15-780 Project Proposal
Use ILP to Evaluate Indirection Scheduling Techniques for Hybrid Circuit/Packet Networks

Conglong Li (andrew ID conglonl), Thomas Kim (andrew ID tskim)
{*conglonl*,*thomas.kim*} @*cs.cmu.edu*

1 Introduction

Traditionally, circuit switching and packet switching have been considered as alternative design choices, each with their own advantages and disadvantages. Packet switches are efficient at multiplexing traffic across a large number of ports, while circuit switches can often service higher line rates in a more cost-effective manner, especially when combined with optical links. However, as optical circuit switching technology advances, the reconfiguration time (i.e., the amount of time it takes to alter the input-to-output mapping) is swiftly decreasing, thus increasingly blurring the division between the packet and circuit regimes.

Indeed, a range of new datacenter switch designs take advantage of this trend, and propose to schedule appropriately large traffic demands via a high-bandwidth circuit switch and handle any remaining traffic with a slower packet switch. However, all recent proposals for such hybrid designs presume the existence of an omniscient scheduling oracle that can compute switch configurations and map traffic to them in an optimal fashion. Recently, Liu *et al.* [7] recently proposed a hybrid switch scheduling algorithm, Solstice, that is both highly effective at scheduling datacenter-like traffic workloads and has practical computational overheads when doing so.

In this proposoal, we propose to improve the Solstice scheduling algorithm by introducing an indirection technique that reduce the number of configuration needed for the scheduling. This proposed technique aims to redirect part of demand traffics to the existing underutilized scheduled links. In this way, those emands are indirectly served by pairs of existing links and won't require additional configuration time. We plan to evaluate this technique by both simulations and comparisons with integer linear programming (ILP) formulations.

The rest of this proposal is organized as follows. Section 2 discusses Solstice and other related work. Section 3 describes the proposed indirection technique. Finally, Section 4 describes the proposed evaluation plan and goals.

2 Related Work

Recently, researchers proposes hybrid datacenter network architectures that offer higher throughputs at lower cost by combining switching technologies. In particular, recent proposals suggest employing highspeed optical [1, 2, 9] or wireless [4, 5, 10] networks configured to service the heavy flows, while passing the remainder of the traffic through a traditional, relatively underprovisioned packet-switched network.

As technology trends usher in dramatically faster reconfiguration times, the distinction between packet and circuit is blurred, and ever smaller flows can take advantage of a hybrid fabric. This trend will soon allow servicing the bulk of the traffic through a rapidly reconfigurable optical switch [8], leaving a relatively minor portion to be serviced by the packet network [6]. While the potential cost savings that hybrid technologies could realize is large, the design space for scheduling resources in the hybrid regime is not yet well understood. What range of traffic demands can be scheduled for a given switch design and how can this schedule be computed efficiently? These questions are not addressed by the existing switch scheduling literature.

Many of the classical approaches to scheduling for switches with non-trivial reconfiguration delays divide the offered demand into two parts: an initial, heavy-weight component that is served by O(N) highly utilized configurations with significant durations, and a second, residual component that is serviced by a similar number of short, under-utilized schedules. The recent Solstice scheduling algorithm [7] exploits the skewed nature of datacenter traffic patterns to create a small number of configurations with long durations that minimize the penalty for reconfiguration and leaves only a small amount of residual demand to be serviced by a low-speed (and lowcost) unconstrained packet switch.

3 The Indirection Technique

In this proposoal, we propose to improve the Solstice scheduling algorithm by introducing a indirection technique that reduce the number of configuration needed for the scheduling. Originally, Solstice will schedule each traffic demand by a direct path, which means that a new configuration is required to fulfill a demand that has different sources or destinations. Our technique proposes to schedule part of demand traffics indirectly. For example, port a has 5 demand to port b and 20 demand to port c; port c has 20 demand to port b. Solstice will require 3 configurations for these demands. However, we could let port a send all 25 demand to port c and let port c send 25 demand to port b. In this way we reduce one configuration time, which is nontrivial for optical switches.

4 Evaluation Plan and Goals

First, we want to evaluate our indirection technique by integer linear programming (ILP) formulations. In the paper, Liu *et al.* [7] provided a ILP formulation of the scheduling problem without indirection considerations. We plan to build a formulation based on Liu's formulation together with indirection consideration. Then we plan to use Gurobi [3] ILP solver to solve scheduling problems with the two formulations in order to find the optimal improvement by indirection. This will be our 75% goal.

Then we plan to evaluate our indirection techinque by simulations. We received the simulator of the Solstice algorithm from the authors. We plan to introduce our techinque into the simulator and compare it with the original algorithm. The simulation results will be comparable in terms of number of configurations, and the total time to serve all demands. This will be our 100% goal. Furthermore, we may deploy our heuristic into a real optical switch and evaluate it with real world workload. This will be our 125% goal.

References

[1] K. Chen, A. Singla, A. Singh, K. Ramachandran, L. Xu, Y. Zhang, X. Wen, and Y. Chen. OSA: An optical switching architecture for data center networks with unprecedented flexibility. In *USENIX NSDI*, 2012.

[2] N. Farrington, G. Porter, S. Radhakrishnan, H. H. Bazzaz, V. Subramanya, Y. Fainman, G. Papen, and A. Vahdat. Helios: a hybrid electrical/optical switch architecture for modular data centers. In *SIGCOMM 2010*.

[3] Gurobi. Gurobi Optimization. `http://www.gurobi.com/`.

[4] D. Halperin, S. Kandula, J. Padhye, P. Bahl, and D. Wetherall. Augmenting data center networks with multi-gigabit wireless links. In *SIGCOMM 2011*.

[5] S. Kandula, J. Padhye, and P. Bahl. Flyways to de-congest data center networks. In *HotNets 2009*.

[6] H. Liu, F. Lu, A. Forencich, R. Kapoor, M. Tewari, G. M. Voelker, G. Papen, A. C. Snoeren, and G. Porter. Circuit Switching Under the Radar with REACToR. In *USENIX NSDI*, 2014.

[7] H. Liu, M. K. Mukerjee, C. Li, N. Feltman, G. Papen, S. Savage, S. Seshan, G. M. Voelker, D. G. Andersen, M. Kaminsky, G. Porter, and A. C. Snoeren. Scheduling Techniques for Hybrid Circuit/Packet Networks. In *CoNEXT*, 2015.

[8] G. Porter, R. Strong, N. Farrington, A. Forencich, P. Chen-Sun, T. Rosing, Y. Fainman, G. Papen, and A. Vahdat. Integrating microsecond circuit switching into the data center. In *SIGCOMM 2013*.

[9] G. Wang, D. G. Andersen, M. Kaminsky, K. Papagiannaki, T. Ng, M. Kozuch, and M. Ryan. c-Through: Part-time optics in data centers. In *SIGCOMM 2010*.

[10] X. Zhou, Z. Zhang, Y. Zhu, Y. Li, S. Kumar, A. Vahdat, B. Y. Zhao, and H. Zheng. Mirror mirror on the ceiling: flexible wireless links for data centers. In *SIGCOMM 2012*.