Edge Cloud Deployment of Enterprise Communication

Zero-Touch Service Management for End-to-End Telecommunication Applications

Opportunity

Cognitive Cloud Control (CCC) assembles AI and ML technologies into a service management framework to control and orchestrate cloud scale virtual network functions (NFV) in heterogeneous cloud infrastructures; these cloud applications autonomously manage themselves, removing cost and time for manual human intervention and providing outstanding user experience. Demands will vary constantly as well as new applications and services will be created and deployed at high frequency into these infrastructures. Bandwidth and latency requirements will complement computing requirements like ultimate scalability and high efficiency. Cognitive technologies will release the burden from the developers to manage such complex systems by learning the relevant parameters and drive optimized control actions during the run-time live in the operational environment.

Our solution consists of:

- Service model-generators identify resource dependencies, needs, and bottlenecks without expensive testing.
- Workload-prediction algorithms anticipate service demand changes to provide enough resources ahead of time.
- Self-management processes and algorithms generate automated configuration scripts for fast, convenient software deployment

Thus, the developers can focus on what they do best - providing new features and optimized experience to their users.

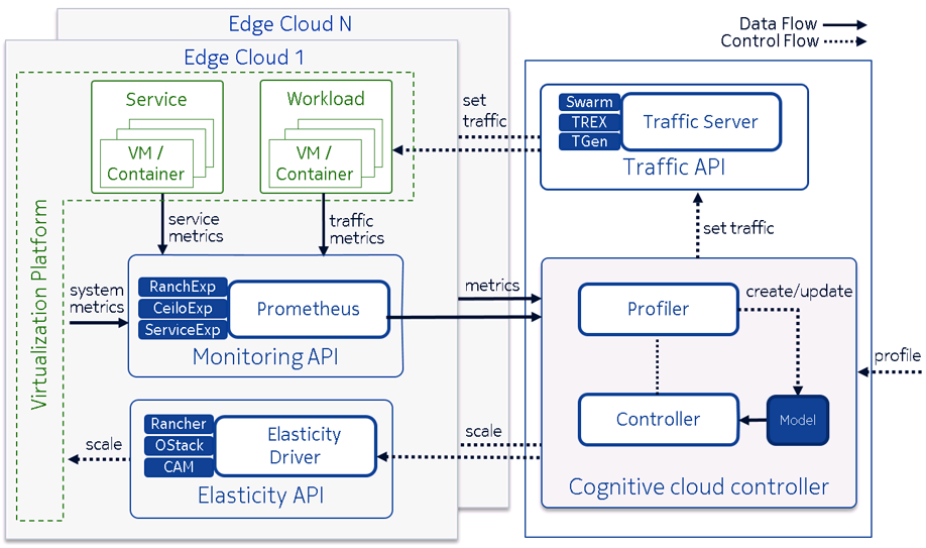

Fig. 1: Cognitive Cloud Control Architecture

Breakthrough Innovations

CCC is implemented by a combination of a metric processing machine learning pipelines in Apache Beam, combined with finite state machines (FSMs) that execute a collection of machine learning and workload forecasting algorithms.

The analytics function detects resource bottlenecks and dependencies by monitoring platform, service, and KPI metrics. In order to learn how KPIs may be affected by resource bottlenecks, CCC employs an innovative bootstrapping phase in the finite state machine which allows safe scale-in/out operations. The information gained during this phase is used for an elastic (scale in/out) model of the service that directly relates workload demands to KPIs.

In contrast to previous approaches based on offline profiling, CCC's models are constantly updating. This not only provides higher robustness against changes in the traffic composition, but also ensures the model adapts easily.

In order to apply the models proactively, CCC forecasts the amount of workload for each relevant type of traffic. CCC focuses on short term predictions, since they are typically accurate enough and can be integrated without significant overhead. Several out off the shelf forecasting techniques have been implemented, and CCC selects the best approach on-the-fly for the current situation.

Description

The demonstration shows the end-to-end setup we have built in the FutureX experience running in Murray Hill, USA. There we are running multiple cloudified products, including Nokia's session border controller (SBC), a video analytics application and a virtualized mobile broadband unit. Surveillance cameras stream video towards the analytics system to detect security and safety breaches. For all these components we are building scaling models by analyzing resource utilization metrics and instantiate the services and update the deployment as demand increases.

Additionally, we visualize how these models are built.

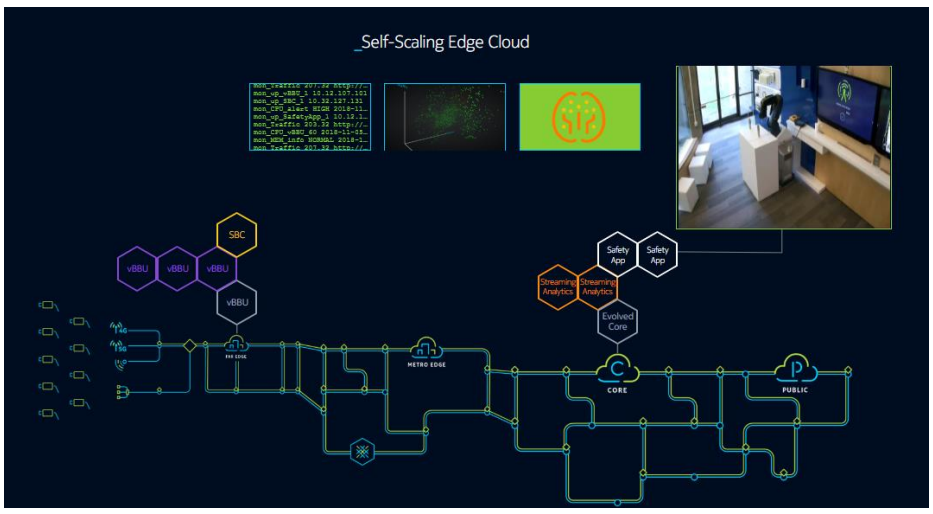

Fig. 2: End-to-end deployed service view

Contact

diego.lugones@nokia-bell-labs.com
juergen.sienel@nokia-bell-labs.com
volker.hilt@nokia-bell-labs.com

© 2019 Nokia Solutions and Networks GmbH & Co. KG - Innovation Days Stuttgart – 26. - 28. November 2019