AT&T Research – March 2015

Title: **Network-aware Data Management Middleware for High Throughput Flows**

Abstract:

Today's data centric environment both in industry and scientific domains depends on the underlying network infrastructure and its performance, to deliver distributed application workloads. As current technology enables faster storage devices and larger interconnect bandwidth, there is a substantial need for novel system design and middleware architecture to address increasing latency, scalability, and throughput requirements. In this talk, I will outline network-aware data management and present solutions based on my past experience in large-scale data migration between remote repositories. I will first introduce a flexible network reservation algorithm for on-demand bandwidth guaranteed virtual circuit services. Flexible reservations find best path in a time-dependent dynamic network topology to support predictable application performance. I will then present a scheduling model with advance provisioning, in which data movement operations are defined with earliest start and latest completion times. The online scheduling of multiple requests relies on a simple but very efficient optimization technique to generate near optimum results on-the-fly, while satisfying users' time and resource constraints.

Within this scope, I will describe my experience in the initial evaluation of 100Gbps network as a part of the Advance Network Initiative project. We needed intense fine-tuning, both in network and application layers, to take advantage of the higher network capacity. End-system bottlenecks and system performance play an important role especially in many-core platforms. I will introduce a special data movement prototype, successfully tested in one of the first 100Gbps demonstrations, in which applications map memory blocks for remote data, in contrast to the send/receive semantics. This prototype was used to stream climate data over wide-area for in-memory application processing and visualization. I will conclude my talk with a brief overview of my other related projects on runtime adaptive tuning and failure recovery of data transfer jobs, and I/O aggregation and performance evaluation of distributed storage for collaborative science.

Bio:

Mehmet Balman is a senior Performance Engineer at VMware Inc. working on hyper-converged storage, virtual volumes, and virtualized solutions. He worked as a research engineer in the Computational Research Division at Lawrence Berkeley National Laboratory (Berkeley Lab). His work focused on data-intensive distributed computing, performance problems in high-bandwidth networks, advance resource provisioning, and data scheduling for large-scale applications. He was also deeply involved in the initial phase of DOE's nationwide 100Gbps network. He has been in the Earth System Grid Federation (ESGF) team for data dissemination in climate research. Before coming to Berkeley Lab, he was in the Center for Computation & Technology (CCT) at Louisiana State University (LSU). At CCT, he worked in various interesting projects, including Stork data scheduler, CyberTools project, and PetaShare state-wide distributed storage. He received his doctoral degree in computer science in 2010 from LSU. He also holds M.S. and B.S. degrees in computer engineering from Bogazici University, Turkey. For the last four years, Mehmet has been co-chairing a workshop series on Network-aware Data Management (NDM) in the SuperComputing (SC) conference.