Power and QoS Performance Characteristics of Virtualized Servers

Takayuki Imada, Mitsuhisa Sato and Hideaki Kimura

University of Tsukuba, Japan

Outline

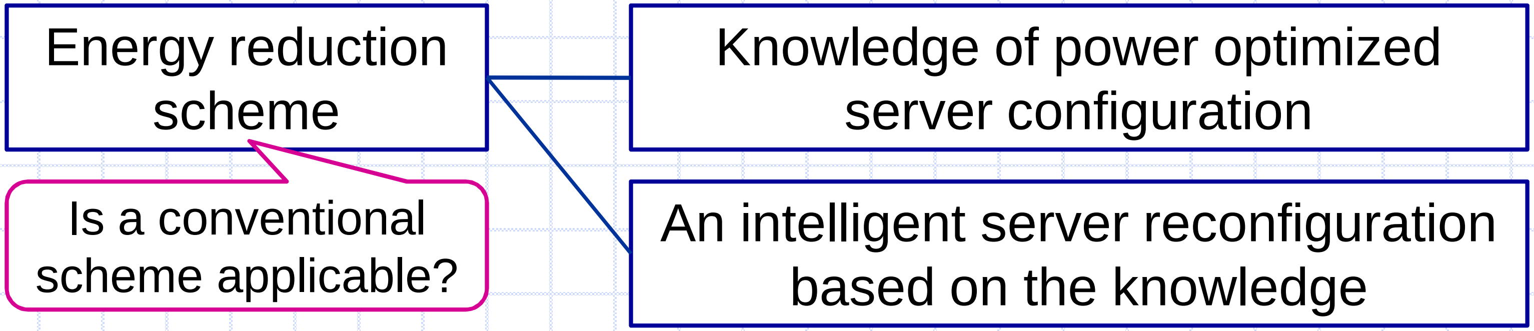
- Motivation and Objective
- Power and QoS performance characteristics on virtualized servers
  - With one virtualized server node case
  - With a load migration method between two virtualized servers
  - With different types of allocated processor cores to VMs and running frequency
- Related work
-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

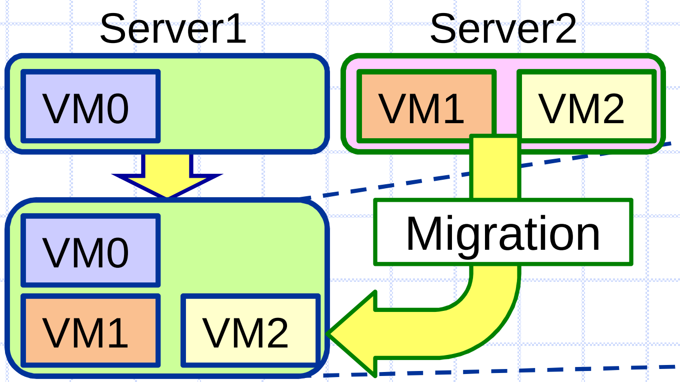
- Server resource management using Virtual Machines（VMs）
  - Merits
    - VM migration and server consolidation
    - Flexible provisioning
    - Green IT!!

Energy reduction on virtualized servers



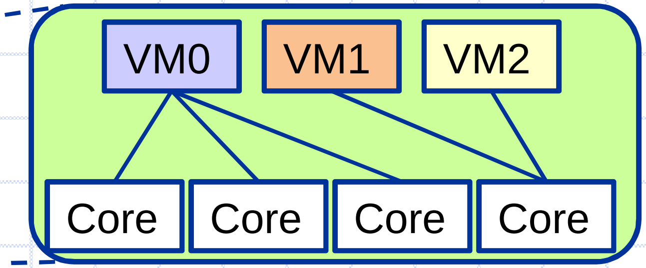
New topics should be considered in virtualized servers

- Server consolidation



  - How is its effect on performance and power?

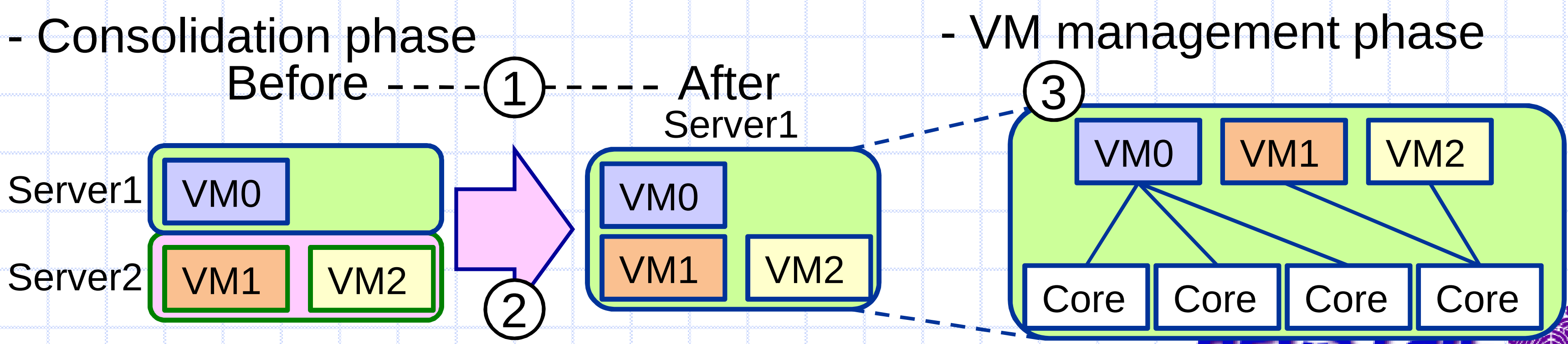
- Multiple VMs running on multi-core processors



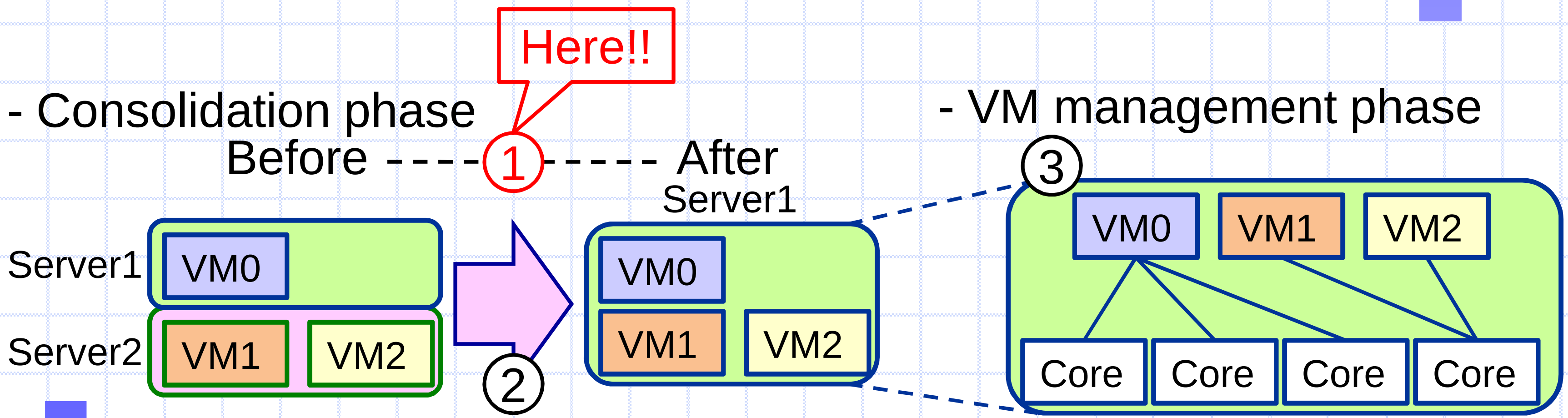
  - How should they be configured for power saving?

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

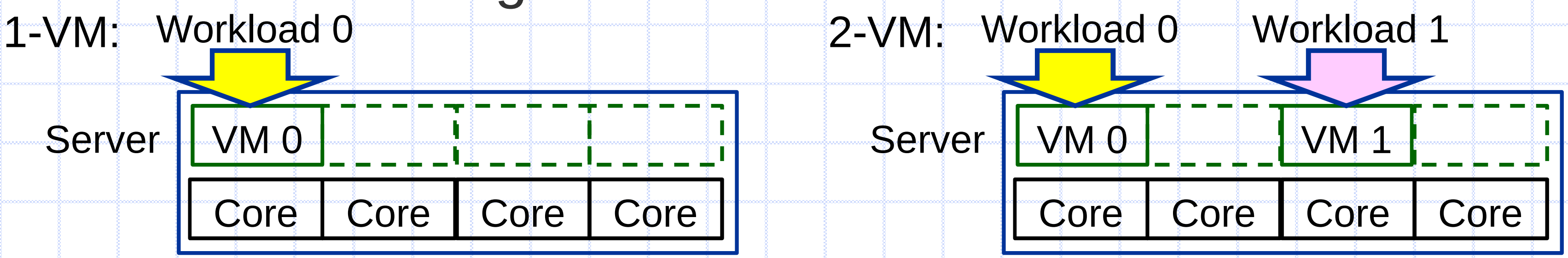


Baseline evaluation



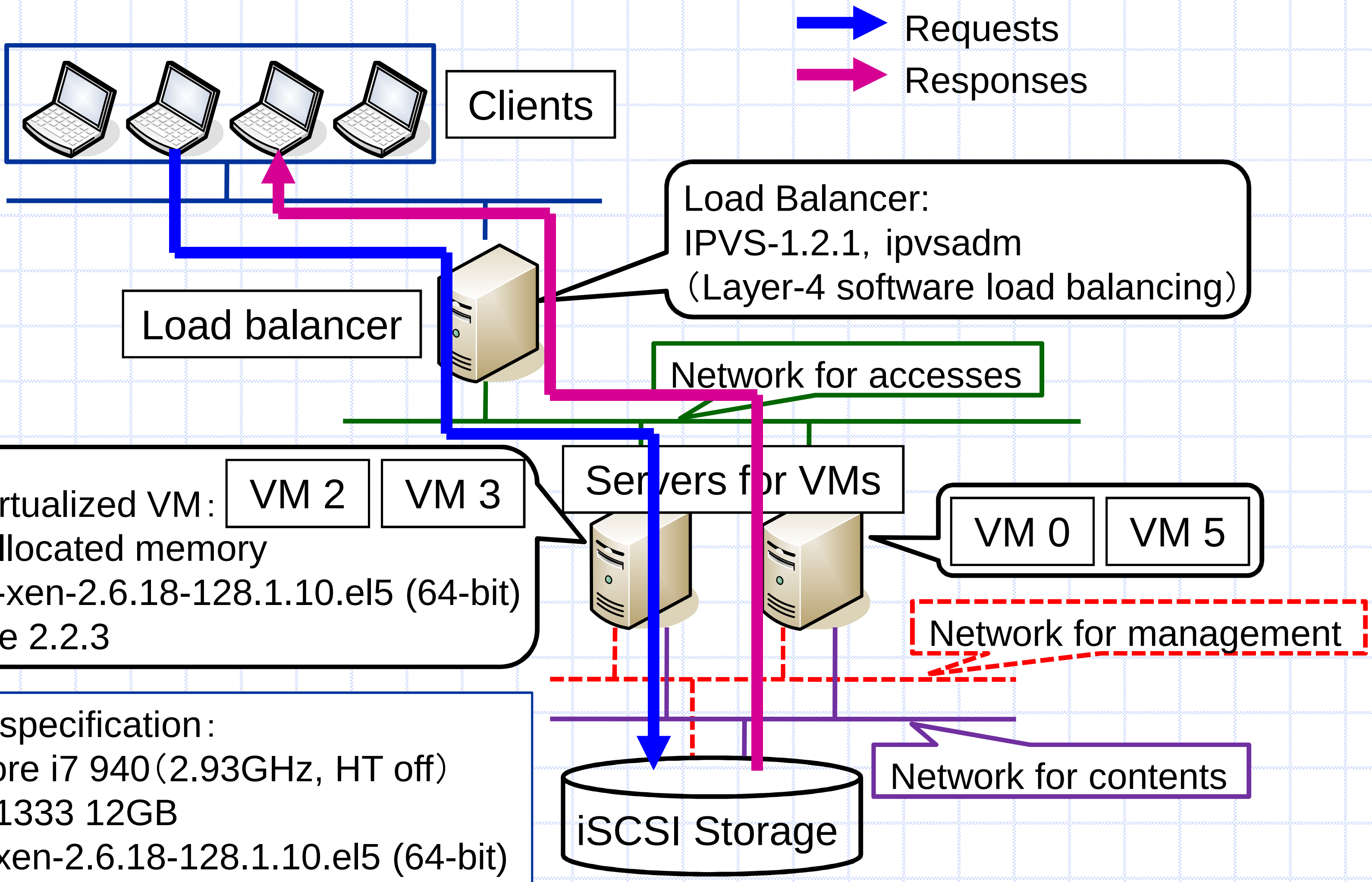
Evaluation of power reduction on a virtualized server

- 1 or 2-VM running on a server node



  - Two different frequencies: fmin = 1.6GHz, fmax = 2.93GHz
- 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



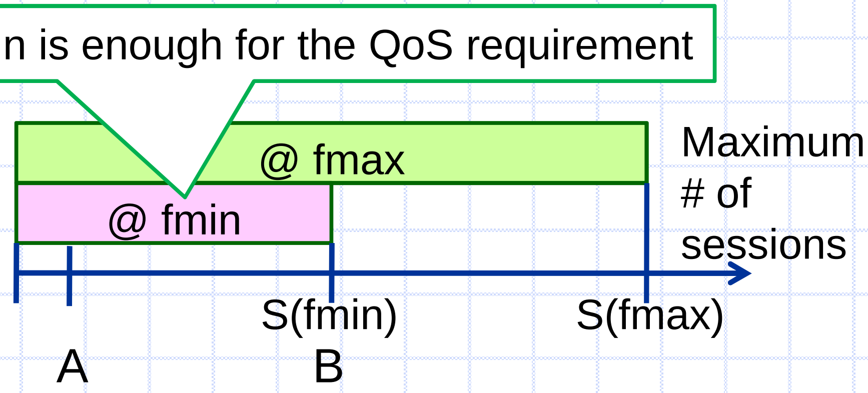
Paravirtualized VM: VM 2, VM 3
1GB allocated memory
kernel-xen-2.6.18-128.1.10.el5 (64-bit)
Apache 2.2.3

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(freq=fmin or fmax):
of the maximum simultaneous sessions
which can satisfy 95% TIME_GOOD QoS



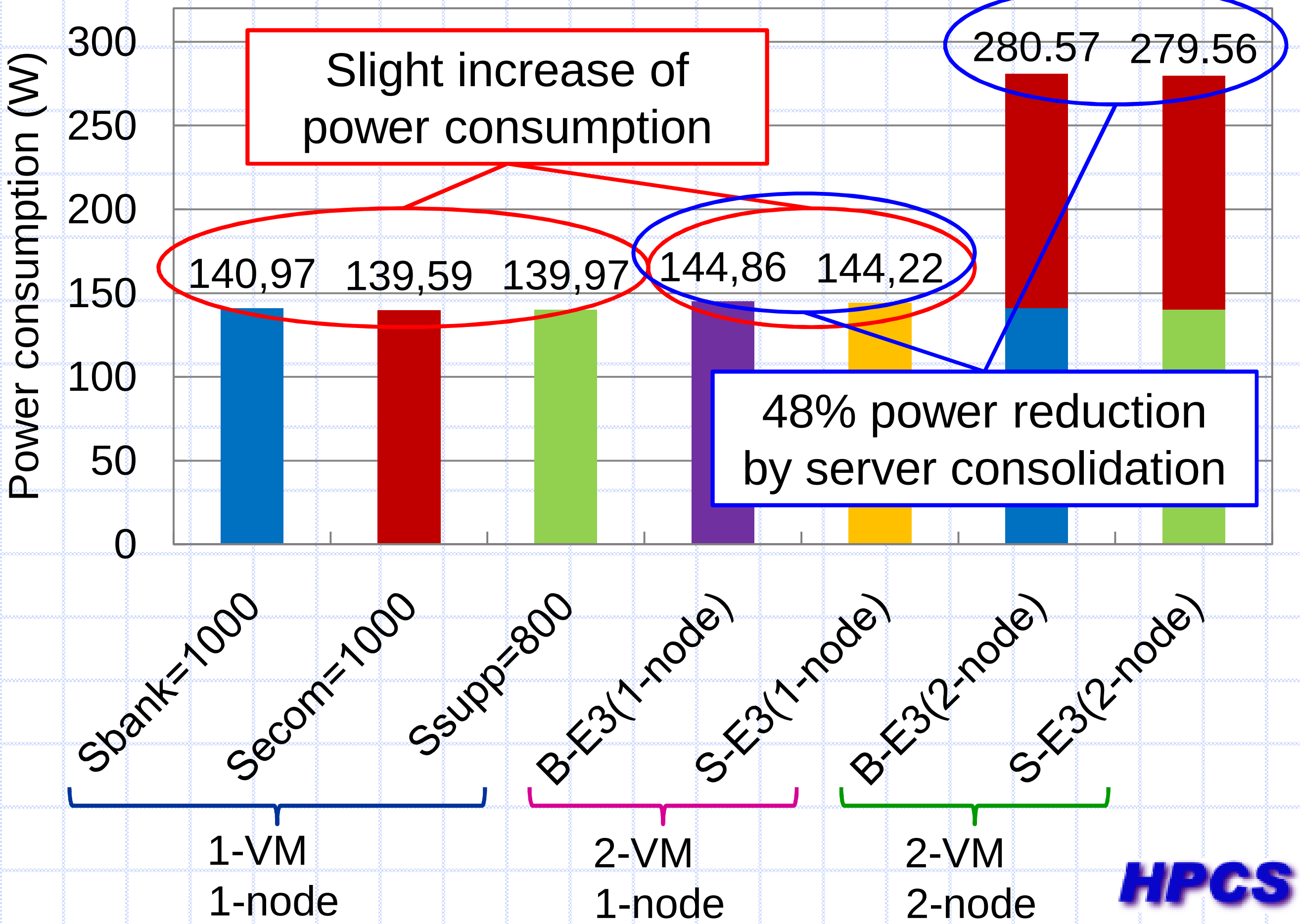
- Power reduction @ A (light-weight) and B (middle-weight)
  - 1-VM: A = 100 simultaneous sessions
    B = S(fmin) on each workload
  - 2-VM: A = B-E1 or S-E1 workload set
    B = B-E2 or S-E2 workload set

Sorry, please see our paper for detailed values

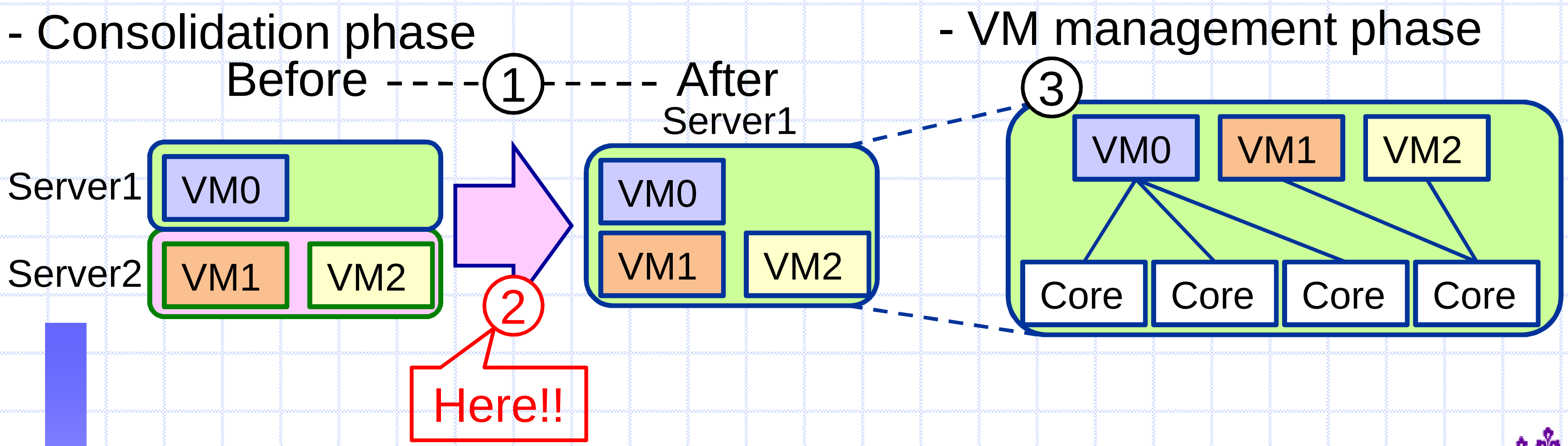
- Results of power reduction: {1 - P(fmin)/P(fmax)} * 100 [%]
  - 1-VM evaluation
    - 1% power reduction @ A
    - 5% power reduction @ B
  - 2-VM evaluation
    - 2% power reduction @ A
    - 8% power reduction @ B

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

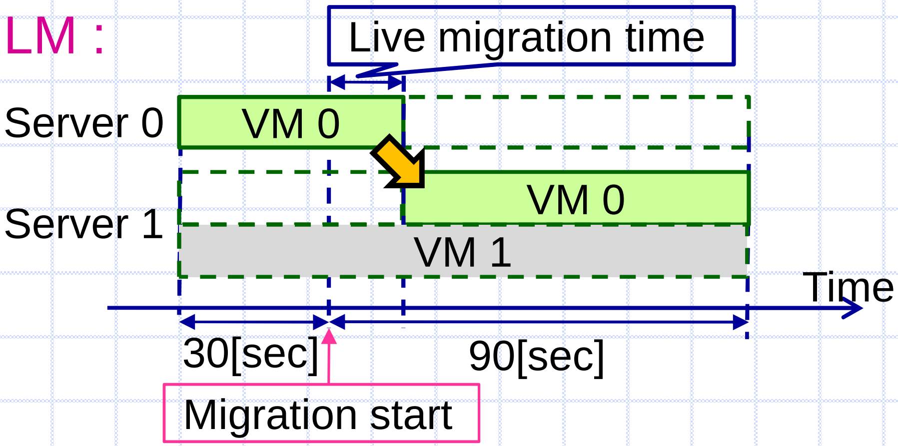


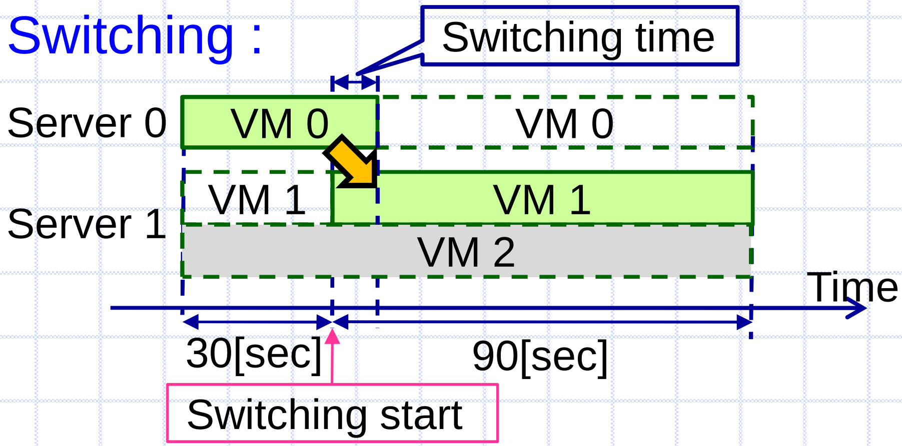
Power and performance evaluation of load migration schemes



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）



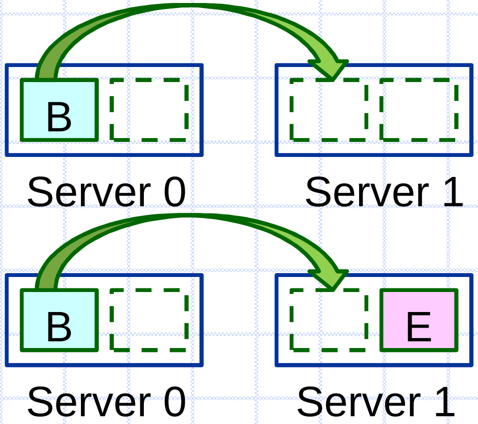


	VM 0	VM 1
A	Banking	None
B	Banking	E-commerce

	VM 0	VM 1	VM 2
A	Banking	Banking	None
B	Banking	Banking	E-commerce

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

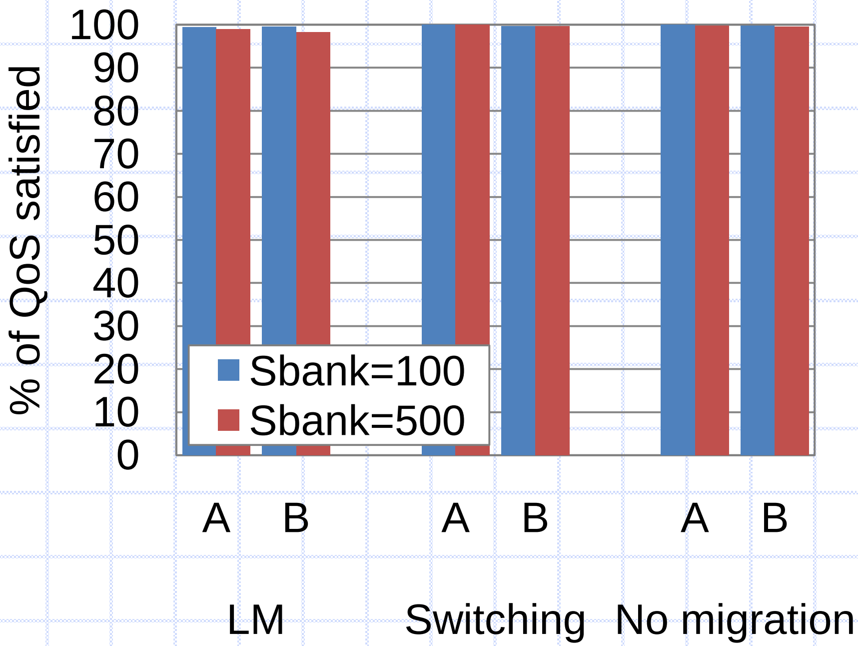


- Workload sets (with different workload levels of Banking)
  - Running @ fmin
    - Sbank = 100 or 500, Secom = 600（static）
  - Running @ fmax
    - Sbank = 100, 500 or 1000, Secom = 1000（static）
- Load migration is started 30 seconds after the benchmark start (Benchmark duration: 120 seconds)

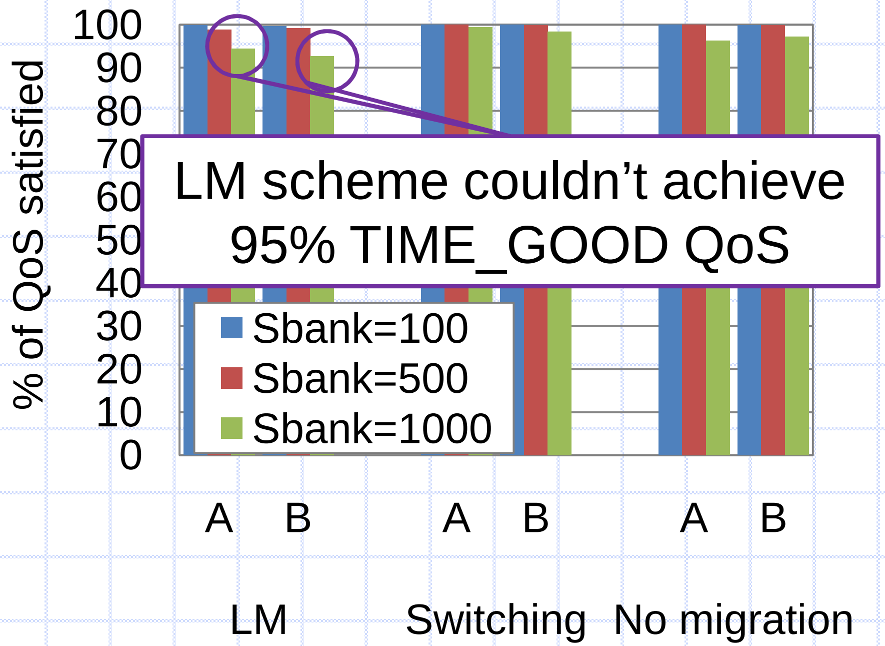
Results: QoS performance

(ratio of requested pages with TIME_GOOD response time)

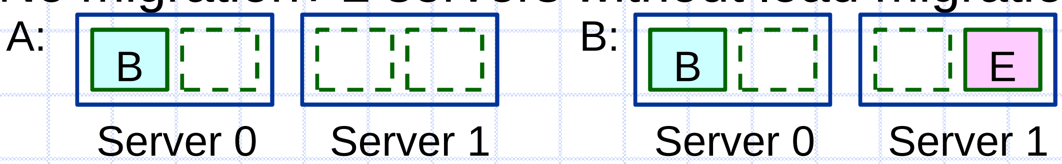
- Running @ fmin



- Running @ fmax



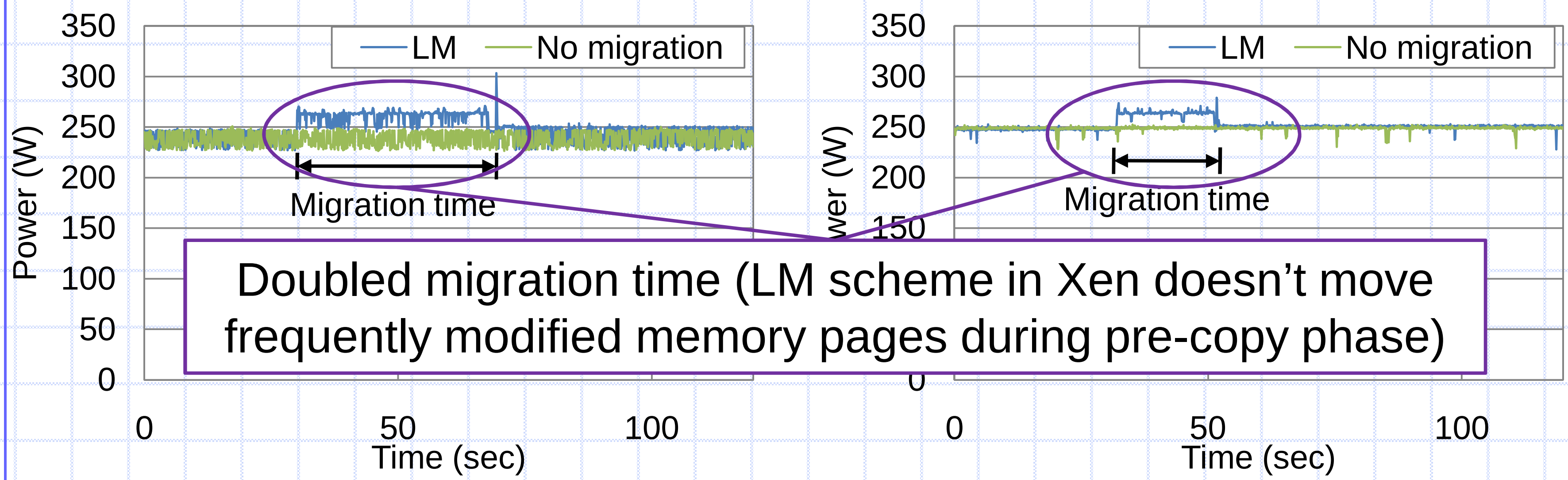
No migration : 2 servers without load migration



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)



Sbank = 500 (4% increase)

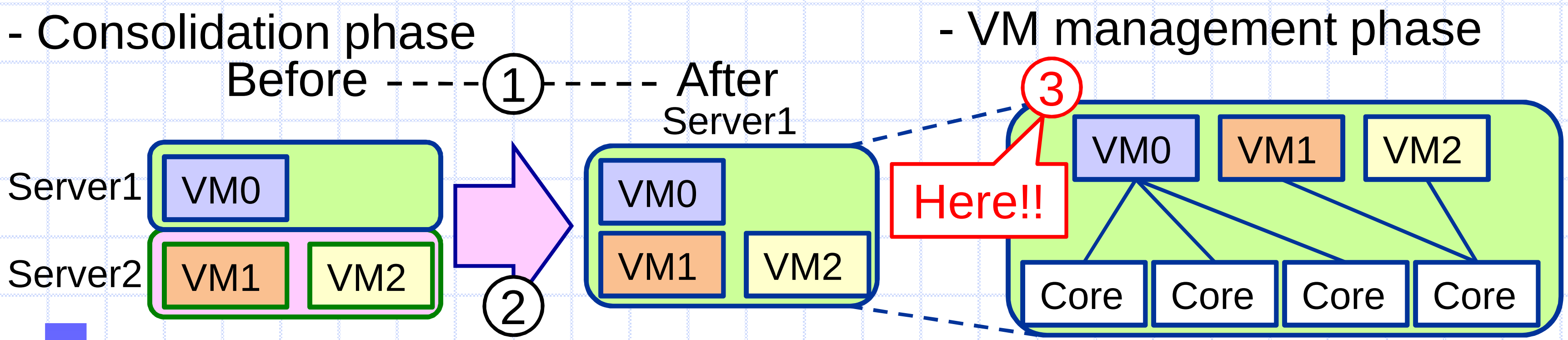
Sbank = 1000 (1% increase)

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)

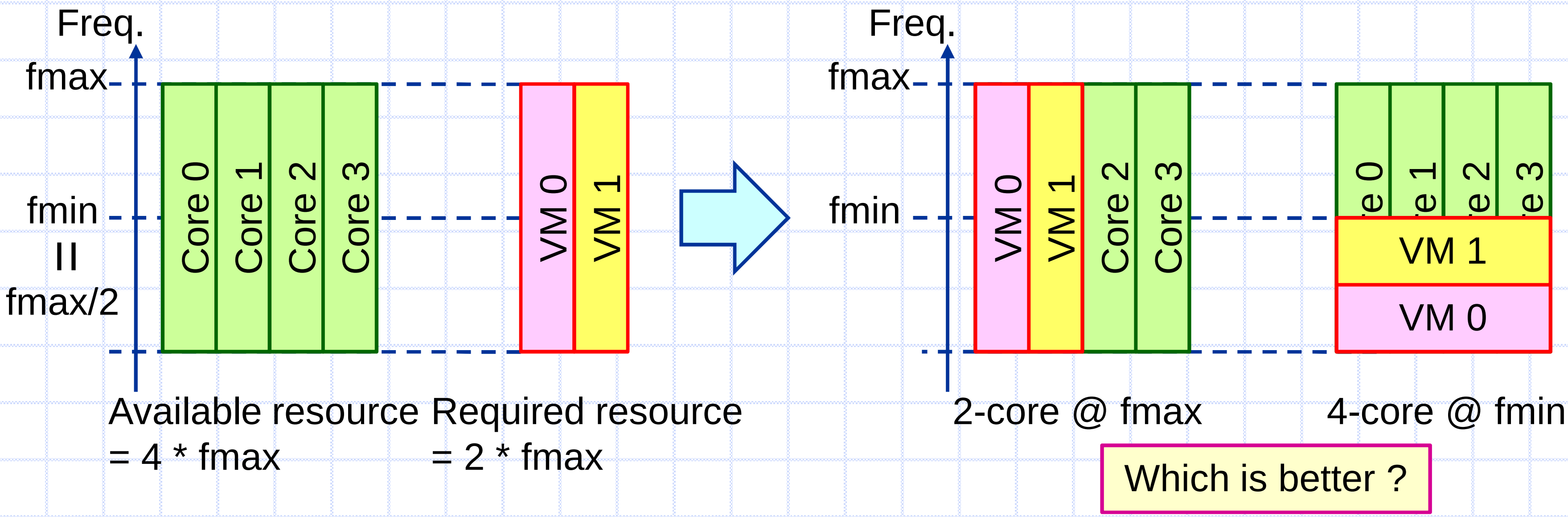
Evaluation with different processor-core allocation schemes



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

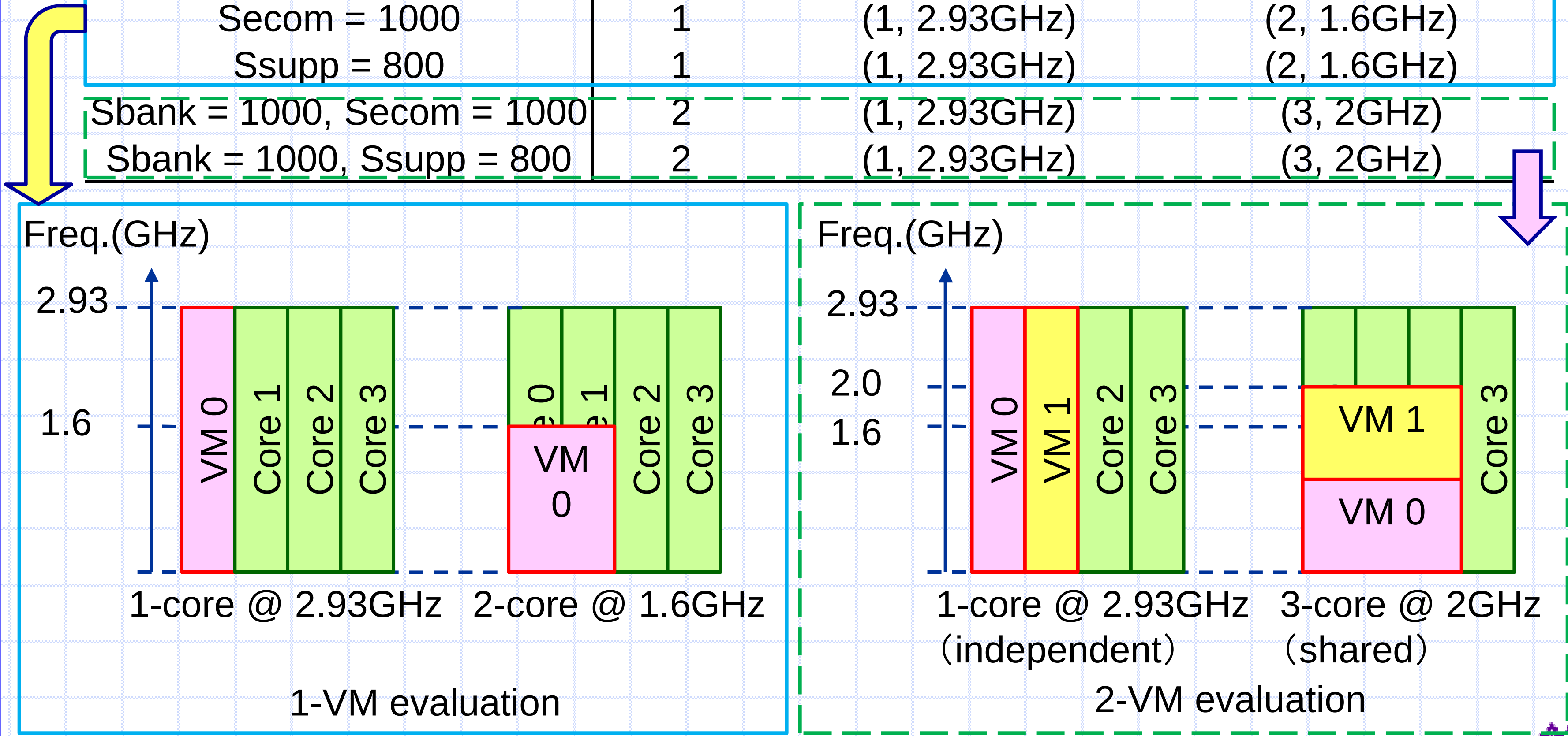
  Ex.) A quad-core processor, fmin = fmax/2



Evaluation overview

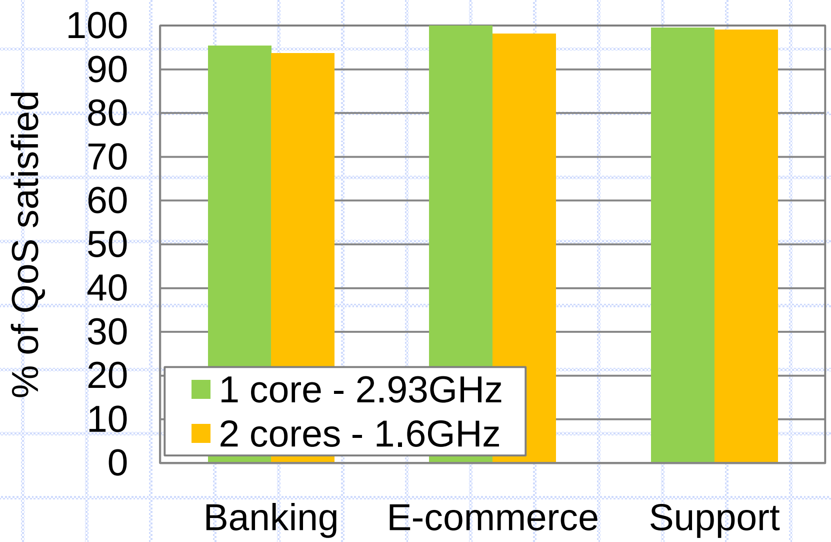
- Configurations (1-VM and 2-VM)

Workload set	# of VMs	(# of processor cores, frequency)	
Sbank = 1000	1	(1, 2.93GHz)	(2, 1.6GHz)
Secom = 1000	1	(1, 2.93GHz)	(2, 1.6GHz)
Ssupp = 800	1	(1, 2.93GHz)	(2, 1.6GHz)
Sbank = 1000, Secom = 1000	2	(1, 2.93GHz)	(3, 2GHz)
Sbank = 1000, Ssupp = 800	2	(1, 2.93GHz)	(3, 2GHz)

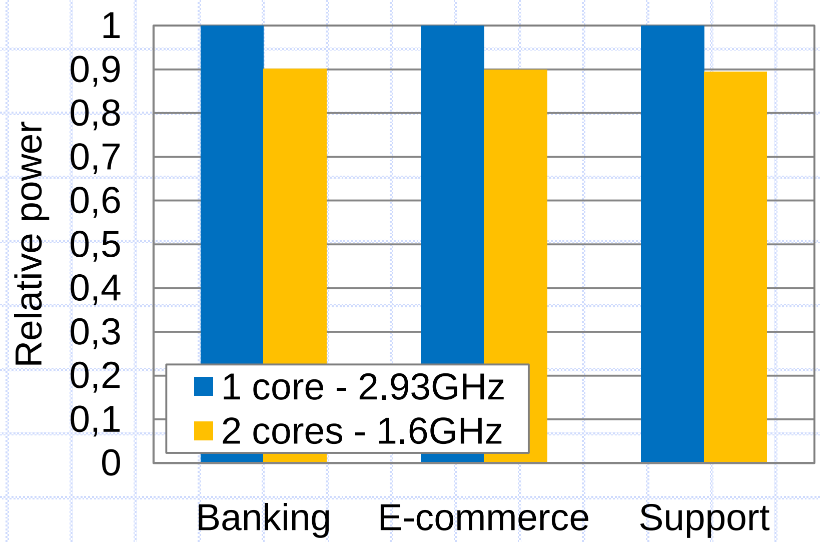


Results: 1-VM evaluation

- QoS performance



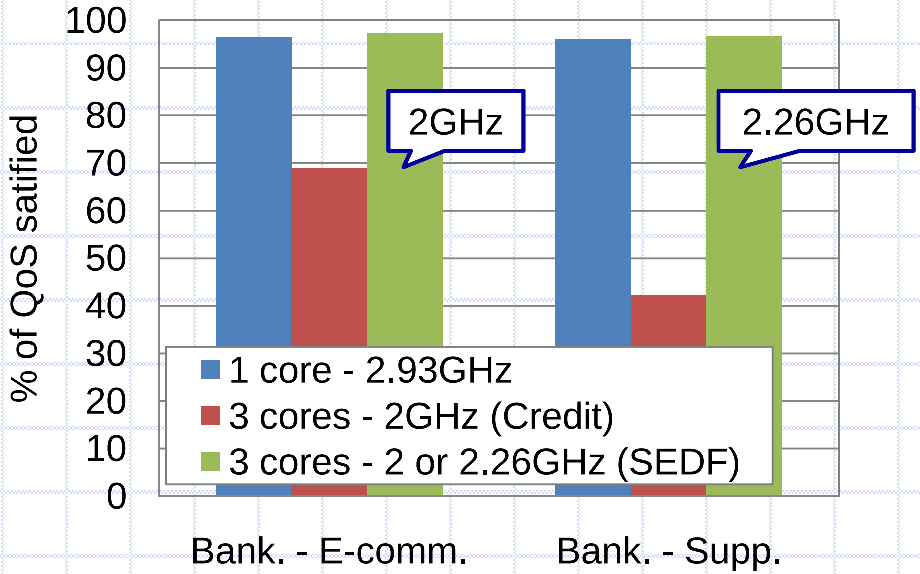
- Relative power



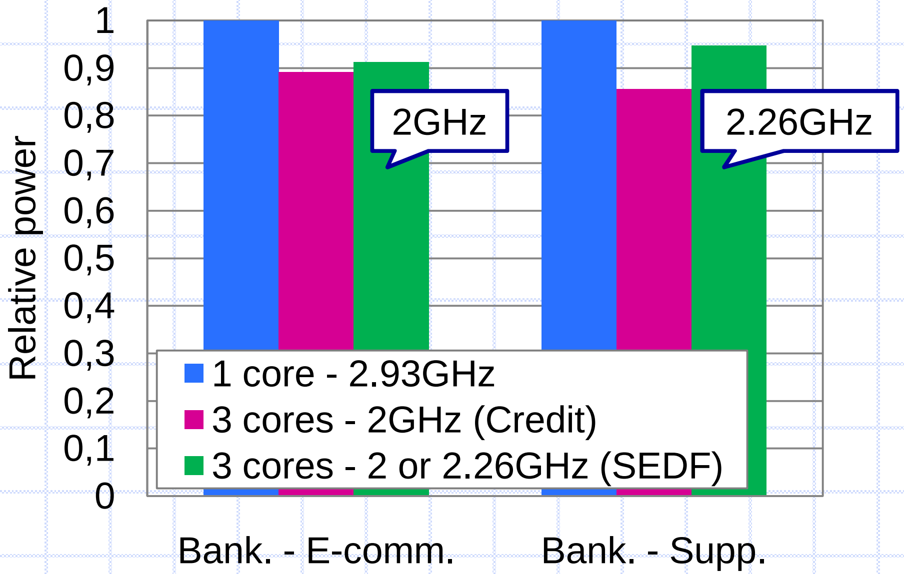
- Slight performance degradation of the QoS performance
  - Relatively increased load by VM scheduling on an additional processor core and the decreased running frequency
- About 10% power reduction in all workloads

Results: 2-VM evaluation

- QoS performance



- Relative power



- 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)

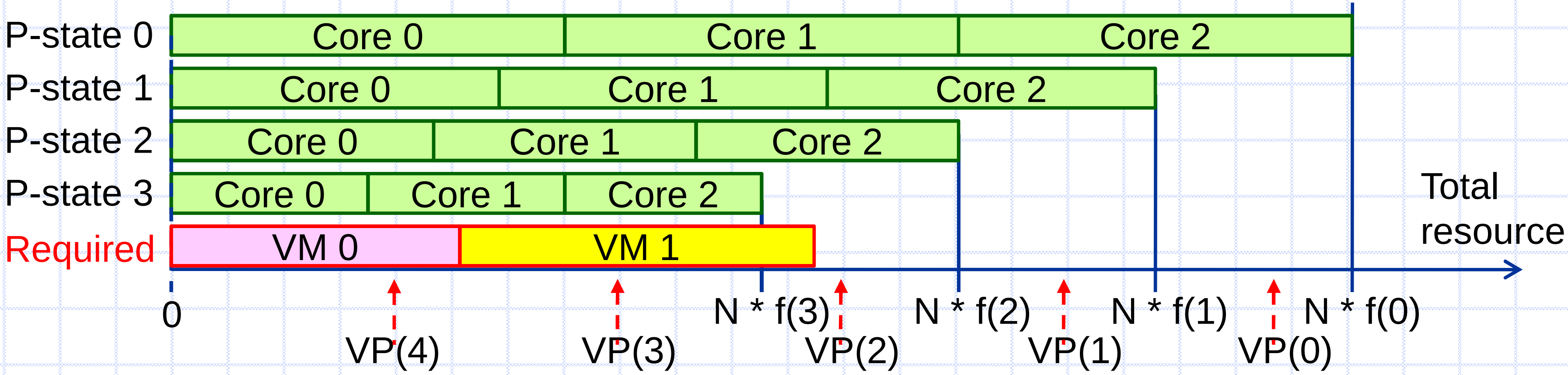
Application to existing methods

- 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 $N * f(i+1) < VP(j) <= N * f(i)$ and time slice management for VMs

Available maximum # of processor cores: N = 3

Frequency of hardware P-state i : f(i) where i = 0,1,2,3 (f(0) > f(1) > f(2) > f(3))



➡ 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!!