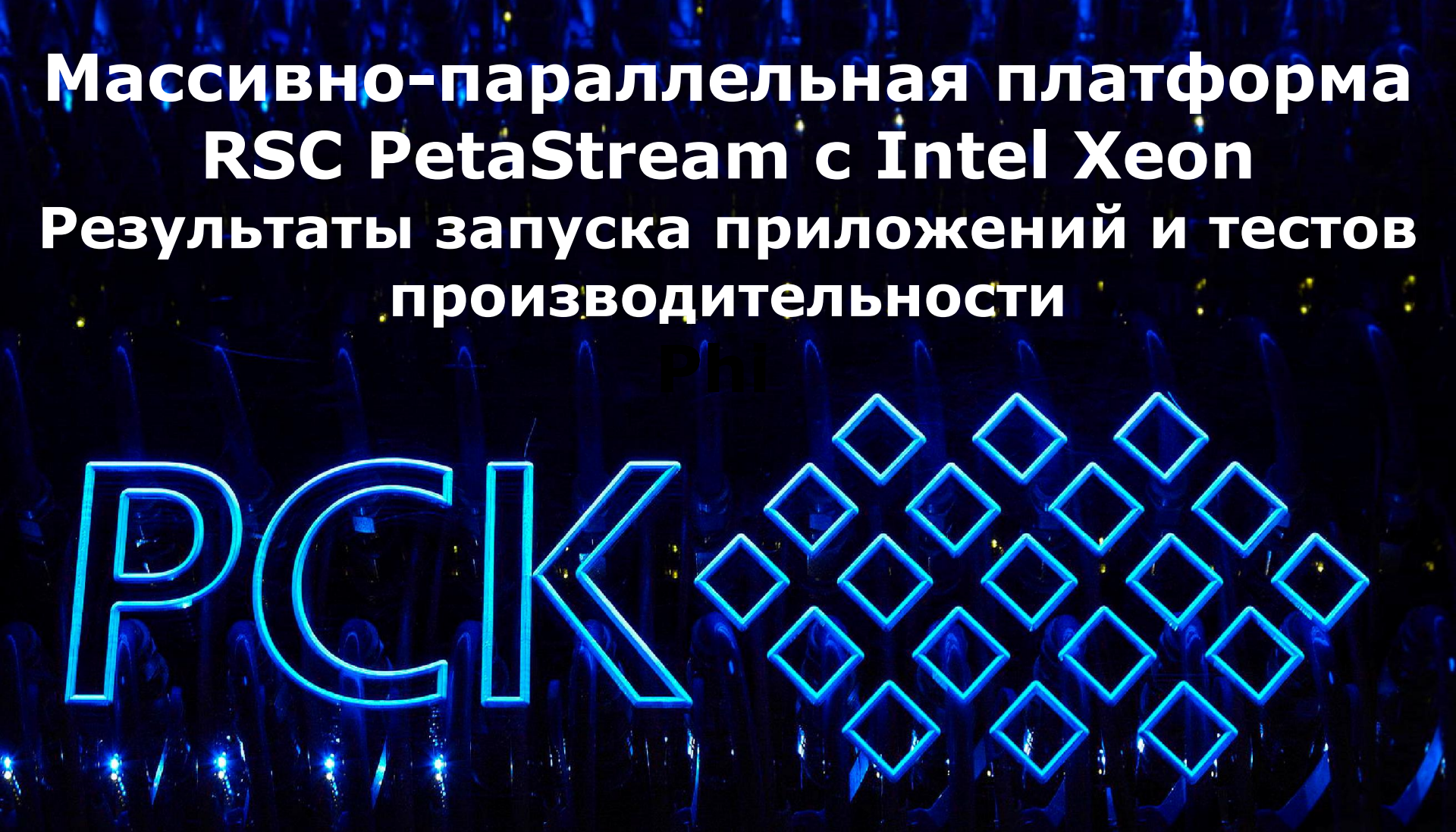
Массивно-параллельная платформа RSC PetaStream с Intel Xeon Phi

Результаты запуска приложений и тестов производительности



Александр Московский, генеральный директор

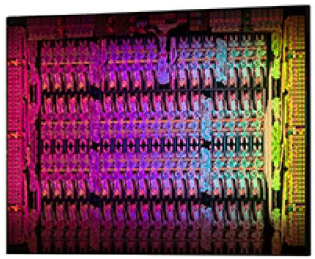
Группа компаний РСК

ПаВТ, Март 2014 г.

Мотивация и требования клиентов

- **Массивно-параллельная система для удовлетворения потребностей экзафлопного уровня**
  - ⇒ Система для процессоров с большим количеством ядер
  - ⇒ Архитектура, поддерживающая масштабирование вычислительных ресурсов, систем хранения данных и сетевых подсистем до решений экзафлопного уровня
- **Защита инвестиций в ПО**
  - ⇒ На базе стандартной архитектуры x86
  - ⇒ Использование существующих моделей программирования, возможность использования имеющихся приложений
  - ⇒ Защита инвестиций в оптимизацию ПО для будущих многоядерных платформ
- **Гибкость для соответствия высоким требованиям клиентов**
  - ⇒ Широкий выбор вариантов сетевых соединений и топологий
  - ⇒ Поддержка возможности инновационного хранения данных и широкие возможности соединений
- **Энергоэффективность**
  - ⇒ Построена на базе проверенной технологии жидкостного охлаждения РСК
  - ⇒ Эффективная и инновационная подсистема питания
  - ⇒ Управление системой и мониторинг

Архитектура RSC PetaStream



Процессор
- 61 ядро x86 / 244 потока
- > 1,2 ТФЛОПС (пик. производительность)
- 352 ГБ/с пик. пропускная способность
- 30 МБ общей кеш-памяти

Вычислительный узел
- Один процессор
- 16 ГБ памяти
- 64 Гбит/с пропускная способность подсистемы ввода/вывода
- Linux µOS

Вычислительный модуль
- Несколько вычислительных узлов
- > 200 Гбит/с пропускная способность подсистемы ввода/вывода
- Жидкостная система охлаждения
- Встроенная система управления
- Эффективное электропитание

Шкаф
- Более 1,2 ПФЛОПС (пик. производительность)
- 249 тысяч потоков / 1024 узла
- До 400 кВт
- Встроенная система управления
- Занимаемая площадь 1 м²

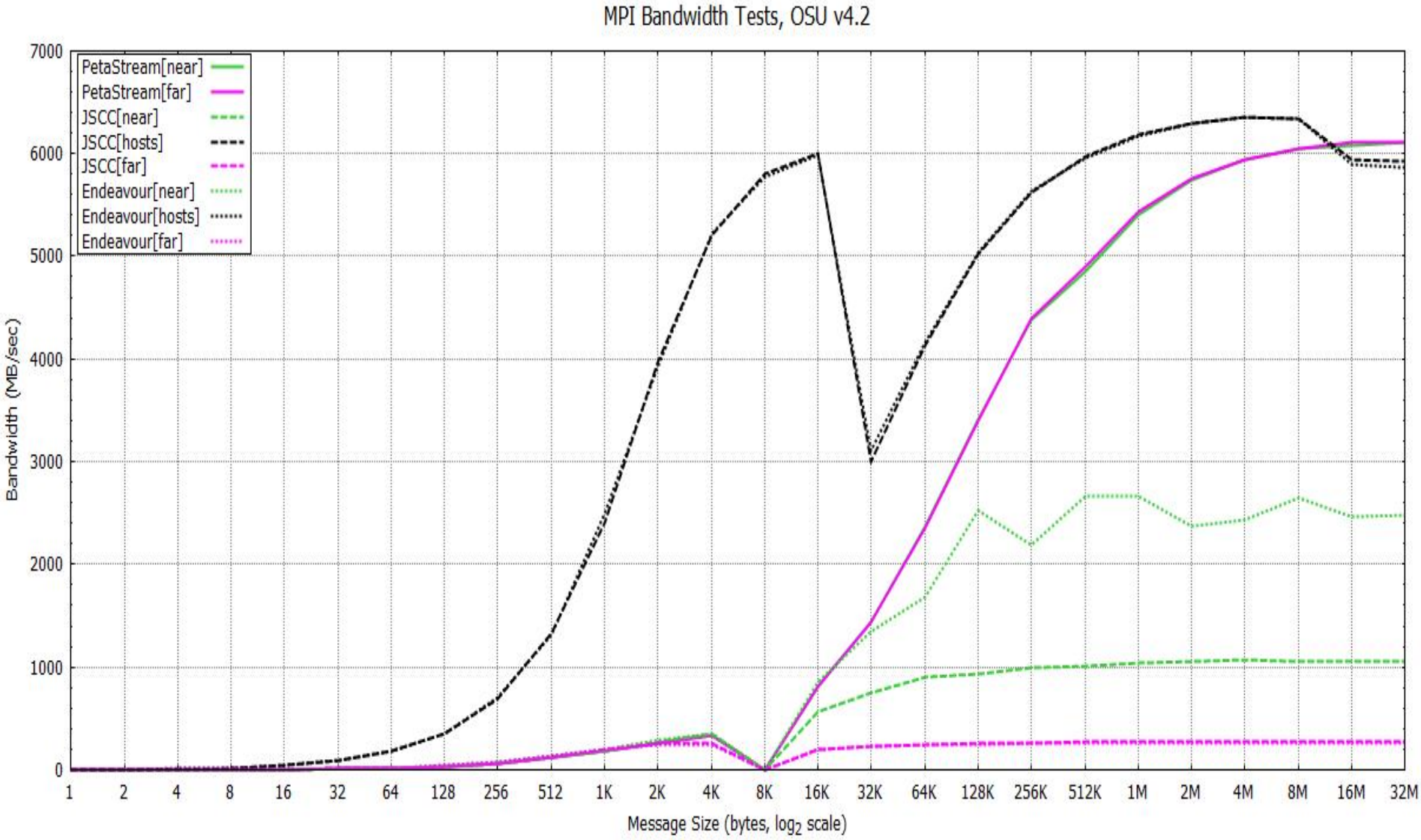
Система
- Идеальна для создания суперкомпьютеров экзафлопного диапазона
- Надежная жидкостная система охлаждения РСК
- Линейно масштабируемая, в соответсвии с требованиями заказчиков
- Гибкие сетевые возможности
- Основана на широко доступных стандартных компонентах

intel inside XEON PHI
intel inside XEON
intel inside SERVER BOARD
Intel Cluster Ready

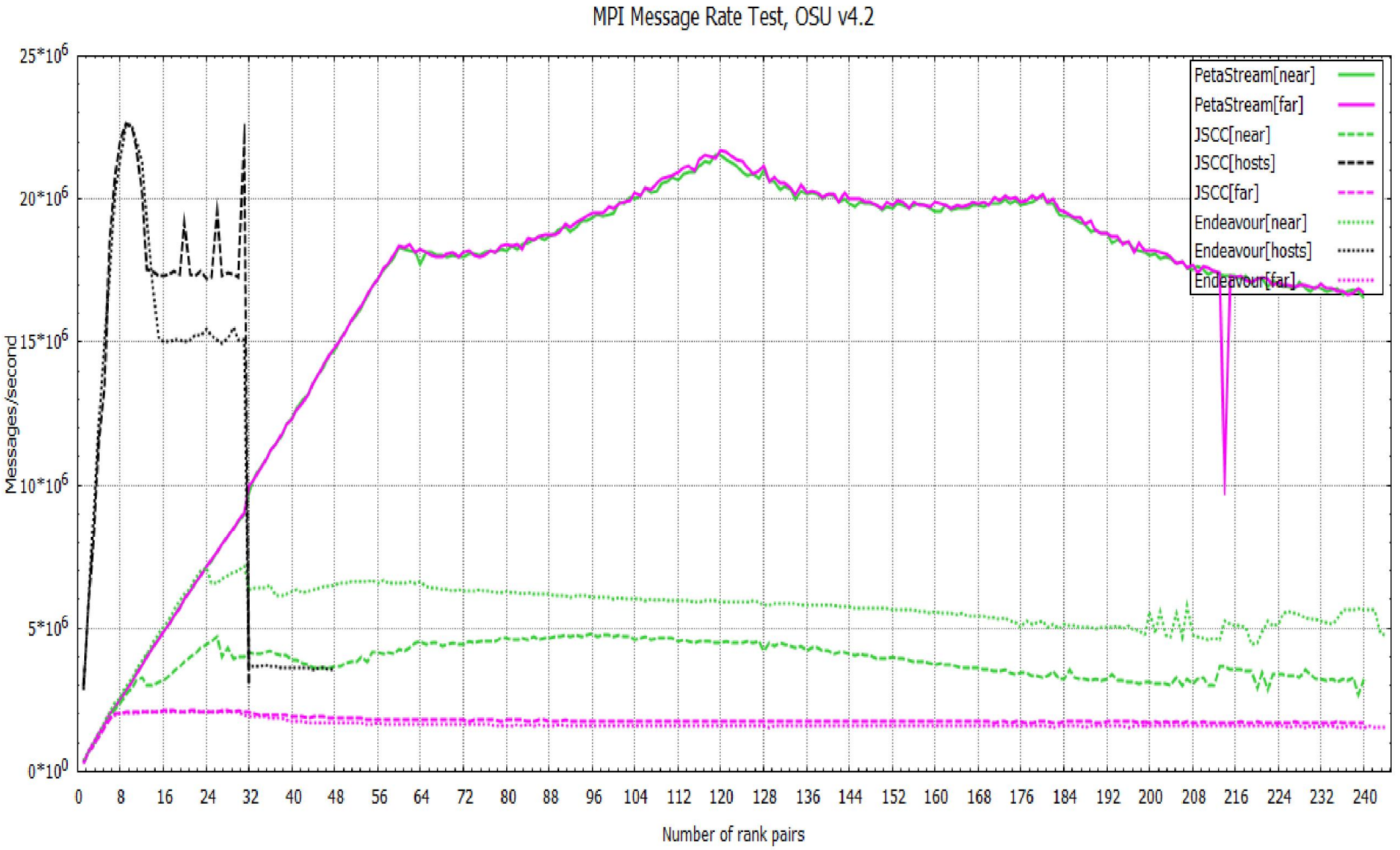
RSC

Производительность приложений на системе RSC PetaStream с Intel® Xeon Phi™

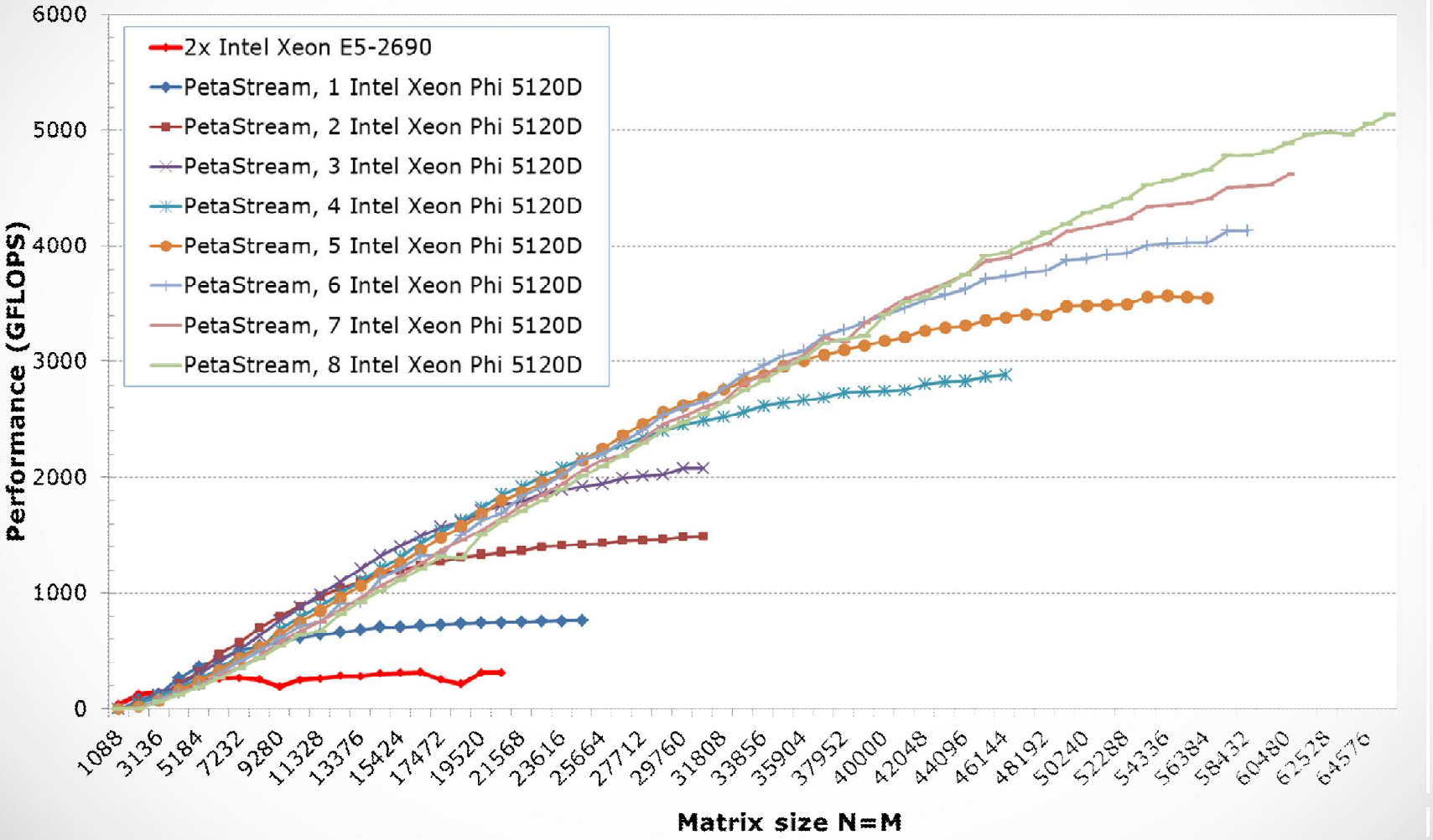
MPI bandwidth



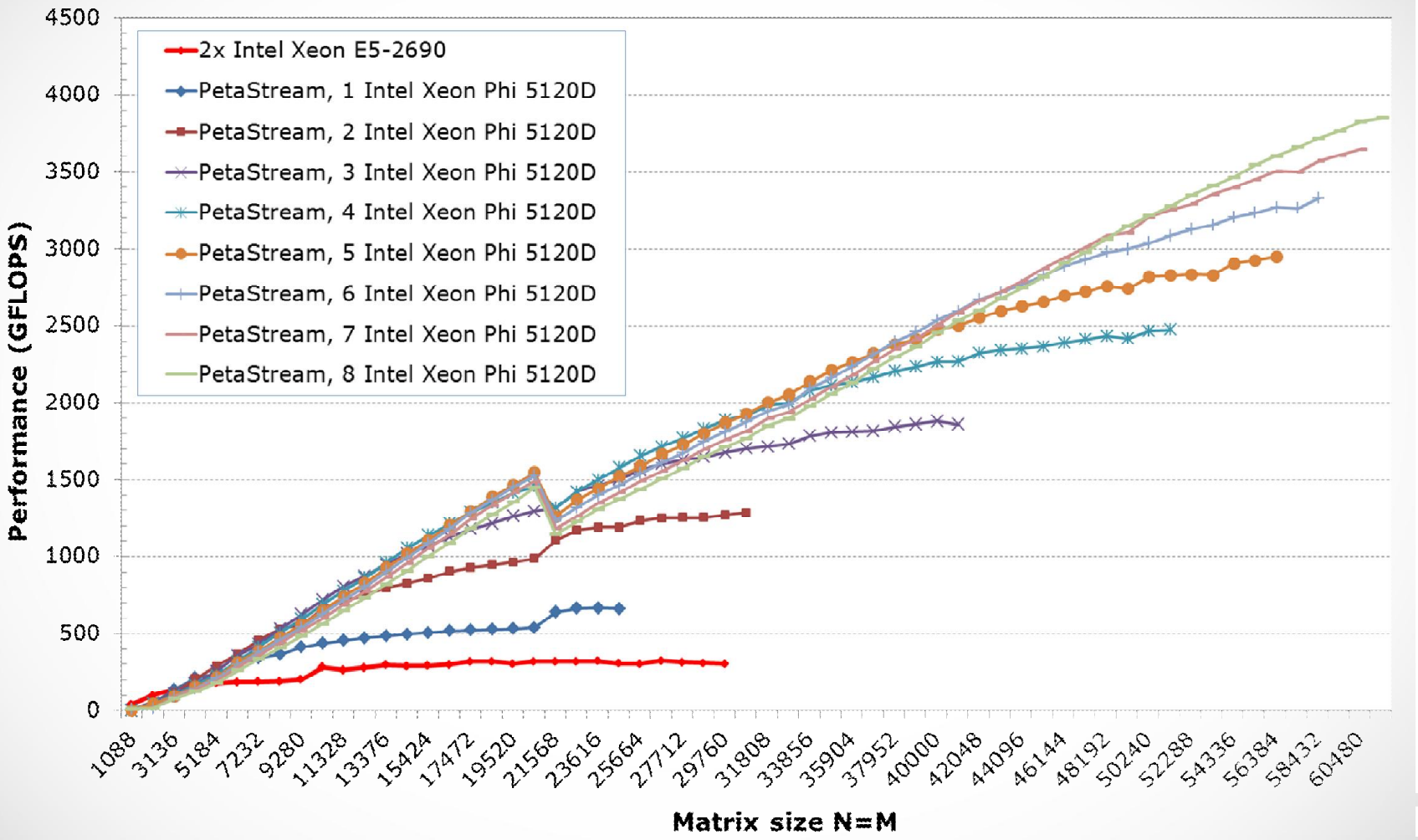
Message rate



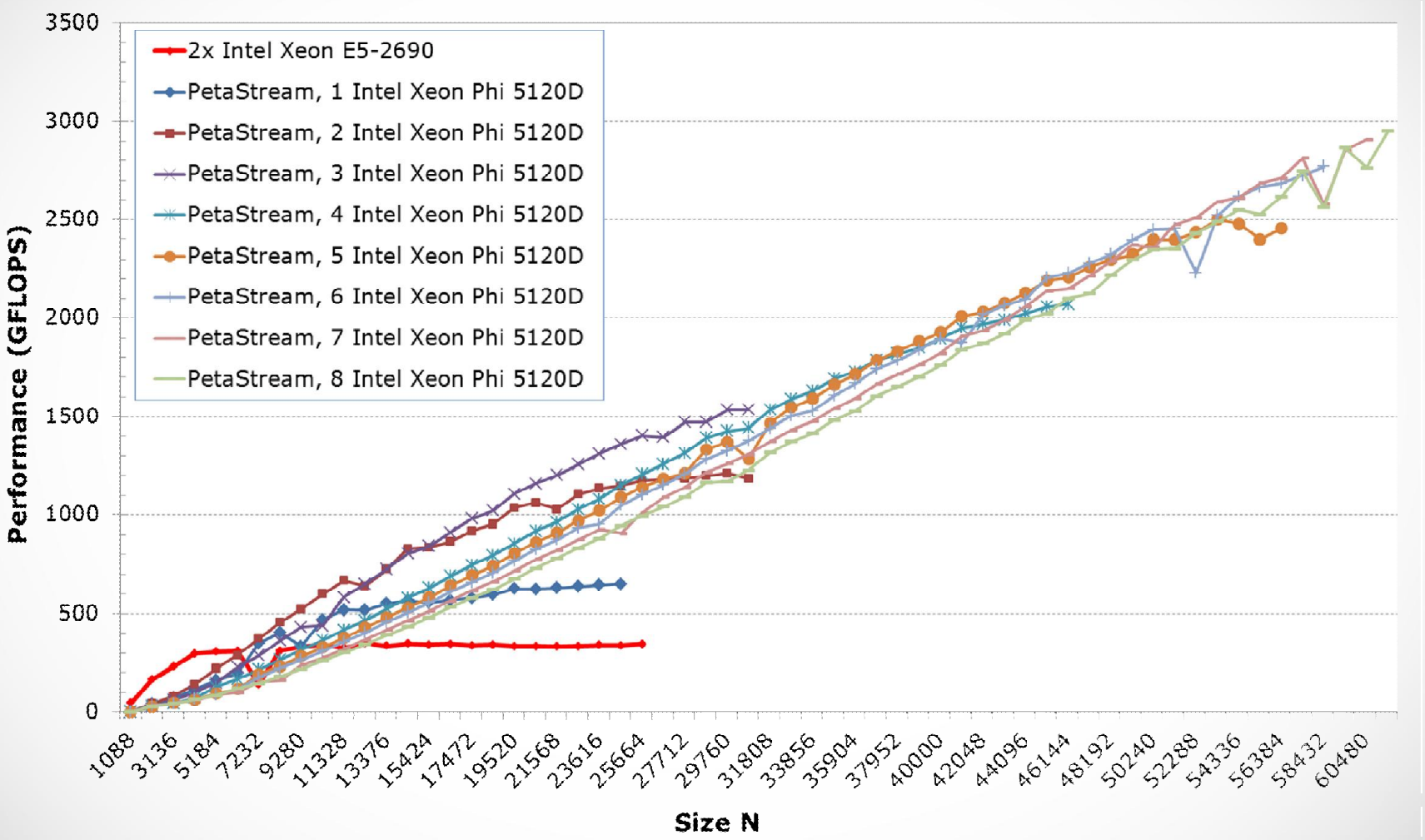
MAGMA MIC Performance, LU



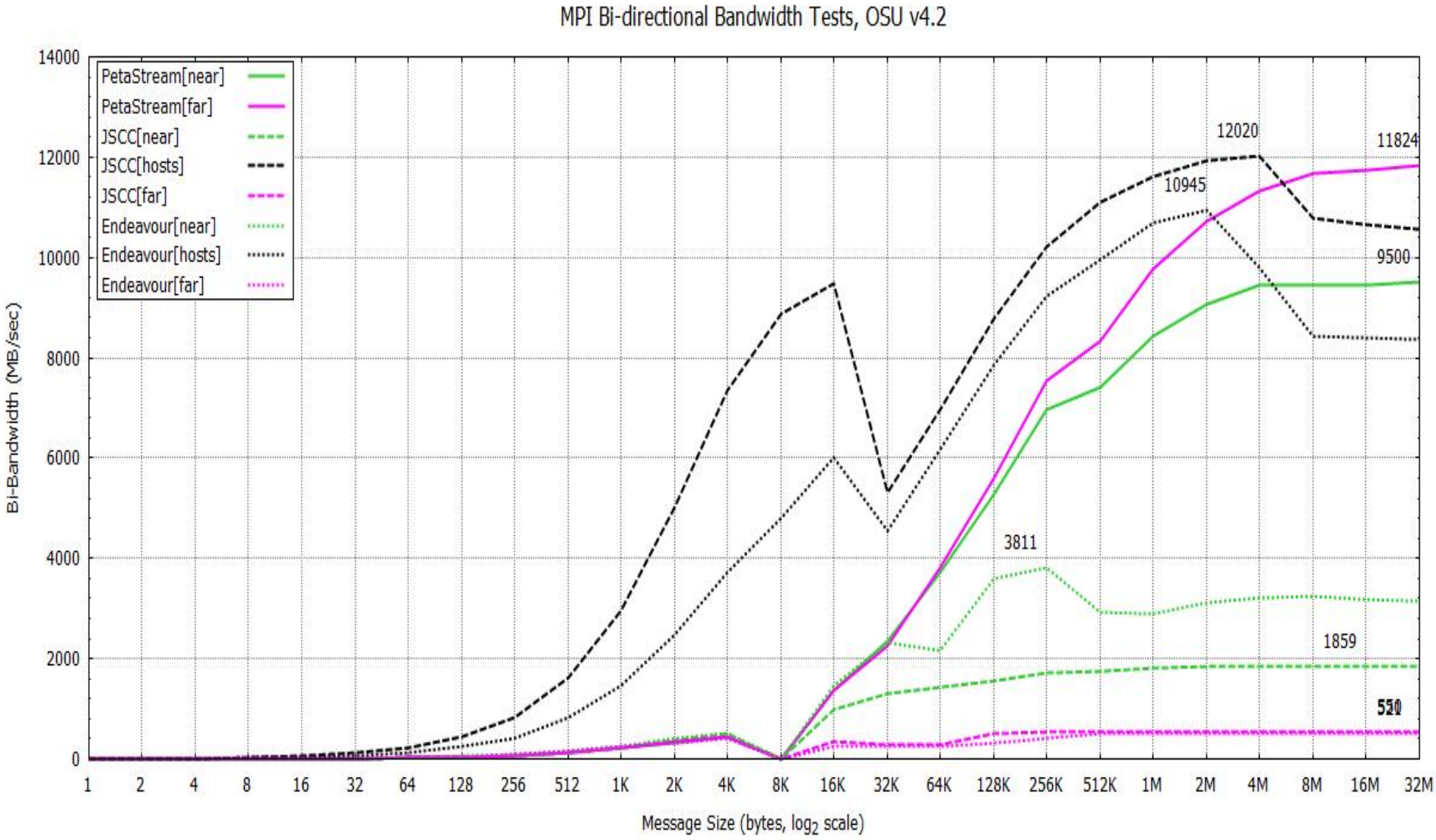
MAGMA MIC Performance, QR



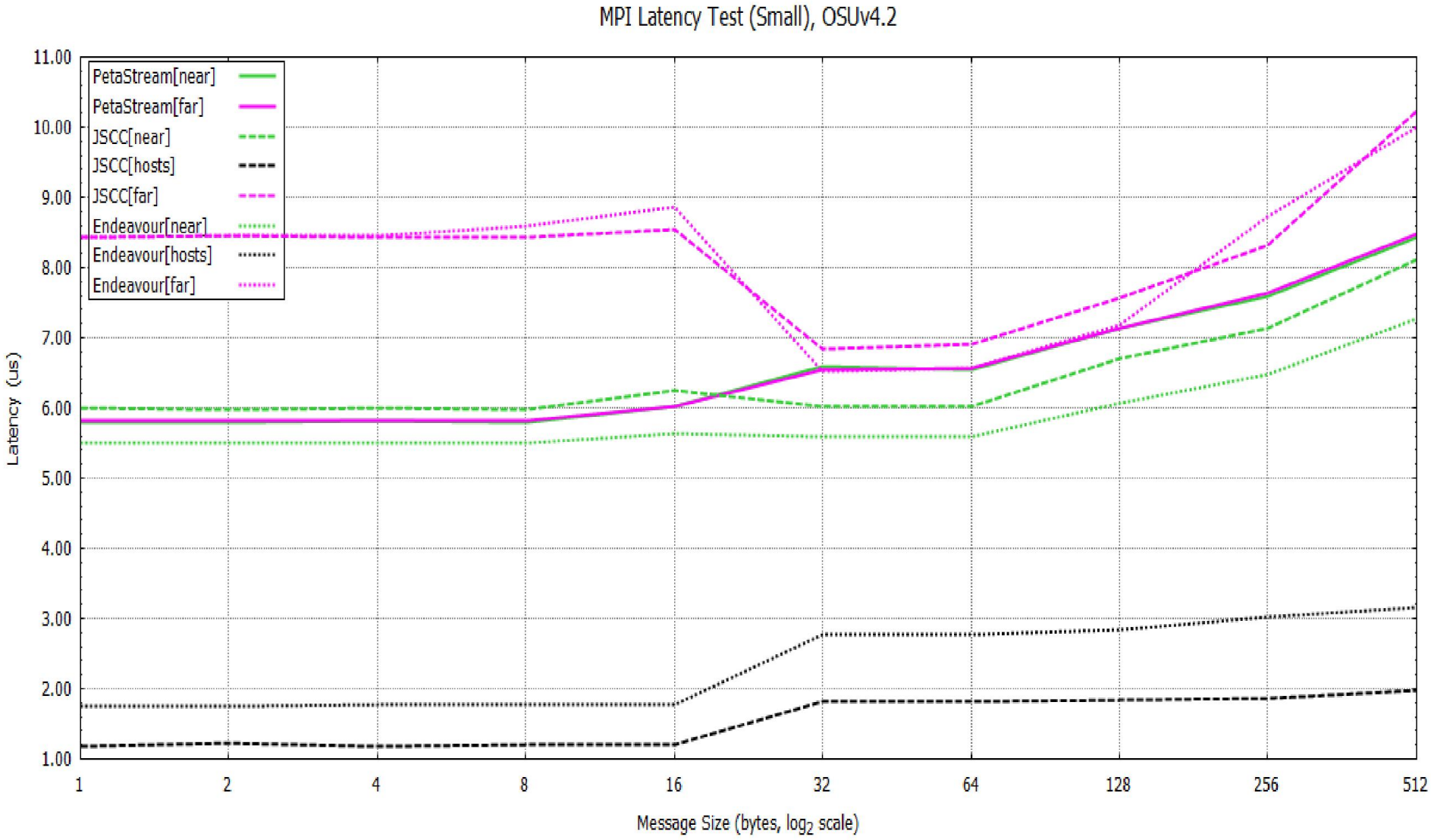
MAGMA MIC Performance, Cholesky



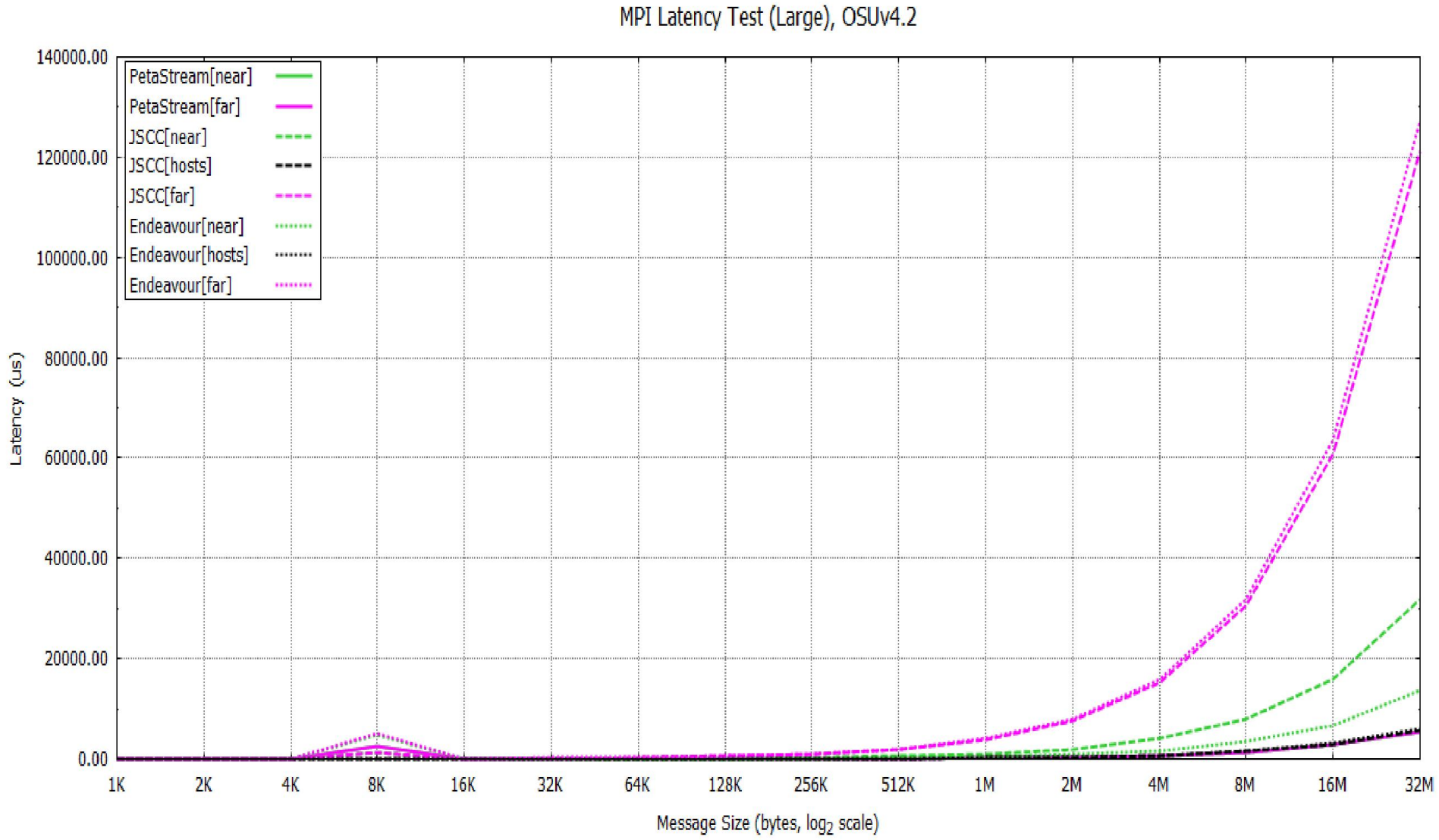
MPI bi-directional bandwidth



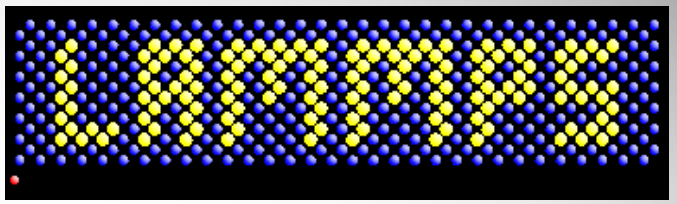
MPI latency – small (1B-512B) messages



MPI latency – large (1 KB-32 MB) messages



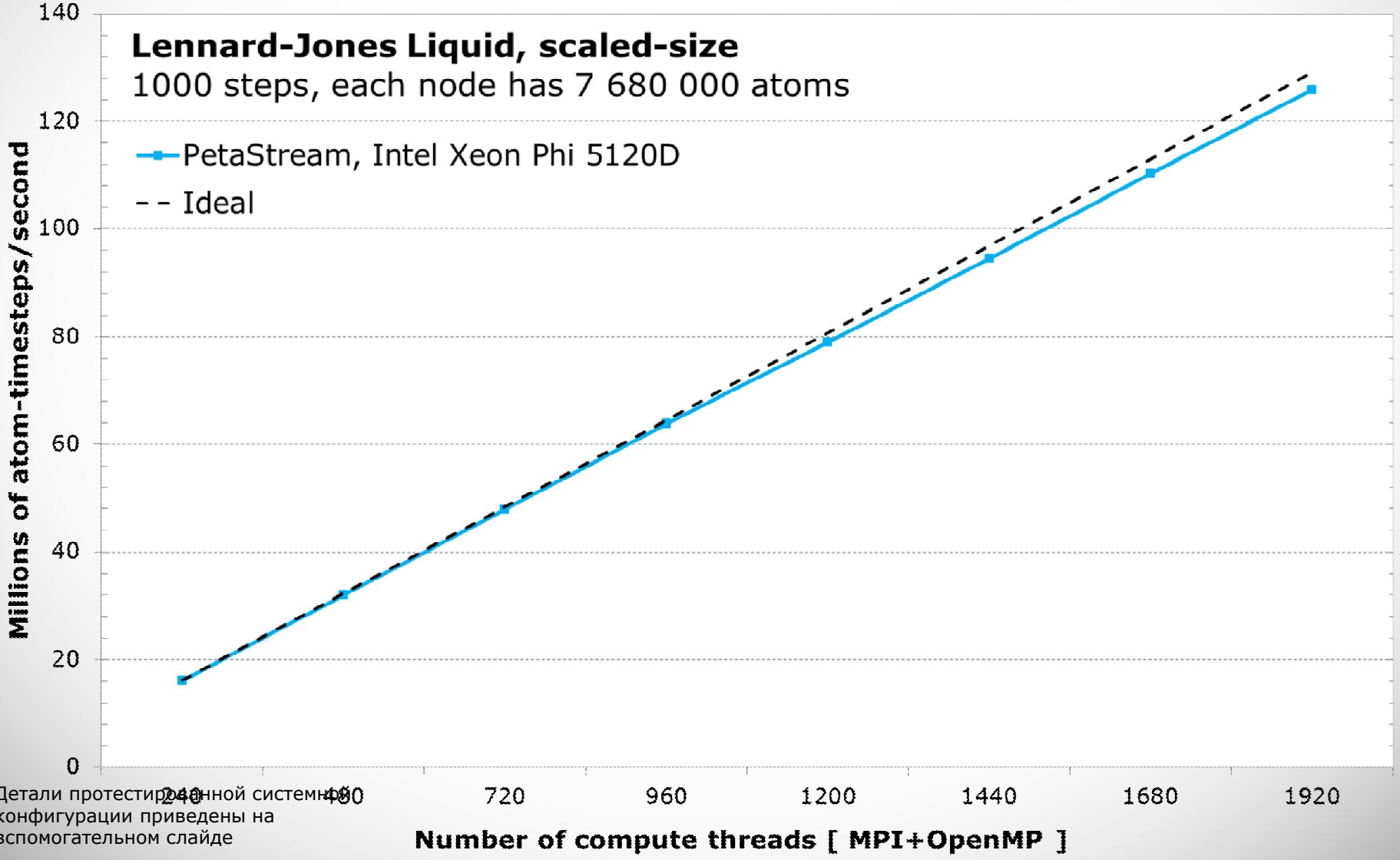
Пример масштабирования: LAMMPS



- Классический молекулярный динамический код (**L**arge-scale **A**tomic/**M**olecular **M**assively **P**arallel **S**imulator - крупномасштабный атомный/молекулярный высокопараллельный симулятор)
- Моделирует мягкие материалы (биомолекулы, полимеры), твердые материалы (металлы, полупроводники) и крупнозернистые или мезоскопические системы
- Распространяется Sandia National Laboratories по лицензии GPL
- Версия от 11 марта 2013 г. с рабочей задачей LJ (3d Lennard-Jones), newton=off, 32 тыс. атомов, масштабируется в соответствии с описанием
- Скомпилировано с помощью Intel Compiler 14.0.0, Intel MKL 11.1.1.106 и Intel MPI 4.1.1.036

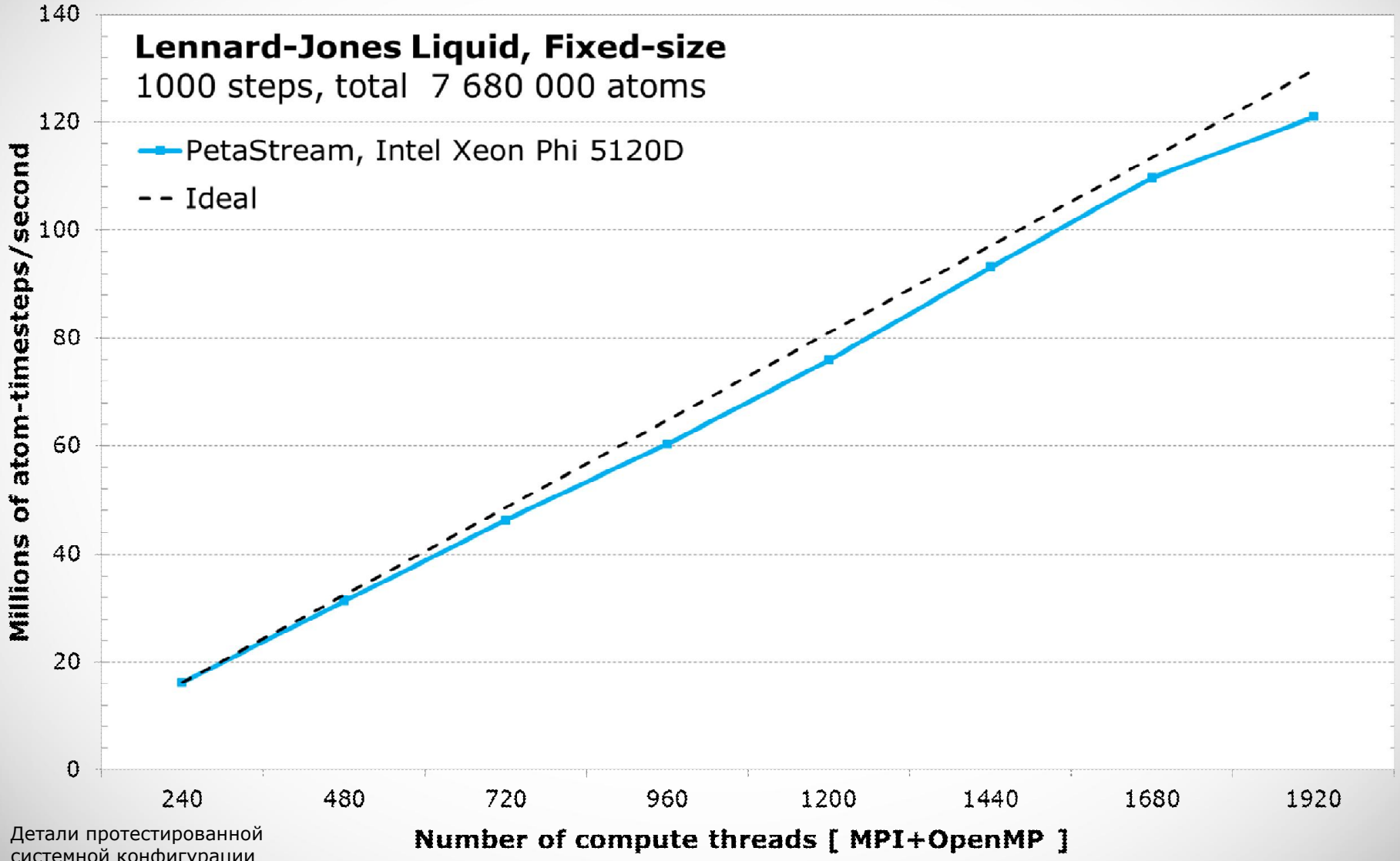
Изображение и описание LAMMPS с сайта: http://lammps.sandia.gov/

LAMMPS: производительность, масштабируемая рабочая нагрузка



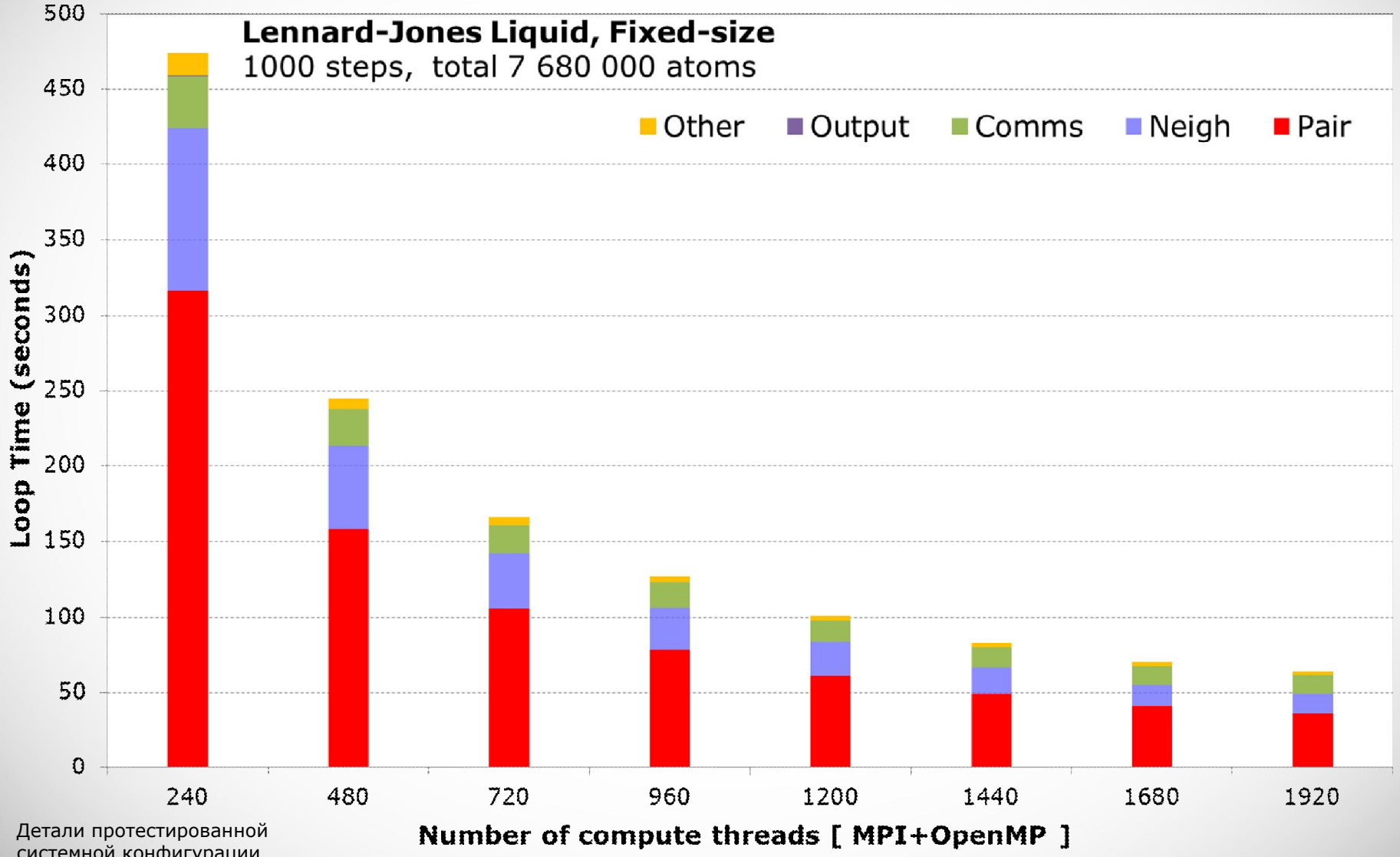
Детали протестированной системной конфигурации приведены на вспомогательном слайде

LAMMPS: производительность, фиксированная рабочая нагрузка



Детали протестированной системной конфигурации приведены на вспомогательном слайде

LAMMPS: профиль времени цикла



Детали протестированной системной конфигурации приведены на вспомогательном слайде

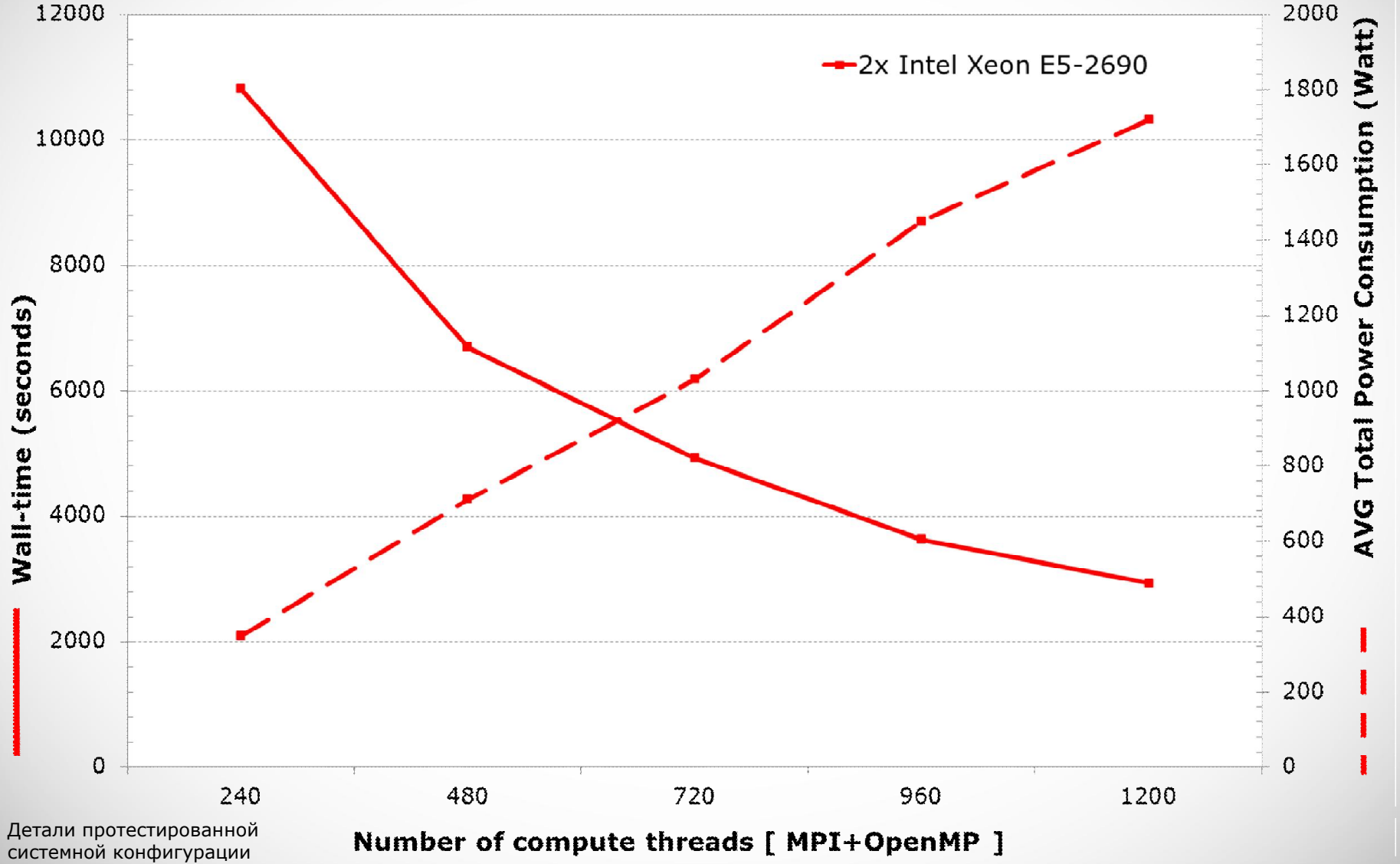
Пример соотношения мощности и производительности: Gromacs

- **GRO**ningen **MA**chine for **C**hemical **S**imulations (Гронингенская машина для химических симуляций)
- В основном предназначается для симуляции протеинов, липидов и нуклеиновых кислот
- Используется в распределенной вычислительной системе Folding@home; в EvoGrid для эволюции искусственной жизни
- Содержит более 1,6 млн. строк исходного кода
- Распространяется по лицензии GNU Lesser General Public License для версий >= 4.6
- Версия 4.6 с рабочей нагрузкой Water (1,536 млн. атомов)
- Скомпилировано с помощью Intel Compiler 14.0.0, Intel MKL 11.1.1.106 и Intel MPI 4.1.1.036

GROMACS FAST. FLEXIBLE. FREE.

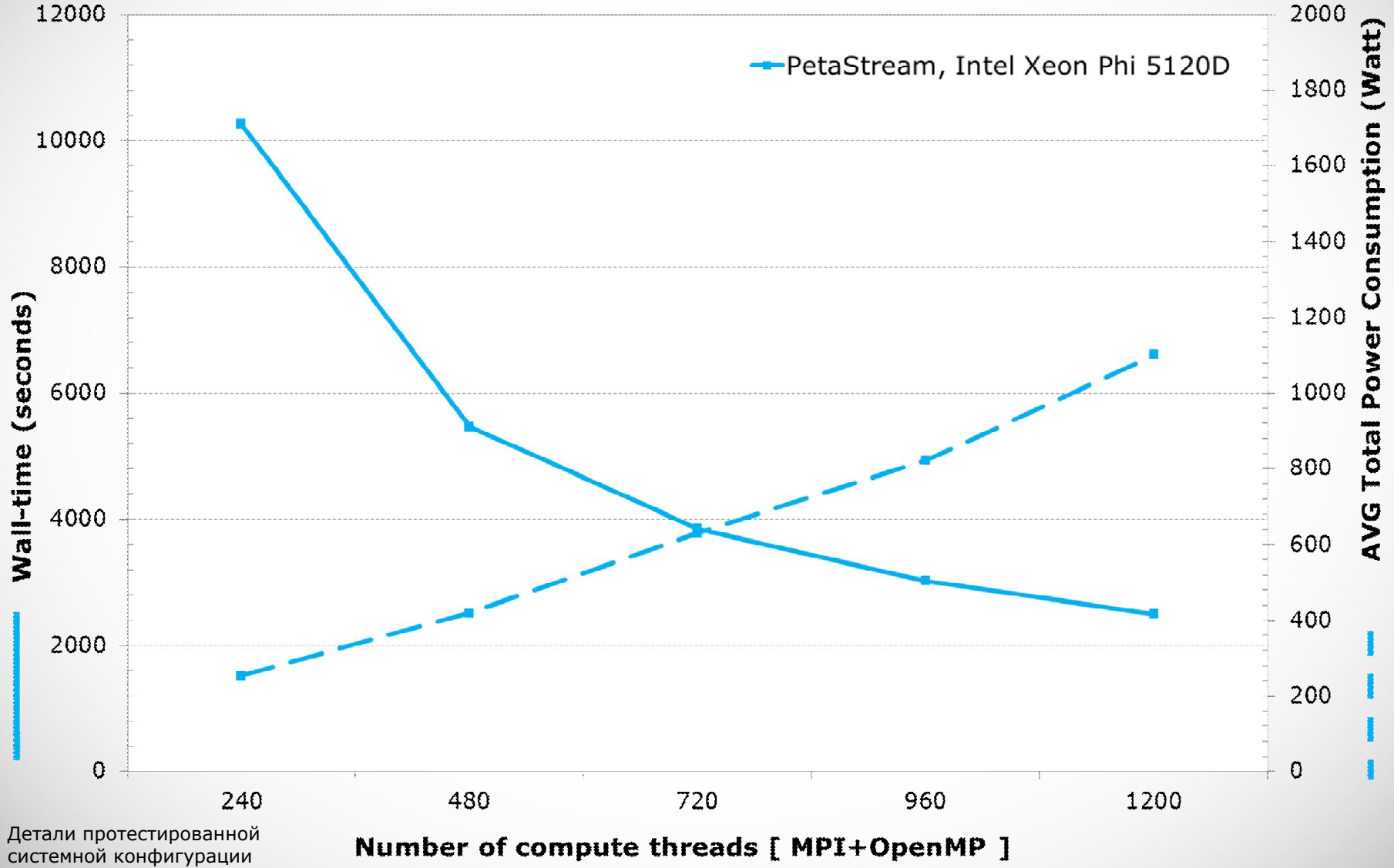
Изображение и описание Gromacs с сайта: http://www.gromacs.org/ и Wikipedia

Gromacs: фактическое время и среднее общее потребление мощности



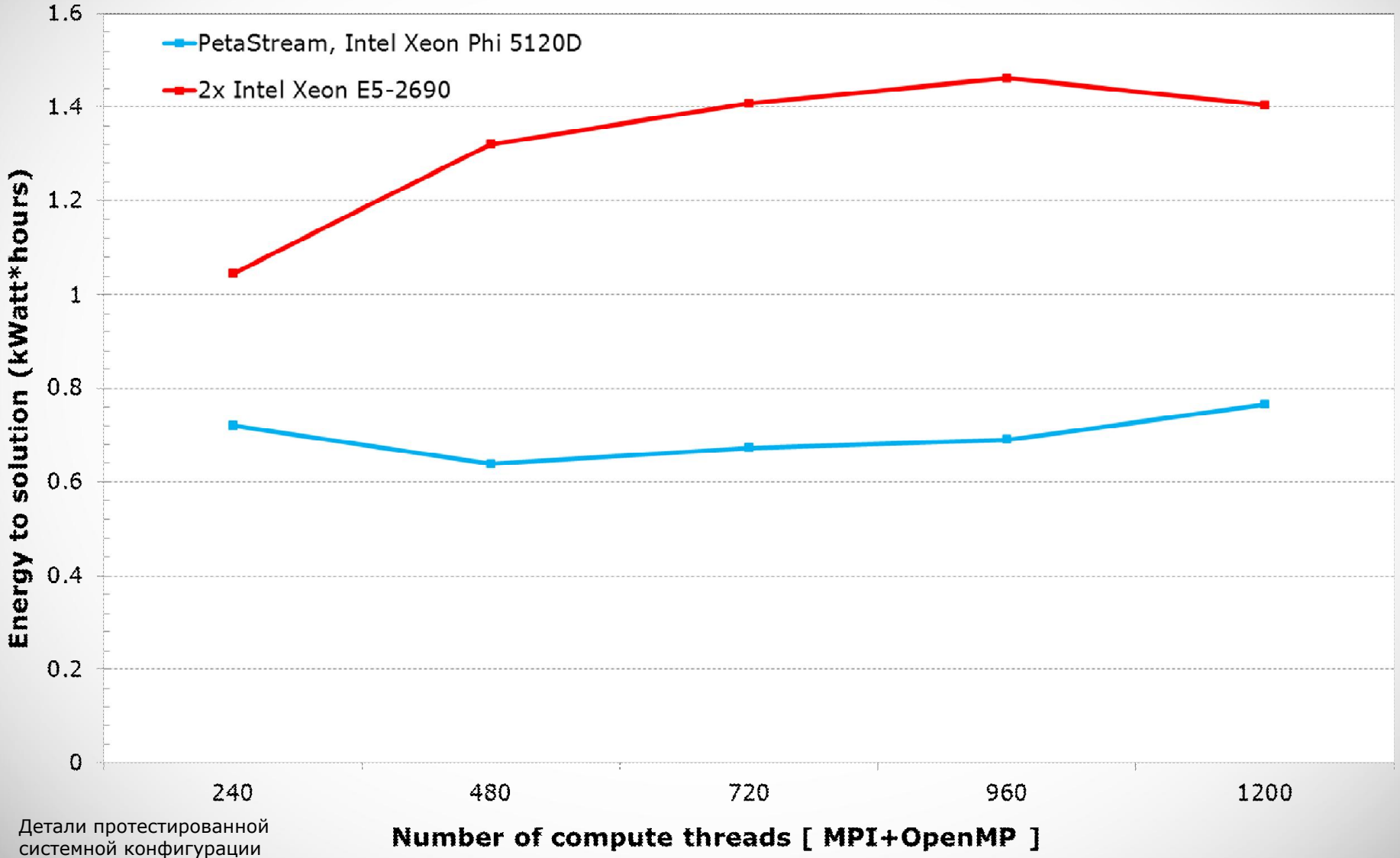
Детали протестированной системной конфигурации приведены на вспомогательном слайде

Gromacs: фактическое время и среднее общее потребление мощности



Детали протестированной системной конфигурации приведены на вспомогательном слайде

Gromacs: энергопотребление решения



Детали протестированной системной конфигурации приведены на вспомогательном слайде

Выводы

- Архитектура Intel® MIC позволяет **легко выполнять миграцию существующих приложений** и обеспечивает **удобную и знакомую среду разработки**
  - Тем не менее, в высокопроизводительных вычислительных системах требуется оптимизация для применяемой архитектуры
- Оптимизированные приложения для сопроцессоров Intel® Xeon Phi™ имеют **отличную производительность и масштабируемость**
- Использование реальных приложений в системах с сопроцессорами Intel® Xeon Phi™ позволяет **снизить энергопотребление решения в более чем 2 раза**

RSC

= 9x

Энергопотребление решения в 2 раза ниже

RSC PetaStream обеспечивает высокую компактность и энергоэффективность вычислительных систем с рекордной производительностью на экзафлопном уровне

Спасибо

www.rscgroup.ru

Контакты

Александр Московский
ЗАО «РСК Технологии»

Россия, Москва,
Кутузовский проспект 36, строение 23
moskov@rsc-tech.ru