

**CHENNAI – PONDICHERRY**

**StarCube: An On-Demand and Cost-Effective Framework for Cloud Data Center Networks with Performance Guarantee**

**Abstract:**

In this paper, we propose a resource management framework called *StarCube*, which guarantees non-blocking resource allocation and topology-preserving reallocation for fat-tree based multi-tenant cloud data centers. With *StarCube*, each cloud service is allocated an isolated non-blocking virtual network topology, and the topology provisioned to each service is guaranteed logically unchanged during and after virtual machine reallocation. This resource management problem is formulated and proved to be NP-complete. To achieve high resource efficiency in acceptable time, we propose a cost-effective algorithm with polynomial-time complexity based on *StarCube* for on-demand resource allocation and reallocation. We demonstrate via extensive simulations that the server resources in *StarCube*-based cloud data centers can be nearly fully utilized with negligible reallocation cost. The results also show that *StarCube* supports a large variety of service provisioning feasibly and efficiently for cloud data centers of various scales and with dynamic demands. To the best of our knowledge, *StarCube* is the first solution to allocating and reallocating cloud services for fat-tree networks with guarantee on non-blocking properties.

**Existing System:**

There are several requirements for allocating resources to latency-sensitive and communication-intensive appli-cations in cloud data centers with network performance guarantees. The provisioned network topology should be consistently non-blocking for each hosting service, hence allowing computing nodes to support arbitrarily intra-service communication patterns at any time. The required network bandwidth and latency of intra-service commu-nications should be guaranteed and kept consistent for the predictability of service response timeand task com-pletion time. Particularly, the consistency of network per-formance guarantees should hold during the entire ser-vice lifetime even when some of the computing nodes are being reallocated. Services operated by different uncoop-erative tenants should be isolated in order to prevent cross-service interference and unfairness, and to ensure the provisioned network is congestion-free.

**Proposed System:**

Virtualization is widely adopted in modern cloud data centers for efficient resource management. Through virtualization, the instances of computing nodes can be hosted by virtual machines (VMs) for flexible allocation and reallocation among servers. Live VM migration is another promising online approach to reallocate work-loads among physical servers with negligibly service downtime. However, without proper schedulingand routing, the migration and workload traffic generated by other services may compete for network bandwidth, and therefore, the total migration time could be prolonged due to the lower transfer rate. In this paper, we will take into account the available bandwidth during the migra-tion process and the migration path construction in real-location scheduling so as to ensure both predictable mi-gration time and network performanceguaranteefor cloud services.