

A Comprehensive Survey on Network Resource Management in SDN Enabled Data Centre Network



Ashish Sharma, Sanjiv Tokekar, and Sunita Varma

Abstract Network Resource Management in the software-defined data centre is achieved through some intelligent software which provides the automated allocation of the resources on the basis of the instantaneous network parameters. In this paper, we present a comprehensive survey of the most relevant research on network resource management in software-defined networking (SDN)-enabled data centre networks. SDN-based resource management can optimize multiple resources in a data centre network. We begin by outlining the SDN and DCN architectures. The survey conducted by the researchers is then discussed in detail. Though there are few survey articles written on resource management there is no detailed and systematic review conducted on resource management challenges and opportunities. The recent growth of an intelligent data centre is confronted with many challenges like scalability, performance, security and energy efficiency. We surveyed the state-of-the-art resource management techniques with advantages and limitations. Cost reduction, power optimization, increased efficiency, and service stability are just a few of the many benefits of data centre resource management. Resource management methods are not entirely dynamic and do not support an unstable network, which is the biggest drawback. This review will aid the researchers in this field to identify the pros and cons associated with each resource management strategy. It will also help them to identify the best suited technique for load distribution in a specific scenario of SDN. The review of the existing literature is conducted on the basis of various parameters like load distribution capability, network efficiency, resource management, computation complexity, static or dynamic controller distribution, etc. These parameters play an immensely vital role in deciding the overall SDN performance in data centre context.

A. Sharma (✉)

Government Women's Polytechnic College, Indore, India

S. Tokekar

IET DAVV, Indore, India

S. Varma

S.G.S.I.T.S, Indore, India

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

M. Kumar et al. (eds.), *6G Enabled Fog Computing in IoT*,

https://doi.org/10.1007/978-3-031-30101-8_14

333

Keywords Software define network · Data Centre network · Service level agreements · Resource management

1 Introduction

The structure of today's modern civilization is based on several sorts of networks. We are already able to observe the repercussions of these shifts in our day-to-day lives, manifesting themselves in spheres as diverse as the economy, technology, and politics. Networks are utilised in all of these ways, and as a result, they play an increasingly important role in the reduction of costs and the enhancement of production. The administration of our wellness and physical fitness has also become increasingly reliant on networks. The exponential growth in scope of today's networks has resulted in an increase in the number and severity of the challenges they face. Complexity and the expense of operation are two of the most significant issues faced by modern networks. We have pushed the limits of traditional networking by implementing low-power, low-capability nodes to build sensor networks. These networks are seen as a very cost-efficient alternative for data collection across a variety of applications, and we have used them to test the boundaries of traditional networking. Because the system does not have a global state or a global time, these nodes frequently function asynchronously, which further complicates the situation. In order to solve these issues, software-defined networks, often known as SDN, were developed. By noting that "In the SDN design, the control and data planes are separated, network intelligence and state are conceptually centralised, and the underlying network infrastructure is abstracted from the applications," the Open Network Foundation defined SDN. In essence, SDNs give users the power to programmatically alter the underlying network's capabilities while causing the least amount of interruption to the applications that depend on the networking infrastructure.

The main focus of SDN is on the separation of the control and data planes, the centralization of the management perspective, the open interfaces between the various planes and vendors, and the programmability of the network and its applications. SDN adds a number of features. The objective is to further separate applications from routers and switches. Point-to-point (P2P) connections between network nodes were the main focus of SDN's architecture for wired networks. Influenced by the next generation of the digital economy, new business models have flourished, paving the way for the meteoric rise of the Internet of ultra-large-scale data centres, which rely on cutting-edge technologies such as 5G, the Internet of Things, cloud computing, big data, artificial intelligence, and similar technologies to support these new services. The data centre is an infrastructure that consists of servers, storage devices, network devices (switch, router, cables, and firewall), a power distribution system, and many more. These infrastructures are interconnected through network links and switches, composed of SDN.

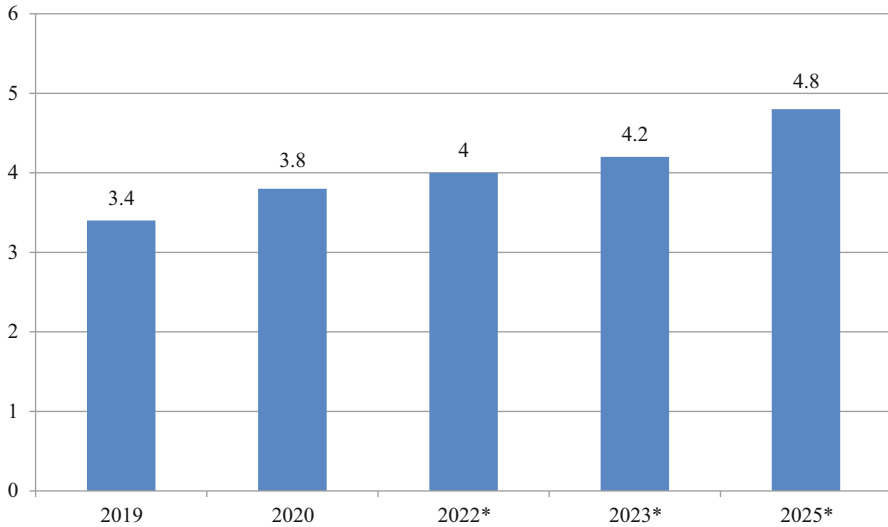


Fig. 1 Value of data centre market investment, India, in USD Billion

In recent years, a data centre has become a vital component of the new internet world [1]. The data centre is an infrastructure that consists of servers, storage devices, network devices (switch, router, cables, and firewall), a power distribution system, and many more. These infrastructures are interconnected through network links and switches, composed of SDN. In recent years, a data centre has become a vital component of the new internet world [1]. The core component of the data centre is Network Infrastructure, Storage Infrastructure, and Computing Resource. The core component of the data centre is Network Infrastructure, Storage Infrastructure, and Computing Resource. According to a report that was just released by NASCOM, the value of the data centre market investment in 2019 was USD 3.4 billion, and it is expected to reach USD 4.8 billion by the year 2025 as can be seen in Fig. 1 below. Between the years 2019 and 2025, it is anticipated that the market for data centre network infrastructure will expand at a compound annual growth rate (CAGR) of 5.5%.

Enterprise DC, Managed Service DC, Colocation DC, and Cloud DC are the four basic forms of DC. The three-layer approach monitors, analyses, and automates all types of DC resource management. We are interested in Cloud DC in this study [2]. Computing as a fifth utility has been made possible by the advent of the “cloud DC,” which allows users to access software and IT infrastructure from anywhere. In the cloud, data centre resource management is still a complex issue that is highly dependent on load on the application. Specific physical devices were used to link the applications. Cloud computing servers in traditional cloud computing environments such as data centres consequently, these servers were

frequently overprovisioned to handle issues related to maximum workload. Wasted resources and energy are the result of this. In terms of resource management, the data centre was prohibitively expensive to run. Dynamically requested services cannot be handled by the Internet's traditional network structure. Fortunately, a solution to this problem has emerged in the form of software defined networking (SDN). SDN has created numerous opportunities for computing and networking researchers in the cloud data centre. The separation of the control and data plans is, as we all know, the most important feature of SDN. The Data plane is in charge of routing and managing packets between the source and destination ports. To operate the packet forwarding mechanism, the control plane is software written in a controller-compatible programming language. The data centre network (DCN), which is made possible by software-defined networking, serves a unique role in a data centre by connecting all network resources. More than that, managing resources is a difficult task because many issues are interconnected, such as resource heterogeneity, asymmetric communication, inconsistent workload and dependency on resources. Additionally, therefore, this study examines the contributions of previous research to SDN-based cloud DCs in terms of network performance and energy efficiency. The major tasks of DCN resource management include improving performance, efficiency, and lowering operational costs [3, 4].

The purposes of our participation in this survey are as follows:

1. We analyze the taxonomy of current trends in resource management strategies, emphasizing their advantages and disadvantages.
2. We discuss the Trend and Opportunities in SDN
3. We identify future research works, which constitute the basis of present and future research recommendations.

This paper's contribution is an analysis of the methods proposed in the literature for managing network resources in SDN-enabled data centre networks. To categorize the resource management options in the SDN-based network including DCN & cloud, Wireless, and WAN, we conduct a literature review of papers published by Springer, IEEE, Elsevier, and ACM between 2015 and 2022. Papers from Springer, IEEE, Elsevier, and ACM, as well as other papers relating to SDN load balancing improvement solutions, are used to supplement this work; their distribution is shown in Fig. 2. SDN, Resource Management, and Data Centre are all terms that are searched for simultaneously. Researchers can use the pie chart in Fig. 2 to pinpoint ACM, Elsevier, IEEE, and Springer journals covering the topic of resource management in SDN.

The remainder of the paper is laid out as follows. In Sect. 2, we discuss the history of SDN and DCN. The researcher's relevant work is presented in Sect. 3. SDN-DCN resource management challenges are covered in Sect. 4. Section 5 depict SDN-prospective DCN's trends and opportunities. The paper came to a close with Sect. 6 which is Conclusion and Future Work.

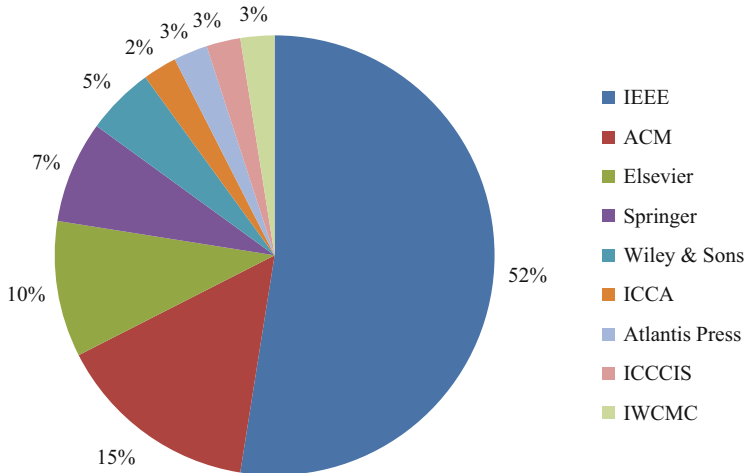


Fig. 2 Percentage of reviewed paper in publication

2 Background

2.1 A Subsection Sample

The separation of the data (also known as forwarding) and control planes is suggested by the potential network design known as software-defined networking (SDN). To coordinate the network, it uses a logically centralised controller supported by a network-wide global view [106]. By switching from network administration to network programming, it fosters creativity. SDN, in theory, aims to solve the legacy networks’ manageability issues by converting static networks into dynamically-programmed ones. In computing and networking, the SDN architecture defines how a system can be designed to use a combination of open, software-based technologies and commodity networking hardware. Open Interface is a fundamental property of SDN. Which is use to connect the network resources and the network traffic. This link is controlled by software that is being developed in response to changing needs. SDN architecture separates the control and data planes of the network stack. SDN architecture comprised a three-layer Application layer, Control layer, and Infrastructure layer, as shown in Fig. 3 [5].

The Infrastructure Layer (Data plane) The data plane, also called the forwarding plane, is present in the hardware of switches and is accessible via the software. Although administrators have some control over the data plane, that control is limited. Because SDN uncouples the data plane and control plane, OpenFlow is needed as a communication route for coordinating these two layers [3]. These planes build routing tables on the routers and switches to identify which packets should travel from A to B and which should go to C. Despite their uniqueness, conven-

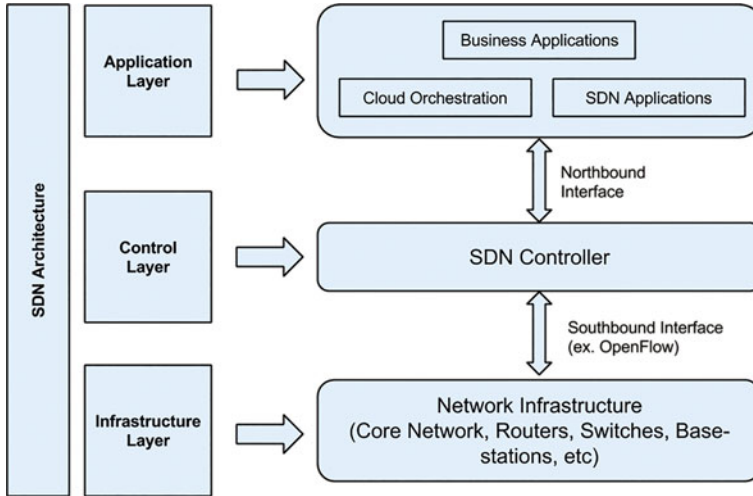


Fig. 3 The traditional SDN architecture [43]

tional routers and switches nonetheless follow IEEE standards and incorporate all necessary planes into their firmware. It comprises multiple interconnected network elements like simple switches, routers, and base stations. The Infrastructure layer connects with SDN southbound interface (or southbound APIs) to the Control layer (Control Plan). However, these APIs are tightly bound to the forwarding elements of the physical infrastructure.

The Control Layer (Control plane) The Control plane transcends the boundaries of a conventional software layer in its entirety. On this level, we have a face-to-face encounter with a concrete member of the hierarchy: the controller. The additional information presented here illustrates how the axes are separated from one another. This plane symbolises the conceptual centre of the argument. One decision point made up the entirety of the control plane's simplest iteration. As a consequence of this, the method makes the process of designing, programming, and resolving logical contradictions simpler. This architecture, on the other hand, is extremely vulnerable to a failure at a single point that might have devastating consequences. Problems with scalability arise when the processing capability of the controller needs to scale up proportionally more as the network increases.

Researchers came up with the idea of physically and conceptually decentralising the SDN's Control plane. To do this, they proposed the concepts of a control hierarchy composed of main controllers and subsidiary controllers as a way to decentralise control. This was done in an effort to mitigate the negative effects of centralised controllers. As a consequence of this, the network was able to divide the load among a number of different physical controllers, hence lowering the probability of experiencing bottlenecks. Comprises of a centralized SDN controller where the network administration takes place. It is responsible for the network

status monitoring, link discovery, policy generation & development with forwarding table management. The Control layer interacts with the application layer through Northbound Interfaces (or northbound APIs).

The Application Layer (application plane) could include network applications and services such as load balancing, routing, security, mobility & wireless, etc., are implemented. The main strength of SDN architecture is that it provides an abstracted holistic view of the entire network. Attributes, services, and rules are all specified at the application layer, which is also the name of this layer. Applications have to have a solid understanding of the infrastructure of the network in order for them to respond effectively. These apps are able to develop capabilities that span end to end and can respond to changes in the underlying network. Applications are able to dynamically adapt the behaviour of the network in order to meet shifts in the network's topology, feature needs, or policy needs. The application programming interfaces (APIs) are what make the previously described layers dependent on one another (APIs). We make the network more intelligent by integrating real-time networking information with the application.

Northbound APIs The programming interface between SDN controllers and the network applications built on top of them is called a Northbound Interface (NBI) API. Network applications can communicate with controllers through a northbound interface and call the services that the controllers make available and that the applications need for proper functioning, such as the global network view. We believe there are a few northbound APIs, but none of them is the industry standard. An ONF project has been launched with the goal of forming a working group for standardising NBIs.

Southbound APIs The interface that enables SDN controllers to talk to switches in order to manage and observe their functioning is known as a Southbound Interface (SBI) API. SBIs, as opposed to NBIs, are standardised to allow the management of network devices from several vendors. OpenFlow [42], which is maintained by the ONF, is the de facto SBI standard at the moment. Other SBIs exist, such ForCES, which was suggested by the IETF. Either in-band or out-of-band connections can be used with SBIs. In-band control includes using the same network (physical connections) to send the controller-switch communication as data traffic. A separate network is used for control and data traffic in out-of-band control. Most SDN deployments choose out-of-band control due to reliability concerns, however in-band control may be preferable due to its financial advantages.

2.2 *Datacentre Architecture*

In order to support the next-generation Computing-as-a-Service (CaaS) and Cloud computing infrastructures, data centres have advanced from using mainframe systems and enterprise networks to sophisticated networks of 100,000 or more

servers. The proliferation of data centres has resulted in a proportional rise in the amount of energy consumed by data centres as well as an increase in the amount of network traffic experienced by data centres. As a result, there is a considerable push for novel study of data centres, with the ultimate goal of developing effective data centre network (DCN) architectures and ways for substantially reducing energy consumption. A data centre is a collection of computing, storage, and networking resources that are connected via a communication network. In a data centre, the Data Centre Network (DCN) is crucial since it connects all of the data centre resources. To satisfy the expanding demands of Cloud computing, DCNs must be scalable and efficient enough to connect tens of thousands, if not millions, of servers. The data centre network design is based on a tried-and-true layered approach that has been tested and refined over the course of several years in some of the world's largest data centre deployments. The layered approach is the fundamental building block of data centre design, and it aims to improve scalability, performance, flexibility, resiliency, and maintenance while also reducing costs. The Data Centre is typically a three-tiered hierarchical internetworking concept. As illustrated in Fig. 2, the model is made up of three layers: access layer, aggregation layer, and core layer [6].

Core Layer Data centres can easily connect to the internet via core layer switches. In and outbound data packets are routed through the data centre's core switch. The Layer 3 routing flexibility of the core switch is complemented by its lightning-fast throughput. This capability enables high-speed packet switching on the data centre's backplane for both incoming and outgoing traffic. The primary function of this layer is to supply a fault-tolerant, fully redundant Layer 3 routed fabric. The core layer also communicates with several aggregation subsystems. Traffic between the campus core and the aggregation layer can be load-balanced with the help of Cisco or other Express Forwarding-based hashing techniques while an interior routing protocol such as OSPF or EIGRP is being run on the core layer.

Aggregation Layer Core layer switches connect the aggregation layer switches to each other. Service modules, Layer 2 domain definitions and spanning tree processing are just some of the important functions that these modules provide. An active-passive high availability mode is used for the aggregation layer. Layer 2 domain definitions, spanning tree processing, and default gateway redundancy are some of the key features that it contributes. Optimizing and securing programmes for multi-tier traffic across servers can be accomplished through the utilisation of services such as server load balancing and firewalls. In Fig. 4, the integrated service modules are represented by the smaller symbols that may be found inside the aggregation layer switch. These modules have the capability of providing a wide variety of services, including but not limited to content switching, firewall defence, SSL offloading, intrusion detection, network analysis, and more. At the aggregation switch, a network's Layer 2 and Layer 3 segments are clearly separated. This layer is also called the Distribution layer.

Access Layer Data centre servers are physically connected to the access switch. All of the servers in a network are connected at the access layer. Various types of

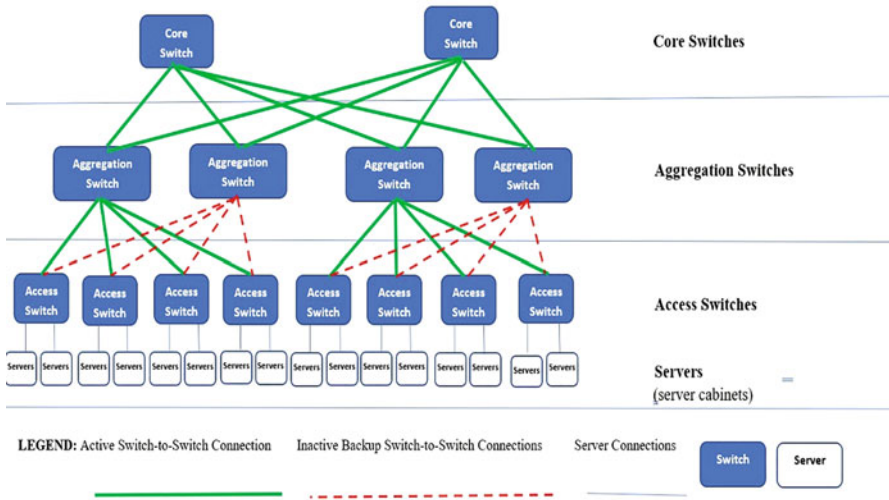


Fig. 4 The traditional data centre architecture [42]

servers, such as standalone machines, blade servers with built-in switches, pass-through blade servers, clustered standalone machines, and mainframes with Open Systems Architecture (OSA) adapters, are included. The access layer of a network might be composed of embedded blade server switches, modular switches with a fixed configuration of 1–2 rack units (RU), or a combination of the three. Switches offer link layer and network layer topologies, making it possible for them to accommodate a wide range of managerial and broadcast domain requirements from servers. Access switches are usually located at the top of the rack. Multiple access layer switches are interconnected with the aggregation switch.

2.3 Data Centre Design Models

Multi-Tier Model A cluster-based multi-tier data-centre has key differences from high-performance computing systems. Data-centres must run a variety of server software, each with its own hardware and software requirements. High-performance computing systems use multiple copies of a programme to complete a task in parallel. Proxy, application/web, and database layers make up a multi-tier data centre’s standard trinity. Due to the different needs and behaviours of each tier, it is difficult to analyse architectural components such as file system, I/O, network protocol, etc. and their impact on a multi-tier data centre. As illustrated in Fig. 5 multi-tier architecture abstracts various functionalities, which is useful as dynamic web content grows. I/O interconnect technologies are crucial for inter- and intra-cluster communication in multi-tier data centres. Distributed web servers improve

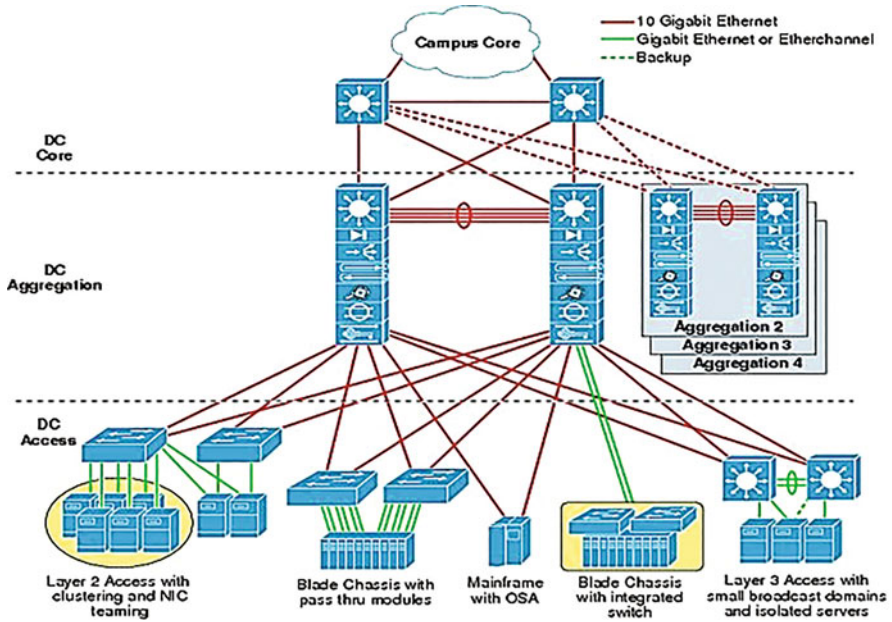


Fig. 5 Data centre multi-tier model topology

throughput and response time. Enterprise resource planning (ERP) and customer relationship management (CRM) systems are supported by a foundation of layered web, application, and database design. This style of design is compatible with a wide variety of web service architectures, such as those based on Microsoft.NET or Java 2 Enterprise Edition. ERP and CRM software like Siebel and Oracle use web service application environments. It is crucial to the multi-tier architecture that services for networked security and application optimization are made available.

Server Cluster Model High availability, load balancing, and more computational power are just a few of the reasons why clusters of servers are used in today’s data centre environment. Clusters are large deployment units made up of tens or hundreds of individual server cabinets connected via large, high-radix cluster switches and outfitted with top-of-rack (TOR) switches. Each cluster is designed to make multiple central processing units (CPUs) function as if they were part of a single, unified; high-performance system by means of specialized software and high-speed network interconnects. Historically, server clusters have been utilized in the context of research at academic institutions, scientific laboratories, and the military for specialized applications. The widespread adoption of server clusters in businesses is a direct result of the widespread application of clustering technology to a growing variety of applications. Recent years have seen the concept of a server cluster spread from academic institutions to large-scale industries like banking, manufacturing, and the media. The concept of a server cluster is useful not only

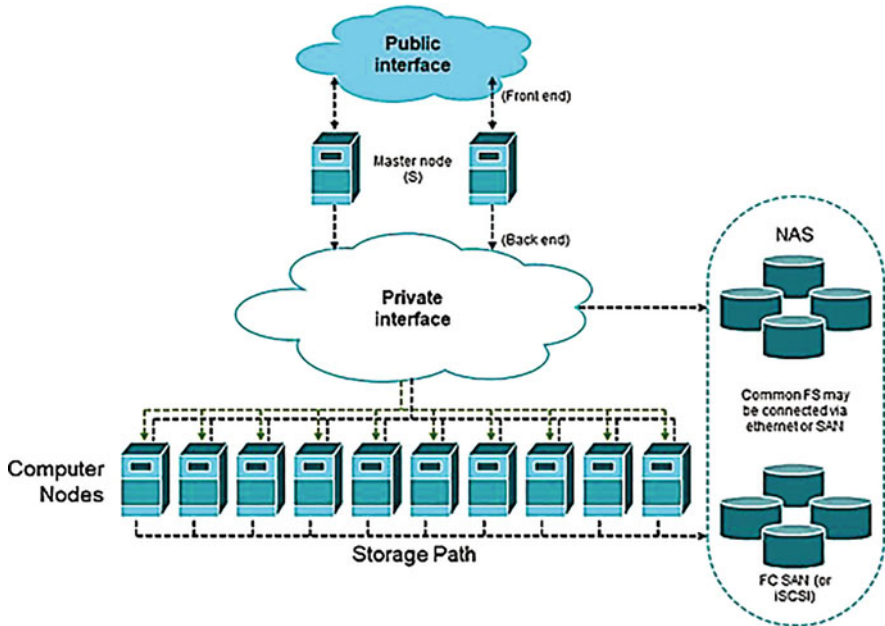


Fig. 6 Server cluster model of data centre layout

for grid and utility computing but also for HPC, parallel computing, and high-throughput computing. These designs as shown in the Fig. 6 are typically based on specialized application architectures created to meet a wide variety of niche market needs in the business world.

3 Related Work

This section presents the variety of important and significant solutions offered by the researchers over the last decade. The objective of discussing the existing works in the field of resource management and load distribution in SDN framework is to present the spectrum of solutions available to attain the most optimal solution. Open research challenges in modern cloud DC systems include resource management and load balancing, as well as issues of security and privacy. This section presents a detailed analysis of some of the relevant previous works related to this resource management of SDN. The potential of Network Resource Management in Data Centre Network with SDN has been recognized and explored by many researched over the last decade.

In a dense network, Yang et al. [7] provided a solution for the video streaming problem and suggested that in order to overcome the problem, video layer selection

and resource allocation must be done correctly. They have applied Lyapunov optimization approach to decompose the problem of resource allocation into two level problem viz. wireless resource allocation and the respective video layer selection. An effective routing scheme for the video streams was also presented by the researchers to utilize the small cell base stations interaction.

Ndikumana et al. [8] proposed the multi-access edge computing concept for reducing the network delay and alleviate the load on the data centre. They proposed an idea of a 4c framework for mobile edge computing. A distributed optimization control algorithm is developed by researchers, which increases bandwidth saving and minimized delay. By collaborative communication computation, the MEC server's caching and control model resource allocation can be done.

Karmoshi et al. [9] developed an application-aware resource method named as virtual physical switch software-defined network (VPS-SDN). This model proposed a novel structure for network virtualization for the multitenant cloud data centre. This scheme achieves high network utilization.

Yahya et al. [10] have suggested that large-size topologies consumed more resources than the small-size topologies in their research work. The number of sources depends upon the type of switches and controller. By using Open v Switch, CPU utilization is more as compared to another switch. As per their study, the CPU utilization of OVS controller was found to be inferior as compared with another controller.

To improve end-to-end system performance and meet different applications' requirements, Chen et al. [11] developed a joint caching computer strategy. They have also proposed a solution to the server selection problem. Energy cost and network usage cost are also minimizing through resource allocation problem solutions.

Chen et al. [12] studied different resource allocation issues by joining networking caching and computing. By balancing the server's load and reducing the network usage, they develop a framework for improving the performance. A discrete stochastic approximation (DSA) algorithm is developed.

Tso et al. [13] surveyed various resource management strategies for the network and server. In their research, they emphasized over the necessity and limitations of adaptive resource provisioning. They have also analysed the challenges and opportunities for adaptive and measurement-based resource allocation. As per their findings, more matrices are required to developing robust resource measurements and better trade-off between network-wide stability and adaptively.

Braiki et al. [14] categorize their survey for resource management issues based on use method, approach and objective model. For better performance, multi-objective models are better in terms of energy consumption and resource utilization. They used a cloudsim simulator. In their proposed method, average energy consumption improves by 12% and average resource utilization by 15%.

Cao et al. [35] reviewed modern data centres, resource management issues. The first issue is figuring out how to combine diverse resources (hardware and software) into a uniform platform (virtual resource). The second is the problem in figuring out how to manage the numerous resources effectively in the organisation. The third

issue is that of resource services in particular, network services, selecting a suitable method of resource management among several resources' platforms. As a result, it is challenging and complex the following conditions should be considered: resource accessibility management, a temporary storage pool, and the ability to be flexible executing network architectures (such as resource allocation).

4 Challenges Associated with the Resource Management in SDN-DCN

Technological changes in the data centre have not delivered their full potential in boosting productivity and economic growth. There are several challenges to managing these technological changes in SDN-DCN. This section discusses some of the difficulties related to SDN-DCN resource management. The comprehensive objective of a DCN is to ensure that it endlessly meets the SLA of the application it hosts, with minimize underutilized and avoiding over-utilized total resources. In DC, resources are diversified and distributed. To manage these resources is a complex system. The major challenge is building an intelligent autonomous resource management system that can self-alleviate and self-heal from any system failures [15].

When in DCN any requests come for one or more resources, the data centre network provides master controller, it schedules the available resources by making them public. Scalability, performance, security, reliability, and energy efficiency are just a few of the resource management concerns that must be addressed [16]

1. **Scalability:** Every day, data centre operators face the challenge of scaling their operations and acquiring enough hardware to meet the demands of complex IT systems (such as increasing compute, storage and network needs). The Internet of Things (IoT), social media platforms, on-demand video, and the global digital revolution are all contributing to a rise in computer density, which in turn is increasing the need for scalable data centres. There has to be systems and strategies in place to manage this expansion that are cost-effective and able to scale. It is one of the most challenging aspects of SDN-DCN. Because of the SDN controller, DC now has the ability to scale up to hundreds of thousands of network resources. It is preferable to create a controller that can handle big dynamic flow tables when dealing with a high number of resources. Designing an intelligent, scalable controller for controlling resources in SDN-enabled data centres requires more research [17]. To meet the tremendous growth in compute power and storage, data centre operators must expand crucial infrastructure (power and cooling) as well as physical space. The reliability of the current operating system must not be compromised in order for MTDC operators to offer electricity and cooling infrastructure promptly, efficiently, and at the lowest cost.
2. **Performance:** The performance guarantee is an essential issue between the data centre service providers and users. User satisfaction and DC cost are

directly linked to performance management. To improve the performance of the DC resource management, the researcher has to study more key performance indicators (KPI) [18]. Data centres typically monitor CPU utilisation as a proxy for performance. However, if application demand remains stable, can we cut the number of machines in the data centre in half? No, probably not at all. This assumes that the average CPU utilisation per machine in the data centre is 50%. Because of the wide variety of workloads and hardware in use in data centres, it can be challenging to conduct a comprehensive performance analysis of the infrastructure's effectiveness. Emulation, automation, and analytics must be utilised by providers at every stage of the delivery process, beginning with design and construction and continuing through deployment, operation, and optimization. Only then can providers guarantee that customer expectations will be satisfied at each stage.

3. **Security:** In SDN-enabled data centre security, associated problems are a significant concern. Effective security must be created and placed at the application plane in SDN for data centre network protection. The controller is the primary target of threats in the SDN-DCN. The attacker targets the controller for serious damage to the data centre network. For the unauthorized access, several alleviate plans of action would be developed. Which reduces the threat and provides high-level security to protect the data centre network. Despite the fact that data centres may be quite complicated, they are only required to adhere to a single comprehensive security policy; yet, each individual component of that policy needs to be carefully evaluated. The term "software security" refers to one category of protection, while "physical security" refers to another. The phrase "physical security" is used to refer to a wide range of different strategies and approaches that are utilised to prevent unwelcome access by individuals from the outside. By using software or virtual security measures, the network is secured against potential invaders who would otherwise be able to circumvent firewalls, crack passwords, or gain access through other security weaknesses. Intruders may be able to do any of these things [19].
4. **Reliability:** It is essential to ensure high reliability in an SDN-enabled data centre network. It is an essential factor in communication within the data centre network. It is very challenging to develop the intelligent and validated SDN controller that manages the network to increase network availability. When any controller fails or is over-utilized, the system must drive through the available alternative controller to maintain a reasonable level of reliability. Reliability-aware capabilities of the controller can improve the reliability of the controller plane. To meet the service level agreement (SLA), SDN enabled DC is more reliable. It is absolutely necessary to have a reliable network in the data centre in order to construct and operate online services that are highly available and scalable. Even while there is a lot of monitoring going on at the device and link levels, the full implications of how stable the network infrastructure is for the software systems that rely on it are still a mystery. The fundamental difficulty lies in identifying the connection that exists between the effect of the software system and problems at the device and link levels. To begin, the redundancy that is built

into the infrastructure of most networks ensures that the majority of network outages do not cause problems with the software systems (including redundant devices, routes, and protocols). Second, automated repair systems are frequently employed in the infrastructure of large-scale networks so that problems can be fixed as soon as they are identified. This makes it possible to resolve issues in a timely manner [20].

5. **Energy Efficiency:** In the recent growth of the Information and Communication Technology (ICT) sector, energy efficiency poses a serious challenge. Data centres in the United States and elsewhere use between 1% and 3% of the world’s total electricity, according to various reports. With the proliferation of IoT devices and AI, these figures might presumably rise dramatically. These days, data centres are an integral part of any discussion on computing or networking. Data centres house computers and networks that collect, organise, and make available vast amounts of information. About \$20 billion is spent annually on their development, and they generate nearly as much carbon dioxide as the airline sector does. The energy usage of network resources within a data centre is predicted to climb to roughly 50% in the next few years in DC. The major challenge is increasing the performance DC uses more redundant resources, which reduces the energy efficiency level. Another challenge is minimizing the underutilization and avoiding over utilized resources for better energy efficiency without disrupting the SDN-enabled DC operation [21].

In this section, a side-by-side comparative analysis of the selected technique is performed in terms of the advantages and limitations as shown in Table 1.

Table 1 Resource Management (RM) Methods and their comparative analysis

Reference	Technique for RM	Advantages	Limitation
[22]	Reinforcement learning	1. Proposed method decreases the SLA (service level agreements) violation time 2. To serve more VM schedule request 3. Migration between resource are decreased	1. More power consumption 2. More dynamic testing environment
[23]	Scalable routing and resource management model (SRRM)	1. Load balancing 2. High degree of scalability	1. Dynamic load balancing 2. Energy efficiency
[24]	Dynamic resource management algorithm (DRMA) Bin packing algorithm	1. Minimize underutilized resources 2. Avoid over-utilized resources	1. Negative impact on the performance 2. Additional energy required
[25]	Resource allocation scheme	1. Less packet loss rate 2. Improvement in link utilization.	1. Congestion management not implemented 2. Only theoretical approach

(continued)

Table 1 (continued)

Reference	Technique for RM	Advantages	Limitation
[26]	Green resource management techniques	1. Decreased the energy consumption 2. Decreased the carbon dioxide emission 3. Decreased dynamic cost	1. More components can handle like green based cloud balancing 2. Green-based match making
[27]	Hybrid virtual machine consolidation approach Beam search algorithm	1. Minimization of energy consumption and SLA violation	1. Workload predicting 2. Reduce the VM migrations
[28]	Traffic analysis-based energy-saving approach	1. Energy efficient 2. Reduce round trip delay 3. More bandwidth utilization	1. Traffic loss is possible 2. Only static resource management
[29]	VPS-SDN (virtual physical switch software defined network)	1. High network utilization 2. Scalable 3. More QoS	1. Only prototype model 2. For less network size
[30]	Overhead reduction scheme	1. Less network latency 2. More throughput 3. More network performance	1. Only for small DCN topology 2. More network scenarios
[32]	Network topology framework	1. Reducing complexities 2. Improving resource administration	1. More VM overheads 2. VM placement issues 3. VM allocation related problems
[33]	CPU-based and CPU memory-based techniques	1. Achieve higher resource utilization and more reliability 2. Shorter response time 3. Higher throughput and transaction rate	1. Only for small network 2. Only theoretical approach
[34]	Chopin framework (an intent-driven framework)	1. Better resource efficiency 2. Better resource allocation	1. Not support for unstable networks
[36]	Intelligent network resource allocation scheme	1. Avoiding congestion and idleness. 2. Automatically assignment of network resource	1. Impact to delay and jitter are not considered 2. Dynamic queue mapping is not done

5 Promising Trend and Opportunities in SDN-DCN

Data Centre Network with SDN-enabled has a new paradigm. Dynamic programming of the controller and integration of the data centre controller with the SDN controller enables optimised network resource management. It also improves network scalability and manageability. Monitoring or predicting the traffic by the controller is also one more advantage of the SDN-enabled data centre network.

The distinguishing characteristics of SDN create more opportunities for network resource management in DCN. The simultaneous use of server resources and network resources can bring more innovations by integration and clustering of DC controller with SDN controller. SDN intelligent controller and DCN controller integration bring DC network resource management innovation [31].

Year-over-year changes in SDN that take DC networking to the next level. There are a few trends in SDN-DCN.

1. **Edge Computing** – In 2022 (and beyond), edge computing will be gaining high importance in the DC network. The demand for edge computing will grow because people have adopted more intelligent technologies in their homes and businesses. In parallel IoT market is also overgrowing. With this growing demand for reliability, speed and connectivity will be managed by edge computing. Edge computing's future is rapidly arriving. There is endless potential to change the world as processors get stronger, storage gets more affordable, and networks access gets better. Edge computing will advance in the future along with cutting-edge networks like 5G, satellite mesh, and artificial intelligence. You've now unlocked some very far-reaching opportunities by having more capacity and power, better access to swift and wider networks (5G, satellite), and smarter computer-based machines (AI) [37].
2. **Green Computing** – On coming year, sustainability issues like water usage, energy emissions, and consumption is growing concerns in the DC network. Renewable resources and management are a big focusing area at the DC network. Look for the DC to find ways to reduce their impact on the environment and help other sectors. The term "green computing," sometimes known as "green IT" (Information Technology) or "green technology," refers to the environmentally responsible use of computing systems and the resources they necessitate. Environmentally friendly computing refers to efforts to develop, deploy, and retire computer systems and components with little impact on the natural world. Some examples of these practises include designing energy-efficient computing equipment, implementing energy-efficient processing units and servers, reducing the use of harmful chemicals, advocating for the recyclable nature of digital products, and disposing of electronic waste in an environmentally responsible manner (e-waste). Green computing strives perpetually to make computing less harmful to the environment. Green Computing was originally called Energy Star when it was introduced in 1992 [38].
3. **Automation** – Due to the coronavirus pandemic, more automation is another aspect of network resource management in SDN-enabled DC networks. More DC shifts to remote monitoring capabilities and routine services like updating and patching to limit contact with other people. Skilled staffing issues are still a concern in DC. The potential for edge computing to automate many existing DC processes is substantial. Artificial intelligence (AI) and robots have progressed so quickly that they have pushed the limits of automation. These days, robots can do a lot of work with hardly any help from humans. Technological automation is not only eliminating repetitive tasks, but also vastly improving workers'

talents. More than half of human labour might be replaced by automated robots, according to some estimates. Automation is utilised in numerous industries, including manufacturing and banking, to improve efficiency, security, profits, and product quality. Automation will enhance reliability and connectivity in a highly competitive market. It seems that in the future, everything will be easily accessible thanks to automation [39].

4. **5G Technology** – DC resource management will require considerable changes to accommodate 5G technology. Primary DC network resources and requirements like QoE, Network inter-operability, performance is resolved by SDN-enabled 5G network. It is mandatory for a data centre to host and stream data at significantly higher speeds, volumes, and lower latencies. Due to better efficiency and bandwidth, 5G network enhances resource management efficiency and can serve faster in the DC network. The next leap forward in wireless technology will be made possible by fifth-generation wireless technologies (5G). There will be more storage space, higher transfer rates, and less lag time with this new technology. 5G has enormous potential for facilitating developments across many sectors, including those related to public safety, transportation, and healthcare. In the emerging IoT ecosystem, where gadgets are increasingly connected to one another, it will also have an impact on sustainability [40].
5. **Hyperscale** – Big data and cloud computing environments are usually used in Hyperscale computing. DC conventional computing architecture structural design is often different from hyper-scale computing architecture. Hyperscale data centre, market size, networking equipment value grows at around 30% from 2018 to 2020. Open architecture, edge computing, and security potentially affect respondents in hyper-scale DC. Massive in size, hyper-scale data centres house thousands of servers, racks of networking hardware, tonnes of cooling and power infrastructure, and more. Demand for data centres has increased as Covid-19 has spread around the world, with strong purchases coming from hyperscale businesses and cloud platforms while spending from many enterprise users has slowed [41].

6 Conclusion and Future Work

Over the last few years, researchers have become increasingly interested in data centre networks that use software-defined networking (SDN). The challenges and opportunities associated with the various techniques for resource management in an SDN-enabled data centre network are thoroughly covered in this paper. Also included are a number of different resource management approaches, as well as their advantages and disadvantages. The bulk of procedures are restricted in their ability to be automated and have low energy efficiency. This paper presents a wider spectrum of load distribution and resource allocation techniques proposed by various researchers over the last decade. The advantages and limitations of these techniques can help the readers to identify the best possible technique for the

scenario under consideration in their research. We then talked about the most recent software-defined network data centre networks trends. A major focus of future research will be on assessing the failure of resources in conjunction with intelligent controller resource management in SDN-DCN. However, there are various concerns and obstacles that must be addressed to enhance the efficiency, reliability, and cost optimization of the data centre network overall.

References

1. Kreutz, D., Ramos, F., Veríssimo, P. E., Esteve, C., Azodolmolky, S., & Uhlig, S. (2015). Software-defined networking: A comprehensive survey. *Proceedings of the IEEE*, 103(1), 14–76.
2. https://www.cisco.com/c/en_in/solutions/data-centre-virtualization/what-is-a-data-centre.html
3. Sun, X., Ansari, N., & Wang, R. (2016). Optimizing resource utilization of a data centre. *IEEE Communications Surveys & Tutorials*, 18(4), 2822–2846.
4. Ma, J. (2016). *Resource management frameworks for virtual data centre embedding based on software define networking, computers and electrical engineering*. Elsevier Ltd.
5. Xu, Y., Sun, Z., & Sun, Z. (2017). SDN-based architecture for big data network. In *2017 IEEE international conference on cyber-enabled distributed computing and knowledge* (pp. 513–516). IEEE.
6. https://www.cisco.com/c/en/us/td/docs/solutions/Enterprise/Data_Centre/DC_Infra2_5/DCInfra_1.html
7. Yang, J., Yang, B., Chen, S., Zhang, Y., Zhang, Y., & Hanzo, L. (2019). Dynamic resource allocation for streaming scalable videos in SDN aided dense small-cell networks. *IEEE Transactions on Communications*, 67(3), 2114–2129.
8. Ndikumana, A., Tran, N. H., Ho, T. M., Han, Z., Saad, W., Niyato, D., & Hong, C. S. (2020). Joint communication, computation, caching, and control in big data multi-access edge computing. *IEEE Transactions on Mobile Computing*, 19(6), 1359–1374.
9. Karmoshi, S., Shuo, W., Saleh, F., Li, J., & Zhu, M. (2019). VPS-SDN: Cloud datacentres network resource allocation. In *HP3C'19: Proceedings of the 3rd international conference on high performance compilation, computing and communications* (pp. 100–105). Association for Computing Machinery.
10. Yahya, E. B., & Al-Somaidai, M. B. (2019). Network parameters effects on system resources in software defined networks. In *ICICT'19: Proceedings of the international conference on information and communication technology* (pp. 89–95). Association for Computing Machinery.
11. Chen, Q., Yu, F. R., Huang, T., Xie, R., Liu, J., & Liu, Y. (2016). Joint resource allocation for software defined networking, caching and computing. In *2016 IEEE global communications conference (GLOBECOM)* (pp. 1–6). IEEE.
12. Chen, Q., Xie, R., Huang, T., Liu, J., & Liu, Y. (2017). Software defined networking, caching and computing resource allocation with imperfect NSI. In *GLOBECOM 2017–2017 IEEE global communications conference* (pp. 1–6). IEEE.
13. Tso, F. P., Jouet, S., & Pezaros, D. P. (2016). Network and server resource management strategies for data centre infrastructures: A survey. *Computer Networks*, 106, 209–225.
14. Braiki, K., & Youssef, H. (2019). Resource management in cloud data centres: A survey. In *15th International wireless communications & mobile computing conference (IWCMC)* (pp. 1007–1012). IEEE.

15. Saeed, N. S. B., & Alenazi, M. J. F. (2020). Utilizing SDN to deliver maximum TCP flow for data centres. In *ICISS'20: Proceedings of the 3rd international conference on information science and systems* (pp. 181–187). Association for Computing Machinery.
16. Xu, G., Yang, J., & Dai, B. (2015). Challenges and opportunities on network resource management in DCN with SDN. In *2015 IEEE international conference on Big Data (Big Data)* (pp. 1785–1790). IEEE.
17. Hossein, A., Watts, M., & Ahmadi, K. (2019). An overview of multi-controller architecture in software-defined networking. In *10th Annual conference of computing and information technology research and education New Zealand (CITREZZ2019) and the 32nd annual conference of the national Advisory committee on computing qualifications* (pp. 9–11). Hamilton, Nelson.
18. Shirmarz, A., & Ghaffari, A. (2020). Performance issues and solutions in SDN-based data centre: A survey. *The Journal of Supercomputing*, 76, 7545–7593.
19. Leng, X., Hou, K., Chen, Y., Bu, K., & Song, L. (2018). SDNKeeper: Lightweight resource protection and management system for SDN-based cloud. In *2018 IEEE/ACM 26th international symposium on quality of service (IWQoS)* (pp. 1–10). IEEE.
20. Abdullah, M. N., Dawood, A. S., & Taqi, A. K. (2017). Network resource management optimization for SDN based on statistical approach. *International Journal of Computer Applications*, 177(6), 5–13.
21. Bahari, H. I., & Shariff, S. S. M. (2016). Review on data centre issues and challenges: Towards the green data Centre. In *2016 6th IEEE international conference on control system, computing and engineering* (pp. 129–134). IEEE.
22. Telenyk, S., Zharikov, E., & Rolik, O. (2018). Modeling of the data centre resource management using reinforcement learning. In *2018 International scientific-practical conference problems of infocommunications. science and technology (PIC S&T)* (pp. 289–296). IEEE.
23. Celenlioglu, M. R., Tuysuz, M. F., & Mantar, H. A. (2018). An SDN-based scalable routing and resource management model for service provider networks. *International Journal of Communication System*, 31(8), e3530.
24. Ngenzi, A., Selvarani, R., & Nair, S. R. (2015). Dynamic resource management in cloud data centres for server consolidation, *arXiv:1505.00577*.
25. Thazin, N., Nwe, K. M., & Ishibashi, Y. (2019). Resource allocation scheme for SDN-Based cloud data centre network. In *Proceedings of the 17th international conference on computer applications, (ICCA)* (pp. 15–22).
26. Surendran, R., & Tamilvizhi, T. (2018). How to improve the resource utilization in cloud data Centre? In *2018 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)* (pp. 1–6). IEEE.
27. Zharikov, E., Telenyk, S., Rolik, O., & Serdiuk, Y. (2019). Cloud resource management with a hybrid virtual machine consolidation approach. In *2019 IEEE international conference on advanced trends in information theory (ATIT)* (pp. 289–294). IEEE.
28. Paliwal, M., & Shrimankar, D. (2019). Effective resource management in SDN enabled data centre network based on traffic demand. In *IEEE Access* (Vol. 7, pp. 69698–69706). IEEE.
29. Pranata, A. A., Jun, T. S., & Kim, D. S. (2019). Overhead reduction scheme for SDN-based data Centre networks. *Computer Standards & Interfaces*, 63, 1–15.
30. Baig, S. R., Iqbal, W., Berral, J. L., Erradi, A., & Carrera, D. (2019). Adaptive prediction models for data centre resources utilization estimation. *IEEE Transactions on Network and Service Management*, 16(4), 1681–1693.
31. Abbasi, A. A., Abbasi, A., Shamshirband, S., Chronopoulos, A. T., Persico, V., & Pescape, A. (2019). Software-defined cloud computing: A systematic review on latest trends and developments. In *IEEE Access* (Vol. 7, pp. 93294–93314). IEEE.
32. Abbasi, A. A., Shamshirband, S., Al-qaness, M. A. A., Abbasi, A., AL-Jallad, N. T., & Mosavi, A. Resource-aware network topology management framework, networking and internet architecture (cs.NI); Machine Learning (cs.LG) *arXiv:2003.00860*

33. Hamed, M. I., ElHalawany, B. M., Fouda, M. M., & Eldien, A. S. T. (2017). A novel approach for resource utilization and management in SDN. In *2017 13th International computer engineering conference (ICENCO)* (pp. 337–342). IEEE.
34. Heorhiadi, V., Chandrasekaran, S., Reiter, M. K., & Carnegie, V. S. (2018). Intent-driven composition of resource-management SDN applications. In *CoNEXT'18: Proceedings of the 14th international conference on emerging networking experiments and technologies* (pp. 86–97).
35. Cao, R., Tang, Z., Li, K., & Li, K. (2020). *Computing, storage, and networking resource management in data centres, data centre handbook: Plan, design, build, and operations of a smart data centre* (2nd ed.). Wiley.
36. Cui, H., Ma, C., Lai, W., Zheng, L., & Liu, Y. (2015). Accurate network resource allocation in SDN according to traffic demand. In *Proceedings of the 4th international conference on mechatronics, materials, chemistry and computer engineering* (Advances in Computer Science Research). Atlantis Press.
37. Balasubramanian, V., Alokaily, M., Reisslein, M., & Scaglione, A. (2021). Intelligent resource management at the edge for ubiquitous IoT: An SDN-based federated learning approach. *IEEE Network*, 35(5), 114–121.
38. Mishra, P., Godfrey, W. W., & Kumar, N. (2021). A green computing-based algorithm in software defined network with enhanced performance. In *2021 International conference on computing, communication, and intelligent systems (ICCCIS)* (pp. 953–958). IEEE.
39. Pokhrel, S. R. (2021). Software defined internet of vehicles for automation and orchestration. *IEEE Transactions on Intelligent Transportation Systems*, 22(6), 3890–3899.
40. Cui, X., Meng, Q., & Wang, W. (2020). A load balancing mechanism for 5G data centres. In *2020 International wireless communications and mobile computing (IWCMC)* (pp. 812–815). IEEE.
41. Nooruzzaman, M., & Fernando, X. (2021). Hyperscale data centre networks with transparent HyperX architecture. *IEEE Communications Magazine*, 59(6), 120–125.
42. <https://blog.leviton.com/new-switch-architectures-and-impact-40100g-migration-data-center>
43. <https://medium.com/@fiberstoreorenda/how-will-sdn-change-the-future-network-a0bbad6a3fld>