

**CHENNAI – PONDICHERRY**

**ATME: Accurate Traffic Matrix Estimation in both Public and Private Datacenter Networks**

**Abstract:**

Understanding the pattern of end-to-end traffic flows in datacenter networks (DCNs) is essential to many DCN designs and operations (e.g., traffic engineering and load balancing). However, little research work has been done to obtain traffic information efficiently and yet accurately. Researchers often assume the availability of traffic tracing tools (e.g., OpenFlow) when their proposals require traffic information as input, but these tools may have high monitoring overhead and consume significant switch resources even if they are available in a DCN. Although estimating the traffic matrix (TM) between origin-destination pairs using only basic switch SNMP counters is a mature practice in IP networks, traffic flows in DCNs show totally different characteristics, while the large number of redundant routes in a DCN further complicates the situation. To this end, we propose to utilize resource provisioning information in public cloud datacenters and the service placement information in private datacenters for deducing the correlations among top-of-rack switches, and to leverage the uneven traffic distribution in DCNs for reducing the number of routes potentially used by a flow. These allow us to develop ATME as an efficient TM estimation scheme that achieves high accuracy for both public and private DCNs. We compare our two algorithms with two existing representative methods through both experiments and simulations; the results strongly confirm the promising performance of our algorithms.

**Existing System:**

Existing proposals in need of detailed traffic flow information collect the flow traces by deploying additional modules on either switches or servers in small scale DCNs. However, both methods require substantial deployments and high administrative costs, and they are difficult to be implemented thanks to the heterogeneous nature of the hardware in DCNs. More specifically, the switchbased approaches, on one hand, need all the ToRs to support flow tracing tools such as OpenFlow, and consume a substantial number of switch resources to maintain the flow entries.1 On the other hand, the server-based approaches, which require instrumenting all the servers or VMs to support data collection, are unavailable in most datacenters and are nearly impossible to be implemented peacefully and quickly while supporting a lot of cloud services in large scale DCNs.

**Proposed System:**

We creatively use resource provisioning information in public datacenters for deriving the prior TM among ToRs. We group all the VMs into several clusters with respect to different users, resulting in the effect that communications only happen within the same cluster and the potential traffic patterns are epitomized among all VMs in turn.

We pioneer in using the service placement information in private datacenters to deduce the correlations of ToR switch pairs, and we also propose a simple method to evaluate the correlation factor for each ToR pair. Our traffic model, assuming that ToR pairs with a high correlation factor may exchange higher traffic volumes, is far more accurate for DCNs than conventional models used for IP networks.