

R7 Architecture Document

- [Blueprint Overview](#)
- [Use Case](#)
- [Business Drivers](#)
- [Overall Architecture](#)
- [Platform Architecture](#)
- [APIs](#)
- [Hardware and Software Management](#)
- [Licensing](#)

Blueprint Overview

CFN, defined as computing force network, is a new information infrastructure that takes computing as the center, and network as the foundation, and deeply integrates Network, Cloud, Big Data, Artificial Intelligence, Security, Edge, Terminal and Blockchain to provide integrated services.

According to the unified ubiquitous scheduling management layer, the computing force of public cloud, on premise, edge cloud, or external third parties is managed to achieve consistent cluster, strategy, configuration, and traffic management, and to achieve resource-level and task-level scheduling.

Use Case

According to the CFN Ubiquitous Computing Force Scheduling, the computing force of public cloud, edge cloud, and external third parties is managed to achieve consistent cluster, policy, configuration and traffic management, and to achieve resource-level and task-level scheduling.

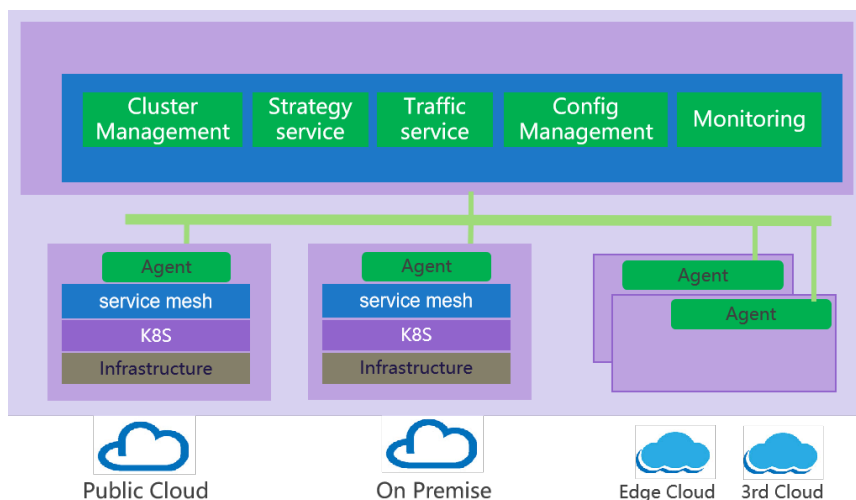
* Ubiquitous Computing Force concept: Logically, the computing force is more three-dimensional, including three levels: center, edge, and terminal. Physically, resources span data centers in different regions. The kernel is heterogeneous, including general computing force (x86/ARM) and special computing force (GPU/DPU...).

In this blueprint,

Business Drivers

Computing Force NetworkCFNputs forward higher requirements for the management and collaboration ability of ubiquitous computing power. Ubiquitous computing power scheduling technology involves cross-layer scheduling of cloud, edge, and end multi-level computing power, and it needs to meet the scene requirements of various heterogeneous computing power management and multi-party social computing power access. At present, the core management framework, key technical solutions, and products of ubiquitous computing scheduling technology are gradually developing, and have not yet entered the mature stage of technology.

Overall Architecture



Road map:

(already done) 1Scheduling computing force by cluster weight. already release in R7;

(already done) 2Rescheduling computing force when a cluster resource is abnormal. already release in R7;

(next step) 3. Designing of integrated computing and network resource scheduling algorithm (next step, envisioned Signpost)

By the end of this stage, the CFN BP focuses on the functional integration of computing force scheduling in the ubiquitous computing environment composed of a number of heterogeneous computing nodes. Computing resource-demanding applications from OTT partner, MIGU, are successfully deployed under the guidance of the scheduling decision in our experiment. However, as CFN BP is still in an early stage, the network topology is simple and the number of nodes is limited. Thus, no modern network measurement techniques are deployed in the system. The scheduler only made scheduling decisions based on coarse information about the network, i.e., accessibility. The computing capability in nodes played a major role in the decision process, which is fine for the present experiment, but not sufficient for the final goal of the CFN BP.

In the future stage, on the one hand, when more nodes are added to the experiment environment, the complexity of network topologies will emerge. Thus, the computing task scheduling results without timely and sufficient information on the network will probably impact the performance of the application. Integrated computing and network resource sensing will be of first importance, which forms the basis for making reasonable scheduling decisions. On the other, as for applications, when micro-service architected or cloud-native applications are introduced in the experiment. The sequential dependency between microservices components in the application should be carefully considered when planning the distributive deployment, adding more complexity to the task scheduling problem. The rapid development of graph learning makes feature extraction from graphic data about networked infrastructure and componentized applications fast and easy, greatly facilitating the training of deep learning models and reinforcement learning agents. Thus, artificial intelligence techniques-based task and resource scheduling approaches are considered in the future stage of the CFN BP.

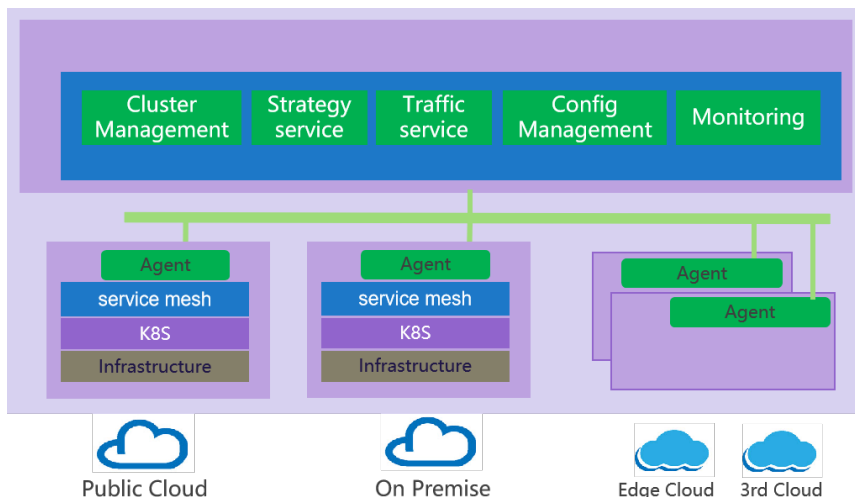
(next step) 4. Scheduling algorithm enhancement considering higher-level factors.next step, envisioned Signpost

As the CFN is intended to provide service to wide consumers, in spite of technique issues, the computing and network resource scheduling will be also influenced by some higher level factors, such as the preference of computing force customers, intents of computing, and network resource suppliers, etc. Practically, consumers will not only care about how fast their computing tasks are finished and how good the performance of the desired service is but also how much is the cost. And different consumers may show different sensitivity to the cost, which deserves further work on mapping from subjective preference to objective measurable metrics. Moreover, some consumers may also care about the robustness of their application, thus duplicate and redundant deployment of applications or just critical exponents should be taken into consideration.

The intent of resource suppliers will also impact resource scheduling. Some suppliers may prefer to use resources in a computing node with first priority to avoid the appearance of fragmented resources. In contrast, other suppliers may want their resources in different nodes to have balanced usage.

Only when these factors are quantitatively analyzed and characterized, can they be adopted in the design of scheduling approaches, and high-level preferences and intents can be reflected in the computing force scheduling. These works will also be considered in the next stage of CFN BP.

Platform Architecture



According to the unified ubiquitous scheduling management layer, the computing force of public cloud, on premise, edge cloud, or external third parties is managed to achieve consistent cluster, strategy, configuration, and traffic management, and to achieve resource-level and task-level scheduling.

APIs

Hardware and Software Management

Licensing