**Derecho: Blindingly Fast RDMA Replication for Cloud and Edge Services.**

**Ken Birman, Matthew Milano and Edward Tremel**

Cloud computing services often replicate data, and consistency can matter too (particularly for IoT applications, where rapidly-changing data is used to coordinate distributed actions). This tutorial centers on Derecho, a library to assist with solving these problems [1, 2]. Derecho is far faster than prior solutions, including the weakly-consistent mechanisms that dominate today’s cloud computing infrastructures.

Derecho’s asynchronous design is key to its performance: the system achieves lock-free data streaming for critical paths (queries and updates occur concurrently but independently), and data movement is out-of-band from the control plane. The biggest payoff is with RDMA, but Derecho outperforms even on TCP.

Part of our goal in the tutorial is to understand the “design pattern” that enables this big performance boost; it involves separating the control and data planes, performing updates on a code path distinct from that used for queries, and transforming the system to use an asynchronous streaming communication model, in which participants often can independently deduce safety properties and perform batches of actions on received data. On 100Gbps RDMA, Derecho can send millions of events per second in each subgroup or shard, and throughput peaks at almost 16GB/s (125Gbps).

We anticipate a few use cases for the technology. The big one is the Edge IoT: Derecho was motivated by the desire to move machine learning tools closer to the IoT edge (the first tier of the cloud, which interacts directly with IoT devices). Existing edge platforms lack support for building complex, structured applications that maintain state and offer coordinated, consistent behavior; Derecho respond to this need. This said, we also see important opportunities in the operating system itself (existing systems have many services that could be improved using Derecho. Some low-hanging fruit includes Zookeeper, pub/sub and DDS technologies, file systems like HDFS or Ceph, and the network infrastructure layer.

The tutorial is structured into two parts. In the first half, based primarily on our TOCS paper [1], situates Derecho relative to prior work, explains how it obtains such high speeds, and reviews experiments that clarify why it outperforms prior solutions. We also learn about features aimed at Edge IoT applications that work with real-time sensor data yet still need strong consistency. The second part teaches attendees to use Derecho: we walk through the process of downloading the library, building services that use it, running them locally, and installing them on a cloud. At the end of the second part, participants modify the Derecho demo (a simple AI Sys application), rebuild it and rerun it.

SOSP participants who only want to hear about the work but not to really learn the system are welcome to drop in just at the beginning, and then can duck out when we get to the hands-on stage.

**Derecho download site:** [**http://GitHub.com/Derecho-Project**](http://GitHub.com/Derecho-Project)**.** The code base is open source.

[1] [Derecho: Fast State Machine Replication for Cloud Services](http://www.cs.cornell.edu/ken/derecho-tocs.pdf). S Jha, J Behrens, T Gkountouvas, M Milano, W Song, E Tremel, R Van Renesse, S Zink, and K Birman. ACM TOCS, 36 :2, April 2019.  
[2] [RDMC: A Reliable Multicast for Large Objects](http://www.cs.cornell.edu/projects/Quicksilver/public_pdfs/RDMC.pdf). Jonathan Behrens, Sagar Jha, Ken Birman, Edward Tremel. IEEE DSN ’18, Luxembourg, June 2018.