

Численные методы в механике сплошной среды.

Основы работы с кластером ИГиЛ

Давыдов М. Н. (ИГиЛ)

01.03.2018

Описание кластера

Кластер ИГиЛ в настоящее время расположен в Центре коллективного пользования «Сибирский суперкомпьютерный центр» и подключен как часть кластера «НКС-1П»

Состав кластера:

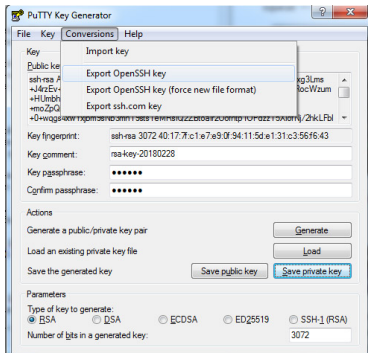
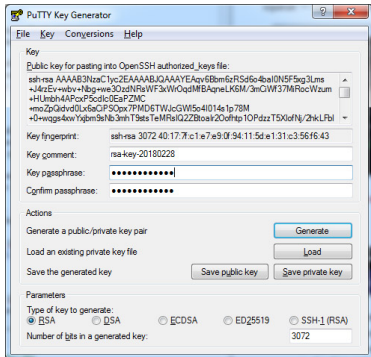
- (16 x) Вычислительные узлы KNL
1 × Xeon Phi 7290, 1.5 (1.7) ГГц (KNL) — 72 ядра × 4 потока — 288 аппаратных потока на узел
96 ГБ памяти на узел, 16 GB памяти MCDRAM (кэш)
3456 GFLOPS пиковая произв-ть
- (20 x) Вычислительные узлы Broadwell
2 × Xeon E5-2697A v4, 2.6 (3.6) ГГц (Broadwell) — 16 ядер × 2 потока — 32 аппаратных потока на узел
128 ГБ памяти на узел
1331.2 GFLOPS пиковая произв-ть
- **(7 x) Вычислительные узлы Broadwell**
2 × Xeon (Broadwell)
16 ядер × 2 потока — 32 аппаратных потока на узел
256 ГБ памяти на узел

- Авторизация не по паролю, а по ключу.
- Ключ состоит из открытой и закрытой части.
- Открытая хранится на сервере, и не является секретной.
- Закрытая хранится у пользователя и является секретной.
- Авторизуется не только клиент на сервере, но и сервер по отношению к клиенту!
- Ключ не передаётся с клиента на сервер, а во время авторизации клиент доказывает серверу, что владеет ключом.
- Первый раз, когда вы заходите на сервер, ssh вас спрашивает, доверяете ли вы ключу. Если отвечаете нет, соединение закрывается.
- Если ключ сервера поменялся (например, сервер переустановили), ssh вопит от подделке ключа.
- Риск атаки «man in the middle» — если нет ошибки с IP или обновления сервера, а ключ поменялся — это повод поднять уровень паранойи.

- Закрытая часть ключа может быть в одном из двух форматов: open-ssh или в формате putty (.ppk). Возможна конвертация из одного формата в другой.
- Пароль не привязан к ключу!
- Пароль защищает только сам файл ключа.
- Пароль можно менять, убирать или устанавливать. Переписывать ключ на сервере не требуется.

Генерация ключей в Windows

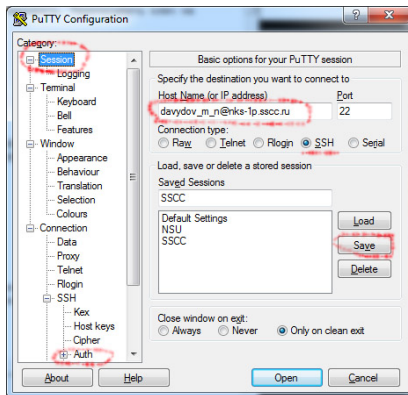
Для работы с ключами используется программа **PuTTYgen**.



Восстановить забытый пароль невозможно!

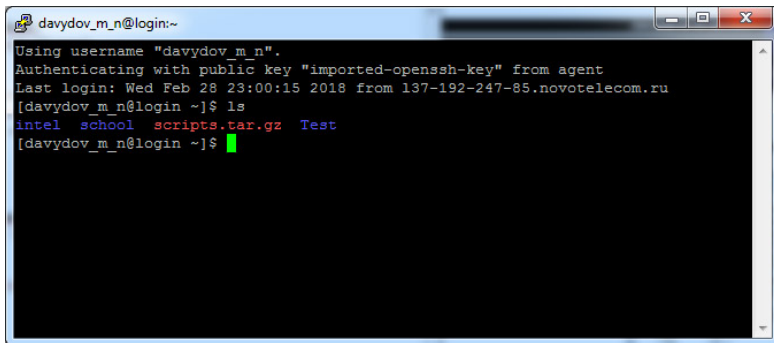
Доступ к кластеру в Windows

Для доступа используется программа *PuTTY*.



Доступ к кластеру в Windows

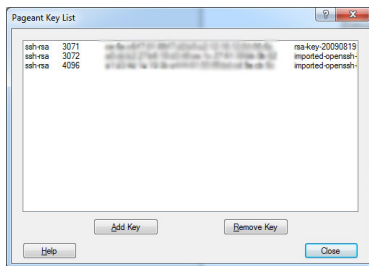
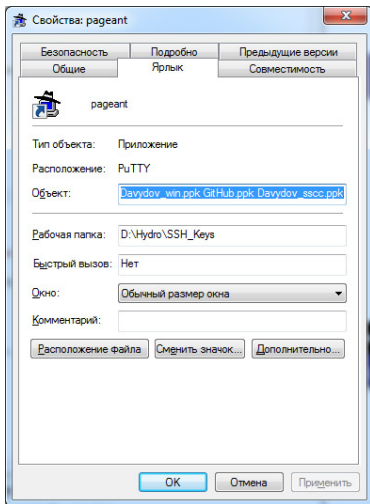
Открытый терминал кластера.



```
davydov_m_n@login:~
Using username "davydov_m_n".
Authenticating with public key "imported-openssh-key" from agent
Last login: Wed Feb 28 23:00:15 2018 from 137-192-247-85.novotelecom.ru
[davydov_m_n@login ~]$ ls
intel  school  scripts.tar.gz  Test
[davydov_m_n@login ~]$
```

Pageant

Для работы с ключами (загрузка, выгрузка) имеется программа **Pageant**.



Microsoft OpenSSH Client

Пользователи Microsoft Windows 10 могут установить Microsoft OpenSSH Client.

Пуск → Параметры → Приложения → Управление дополнительными компонентами → Добавить компонент → "OpenSSH Client".

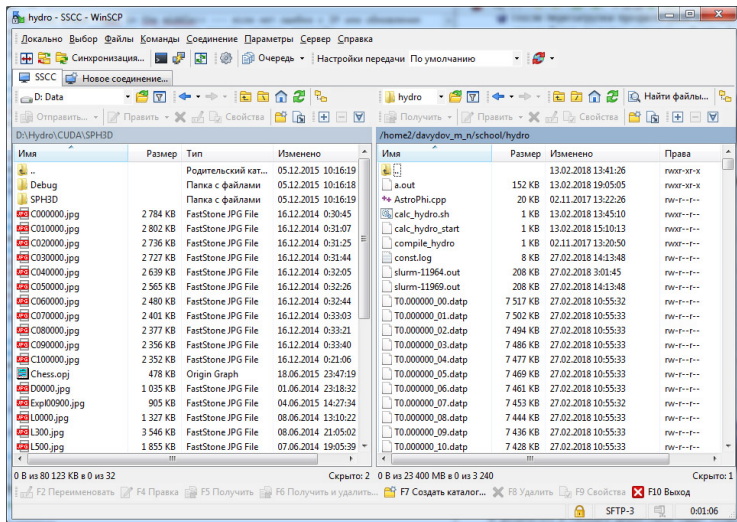
После перезагрузки профиля пользователя (или компьютера) в командной строке будут доступны команды: `ssh`, `ssh-add`, `ssh-agent`, `ssh-keygen`, `scp`, `sftp`

Пользователи Microsoft Windows 10 могут установить Linux для Windows (доступны Ubuntu и SUSE). Предоставляет полный функционал Linux.

- 1 Установка компонента «Подсистема Windows для Linux»: «Панель управления» → «Программы и компоненты» → «Включение или отключение компонентов Windows» → «Подсистема Windows для Linux».
- 2 Через магазин приложений Windows (Microsoft Store) установить приложение "Ubuntu" (или "openSUSE" или "SUSE Linux Enterprise Server").
- 3 После перезагрузки профиля (или компьютера) в командной строке будут доступны команды: `ubuntu`, `bash`, и т.д.

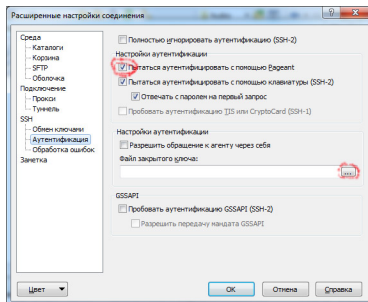
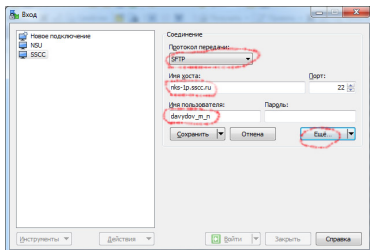
Передача файлов в Windows

Для передачи файлов используется программа **WinSCP**



Передача файлов в Windows

Настройка соединения в *WinSCP*



Доступ к кластеру в Linux

Генерация ключей:

```
ssh-keygen -t rsa -b 3072 -f newkey
```

В результате в текущей директории будут созданы файлы `newkey` — закрытая часть в формате `open-ssh` и `newkey.pub` — открытая часть.

Соединение с кластером:

```
ssh -i newkey login@nks-1p.sssc.ru
```

Копирование файлов в терминале:

Копируем файл `"file.txt"` с удаленного сервера на локальный компьютер.

```
scp -i newkey login@nks-1p.sssc.ru:file.txt /local/directory
```

Копируем файл `"file.txt"` с локального компьютера на удаленный сервер.

```
scp -i newkey file.txt login@nks-1p.sssc.ru:/remote/directory
```

Сменить пароль на ключ:

```
ssh-keygen -p
```

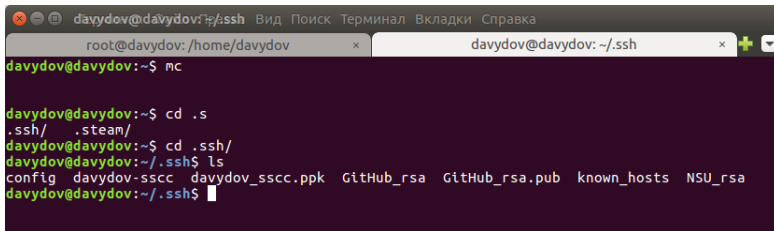
ssh-agent позволяет вводить пароль к ключу только один раз.

`ssh-agent roxterm` — запускаем через него терминал

`ssh-add ~/.ssh/id_dsa` — подгружаем ключ.

Доступ к кластеру в Linux

Закрытые части ключей можно поместить в каталог `.ssh` (скрытый каталог в домашней папке пользователя)

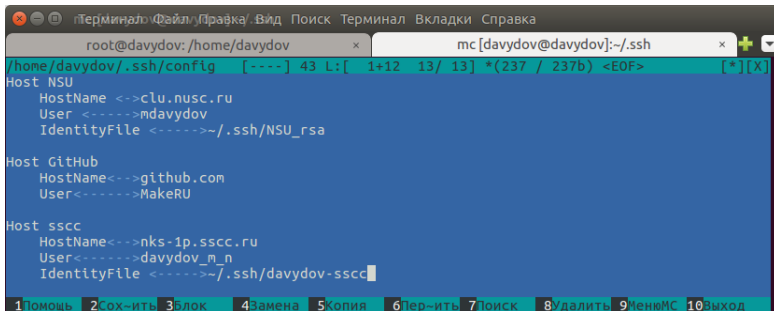


```
davydov@davydov:~$ mc
davydov@davydov:~$ cd .
.dash/ .steam/
davydov@davydov:~$ cd .ssh/
davydov@davydov:~/.ssh$ ls
config davydov-sscc davydov-sscc.ppk  Github_rsa  Github_rsa.pub  known_hosts  NSU_rsa
davydov@davydov:~/.ssh$
```

Права доступа к файлам ключей должны быть только на чтение и запись владельцу!

Доступ к кластеру в Linux

Написать файл `~/.ssh/config`



```
root@davydov: /home/davydov
mc [davydov@davydov]: ~/.ssh
/home/davydov/.ssh/config [----] 43 L: [ 1+12 13/ 13] *(237 / 237b) <EOF> [*][X]
Host NSU
  HostName <-->clu.nusc.ru
  User <----->mdavydov
  IdentityFile <----->~/.ssh/NSU_rsa

Host GitHub
  HostName<-->github.com
  User<----->MakeRU

Host sssc
  HostName<-->nks-1p.sssc.ru
  User<----->davydov_m_n
  IdentityFile <----->~/.ssh/davydov-sssc
```

Теперь для доступа достаточно команды:

```
ssh sssc
```

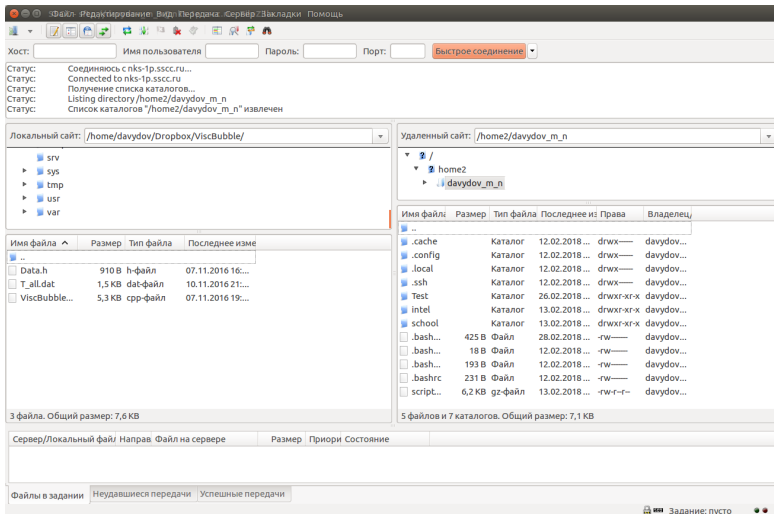
Или с использованием скрипта:

```
#!/bin/bash
```

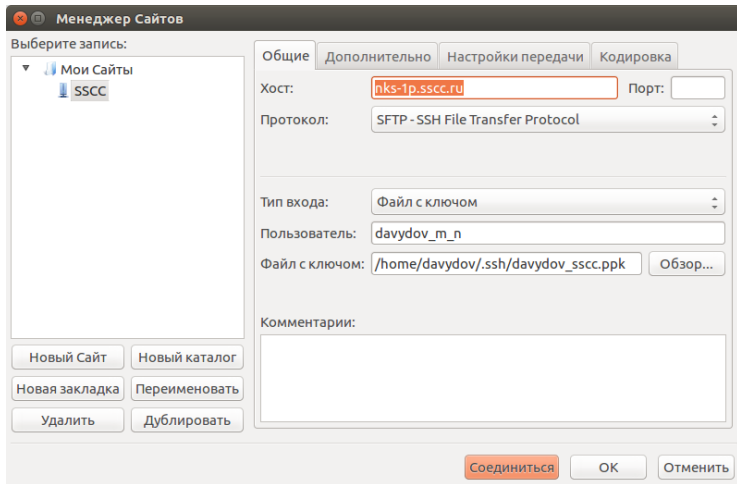
```
gnome-terminal --tab -e "ssh sssc"
```

Передача файлов в Linux

Для передачи файлов используется программа **FileZilla**



Передача файлов в Linux



Для доступа FileZilla требует ключ в формате ppk.

Этапы работы с кластером

- ❶ Интерактивный вход пользователя на консоль сервера управления по протоколу ssh.
- ❷ Получение доступа к вычислительным узлам через планировщик задач SLURM:
 - В интерактивном режиме пользователь запрашивает у планировщика требуемое количество вычислительных узлов, после чего ожидает их выдачи. В случае успешного выделения узлов планировщиком (о чем система сообщит в консоли текущей сессии), пользователь может получить прямой ssh-доступ к выданным узлам на запрошенное время;
 - В пакетном режиме запуск осуществляется с помощью сценария, представляющего собой shell скрипт. Планировщик размещает сценарий в очередь планирования и сам принимает решение о дате и месте ее запуска. Скрипт будет запущен на первом из выделенных узлов.

Каждый узел выделяется пользователю в единоличное пользование в рамках конкретной задачи.

Пользовательская директория находится по адресу
/home/fano.hydro/<логин пользователя>.

Домашние директории пользователей hydro находятся на отдельной
СХД (системе хранения данных).

***Подключать отдельно рабочие области для временных
файлов, как на кластере НГУ не нужно.***

На кластере существуют следующие общие директории, в которых
установлены пользовательские программы:
/opt/

Environment Modules

Для управления множественными версиями различных прикладных программных пакетов и библиотек на вычислительной системе установлен пакет ***Environment Modules***.

Каждый модуль содержит информацию, необходимую для настройки окружения под конкретное приложение. Настройка осуществляется через задание переменных PATH, MANPATH, INCLUDE, LD_LIBRARY_PATH и т.д.

\$ module list — список подгруженных модулей.

\$ module avail — получение списка доступных модулей.

\$ module load <modulefile> — загрузка модуля.

\$ module unload <modulefile> — выгрузка модуля.

\$ module initadd <modulefile> — добавление модуля в автозагрузку.

\$ module initrm <modulefile> — удаление модуля из автозагрузки.

Environment Modules

```
davydov_m_n@login/opt
Using username "davydov_m_n".
Authenticating with public key "imported-openssh-key" from agent
Last login: Tue Feb 27 10:53:24 2018 from 137-192-247-85.novotelecom.ru
[davydov_m_n@login ~]$ mc

[davydov_m_n@login opt]$ module list
No Modulefiles Currently Loaded.
[davydov_m_n@login opt]$ module avail

----- /usr/share/Modules/modulefiles -----
dot          module-git  module-info  modules      null          use.own

----- /etc/modulefiles -----
mpi/openmpi-x86_64

----- /opt/basis/modules -----
compilers/intel/2017.4.196      libraries/mkl/2018.1.163
compilers/intel/2018.1.163     libraries/tbb/2017.4.196
debuggers/inspector/2017.1.3.510645  libraries/tbb/2018.1.163
debuggers/inspector/2018.1.1.535159  module-info
debuggers/intel/2017            modules
debuggers/intel/2018.1.163      mpi/intel/2018.1.163
intel/2017.4.196(default)      null
intel_license                  parallel/mpi.intel/2017.4.196
intel_license_2018             parallel/mpi.intel.broadwell/2017.4.196
intelpython2                   parallel/mpi.intel.knl/2017.4.196
intelpython3                   profilers/advisor/2017.1.3.510716
libraries/daal/2017.4.196      profilers/advisor/2018.1.1.535164
libraries/daal/2018.1.163      profilers/amplifier/2017.3.0.510739
libraries/ipp/2017.4.196       profilers/amplifier/2018.1.0.535340
libraries/ipp/2018.1.163       profilers/tac/2018.1.163
libraries/mkl/2017.4.196       profilers/tac/intel64/2017.3.030
[davydov_m_n@login opt]$
```

Для компиляции прикладного программного обеспечения на кластере установлен стандартный набор библиотек и компиляторов из пакета ***Intel Parallel Studio Cluster Edition***.

Для его использования необходимо подгрузить соответствующий модуль.

В частности, и си и фортран подключаются так:

```
$ module load compilers/intel/2018.1.163
```

Настройки компиляции зависят от того на каких узлах будет запущена программа.

- Для архитектуры Broadwell:

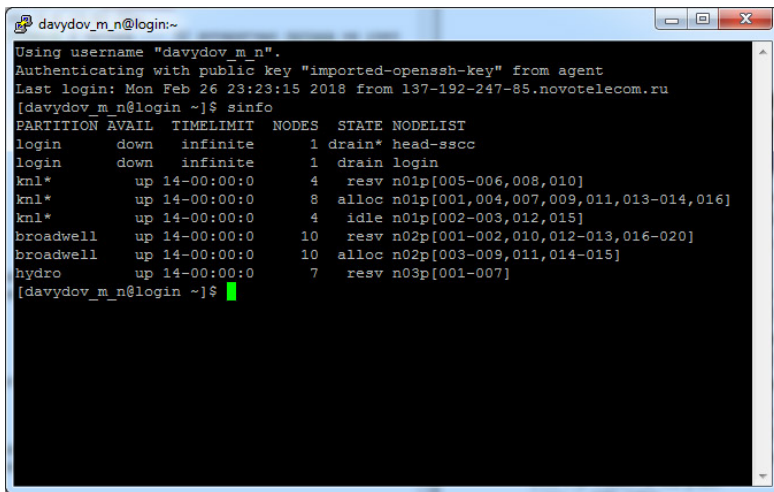
```
1. module purge
2. module load intel/2017.4.196
   parallel/mpi.intel.broadwell/2017.4.196
   compilers/intel/2017.4.196
3. mpiicc -qopenmp -O3 -o a.out AstroPhi.cpp -lm
```

- Для архитектуры KNL:

```
1. module purge
2. module load intel/2017.4.196
   parallel/mpi.intel.knl/2017.4.196
   compilers/intel/2017.4.196
3. mpiicc -xMIC-AVX512 -qopenmp -O3 -o a.out
   AstroPhi.cpp -lm
```

Основные возможности Slurm

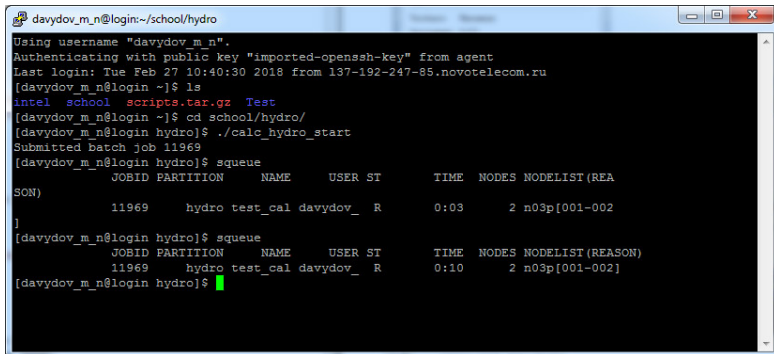
sinfo — просмотр ресурсов.



```
davydov_m_n@login:~  
Using username "davydov_m_n".  
Authenticating with public key "imported-openssh-key" from agent  
Last login: Mon Feb 26 23:23:15 2018 from 137-192-247-85.novotelecom.ru  
[davydov_m_n@login ~]$ sinfo  
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST  
login      down    infinite     1 drain* head-sscc  
login      down    infinite     1 drain login  
knl*       up      14-00:00:0     4  resv n01p[005-006,008,010]  
knl*       up      14-00:00:0     8  alloc n01p[001,004,007,009,011,013-014,016]  
knl*       up      14-00:00:0     4  idle  n01p[002-003,012,015]  
broadwell  up      14-00:00:0    10  resv n02p[001-002,010,012-013,016-020]  
broadwell  up      14-00:00:0    10  alloc n02p[003-009,011,014-015]  
hydro      up      14-00:00:0     7  resv n03p[001-007]  
[davydov_m_n@login ~]$
```


Основные возможности Slurm

squeue — просмотр своей очереди задач.



```
davydov_m_n@login:~/school/hydro
Using username "davydov_m_n".
Authenticating with public key "imported-openssh-key" from agent
Last login: Tue Feb 27 10:40:30 2018 from 137-192-247-85.novotelecom.ru
[davydov_m_n@login ~]$ ls
intel  school  scripts.tar.gz  Test
[davydov_m_n@login ~]$ cd school/hydro/
[davydov_m_n@login hydro]$ ./calc_hydro_start
Submitted batch job 11969
[davydov_m_n@login hydro]$ squeue
      JOBID PARTITION     NAME     USER ST       TIME  NODES NODELIST(REASON)
      11969      hydro test_cal davydov_ R        0:03      2 n03p[001-002]
[davydov_m_n@login hydro]$ squeue
      JOBID PARTITION     NAME     USER ST       TIME  NODES NODELIST(REASON)
      11969      hydro test_cal davydov_ R        0:10      2 n03p[001-002]
[davydov_m_n@login hydro]$
```

В отличие от кластера НГУ просмотр всей очереди недоступен.

Основные возможности Slurm

- `scancel N` — удаление задачи с номером `N` из очереди задач.
- `salloc` — интерактивное выделение вычислительных узлов.
 - `-N`, `-nodes` указывает количество необходимых узлов;
 - `-t`, `-time` время доступности выделенных ВУ (в минутах);
 - `-p`, `-partition` выделение ресурсов в указанной партии.
- `sbatch` — пакетный запуск задач.
 - `-N`, `-nodes` указывает количество необходимых узлов;
 - `-n`, `-ntasks` общее количество запущенных процессов;
 - `-ntasks-per-node` задает количество процессов, запускаемых на каждом вычислительном узле;
 - `-t`, `-time` время доступности выделенных ВУ (в минутах);
 - `-p`, `-partition` выделение ресурсов в указанной партии.
- `srun` — интерактивный запуск задач.
- `scontrol show job <JOBID>` — подробная информация о задаче.

Для удобства использования планировщика задач SLURM на кластере используется модуль автодополнения команд (`bash completion`). Работает по двойному `<tab>`.

Запуск программ (Broadwell)

Для запуска программ рекомендуется подготовить скрипт с настройками запуска и скрипт для постановки в очередь.

- Скрипт с настройками запуска (calc_hydro.sh):

```
#!/bin/bash
# set the number of nodes
#SBATCH --nodes=2
# hyperthreading off
#SBATCH --threads-per-core=1
# set max wallclock time
#SBATCH --time=6-0
# set name of job
#SBATCH --job-name=test_calculation
# set queue name
#SBATCH -p hydro
#SBATCH --reservation=root_6
# run the application
mpirun -n=32 ./a.out
```

Запуск программ (Broadwell)

Для запуска программ рекомендуется подготовить скрипт с настройками запуска и скрипт для постановки в очередь.

- Скрипт постановки в очередь (calc_hydro_start):

```
module purge
module load intel/2017.4.196
parallel/mpi.intel.broadwell/2017.4.196
sbatch ./calc_hydro.sh
```

Запуск задач на узлах архитектуры Intel Xeon Phi (KNL)

Поддерживаются различные режимы работы на процессорах KNL.
MCDRAM:

- cache — Вся память MCDRAM используется как кэш
- equal — Память MCDRAM будет объединена с основной памятью
- flat — MCDRAM будет объединена с основной памятью в flat режиме

NUMA:

- a2a — Все ко всем
- hemi — Полусфера
- snc2 — Sub-NUMA кластера 2
- snc4 — Sub-NUMA кластера 4
- quad — Квадрант

Запуск задач на узлах архитектуры Intel Xeon Phi (KNL)

```
davydov_m_n@login:~
Using username "davydov_m_n".
Authenticating with public key "imported-openssh-key" from agent
Last login: Tue Feb 27 14:53:00 2018 from 137-192-247-85.novotelecom.ru
[davydov_m_n@login ~]$ sinfo -p knl
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
knl*      up 14-00:00:0    2  drain* n01p[006,008]
knl*      up 14-00:00:0    2   resv n01p[005,010]
knl*      up 14-00:00:0   12  alloc n01p[001-004,007,009,011-016]
[davydov_m_n@login ~]$ sinfo -o "%30N %20b %f" -p knl
NODELIST
n01p[005-006,008,010,012]  ACTIVE_FEATURES  AVAIL_FEATURES
n01p001                  snc4,hybrid,knl  a2a,hemi,quad,snc2,snc4,cache,flat,hybrid,auto,knl
n01p[002-004,011,013,016] snc2,flat,knl    a2a,hemi,quad,snc2,snc4,cache,flat,hybrid,auto,knl
n01p[007,009,014-015]    hemi,cache,knl   a2a,hemi,quad,snc2,snc4,cache,flat,hybrid,auto,knl
n01p[007,009,014-015]    quad,cache,knl   a2a,hemi,quad,snc2,snc4,cache,flat,hybrid,auto,knl
[davydov_m_n@login ~]$
```

Для запуска задач на узлах с требуемой конфигурацией при аллоцировании необходимо указать опцию `--constraint`.

Выделенные для аллоцирования узлы будут перезагружены для применения необходимой конфигурации.