Hochschule Reutlingen
Reutlingen University

Migrating Parallel Applications to the Cloud

Assessing Cloud Readiness based on Parallel Design Decisions

12th Symposium and Summer School On Service-Oriented Computing,
June 24 – June 29, 2018 in Crete, Greece

Stefan Kehrer, Wolfgang Blochinger
Parallel and Distributed Computing Group
Reutlingen University

Universität Stuttgart
Herman Hollerith Zentrum
Hochschule Reutlingen
Reutlingen University

Agenda

- Introduction
- Problem Statement
- Meta Model & Cloud Readiness Assessment
- Teaser: Case Study
- Conclusion

Agenda

- **Introduction**
- Problem Statement
- Meta Model & Cloud Readiness Assessment
- Teaser: Case Study
- Conclusion

Introduction

» High Performance Computing adopts the Cloud!

- Applications for domains such as artificial intelligence, engineering and scientific simulations, …
- Application design is a hard task and requires a deep understanding of parallel processing techniques
- Cloud providers provide offerings optimized for HPC workloads:

"AWS provides unique benefits for **entirely new categories of applications**"

"As the capabilities and performance of AWS have continued to advance, the types of **HPC applications that are running on AWS have also evolved**"
HPC Whitepaper, AWS

"There are **plenty of questions to be answered** in HPC cloud"
HPC Cloud for Scientific and Business Applications: Taxonomy, Vision, and Research Challenges, Netto et al. (2018)

How to migrate parallel applications to the cloud?

HPC Whitepaper, AWS: https://d0.awsstatic.com/whitepapers/Intro_to_HPC_on_AWS.pdf
Netto M, Calheiros R, Rodrigues E, Cunha E, and Buyya R. 2018. HPC Cloud for Scientific and Business Applications: Taxonomy, Vision, and Research Challenges. ACM Comput. Surv. 51, 1, Article 8

Introduction

- Many parallel applications can be deployed to cloud environments without modifications
- However, copying existing applications to the cloud does not enable them to exploit cloud-specific characteristics

High Performance Computing

- Applications to solve large problems
- Focus on parallel performance in terms of speedup & parallel efficiency
- Limited scalability; parallel efficiency is a major issue

Cloud Computing

- Typically web-based multi-tier applications
- Focus on cloud-specific properties to maximize benefits
- Tend to be perfectly scalable; parallel efficiency is not an issue

HPC and Cloud Computing follow different goals!

Maximize parallel efficiency vs. maximize cloud-specific benefits

Agenda

- Introduction
- **Problem Statement**
- Meta Model & Cloud Readiness Assessment
- Teaser: Case Study
- Conclusion

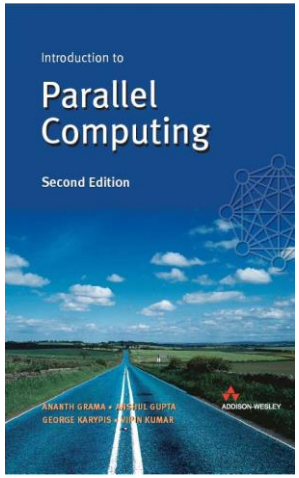
Problem Statement

– Different goals lead to different design guidelines:

High Performance Computing

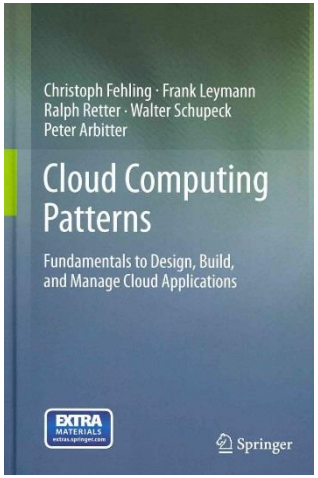
» Parallel computing principles guide application design

» Several design decisions have to be made



Cloud Computing

» Cloud-native application design: Isolated state, Distribution, Elasticity, Automated management, and Loose coupling (IDEAL)



http://www-users.cs.umn.edu/~karypis/parbook/Images/bcover.jpg
https://images-na.ssl-images-amazon.com/images/I/4124esUdZPL._SX330_BO1,204,203,200_.jpg

Problem Statement

– Different goals lead to different design guidelines:

High Performance Computing

» Parallel computing principles guide application design

» Several design decisions have to be made

Cloud Computing

» Cloud-native application design: Isolated state, Distribution, Elasticity, Automated management, and Loose coupling (IDEAL)

How to maximize parallel efficiency and cloud readiness?

How to identify conflicting properties?

http://www-users.cs.umn.edu/~karypis/parbook/Images/bcover.jpg
https://images-na.ssl-images-amazon.com/images/I/4124esUdZPL._SX330_BO1,204,203,200_.jpg

Problem Statement

– Different goals lead to different design guidelines:

High Performance Computing

» Parallel computing principles guide application design

» Several D

Cloud Computing

» Cloud-native application design: Isolated

ated

DEAL)

Understanding the implications of parallel design decisions on an application's cloud readiness is a key requirement for successful cloud migration!

Design decisions lead to specific properties that required for maximu parallel efficiency!

h these properties your application reaches imum cloud readiness!

Assess the cloud readiness of parallel applications!

How to maximize parallel efficiency and cloud readiness?

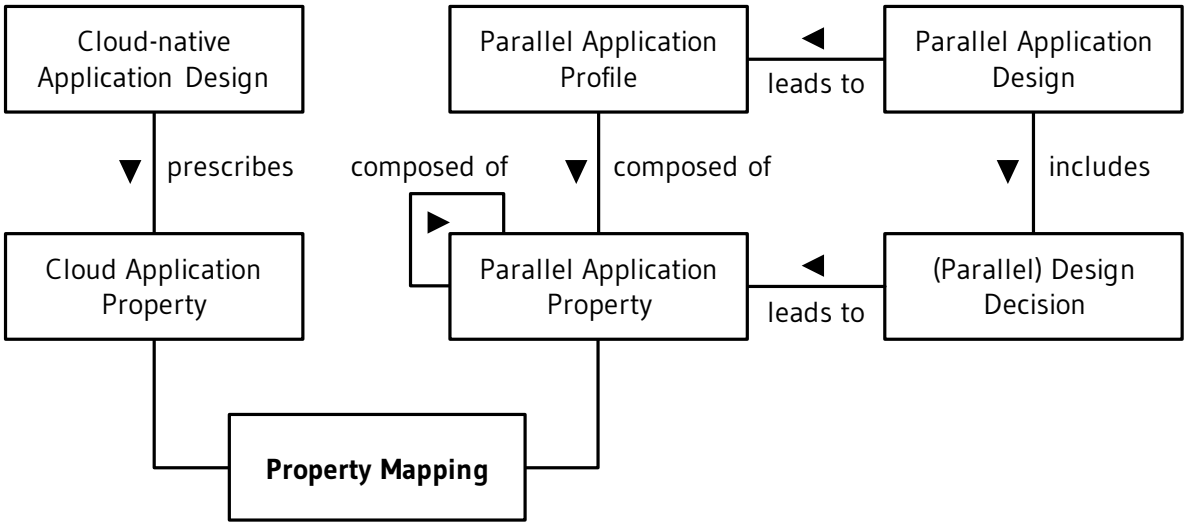
How to identify conflicting properties?

Agenda

- Introduction
- Problem Statement
- **Meta Model & Cloud Readiness Assessment**
- Teaser: Case Study
- Conclusion

Meta Model & Cloud Readiness Assessment

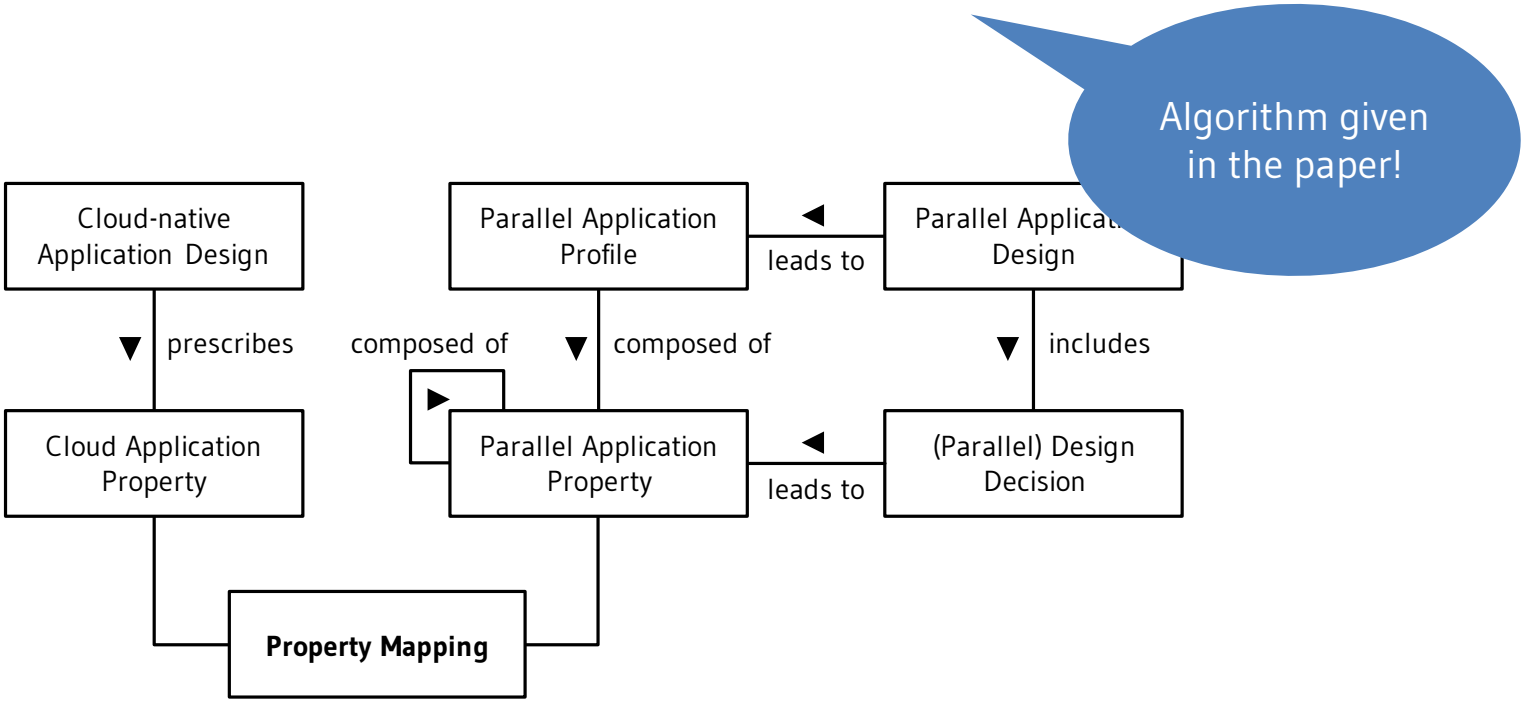
- Cloud-native design guidelines and cloud pattern languages are driven by prescriptive *cloud application properties* (IDEAL properties)
- *Cloud-native application design* describes an ideal cloud application, i.e., an application with maximum cloud readiness
- *Parallel design decisions* lead to *parallel application properties*
- A parallel application can be characterized by a set of *parallel application properties*
- Parallel applications are described as *parallel application profile*



Meta model that enables the assessment of a parallel applications

Meta Model & Cloud Readiness Assessment

- *Property mappings* allow us to qualify the impact on cloud readiness as positive (+) of negative (-)
- Property mappings describe the conceptual fitness of mapping a *cloud application property* to a *parallel application property*
- Instantiating the meta model allows us to assess the cloud readiness of a parallel application
- Cloud readiness assessment procedure calculates the cloud readiness & provides hints for optimization



Meta model that enables the assessment of a parallel applications

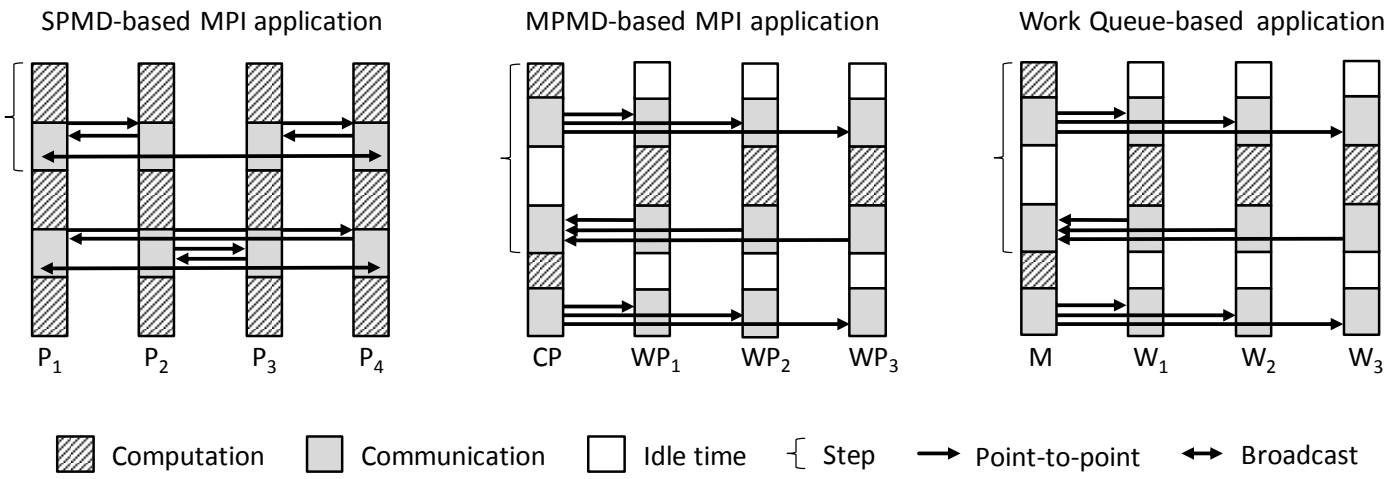
Agenda

- Introduction
- Problem Statement
- Meta Model & Cloud Readiness Assessment
- **Teaser: Case Study**
- Conclusion

Teaser: Case Study

Applied in cancer research and research into Alzheimer's disease.

- 3 parallel applications for *Replica Exchange Molecular Dynamics (REMD)* are assessed:



Computational steps and interaction patterns of exemplary parallel applications

Table 1 Parallel application properties of exemplary applications and their impact on cloud readiness

Design Decision	SPMD-based MPI app.	MPMD-based MPI app.	Work Queue-based app.
Task Generation	Static (e^-)	Dynamic (e^+)	Dynamic (e^+)
Task Mapping	Static (e^-, i^-)	Cent. Dynamic (e^+, i^+)	Cent. Dynamic (e^+, i^+)
Task Interaction	Unstructured (l^-)	Structured (l^+)	Structured (l^+)
Communication Model	Synchronous (l^-)	Synchronous (l^-)	Asynchronous (l^+)
Size of Associated Data	Small (d^+, e^+, i^+)	Small (d^+, e^+, i^+)	Small (d^+, e^+, i^+)

Agenda

- Introduction
- Problem Statement
- Meta Model & Cloud Readiness Assessment
- Teaser: Case Study
- **Conclusion**

Conclusion

Conclusion:

- We presented a systematic approach to assess the cloud readiness of parallel applications
- We analyzed application properties from both domains with respect to their conceptual fitness
  - The results are captured in in a reusable manner
- Parallel applications can be analyzed with respect to their cloud readiness at the design level
  - We enable better decisions if a certain application should be ported to the cloud or not (and the effort required for migration / benefits gained)

Future work:

- We plan to investigate several parallel application classes with respect to their cloud readiness
- Specifically, parallel applications with unpredictable resource requirements benefit from elasticity and on-demand resource provisioning

Thank You

Thank You

Stefan Kehrer
Parallel and Distributed Computing Group
Reutlingen University
stefan.kehrer@reutlingen-university.de