FairCloud - CS294 Class Project

Gautam Kumar
Mosharaf Chowdhury

Problem Formulation

Problem Definition & Importance Existing cluster resource sharing mechanisms provision CPU and memory resources to the cloud tenants depending on their individual payments. However, the network often becomes the bottleneck in cloud environments [1]. In this project, we want to address the issues related to sharing the network while ensuring payment-proportional allocations, certain guarantees on minimum bandwidth, and additional properties.

Challenges The definition of "fair sharing" while treating network as a resource is still amenable to several definitions. Popa et al. [4] enlist a set of properties that fair allocations should adhere to and identify two properties, weight fidelity and bandwidth guarantee, that seem to have an inherent tradeoff. However, it is uncertain how strong this tradeoff is and which factors affect it. Sharing network is different from allocation of CPU or memory because it cannot be attributed to just a single entity. It is also not clear, even with the formal properties to dictate fairness on individual links, if an allocation is better than the other on a global scale.

Approach

Related Work Current approaches to this problem include per-flow allocation, allocation based on per source-destination pair, and attributing the usage to only the source (or the destination). These approaches, however, defy at least one of the required properties outlined in [4] along with presenting a mediocre point in the weight fidelity–bandwidth guarantee tradeoff.

Idea We build on the sharing scheme proposed in [4] that satisfy all the desirable properties along with giving control knobs to achieve different price points. The key idea is to share a link based on the participating entities on both sides of the link based on their respective weights, while ensuring guarantees on bandwidth by scaling these weights down by the number of other entities that one communicates with. However, their current formulation deals only with tree-based datacenter topologies, and it is non-trivial to extend them to other topologies.

Deliverables / Milestones

Estimated Accomplishments

- Simulation is the first step to make meaningful observations and solve some of the questions posed earlier in this proposal. The simulation framework will also be useful to implement other strategies (e.g., [1]) to observe the global consequences of such strategies on different workloads and their convergence properties. Moreover, we want to develop a global metric that can differentiate between network sharing strategies by viewing the global picture.
- Implementing a working prototype is next logical step after simulation. We are considering the following three approaches:
  - Implementation of fair queueing in *Click* [2].
  - Implementation in *Open vSwitch* [3] that can maintain flow tables and handle expirations. We know about implementations of distributed rate-limiting in Open vSwitch but not of fair queueing.
  - Implementing as a shim layer, similar to [1], but approximating some of the feedback control information without changes to switches.
- Formalizing the tradeoff between weight-fidelity and minimum resource guarantees.
- Determining the topologies that are amenable to the proposed solution.

Other Objectives We plan on working on this project even beyond this course. In addition to the objectives listed earlier, we would like to ask if the properties presented in the paper [4] form an exhaustive set. We also want to answer whether their exists an algorithm that can work on different points in the tradeoff space. The proposed scheme is demand-agnostic; we need to address this as well. Finally, we want to learn if network share assignments can in turn help us in better VM placement across servers.

Intermediate Milestone We would like to complete at least the simulator and a rough prototype implementation by October 26th.

.

Requirements

Servers We need at least 2 servers to test our prototype implementation.

References

1. A. Shieh, S. Kandula, A. Greenberg, C. Kim, and B. Sahak. Sharing the Data Center Network, In *Usenix NSDI*, 2011.
2. E. Kohler, R. Morris, B. Chen, J. Jannotti, and M. Kaashoek. The Click Modular Router. In *ACM Trans. Comp. Sys.*, August 2000.
3. B. Pfa, J. Pettit, T. Kopenen, K. Amdion, M. Casado, and S. Shenker. Extending Networking into the Virtualization Layer. In *HotNets-VIII*, 2009.
4. L. Popa, A. Krishnamurthy, S. Ratnasamy, I. Stoica. *FairCloud*: Sharing the Network in Cloud Computing. In *HotNets-X*, 2011.