

Towards Network-Aware Resource Provisioning in Kubernetes for Fog Computing applications

Abdul Ahad Ayaz

Summer term 2020

Abstract -

1 Introduction

In recent years with the evolution of technology, Internet of Things (IoT) devices are increasing day by day. According to Ericsson mobility report[], there will be 17% (approx. 22.3 billion) increase in IoT devices by 2024. Functionally IoT is defined as *"The Internet of Things allows people and things to be connected Anytime, Anyplace, with Anything and Anyone"* [European commission 2008]. IoT devices have served mankind in many ways such as from smart houses to smart cities, smart transportation systems and many medical applications. These IoT applications enables many devices connected to network and generates alot of heterogeneous data also known as BigData which requires special data processing models and Infrastructure support. Processing BigData required alot of resources and cloud computing theoretically provides it unlimited resources[fog-comp-survey]. But there is downside of using cloud computing for such complex computation as it is more costly when it comes to computation power, storage and bandwidth. Computation need to be performed at the node level and only the aggregated data need to send to central node for further computations and analysis. This de-centralized approach will save alot of computation power as well as bandwidth requirments[fog-comp-survey]. To overcome the downside of cloud computing, the terminology fog computing is used. Fog computing allows the computation at the egde of network in-

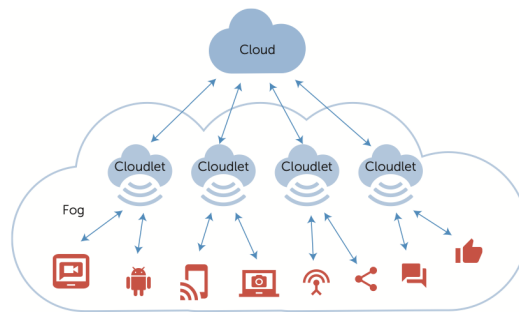


Figure 1: Fog computing: Top-level overview[5]

stead of central core.

Fog computing is defined as *"an architecture that uses one or a collaborative multitude of end-user clients or near-user edge devices to carry out a substantial amount of storage (rather than stored primarily in cloud data centers), communication (rather than routed over the internet backbone), and control, configuration, measurement and management (rather than controlled primarily by network gateways such as those in the LTE (telecommunication) core)"*[Chiang 2015; Aazam and Huh 2014].

Traditionally, user applications running in cloud access the cloud core network through access points for data exchange to fetch data from data-centers [5]. In fog computing these access points also serves as resource providers such as computation power and storage etc. and are called "cloudlets"[5]. Figure 1 show the top-level architecture of the fog computing.

Fog computing is responsible for providing resources to IoT devices for processing. Traditionally these resources are allocated as VMs from different cloud infrastructures such as AWS, Google, Open-Stack, etc. to run the applications. VMs are considered resource greedy and require more computational resources. Alternate is to use the Containers such as Docker which are light-weight, requires less resources and based on micro-service architecture. Large applications are split into containers based on the main processes of the application. This increasing number of containers per application required the proper monitoring for health check and resource consumption. The most commonly used orchestrator for containers is Kubernetes.

Kubernetes act as IaaS for fog computing to provide resource for IoT applications. Kubernetes is an open-source platform for management, deployment and scaling of containers. In Kubernetes, applications are deployed as pod consisting of multiple containers. When the configuration of deploying application is passed to Kubernetes, it checks for the availability of resources and deploys afterward. Kubernetes default resource scheduler monitor and deploys the pod using computation power-based scheduling mechanism and does not consider latency and available bandwidth, which is considered important while dealing with data-centric application. Example of data-centric application is weather forecast that receives data from scattered IoT devices and provide prediction. If the data is lost or delayed due higher latency and poor bandwidth, timely decisions cannot be made that leads to disaster. To overcome this drawback of Kubernetes, author proposed an alternate Kubernetes scheduler that consider network resources along with computational resources.

2 Backgroud

- Kubernetes Internal Archietecture and Main Components
- Kubernetes works as an Orchestrator
- Kubernetes resource provisioning
- Concluding the section with pitfals of default scheduler of Kubernetes

2.1 Kubernetes Main Components

- Write about the Archietecture of Kubernetes with diagram
- Write about the building blocks of Kubernetes and their working

2.2 Kubernetes as Orchestrator

- Orchestrator main functions
- Comparison of available Orchestrator (Open-Stack vs Kubernetes)
- Workflow of Kubernetes as an Orchestrator (steps)

2.3 Kubernetes Resource Provisioning

- write about the default Kubernetes scheduler
- its main Components
- workflow of default scheduler

3 Kubernetes Network-based Resource Provisioning

- write about why we need network-based resource provisioning
- main factors consideration (e.g bandwidth and latency)
- workflow of network-based scheduler

4 Performance Evaluation

- Write about the considered use-case of Fog Computing for Evaluation

4.1 Expermentation Setup

- setup of Kubernetes base on the mentioned use-case of Fog Computing with diagram

4.2 Analysis of Kubernetes Default and Network-based Resource Provisioning

- write about the Performance difference between default Kubernetes scheduler and network based scheduler with supporting result tables and graphs

5 Comparison of Network-based Resource Provisioning Solutions

- Compare different solutions based on the following criteria:

5.1 Orchestrator

- write about the differences between Kubernetes(main-paper)[3] and other available cloud solutions such as Fogernetes[4] and [2].

5.2 Resource Provisioning Techniques

- difference between different resource scheduling techniques such as [5], [1] etc.

[2] A. Reale, P. Kiss, M. Tóth, and Z. Horváth. *Designing a decentralized container based Fog computing framework for task distribution and management*. Tech. rep. <http://www.naun.org/main/UPress/cc/2019/a022012-044.pdf>. 2019.

[3] J. Santos, T. Wauters, B. Volckaert, and F. De Turck. “Towards network-Aware resource provisioning in kubernetes for fog computing applications”. In: <http://physics.nist.gov/Document/sp811.pdf>. IEEE Conference on Network Softwarization Unleashing the Power of Network Softwarization, NetSoft 2019, 2019.

[4] C. Wöbker, A. Seitz, H. Mueller, and B. Bruegge. “Fogernetes: Deployment and management of fog computing applications”. In: <https://ieeexplore.ieee.org/document/8406321>. IEEE/IFIP Network Operations and Management Symposium: Cognitive Management in a Cyber World, NOMS 2018, 2018.

[5] L. F. Bittencourt, J. Diaz-Montes, R. Buyya, O. F. Rana, and M. Parashar. “Mobility-Aware Application Scheduling in Fog Computing”. In: (2017). <https://ieeexplore.ieee.org/document/7912261>.

6 Conclusion

7 Further Research Topics

- after writing the seminar, if there is any improvement that can be done, will be added in this section.

References

- [1] D. Haja, M. Szalay, B. Sonkoly, G. Pongracz, and L. Toka. “Sharpening Kubernetes for the Edge”. In: <https://dl.acm.org/doi/10.1145/3342280.3342335>. SIGCOMM 2019 - Proceedings of the 2019 ACM SIGCOMM Conference Posters and Demos, Part of SIGCOMM 2019, 2019.