Resource and Risk Management in Datacenters

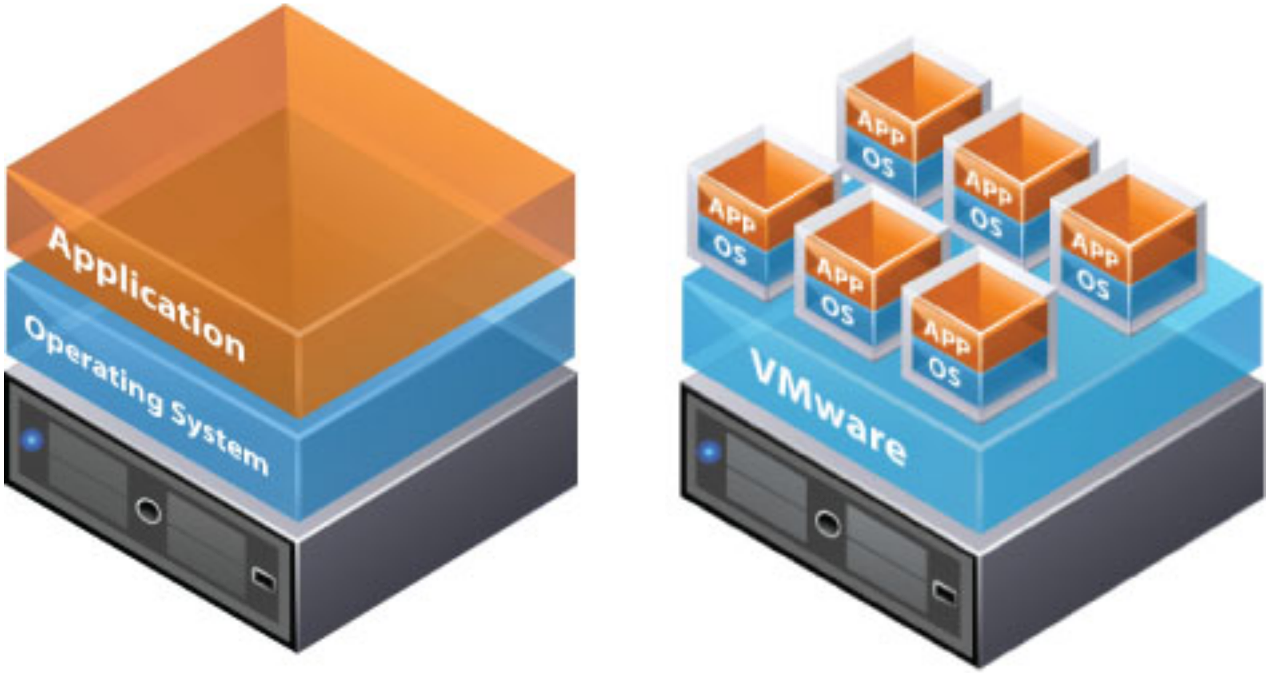
- Vincent van Beek
- Distributed Systems Group TU-Delft
- Supervisor: Prof. Dr. ir. Alexandru Iosup
- Promotor: Prof. Dr. ir. Dick Epema

TU Delft

Why Resource and Risk Management in Datacenters

- Datacenters are at the center of all IT systems
- Hosting Business Critical Applications
- New technology is introduced at a rapid rate
- Consolidation is driving costs down
- Many enterprise customers are risk averse and want guarantees

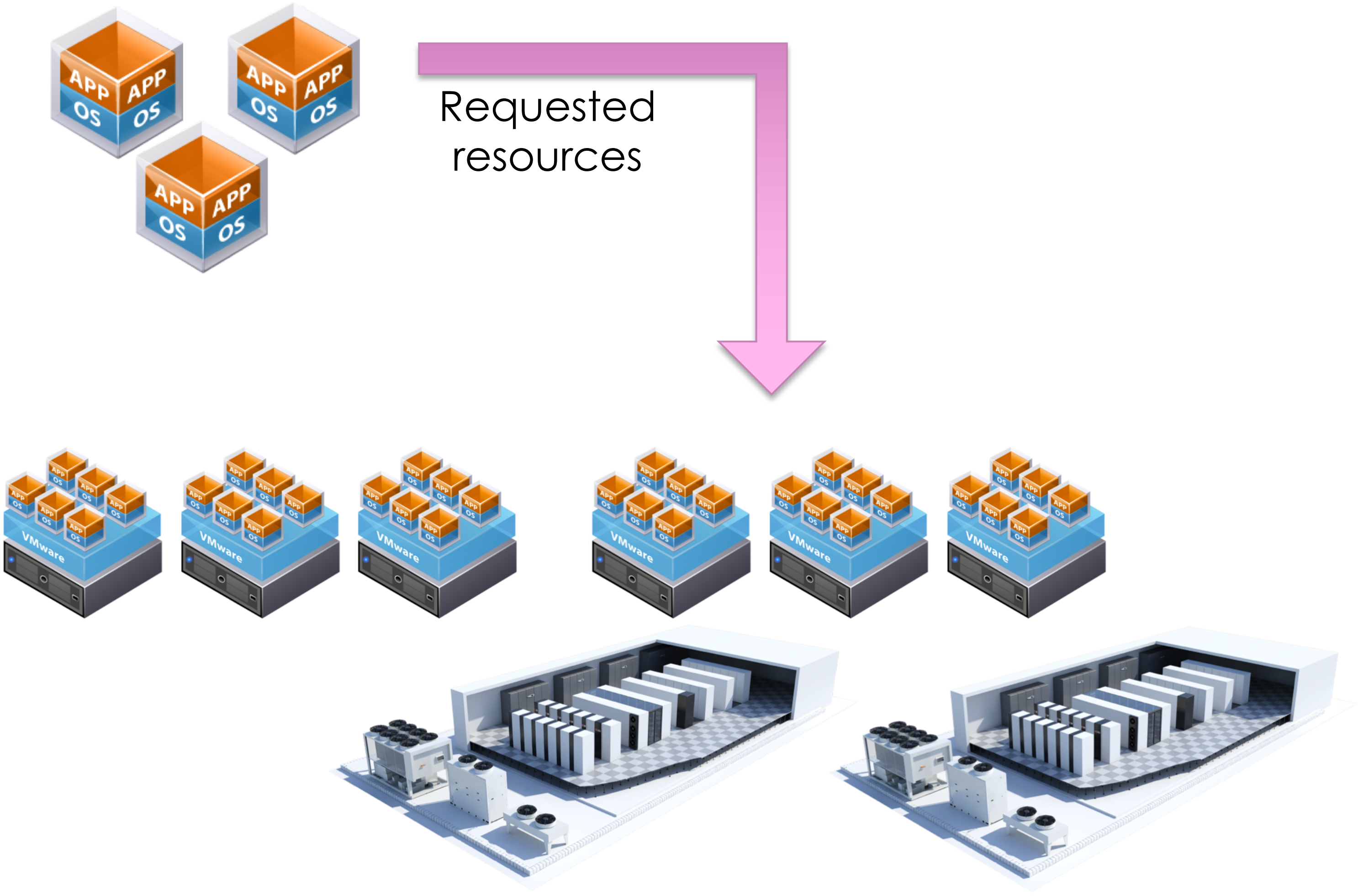
Virtualization in Datacenters



Traditional Architecture

Virtual Architecture

VM Placement



Requested resources

Used resources

Problem statement

Affinity

SLAs

Dynamic Characteristics

Multiple Resources

Where to put which VM?

Time Component

Many Single Resource Solutions

Scheduling

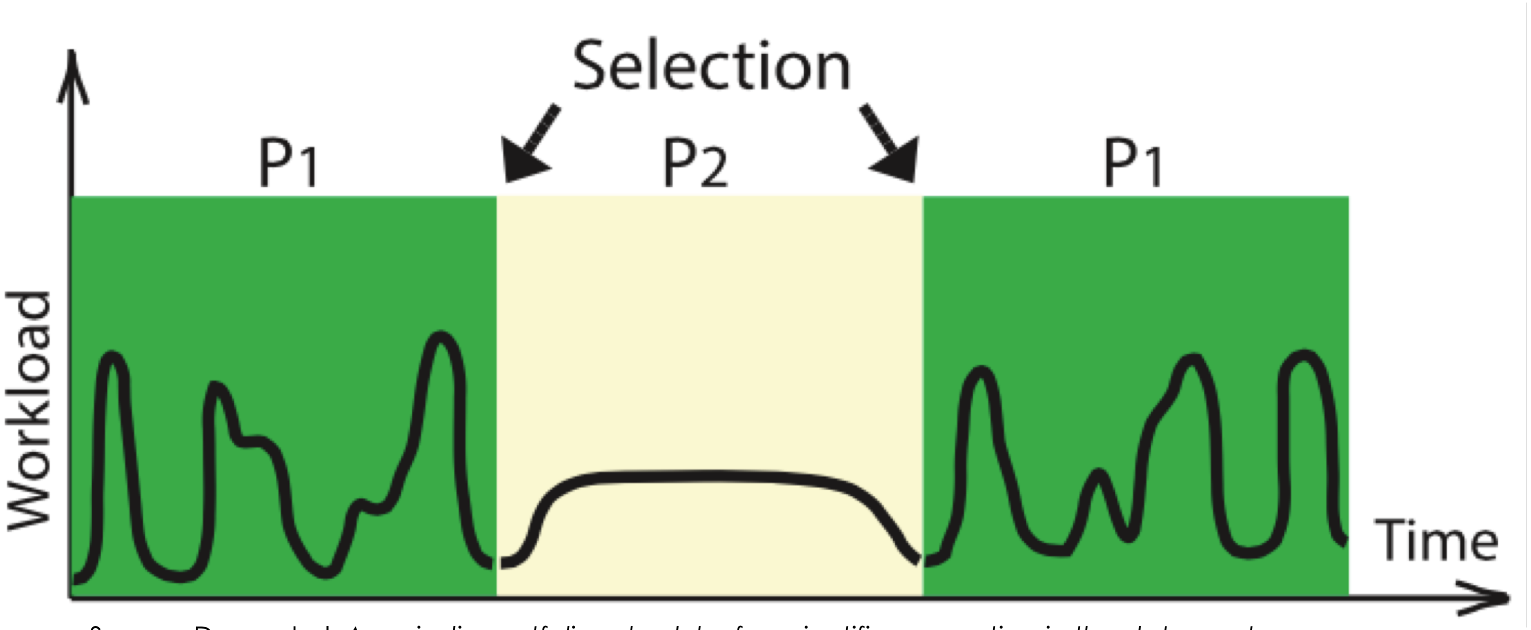
Workloads

Many Point Solutions

Anti-Affinity

BlackBox Problem

Why Portfolio Scheduling

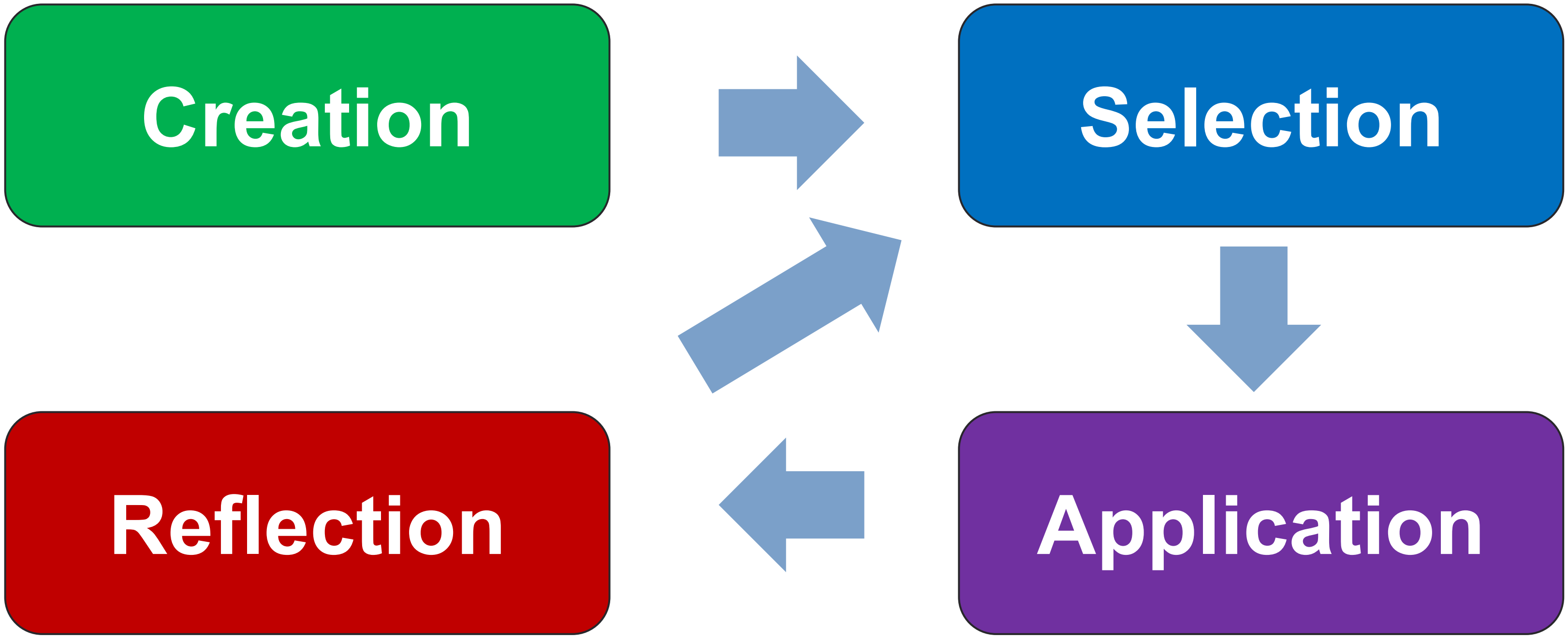


Source: Deng et al. *A periodic portfolio scheduler for scientific computing in the data center*

What is Portfolio Scheduling?

Which policies to include?

Which policy to activate?
Explain to sysadmin



Creation

Selection

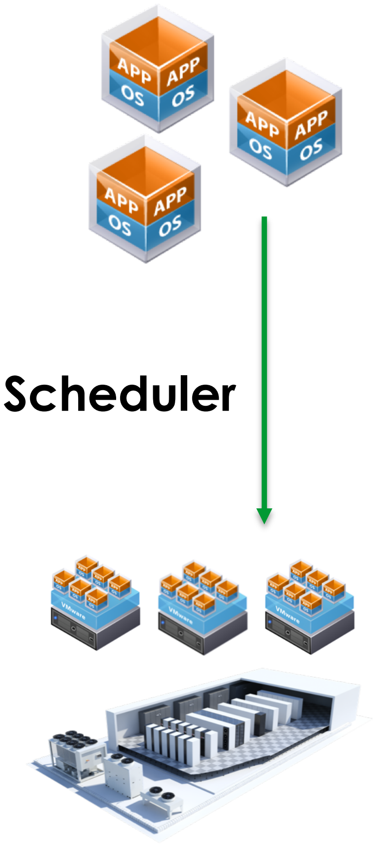
Reflection

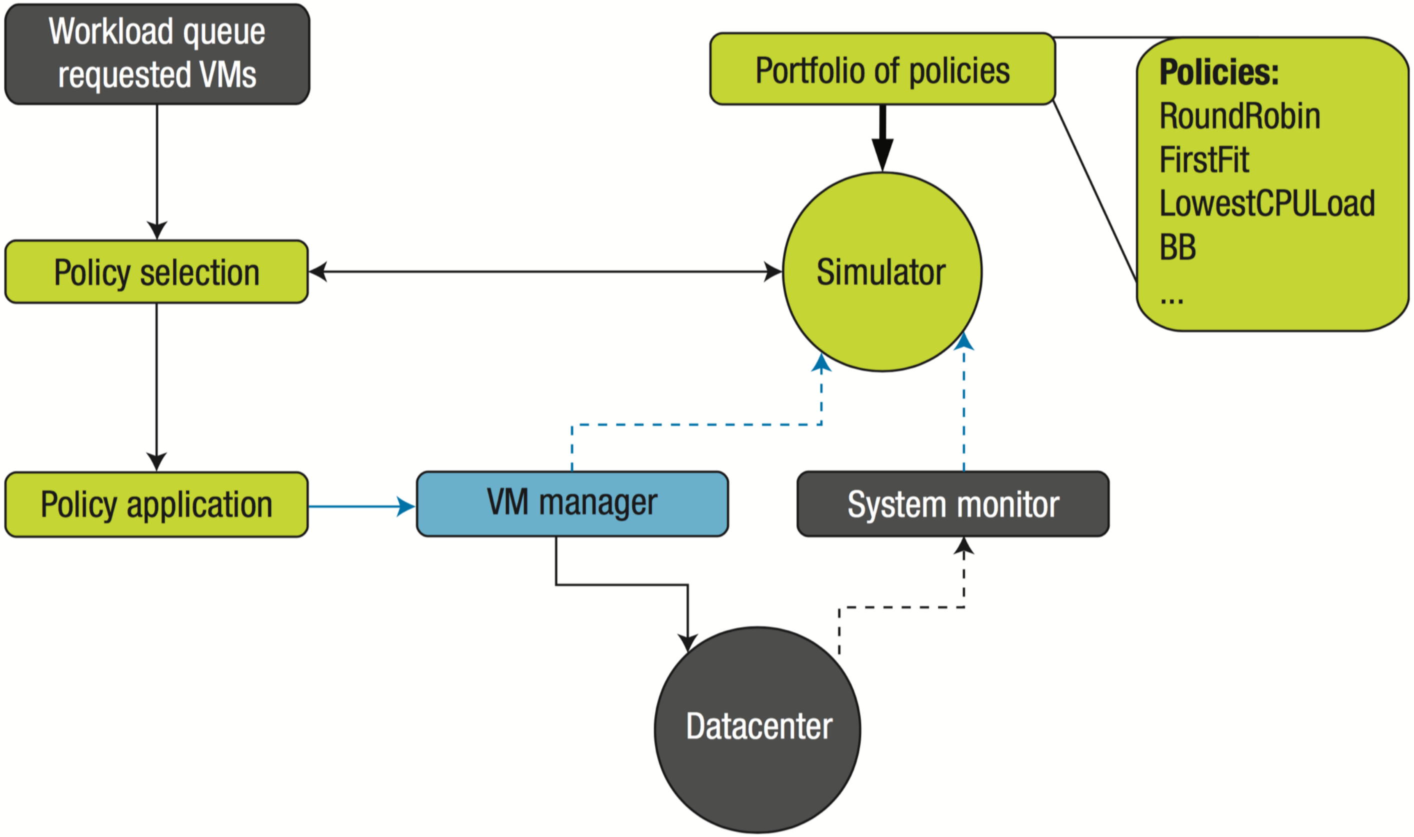
Application

Which changes to the portfolio?

Validate selection immediately

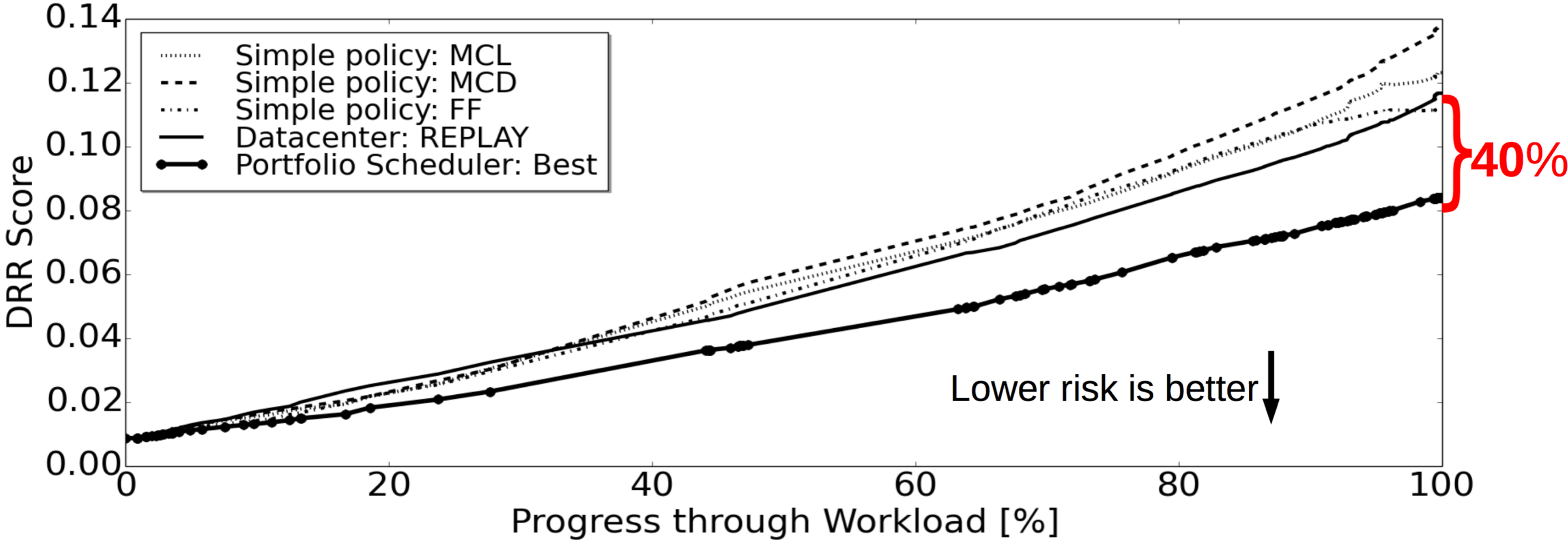
Portfolio Scheduling for VM placement In Datacenters





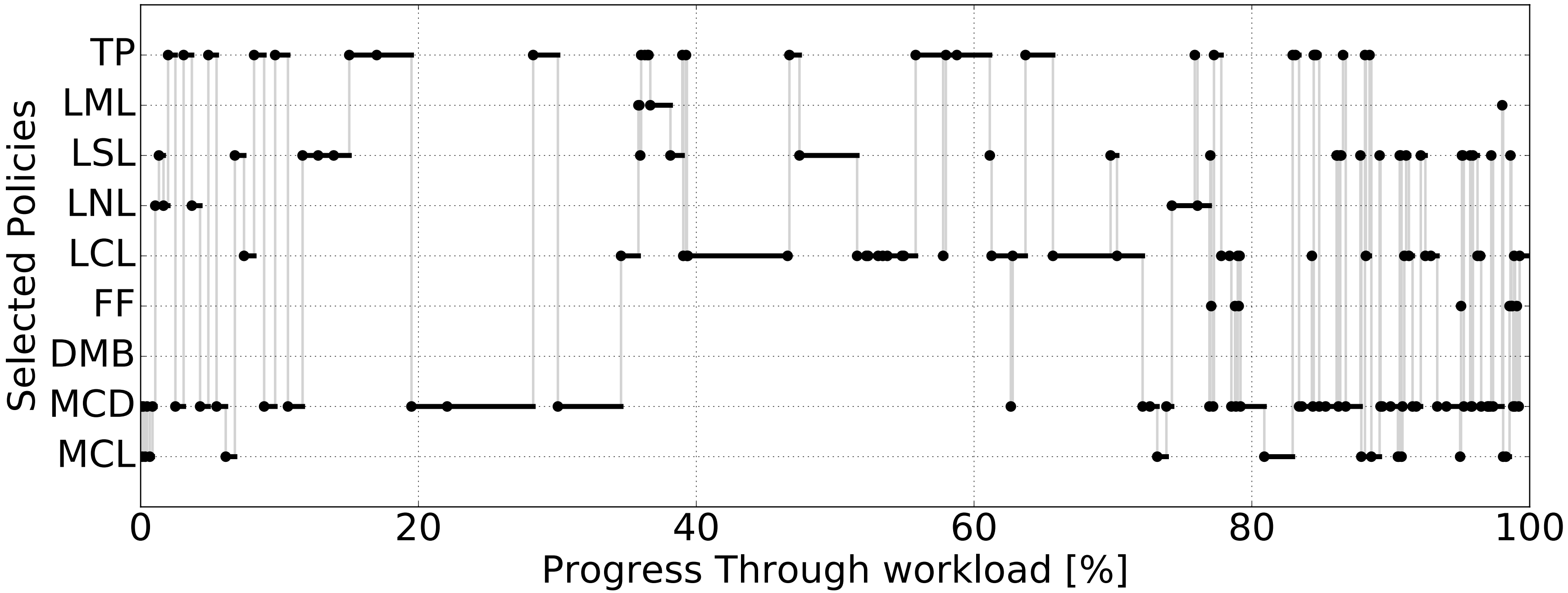
Mnemos: Self-Expressive Management of Business-Critical Workloads in Virtualized Datacenters
In IEEE Computer, 2015.

PS vs its constituent policies



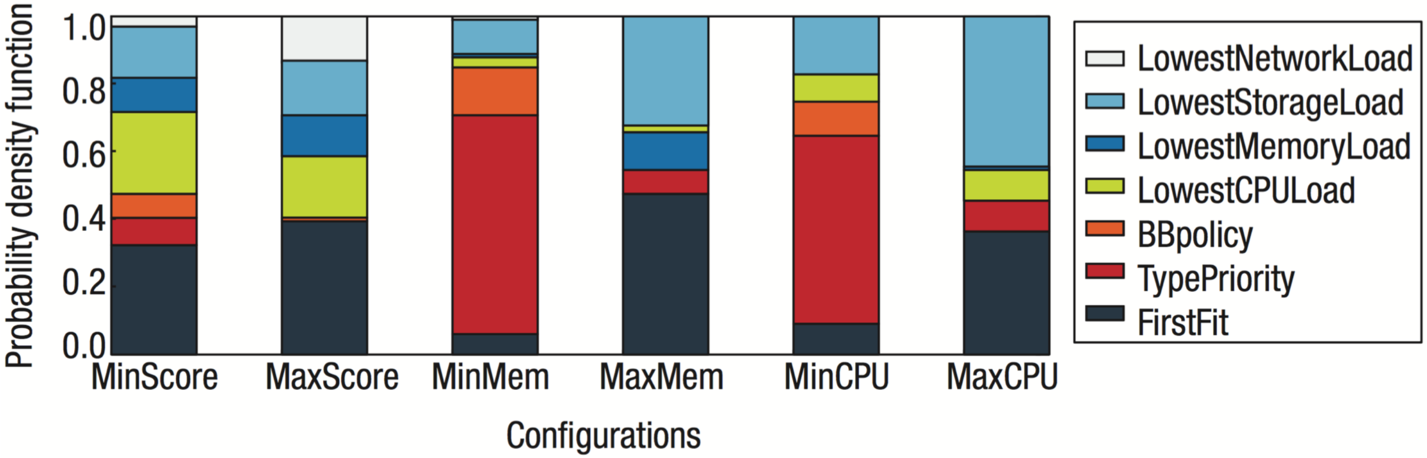
1. Portfolio scheduler achieves the lowest risk of all scenarios.
2. Portfolio scheduler achieves at least **35%** lower DRR compared to individual policies.
3. **40%** lower DRR than commercial production system (REPLAY).

Policy Selection over Time



Portfolio Scheduling for Business-Critical Workloads

- **Policy selection for 6 different utility functions**



Roadmap

Near future

- Implementation of Portfolio Scheduling in OpenDC
- SLA based scheduling

Long term

- Multi Cloud
- Reflection
- Theoretic Boundaries
- Practical Limitations

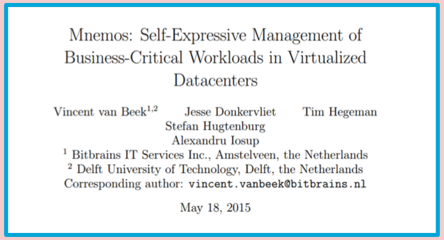
What does OpenDC bring to the table?

1. Datacenter Technology & Methods

Risk Analysis + Management

Heterogeneity

Efficiency → SME Availability

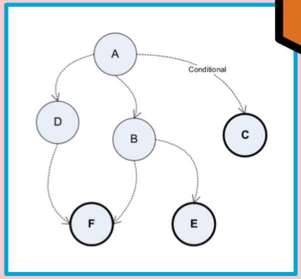


Mnemos: Self-Expressive Management of Business-Critical Workloads in Virtualized Datacenters

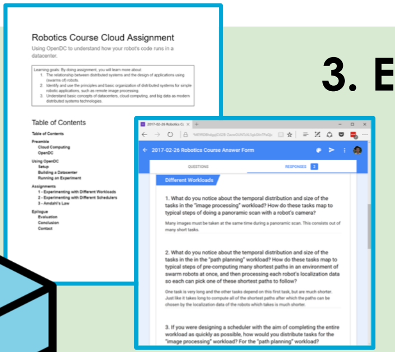
Vincent van Beek[1,2] Jesse Donkervliet Tim Hegeman
Stefan Hugtenburg
Alexandru Iosup

[1] Bitbrains IT Services Inc., Amstelveen, the Netherlands
[2] Delft University of Technology, Delft, the Netherlands
Corresponding author: vincent.vanbeek@bitbrains.nl

May 18, 2015



2. Scientific Methods



Robotics Course Cloud Assignment

Using OpenDC to understand how your robot's code runs in a datacenter.

Learning goals: By doing assignment, you will learn more about:
1. The relationship between distributed systems and the design of applications using (swarms of) robots.
2. Identify and use the principles and basic organization of distributed systems for simple robotic applications, such as remote image processing.
3. Understand basic concepts of datacenters, cloud computing, and big data as modern distributed systems technologies.

Table of Contents

Table of Contents
Preamble
Cloud Computing
OpenDC
Using OpenDC
Setup
Building a Datacenter
Running an Experiment
Assignments
1 - Experimenting with Different Workloads
2 - Experimenting with Different Schedulers
3 - Amdahl's Law
Epilogue
Evaluation
Conclusion
Contact

2017-02-26 Robotics Course Answer Form

Different Workloads

1. What do you notice about the temporal distribution and size of the tasks in the "image processing" workload? How do these tasks map to typical steps of doing a panoramic scan with a robot's camera?

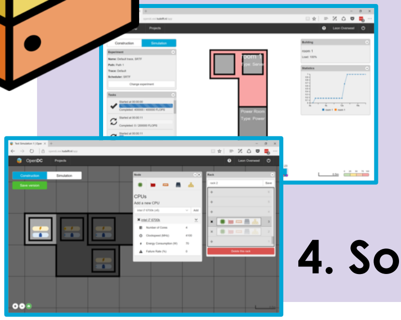
Many images must be taken at the same time during a panoramic scan. This consists out of many short tasks.

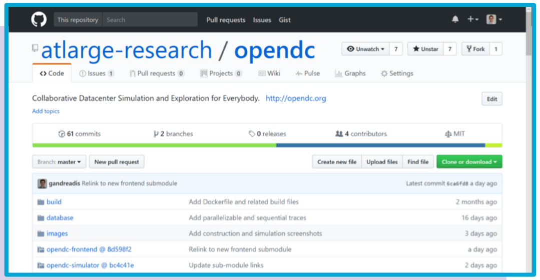
2. What do you notice about the temporal distribution and size of the tasks in the in the "path planning" workload? How do these tasks map to typical steps of pre-computing many shortest paths in an environment of swarm robots at once, and then processing each robot's localization data so each can pick one of these shortest paths to follow?

One task is very long and the other tasks depend on this first task, but are much shorter. Just like it takes long to compute all of the shortest paths after which the paths can be chosen by the localization data of the robots which takes is much shorter.

3. If you were designing a scheduler with the aim of completing the entire workload as quickly as possible, how would you distribute tasks for the "image processing" workload? For the "path planning" workload?

3. Education Practices





atlarge-research / opendc

Collaborative Datacenter Simulation and Exploration for Everybody. http://opendc.org

Add topics

61 commits | 2 branches | 0 releases | 4 contributors | MIT

gandreadis Relink to new frontend submodule — Latest commit 6ca6fd8 a day ago

build	Add Dockerfile and related build files	2 months ago
database	Add parallelizable and sequential traces	16 days ago
images	Add construction and simulation screenshots	3 days ago
opendc-frontend @ 8d598f2	Relink to new frontend submodule	a day ago
opendc-simulator @ bc4c41e	Update sub-module links	2 days ago

4. Software & Data Artifacts

Take home message

Datacenter research is important!

Ask members of the @large team about OpenDC

https://opendc.org