{ Considering I/O Processing in CloudSim for Performance and Energy Evaluation }

Hamza Ouarnoughi

Plan

- Context
- Introduction & Problem Statement
- Contribution
- Evaluation
- Conclusion & Future Work

Context

General Context

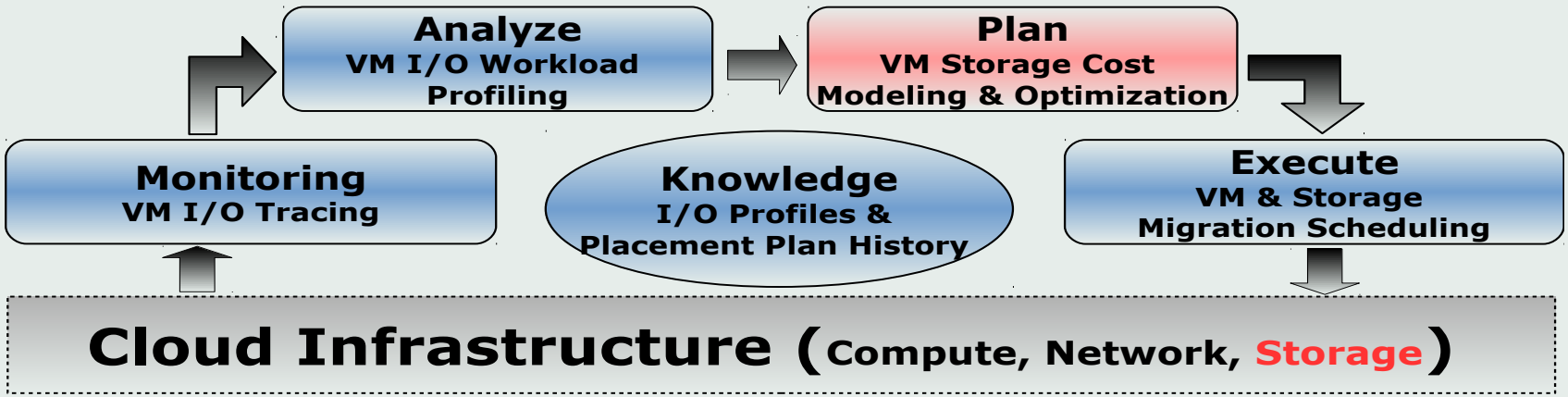
This work is part of a project that aims to optimize performance and energy consumption of a hybrid storage system in the context of IaaS Cloud

General Context

This work is part of a project that aims to optimize performance and energy consumption of a hybrid storage system in the context of IaaS Cloud

Storage Management Model

The storage management is an autonomic loop that obeys the MAPE-K (Monitor, Analyze, Plan, Execute - Knowledge) model :



Introduction & Problem Statement

Introduction & Problem Statement

Cloud Infrastructure Power Consumption

Introduction & Problem Statement

Cloud Infrastructure Power Consumption

- **Minimizing infrastructure power consumption** : an important concern for Cloud providers

Introduction & Problem Statement

Cloud Infrastructure Power Consumption

- **Minimizing infrastructure power consumption** : an important concern for Cloud providers
- **VM placement optimization methods** : popular approaches to minimize data centers power consumption

Introduction & Problem Statement

Cloud Infrastructure Power Consumption

- **Minimizing infrastructure power consumption** : an important concern for Cloud providers
- **VM placement optimization methods** : popular approaches to minimize data centers power consumption

State of the art

Introduction & Problem Statement

Cloud Infrastructure Power Consumption

- **Minimizing infrastructure power consumption** : an important concern for Cloud providers
- **VM placement optimization methods** : popular approaches to minimize data centers power consumption

State of the art

- **VM placement optimization** : usually based on CPU load

Introduction & Problem Statement

Cloud Infrastructure Power Consumption

- **Minimizing infrastructure power consumption** : an important concern for Cloud providers
- **VM placement optimization methods** : popular approaches to minimize data centers power consumption

State of the art

- **VM placement optimization** : usually based on CPU load
- **Storage activities** : may greatly contribute to the overall data center power consumption (up to 40% [1])

Introduction & Problem Statement

Cloud Infrastructure Power Consumption

- **Minimizing infrastructure power consumption** : an important concern for Cloud providers
- **VM placement optimization methods** : popular approaches to minimize data centers power consumption

State of the art

- **VM placement optimization** : usually based on CPU load
- **Storage activities** : may greatly contribute to the overall data center power consumption (up to 40% [1])
- **CloudSim** : the most used simulator for VM placement approaches implementation and evaluation

[1] Z. Li, K. M. Greenan, A. W. Leung, E. Zadok "Power Consumption in Enterprise-Scale Backup Storage Systems"

Problem Statement : CloudSim

Power Consumption & I/O Workload Execution

Problem Statement : CloudSim

Power Consumption & I/O Workload Execution

During I/O workload execution, power is mainly consumed by :

Problem Statement : CloudSim

Power Consumption & I/O Workload Execution

During I/O workload execution, power is mainly consumed by :

- Storage system components

Problem Statement : CloudSim

Power Consumption & I/O Workload Execution

During I/O workload execution, power is mainly consumed by :

- Storage system components
- CPU and memory when processing I/Os (i.e. executing the I/O software stack)

Problem Statement : CloudSim

Power Consumption & I/O Workload Execution

During I/O workload execution, power is mainly consumed by :

- Storage system components
- CPU and memory when processing I/Os (i.e. executing the I/O software stack)

CloudSim & Storage Capabilities

Problem Statement : CloudSim

Power Consumption & I/O Workload Execution

During I/O workload execution, power is mainly consumed by :

- Storage system components
- CPU and memory when processing I/Os (i.e. executing the I/O software stack)

CloudSim & Storage Capabilities

- **VM I/O workload execution** : not considered (i.e. its related time and power consumption)

Problem Statement : CloudSim

Power Consumption & I/O Workload Execution

During I/O workload execution, power is mainly consumed by :

- Storage system components
- CPU and memory when processing I/Os (i.e. executing the I/O software stack)

CloudSim & Storage Capabilities

- **VM I/O workload execution** : not considered (i.e. its related time and power consumption)
- **Storage system** : limited and consists of an HDD-based SAN (Storage Area Network)

Contribution

Contributions Summary

Contribution

Contributions Summary

We propose an extension for CloudSim storage capabilities in form of three contributions :

Contribution

Contributions Summary

We propose an extension for CloudSim storage capabilities in form of three contributions :

- Considering VMs I/O workload execution and related time and power consumption

Contribution

Contributions Summary

We propose an extension for CloudSim storage capabilities in form of three contributions :

- Considering VMs I/O workload execution and related time and power consumption
- Adding Flash-based SSD and local storage support, and power related capabilities

Contribution

Contributions Summary

We propose an extension for CloudSim storage capabilities in form of three contributions :

- Considering VMs I/O workload execution and related time and power consumption
- Adding Flash-based SSD and local storage support, and power related capabilities
- Adding a model to compute the CPU time related to I/O workload processing

Time Computation

Before the Extension

Before the Extension

- **CloudSim I/O activity time** : only affected by the Cloudlet [1] transfer during its submission or migration

Before the Extension

- **CloudSim I/O activity time** : only affected by the Cloudlet [1] transfer during its submission or migration

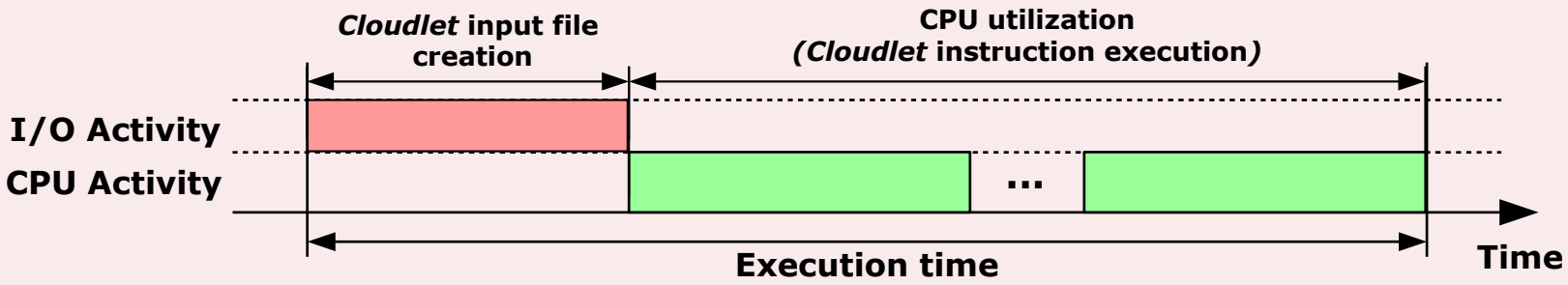


Figure – Time model before the extension

1. In CloudSim, a Cloudlet is a process representing a CPU workload and executed by a VM

Time computation

After the Extension

Time computation

After the Extension

Our extension adds two events affecting the simulation time :

Time computation

After the Extension

Our extension adds two events affecting the simulation time :

- I/O workload execution (i.e. storage device and CPU time)

Time computation

After the Extension

Our extension adds two events affecting the simulation time :

- I/O workload execution (i.e. storage device and CPU time)
- VM image (virtual disk) creation (e.g. copy from a repository to the local storage)

After the Extension

Our extension adds two events affecting the simulation time :

- I/O workload execution (i.e. storage device and CPU time)
- VM image (virtual disk) creation (e.g. copy from a repository to the local storage)

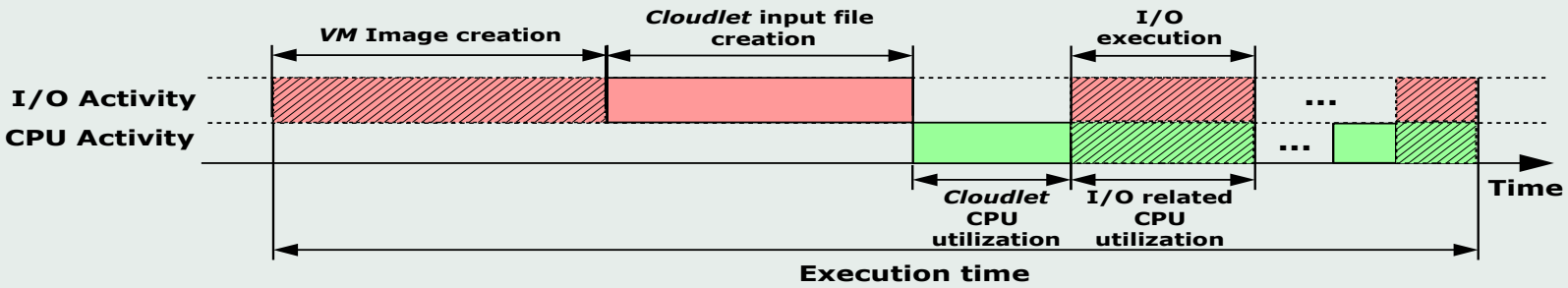


Figure – Time model after the extension

Power Computation

Before the Extension

Power Computation

Before the Extension

- **CloudSim power model** : uses only hosts CPU utilization to determine their power consumption

Power Computation

Before the Extension

- **CloudSim power model** : uses only hosts CPU utilization to determine their power consumption

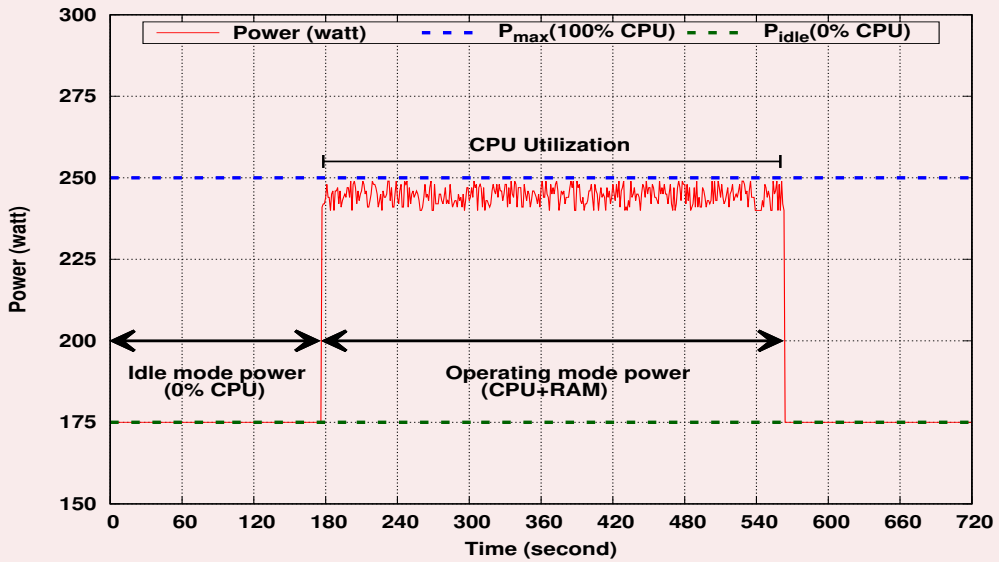


Figure – Power model before the extension

After the Extension

Power Computation

After the Extension

- Our extension adds to the existing power model the power consumption generated by the I/O workload execution [2]

Power Computation

After the Extension

- Our extension adds to the existing power model the power consumption generated by the I/O workload execution [2]

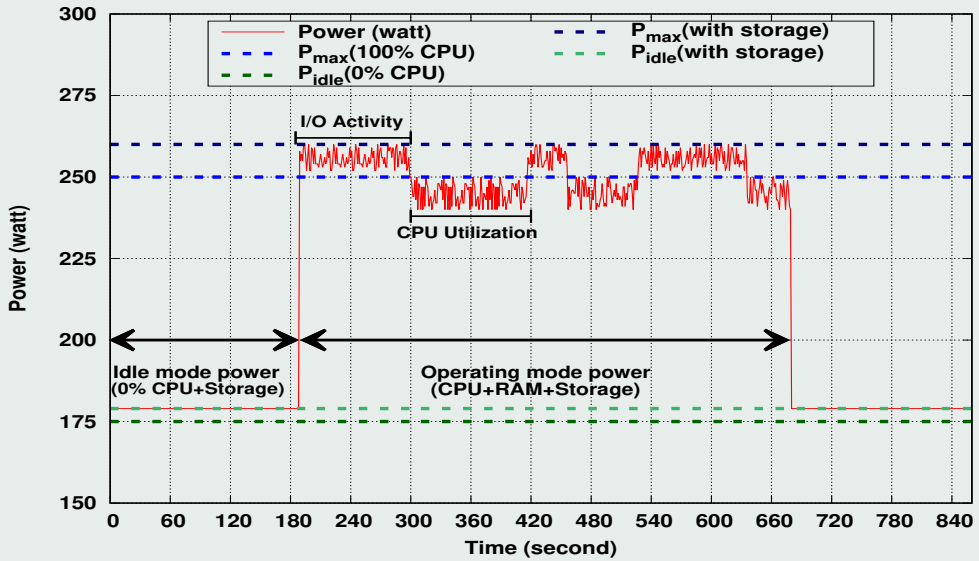


Figure – Power model after the extension

[2] H. Ouarnoughi, J. Boukhobza, F. Singhoff, S. Rubini : "A Cost Model for Virtual Machine Storage in Cloud IaaS Context"

Before the Extension

Storage Device Support

Before the Extension

- **Native CloudSim storage system** :an HDD-based SAN, shared with all data center's hosts

Storage Device Support

Before the Extension

- **Native CloudSim storage system** :an HDD-based SAN, shared with all data center's hosts
- **Storage system performances** : $f(max(HDD_{Performance}, SAN_{Throughput}))$

Before the Extension

- **Native CloudSim storage system** :an HDD-based SAN, shared with all data center's hosts
- **Storage system performances** : $f(max(HDD_{Performance}, SAN_{Throughput}))$

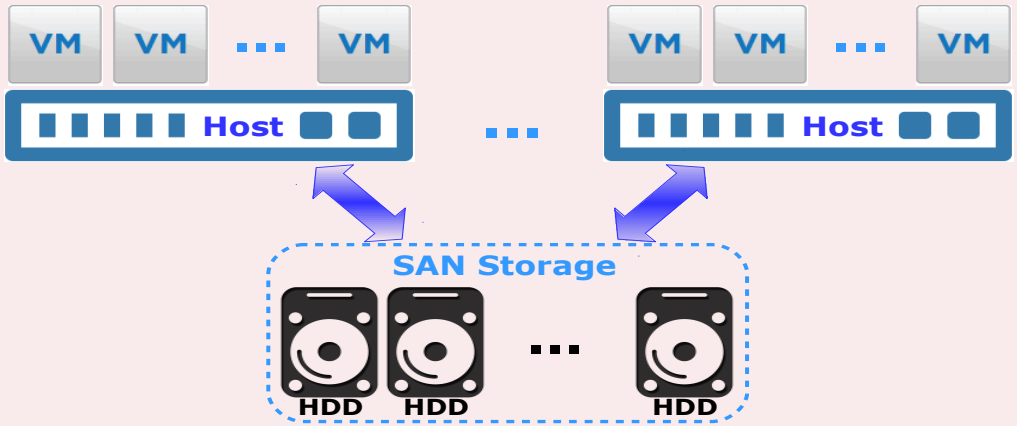


Figure – Data center storage system before the extension

After the Extension

After the Extension

Three main storage system capabilities :

After the Extension

Three main storage system capabilities :

- Support for flash-based SSD

After the Extension

Three main storage system capabilities :

- Support for flash-based SSD
- Support for local storage (i.e. DAS storage)

Storage Device Support

After the Extension

Three main storage system capabilities :

- Support for flash-based SSD
- Support for local storage (i.e. DAS storage)
- More storage device attributes (e.g. power, performance, reliability, etc)

After the Extension

Three main storage system capabilities :

- Support for flash-based SSD
- Support for local storage (i.e. DAS storage)
- More storage device attributes (e.g. power, performance, reliability, etc)

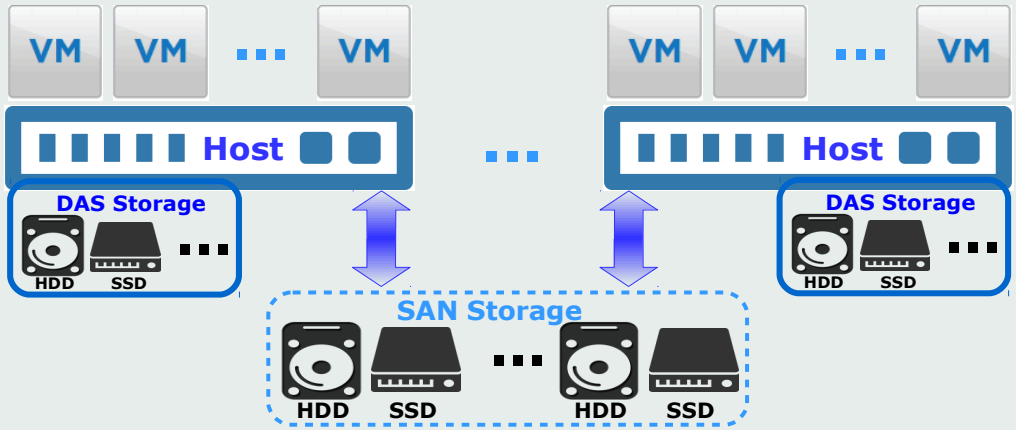


Figure – Data center storage system after the extension

I/O Workload and CPU Time

CPU time for I/O Workload

I/O Workload and CPU Time

CPU time for I/O Workload

- **VM I/O workload execution** : storage system activities + CPU utilization due to I/O software stack execution

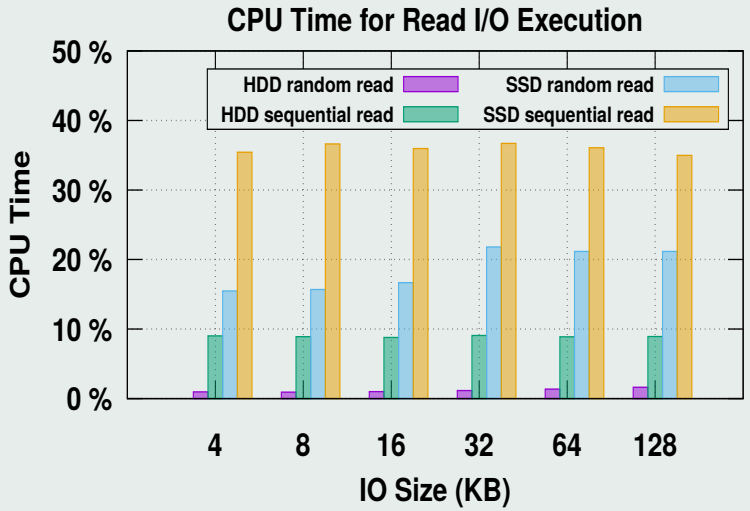
I/O Workload and CPU Time

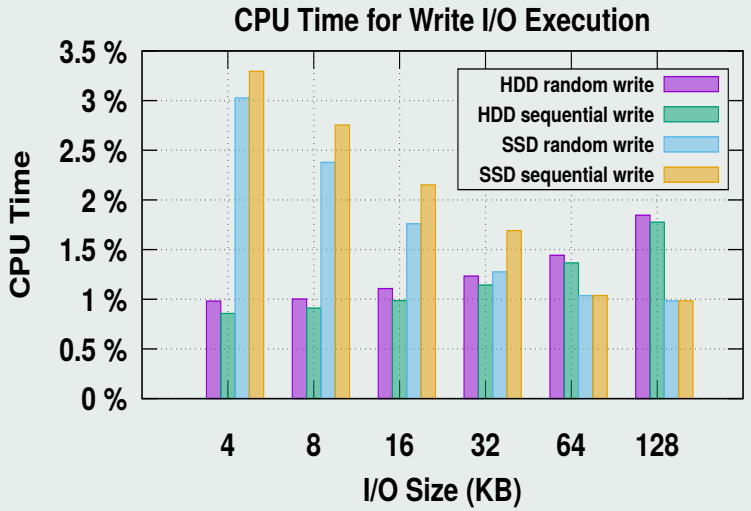
CPU time for I/O Workload

- **VM I/O workload execution** : storage system activities + CPU utilization due to I/O software stack execution
- **I/O workload** $\rightarrow$ **CPU utilization** : a correlation model that gives CPU time depending on I/O workload characteristics

CPU time for I/O Workload

- **VM I/O workload execution** : storage system activities + CPU utilization due to I/O software stack execution
- **I/O workload** → **CPU utilization** : a correlation model that gives CPU time depending on I/O workload characteristics





Evaluation

Evaluation : Summary

Evaluation Methodology

Evaluation : Summary

Evaluation Methodology

Two steps to evaluate CloudSim storage extensions :

Evaluation : Summary

Evaluation Methodology

Two steps to evaluate CloudSim storage extensions :

- **Real workload execution** : running experiments in a real environment + collecting CPU, memory, and I/O measures.

Evaluation : Summary

Evaluation Methodology

Two steps to evaluate CloudSim storage extensions :

- **Real workload execution** : running experiments in a real environment + collecting CPU, memory, and I/O measures.
- **Workload simulation** : replaying the same scenario in CloudSim + turning the storage system ON/OFF

Evaluation : Summary

Evaluation Methodology

Two steps to evaluate CloudSim storage extensions :

- **Real workload execution** : running experiments in a real environment + collecting CPU, memory, and I/O measures.
- **Workload simulation** : replaying the same scenario in CloudSim + turning the storage system ON/OFF

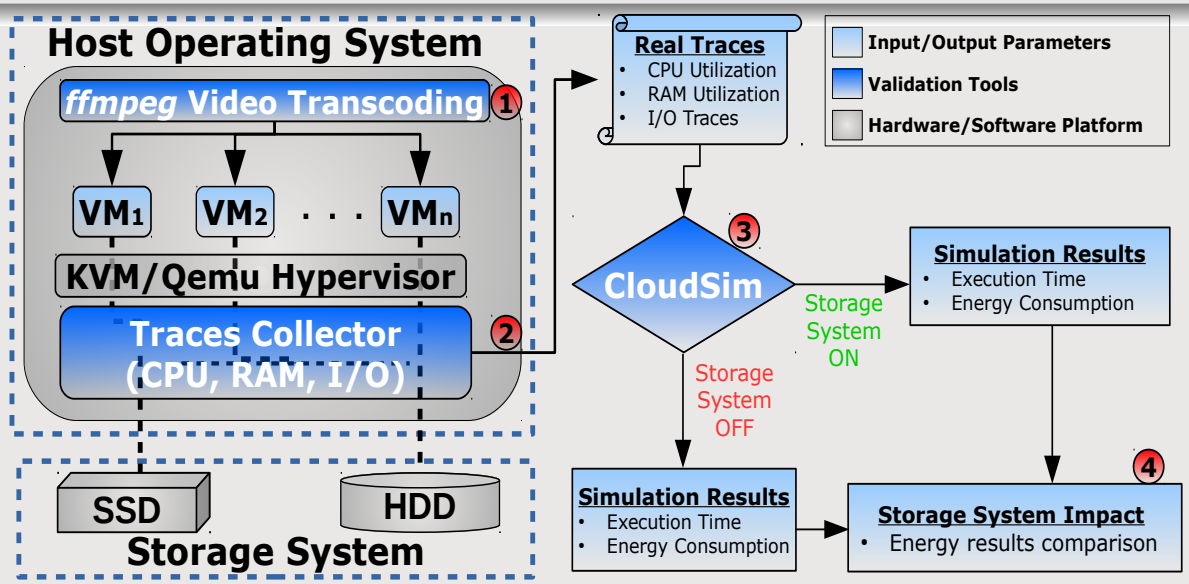


Figure – Evaluation methodology

Workload execution

Real workload execution

The real workload execution phase includes two steps :

Workload execution

Real workload execution

The real workload execution phase includes two steps :

- **Workload execution** : VMs running encoding video benchmark as a real use case

Workload execution

Real workload execution

The real workload execution phase includes two steps :

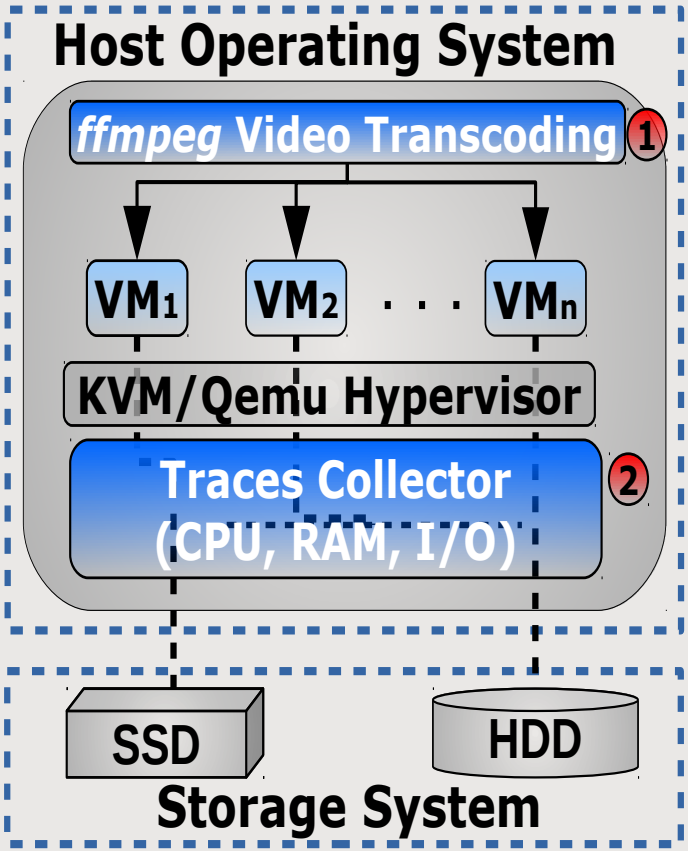
- **Workload execution** : VMs running encoding video benchmark as a real use case
- **Trace collection** : monitoring the host CPU, memory utilization + tracing I/Os

Workload execution

Real workload execution

The real workload execution phase includes two steps :

- **Workload execution** : VMs running encoding video benchmark as a real use case
- **Trace collection** : monitoring the host CPU, memory utilization + tracing I/Os



Workload simulation

Real workload execution

The workload simulation phase includes two steps :

Workload simulation

Real workload execution

The workload simulation phase includes two steps :

- **Simulation using real traces** : replaying traces in CloudSim + varying parameters (i.e. number of VMs/Host, storage system ON/OFF, and storage device HDD/SSD)

Workload simulation

Real workload execution

The workload simulation phase includes two steps :

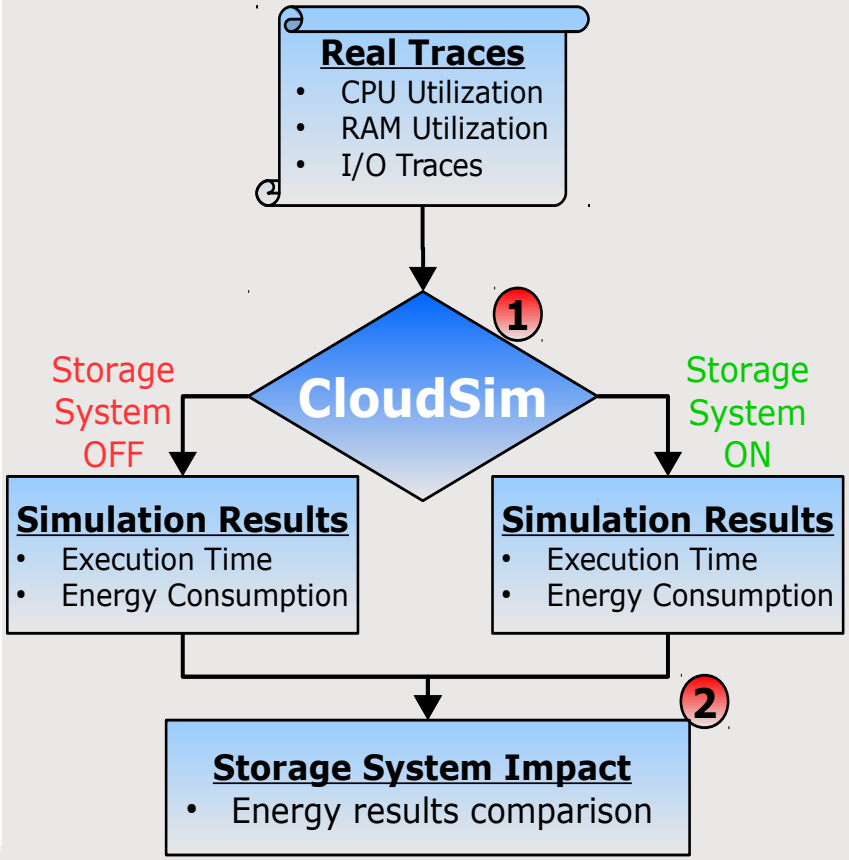
- **Simulation using real traces** : replaying traces in CloudSim + varying parameters (i.e. number of VMs/Host, storage system ON/OFF, and storage device HDD/SSD)
- **Simulation results comparison** : power consumption to quantify I/O workload execution impact

Workload simulation

Real workload execution

The workload simulation phase includes two steps :

- **Simulation using real traces** : replaying traces in CloudSim + varying parameters (i.e. number of VMs/Host, storage system ON/OFF, and storage device HDD/SSD)
- **Simulation results comparison** : power consumption to quantify I/O workload execution impact



Evaluation Metrics

Evaluation Metrics

Our evaluation results show two evaluation metrics :

Evaluation Metrics

Evaluation Metrics

Our evaluation results show two evaluation metrics :

- Power consumption related to I/O processing $\rightarrow$ validate the impact of I/Os on the overall power consumption

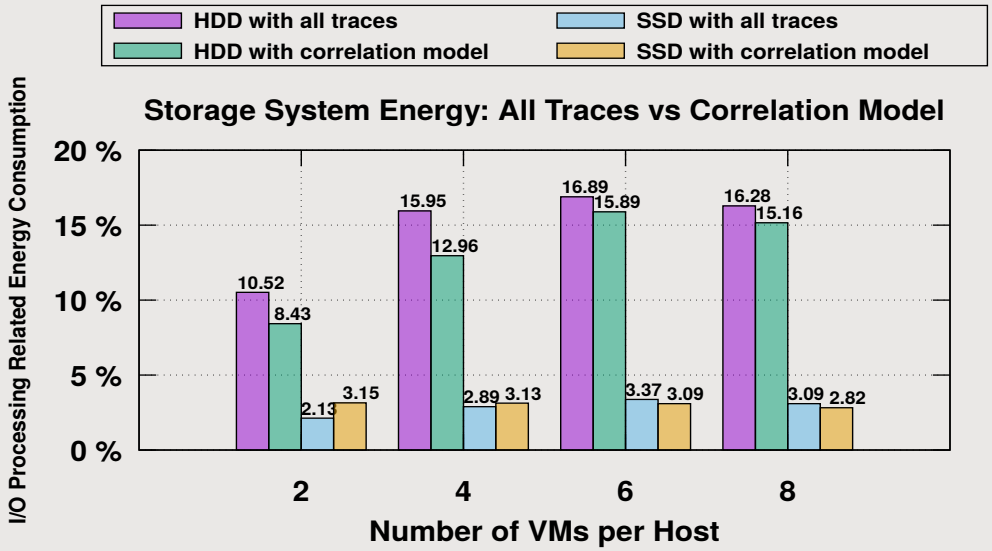
Evaluation Metrics

Evaluation Metrics

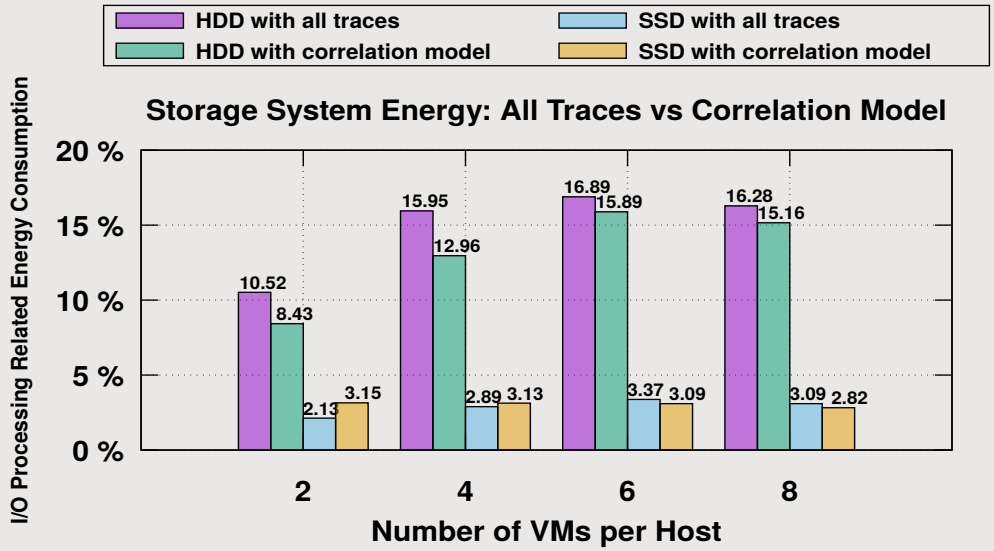
Our evaluation results show two evaluation metrics :

- Power consumption related to I/O processing $\rightarrow$ validate the impact of I/Os on the overall power consumption
- Difference between simulations using real traces and using correlation model $\rightarrow$ validate its accuracy

Evaluation Results & Discussion

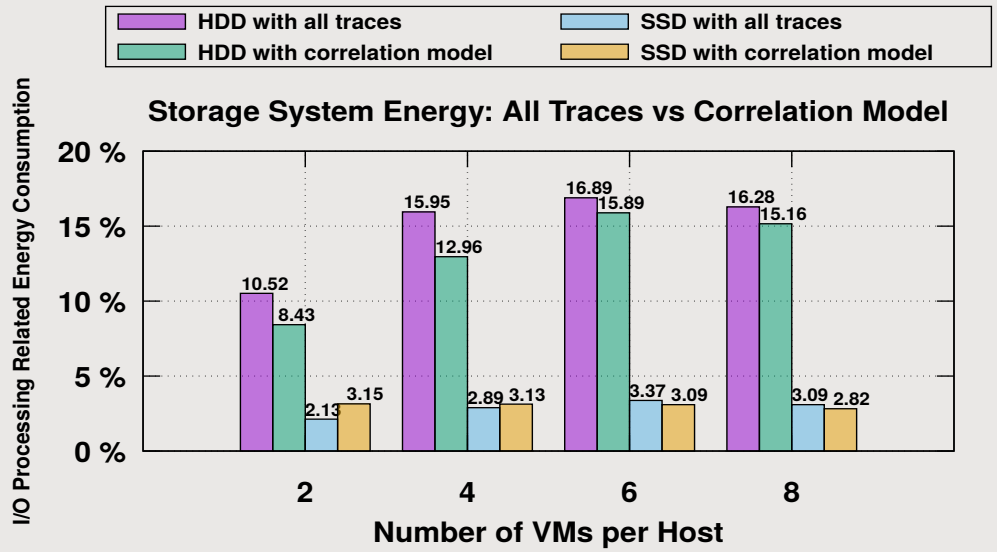


Evaluation Results & Discussion



Discussion

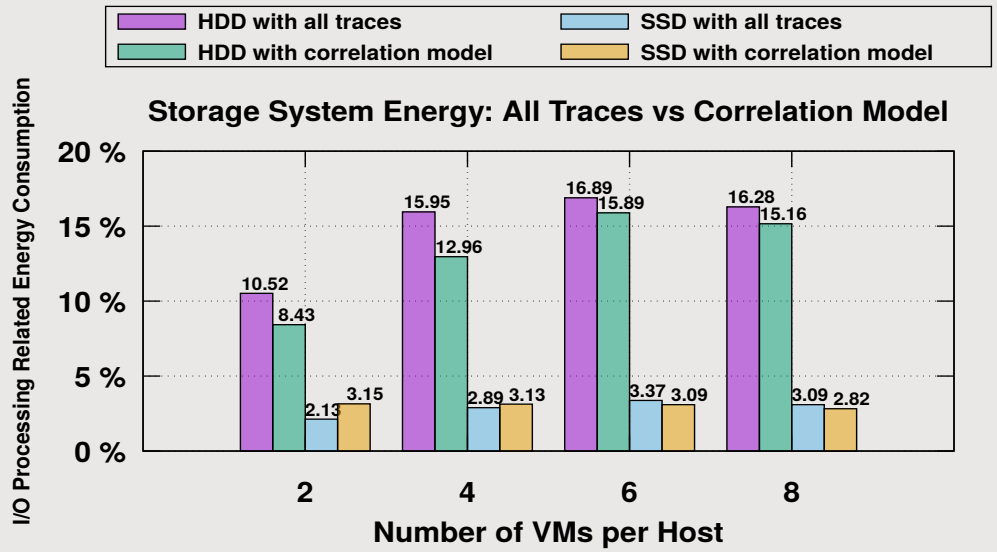
Evaluation Results & Discussion



Discussion

- Energy consumption ↗ when to the number of VMs/Host ↗

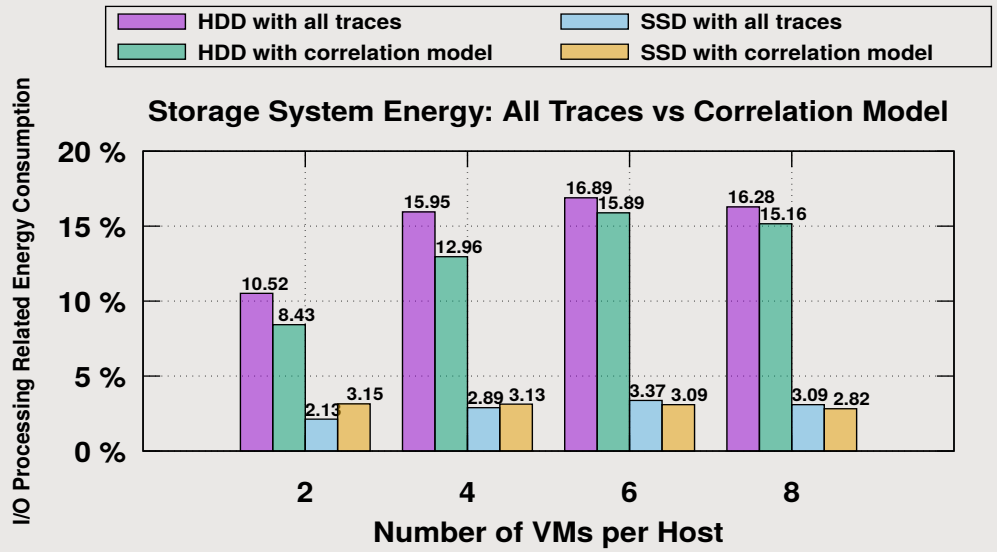
Evaluation Results & Discussion



Discussion

- Energy consumption ↗ when to the number of VMs/Host ↗
- Difference between HDD and SSD energy consumption

Evaluation Results & Discussion



Discussion

- Energy consumption ↗ when to the number of VMs/Host ↗
- Difference between HDD and SSD energy consumption
- Correlation model vs real traces : ∼12% average error rate

Conclusion & Future Work

Conclusion & Future Work

Conclusion & Future Work

Conclusion

We propose an extension of CloudSim to consider VM I/O workload processing in form of three contributions :

Conclusion & Future Work

Conclusion

We propose an extension of CloudSim to consider VM I/O workload processing in form of three contributions :

- Updating the time and energy models of CloudSim

Conclusion & Future Work

Conclusion

We propose an extension of CloudSim to consider VM I/O workload processing in form of three contributions :

- Updating the time and energy models of CloudSim
- Considering different storage devices (i.e. HDD and SSD)

Conclusion & Future Work

Conclusion

We propose an extension of CloudSim to consider VM I/O workload processing in form of three contributions :

- Updating the time and energy models of CloudSim
- Considering different storage devices (i.e. HDD and SSD)
- Including a CPU correlation model that depends on I/O workload characteristics and storage device type

Conclusion & Future Work

Conclusion

We propose an extension of CloudSim to consider VM I/O workload processing in form of three contributions :

- Updating the time and energy models of CloudSim
- Considering different storage devices (i.e. HDD and SSD)
- Including a CPU correlation model that depends on I/O workload characteristics and storage device type

Future Work

Conclusion & Future Work

Conclusion

We propose an extension of CloudSim to consider VM I/O workload processing in form of three contributions :

- Updating the time and energy models of CloudSim
- Considering different storage devices (i.e. HDD and SSD)
- Including a CPU correlation model that depends on I/O workload characteristics and storage device type

Future Work

- Using this implementation in a VM placement optimization approach

Conclusion & Future Work

Conclusion

We propose an extension of CloudSim to consider VM I/O workload processing in form of three contributions :

- Updating the time and energy models of CloudSim
- Considering different storage devices (i.e. HDD and SSD)
- Including a CPU correlation model that depends on I/O workload characteristics and storage device type

Future Work

- Using this implementation in a VM placement optimization approach
- Studying and integrating the problem of interference between VMs sharing the same storage device

My wish...

More research experience, so I am looking for a postdoctoral position !

Thanks

{ hamza.ouarnoughi@b-com.com }