Global scheduling against waste

Laércio LIMA PILLA

pilla@lri.fr
Agenda

Motivation
Distributed scheduling
Scheduling library
Scheduling benchmarks
Motivation

Scheduling is "[...] a mechanism or policy used to **efficiently** and **effectively** manage the access to and use of a **resource** by its various **consumers**" [Casavant & Kuhl, 1988]

Other names: *load balancing, mapping, orchestration, resource allocation*...
Motivation

More complex and larger applications and platforms → more performance problems

More parallelism → more chances for load imbalance issues

More complex hierarchies and networks → more chances for locality issues
Why is the efficient utilisation of resources important?

Example: 5% of 20% of the global energy consumption = 1% of the global energy consumption

Source: "How to stop data centres from gobbling up the world’s electricity"
https://www.nature.com/articles/d41586-018-06610-y
Motivation

Global scheduling against waste

Improve the utilisation of computing resources
Tackle the challenges related to scalability, locality, and heterogeneity

Against global scheduling waste

Improve the ways we do our research in the domain
Reuse and adapt the knowledge, ease testing, reuse code
Distributed scheduling

Vinícius's talk +

Differences between applied and theoretical research in [distributed] scheduling

**HPC:** take a load imbalanced scenario, propose new algorithms [without convergence guarantees], experiment in a real system

**Theory:** define a model, define their bounds and propose algorithms, maybe run simulations

Problem: wasting part of the state of the art
Distributed scheduling

Internal LRI project (+ Johanne Cohen): Explore/Trade/Sched

Explore the common and different parts of these two different scheduling contexts in order to find their trade-offs and bring them closer together

Objectives

To understand the differences.
To delineate the limits of current algorithms.
To adapt scheduling algorithms to scale.
To design new algorithms and models.
To make available all code & datasets.
Problem: Complex scheduling algorithms → Hard to reuse between [runtime] systems, to adapt, to debug, to reproduce results

MOGSLib

Meta-programming-Oriented Global Scheduler Library

Extensible framework for modular schedulers

https://github.com/ECLSScheduling/MOGSLib
Before MOGSLib

Runtime System

- Global Scheduler
- Scheduling Context
- Scheduling Policy

- Functionalities
- Data Structures
After MOGSLib

Global Scheduler Library

- Scheduling Policy
- Global Scheduler
- Scheduling Context
- Runtime Adapter

Requires functionalities

Employs

Runtime System

Native Global Scheduler

Functionalities

Data structures

Initialize

Forward

Indirect Access

Scheduling library
Scheduling library

Reuse with MOGSLib

Schedulers
- LPT
- BinLPT
- Numa+LPT
- CommAware

Contexts
- Charm LPT
- Charm Comm
- OpenMP Workload
- OpenMP HwLoc + Workload
- Test DummyComm

Adapters
- Charm Centralized
- Charm Distributed
- OpenMP on-demand
- Testing Benchmark
Scheduling library

Performance results
Running experiments with scheduling algorithms can be hard

- trade-off between abstraction and complexity (from benchmarks to applications)
- lack of reproducibility (unavailable code, missing experimental methodology)
- scarcity
- absence of scheduling control

Problem: waste of time and effort to validate and test scheduling algorithms
3BEARS: Broad Bundle of BEnchmarks for Allocation of Resources and Scheduling in Parallel and Distributed Computing

Submitted PHC Project with Oguz Kaya (LRI), Florina Ciorba (Unibas, CH)

Objectives

- Provide a shared resource allocation and scheduling knowledge base.
- Facilitate the comparison of scheduling algorithms and the reproduction of scheduling results.
- Reduce the entry barrier for young scientists into the scheduling domain.
- Improve the efficiency of applications and the utilisation of parallel computing resources.
- Serve as a means to demonstrate the benefits of scheduling algorithms.
Concluding remarks

5 ways to work together

1. Integration of "old" scheduling algorithms
2. Proposal of new scheduling algorithms
3. New sources of scheduling information (runtime, compilation, tracing)
4. Use cases (benchmarks, applications)
5. Experimental methodologies
Global scheduling against waste
Against global scheduling waste

Laércio LIMA PILLA
pilla@lri.fr

Thank you.
Icons from Noun Project

Person - Created by mikicon - https://thenounproject.com/icon/737677/
Milky Way - Created by Setyo Ari Wibowo - https://thenounproject.com/icon/1217198/
Atom - Created by Setyo Ari Wibowo - https://thenounproject.com/icon/1224254/
Sun - Created by ProSymbols - https://thenounproject.com/icon/641671/
Raining - Created by ProSymbols - https://thenounproject.com/icon/641687/
Earthquake - Created by arif fajar yulianto - https://thenounproject.com/icon/1240012/
Ram - Created by Setyo Ari Wibowo - https://thenounproject.com/icon/968177/
Processor - Created by Setyo Ari Wibowo - https://thenounproject.com/icon/968171/

Emojis from Microsoft @ https://emojipedia.org/