Federate Orchestration for Cloud-Fog Infrastructure

Dominic Lindsay

Lancaster Univeristy

February 5, 2020

Overview
Todays Presentation...

- Introduction
- *Fog-Cloud Computing*
- Orchestration at the edge
- Quality of Service
- Problem Statement
- Proposal
- System Design
- HotCloud'20
- Questions

Cloud-Fog Computing
Cloud-Fog Continuum

Hierarcichal Network

- *Cloud-Fogs* are composed of hierarchal networks.
- Nodes top possess more *resources* and exhibit higher *availability*.

Characteristics

1. Distributed Infrastructure
2. Non-Neglible Latencies
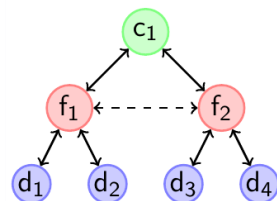3. Limited Availability and Reachability

Figure: Cloud-Fog Network. Cloud Clusters possessing the most resources are found at the top with high availability, whilst Fog Node/Cloudlets are in middle tier, posess less resources but possess lower latencies and higher aggregate bandwidth.

- Orchestration
  - Scheduling and Resource Management
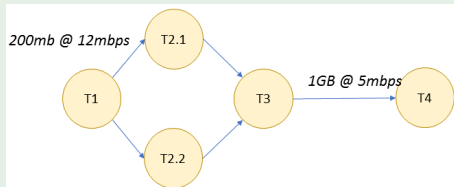 - Cloud Orchestration

Application Contexts
Applications DAGs

Application DAG

- Application submit jobs as DAGs.
- Vertices identify tasks and their requirements
  e.g. spark executor ⟨2GB, 1CPU⟩
- Edges identify both *time and space* constraints.
- Task are executed by workers on behalf of an *Application Master*.
- *or clusters*

Example *DAG*

Quality of Service

Application Performance

Applications are composed of individual services
Workers execute work on behalf of Application Masters

Intra-Service Quality of Service

- Application depends on multiple services for execution of tasks.
- Intra-Node *Bandwidth* and *Latency* dictates *QoS* of task executors.

Fog-Cloud Orchestration

Centralised Orchestration

Orchestration Systems typically manage only a single cluster.
As such resources are manage as a single flat hierarchy.

False Assumptions

Centralised orchestration systems exhibit several limitations when managing *Cloud-Fogs*

- Non-Negligible Intra-Cluster Latencies
- Unlimited intra-cluster bandwidths
- Uniform Sub-Cluster Resource Utilization
- Uniform Sub-Cluster Workload Affinity

Leading to poor quality placement and degraded application performance

Proposal
Federate Orchestration Framework

Hierarchical Resources
Fog-Cloud computing require orchestration mechanisms capable of adapting to dynamic sub-cluster utilization and dynamic bandwidth and latency constraints.

Objective
The objective of this work is to investigate new orchestration mechanisms for resource management in decentralised Cloud-Fogs infrastructures, capable of making *QoS* aware placements in dynamic hierarchical infrastructures.

Contributions
- Two-Layer Federate, Mobility Aware Orchestration Platform
- Bandwidth and Latency Migration Policy
- High Level Application Level Constraints.

Requirements

Cluster Aware

Workload posses affinities to differnt scheduling archetectures
Cluster Utilization impacts workload performance

Quality of Service Aware

Quality of Service is dictated not only by resource (CPU, mem, etc) but by
bandwidth and latency across sub-clusters

Application Aware

Careful placement of a workload can lead to increased performance.
Scheduler should be aware of executor population.
Provide Cluster Affinity/Anti-Affinity and Cardinality.

System Diagram
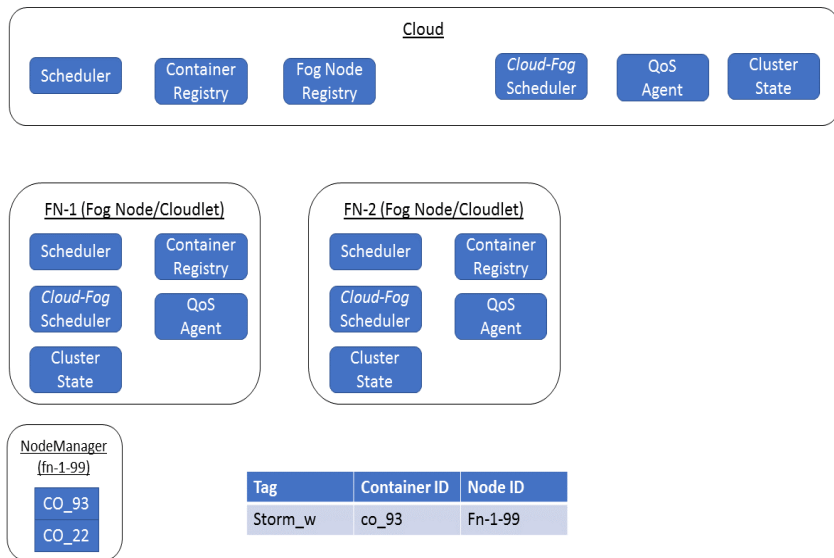
- USENIX workshop
- Fancy Companies
- Deadline: 17th March

Questions?