

Intelligent Data Placement in a Home Environment

Brandon Salmon, CMU (bsalmon@ece.cmu.edu)

Greg Ganger, CMU

The recent increase in the power and connectivity of consumer devices has set the stage for a rapid growth in the number of these devices found within the average home. Television, music, photo and video are all transitioning to digital formats that can be stored on and accessed from these devices. These devices are very heterogeneous in terms of mobility, power, and storage capacity.

Current technology requires users to specify where data is placed, how many replicas are created, and how these replicas are kept consistent. This will become time consuming and unsustainable as the number of devices and amount of data increases. The ubiquity of small, limited capacity devices, continued increases in user demand for data, and consistency necessities will make replication on all devices infeasible.

As a first entry into the problem, I plan to explore the use of content metadata to make intelligent predictions about data placement. In a home environment, much of the data contains system readable metadata. For example, music files contain information about the artist and genre, while word documents contain a large number of searchable tags. This information can be used in combination with observed access patterns from devices to decide where data should be replicated.

While this approach has similarities to hoarding, it is unique in that it uses information about data and device access patterns to match data to a device, rather than using access proximity relationships between pieces of data to predict the next data to be accessed.

In combination with these predictions, I plan to explore the use of information about the amount of time devices are accessible from one another to make smart placement decisions. For example, this could allow a smaller device (such as a music player) to use a frequently accessible larger device (such as a laptop) as a backup cache.