Resource and task management tools for physics applications

A.V. Bogdanov bogdanov@csa.ru
A.B. Degtyarev deg@csa.ru
V.Yu. Gaiduchok gajduchok@cc.spbu.ru
I.G. Gankevich igankevich@cc.spbu.ru
V.V. Korkhov vladimir@csa.ru
N.V. Yuzhanin yuzhanin@cc.spbu.ru

NEC'2015

Agenda

Introduction

Examples

Approaches

Traditional approach

New approach

New API

Conclusions

Introduction

- The task of heterogeneous resources management is one of the most difficult and important tasks today.
- Administrators should set resource sharing policies that will meet different requirements of different groups of users.
- Users want to compute their tasks fast while organizations want their resources to be utilized efficiently.
- Traditional schedulers do not allow administrator to efficiently solve these problems in that way.

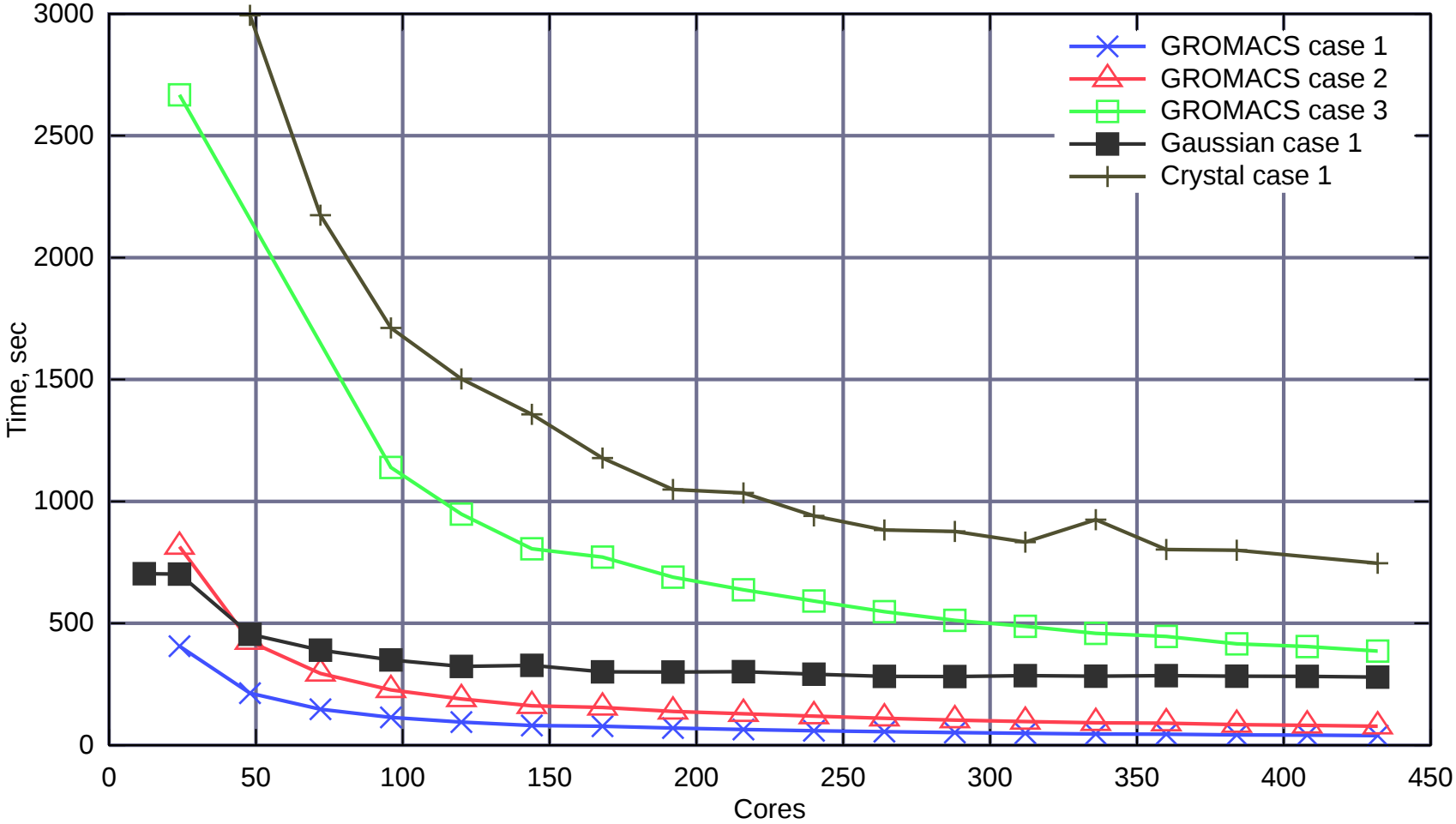
Introduction

- Users run different applications with different scalability.
- They can use different libraries which have different performance.
- Tasks of the same application can use diffenet modules and vary greatly in requirements for computational hardware.

OpenFOAM

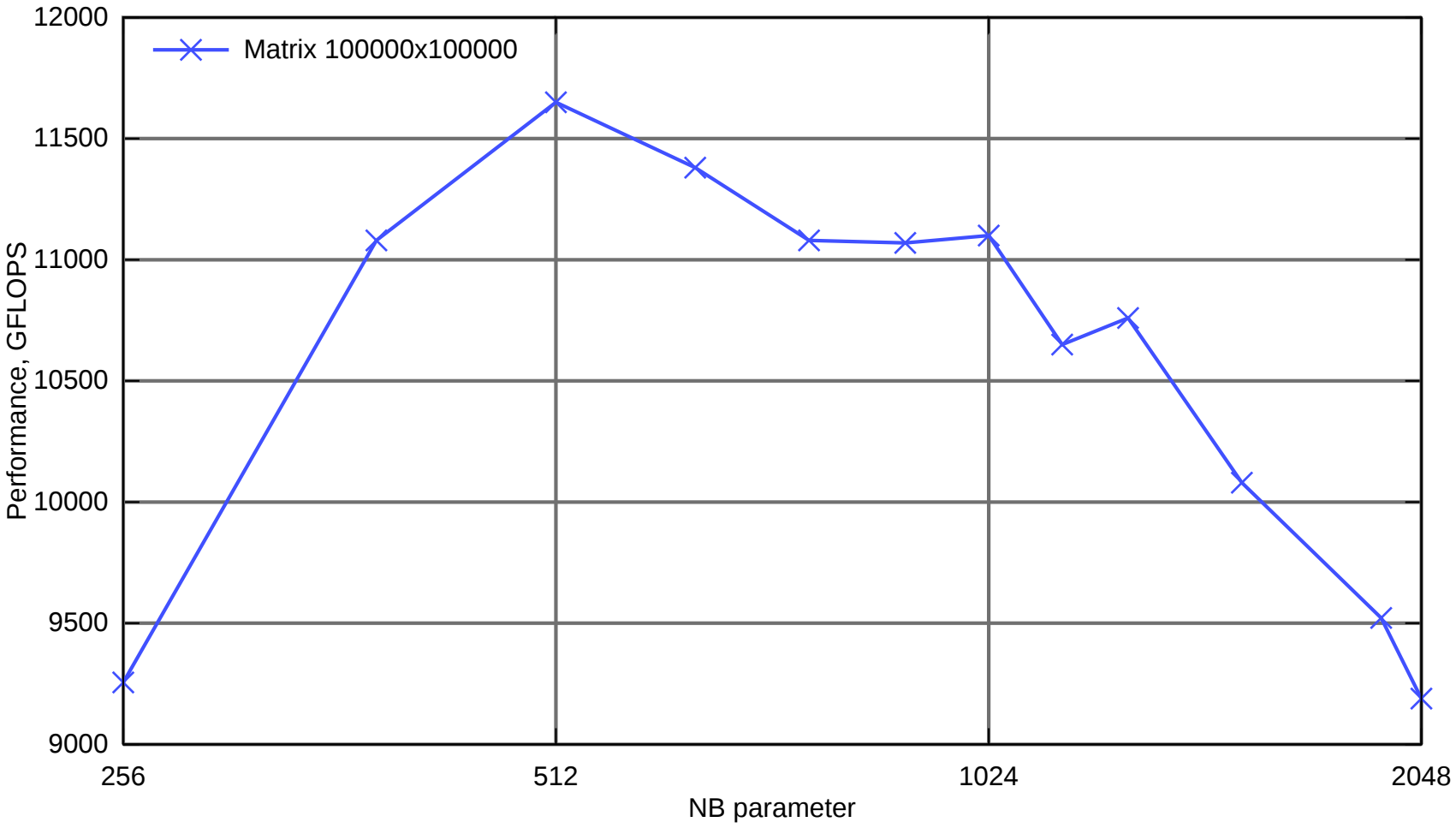
GROMACS FAST. FLEXIBLE. FREE.

Examples



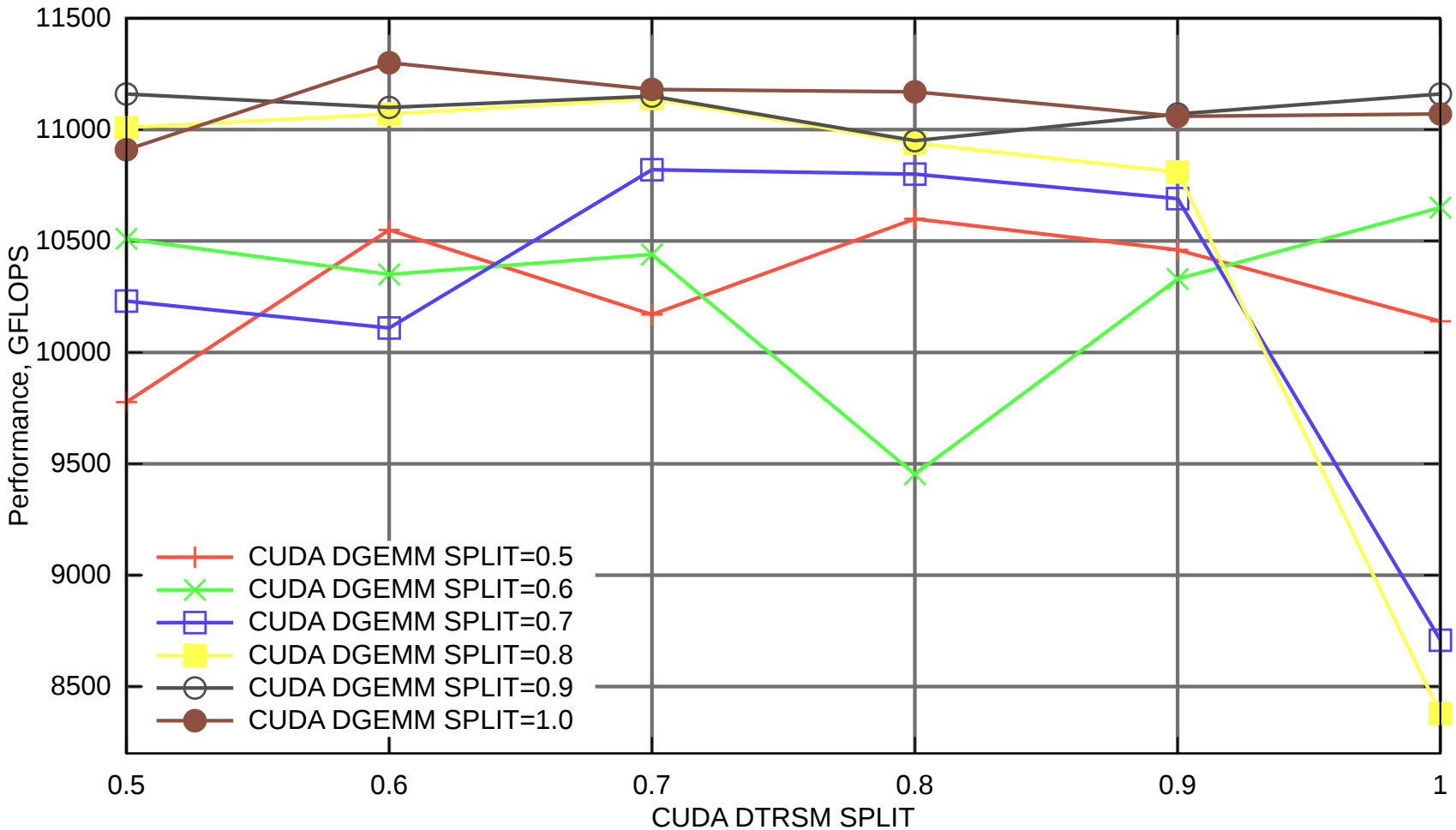
Application tests

Examples



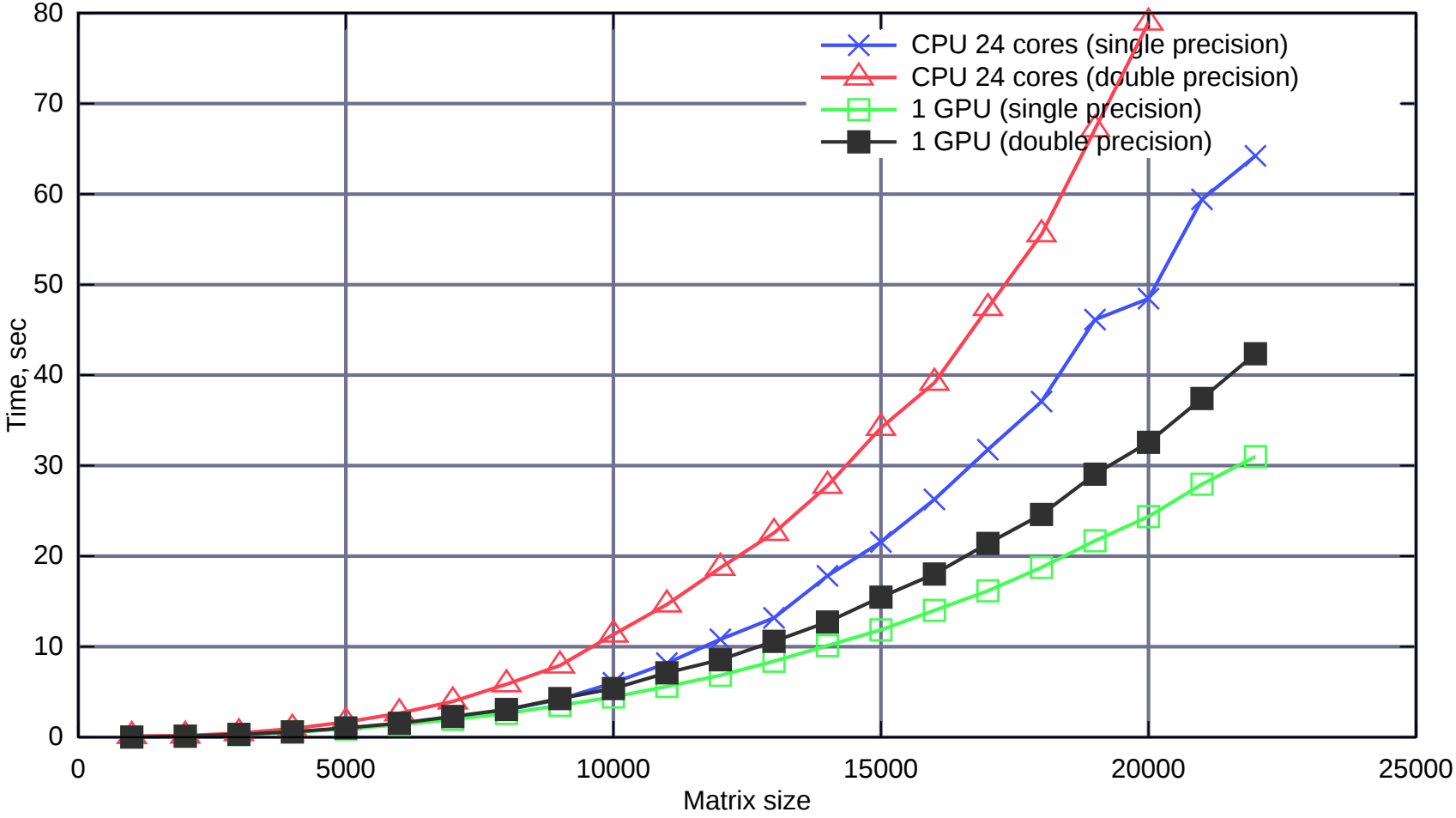
GPGPU LINPACK benchmark

Examples



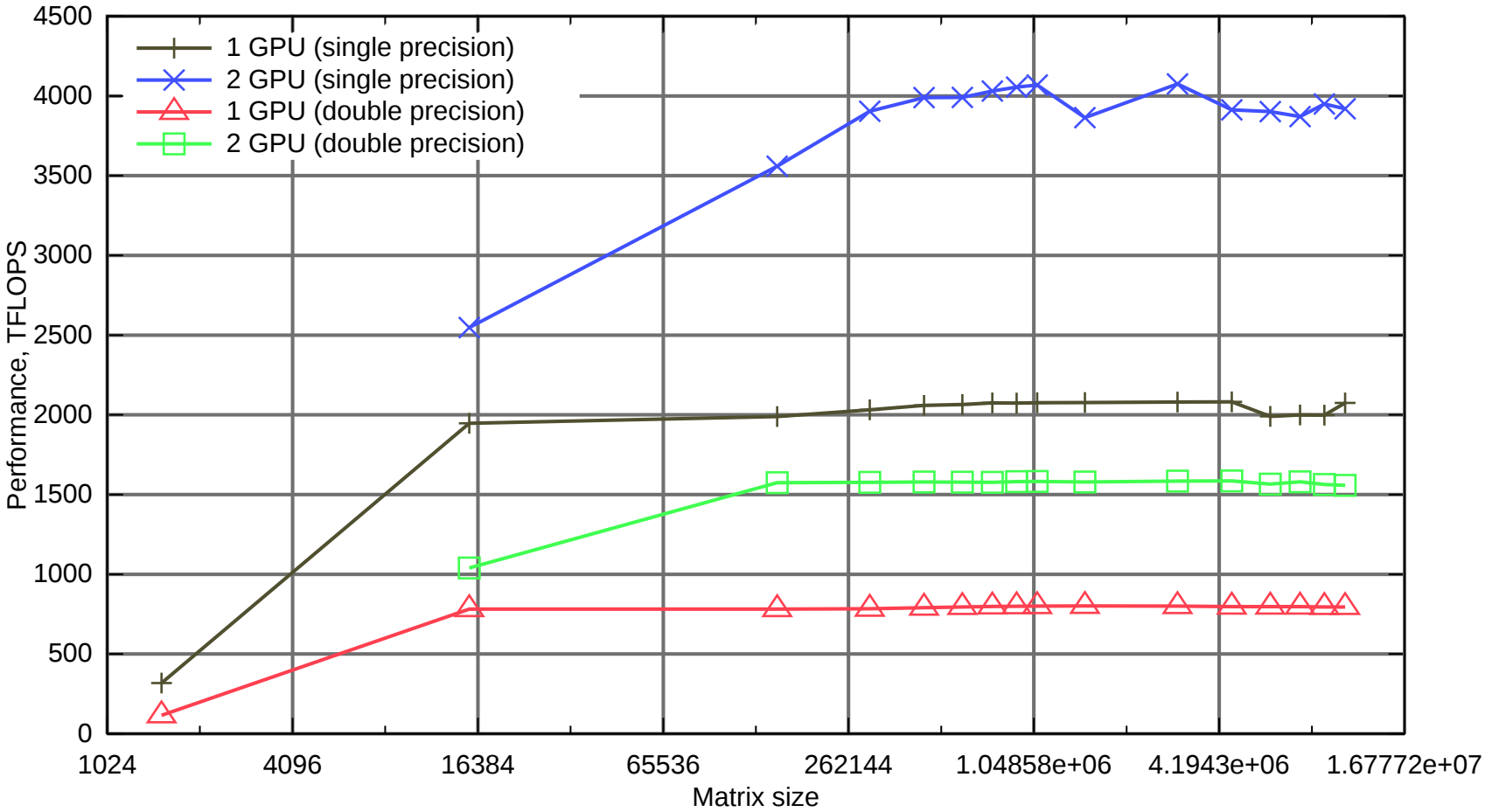
GPGPU LINPACK benchmark

Examples



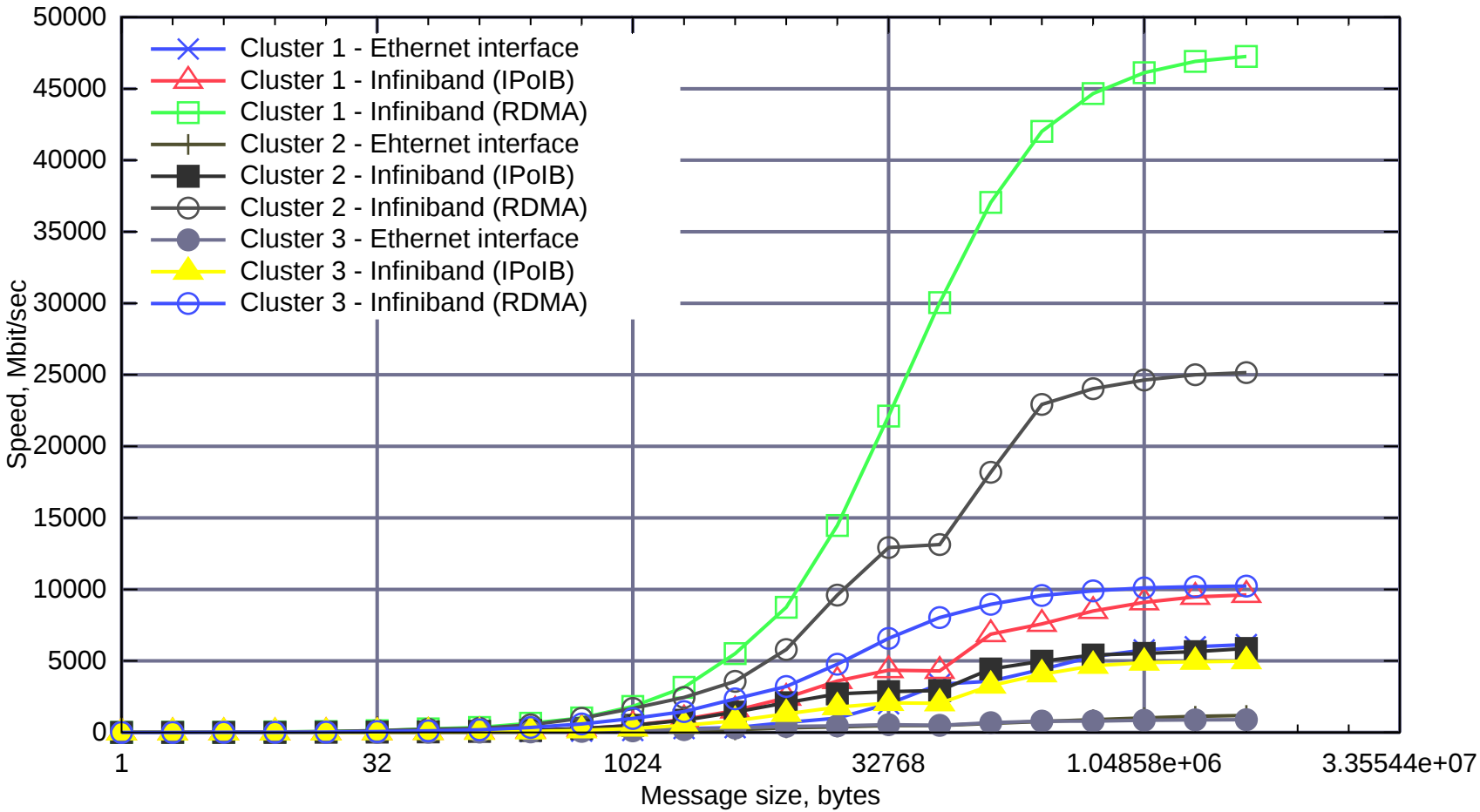
MATLAB task

Examples



Nbody test

Examples

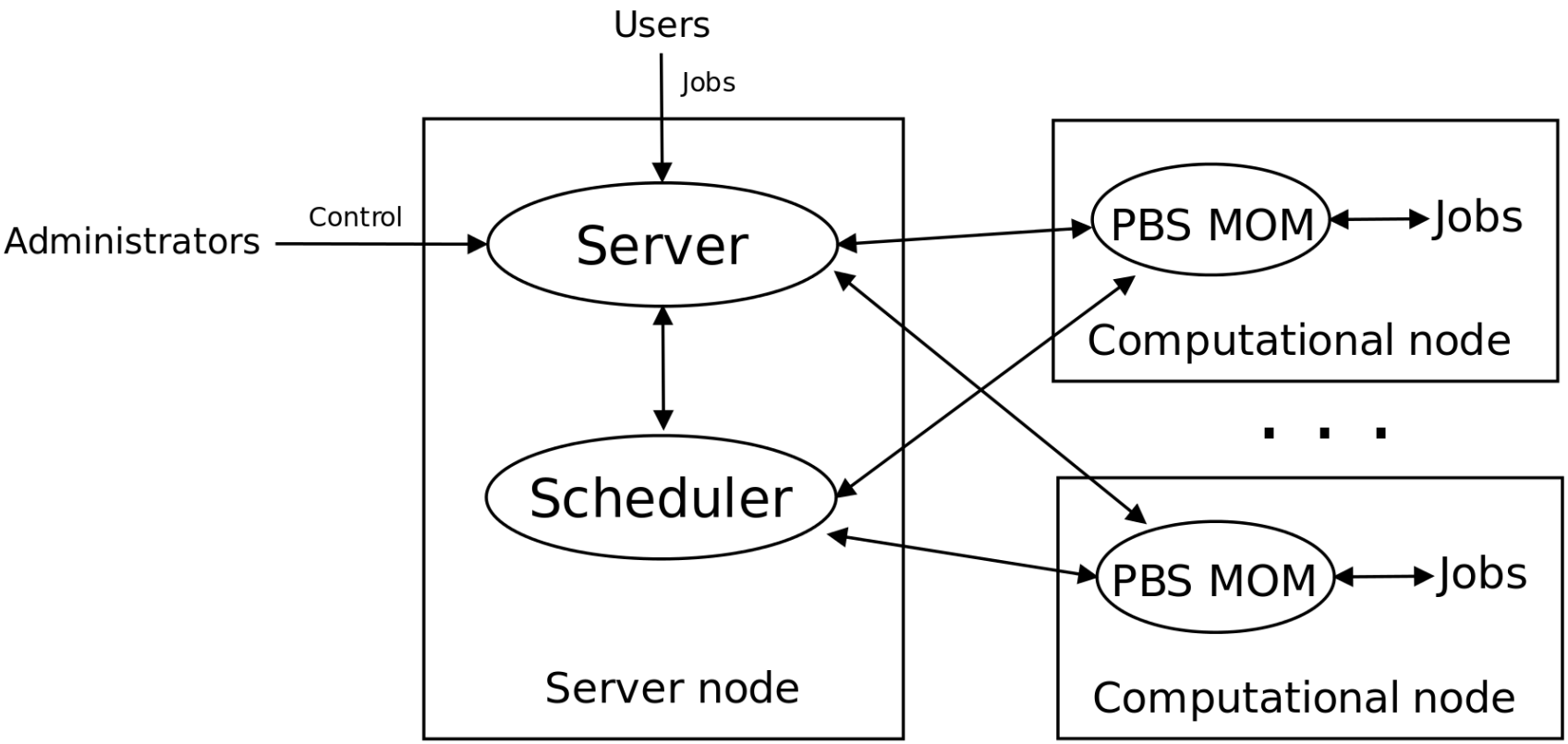


Networking tests

Approaches

- Classical management system. *Dynamic reallocation.*
- No management system. *Maintenance.*
- Single system image. *Fault tolerance.*
- Cloud. *Overheads.*
- Other approaches.

Traditional approach

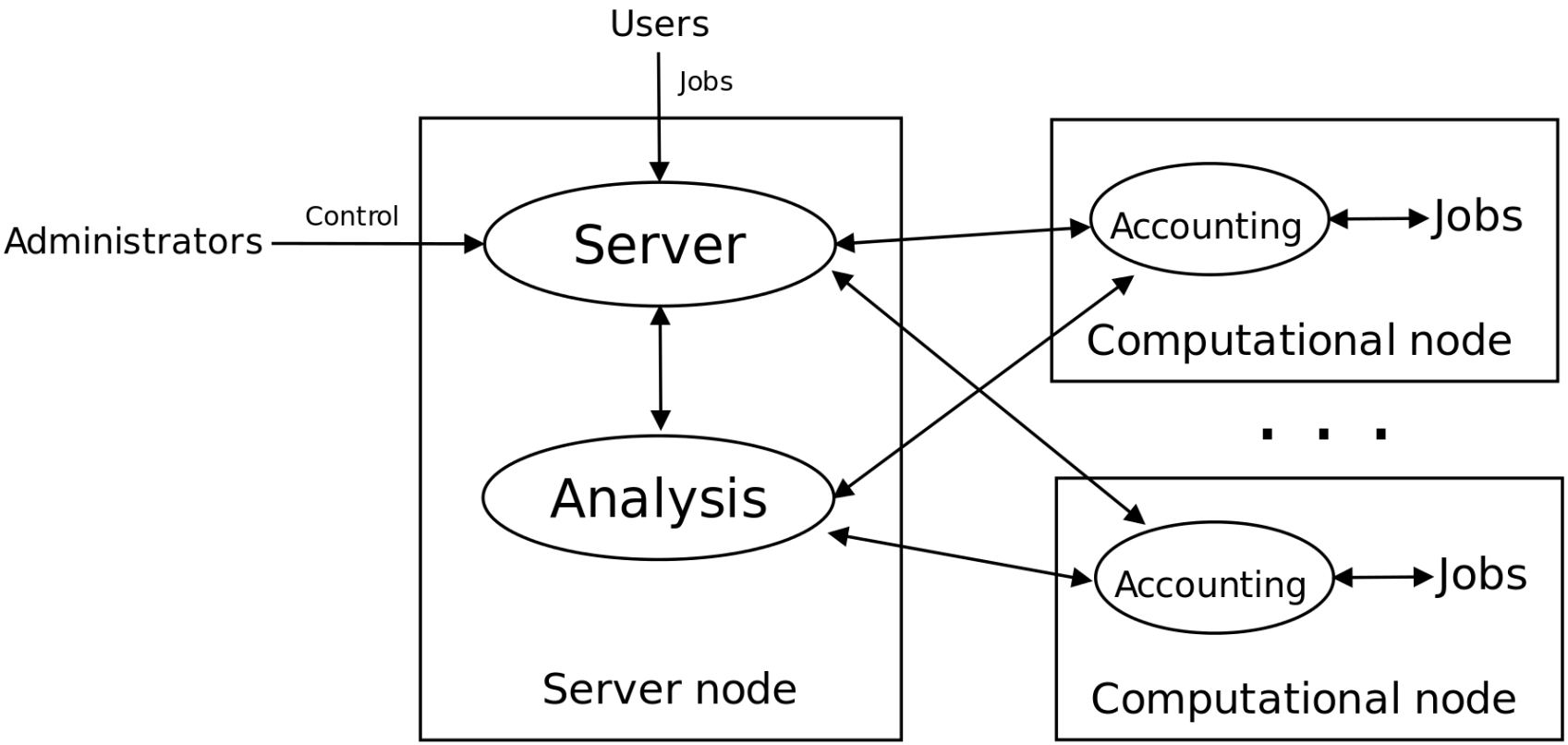


PBS scheme

Traditional approach

- Implements the classical scheme (Portable Batch system).
- Several implementations (TORQUE, PBS Professional, etc).
- It usually does not monitor dynamic resource load.
- Resource reservation can be changed only by user or administrator.
- Scarce accounting information.

New approach



New approach based on PBS scheme

New approach

- Dynamic resource reallocation.
- Profiling, detailed accounting and monitoring.
- Flexible resource reservation.
- Predictions module.
- User rating.
- Native API.
- Modules can be used within existing PBS.
- Small overheads.
- Especially beneficial in case of underload or overload.

New API

- The described system is designed for applications that use traditional APIs.
- Effective solution is a new API.
- Such API implies tight cooperation with the scheduler.
- Dynamic resource reallocation could solve the resource utilization problem.
- API implies step by step resource allocation up to allowable maximum with detailed monitoring.
- A special algorithm is used in order to orchestrate the nodes of the cluster for efficient network communication.

Conclusions

- Resource management for scientific computations sometimes can be challenging.
- Dynamic resource reallocation could lead to effective resource utilization.
- Scheduling applications can be done using the described approach.
- Transition to the new API can take time, but it could be considered as a way to improve utilization of a heterogeneous complex.

Questions

Thank you!