

Международная лаборатория суперкомпьютерного атомистического
моделирования и многомасштабного анализа

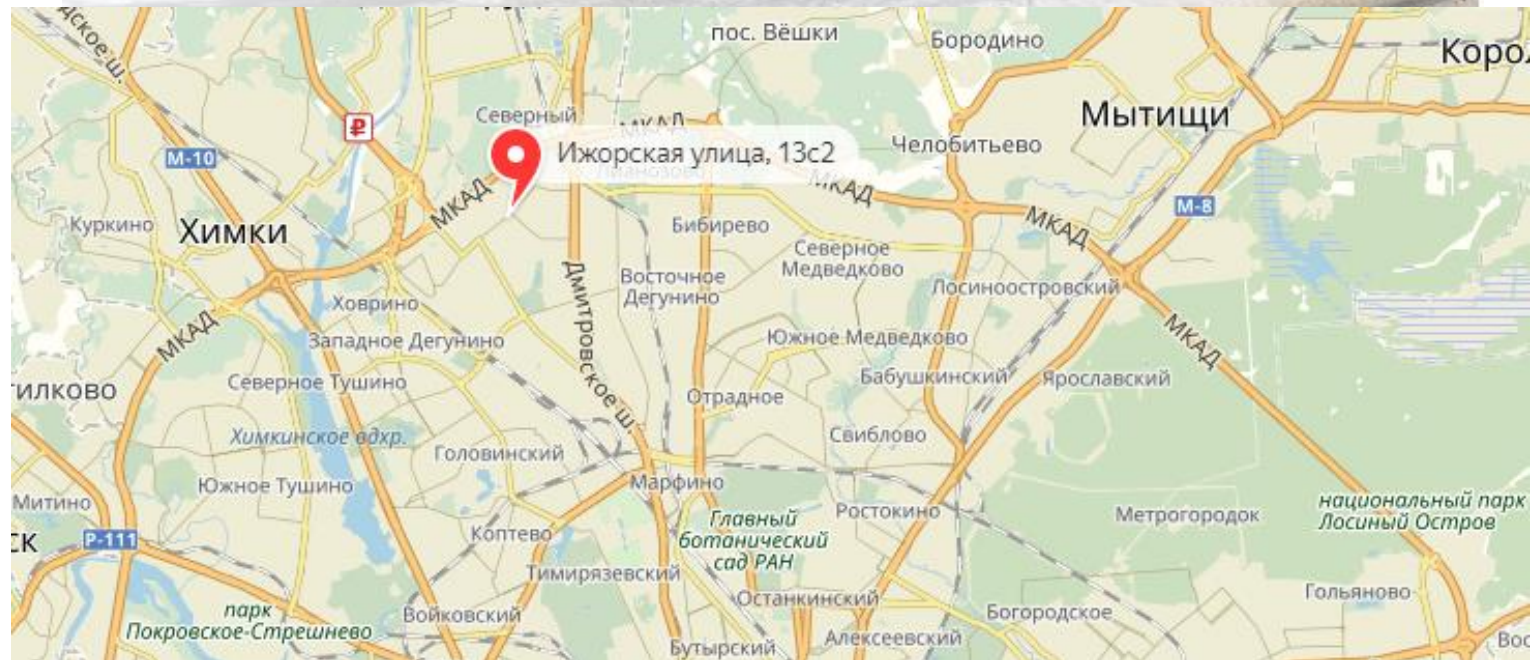
Школа-семинар «Поиск эффективных суперкомпьютерных архитектур в пост-Муровскую эру»

Работа суперкомпьютера Десмос: политика работы с пользователями

Г.С.Смирнов



NATIONAL RESEARCH
UNIVERSITY



История центра

Центр коллективного пользования (ЦКП)
"Суперкомпьютерный Центр ОИВТ РАН" был
создан в 2016 году на базе информационно-
вычислительного центра ОИВТ РАН.

История центра



Суперкомпьютер MBC-1000 был поставлен в ОИВТ РАН в 2001 году НИИ «Квант».

Суперкомпьютер имел 16 вычислительных узлов и один управляющий, на каждом узле было установлено по 2 процессора с тактовой частотой 1 ГГц и 1 Гб оперативной памяти.

Пиковая производительность кластера составляла 32 Гфлопс.

История центра



Суперкомпьютер NWO5 был поставлен в ОИВТ РАН в 2005 году на средства РФФИ.

На кластере настроено программное обеспечение для проведения расчетов в GRID-системах.

Суперкомпьютер состоит из 13 вычислительных узлов, имеющих следующие характеристики: 2 процессора Intel Xeon 3.0 ГГц, 2 Гб оперативной памяти на узел, жесткий диск на 160 Гб, сетевые интерфейсы Fast Ethernet и Gigabit Ethernet. Производительность кластера на тесте LINPACK составляет 108 Гфлопс.

История центра



Суперкомпьютер Т-Платформы TEdge-48 компании Т-Платформы был поставлен в ОИВТ в 2008 году и состоит из 24 вычислительных модулей.

Каждый модуль содержит 2 четырехядерных процессора Intel Xeon 5445 с тактовой частотой 2.33 ГГц и 8 Гбайт оперативной памяти.

Производительность кластера на тесте Linpack составляет 1.4 Tflops.

В 2017 году суперкомпьютер модернизируется (приобретаются 18 узлов на основе процессоров AMD Epyc).

История центра



Суперкомпьютер DESMOS введен в эксплуатацию в конце 2016 г.

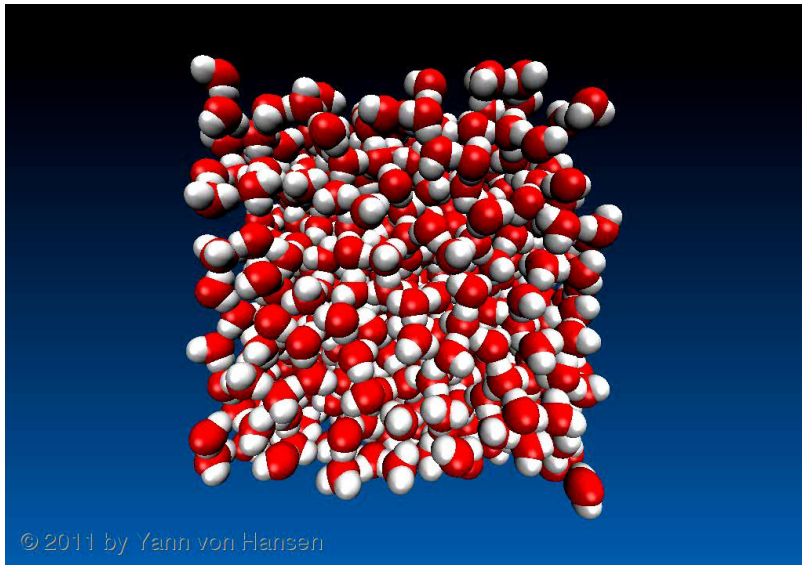
Кластер состоит из 32 гибридных вычислительных узлов и одного головного.

Для объединения узлов используется сеть Ангара.

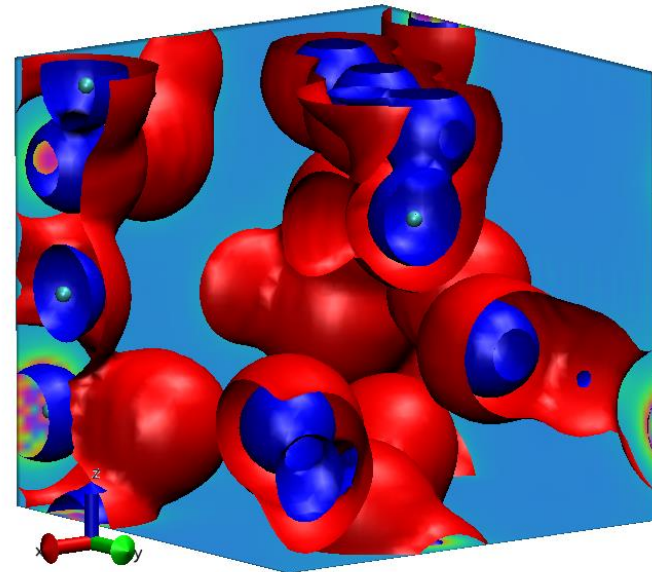
Основные цели создания суперкомпьютера Desmos

- Создание суперкомпьютера для задач классической молекулярной динамики
- Использование интерконнекта Ангара (и максимизация числа узлов при фиксированном бюджете)

LAMMPS, GROMACS



VASP, CP2K

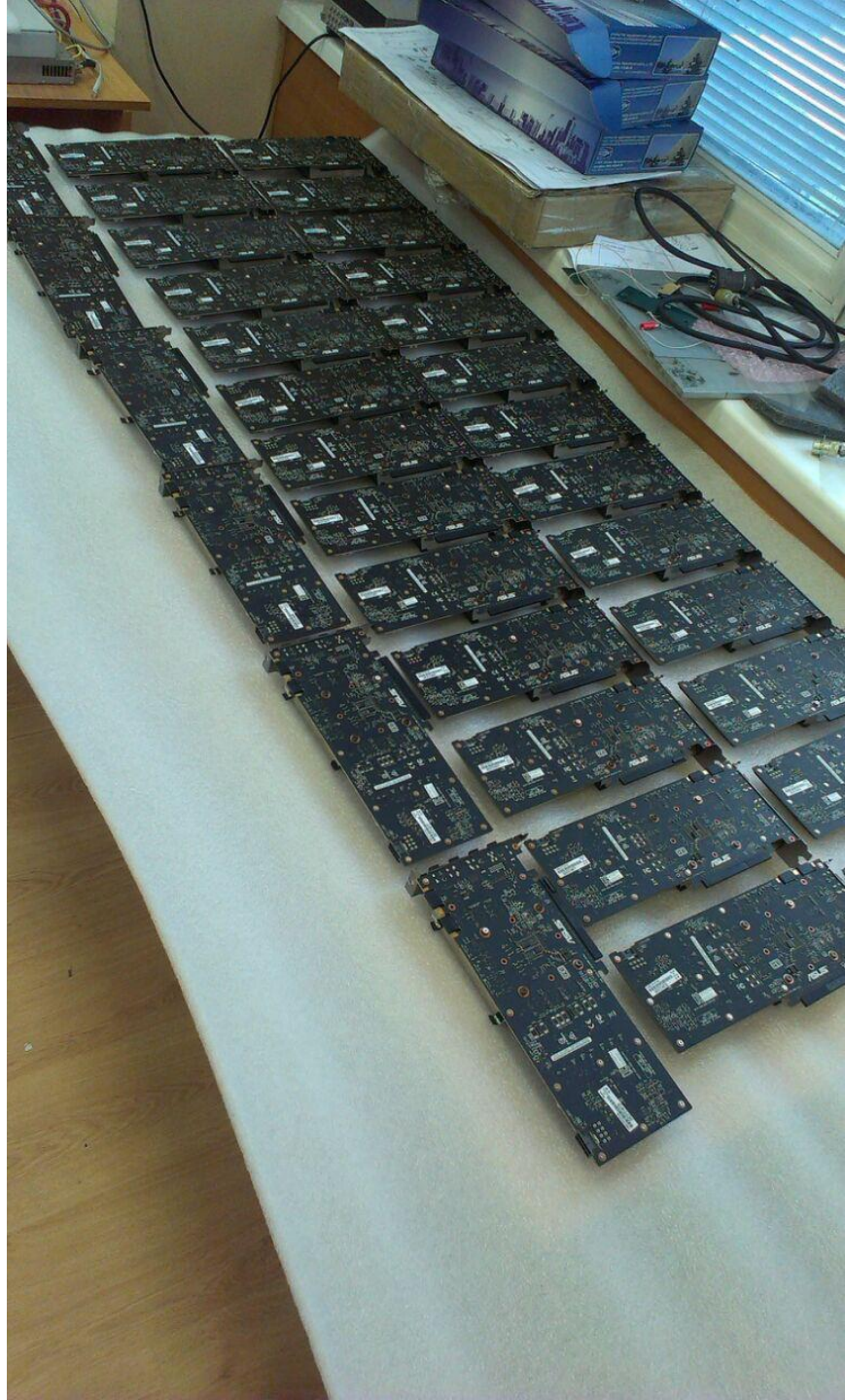


Конфигурация и стоимость компонентов одного узла суперкомпьютера Desmos (московские цены октября 2016 г.)

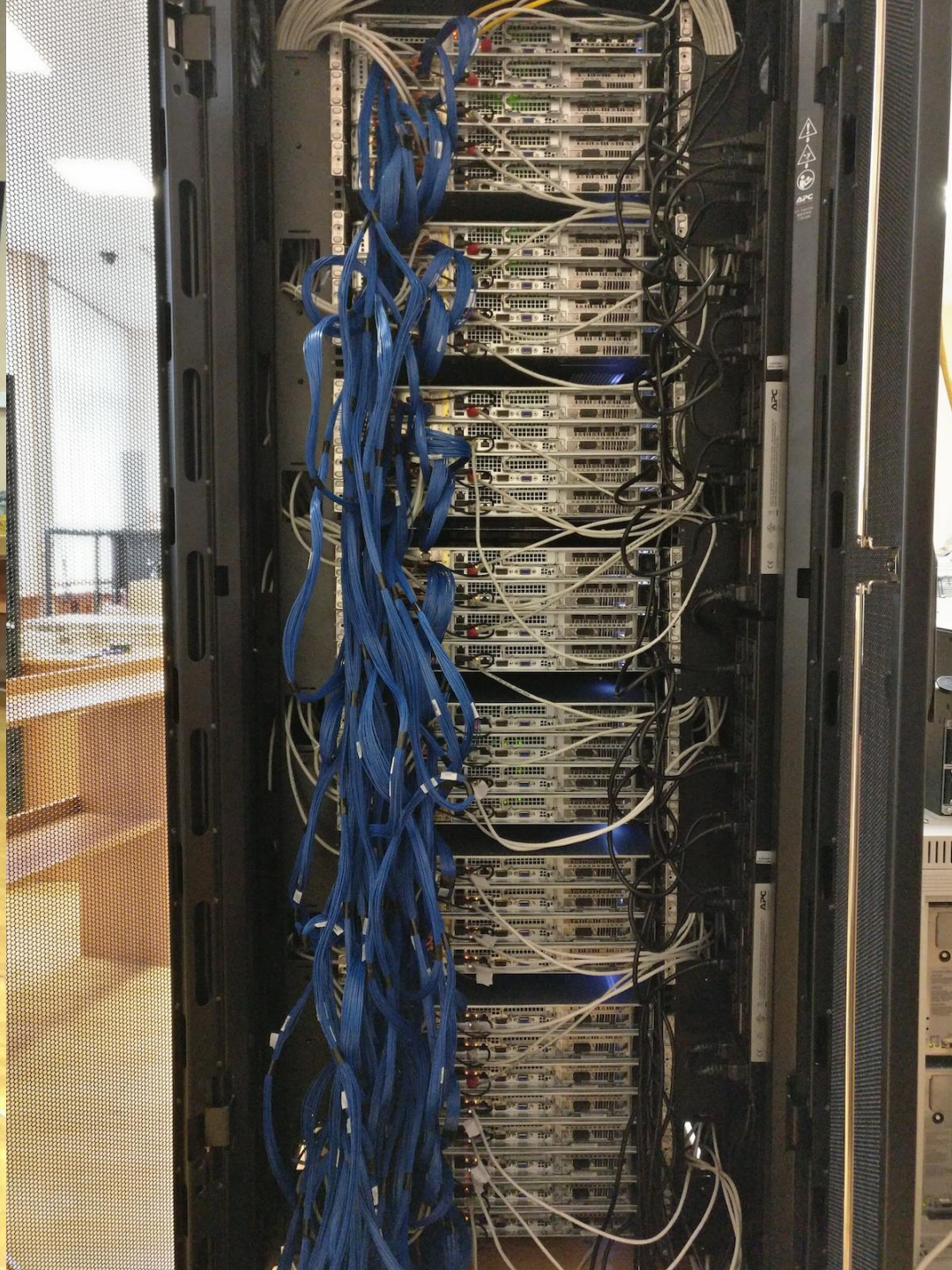
Компонент	Наименование	Цена
Chassis	SuperMicro SuperServer 1018GR-T	955 USD
CPU	Intel Xeon E5-1650 v3 (6 cores, 3.5 GHz)	700 USD
GPU	Nvidia GeForce GTX 1070 (8 Gb GDDR5)	550 USD
DRAM	DDR4 16 Gb	140 USD
HDD	HDD 500 Gb	44 USD
	В сумме (без интерконнекта):	2389 USD

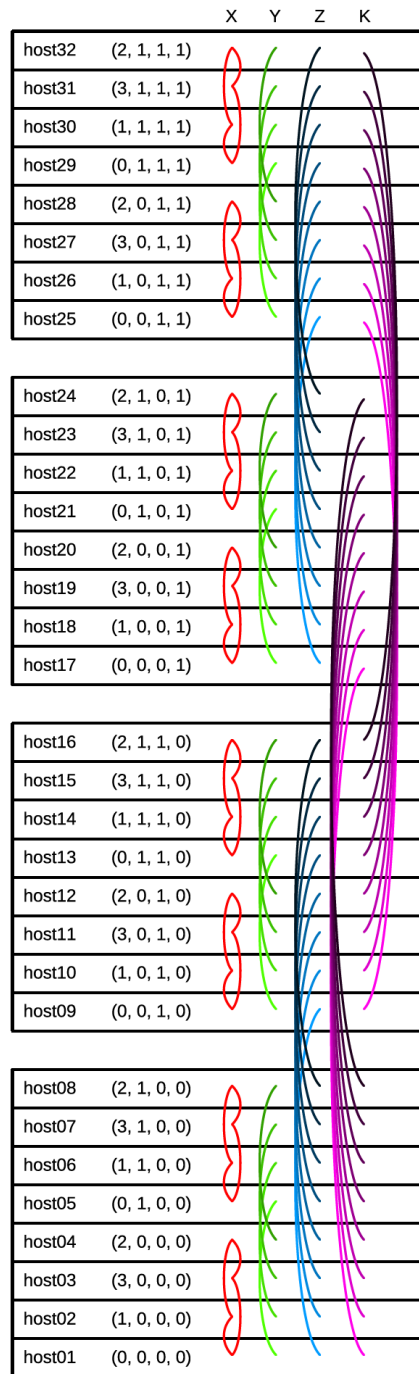
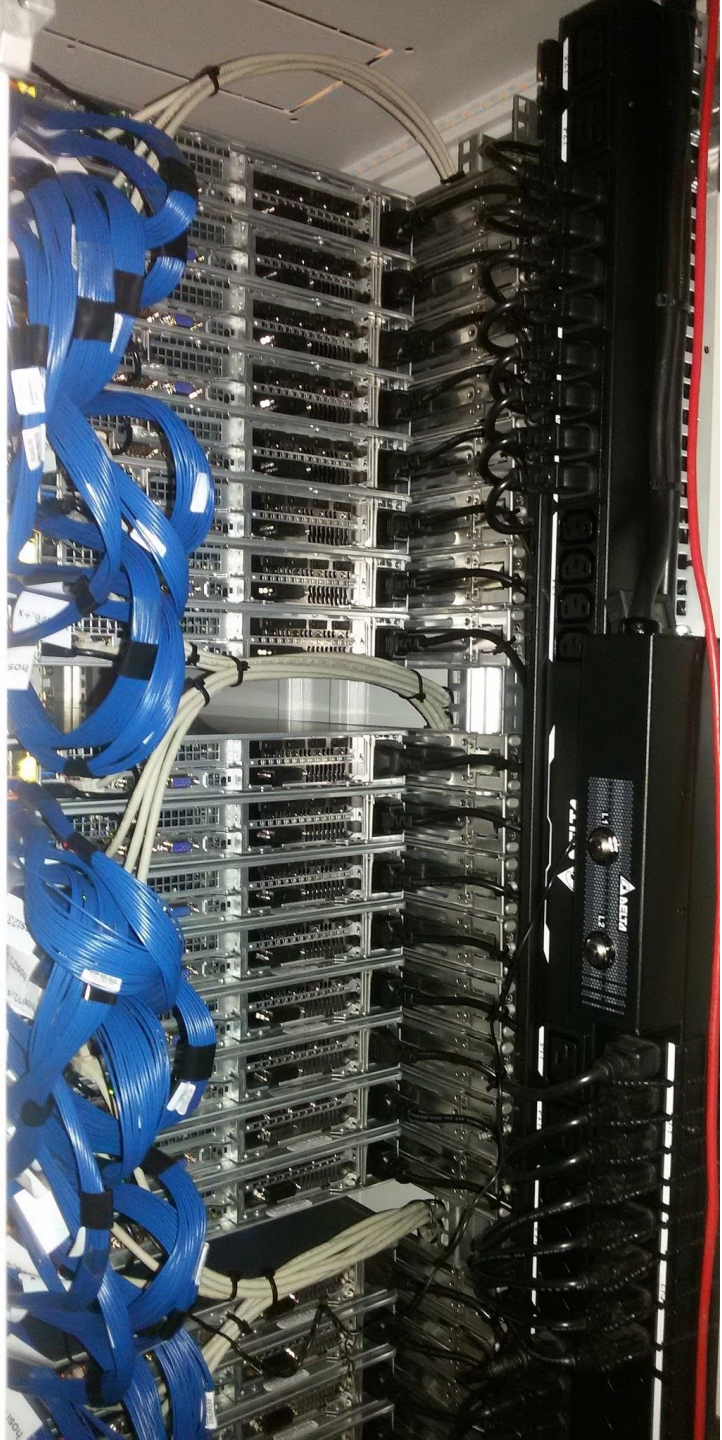
Карты Nvidia GeForce имеют высокую производительность в одинарной точности (4-10 Tflops/sec), которая может быть эффективно использована в МД-кодах LAMMPS, GROMACS и др.

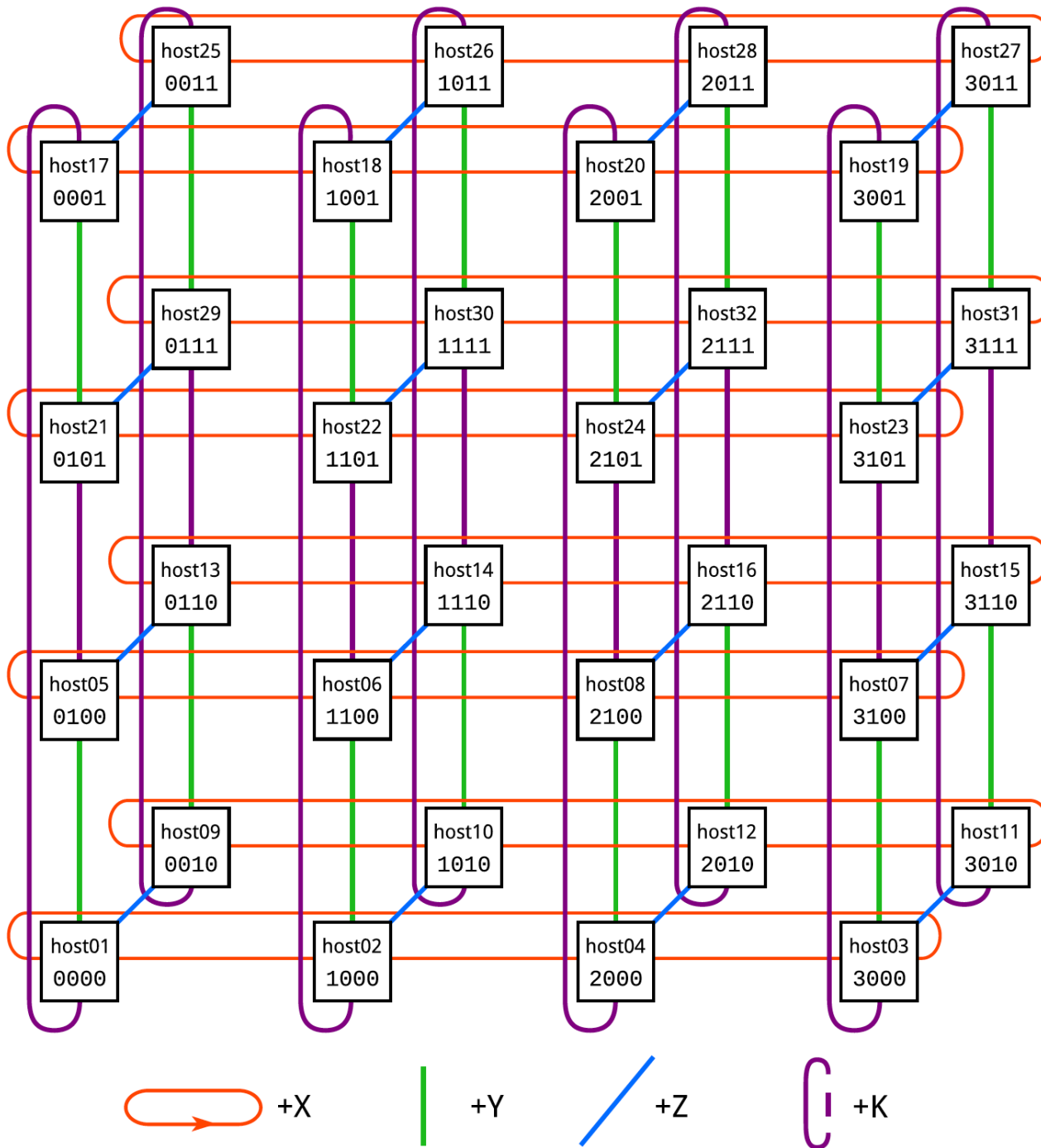
ПРОЦЕСС СБОРКИ











Некоторые характеристики суперкомпьютера Desmos

- Энергопотребление 6.5 – 14.4 кВт.
- Производительность HPL (без GPU) – 9.12 Тфлопс
- Операционная система SLES 11 SP4
- Angara MPI на базе MPICH 3.0.4
- Система управления очередью SLURM
- Ведется сбор статистики при помощи БД SLURM
- Для управления окружением компиляции и запуска используется пакет lmod

Политика системы очередей

- Установлены лимиты времени
- Введены разделы max1n , max8n , max16n , max32n
- Задачи на один узел вынесены в отдельный раздел
- Введены ограничения:
 - max1n : макс. время=1440 минут, мин. и макс. число узлов=1
 - max8n : макс. время=1440 минут, мин. узлов=4, макс. узлов=8
 - max16n : макс. время=720 минут, мин. узлов=4, макс. узлов=16
 - max32n : макс. время=360 минут, мин. узлов=4, макс. узлов=32

Сбор статистики

- Slurm MySQL plugin
- Контроль использования gpu (Generic resource (GRES) scheduling)
- Сбор статистики по подразделениям ОИВТ и внешним организациям

Сбор статистики

```
> sacct -j 28300
```

JobID	JobName	Partition	Account	AllocCPUS	State	ExitCode
28300	lu.C.8	max32n	hse	96	COMPLETED	0:0
28300.0	lu.C.8		hse	32	COMPLETED	0:0

Сбор статистики

```
> sreport cluster AccountUtilizationByUser END=2017-12-10T23:59:59
```

```
-----  
Cluster/Account/User Utilization 2017-12-10T00:00:00 - 2017-12-10T23:59:59 (86400 secs)  
Use reported in TRES Minutes  
-----
```

Cluster	Account	Login	Proper Name	Used	Energy
desmos	root			471160	0
desmos	jiht			471160	0
desmos	levashov_lab			138237	0
desmos	levashov_lab	mparamon+	mparamonov	138237	0
desmos	stegailov_lab			282073	0
desmos	stegailov_lab	dfleyta	dfleyta	5593	0
desmos	stegailov_lab	kondraty+	kondratyuk	276480	0
desmos	usachev_lab			50850	0
desmos	usachev_lab	vfilinov	vfilinov	50850	0

Управление окружением компиляции и запуска: Imod

- Использование файлов модулей для динамического управления переменными окружения пользователя
- Язык программирования Lua
- Контроль совместимости разных версий ПО
- Возможность создания пользовательских файлов модулей без участия администратора

Управление окружением КОМПИЛЯЦИИ И ЗАПУСКА

- binutils/2.29
- gcc/5.3.0
- make/4.2.1
- mpich/3.0.4
- fftw/3.3.6-pl2
- OpenBLAS/0.2.20
- scalapack/2.0.2
- elpa-2016.05.004
- elpa-2017.05.002
- libint/1.1.6
- libxc/3.0.0
- libxsmm/1.8.1
- abinit/8.4.3
- lammps/11Aug17

Выводы

Основной вектор развития СКЦ ОИВТ РАН – апробация новых высокоэффективных вычислительных инструментов для решения конкретных задач математического моделирования

Суперкомпьютер «Десмос» - первый и успешный опыт внедрения сети Ангара для гибридного кластера (CPU+GPU)