A Quantitative Approach for Guiding Data Management on Complex Memory Machines

Brandon Kammerdiener and Michael R. Jantz
University of Tennessee, Knoxville
(423) 863-0807, [bkammerd,mrjantz]@vols.utk.edu

1. Introduction

Proliferation of real-time and AI-driven decision making continues to fuel the need for ever faster access to larger sets of data in memory. At the same time, increasing demands for high-density sharing are leading to more supercomputing configurations with large amounts of memory attached to each node and connected through efficient networking resources. New media technologies, such as high bandwidth memories, phase change memories, spin-torque transfer RAMs, and many others, and new processing-memory interconnect options, including the Compute Express Link (CXL), are bringing rich opportunities for addressing the diverse needs of applications under various cost, performance, and power constraints. In response to these trends, many supercomputing systems now include a heterogenous mix of memory devices and organizations, which can enable the combined benefits of their unique capabilities and support such diverse and multi-tenant workloads in modern computing centers.

Despite their potential benefits, heterogeneous memory architectures present new challenges for data management. Computing systems have traditionally viewed memory as a single homogeneous address space, sometimes divided into different non-uniform memory access (NUMA) domains but consisting entirely of the same storage medium (i.e., DDR* SDRAM). To utilize heterogeneous resources efficiently, alternative strategies are needed to match data to the appropriate technology in consideration of hardware capabilities, application usage, and in some cases, NUMA domain.

Spurred by this problem, the architecture and systems communities have proposed a range of hardware and software techniques to manage data efficiently on heterogeneous memory systems. For example, some systems choose to utilize high performance memories as a large, hardware-managed cache. Such configurations are not only inflexible, but they are also often inefficient because they require storage for very long tag values and additional bandwidth to distinguish cache blocks at DRAM scale.

Alternatively, many systems employ a software-based approach, where the operating system (OS), or the OS in collaboration with the applications, assign and move data into different memory and storage devices. Many modern platforms also provide system-level interfaces and custom allocators that allow the applications themselves to control the placement of data across the memory hierarchy. While software-based controls increase flexibility, they are still limited because they either proceed with little or no knowledge of how the applications intend to use memory resources or they require developers with detailed knowledge of complex memory resources and the capability and resources to update existing applications. These constraints are particularly problematic for scientific and high-performance computing (HPC) applications due to the frequency and scale of data usage as well as the need for performance portability in the face of a rapidly changing architectural landscape.

2. Our Approach

Building upon our own prior work in heterogeneous memory management [1 – 6], we propose to address these challenges through a quantitative approach that leverages detailed profiling and analysis of application behavior to steer data management of complex memory platforms. Our proposed approach is fully automatic and enables guided data tiering of individual program data objects, without requiring application modification or even recompilation. Figure 1 depicts the main components of our approach, which primarily consists of two new pieces of user-level software:

1) Custom runtime facilities, which enable applications to invoke the monitoring and management capabilities of our framework without updates to program source code, and
2) A systemwide monitoring and data management process, which runs alongside the applications and conducts object tiering through a series of complementary activities, including (1) monitoring and structuring profiles of object allocation and usage, (2) automated heuristics to prioritize objects for placement in fast memory, and (3) mechanisms to enforce tier recommendations when a particular event occurs.

By providing automated monitoring and analysis along with customizable controls, this approach will enable applications to define custom quantitative strategies for guiding data tiering on heterogeneous memory machines. Moreover, it provides these capabilities by leveraging standard tools and system interfaces and can therefore be deployed in real and diverse supercomputing environments, with minimal effort. Thus, we believe that this approach has significant potential to address the challenges presented by complex memory hierarchies and will enable more efficient execution for a wide range of supercomputing applications.

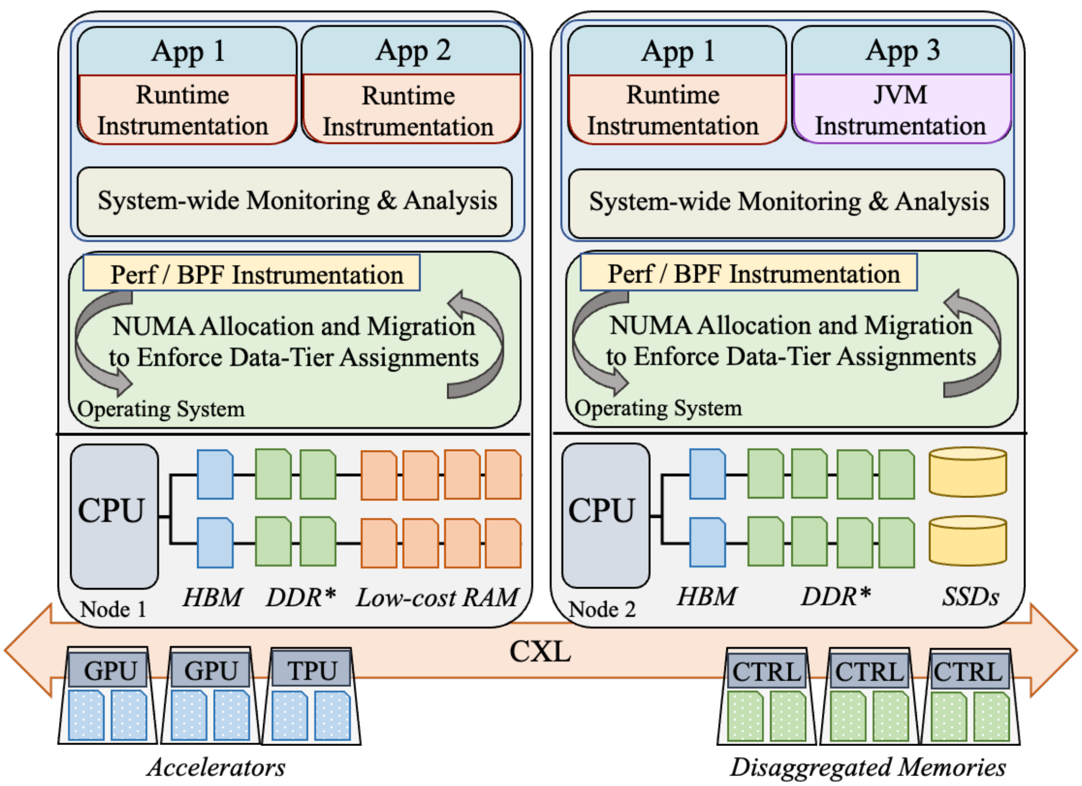

Figure 1. Design overview of our approach

Table 1: Complete author list and institutional affiliations

Position Paper Author List			**Institution**
Last Name	First Name	Title	Institution Name
Kammerdiener	Brandon	Graduate Research Assistant	University of Tennessee
Jantz	Michael R.	Associate Professor	University of Tennessee

Optional References Here

1. Effler, T. C., Howard, A. P., Zhou, T., Jantz, M. R., Doshi, K. A., and Kulkarni., P. A. On automated feedback-driven data placement in hybrid memories. In LNCS International Conference on Architecture of Computing Systems (April 2018).
2. Effler, T. C., Jantz, M. R., and Jones, T. Performance potential of mixed data management modes for heterogeneous memory systems. In 2020 IEEE/ACM Workshop on Memory Centric High Performance Computing (MCHPC) (Washington, DC, USA, 2020), pp. 10–16.
3. Effler, T. C., Kammerdiener, B., Jantz, M. R., Sengupta, S., Kulkarni, P. A., Doshi, K. A., and Jones, T. Evaluating the effectiveness of program data features for guiding memory management. In The International Symposium on Memory Systems (2019), MEMSYS '19, p. 383–395.
4. Olson, M., Kammerdiener, B., Jantz, M., Doshi, K., and Jones, T. Portable application guidance for complex memory systems. In The International Symposium on Memory Systems (09 2019), MEMSYS '19, ACM, pp. 156–166.
5. Olson, M. B., Kammerdiener, B., Doshi, K. A., Jones, T., and Jantz, M. R. Online application guidance for heterogeneous memory systems. To appear in the ACM Transactions on Architecture and Code Optimization (May 2022).
6. Olson, M. B., Zhou, T., Jantz, M. R., Doshi, K. A., Lopez, M. G., and Hernandez, O. R. Membrain: Automated application guidance for hybrid memory systems. In IEEE Intl. Conf. on Networking, Arch. and Storage, NAS, October 11-14, 2018. Best Paper Award. (2018), pp. 1–10.