USE CASES (5) SE COMPUTING

DILIP KRISHNASWAMY (DILIP@IEEE.ORG) PRESENTED IN THE NFVRG INTERIM MEETING ON "MODEL-DRIVEN AND AI-ENABLED INTER-CLOUD OPTIMIZATION, APRIL 24, 2018



INTER-DC VIRTUAL FUNCTION PLACEMENT & RESOURCE MGMT

Map users to VNFs on DCs based on latency, availability, DC load, energy, mobility Dynamically direct new user flows to utilize VNFs at the most appropriate DCs Dynamically divert new VNF resource requirements away from a DC if heavily loaded Example Depiction – Serving users in Sunnyvale and Cupertino CA

> Arizona DC (Solar Powered)

San Francisco DC

Low Energy cost, Higher Latency

High Energy cost, Medium Latency, Higher Capacity compared to Small DC

Cupertino Small DC Sunnyvale Small DC Lower Latency, Lower Capacity High Energy Cost

IEEE NFV-SDN 2015

DISTRIBUTED FUNCTIONS VIRTUALIZATION (DFV)

With the availability of compute & storage in-networks, explore the placement of network / service / application functions across DCs



How can we best place VNFs in hierarchical data centers taking care of latency constraints associated with VNFs, user mobility, energy cost of utilization, resource availability

3

How can we best utilize the dynamic availability of such distributed compute/network/storage/energy resources?

DISTRIBUTED SDN FOR DISTRIBUTED NFY Distributed SDN and NFV will enable smart distributed processing of functions acrossidata centers



 \rightarrow Partition / Collapse / Replicate functions across data centers

- Address latency constraints, user mobility, dynamic resource availability, security (compute/network/storage/energy)
- \rightarrow Dynamic Monitoring, Analytics, Optimization, Orchestration, Scaling

HIERARCHICAL COLLAPSED FUNCTIONS (IEEE NFV-SDN'15)



GENERALIZED DFV (IEEE NFV-SDN 2015)

NETWORK, SERVICE, AND APPLICATION FUNCTION VM PARTITIONING



IOBILITY AWARE VNF-PLACEMENT (IEEE NFV-SDN 2017) (COLLABORATION WITH AKANKSHA PATEL PROF MYTHILI VUTUKURU)



Fig. 1. Physical substrate network of a telecom operator



Fig. 8. Comparison of time taken to obtain the optimal placement by running model on complete graph vs subgraphs



Fig. 4. Division of graph into subgraphs



Fig. 6. Comparison of the average handover latency for different approaches (min, median, and max values have been plotted)



THANK YOU

• Questions?