

Центр параллельных вычислительных технологий
ГОУ ВПО “Сибирский государственный университет
телекоммуникаций и информатики”
Новосибирск, Россия

АРХИТЕКТУРА, АНАЛИЗ И ОРГАНИЗАЦИЯ ФУНКЦИОНИРОВАНИЯ РАСПРЕДЕЛЕННЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

**В.Г. Хорошевский,
М.Г. Курносков, С.Н. Мамоиленко**

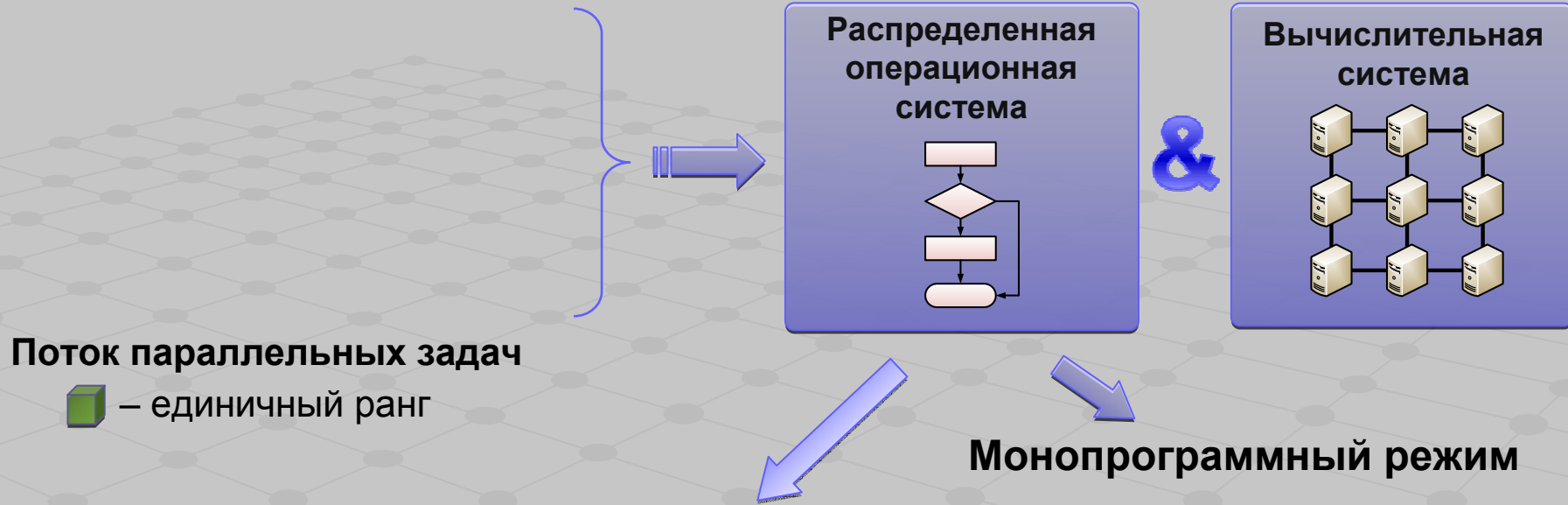
E-mail: {khor, mk, msn}@cpct.sibsu.ru

Международная конференция “Облачные вычисления.
Образование. Исследования. Разработка”
Президиум РАН, г. Москва, 15 - 16 апреля 2010 г.

НАПРАВЛЕНИЯ ИССЛЕДОВАНИЙ В РАМКАХ ПРОГРАММЫ “Университетский кластер”

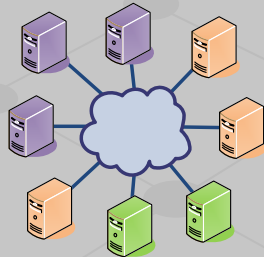
1. Разработка инструментария параллельного мультипрограммирования для распределенных вычислительных систем (ВС).
2. Создание моделей и алгоритмов для организации эффективного решения параллельных задач на распределенных ВС.
3. Развитие средств самоконтроля, самодиагностики распределенных ВС и организации отказоустойчивого выполнения параллельных программ.
4. Создание методов анализа функционирования распределенных ВС: надежности и живучести ВС; осуществимости параллельного решения задач.
5. Отработка архитектурных решений в области пространственно-распределенных вычислительных и GRID систем.

ПАРАЛЛЕЛЬНОЕ МУЛЬТИПРОГРАММИРОВАНИЕ



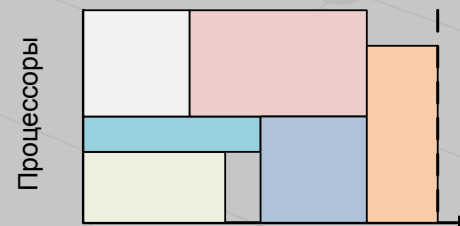
Мультипрограммные режимы

Обслуживание потоков задач
Генерация в пределах ВС подсистем



- Техника теории игр
- Стохастическое программирование

Обработка наборов задач
Формирование расписаний решения параллельных задач



Эвристические методы

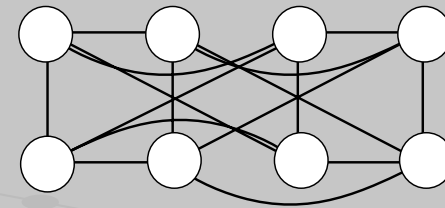


МОНОПРОГРАММНЫЙ РЕЖИМ: ВЛОЖЕНИЕ ПАРАЛЛЕЛЬНЫХ ПРОГРАММ В ВС

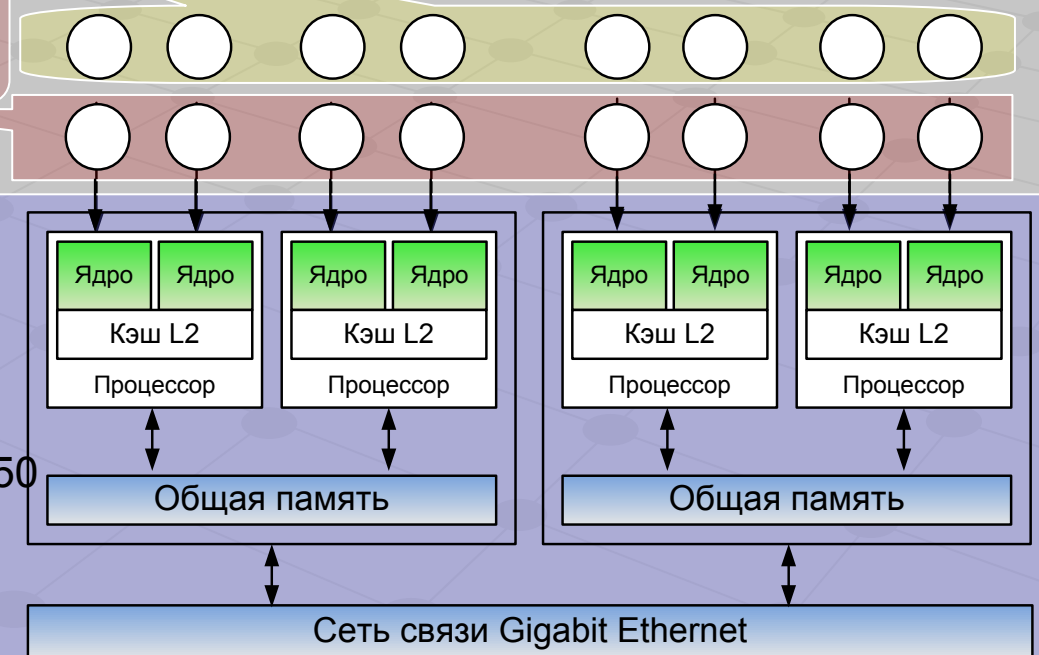
Вложение теста High Performance
Linpack в подсистему:

Стандартными MPI-утилитами (MPICH2):
время выполнения **118 сек.** (44 GFLOPS)

Разработанными средствами:
время выполнения **100 сек.** (53 GFLOPS)



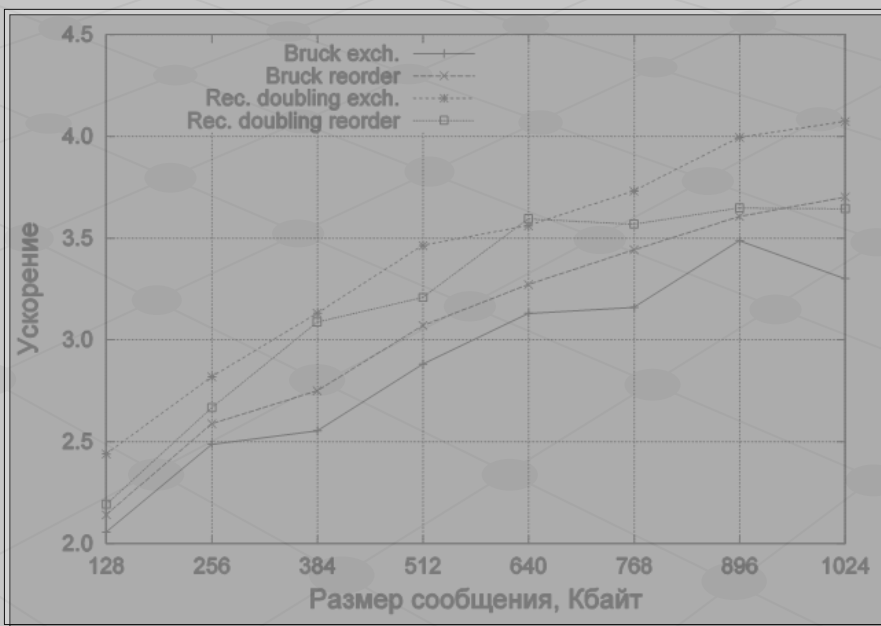
High Performance Linpack



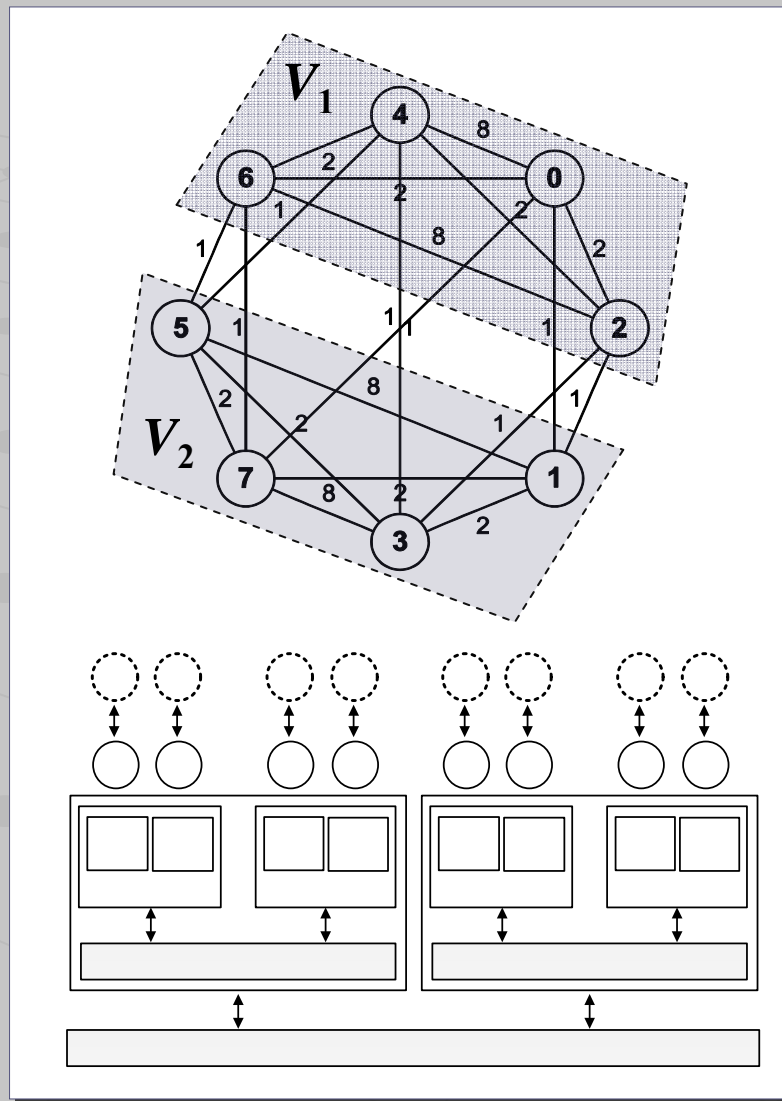
СТРУКТУРНО-ОРИЕНТИРОВАННЫЕ АЛГОРИТМЫ КОЛЛЕКТИВНЫХ ОБМЕНОВ ИНФОРМАЦИЕЙ МЕЖДУ ВЕТВЯМИ ПАРАЛЛЕЛЬНЫХ ПРОГРАММ

Предложен метод оптимизации коллективных операций (MPI_Allgather, MPI_Vcast и др.) для распределенных ВС с иерархической структурой

Параллельные ветви программы, обменивающиеся большими объемами данных при реализации коллективной операции заданным алгоритмом (ring, recursive doubling, Bruck, и др.), динамически перераспределяются на процессорные ядра одного вычислительного узла.



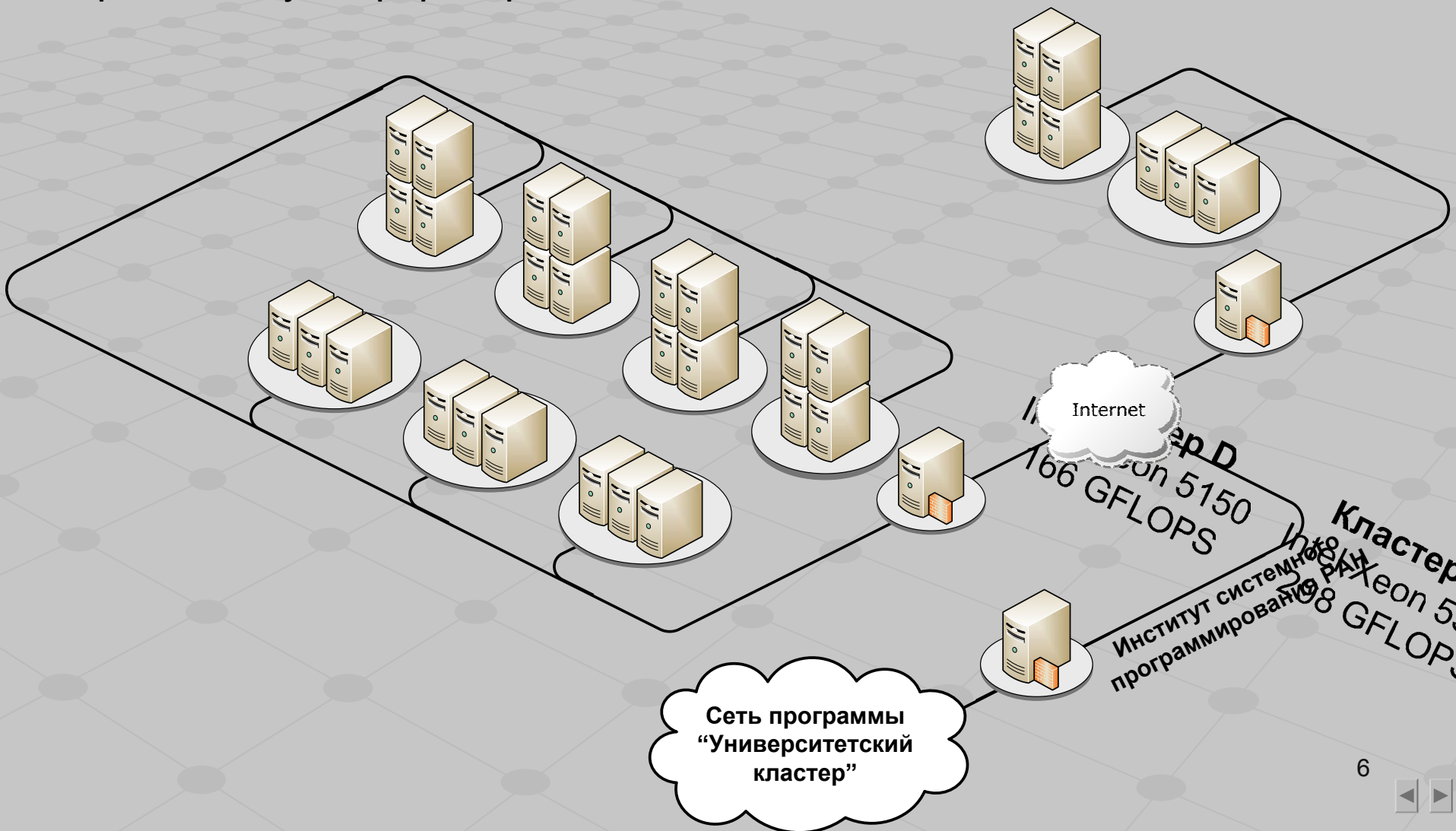
Ускорение оптимизированных алгоритмов Allgather относительно исходных: 8 узлов по 8 процессорных ядер, сеть связи - Gigabit Ethernet



ПРОСТРАНСТВЕННО-РАСПРЕДЕЛЁННАЯ МУЛЬТИКЛАСТЕРНАЯ ВЫЧИСЛИТЕЛЬНАЯ СИСТЕМА

GRID-модель

- Программируемость структуры, масштабируемость, живучесть
- Параллельное мультипрограммирование



ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ МУЛЬТИКЛАСТЕРНОЙ ВС

Удаленный доступ и мониторинг
(SSH, Globus Toolkit; Ganglia)

Средства разработки параллельных программ

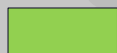
- **MPI**: **TopoMPI**, MPICH2, OpenMPI
- **PGAS**: Unified Parallel C
- **OpenMPI**: GNU Compilers, Intel Compilers, Sun Studio Compilers
- Средств анализа MPI-программ: **mpistat**, **otfstat**, VampirTrace

Средства организации распределенной очереди задач
(**Gbroker**, **dqueued**, GridWay)

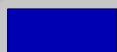
Подсистема параллельного мультипрограммирования
(TORQUE, MAUI, mpiexec)

Подсистема самоконтроля, самодиагностики ВС и организации отказоустойчивого выполнения параллельных программ
(**DMTCP** – Distributed MultiThreaded CheckPointing)

Операционная система GNU/Linux



Подсистема параллельного мультипрограммирования



Разрабатываемые в ЦПВТ ГОУ ВПО “СибГУТИ” компоненты



**СПАСИБО
ЗА ВНИМАНИЕ**