C1-1 MORDIA – Shifting Data Center Optical Circuit Switching from Milliseconds to Microseconds

Nathan Farrington[1], George Porter[1], Pang-Chen Sun[1], Alex Forencich[1], Joseph Ford[1], Yeshaiahu Fainman[1], George Papen[1], Amin Vahdat[1,2]

[1]UC San Diego, [2]Google, Inc.

Thrust 1, Working Group 1 – Data Centers

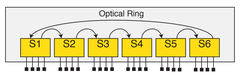
Statement of the Problem

Data center networks are straining to keep up with the explosive growth of network traffic needed to support cloud computing. Each web browser request is replicated over 1,000 times within the data center. It is possible, but prohibitively expensive to build a fully non-blocking electrically packet switched (EPS) data center network to handle the load. *What is needed is a cheaper source of bandwidth.* Optical circuit switching (OCS) promises to deliver faster, more scalable data center networks.

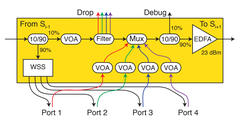
Technical Approach

We have built a fast 24-port OCS prototype using 2D MEMS switches from Texas Instruments. We are evaluating this OCS in a CIAN datacenter testbed with real workloads.

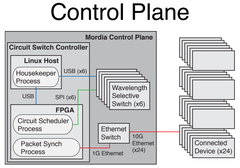
Data Plane

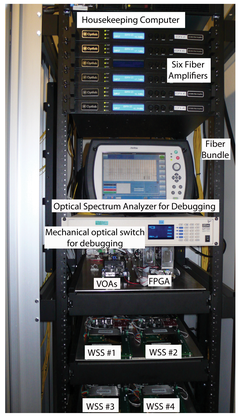


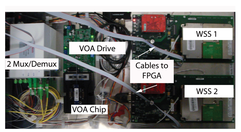
Station in Data Plane



Control Plane







Related Work & State of the Art

[1] Farrington, N., Porter, G., Radhakrishnan, S., Bazzaz, H.H., Subramanya, V., Fainman, Y., Papen, G., and Vahdat, A. Helios: A Hybrid Electrical/Optical Switch Architecture for Modular Data Centers. In *SIGCOMM '10*.

[2] Wang, G., Andersen, D.G., Kaminsky, M., Papagiannaki, K., Ng, T.S.E., Kozuch, M., and Ryan, M. c-Through: Part-time Optics in Data Centers. In *SIGCOMM '10*.

[3] Chen, K., Singla, A., Singh, A., Ramachandran, K., Xu, L., Zhang, Y., Wen, X., and Chen, Y. OSA: An Optical Switching Architecture for Data Center Networks with Unprecedented Flexibility. In *NSDI '12*.

[4] Farrington, N., Fainman, Y., Liu, H., Papen, G., and Vahdat, A. Hardware Requirements for Optical Circuit Switched Data Center Networks. In *Optical Fiber Conference (OFC/NFOEC) '11*.

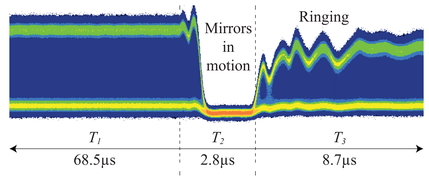
Our work with Helios [1] along with c-Through [2] introduced the concept of hybrid EPS/OCS data center networks. OSA [3] extended this concept with additional degrees of freedom. We analyze the Helios architecture in [4].

Mordia is 3 orders of magnitude faster than Helios, meaning that it is useful for a wider range of data center workloads.

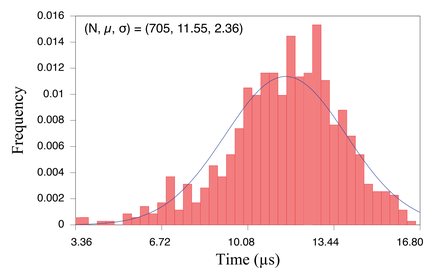
	Helios	Mordia	Speedup
T_1	5 ms	68.5 μs	73x
T_2	12 ms	2.8 μs	4,286x
T_3	15 ms	8.7 μs	1,724x

Results

In [4], we define programming time (T_1), switching time (T_2), and receiver electronics initialization time (T_3).

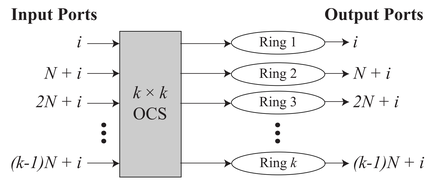


Measurements from an oscilloscope (top) and software running on hosts (bottom) clearly show these periods. Application-level experiments show the proper operation of the OCS prototype.



Future Plans

We have designed a 704-port OCS using an 8×8 *ring selection* OCS and are seeking to prototype it.



We are building out a more interactive demo to highlight the advantages of data center OCS networks (Aug 2012). We are developing a real-time NIC (network interface controller) to isolate the host software from the OCS and provide good TCP performance (Sep 2012). We are integrating a fast electrical circuit switch (ECS) into the testbed as another point of comparison (2013).

The authors would like to acknowledge support from the National Science Foundation through CIAN NSF ERC under grant #EEC-0812072.