Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

44

Cloud Computing: Issues and Existing Solutions

Mayur Patil
mayurpatil4339@gmail.com

Rutuja Parkale
rutuparkale123@gmail.com

Soham Deshpande
s17deshpande@gmail.com

Suraj Bhogulkar
surajbhogulkar20@gmail.com

Dr. Vishwanath Karad MIT World Peace University, Pune

Abstract

The computational universe has grown in size and complexity. The cloud is a new technology in the Information Technology (IT) industry that provides IT resources over the internet. Cloud-based services are available on demand, scalable, device agnostic, and dependable. Virtualization is the foundation of cloud computing. Due to cost, virtualization, elasticity, broad network access, metered service features cloud computing becomes integral part of almost all small to medium to large businesses. Cloud services providers and cloud users experiences several key issues of cloud computing such as vendor lock-in, security, incast, scalability, availability, load balancing, performance, data lock-in and many more. In this paper, authors have reviewed security, incast, load balancing, scalability issues of cloud computing. Furthermore, solutions to security, incast, load balancing and scalability issues are analyzed comprehensively.

Keywords: Cloud computing, security, incast, scalability, load balancing

1. Introduction

The cloud computing framework has been anticipated since nearly the middle of the last century. It provides scalable or elastic computer technology on virtually all platforms. Computing devices that support all existing and archaic software tools and technologies and are served via disparate networks, allowing it to be platform independent, portable, and

ubiquitous. Similarly, the capacity to provide service on-demand, share, and instantly commission and decommission configurable computing resources makes it resilient, sustainable, and near-utility computing. It provides software as a service (SaaS), platform as a service (PaaS), and infrastructure as a service (IaaS) service. Application software services are referred to as SaaS, application infrastructure services are referred to as PaaS, and system infrastructure services are referred to as IaaS. Cloud computing technology is available in four flavors: private, community, public, and hybrid cloud [1].

The reduced cost is frequently what draws organizations to the cloud. Customers do not have to worry about hardware maintenance and upgrade costs, or the additional cost associated with underutilized physical systems, because they are charged per execution-hour or gigabyte of storage. The use of virtualization enables easy scalability, whether through the duplication of instances or by changing the amount of CPU and memory available on a virtual machine. There are several advantages to mobility. The location and placement of resources in the cloud has no bearing on information access. The benefit of cloud computing is the execution environment and information can be located near to the location of highest demand, which is a benefit. The cloud computing environment shifts physical system administration to the cloud provider, resulting in centralized administration of cloud services. This enables customers' IT departments to concentrate on the solutions of their organizations. Most cloud service providers host customer data in multiple locations. This distributed resource approach results in system redundancy, if some of the resources are depleted, the effect on the remaining resources will be minimal.

Cloud computing is accessed remotely, it presents a number of challenges, including high security and a plethora of other complex technical demands. Common cloud computing issues include common infrastructure, safety, data lock-in, unpredictable outcomes, data transfer bottlenecks, lack of control, insiders, out-of-scope employee errors, reduced OS customizability, repairing, physical data location, internet risks, audit, traceability, encryption, authentication and authorization, and confidentiality and privacy [11,30-34]. This paper reviews a few important issues and their existing solutions made available by researchers. Section 2 describes major issues in cloud computing. Section 3 contains existing

solutions to issues described in section 2. Analysis about existing solutions is given in section 4. Finally, the paper is concluded in section 5.

2. Issues in cloud computing

Cloud computing has several issues and challenges. In this paper we have studied 4 major issues of cloud computing viz security, incast, load balancing and scalability.

2.1 Security issue

Depending on a business's needs, privacy and security issues may arise. For instance, in order to protect the confidentiality and integrity of its customers, the banking industry depends heavily on data. Access, trust, virtualization, software, and computation are some of the concerns and difficulties associated with cloud computing security. Access and computation-related threats have been recognized as the most significant ones, making up around 51% of all threats. Cryptography, virus, storage, and sanitization provide computational challenges. Physical safety issues and authentication are difficulties with access. Due to the possibility of data being used by unauthorized individuals, confidentiality and data security are also top security issues. Another significant security problem with the cloud is lack of control. Furthermore, XML signature element wrapping, cloud malware injection attacks, hostile insiders, and cyberattacks are all possible in a cloud setting. The cloud layer is added with a malicious program by an attacker, who then treats it as a legitimate instance. In addition, a Trojan horse or malware could be uploaded to the cloud [25].

In a cloud environment, data security threats can be classified as either external or internal. Internal threats are primarily the result of insider attacks, while external threats are the result of outside attacks when data is accessed by a third party. Attackers have the ability to obtain a user's personal information. To ensure data availability, the cloud infrastructure must be scalable. Data dependability, privacy and confidentiality of data, data availability, data location, backup and recovery of data, data encryption are major security issues [28, 29].

2.2 Incast issue

Cloud data center networks (DCNs) use ToR switches with small-sized buffers, low propagation delay (with hundreds of microseconds of RTT) and high-bandwidth (1 Gbps and onwards) links [4]. Several data center applications follow a many-to-one communication

pattern at the access layer of DCN, in which numerous workers transfer data towards the same aggregator through a common ToR switch at the same time. This simultaneous busty transmission may overload the buffer of a ToR switch connected to the aggregator and cause timeout events after frequent packet drops. Due to this, goodput collapses and it is termed as an incast problem in DCN [5].

There are two types of incast problems i.e., Transmission Control Protocol (TCP) incast problem and Multipath TCP (MPTCP) incast problem. Under many-to-one communication pattern, multiple single-homed concurrent workers use single path TCP protocol to simultaneously transmit rack-local short flows data towards single-homed aggregator connected with common ToR switch. This creates huge congestion at bottleneck ToR switch connected to single-homed aggregator. This results into TCP incast at access layer of single-homed DCN [6]. Under many-to-one communication pattern, several multi-homed concurrent workers use MPTCP protocol to simultaneously transmit rack-local short flows data towards same multi-homed aggregator through their multiple subflows. This creates huge congestion at bottleneck ToR switches connected to multi-homed aggregator, causes severe packet loss and results into MPTCP incast at access layer of multi-homed DCN [7]. TCP incast and MPTCP incast problems have various consequences such as drop in goodput, increase in flow completion time and minimum fan-out in DCN. This eventually deteriorates the data center user's experience.

2.3 Load Balancing issue

In order to prevent any one server from becoming overloaded and experiencing performance problems or outages, load balancing is a technique that divides incoming network traffic among several servers. Load balancing enhances application efficiency and dependability by maximizing throughput, reducing reaction times, and preventing service interruptions. According to numerous techniques, including round-robin, least connections, IP hash, and others, load balancing distributes traffic requests among various servers. According to the selected methodology, the load balancer serves as a single-entry point for incoming traffic and distributes the requests to the servers.

The application layer, transport layer, and network layer are just a few of the network layers where load balancing can be used. With HTTP(S) load balancers, load balancing can be accomplished at the application layer by distributing traffic according to URLs, cookies, or other application-specific parameters. TCP/UDP load balancers can be used for load balancing at the transport layer, distributing traffic according to IP addresses, ports, or other transport layer parameters. DNS load balancers, which distribute traffic based on DNS requests, can be used for load balancing at the network layer. In web applications, e-commerce sites, online gaming, and other high-traffic services, load balancing is frequently used to increase the scalability, availability, and performance of programs.

While load balancing can significantly improve the performance and availability of an application, there can also be certain difficulties and problems. Many typical load balancing problems include the following:

Single Point of Failure: If load balancers malfunction or get overloaded, they may become a single point of failure in and of itself, which may result in system outage.

Insufficient Scaling: If the load balancer is improperly configured, it may not be able to handle the load or scale up to meet rising traffic demands, which would negatively affect the performance and response times of the application.

Inefficient Traffic Distribution: Load balancing algorithms occasionally distribute traffic in an uneven or inefficient manner, resulting in a less-than-ideal distribution of resources and lengthier response times

Lack of Monitoring: If the load balancer is not sufficiently monitored, it may fail to identify difficulties or issues with the servers, resulting in slower response times and even downtime.

Security Concerns: Load balancers can become a target for attacks, and hackers may try to take advantage of flaws in the load balancer to access the network or interrupt services.

Configuration Complexity: Setting up and operating a load balancer can be difficult and need specific knowledge and skills, which might result in mistakes and incorrect configurations that can create issues.

Selecting a trustworthy load balancer is crucial, as is making sure it is properly configured, well-monitored, and consistently patched with security updates. Moreover, routine load testing can aid in seeing and resolving possible problems before they arise. Also, if there is a breakdown or overload, having a backup or failover load balancer can lessen the chance of downtime.

2.4 Scalability issue

Scalability issues in cloud computing refer to the challenges that arise when expanding or shrinking cloud resources to accommodate changes in workload demands. These issues can impact the performance, availability, and cost of cloud services. Scalability is one of the primary benefits of cloud computing, which allows businesses to easily expand or shrink their computing resources according to their needs. However, scalability can also present certain challenges, particularly related to performance and cost, which are known as scalability issues.

As the workload on cloud resources increases, it can lead to performance degradation, resulting in slower response times and decreased availability of the application [2]. While scaling up cloud resources can improve performance, it can also lead to increased costs [3]. Allocating the right number of resources to meet changing demand can be a challenge. Under-provisioning can lead to poor performance, while over-provisioning can result in higher costs [8]. Scaling a complex, distributed system can create integration challenges, particularly when integrating with legacy systems or third-party services [9]. As the number of cloud resources increases, the risk of security breaches and data leaks can also increase. This can be caused by a number of factors, including inadequate security controls, misconfigured resources, or insider threats [10].

3. Existing solutions to cloud computing issues

3.1 Existing solutions to Security

Virtualization of Computer Systems: A number of security risks involved with guest virtualization, such as hypervisors or VMMs, and host virtualization, such as Virtual Machines VMs. The problems occur if the hypervisors are affected as a result of some of the VM gaining privileged access. This malicious VM can then perform malicious operations on

other VMs in the multi-tenant environment. This occurs when hackers exploit loopholes in the hypervisor's software [36].

Programming for Applications: Cloud computing heavily relies on web applications or web services, as well as SOA. The OWASP list of the top most critical web application security risks are crucial [20]. The study also discusses the exploitability, prevalence and delectability, and technical impact of each risk. Injection (Structured Query Language SQL, Operating System OS, and Lightweight Directory Access Protocol (LDAP)), Insecure Direct Object Reference (DOR), Security Misconfiguration, and Missing Function Level Access Control are all easily exploitable [35].

Integrated Security on Multiple Levels: There is a large body of literature debating and promoting the concept of security-as-a-service. However, there are clear negative implications to this thought that are extremely detrimental to the spread of cloud computing. For starters, it implies that one must pay directly for security, which is a significant deterrent in an era of scarce competitive resources. Second, it implies that security is a luxury of effluents, undermining efforts to close the digital divide. The worst implications are that it sends the message that without a security-as-a-service subscription, your resources are not secure, which is a major deterrent for potential adopters [37].

Many security measures can be taken, including standardizing APIs, establishing a public key infrastructure, and disseminating data. In order to increase security levels in cloud computing, access control, authentication, and authorization are crucial. Cloud computing is one of the most serious security issues currently being addressed. A lack of security measures, or their ineffective implementation, can pose significant data-transmission risks [26, 27].

3.2 Existing solutions to Incast

3.2.1 Existing solutions to TCP Incast

Reducing RTO_{min} Timer: A smaller RTO_{min} value allows each worker to quickly perform retransmissions after immediately detecting packet loss. Reducing RTO_{min} timer method ensures efficient utilization of aggregator's link capacity [12].

Incast Congestion Control for TCP (ICTCP): ICTCP controls transmission rates of workers by adjusting advertised window size. This helps to prevent bottleneck switch buffer overflow.

ICTCP uses measured available bandwidth at aggregator and RTT values to adjust advertised window size. ICTCP necessitates per RTT calculation of throughput for each TCP flow. In ICTCP, at the aggregator side, on top of TCP layer, an additional shim-layer implementation is required [13].

Data Center TCP (DCTCP): DCTCP keeps switch buffer occupancy under threshold value. Switch marks new incoming data packet with CE codepoint, if switch buffer is occupied beyond marking threshold K. DCTCP causes premature indication of congestion as it marks packets very early. DCTCP requires worker TCP, aggregator TCP and switch operation modifications [14].

Adaptive Application-layer Incast Control scheme (AAIC): Depending upon current network situation and number of concurrent flows, AAIC equally sets the advertised window size of each concurrent flow. Furthermore, AAIC dynamically regulates the number of concurrent flows using a sliding-connection-window mechanism [15].

Proactive Incast Congestion Control system (PICC): Frequently requested data objects (i.e., popular data objects) are placed into selected workers. Such kind of data placement into a limited number of workers avoids incast congestion. Moreover, PICC identifies data objects that are concurrently requested and are re-allocated into the same workers [16].

Cross-Layer Flow Schedule with Dynamic Grouping (CLFS-DG): Aggregator use application-level information to organize data transmission schedules. Grouping of multiple workers is done for performing concurrent transmissions but simultaneously flow completion deadline is also satisfied. Finishing flow transmission before exceeding switch buffer capacity and forming optimal groups of workers are the objectives of CLFS-DG [17].

Enhanced DSRAL (EDSRAL): This application layer technique allows consecutive workers to overlap their SRU transmissions by using Flow Overlapping Factor (FOF). This helps EDSRAL to efficiently utilize DCN links at maximum level, offer higher application goodput and reduce flow completion time [18].

3.2.2 Existing solution to MPTCP Incast:

Equally-Weighted Multipath TCP (EW-MPTCP): EW-MPTCP alleviates MPTCP incast by weighting subflows of each MPTCP connection. That is, EW-MPTCP controls

aggressiveness of each MPTCP subflow competing at a shared bottleneck irrespective of number of subflows and concurrent workers [19].

Maximum Multipath TCP (MMPTCP): MMPTCP attempts to decrease latency-sensitive short flows completion time and to increase goodput of long flows. MMPTCP uses two protocols i.e., Packet Scatter (PS) protocol to transmit initial specific number of bytes under single congestion window and MPTCP protocol to transmit remaining bytes through multiple congestion windows [20].

Queuing Cache Balance Factor (QCBF): In QCBF, initially, ToRs cluster cache balance queue is built by proposing buffer pool balance factor. Then based on this buffer pool balance factor, congestion window of each concurrent subflow is determined [21].

Advanced MPTCP (AMP): AMP strives to avoid MPTCP incast by decreasing short flows delay and improving large flows throughput. This is done by detecting and controlling congestion through adjusting time granularity. AMP adjusts the subflow congestion window by tracking arrived ECN-marked packet position in congestion window [22].

Datacenter MPTCP (DCMPTCP): DCMPTCP reduces overhead by removing unnecessary subflows. This is done after identifying many-to-one rack local traffic. Upon detecting many-to-one rack local traffic, FACT algorithm allows MPTCP to switch to standard TCP by removing additional subflows [23].

Balanced Multipath TCP (BMPTCP): BMPTCP protocol is proposed to efficiently mitigate MPTCP incast in multi-homed data center networks. BMPTCP computes and controls subflow congestion window sizes by considering ToR switch buffer size, total header size of data packets and count of total flows traversing through bottleneck interface of ToR switch. It maintains identical congestion window size for all concurrent subflows to avoid timeout events due to full window loss at ToR switch [24].

3.3 Existing solutions to Load balancing

Load balancing: Load balancing distributes workloads across multiple servers, reducing the load on any single server and improving performance and availability.

There are several solutions for load balancing, each with its own advantages and disadvantages. Here are some popular load balancing solutions:

Hardware Load Balancers: These are specialized hardware devices designed to handle high traffic loads and provide advanced features such as SSL offloading, caching and security features. Hardware load balancers can be expensive, but they offer excellent performance and scalability.

Software Load Balancers: These are software-based load balancing solutions that can be used on regular servers or virtualized environments. They are more cost effective than hardware load balancers and offer many of the same features.

Cloud Load Balancers: These are load balancing solutions offered by cloud service providers such as Amazon Web Services (AWS) Elastic Load Balancing (ELB), Microsoft Azure Load Balancer and Google Cloud Load Balancing. These services are scalable, reliable and offer advanced features such as auto-scaling, health checks and geo-routing.

DNS Load Balancing: DNS load balancing means using DNS servers to distribute traffic to different servers based on DNS queries. This solution can be cost-effective, but may not be as reliable or offer as many features as hardware or software load balancers.

Each of these load balancing options has advantages and disadvantages of its own, and the best option will be determined by a number of variables, including the application type, traffic load, budget, and scalability needs.

3.4 Existing solutions to Scalability

Auto-scaling techniques: Auto-scaling is a technique used in cloud computing to automatically adjust the number of resources based on the workload. This technique can help to optimize the performance of the system and reduce the cost of the infrastructure [38].

Caching: Caching is a technique used to store frequently accessed data in memory or on disk, in order to reduce the amount of time it takes to access the data. This technique can help to improve the performance of the system and reduce the load on the database [39].

Database optimization: Database optimization is a technique used to improve the performance and scalability of a database in a cloud computing environment. This technique

can involve optimizing the database schema, tuning the database configuration, and using techniques such as sharding and replication [40].

Predictive scaling: Predictive scaling is a technique used to predict future resource usage based on historical data, in order to proactively scale the infrastructure. This technique can help to ensure that the system is always available and responsive, even during periods of high demand [41].

Cloud-native architectures: Cloud-native architectures are a set of design principles and practices that enable applications to be developed and deployed in a cloud computing environment. This approach can help to improve the scalability, availability, and resilience of the system.

Multi-cloud and hybrid cloud: multi-cloud and hybrid cloud architectures involve using multiple cloud providers or combining on-premises and cloud-based resources to provide greater scalability, flexibility, and resilience.

4. Analysis and discussion

This section analyses few of the existing solutions to security, incast, load balancing and scalability in cloud computing technology. Table-1 shows issues and their existing solutions.

Table-1: Cloud computing issues and existing key solutions

Cloud computing issues	Existing key solutions
Security	Strong access control
	Regular security assessments
	Keeping the hypervisor software up to date with security patches
	Establishing a public key infrastructure
	Authentication and authorization
	Use of security controls such as firewalls and intrusion detection systems

Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

TCP Incast	Reducing RTO_{min} Timer
	Incast Congestion Control for TCP (ICTCP)
	Data Center TCP (DCTCP):
	Adaptive Application-layer Incast Control scheme (AAIC)
	Proactive Incast Congestion Control system (PICC)
	Cross-Layer Flow Schedule with Dynamic Grouping (CLFS-DG)
	Enhanced DSRAL (EDSRAL)
MPTCP Incast	Equally-Weighted Multipath TCP (EW-MPTCP)
	Maximum Multipath TCP (MMPTCP)
	Queuing Cache Balance Factor (QCBF)
	Advanced MPTCP (AMP)
	Datacenter MPTCP (DCMPTCP)
	Balanced Multipath TCP (BMPTCP)
Load balancing	Hardware Load Balancers
	Software Load Balancers
	Cloud Load Balancers
	DNS Load Balancing
Scalability	Auto-scaling techniques
	Caching

	Database optimization
	Predictive scaling
	Cloud-native architectures

Encryption technique is implemented at application layer hence it becomes an efficient technique. Access control mechanism is implemented at all layers that is physical, link, network, transport and application layer of TCP/IP protocol suit. Due to this, Access control mechanism is considered as robust to ensure security in cloud environment as it can be implemented at all the layers of TCP/IP protocol suite. Multi-factor authentication is one of the important techniques to ensure robust security in cloud environment as multiple forms of authentications are required for attackers to get unauthorized access to cloud resources.

Application layer solutions such as AAIC, PICC, CLFS-GD, DSRAL mitigates TCP incast and considered as an easy technique compared to transport layer solutions as they do not demand modification to TCP/IP protocol stack and switch operations. Compared to application layer solutions transport layer solutions support large number of servers under many-to-one communication pattern. ECN-based solutions (AMP) are robust solutions to mitigate MPTCP incast as compared to window-based solutions (EW-MPTCP, BMPTCP, QCBF) and dynamic subflow management solutions (MMPTCP, DCMPTCP).

Software load balancers are cost effective solutions to load balancing as compared to hardware load balancers. But hardware load balancers efficiently perform load balancing as compared to software load balancers.

Many organizations address scalability issues by adopting various strategies, such as using auto-scaling tools to adjust resource allocation automatically, implementing caching mechanisms to improve performance, and using cloud-native security tools to enhance security. It is important to carefully plan and test cloud infrastructure to ensure it scales effectively while maintaining optimal performance and cost efficiency. In summary, there are a range of existing and suggested solutions to address scalability issues in cloud computing,

ranging from auto-scaling and load balancing to predictive scaling and cloud-native architectures. By adopting these solutions, organizations can improve the performance, availability, and cost-efficiency of their cloud services.

5. Conclusion

Cloud computing model is well adopted by large community of service providers, small to medium to large organizations and individuals. Cost effectiveness, metered-service, elasticity, broad network features of cloud computing make it exponentially popular. Users' experiences some of the potential issues while accessing services of cloud computing. These issues include security, incast, load balancing, scalability. Multi-factor authentication, access control, encryption, intrusion detection systems, firewalls are some of the popular solutions to ensure security in cloud environment. Rate-based and ECN-based transport layer solutions have potential to effectively mitigate TCP and MPTCP incast issue in cloud data centers. Hardware load balancers are widely used in cloud environment to ensure load balancing. Scalability issue competently mitigated through auto-scaling techniques.

References:

[1] P. Mell and T. Grance, "The NIST Definition of Cloud Computing," National Institute of Standards and Technology Draft (NIST) Special Publication 800-145, 2011. [Online]. Available: http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecial-publication800- 145.pdf

[2] Somasundaram, Thamarai Selvi, V. Prabha, and Mahesh Arumugam. "Scalability issues in cloud computing." In *2012 Fourth International Conference on Advanced Computing (ICoAC)*, pp. 1-5. IEEE, 2012.

[3] Coutinho, Emanuel Ferreira, Flávio Rubens de Carvalho Sousa, Paulo Antonio Leal Rego, Danielo Gonçalves Gomes, and José Neuman de Souza. "Elasticity in cloud computing: a survey." *annals of telecommunications-annales des télécommunications* 70 (2015): 289-309.

[4] Yu, Ye, and Chen Qian. "Space shuffle: A scalable, flexible, and high-bandwidth data center network." In *2014 IEEE 22nd International Conference on Network Protocols*, pp. 13-24. IEEE, 2014.

[5] Phanishayee, Amar, Elie Krevat, Vijay Vasudevan, David G. Andersen, Gregory R. Ganger, Garth A. Gibson, and Srinivasan Seshan. "Measurement and analysis of TCP throughput collapse in cluster-based storage systems." In *FAST*, vol. 8, pp. 1-14. 2008.

[6] Chen, Yanpei, Rean Griffith, Junda Liu, Randy H. Katz, and Anthony D. Joseph. "Understanding TCP incast throughput collapse in datacenter networks." In *Proceedings of the 1st ACM workshop on Research on enterprise networking*, pp. 73-82. 2009.

[7] Li, Ming, Andrey Lukyanenko, Sasu Tarkoma, and Antti Ylä-Jääski. "MPTCP incast in data center networks." *China Communications* 11, no. 4 (2014): 25-37.

[8] Ab Rashid Dar, Ravindran D. "Survey on Scalability in Cloud Environment." *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)* 5, no. 7 (2016): 2124-2128.

[9] Dillon, Tharam, Chen Wu, and Elizabeth Chang. "Cloud computing: issues and challenges." In *2010 24th IEEE international conference on advanced information networking and applications*, pp. 27-33. Ieee, 2010.

[10] Kuyoro, S. O., F. Ibikunle, and O. Awodele. "Cloud computing security issues and challenges." *International Journal of Computer Networks (IJCN)* 3, no. 5 (2011): 247-255.

[11] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, et al., "A view of cloud computing," Communications of the ACM, vol. 53, pp. 50-58, 2010.

[12] Vasudevan, Vijay, Amar Phanishayee, Hiral Shah, Elie Krevat, David G. Andersen, Gregory R. Ganger, Garth A. Gibson, and Brian Mueller. "Safe and effective fine-grained TCP retransmissions for datacenter communication." *ACM SIGCOMM computer communication review* 39, no. 4 (2009): 303-314.

[13] Wu, Haitao, Zhenqian Feng, Chuanxiong Guo, and Yongguang Zhang. "ICTCP: Incast congestion control for TCP in data center networks." In *Proceedings of the 6th International COnference*, pp. 1-12. 2010.

[14] Alizadeh, Mohammad, Albert Greenberg, David A. Maltz, Jitendra Padhye, Parveen Patel, Balaji Prabhakar, Sudipta Sengupta, and Murari Sridharan. "Data center tcp (dctcp)." In *Proceedings of the ACM SIGCOMM 2010 Conference*, pp. 63-74. 2010.

[15] Luo, Jintang, Xiaolong Yang, Jie Xu, and Jian Sun. "AAIC: adaptive-sliding-connection-window solution to TCP incast from application layer." *IEEE Communications Letters* 20, no. 10 (2016): 1967-1970.

[16] Wang, Haoyu, and Haiying Shen. "Proactive incast congestion control in a datacenter serving web applications." In *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*, pp. 19-27. IEEE, 2018.

[17] Tseng, Hsueh-Wen, Wan-Chi Chang, I-Hsuan Peng, and Pei-Shan Chen. "A cross-layer flow schedule with dynamical grouping for mitigating larger-scale TCP incast." *ACM SIGAPP Applied Computing Review* 17, no. 1 (2017): 15-25.

[18] Suryavanshi, Mahendra, and Jyoti Yadav. "Mitigating TCP incast in data center networks using enhanced application layer technique." *International Journal of Information Technology* 14, no. 5 (2022): 2523-2531.

[19] Li, Ming, Andrey Lukyanenko, Sasu Tarkoma, and Antti Ylä-Jääski. "MPTCP incast in data center networks." *China Communications* 11, no. 4 (2014): 25-37.

[20] Kheirkhah, Morteza, Ian Wakeman, and George Parisis. "MMPTCP: A multipath transport protocol for data centers." In *IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications*, pp. 1-9. IEEE, 2016.

[21] Pang, Shanchen, Jiamin Yao, Xun Wang, Tong Ding, and Li Zhang. "Transmission control of MPTCP incast based on buffer balance factor allocation in data center networks." *IEEE Access* 7 (2019): 183428-183434.

[22] Ye, Jin, Luting Feng, Ziqi Xie, Jiawei Huang, and Xiaohuan Li. "Fine-grained congestion control for multipath TCP in data center networks." *IEEE Access* 7 (2019): 31782-31790.

[23] Dong, Enhuan, Xiaoming Fu, Mingwei Xu, and Yuan Yang. "Dcmptcp: Host-based load balancing for datacenters." In *2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS)*, pp. 622-633. IEEE, 2018.

[24] Suryavanshi, Mahendra, Ajay Kumar, and Jyoti Yadav. "Balanced Multipath Transport Protocol for Mitigating MPTCP Incast in Data Center Networks." *International Journal of Next-Generation Computing* 12, no. 3 (2021).

[25] D. Jamil and H. Zaki, "Security issues in cloud computing and countermeasures," International Journal of Engineering Science and Technology (IJEST), vol. 3, no. 4, pp. 2672–2676, 2011.

[26] M. S. Almutairi, "Cloud computing: Securing without losing control," Journal of Advances in Mathematics and Computer Science, vol. 31, no. 2, pp. 1–9, 2019.

[27] R. V. Rao and K. Selvamani, "Data security challenges and its solutions in cloud computing," Procedia Computer Science, vol. 48, pp. 204–209, 2015.

[28] M. A. Razzaq, J. A. Mahar, M. A. Qureshi and Z. U. Abidin, "Smart campus system using internet of things: Simulation and assessment of vertical scalability," Indian Journal of Science and Technology, vol. 13, no. 28, pp. 2902–2910, 2020.

[29] S. A. Hussain, M. Fatima, A. Saeed, I. Raza and R. K. Shahzad, "Multilevel classification of security concerns in cloud computing," Applied Computing and Informatics, vol. 13, no. 1, pp. 57–65, 2017.

[30] Attiya and X. Zhang, "Cloud Computing Technology: Promises and Concerns," International Journal of Computer Applications, vol. 159, 2017.

[31] S. Singhal and J. Grover, "Hybrid biogeography algorithm for reducing power consumption in cloud computing," in Advances in Computing, Communications and Informatics (ICACCI), 2017 International Conference on, 2017, pp. 121-124.

[32] P. A. Garcıa, J. M. M. Fernández, J. L. A. Rodrigo, and R. Buyya, "Proactive power and thermal aware optimizations for energyefficient cloud computing," 2017.

[33] K.-H. Park and H.-S. Jang, "A Study of Cloud Computing-based Disaster Recovery System for Securing High Availability of Academic Affairs Information Service," International Information Institute (Tokyo). Information, vol. 20, p. 567, 2017.

[34] M. T. Amron, R. Ibrahim, and S. Chuprat, "A Review on Cloud Computing Acceptance Factors," Procedia Computer Science, vol. 124, pp. 639-646, 2017.

[35] OWASP Top 10, "The Ten Most Critical Web Application Security Risks," 2013.

[36] H. Orman, "Both Sides Now: Thinking about Cloud Security," IEEE Internet Comput., vol. 20, no. 1, pp. 83–87, 2016.

[37] M. Ali, S. U. Khan, and A. V. Vasilakos, "Security in cloud computing: Opportunities and challenges," Inf. Sci. (Ny)., vol. 305, pp. 357–383, 2015.

[38] Arvindhan, M., and Abhineet Anand. "Scheming an proficient auto scaling technique for minimizing response time in load balancing on Amazon AWS Cloud." In *International Conference on Advances in Engineering Science Management & Technology (ICAESMT)-2019, Uttaranchal University, Dehradun, India*. 2019.

[39] Nishtala, Rajesh, Hans Fugal, Steven Grimm, Marc Kwiatkowski, Herman Lee, Harry C. Li, Ryan McElroy et al. "Scaling memcache at facebook." In *Presented as part of the 10th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 13)*, pp. 385-398. 2013.

[40] Shute, Jeff, Radek Vingralek, Bart Samwel, Ben Handy, Chad Whipkey, Eric Rollins, Mircea Oancea et al. "F1: A distributed SQL database that scales." (2013).

[41] Al-Dulaimy, Auday, Javid Taheri, Andreas Kassler, M. Reza HoseinyFarahabady, Shuiguang Deng, and Albert Zomaya. "MULTISCALER: A multi-loop auto-scaling approach for cloud-based applications." *IEEE Transactions on Cloud Computing* 10, no. 4 (2020): 2769-2786.