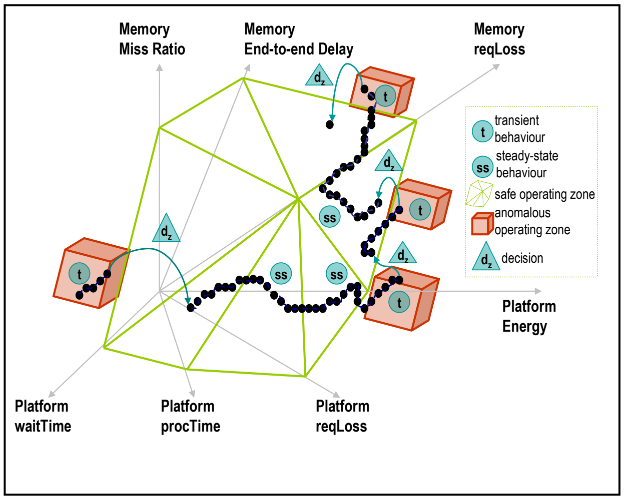
Autonomic Performance/Power Optimizations in Next Generation Datacenters

by Salim Hariri

The goal of this project is to design innovative autonomic framework and architecture to integrate in traditional server platforms and enclosures to intelligently optimize their performance/watt. To achieve this goal, we plan to extend our performance/watt optimization at the resource-level to the platform-, enclosure- and datacenter-levels. The central ideas behind this research, which includes power and thermal optimizations are: (i) to proactively detect and reduce resource over-provisioning in server platforms such that it is just right-sized to handle the requirements of the application; and (ii) migrate virtual machines from one physical server in an enclosure to another server to eliminate thermal hot spots and smooth thermal gradients within the enclosure, reducing cooling costs. This holistic multi-variable mathematically rigorous optimization approach for determining optimal performance/watt lends itself best to achieve the desired goals.

Most early work on server power management has either focused on specific components such as the processor or used heuristics to address base power consumption in server clusters or have ignored thermal ramifications completely. This motivated us to adopt a holistic approach for system-level power management within a server farm or enclosure where we exploit interactions and dependencies among different resources and platforms.

We consider an enclosure with multiple servers – each consists of multi-core processors and multi-rank memory subsystems plus other resources. The Autonomic Enclosure consists of three hierarchies of management, as shown the first Figure – the Enclosure Autonomic Manager (EAM) at the enclosure level; the Platform Autonomic Manager (PAM) at the platform-level; and Core Manager (CM) and Rank Manager (RM) at the individual processor core and memory rank, respectively. EAM ensures that all platforms within the enclosure operate within the pre-determined thermal gradient and thermal/power envelope by migrating virtual machines running specific workloads from one platform to another. PAM's objective is to ensure that platform resources (processor/memory) are configured to meet the dynamic application resource requirements such that additional platform capacity can be transitioned to low-power states. In this manner both the EAM and PAM save total power without hurting application performance. The platform power and performance parameters together determine the platform operating point in an n-dimensional space at any instant of time during the lifetime of the application. PAM manages the platform power and performance by maintaining the platform operating point within a predetermined safe operating zone, as the example shown in the second Figure. PAM predicts the trajectory of the operating point as it changes in response to changes in the nature and arrival rate of the incoming workload and triggers a platform reconfiguration whenever the operating point drifts outside of the safe operating zone. See how we used this approach for memory performance/watt optimizations, where we specifically monitor additional parameters such as memory miss ratio, memory end-to-end delay and memory request loss to determine the best memory configuration that would maintain the platform response time within the safe operating region .



Platform state is defined by the number of processor cores in active state, the number of memory ranks in the active state and I/O devices. Since the performance of the platform state depends on the physical configuration of the platform, PAM platform reconfiguration decision actually involves a platform state transition from the current state to a target state that would maintain the performance while giving the smallest power consumption. The search for this ideal target state is formulated as an optimization problem.

EAM will rely on thermal sensors that will be placed in critical positions within the enclosure and platforms. Data from all these sensors will be collected by EAM, which will make decisions based on the thermal gradient and presence/absence of thermal hot spots within the enclosure, the set of resources required by the running workloads and the decisions PAM has undertaken. One interesting research challenge is the sensitivity to data inaccuracies when moving across the hierarchy from the component managers to the platform managers to the enclosure manager.

This project will develop innovative management techniques at the resource level to address the following research challenges: 1) How to efficiently and accurately model power and energy consumption from a system-level perspective that involves complex interactions of different classes of devices such as processor, memory, network and I/O? A system-level view of these components would present more opportunities for power savings since we can exploit the non-mutually exclusive behaviors of these components to set them at power states such that the global system power consumption is minimal; 2) How to predict, in real-time, the behavior of system resources and their power consumptions, as workloads change dynamically by several order of magnitude within a day or a week; and 3) How to design efficient and self-adjusting optimization mechanisms that can continuously learn, execute, monitor, and improve themselves in meeting the collective objectives of power, thermal and performance optimizations? Game theory and data mining techniques will be exploited to address this research challenge.