The Federated Data Furnace (FDF) – A Platform for Scalable Urban Computing

Kamin Whitehouse, University of Virginia
Jie Liu, Microsoft Research

The explosion of Internet connected sensors is changing the computing landscape by generating enormous quantities of data at the network edge. Applications that use this data often have stringent latency requirements because they involve real-time automation or human interaction, creating increased demand for edge computing infrastructure. However, moving infrastructure to the network edge is not only a topological migration among routers, it is also a geographic migration: computers must move from data centers to population centers, i.e. out of remote, rural locations and into the world's cities. But how will this infrastructure be built? Urban areas pose many new challenges for scaling up computing infrastructure, e.g.: space is difficult to find, electricity prices are higher, and reliable power is more difficult to achieve. Geographically distributing the computing infrastructure in small units throughout a city is even more difficult because the compute elements also require infrastructure for reliable power, high-bandwidth networking, and cooling. These issues can be called *urban computing*: the physical issues that arise when edge computing meets the urban environment.

It is time to re-envision computing infrastructure in the context of the urban environment, including its unique needs, challenges, and opportunities. We are currently developing new architectures based on what we call a *data furnace* (DF) to enable scalable urban computing infrastructure at lower cost, with less pollution and waste, and with less stress on the nation's power infrastructure than conventional approaches. The basic approach is to put servers into homes, restaurants, offices, and industrial buildings to serve as a secondary heat source for water heating, industrial heating needs, and space heating. The DF saves energy by piggybacking on the energy that would have been used for heating anyway. It also dramatically reduces cost by piggybacking on the real estate, brick and mortar construction, power infrastructure, and air/water circulation that is already committed to heating anyway. Our preliminary results indicate that DFs can reduce energy usage by up to 40% and reduce total cost of ownership by up to 30-62% [14], depending on the climate zone. Most importantly, the geo-spatial distribution and natural co-location of DFs with population centers make them ideally suited for edge computing demands.

Since the authors and their colleagues originally proposed the data furnace computing model [14], several startup companies, including Nerdalize, CloudAndHeat, ARK labs, and Cline Property Management, have designed and sold data furnaces with various form factors and capacities, and have demonstrated the concept's viability in terms of security, maintenance costs, and resource availability. However, several systems challenges must still be solved before DFs can become a scalable solution for urban computing. Specifically, the cost and energy efficiency of the data furnace changes over time depending on its heating load: it is much more efficient to service workloads on a data furnace that is heating a room or water tank than on one that is not. If computing jobs are serviced by DFs that are venting heat directly outside, other DFs in the same region may need to run in a spin loop to produce enough heat, thereby wasting energy. Therefore, we are developing new techniques to combine multiple DFs into *federated data furnaces* (FDFs) that manage storage, network and computing resources to service future workloads based on unknown future efficiencies. This requires anticipatory techniques for job scheduling and data migration to satisfy quality of service (QoS) constraints with lower cost and energy usage than a group of independent DFs.

[1] The Data Furnace: Heating Up with Cloud Computing. In The 3rd USENIX Workshop on Hot Topics in Cloud Computing, June 2011.