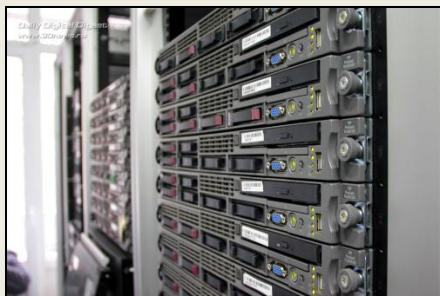


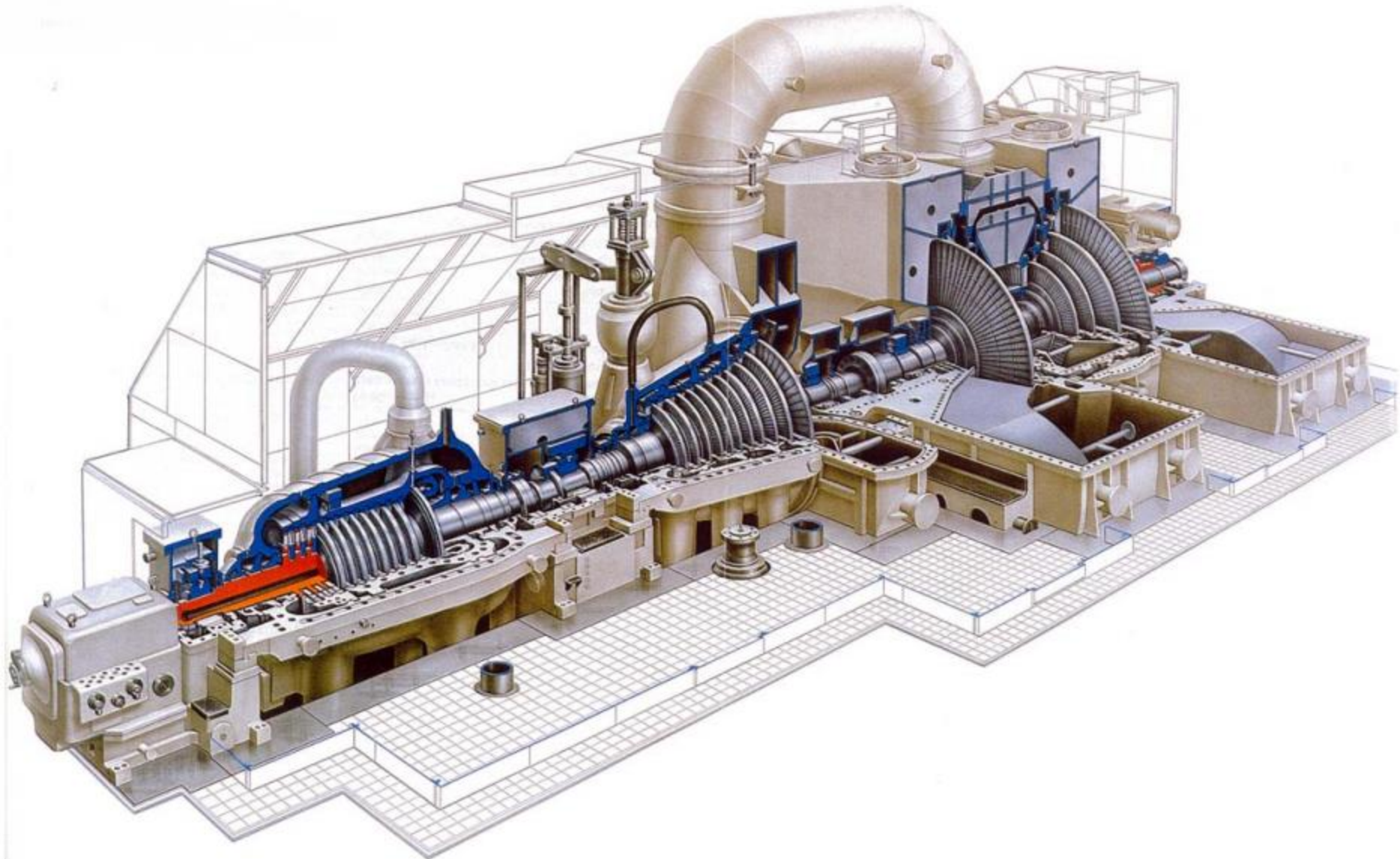
Организация управления распределенными вычислениями в интегрированной кластерной системе

Костромин Роман Олегович, Институт
динамики систем и теории управления имени
В.М. Матросова, Иркутск



Введение

Процессы в теплоэнергетических установках



Время расчета

> 30 минут



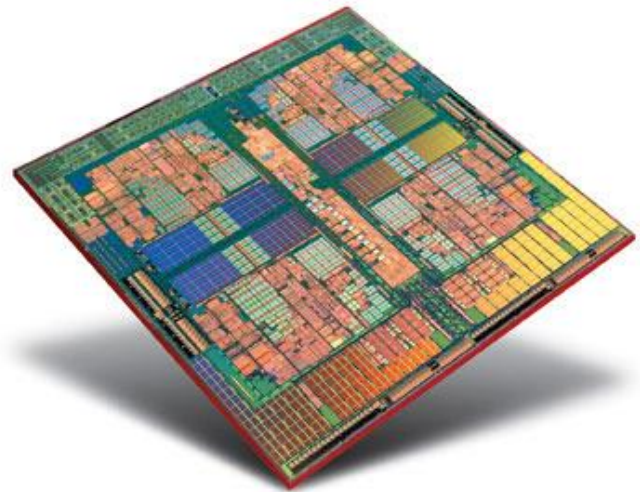
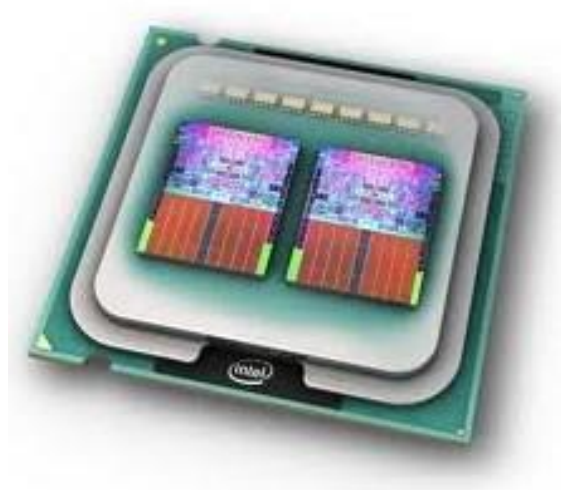
Количество расчетов

> 100



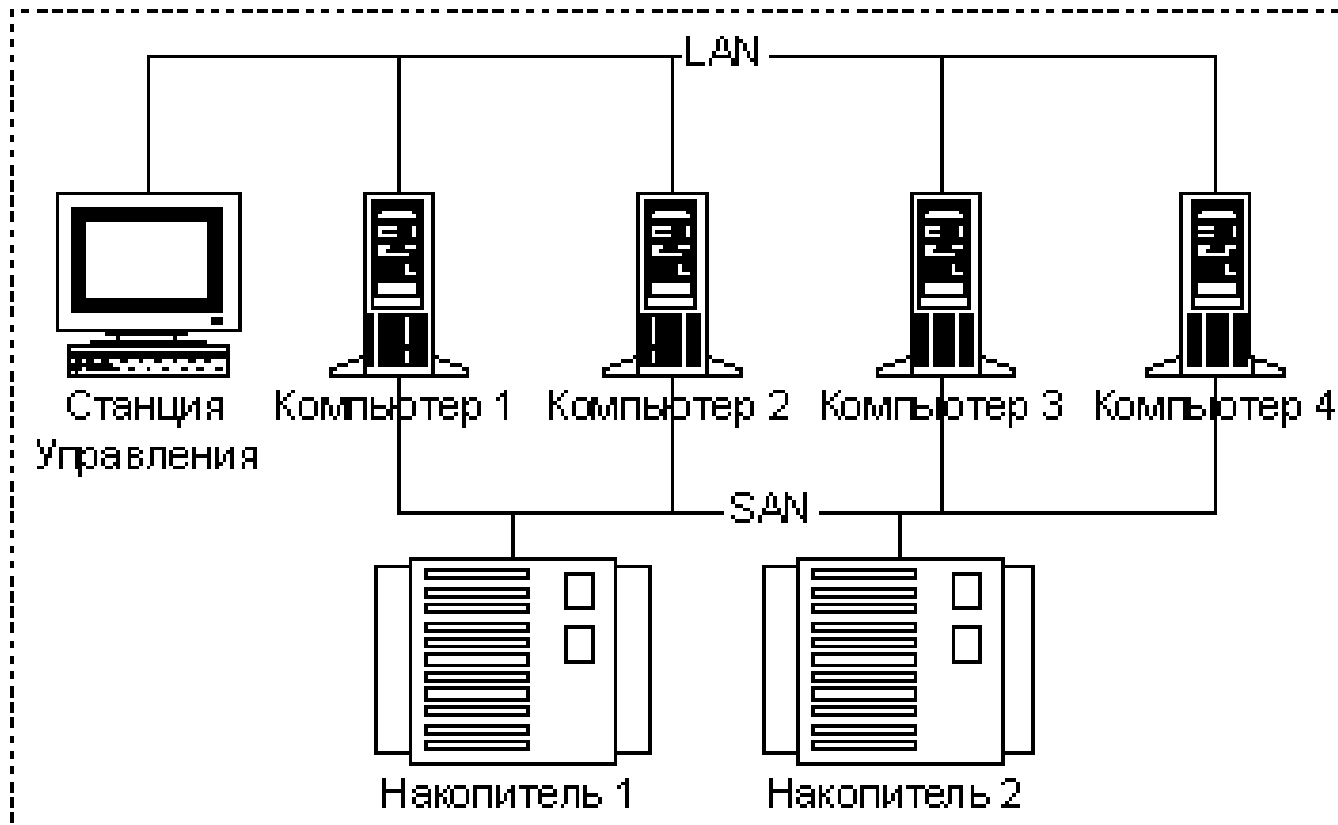
Частота – меньше, ядер – больше

- ➔ параллельные алгоритмы;
- ➔ многопроцессорные системы;



Кластер

Кластер – группа компьютеров (вычислительных узлов), объединенных в локальную вычислительную сеть и способных работать в качестве единого вычислительного ресурса.



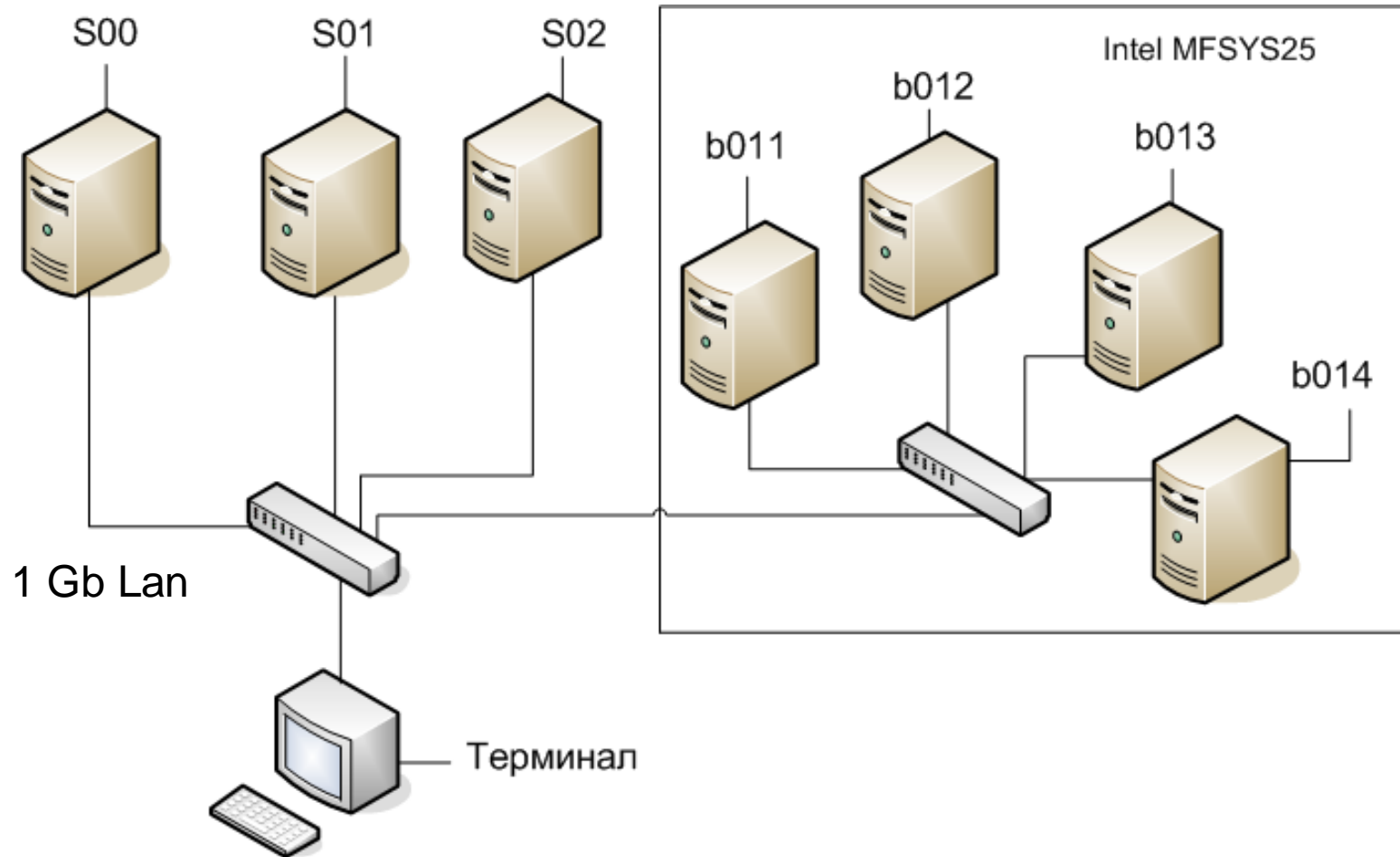
Вычислительные узлы

Узел	Платформа	Процессор	Частота Ghz / Ядер	ОЗУ
S00	Intel SR2520SAXS(R)	2x Xeon (5345)	2.65 / 8	16 Гб
S01	Intel SR2520SAXS(R)	2x Xeon (5430)	2.32 / 8	16 Гб
S02	Intel S5000XVN	2x Xeon (5345)	2.65 / 8	12 Гб
B011	Intel SR1625URSAS	2x Xeon (E5640)	2.65 / 8	12 Гб
B012	Intel SR1625URSAS	2x Xeon (E5649)	2.52 / 12	12 Гб
B013	Intel SR1625URSAS	2x Xeon (E5649)	2.52 / 12	12 Гб
B014	Intel SR1625URSAS	2x Xeon (E5649)	2.52 / 12	12 Гб



Intel MFSYS25

Схема соединения узлов



Выполненные задачи

Выбрано, установлено и настроено системное ПО:

- ОС Gentoo Linux – для управления ресурсами кластера;
- библиотека Open MPI (Open Message Passing Interface) – для поддержки выполнения параллельных расчетов;
- система BPS Torque (Terascale OpenSource Resource and QUEue Manager) - для управления вычислительными заданиями.

Разработан набор скриптов на языке BASH (Bourne Again Shell) для автоматизации процессов установки операционной системы Gentoo Linux в узлах кластера:

- создания/удаления разделов жесткого диска;
- создания необходимых файловых систем и их монтирования;
- настройки сетевых подключений;
- загрузки и установки программных пакетов;
- конфигурирования необходимых параметров системы.

Обеспечена расширяемость кластера - оперативное добавление включение в его инфраструктуру новых узлов;

Обеспечена поддержка отказоустойчивости кластера - быстрое восстановление его работоспособности в случае сбоев программно-аппаратных средств;

Произведена оценка производительности кластера при помощи набора эталонных тестов NASA Advanced Supercomputing Parallel Benchmarks (NPB).

Оценка производительности

NAS Parallel Benchmark (NPB) - Ep.C.48

(EP, класс C, 48 потоков). Ядро EP («**Embarrassing Parallel**») основано на порождении пар псевдослучайных, нормально распределенных чисел (Гауссово распределение)

Узлы	Процессов на узел	Время выполнения, сек	Ускорение
S01	1	430,01	1
S01	2	214,98	2,0
S01	4	107,46	4,0
S01	8	53,94	7,9
S01, S02	8	27,02	15,9
S01, S02, B011, B012	8	13,70	31,38
S00, S01, S02, B011, B012, B013	8	9,43	45,60

NAS Parallel Benchmark (NPB) - Ep.C.48



Linpack

Узел	Gflops
S00	72,73
S01	51,81
S02	78,60
B011	79,38
B012	78,55
B013	78,66
B014	78,71

$\approx 500Gflops$

Создание вычислительного кластера обеспечило возможность решения ряда сложных задач оперативного управления теплоэнергетическими установками за приемлемые промежутки времени.

Интеграция вычислительных кластеров



Суперкомпьютер «В.М. Матросов»
ИДСТУ СО РАН

Производительность - 33,7 Тфлопс



Кластеры
выделенных и
невыделенных
рабочих станций



Интегрированная кластерная система

Потоки заданий пользователей интегрированной кластерной системы

Потоки заданий пользователей сегментов кластерной системы

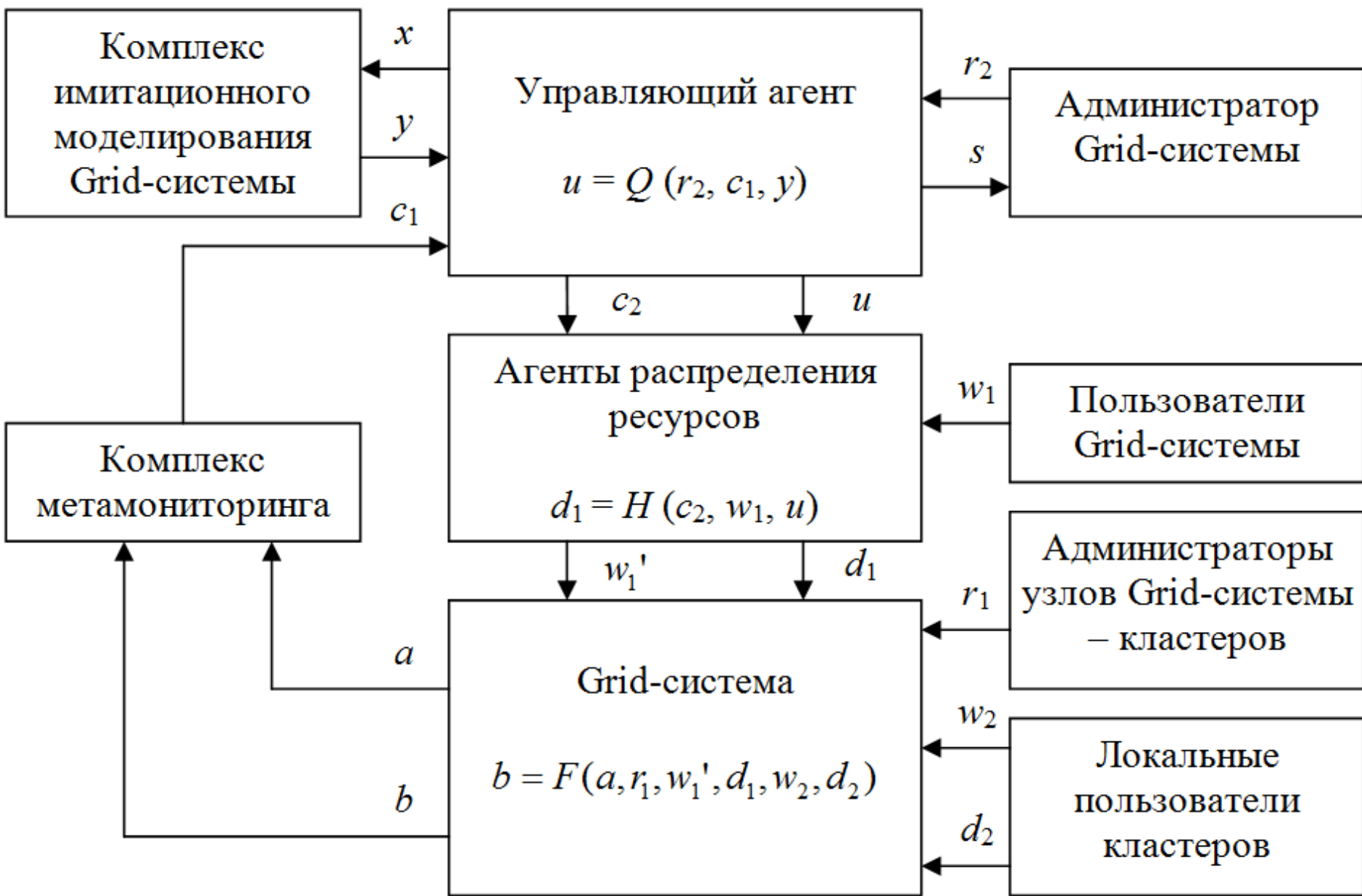
Интегрированная кластерная система



- интеграция вычислительных кластеров позволяет существенно расширить спектр решаемых с их помощью задач и предполагает использование специального промежуточного программного обеспечения (middleware);
 - процесс решения задачи специфицируется в виде задания вычислительной среде, содержащего информацию о требуемых вычислительных ресурсах, исполняемых прикладных программах, входных/выходных данных, а также другие необходимые сведения;
 - традиционные системы управления заданиями промежуточного программного обеспечения недостаточно полно учитывают специфику спроса и предложения ресурсов при распределении этих ресурсов для выполнения заданий
=>
- => проблема эффективного планирования вычислений и распределения ресурсов для их выполнения\$

Решение => применение мультиагентных систем для управления вычислениями.

Мультиагентная система с заданной организационной структурой



- прототип системы агентов реализуется с помощью инструментария Java Agent DEvelopment framework (JADE);
- алгоритмы функционирования агентов разрабатываются на основе конечно-автоматной модели в соответствии со спецификой действий, выполняемых этими агентами в системе управления вычислениями;
- разработка системы агентов осуществляется в рамках реализации мультиагентного подхода к управлению распределенными вычислениями в интегрированной кластерной системе.

Планирование заданий

Процесс планирования осуществляется в четыре этапа агентами, представляющими узлы вычислительного кластера

- формирование всего множества доступных узлов;
- конкретизация сформированного множества путем исключения из него перегруженных узлов (относительно текущей средней загрузки узлов с учетом имеющихся очередей заданий);
- построение поливариантного плана выполнения задания в узлах;
- извлечение из построенного поливариантного плана специализированного плана (осуществляется на основе экономического механизма регулирования спроса и предложения вычислительных ресурсов).

- создан вычислительный кластер;
 - проведено тестирование производительности;
- предоставлена возможность решать сложные задачи оперативного управления теплоэнергетических установок за приемлемые промежутки времени.
- получен опыт организации кластера и работы с облачными вычислениями
- разработана система скриптов для автоматизированного разворачивания узлов (быстрое масштабирование)
- организованный кластер используется в ИСЭМ СО РАН для выполнения следующих задач:
- расчеты технологических схем теплоэнергетических установок
 - оптимизация параметров теплоэнергетических установок
 - долгосрочное прогнозирование природообусловленных факторов
 - глобальная оптимизация

Спасибо за внимание!

Организация управления распределенными вычислениями в интегрированной кластерной системе

Костромин Роман Олегович, Институт
динамики систем и теории управления имени
В.М. Матросова, Иркутск

