Resource and Service Management in Fog Computing

Shehenaz Shaik
Dept. of CSSE, Samuel Ginn College of Engineering
Auburn University
Auburn, AL, USA
szs0117@auburn.edu

Sanjeev Baskiyar
Dept. of CSSE, Samuel Ginn College of Engineering
Auburn University
Auburn, AL, USA
baskisa@auburn.edu

ABSTRACT

Fog computing paradigm is emerging as complementary to cloud computing paradigm to realize deployment of large scale IoT environments supporting widely dispersed IoT devices, users, and corresponding applications, leveraging fog nodes of varied resource configurations located in vicinity. Researchers have shown the need of fog computing towards deployment of latency-critical and bandwidth-intensive applications. Management of resources and services is critical in widely dispersed fog environment with heterogeneous fog nodes to ensure optimal utilization of available infrastructure and energy resources on fog nodes and IoT devices. Towards efficient management of fog, we proposed Hierarchical and Autonomous Fog Architecture (HAFA) to organize heterogeneous fog nodes into a multi-layered connected hierarchy based on several parameters such as physical location, distance from IoT devices and/or users, node resource configuration, privacy and security. The initial results show the ease of search for an optimal fog node with required resource configuration towards deployment of application services.

CCS CONCEPTS

• **Computing methodologies → Distributed computing methodologies**

KEYWORDS

Internet of Things; Cloud Computing; Fog Computing; Architecture; Resource management; Service Management;

1 INTRODUCTION

The variable latency introduced by WAN limits the deployment of IoT applications with strict QoS needs. As cloud computing paradigm falls short to deploy IoT applications with such requirements, fog computing paradigm holds promise. Fog computing extends cloud computing to include physical nodes of varied resource configurations and capabilities distributed over vast geographical areas to support latency-sensitive and data-intensive applications. In addition, fog environments allow deployment of location-sensitive applications i.e. location-aware, and those of only local value.

Research Challenges. Applications as well as infrastructure resources such as compute, network, and storage can be offered in fog as services similar to those in cloud. In spite of this similarity, management of fog differs from that of cloud for several reasons. Contrary to cloud, fog environment includes heterogeneous nodes dispersed over wide geographic regions, possibly at unmanaged sites, and identified by their physical locations, which is significant in delivering location-sensitive, context-aware services and those only of local value. Additionally, there is need for efficient placement of application services to satisfy their resource requirements as well as optimize various factors such as node utilization, network utilization, service execution cost, energy consumption, performance, availability, and load balancing. Considering the huge number of users, applications, service requests, IoT devices, and fog nodes, service deployment approach complexity should be independent of these factors. Additionally, solution approach should support mobility of users, and IoT devices, as well as heterogeneity in fog. Wide geographic dispersion of fog nodes make it difficult and sometimes infeasible to maintain entire system state at a centralized authority. Thus, there is need for efficient management and allocation of fog infrastructure resources such that they can be allocated in an optimal fashion to application services with minimum overhead, reducing overall system maintenance costs and maximizing utilization of fog nodes assuming no knowledge of complete system state. The solution should be preferably a distributed approach, owing to the dispersed nature of fog environment.

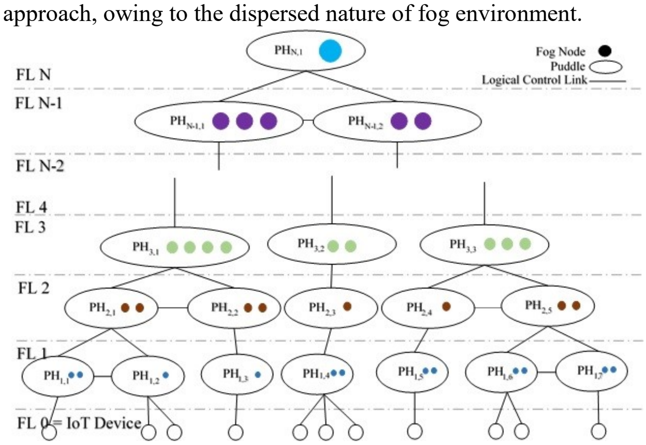

Figure 1: Logical organization of heterogeneous fog nodes into multi-layered hierarchy.

Related Work. Alsaffar et. al. [1] proposed a three-layer architecture of IoT service delegation and resource allocation with device, fog and cloud layers. Taneja and Davy [4] considered a three-layer edge-fog-cloud architecture for IoT environments with fog layer being a flat layer, and assumed that the system is static and knowledge of entire environment is available. Chang, Sriram, and Buyya [2] have proposed Indie-Fog infrastructure for fog deployment leveraging consumers' networking devices. The paper did not discuss organization of nodes to facilitate efficient management of large-scale fog environment. On the contrary, our proposed fog architecture allows inclusion of any node with sufficient compute, storage, and/or network resources to take the role of a fog node, which is logically assigned to a specific layer in fog hierarchy based on pre-defined criteria.

4 HIERARCHICAL AND AUTONOMOUS FOG ARCHITECTURE (HAFA)

Towards distributed management and efficient utilization of fog nodes, we have proposed HAFA [3] to organize them into a logically layered hierarchy. Heterogeneous fog nodes are classified into a set of categories as shown in Fig. 1, and each category of fog nodes forms a fog layer. Nodes belonging to a given fog layer are considered homogeneous w.r.t. computational power, energy resources, network connectivity characteristics, as well as deployment and execution cost. Homogeneous nodes in close vicinity form groups, called Puddles, for purpose of resource pooling and local control. Heterogeneous groups of fog nodes i.e. Puddles belonging to different layers, which are in close vicinity and belonging to immediate upper and lower layers, are logically connected to provide extended control and autonomy. These logical control links are maintained by local management authority of Puddle, referred as PuddleHead. PuddleHead maintains parent/child control links with Puddles belonging to other fog layers and east/west control links with Puddles belonging to same fog layer. East/west control links are used to share workload information with neighboring PuddleHeads, which helps lateral handoff of workload during overflow and failover during disaster scenarios. The set of control links maintained independently by all the Puddles in system together form a tree-like structure as shown in Fig. 1, referred as PuddleTree. Note that PuddleTree itself is not maintained in its entirety by any entity. Instead, it is dispersed in the form of logical control links, which are maintained by individual PuddleHeads. This organization facilitates distributed control and helps in efficient search for fog nodes of required resource and QoS characteristics [3]. Shown in Fig. 1 is a pictorial representation of layering, grouping, and interconnection of a set of fog nodes with varied resource configurations. The size of fog nodes represents the amount of resources available on node i.e. larger the dot, higher the resources. Nodes belonging to a given layer are represented by points of same color. Nodes belonging to a specific Puddle are shown to be enclosed in an ellipse.

Preliminary Evaluation. To demonstrate the features of HAFA, we created a hypothetical data set representing various types of fog nodes deployed in a prospective smart city environment. The dataset includes five types of fog nodes, each represented by a fog layer, and vary in resource configuration and mobility. Shown in Fig. 2 are the fog nodes belonging to each layer represented by a unique color. Higher layer nodes are depicted with larger size vertices to represent their larger resource capacities. Nodes belonging to same fog layer dispersed over the given area are grouped into Puddles using agglomerative complete linkage hierarchical clustering approach. Parent-child relationships among Puddles in adjacent layers are formed using Complete Linkage method. From Fig. 2, we see that parent-child relationships among Puddles from different layers can be leveraged to efficiently find a Puddle with fog nodes of required resource characteristics located closest to user initiating the request. Note that the parent-child relationship information is dispersed at PuddleHeads of individual Puddles and is not available in its entirety with any entity. Thus, our proposed fog architecture facilitates distributed resource management and allocation. This solution assumes that the fog node of required characteristics located nearest to the user has sufficient resources to deploy a given service, which may not be true at all times. Thus, we are developing efficient strategies for service deployment and management in fog with varied resource characteristics and service requirements. We are working on an event-based simulator, which reflects our proposed architecture to show how fog nodes in various IoT environments can be organized in an efficient manner. We plan to use it to test service deployment strategies in fog for varied scenarios.

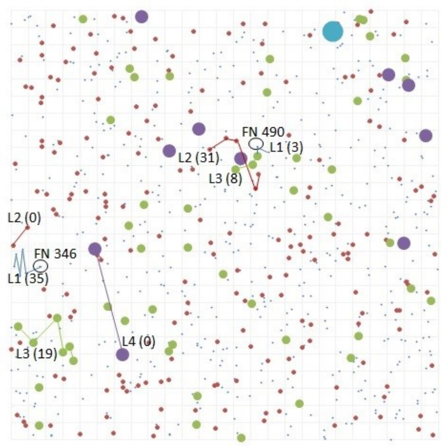

Figure 2: Search for nodes.

REFERENCES

[1] A. A. Alsaffar, H. P. Pham, C. -S. Hong, E. -N. Huh, and M. Aazam. 2016. An Architecture of IoT Service Delegation and Resource Allocation Based on Collaboration between Fog and Cloud Computing. *Mobile Information Systems (Aug. 2016)*.

[2] C. Chang, S. N. Srirama, and R. Buyya. 2017. Indie Fog: An Efficient Fog-Computing Infrastructure for the Internet of Things. *Computer. 50, 9 (Sep. 2017)*, 92-98.

[3] S. Shaik, and S. Baskiyar. 2018. Hierarchical and Autonomous Fog Architecture. *ICPP '18 Comp, August 13–16, 2018, Eugene, OR, USA*. In press. DOI: https://doi.org/10.1145/3229710.3229740.

[4] M. Taneja, and A. Davy. 2017. Resource aware placement of IoT application modules in Fog-Cloud Computing Paradigm. *IFIP/IEEE Symposium on Integrated Network and Service Management (IM, 2017)*. 1222-1228.