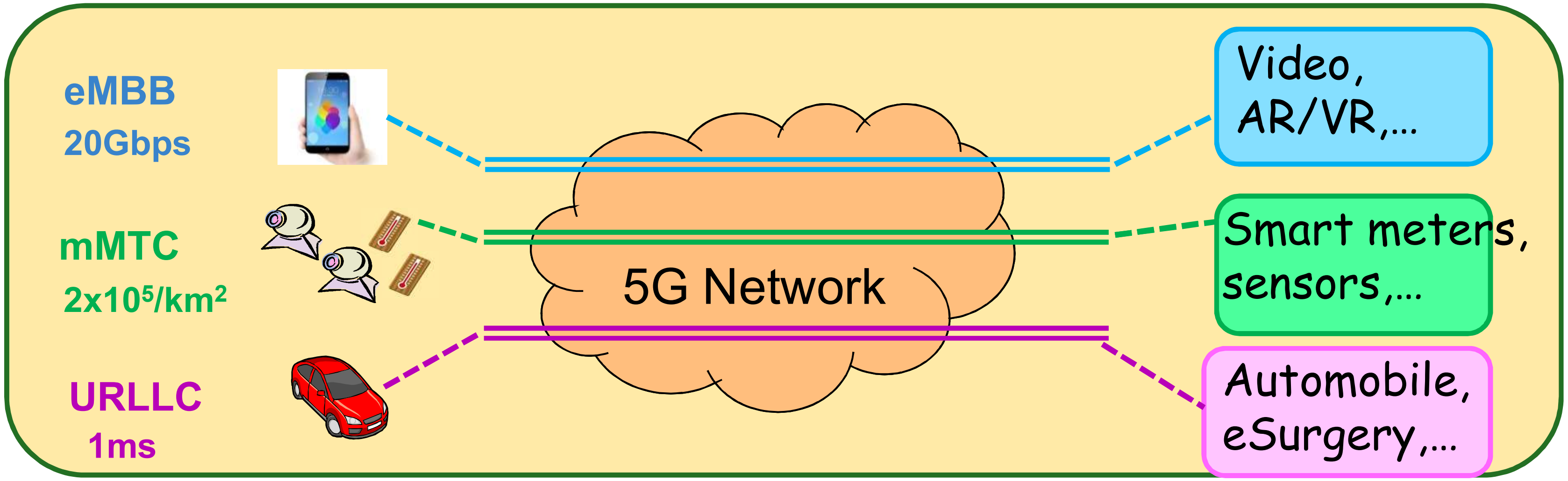
・Explosive growth of control data, complicated operation and system configuration

・Requiring advanced skill in software

We need to develop automation technologies that require less human intervention for network design, construction and operation.

5G service scenarios overview

- Various services in 5G/IMT-2020 networks, diverse requirements:
  - eMBB: very high throughput
  - mMTC: large connection density
  - URLLC: ultra-low latency



- Different services be served through network slices

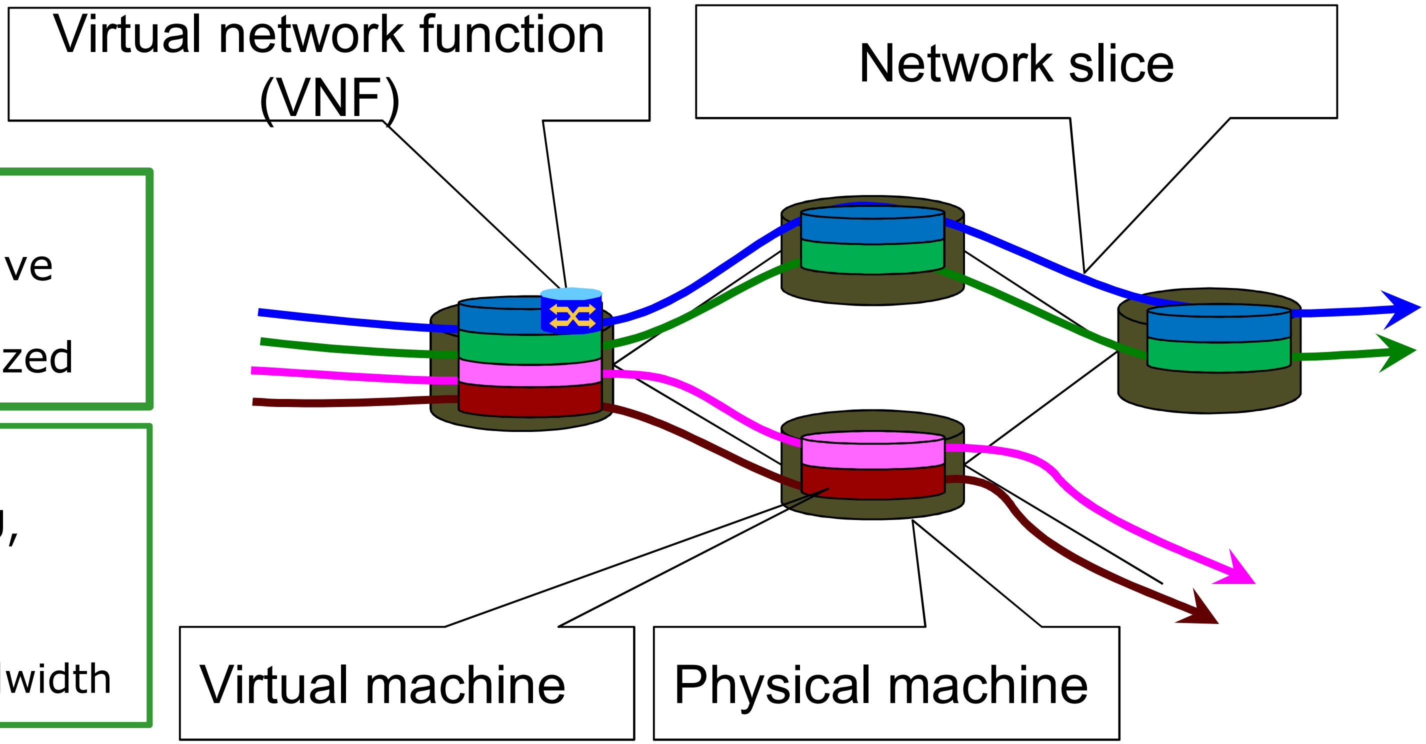
ITU-R M.2083-0 (09/2015): IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond.

Network slicing through NFV, SDN

- Creation of multiple virtual network slices over the same physical network
  - SDN and NFV are supporting technologies

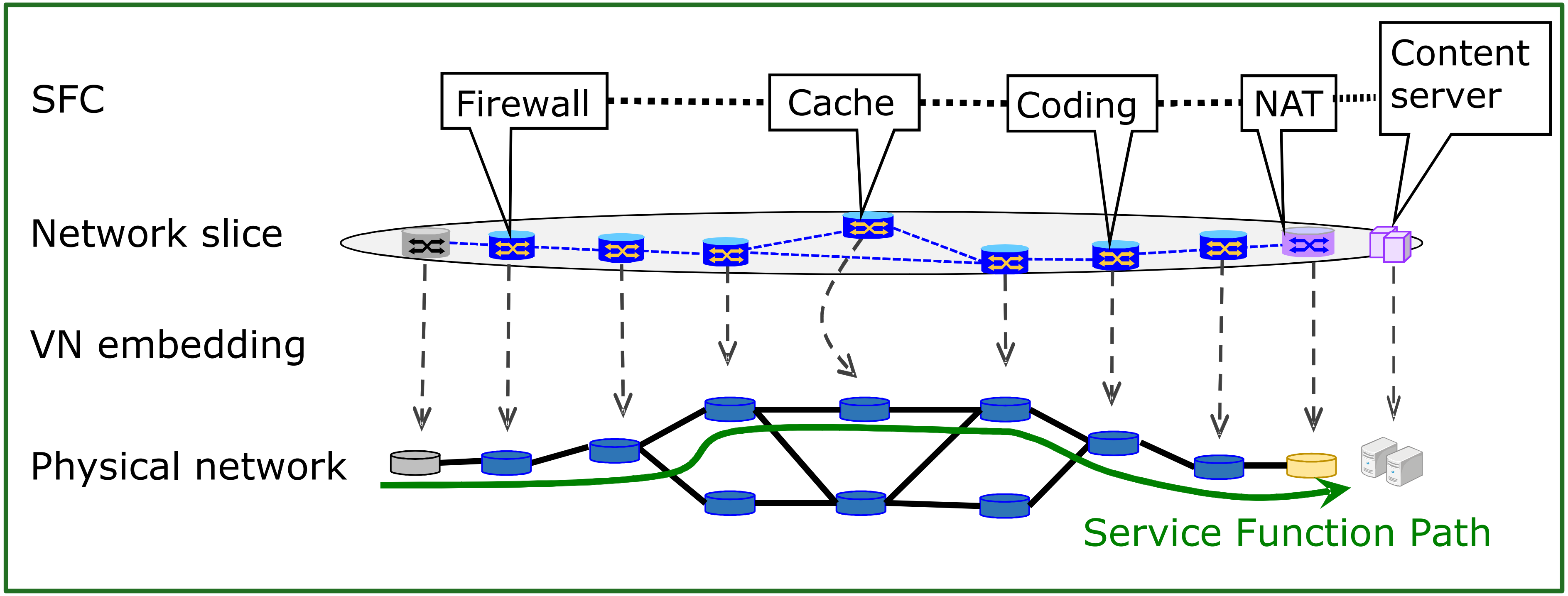
- VNF
  - cloud-native function, containerized

- VN resource
  - Node: CPU, memory, storage
  - Link: bandwidth

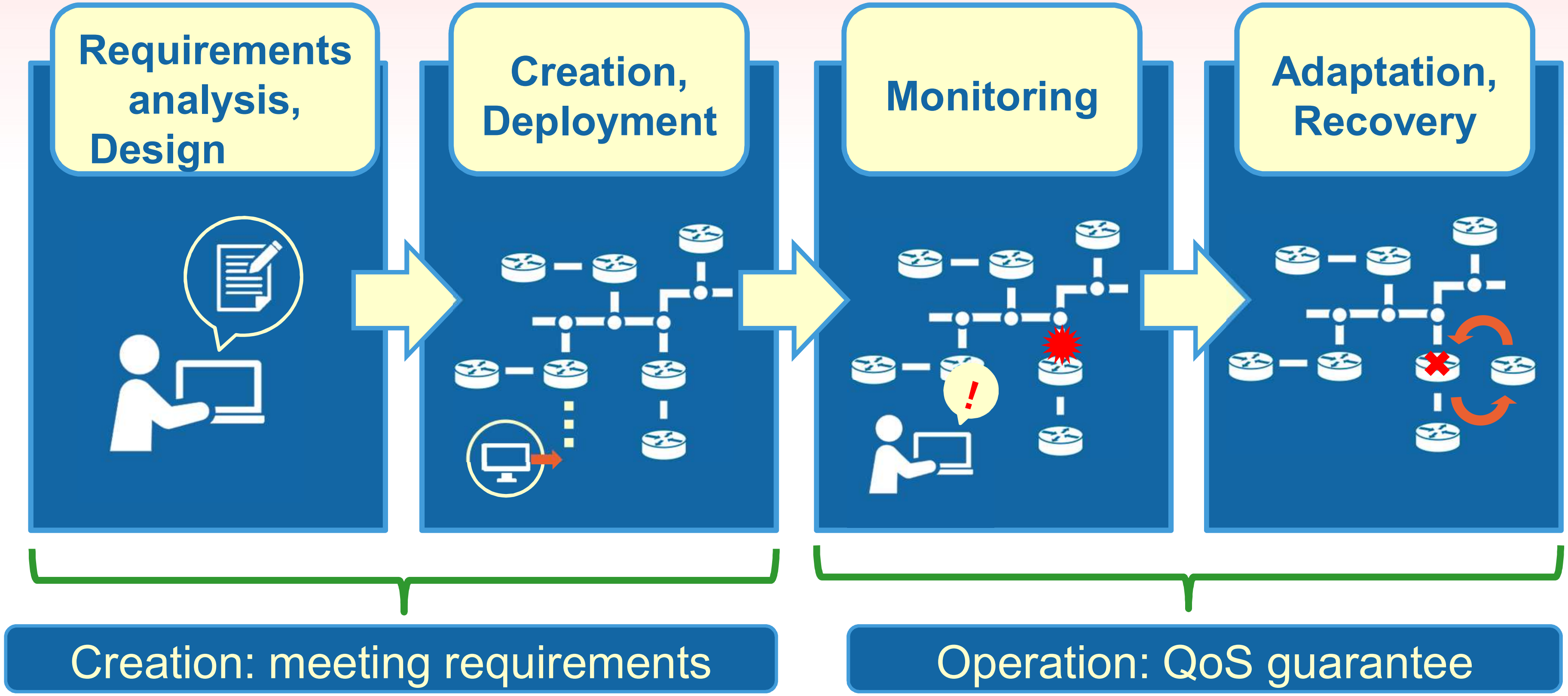


Network slicing: SFC

- Service Function Chaining (SFC):
  – Ordered placement of VNFs for a given network service



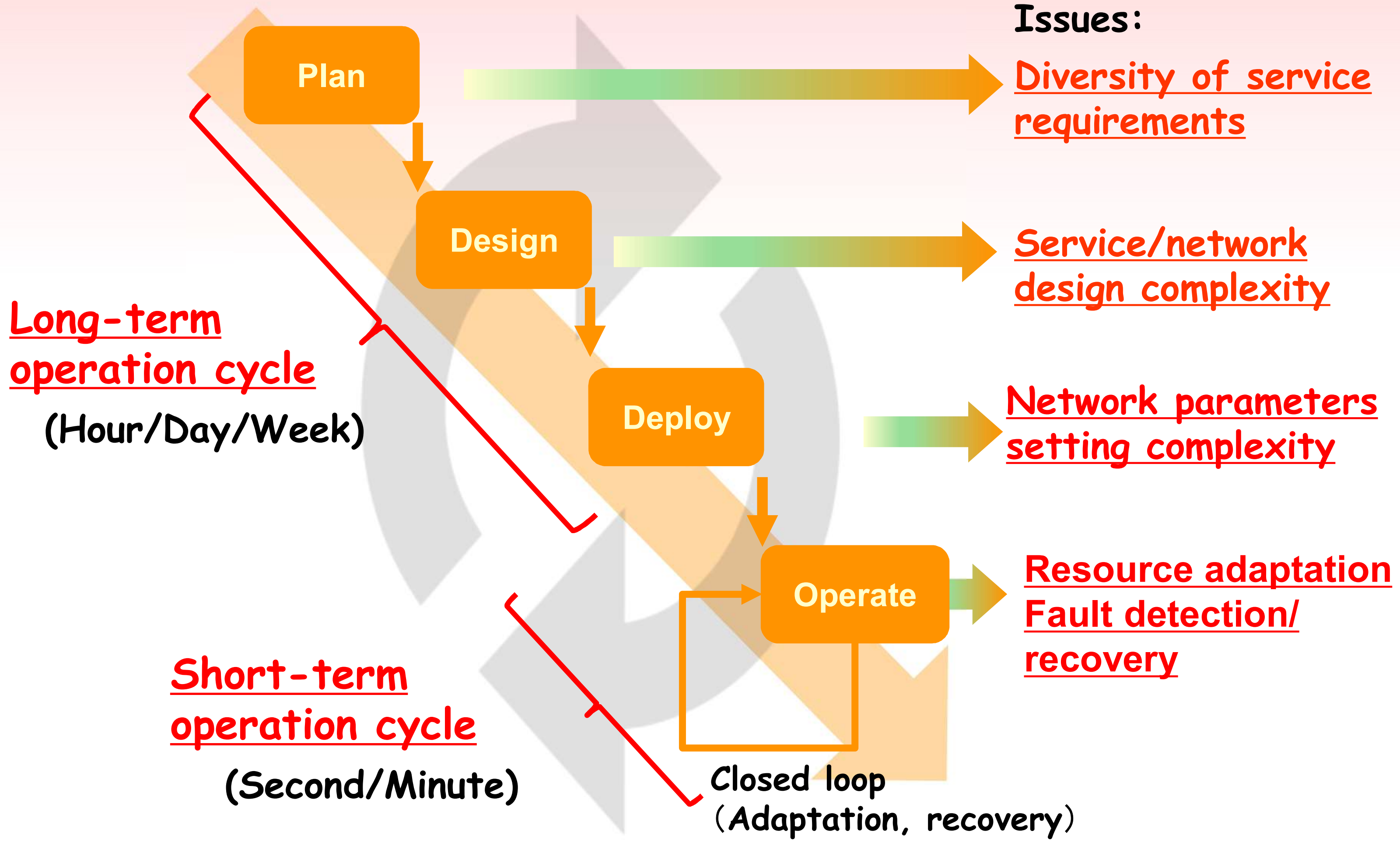
Network slicing – creation to operation steps



On-demand network provisioning

Agile adaptation, faster recovery

Network slicing: Long-term and short-term operation



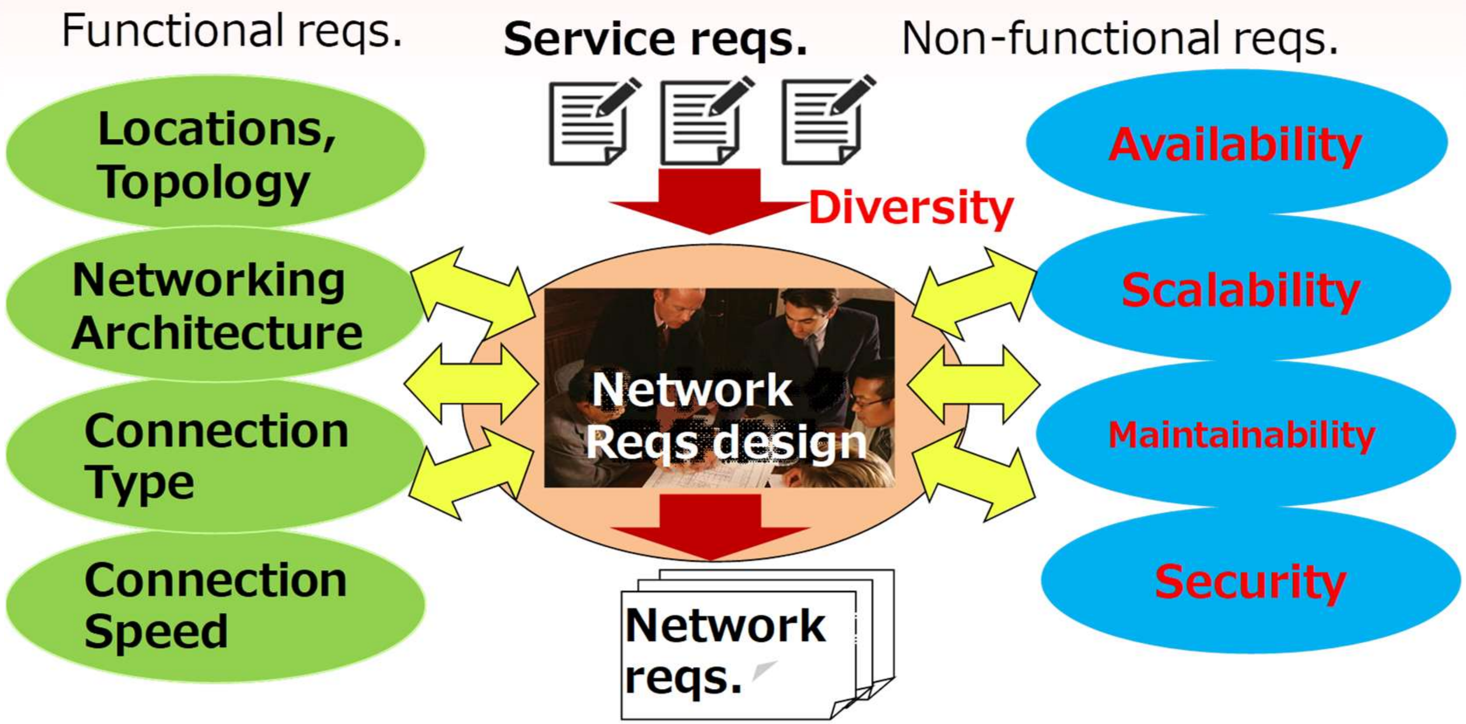
Developing AI-based network control technology for long- and short-term cycle operation of stable communication infrastructure

Outline

- Research motivation
  - 5G service scenarios
  - On-demand network slicing
- Network service design
  - Diverse requirements analysis
- Network design
  - Parameters setting
  - NW construction & deployment
- Resource adaptation
  - Agile control, QoS guarantee
- Failure detection and recovery
  - Dependable service
  - Fast recovery

Network requirements design challenge

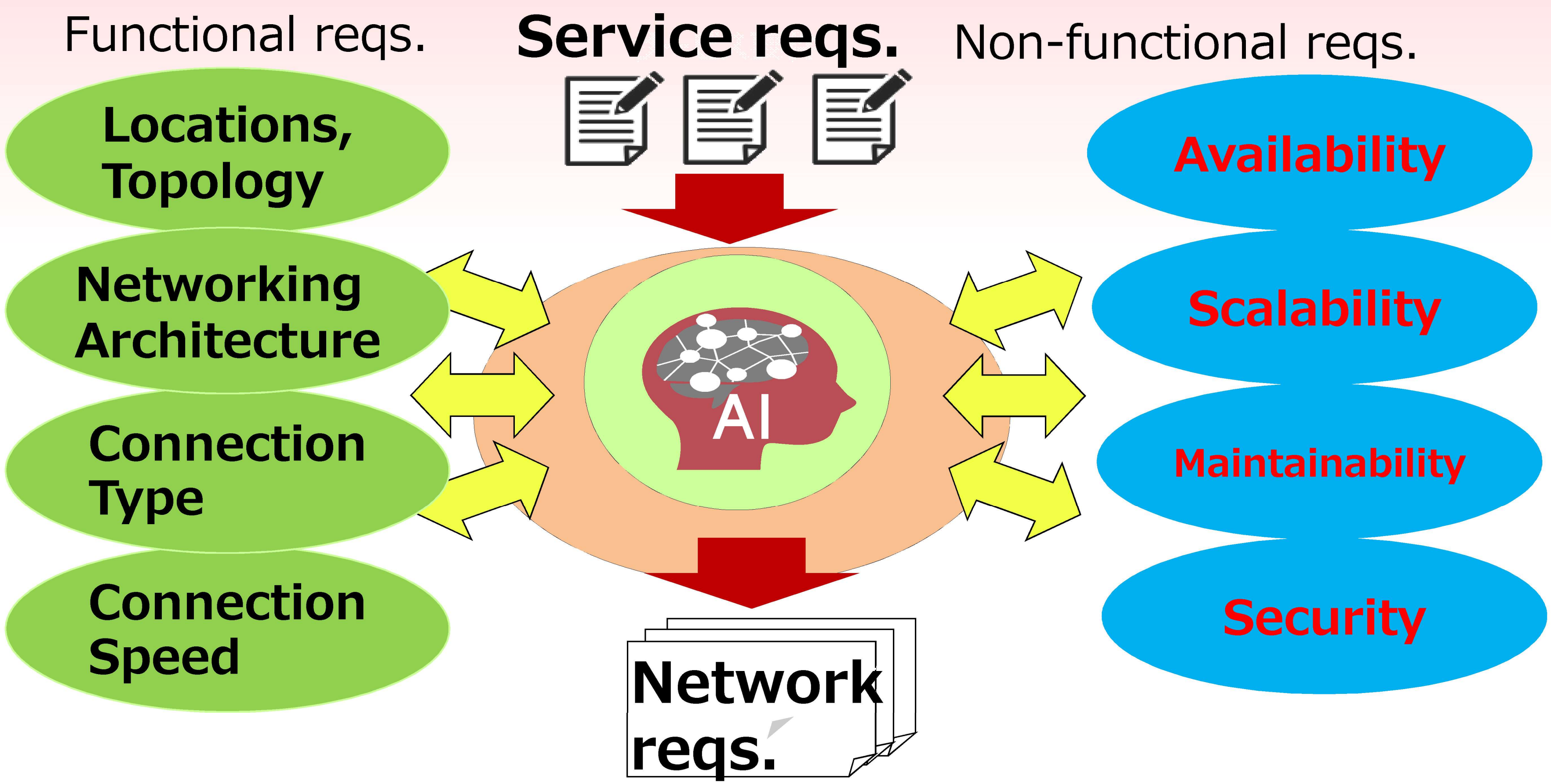
- **Diverse functional and non-functional requirements**
- **Analysis & design skill of experts required**



- **<u>Labor intensive, case-by-case basis</u>**

  ⇒Bloated design- & maintenance-operations
  ⇒Knowledgeable expert shortage

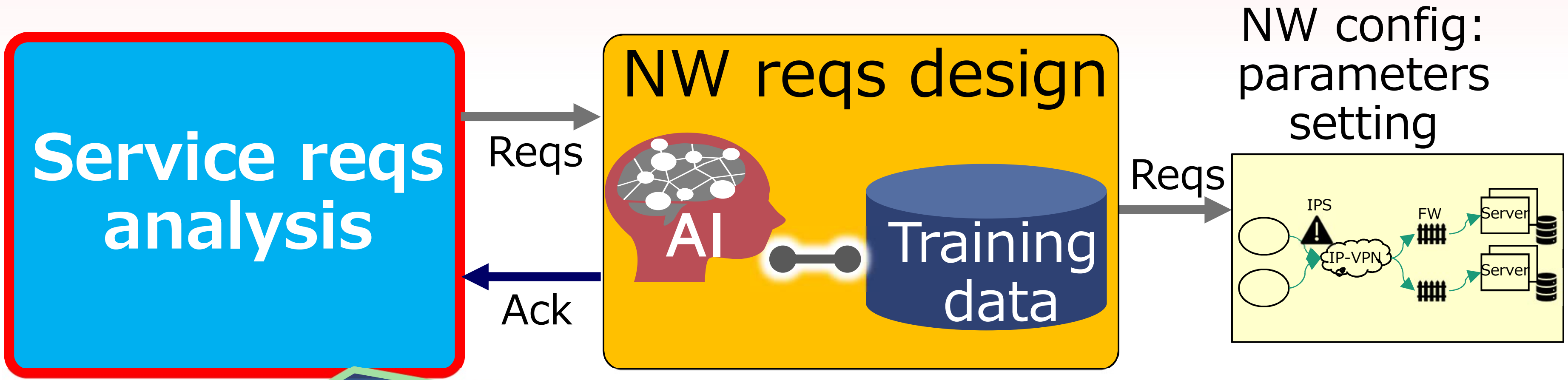
AI-based approach to network design



- **Automate translation from service reqs. to network reqs.**
  - **⇒Sharing the expert knowledge**
  - **⇒Agility in network construction and operation**

Service requirements collection and analysis

- **Systematic service-requirements analysis & translation through modeling**



1）Service-requirements collection:

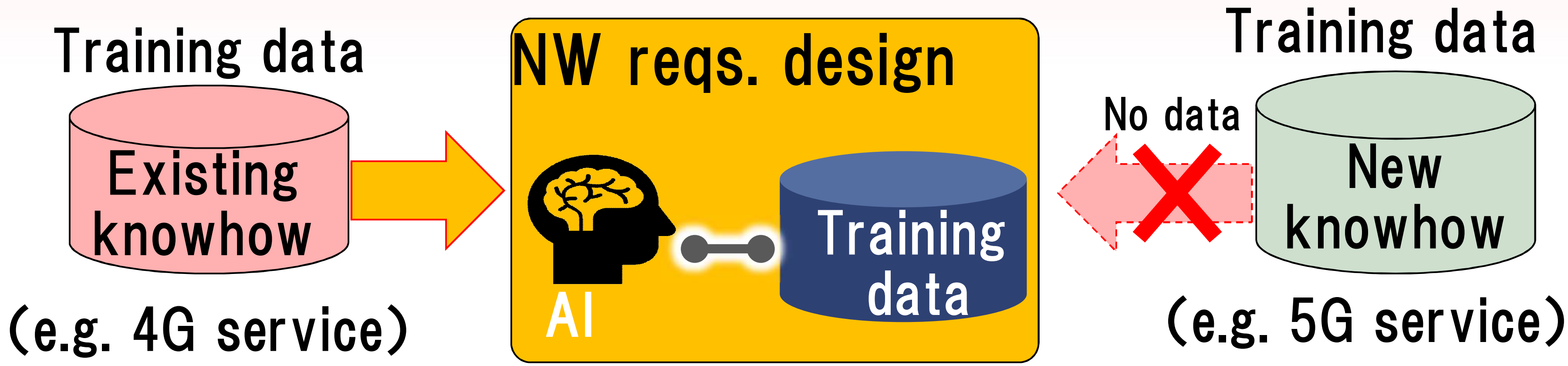
・Via free format description, conversation, & GUI

2）Model-based analysis of service-requirements:

・Translating service requirements into NW requirements

Adapting to emerging requirements

▪ **AI scheme selection based on the available training data**



- Massive training data for existing networks ⇒ Supervised
- Less for the emerging networks ⇒ Unsupervised, transition

+

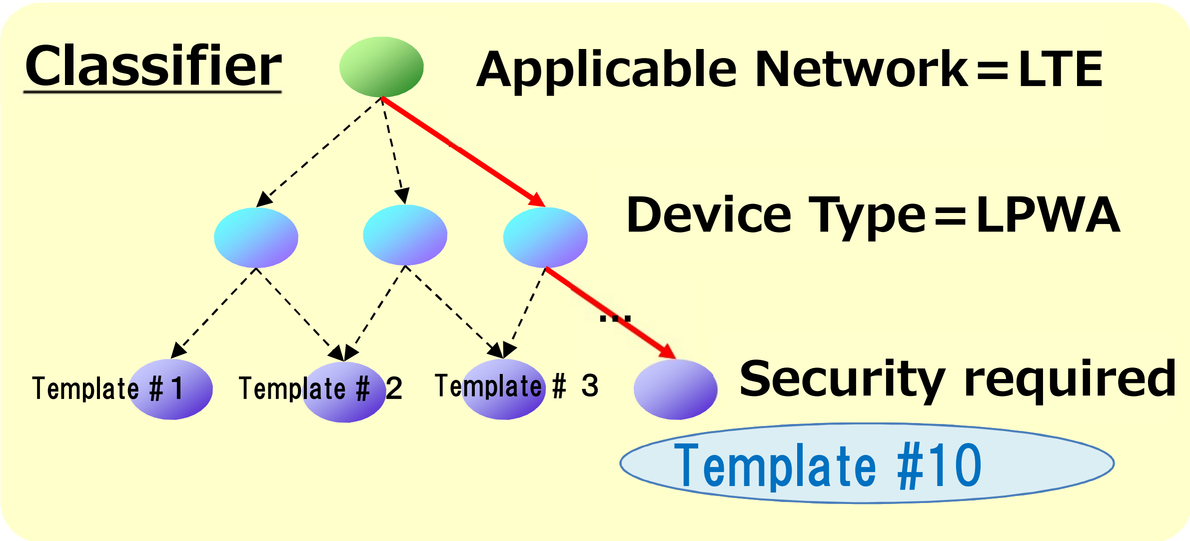
- Human interactive generation of training data for emerging service reqs.

Evolutional AI with human in the loop

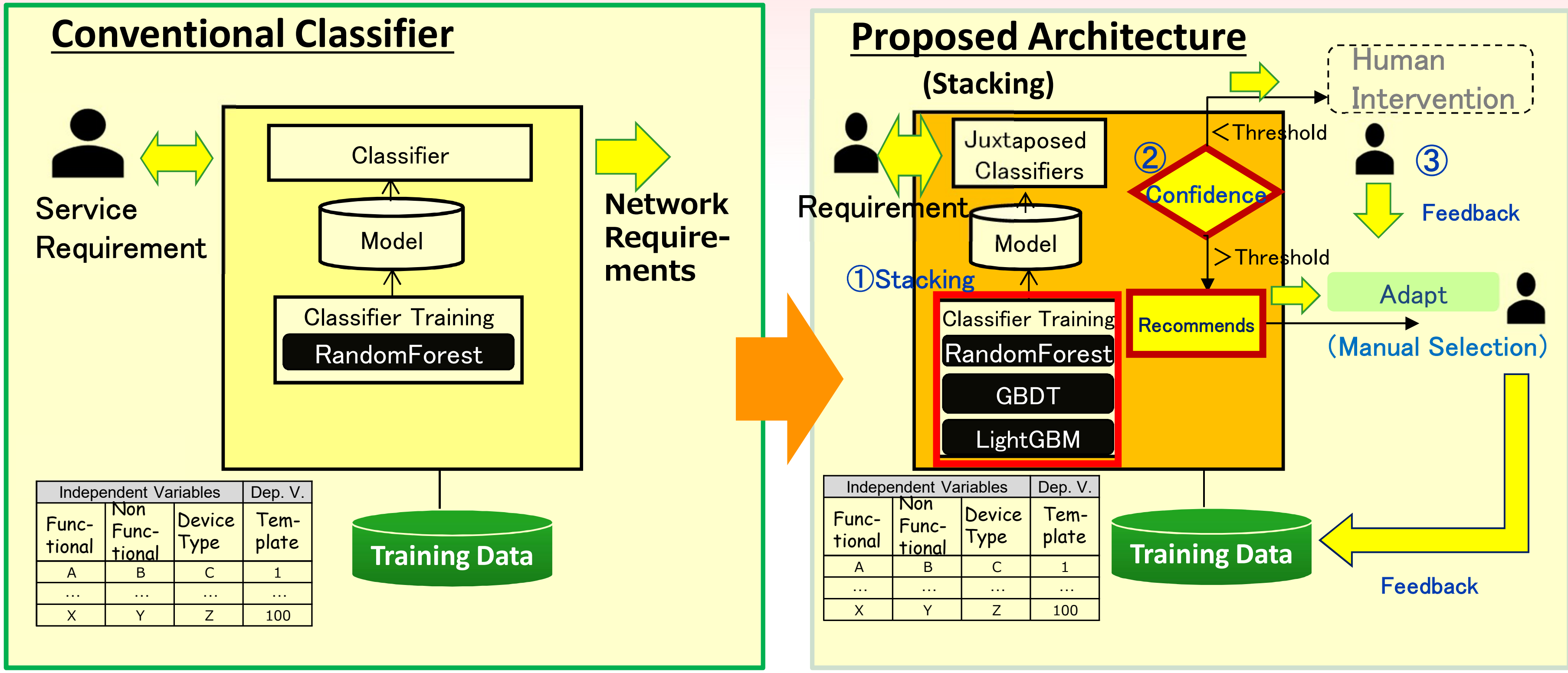
Applicable AI techniques

・Classification of service requirements
・Selection of a juxtapositional network requirement

- Classifier for decision tree with the training data from:
  - Networking system specifications (independent)
  - Service templates (dependent)



Scalable AI: Proposed architecture



- Issues: Noisy results caused by the characteristics of each classifier architecture

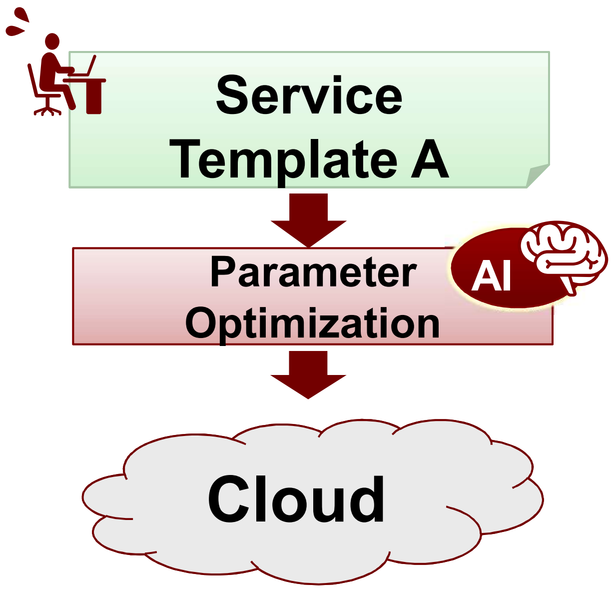
- **Improvement by "Stacking"**
- **Human intervention based on "Confidence"**
- **Transition learning from domain experts**

Outline

- **Research motivation**
  - 5G service scenarios
  - On-demand network slicing
- **Network service design**
  - Diverse requirements analysis
- **Network design**
  - Parameters setting
  - NW construction & deployment
- **Resource adaptation**
  - Agile control, QoS guarantee
- **Failure detection and recovery**
  - Dependable service
  - Fast recovery

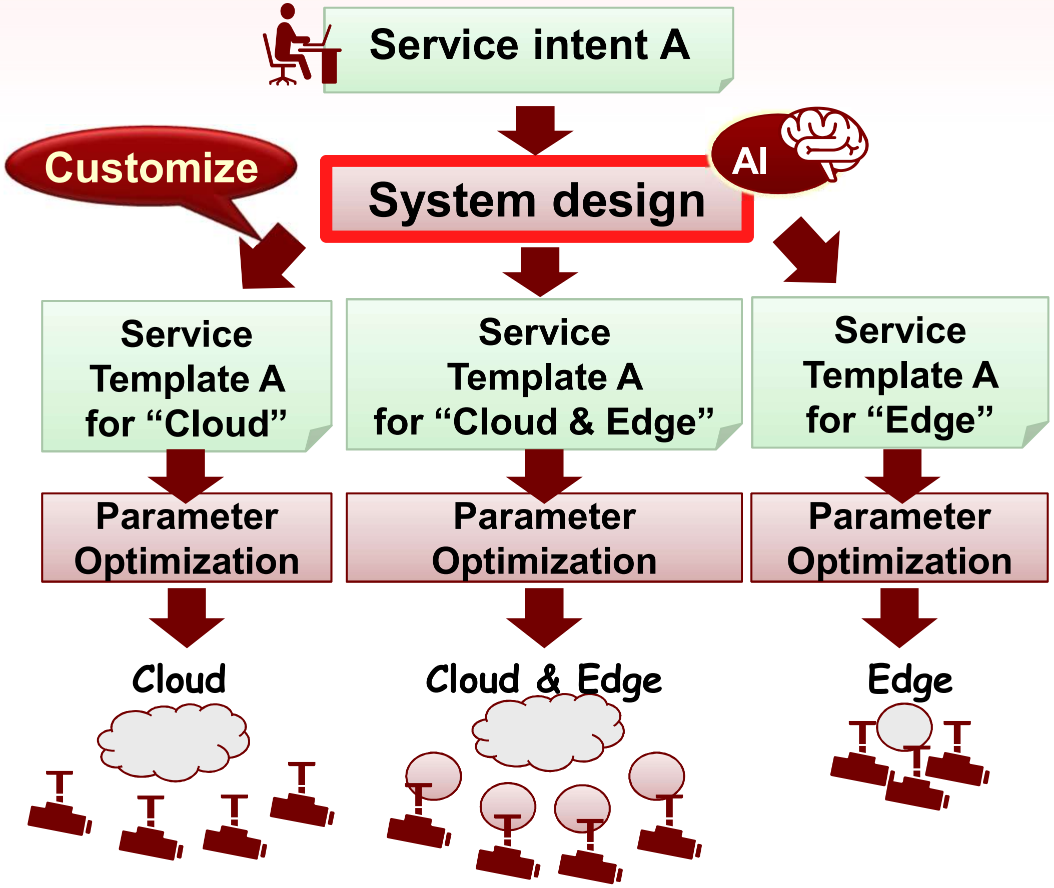
Challenge : Intent-based system design & deployment

System deployment on a fixed infrastructure



AI is being applied to mathematical optimization

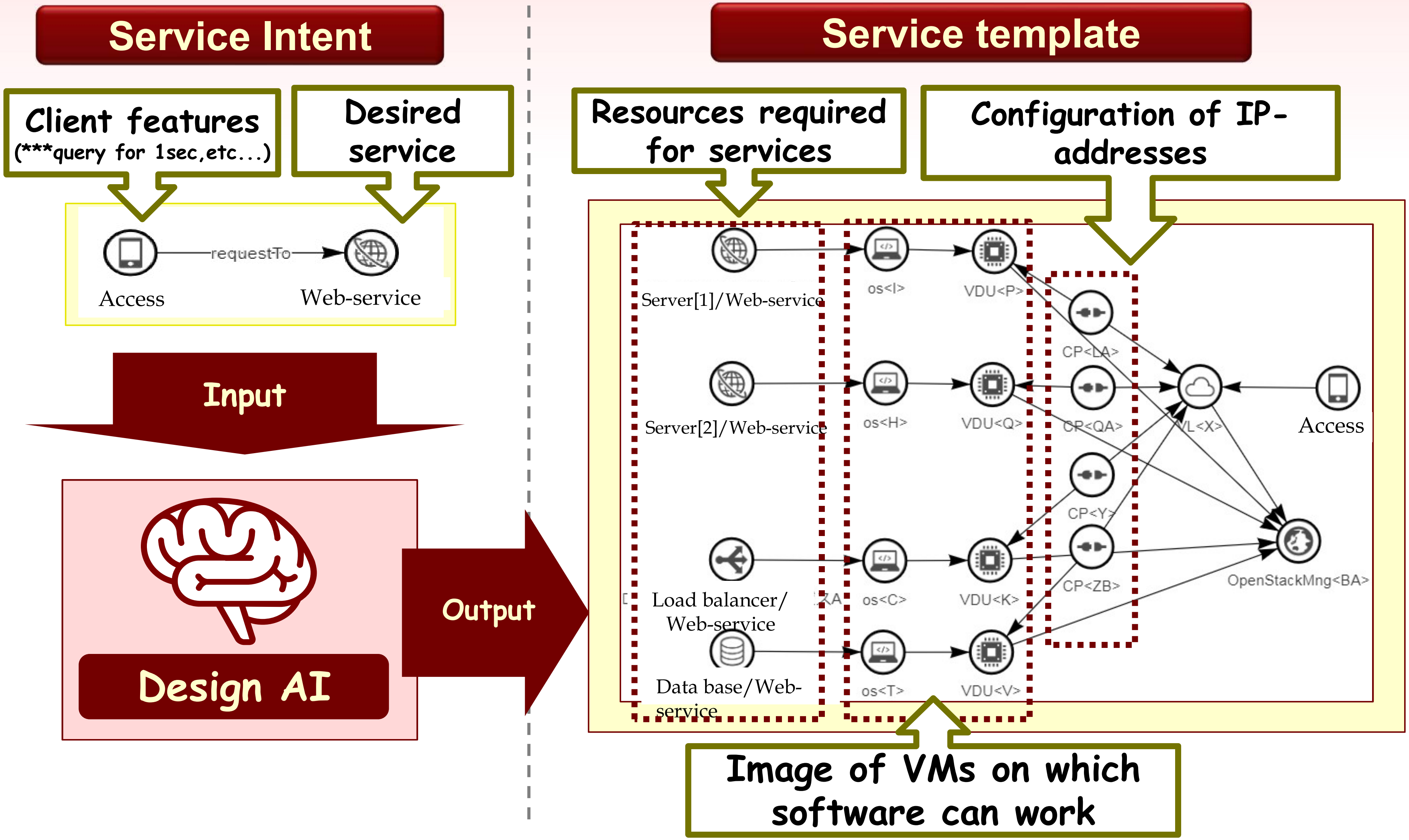
System deployment on variety of infrastructures



Appling AI for template customization (= system design) is our challenge

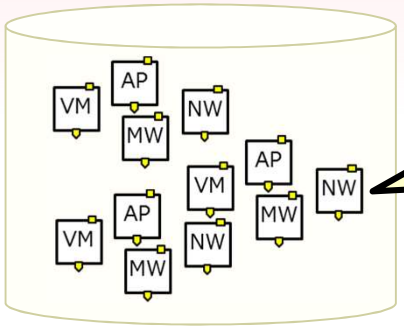
Automated system design with AI

Refining abstract service intent to concrete service template



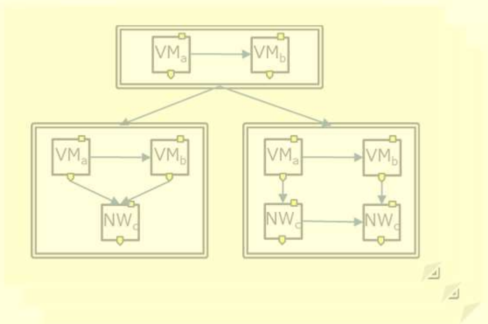
Procedure of automated system design (1/4)

Step 1: Preparing component models



Component models

VM:
hypervisor: xxx
flavor: yyy
image-id: zzzz-zzzz-zzzz
etc...



Fine-grained refinement patterns

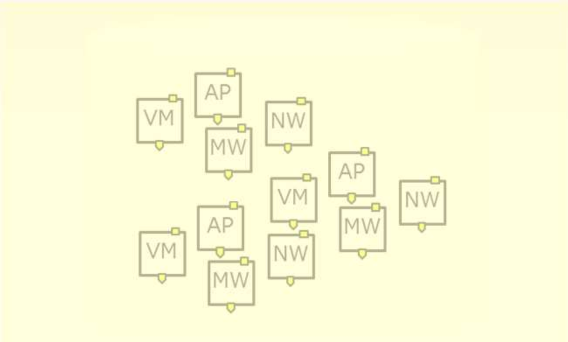
Model of system components which include...

- parameters
- connection constraints
- orchestration tasks

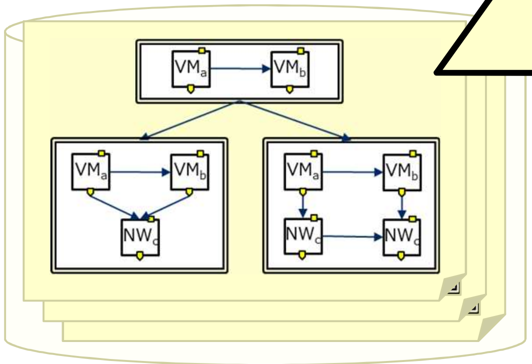
Procedure of automated system design (2/4)

Step 2. Preparing refinement patterns

Refinement options for every abstract part of intents

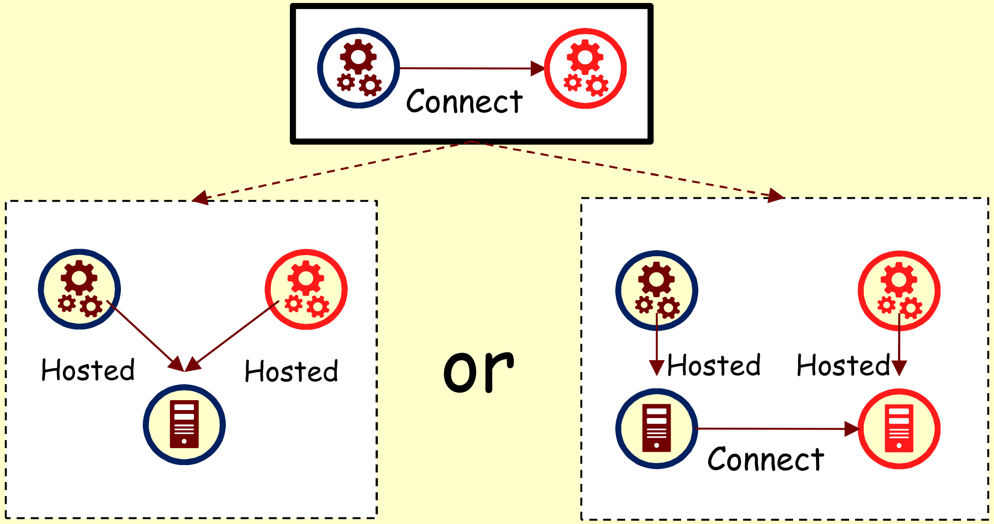


Component models



Fine-grained refinement patterns

How can I refine
"**HTTP connection between Apps**" ?

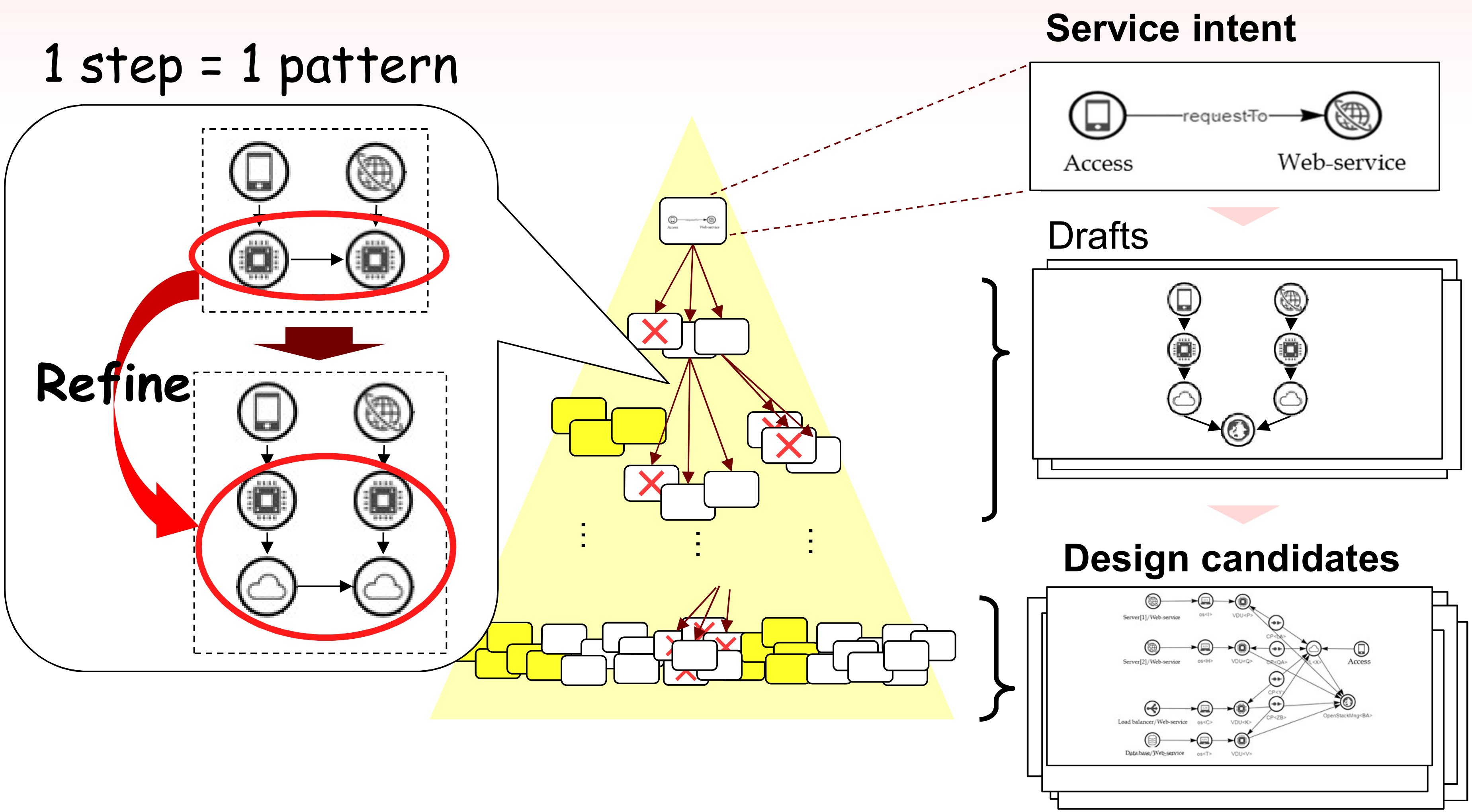


Assign them to the same machine

Connect the machines they hosted on

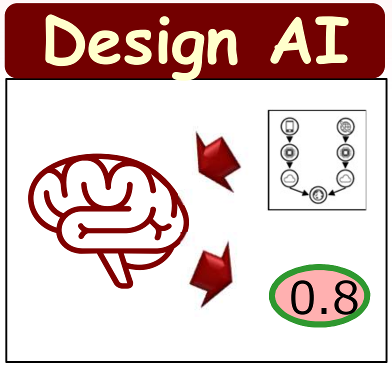
Procedure of automated system design (3/4)

Step 3: Generating various design candidates by patterns

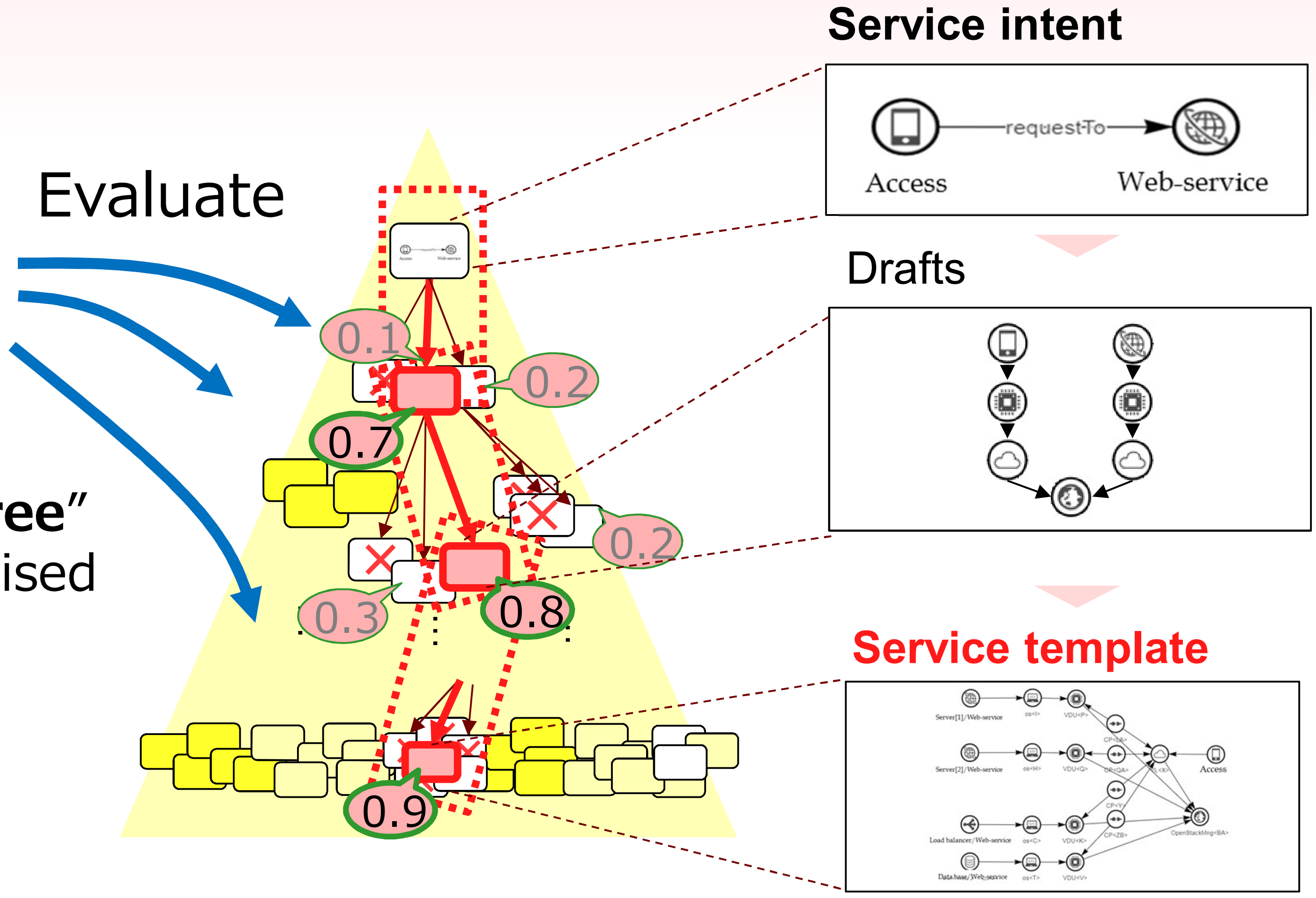


Procedure of automated system design (4/4)

Step 4: Selecting an appropriate design candidate by AI/ML



Evaluate
***"promise* degree"**
to choose promised
branches



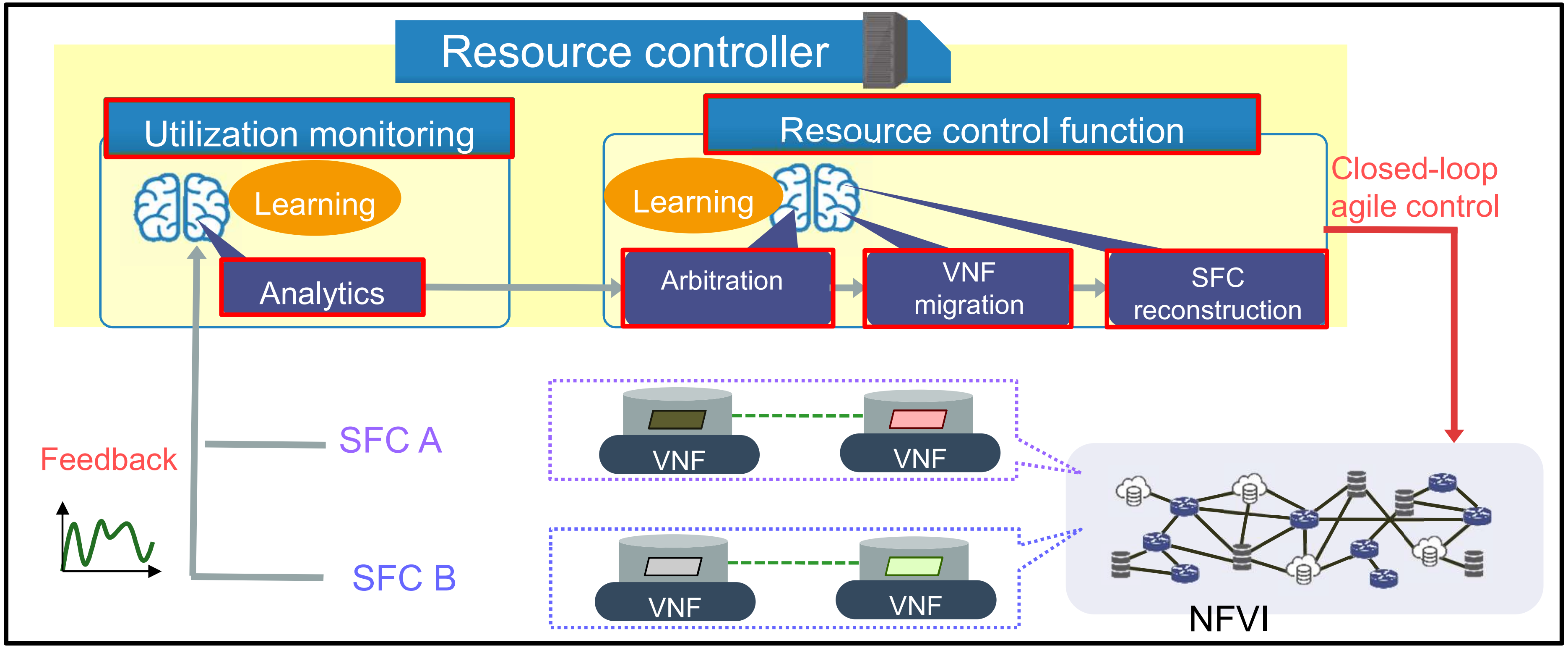
Outline

- **Research motivation**
  - 5G service scenarios
  - On-demand network slicing
- **Network service design**
  - Diverse requirements analysis
- **Network design**
  - Parameters setting
  - NW construction & deployment
- **Resource adaptation**
  - Agile control, QoS guarantee
- **Failure detection and recovery**
  - Dependable service
  - Fast recovery

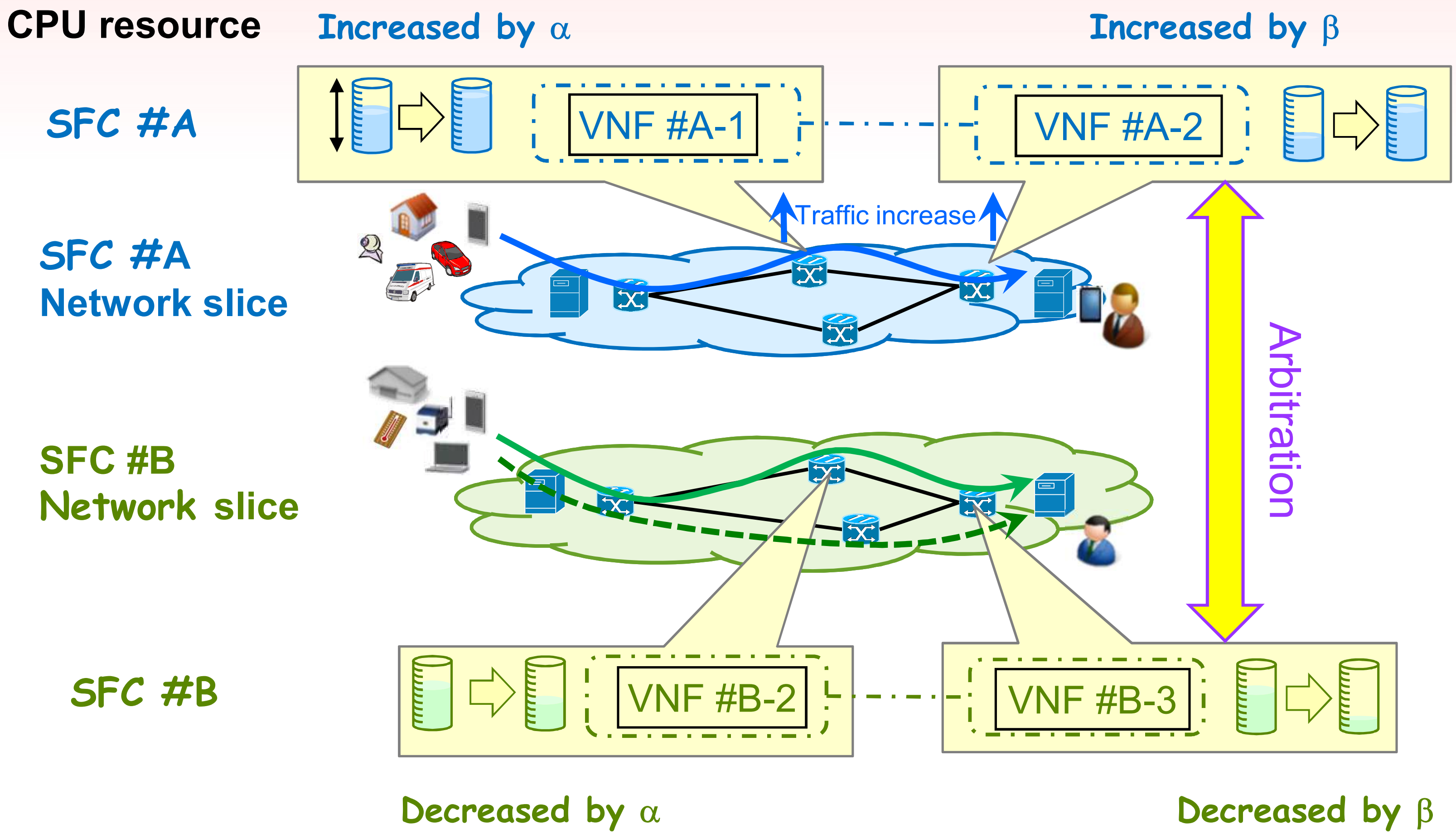
SFC dynamic resource adaptation framework

Objective: With AI-based network control、maintaining high-quality service while achieving agile resource arbitration and VNF migration

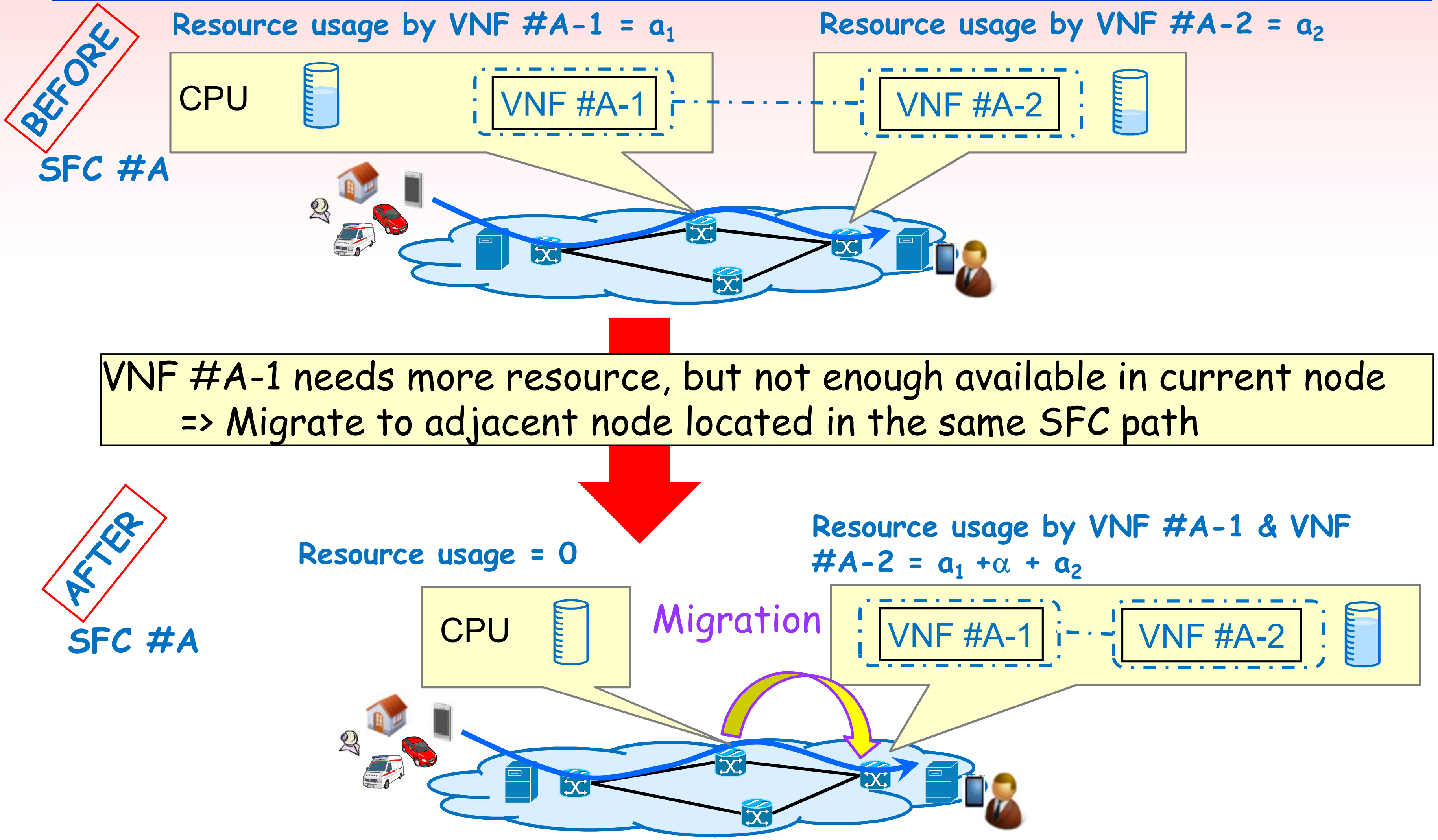
QoS/QoE × Utilization × Agility ・・Realizing simultaneously！



Dynamic resource arbitration among SFCs



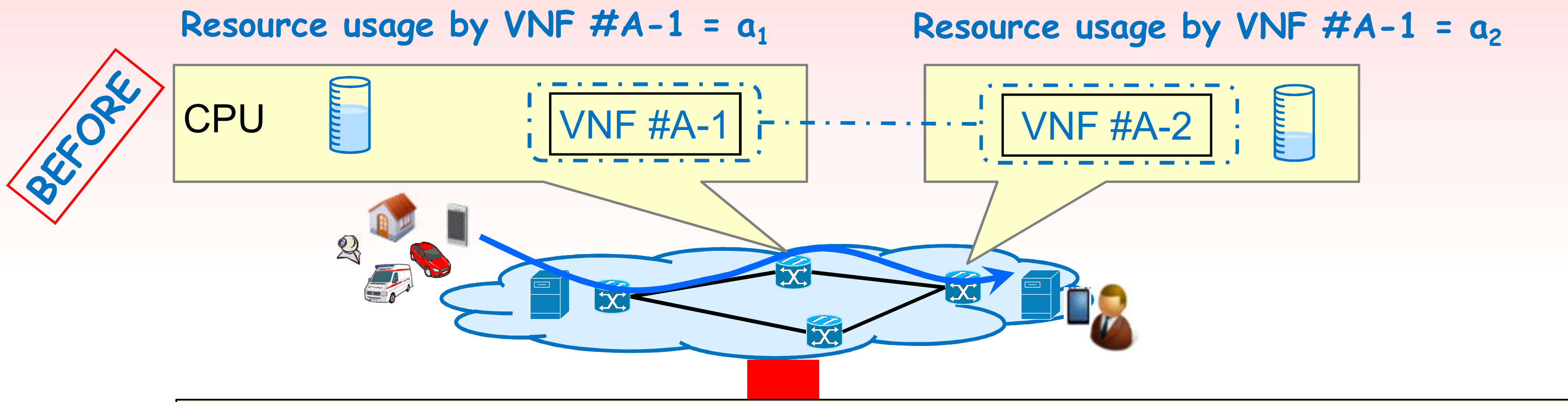
VNF migration in SFC



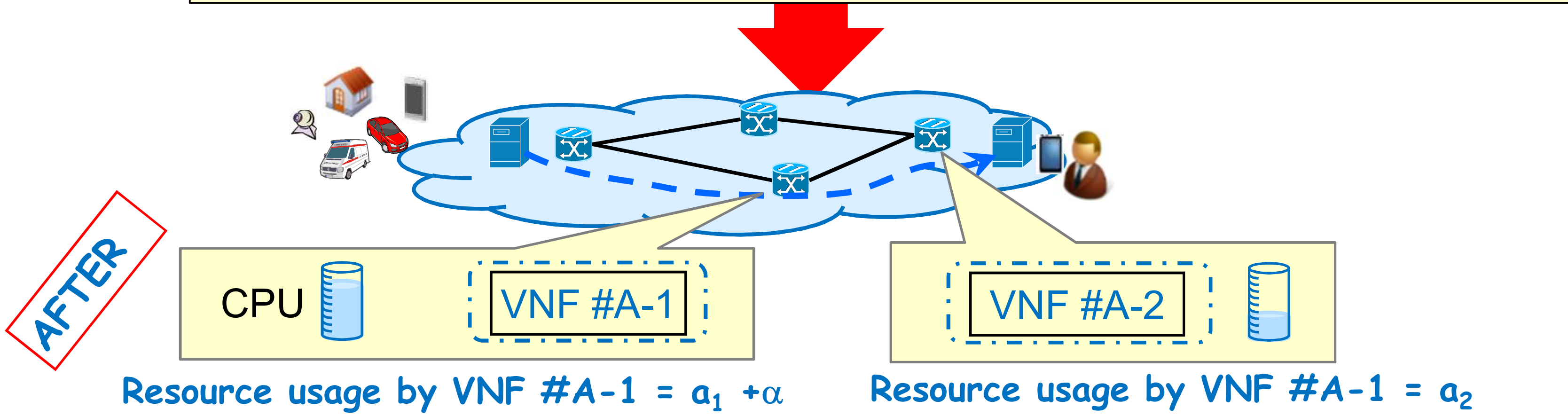
VNF #A-1 needs more resource, but not enough available in current node
=> Migrate to adjacent node located in the same SFC path

No change in SFC path => faster, agile adjustment

SFC reconstruction



VNF #A-1 needs more resource not available in the same SFC path
=> Reconstruct SFC

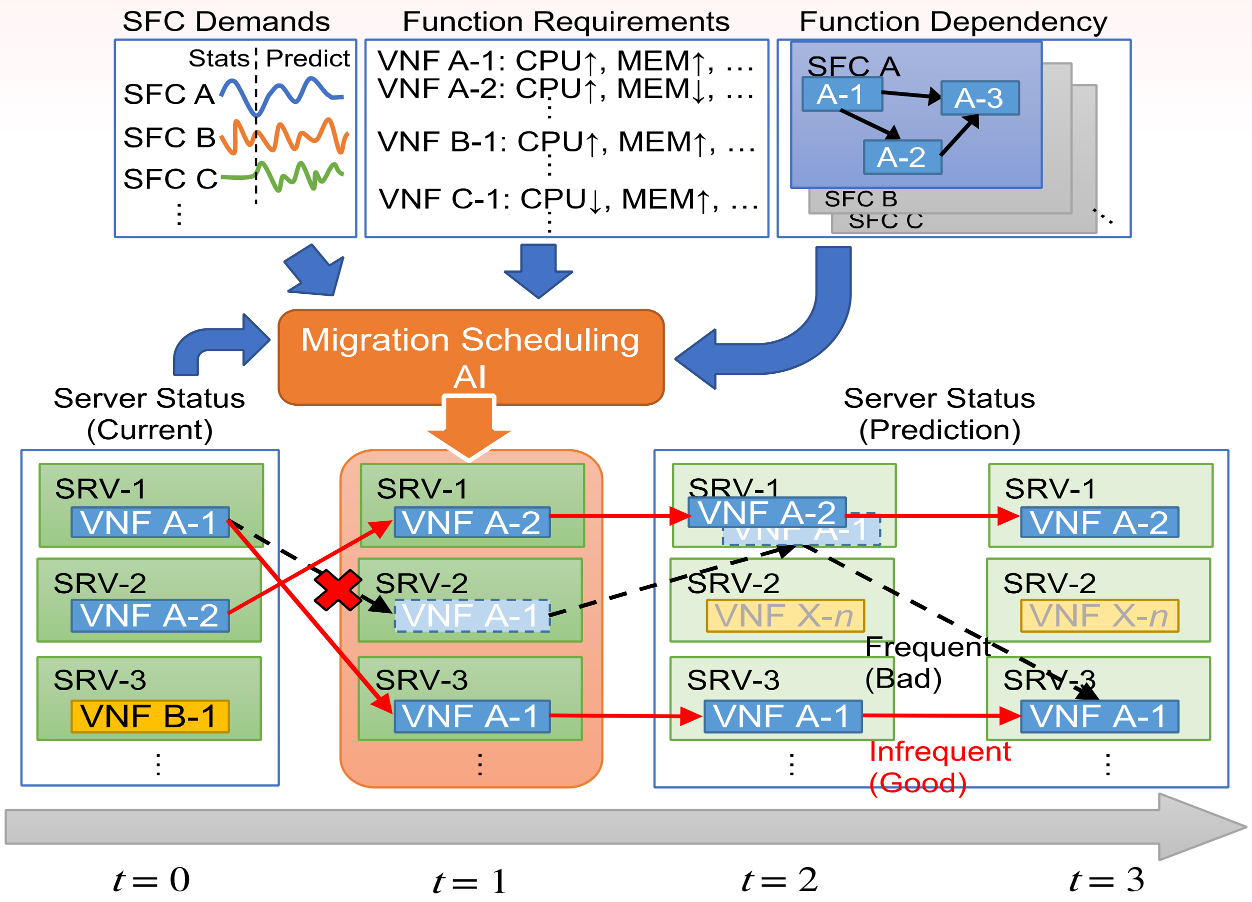


Realtime operation takes longer time.
=> Proactive machine learning approaches are being explored.

AI/ML for VNF auto-migration and SFC reconstruction (1/2)

VNF Migration Planning Based on Resource Demand Prediction

- **Objectives**:
  1. Meet resource requirements
  2. Minimize migration frequency



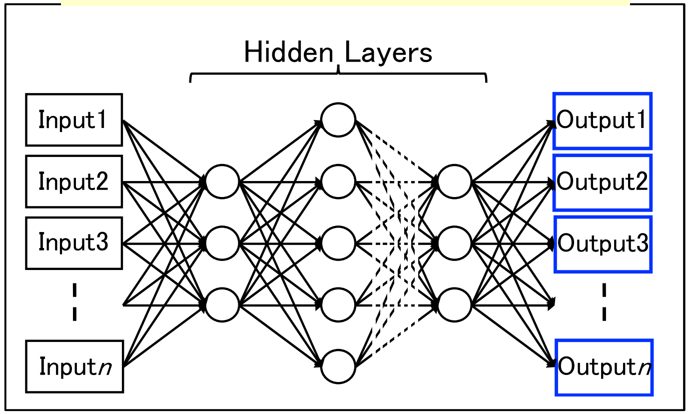
AI/ML for VNF auto-migration and SFC reconstruction (2/2)

Related work limitations

Integer Linear Programming

- Optimization for instance
  →Too many VNF migrations
- Optimization for long time-scale
  →Too much time to solve
  (several hours)

DNN: Deep Neural Network

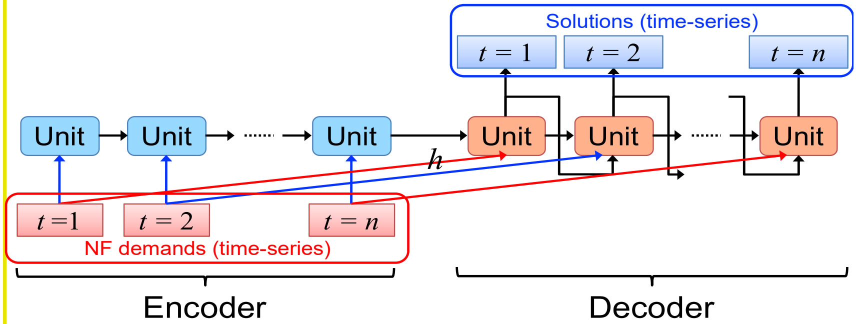


It takes 4~5 hours to train DNN
(100 cycles with 15,000 data)

Proposed approach

ED-RNN

Suitable for dynamic VNF migration technique by taking given time-series data of NFs resource demands as input.



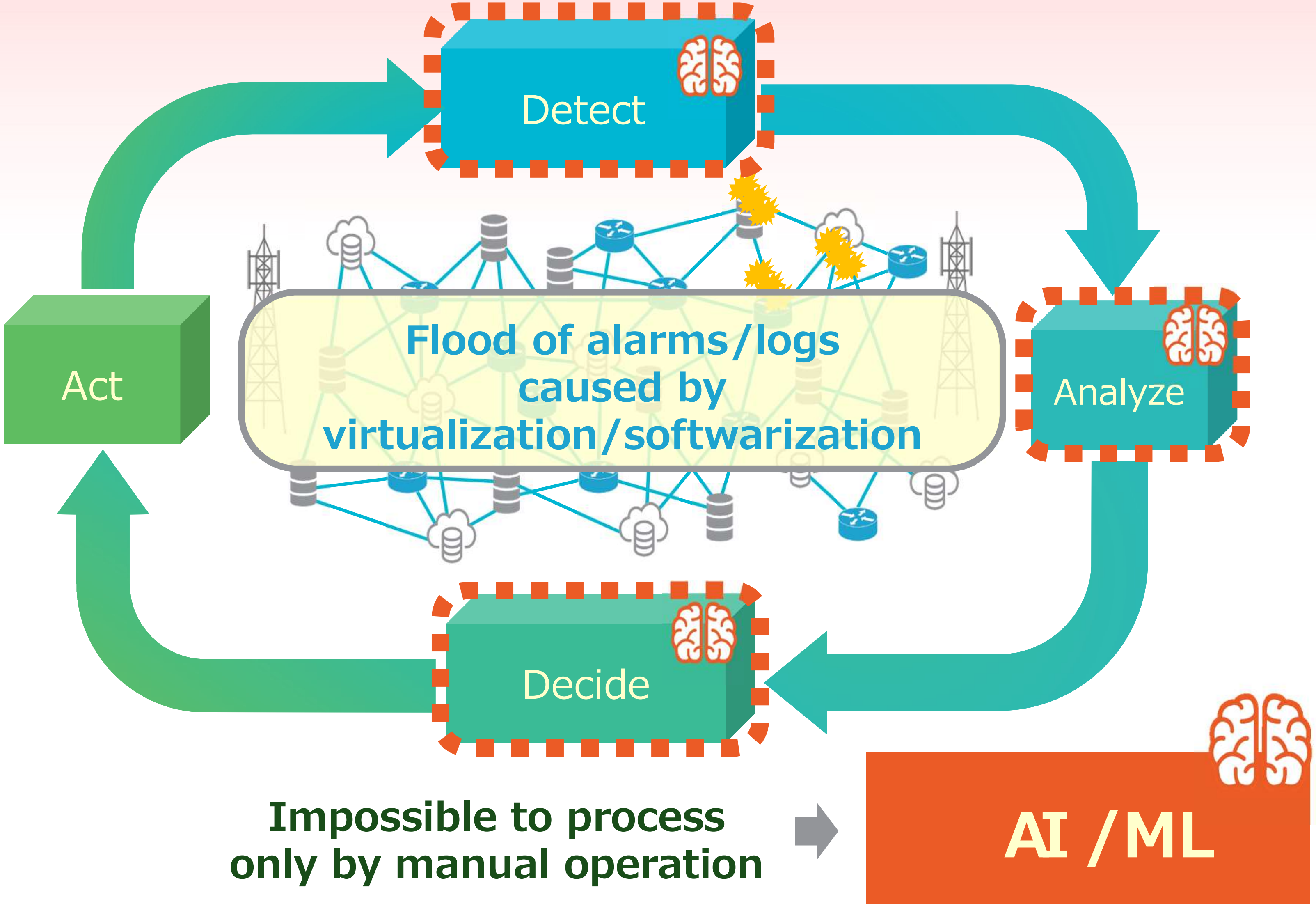
- It minimizes the situation of resource shortage and occurrences of VNF migration.
- It determines migration schedule quickly.
- It takes 1~2 hours to train ED-RNN (100 cycles with 15,000 data)

ED-RNN: Encoder-Decoder Recurrent Neural Network

Outline

- Research motivation
  - 5G service scenarios
  - On-demand network slicing
- Network service design
  - Diverse requirements analysis
- Network design
  - Parameters setting
  - NW construction & deployment
- Resource adaptation
  - Agile control, QoS guarantee
- Failure detection and recovery
  - Dependable service
  - Fast recovery

Closed-loop automation of network operation using AI

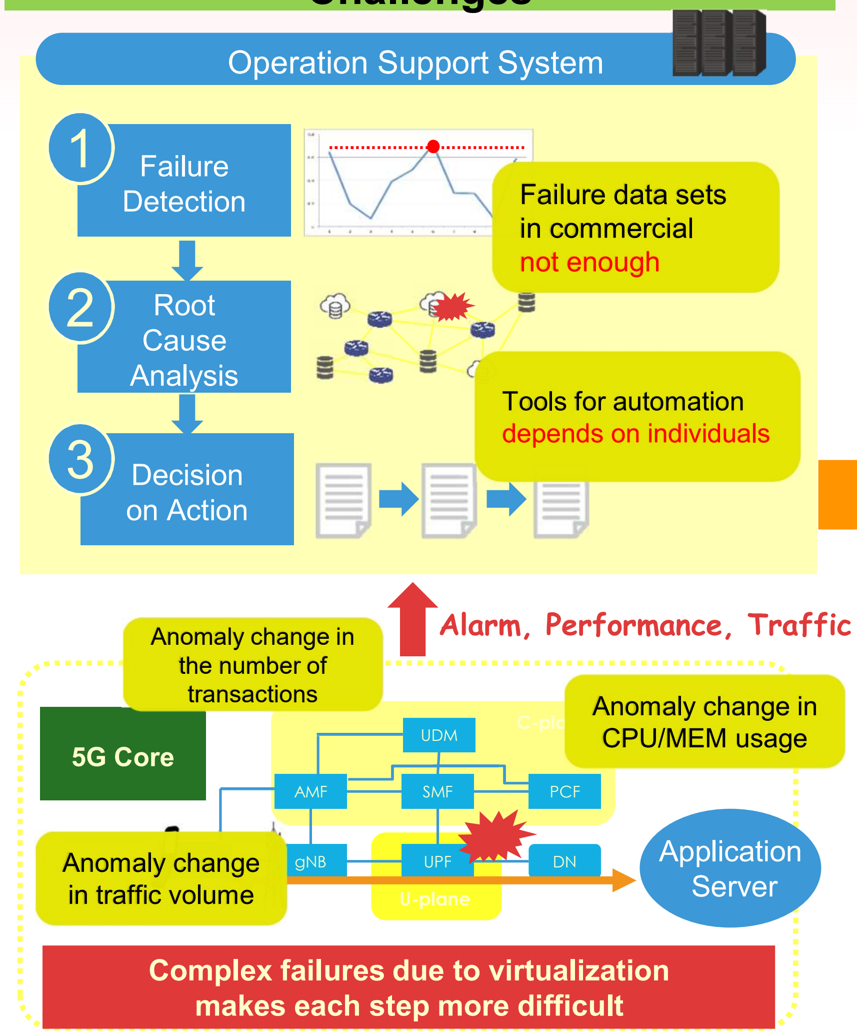


Impossible to process only by manual operation

Challenge and solution in network operation

The conventional automation based on thresholds/rules results in huge workload for maintenance

Challenges



Solution

Test network (NFV)

Repeated failures in test NW to collect enough training data sets

Data sets collected in various failures

Training Data Sets

Various AI engines using stored training data sets

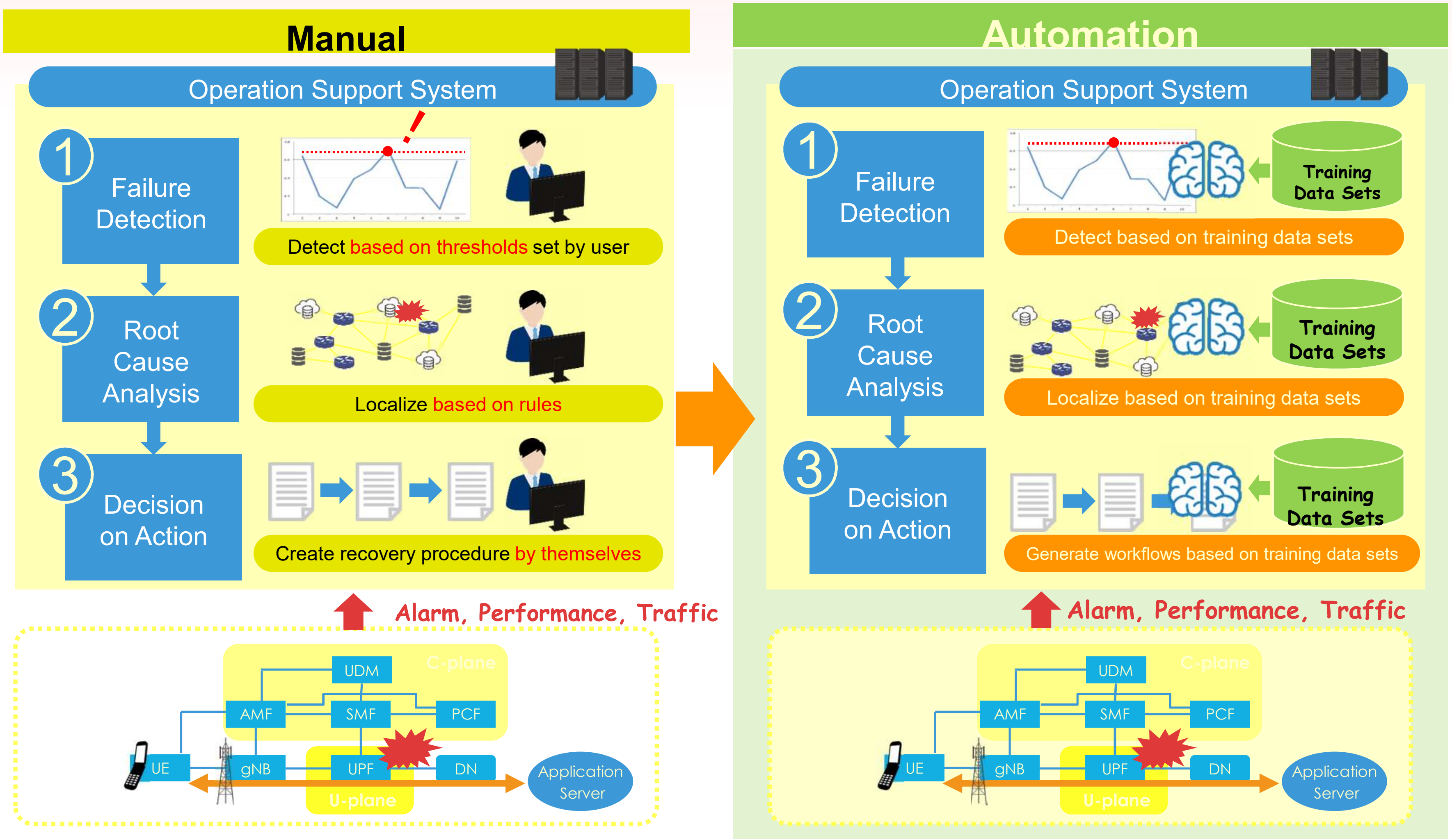
Commercial NW

Trained AI engines perform predictions for trouble-shooting in commercial NW

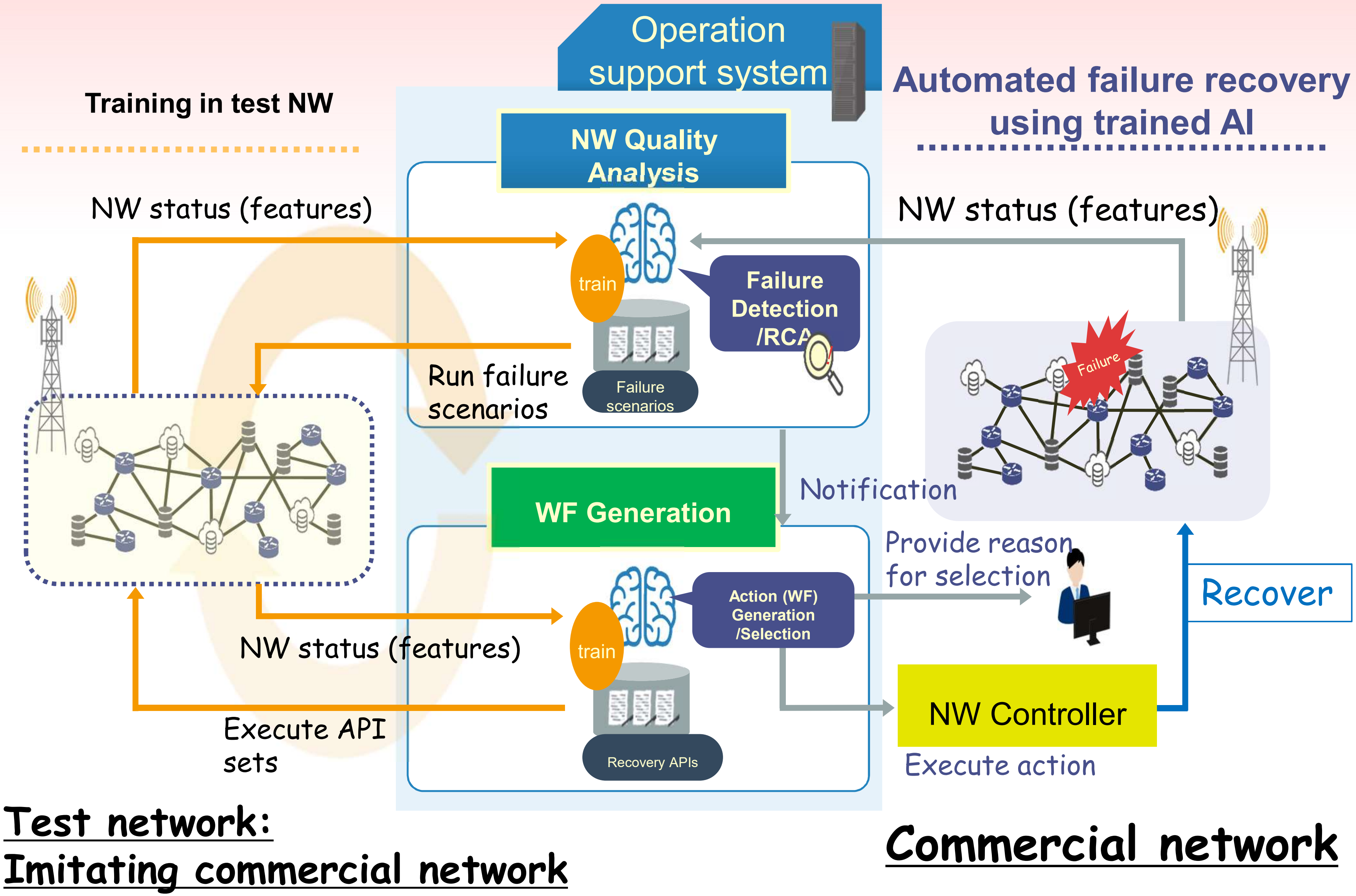
Develop AI for (1) failure detection, (2) RCA, (3) decision on action using big data collected from test NW

AI-supported automated failure recovery

Our proposed system utilizes AI engines to achieve automation of each step



Network operation support system using AI



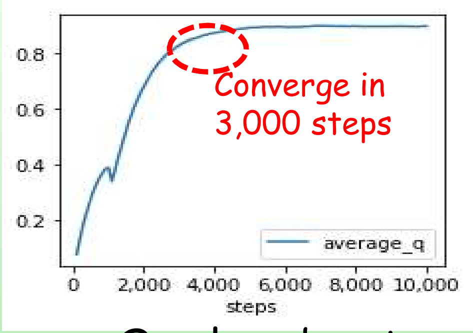
Training mechanism

Network Quality Analysis

		Prediction			
		Re-boot	Nor-mal	vCPU up	vCPU Rising
Actual	Re-boot	209	0	0	0
	Nor-mal	0	28,044	0	1
	vCPU Up	0	11	26	2
	vCPU rising	0	7	0	43

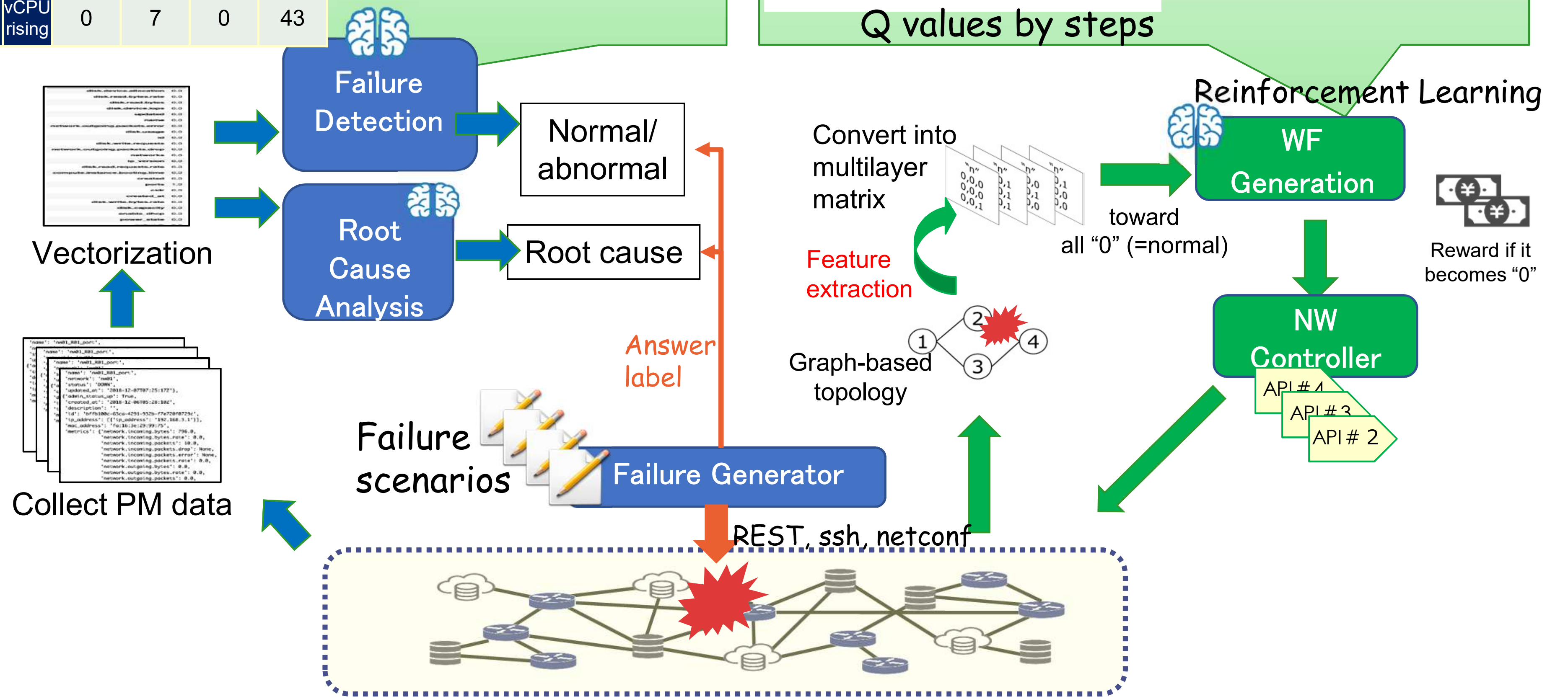
Example result
RCA using random forest achieved high performance

Work Flow Generation



Q values by steps

Example result
Training by DQN established correct actions in 3,000 steps



Conclusion

- Highlighted the need and applicability of AI techniques for the automation of network service design, deployment, adaptation and failure recovery.

- Covered four areas:
  - Service design
  - Network design
  - Resource adaptation
  - Failure detection and recovery

  - These four scenarios are also included in the use-cases and requirements deliverable (Sections 4.12 - 4.15) produced by FG ML5G.

Future work, acknowledgement

- Future work
  - Design of AI-based function architecture for network automation
  - Investigation of various AI-algorithms through experiments
  - Design of interfaces between functional components
  - Bringing contributions to ITU-T FG ML5G and SG13

- Acknowledgement
  - This work was conducted as part of the project entitled “Research and development for innovative AI network integrated infrastructure technologies” supported by the Ministry of Internal Affairs and Communications, Japan.