Carl Waldspurger, Research and Development, CloudPhysics, Inc.

Irfan Ahmad, CTO, CloudPhysics, Inc.

TITLE

Practical Online Cache Analysis and Optimization

ABSTRACT

The benefits of storage caches are notoriously difficult to model and control, varying widely by workload, and exhibiting complex, nonlinear behaviors. However, recent advances make it possible to analyze and optimize high-performance storage caches using lightweight, continuously-updated miss ratio curves (MRCs). Previously relegated to offline modeling, MRCs can now be computed so inexpensively that they are practical for dynamic, online cache management, even in the most demanding environments.

After reviewing the history and evolution of MRC algorithms, we will examine new opportunities afforded by recent techniques. MRCs capture valuable information about locality that can be leveraged to guide efficient cache sizing, allocation, and partitioning, in order to support diverse goals such as improving performance, isolation, and quality of service. We will also describe how multiple MRCs can be used to track different alternatives at various timescales, enabling online tuning of cache parameters and policies.

BIOGRAPHIES

Carl Waldspurger has been leading research at CloudPhysics since its inception. He is active in the systems research community, and serves as a technical advisor to several startups. For over a decade, Carl was responsible for core resource management and virtualization technologies at VMware. Prior to VMware, he was a researcher at the DEC Systems Research Center. Carl holds a PhD in computer science from MIT.

Irfan Ahmad is the Chief Technology Officer of CloudPhysics which he cofounded in 2011. Prior to CloudPhysics, Irfan was at VMware, where he was R&D tech lead for the DRS team and co-inventor for flagship products, including Storage DRS and Storage I/O Control. Irfan worked extensively on interdisciplinary endeavors in memory, storage, CPU, and distributed resource management, and developed a special interest in research at the intersection of systems. Irfan also spent several years in performance analysis and optimization, both in systems software and OS kernels. Before VMware, Irfan worked on a software microprocessor at Transmeta.