PERFUME: Power and Performance Guarantee with Fuzzy MIMO Control in Virtualized Servers

Palden Lama and Xiaobo Zhou
University of Colorado at Colorado Springs

Outline

• Background and Motivation
• Challenges
• Related Work
• Proposed Approach
• Performance Evaluation
• Conclusion
• Q & A
Data Center: Key Issues

- Multi-facet Challenges
  - Performance Assurance, Server Utilization, Power Consumption.

- Server utilization
  - built on the over-provisioning model.
  - dedicated servers for different applications.
  - most servers in a typical data center run at only 5-10 percent utilization

- Power consumption and Carbon footprint
  - According to the U.S. Department of Energy datacenters are the fastest-growing energy consumers in the United States today.
  - IEA (International Energy Agency) updated a warning in 5/2009 that information and communication technology energy use could double by 2022, and triple by 2030
  - Data centers are responsible for the tens of millions of metric tons of carbon dioxide emissions annually more than 5% of the total global emissions.

Virtualized Data Centers

- Virtualization
  - abstracts physical resources into virtual machines (VMs).
  - diverse OS and applications share underlying server resources.

- Consolidation
  - improves server utilization
  - reduces power consumption

- Platform for Cloud Computing
  - Flexible and Fine-grained Resource Allocation
  - On-demand, pay-per-use service
Power Management in Data Centers

- **Power over-subscription**
  - the sum of the possible peak power consumptions of all the servers combined is greater than the provisioned capacity
  - Power budgeting mechanism (DVS) on each server, to ensure total power stays below capacity.

- **Hardware power budgeting**
  - does not respect the isolation among virtual machines with different performance requirements.

- **Need for a holistic view of power and performance management in data centers.**

Joint Power and Performance Control

- **Power oriented vs. performance oriented**
  - Controlling either power or performance while achieving the other objective in best-effort manner.
  - No explicit co-ordination between power and performance.

- **Effect of workload dynamics (highly dynamic and bursty)**
  - Control accuracy
  - System stability

- **Percentile based performance metric**
  - Most previous works focus on average performance guarantee. (not suitable for interactive applications)
  - A percentile response time introduces much stronger nonlinearity to the system, making it difficult to derive an accurate performance model.
Challenges: Workload Dynamics

- Workload Variation at multiple time scales demands self-adaptive and robust techniques for power and performance management.

- System stability should be guaranteed to avoid oscillatory behavior in system states that result in poor power and performance assurance.

Challenges: Multi-tier architecture

- Cross-tier dependencies
- Bottleneck switching

- Performance is the result of a complex interaction of workloads in a very complex underlying computer system.

- Power usage of different tiers of one application may vary with workload.
RELATED WORK

  - MIMO control for cluster-level power control using DVFS.
  - Not applicable to virtualized servers
  - Power-oriented: no performance guarantee

  - Coordination of power controllers at various levels (Enclosure, Server & VM)
  - Power-oriented: no performance guarantee

RELATED WORK

  - Maps workloads to best suited platforms for power efficiency
  - Primary objective: meeting service level agreement of applications.
  - Lacks explicit control on power consumption.

  - Two-layer control architecture
    • primary control: VM resource allocation for balancing their relative perf. level.
    • secondary control: reducing power consumption by manipulating CPU frequency.
  - Power consumption is reduced in best-effort manner.
**RELATED WORK**

  - Co-ordinated two-level controller for power and performance control
  - May not adapt to workload changes.

  - Allows flexible tradeoff between power and performance objectives
  - Reduces performance relative deviation by 17% compared with two layer feedback controller (as in Co-Con).
  - Lacks the guarantee on stability and performance of the server system especially in the face of highly dynamic and bursty workloads.
  - Performance relative deviation may degrade in case of percentile-based performance metric.

**PERFUME System**

- Flexible tradeoffs
  - It guarantees both power and performance targets with user specified tradeoffs.

- Well-suited to virtualized environments
  - It enforces power budgeting by controlling CPU usage limits of VMs instead of throttling CPU frequency of physical server.

- Stability and control accuracy (Fuzzy MIMO Control)
  - FUMI applies Model Predictive control (MPC) technique to control CPU usage limits of various multi-tier applications hosted in virtualized servers.
  - To apply MPC technique, it generates fuzzy models that capture power and performance behavior of multi-tier applications hosted in virtualized servers.
  - It adapts the fuzzy models at run-time in response to changes in workload.

- It is able to control both average and percentile-based performance metric due to its Fuzzy modeling
PERFUME System Architecture

Testbed:
- HP ProLiant BL460C G6 blade server modules with a HP EVA storage area network.
- 10 Gbps Ethernet and 8 Gbps Fibre/iSCSI dual channels.
- Virtualized with Vmware ESX 4.1
- Hosting multi-tier application benchmark, RUBiS

FUMI Control Interface

- Fuzzy model:
  - represents an arbitrarily complex system by a combination of inter-linked subsystems with simple functional dependencies.
  - Accurately capture the non-linearity of computer systems: response time vs. resource allocation
- Optimizer:
  - Formulates MIMO control problem as a constrained optimization.

Minimize:
\[ V(k) = \sum_{i=1}^{n} (y_i(k)-y_i(k-1))^2 + \sum_{i=1}^{n} (y_i(k)-y_i(k-2))^2 + \sum_{j=1}^{m-1} (\Delta y_j(k))^2 + \sum_{j=1}^{m-1} (\Delta u_j(k))^2 \]
Constraint:
\[ \sum_{i=1}^{M} (\Delta u_i(k) + u_i(k)) \leq U_{max} \]
Fuzzy Modeling

- Each controlled variable (power, performance) is represented by a fuzzy model

\[ y(k + 1) = R(\xi(k), u(k)). \]

- Fuzzy model (R) is composed of a set of fuzzy rules.

\[ R_i: \text{If } \xi_1(k) \text{ is } \Omega_{i,1} \text{ and } \ldots \xi_p(k) \text{ is } \Omega_{i,p} \text{ and } u_1(k) \text{ is } \Omega_{i,p+1} \text{ and } \ldots u_m(k) \text{ is } \Omega_{i,m+p} \text{ then} \]

\[ y_i(k + 1) = \sum_{i=1}^{m+p} \alpha_i \xi_i(k) + \beta_i u_i(k) + \gamma_i. \]

- Model’s final output is sum of output given by each rule, weighted by its activation strength.

- Initial fuzzy model obtained by subtractive clustering and ANFIS (Artificial Neural Network Fuzzy Inference System) technique.

- At run time, wRLS method updates the fuzzy model parameters.

Control Solution

- Express MIMO Control Objective as Quadratic Programming Problem

\[
\min_{\Delta u} \left\{ \frac{1}{2} \Delta u^T \cdot H \cdot \Delta u + f^T \cdot \Delta u \right\}
\]

- Linearize fuzzy model at each sampling interval, to extract state space model

\[ \begin{align*}
\dot{x}_{\text{lin}}(k + 1) &= \bar{A}(k) x_{\text{lin}}(k) + \bar{B}(k) u(k) \\
y_{\text{lin}}(k) &= C(k) x_{\text{lin}}(k).
\end{align*} \]

- The matrices \( A(k), B(k) \) and \( C(k) \) are constructed by freezing the parameters of the fuzzy model at a certain operating point \( y(k) \) and \( u(k) \).

- Solve using any Quadratic solver software or MATLAB.
Prediction Accuracy

<table>
<thead>
<tr>
<th>throughput</th>
<th>Power</th>
<th>95th percentile response time</th>
</tr>
</thead>
<tbody>
<tr>
<td>NRMSE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.5%</td>
<td>15.2%</td>
<td>17.6%</td>
</tr>
</tbody>
</table>

Self-Adaptiveness of Power/Perf Model

- Browsing mix of 1000 users to bidding mix of 500 concurrent users & vice versa
- Comparison with ARMA
Flexible Tradeoffs

- Control accuracy for various tradeoffs.

System Stability

(a) A highly dynamic workload.

Percentile-based Response Time Guarantee

- Improvement of 40% in terms of relative deviation
Conclusion

• PERFUME provides holistic and self-adaptive performance and power control in a virtualized server cluster.
• Testbed implementation demonstrates
  – precise control of power consumption of virtualized blade servers
  – effective control of throughput and percentile-based response time of multi-tier applications.
  – flexible tradeoffs
  – control accuracy and system stability

ANY QUESTIONS?