Power and QoS Performance Characteristics of Virtualized Servers

Takayuki Imada, Mitsuhisa Sato and Hideaki Kimura

University of Tsukuba, Japan
Outline

• Motivation and Objective
• Power and QoS performance characteristics on virtualized servers
  - With one virtualized server node case
  - With a load migration method between two virtualized servers
  - With different types of allocated processor cores to VMs and running frequency
• Related work
• Conclusion
Motivation

• Increasing power consumption in datacenters
  ■ For server nodes, network equipments and cooling facilities
  ■ A problem to be solved as soon as possible
  ■ From 2000 to 2005, electricity use on servers became roughly doubled in the U.S. (Jonathan, 2007)

• Server resource management using Virtual Machines (VMs)
  ■ Merits
    ◆ VM migration and server consolidation
    ◆ Flexible provisioning
    ◆ Green IT!!
Energy reduction on virtualized servers

New topics should be considered in virtualized servers

- **Server consolidation**
  - How is its effect on performance and power?

- **Multiple VMs running on multi-core processors**
  - How should they be configured for power saving?
Objective

• To characterize power and QoS performance on virtualized servers for developing an energy saving scheme

① Baseline characteristics
  • Comparison of power consumption before/after server consolidation
  • Processor’s DVFS control effects in a virtualized server node

② Effects of workload migration
  • Different migration schemes and workload levels

③ How to allocate processor cores to VMs
  • Different # of processor cores allocated to VMs and running frequencies

- Consolidation phase
  Before
  Server1: VM0, VM1, VM2
  Server2: VM1, VM2

  After
  Server1: VM0, VM1, VM2

- VM management phase
  VM0, VM1, VM2
  Core: Server1, Server2
Baseline evaluation

- Consolidation phase
  Before
  Server1
  VM0
  VM1
  VM2
  After
  Server1
  VM0
  VM1
  VM2
  Here!!

- VM management phase
  Core
  Core
  Core
  Core
  VM0
  VM1
  VM2
Evaluation of power reduction on a virtualized server

• 1 or 2-VM running on a server node
  1-VM: Workload 0
  2-VM: Workload 0
  2-VM: Workload 1

  Two different frequencies: $f_{\text{min}} = 1.6\text{GHz}$, $f_{\text{max}} = 2.93\text{GHz}$

• Benchmark: SPECweb2005
  - Banking (B), E-commerce (E) and Support (S) workloads
  - Required QoS in SPECweb2005
    - Ratio of requested pages each of which is returned within a defined time (=TIME_GOOD) should be more than 95% (95% TIME_GOOD QoS)

• Workload sets
  - 1-VM evaluation: each B, E, S workload
  - 2-VM evaluation: B-E, S-E simultaneous workloads
Test-bed environment

- Load balancer
- Clients
- iSCSI Storage
- VMs: VM 0, VM 5
- Servers for VMs
- Requests
- Responses

Load Balancer: IPVS-1.2.1, ipvsadm (Layer-4 software load balancing)

Network for accesses
Network for management
Network for contents

Paravirtualized VM: 1GB allocated memory
kernel-xen-2.6.18-128.1.10.el5 (64-bit)
Apache 2.2.3

Server specification:
Intel Core i7 940 (2.93GHz, HT off)
DDR3-1333 12GB
kernel-xen-2.6.18-128.1.10.el5 (64-bit)
Xen 3.3.1
Power reduction by DVFS control

• Overview

\[
S(f_{\text{freq}=\text{fmin} \text{ or fmax}}): \quad \text{# of the maximum simultaneous sessions which can satisfy 95\% TIME\_GOOD QoS}
\]

• Power reduction @ A (light-weight) and B (middle-weight)
  
  1-VM: A = 100 simultaneous sessions
  B = S(f_{\text{fmin}}) on each workload
  
  2-VM: A = B-E1 or S-E1 workload set
  B = B-E2 or S-E2 workload set

• Results of power reduction: \( \{1 - \frac{P(\text{fmin})}{P(\text{fmax})}\} \times 100 \% \)
  
  1-VM evaluation
  ✓ 1% power reduction @ A
  ✓ 5% power reduction @ B

  2-VM evaluation
  ✓ 2% power reduction @ A
  ✓ 8% power reduction @ B

Sorry, please see our paper for detailed values
Power reduction by server consolidation

- (2-VM on 1-node) v.s. 2 * (1-VM on 1-node)
- B-E3 and S-E3 workload sets

![Bar chart showing power consumption comparison between 1-VM and 2-VM configurations on 1-node and 2-node setups.]

- Slight increase of power consumption
- 48% power reduction by server consolidation
Power and performance evaluation of load migration schemes

- Consolidation phase
  Before
  Server1
  VM0
  VM1
  VM2

  Server2

  VM0
  VM1
  VM2

- VM management phase
  VM0
  VM1
  VM2
  Core
  Core
  Core
  Core

Here!!

HPCS Lab.
High Performance Computing System Lab., Univ. of Tsukuba
Migration schemes

• Two different schemes
  - LM (Live Migration): Live migration method provided by Xen
  - Switching: VM 1 exists at the beginning of benchmark run (the workload on VM 0 is switched by a load balancer)
Evaluation overview

• Two different background load levels
  ■ Situation A
    * Banking workload migration to a no-loaded server
  ■ Situation B
    * Banking workload migration to a server which handles E-commerce workload

• Workload sets (with different workload levels of Banking)
  ■ Running @ fmin
    * $S_{bank} = 100$ or $500$, $S_{ecom} = 600$ (static)
  ■ Running @ fmax
    * $S_{bank} = 100$, $500$ or $1000$, $S_{ecom} = 1000$ (static)

• Load migration is started 30 seconds after the benchmark start (Benchmark duration: 120 seconds)
Results: QoS performance
(ratio of requested pages with TIME_GOOD response time)

- Running @ fmin
- Running @ fmax

LM scheme couldn’t achieve 95% TIME_GOOD QoS

No migration: 2 servers without load migration

A: B
Server 0
Server 1

B: B
Server 0
Server 1

HPCS Lab.
High Performance Computing System Lab., Univ. of Tsukuba
Results: Power consumption

- Relative power consumption compared to the No migration
  - Switching: almost same in all workload sets
  - LM: about 1-4% increase because of VM memory data transfer

Power profiles during Banking workload (LM, fmax)

Doubled migration time (LM scheme in Xen doesn’t move frequently modified memory pages during pre-copy phase)

Sbank = 500 (4% increase)  Sbank = 1000 (1% increase)
Discussion

• File size characteristics to be considered
  - LM scheme can be applicable to large files like movie or music files
  - Switching scheme is more suitable for small files (if lasting communications exist at the Switching call, the start of Switching scheme will be delayed because the communications should not be lost)

• Used resources
  - LM scheme needs only the minimum number of required VMs
  - Switching scheme needs redundant VMs (but turning on and off VMs can be saved if we can employ Suspend To RAM on a host server)
Evaluation with different processor-core allocation schemes

- Consolidation phase
  Before
  Server1
  VM0  VM1  VM2
  Server2

- VM management phase

Here!!

Consolidation phase
Before
After

VM0  VM1  VM2

VM0  VM1  VM2

- Server1

- Core

VM0  VM1  VM2

VM0  VM1  VM2

Core  Core  Core
VM management in a server

- For further power optimization, we should consider VM scheduling in each server node.
- **Factors:** the number of allocated processor cores and running frequency
  
  Ex.) A quad-core processor, $f_{min} = \frac{f_{max}}{2}$

<table>
<thead>
<tr>
<th>Core 0</th>
<th>Core 1</th>
<th>Core 2</th>
<th>Core 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>VM 0</td>
<td>VM 1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Available resource = 4 * $f_{max}$
- Required resource = 2 * $f_{max}$

Which is better?

2-core @ $f_{max}$
4-core @ $f_{min}$

Which is better?
Evaluation overview

- Configurations (1-VM and 2-VM)

<table>
<thead>
<tr>
<th>Workload set</th>
<th># of VMs</th>
<th>( # of processor cores, frequency)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sbank = 1000</td>
<td>1</td>
<td>(1, 2.93GHz) (2, 1.6GHz)</td>
</tr>
<tr>
<td>Secom = 1000</td>
<td>1</td>
<td>(1, 2.93GHz) (2, 1.6GHz)</td>
</tr>
<tr>
<td>Ssupp = 800</td>
<td>1</td>
<td>(1, 2.93GHz) (2, 1.6GHz)</td>
</tr>
<tr>
<td>Sbank = 1000, Secom = 1000</td>
<td>2</td>
<td>(1, 2.93GHz) (3, 2GHz)</td>
</tr>
<tr>
<td>Sbank = 1000, Ssupp = 800</td>
<td>2</td>
<td>(1, 2.93GHz) (3, 2GHz)</td>
</tr>
</tbody>
</table>
Results: 1-VM evaluation

- **QoS performance**
  - Slight performance degradation of the QoS performance
    - Relatively increased load by VM scheduling on an additional processor core and the decreased running frequency
  - About 10% power reduction in all workloads

- **Relative power**
Results: 2-VM evaluation

• QoS performance

- Large performance degradation (3 cores - 2GHz, Credit)
  - A default scheduler (Credit) in Xen (non preemptive, scheduling period = 30[msec]) couldn’t achieve real-time processing

- 4-9% power reduction by boosted processor frequency and a preemptive scheduler (SEDF)

• Relative power
Application to existing methods

- How to compromise multiple P-states transition calls from each different VM?
  - Independently different P-states on each core may not be allowed

Locate VP(j) regularly, the VP(j) can be realized with \( f(i) \) satisfies
\[
N \cdot f(i+1) < VP(j) \leq N \cdot f(i)
\]
and time slice management for VMs

Available maximum # of processor cores: \( N = 3 \)
Frequency of hardware P-state \( f(i) \) where \( i = 0,1,2,3 \) (\( f(0) > f(1) > f(2) > f(3) \))

P-state 0
- Core 0
- Core 1
- Core 2

P-state 1
- Core 0
- Core 1
- Core 2

P-state 2
- Core 0
- Core 1
- Core 2

P-state 3
- Core 0
- Core 1
- Core 2

Required
- VM 0
- VM 1

VP(2) can satisfy each given QoS requirement imposed on a corresponding VM.
Related work

• VirtualPower: Virtualized processor P-states for VMs [Nathuji et al., 2007]
  - realized by controlling of DVFS on a processor and CPU time slice
  - doesn’t consider a VM running on multiple processor cores

• A virtualized server reconfiguration scheme for energy reduction [Kusic et al., 2008]
  - Server resource management based on a queuing theory
  - Live migration of VMs is not considered
  - DVFS control for a VM across multiple processor cores is not considered
Conclusion

• Power and QoS performance characteristics on virtualized servers for developing an energy saving scheme
  - Drastic power reduction by server consolidation is possible
    - Slightly increased power consumption by an added VM on a server
    - DVFS control for a processor provides further power reduction
  - VM live migration needs considerations on power and performance
    - Live migration time can be varied with workload level for a VM to be moved
  - Further power reduction can be achieved by both multiple-core allocation to VMs and using lower frequency
    - Such processor management can be implemented by extending an existing method

• Future work
  - To develop an intelligent reconfiguration algorithm based on the obtained results and knowledge
Thanks!!