Chapter 7
Computational Intelligence for Green Cloud Computing and Digital Waste Management:
Intelligent Computing Resource Management in Cloud/Fog/Edge Distributed Computing

Sana Dahmani
Independent Researcher, Germany

ABSTRACT

Intelligent resource management across fog, edge, and cloud computing entails dynamic allocation, optimal utilization, and robust security measures. Cloud computing adopts a centralized approach to provision resources, whereas edge and fog computing allocate resources at the periphery, strategically minimizing latency. The incorporation of AI/ML algorithms has a pivotal role, enabling resource prediction for anticipating demand, detecting anomalies, and optimizing allocation efficiently. The self-organizing management aspect facilitates autonomous adaptation. Resource virtualization abstracts physical resources into flexible virtual counterparts, complemented by meticulous accounting that tracks consumption and costs. The inclusion of security-aware measures ensures protection against unauthorized access. This comprehensive approach not only enhances performance, scalability, and security but also promotes adaptive scaling and proactive decision-making. Additionally, the implementation of green IT practices optimizes resource utilization, effectively reducing environmental impact.

DOI: 10.4018/979-8-3693-1552-1.ch007

I. INTRODUCTION

As the Internet of Things (IoT) paradigm continues to expand, there has been a notable transformation in the operation of distributed devices towards a common goal. It remains imperative for these devices to function collaboratively to achieve collective objectives. Cloud computing has emerged as the predominant approach for managing extensive computing tasks on a large scale, providing a variety of services over the internet through distributed infrastructure. This infrastructure comprises virtualized computing nodes that can dynamically allocate resources based on negotiated service-level agreements (SLAs) between providers and consumers. The evolution of cloud computing from earlier paradigms such as distributed computing, cluster computing, grid computing; to its current state involves utility computing services that cover essential requirements such as computing power, data storage, memory, licensed software, and modern software development platforms. Cloud computing services are generally classified into three main types: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS) and more services such as data as a services or development as a service may be defined. However, the proliferation of Internet of Things (IoT) devices has generated a demand for reduced response times and real-time data processing, which is a challenge traditional cloud computing struggles to tackle. This has prompted the emergence of fog and edge computing, aiming to complement the centralized cloud model by bringing computing and storage closer to IoT devices (Elhoseny et al, 2018; Choi et al, 2019).

Traditionally, data originating from user devices such as smartphones, wearables, or sensors in smart city environments and factories is transmitted to distant cloud locations for processing and storage. However, this prevailing computing paradigm may prove impractical for the future as the connection of billions of devices to the internet is likely to escalate communication latencies. Such an increase in latencies could have adverse effects on applications, leading to a degradation of overall Quality-of-Service (QoS) and Quality-of-Experience (QoE). This highlights the need for alternative models which address potential latency issues and enhance the performance of internet-connected devices. Proposing a remedial computing model entails the proximate allocation of computing resources to user devices and sensors, utilizing them for data processing, albeit in a partial capacity. This strategy is designed to curtail the volume of data transmitted to the cloud, thereby mitigating communication latencies. The underlying principle involves the decentralization of a subset of computing resources from extensive data centers, disseminating them towards the periphery of the network, closer to end-users and sensors. These resources can manifest as dedicated 'micro' data centers strategically positioned within public or private infrastructure, or as Internet nodes—such as routers, gateways, and switches—enhanced with computing capabilities. Edge computing demonstrates a computing model reliant on resources situated at the network's periphery. Moreover, a model that amalgamates both edge resources and cloud resources is termed Fog computing. The Fog and Edge conceptual frameworks seek to optimize data processing by strategically distributing computational resources across the network hierarchy (Hong & Varghese, 2018 ; Varghese & Buyya, 2018).

Effective resource management is crucial for Fog Computing due to the resource constraints and heterogeneity of fog devices alternatively referred to as fog nodes or foglets, which are tangible entities, responsible for executing computational functions, storing data, and facilitating communication services at the periphery of the network, in proximity to the data origins. These devices serve as intermediaries, connecting the cloud and end-user devices, thereby minimizing the distance between the two. Resource management strategies for Fog Computing include resource allocation, provisioning, scheduling, and task offloading (Rezazadeh et al, 2023; Fahimullah et al, 2022; Tadakamalla & Menascé, 2022).