Построение и администрирование кластерных систем

Жуматий С.А.
serg@parallel.ru
в.н.с. НИВЦ МГУ им.М.В.Ломоносова
2017

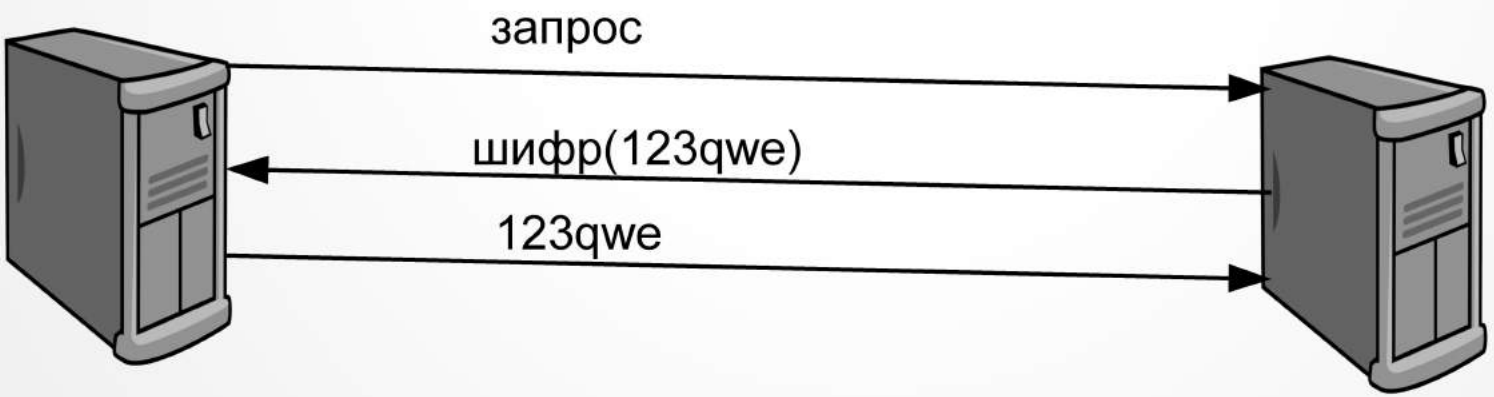
План

- ssh, sftp, putty
- Учётные записи
- Управление узлами
- Загрузка
- Файловые системы
- Управление пользователями

SSH

закрытый

открытый



OpenSSH

/etc/ssh/sshd_config

- AllowUsers / AllowGroups (DenyUsers/...)
- X11forward
- UseLogin
- Banner
- PermitRootLogin without-password
- PubkeyAuthentication yes
- PasswordAuthentication no

OpenSSH 2

Match:

User, Group, Host, Address

AuthorizedKeysFile, Banner, ChrootDirectory,

ForceCommand,MaxAuthTries, **MaxSessions**,

PasswordAuthentication,

PermitEmptyPasswords, **PermitRootLogin**,

PubkeyAuthentication, X11Forwarding

OpenSSH 3

- AcceptEnv / SendEnv
- Match Address 10.0.0.0/8
  - PasswordAuthentication yes
  - KbdInteractiveAuthentication yes
  - PermitRootLogin yes

OpenSSH 4

.ssh/authorized_keys:

command="....." ssh-rsa r123hbedf123h12g...

.ssh/config:

```
Host work
    Address 1.2.3.4
    User pupkin
    IdentityFile ~/.ssh/work
```

OpenSSH 5

- Доступ без паролей:

ssh-keygen -N '' -q [-o ~/.ssh/id_rsa]

id_rsa = закрытый ключ,

id_rsa.pub = открытый ключ

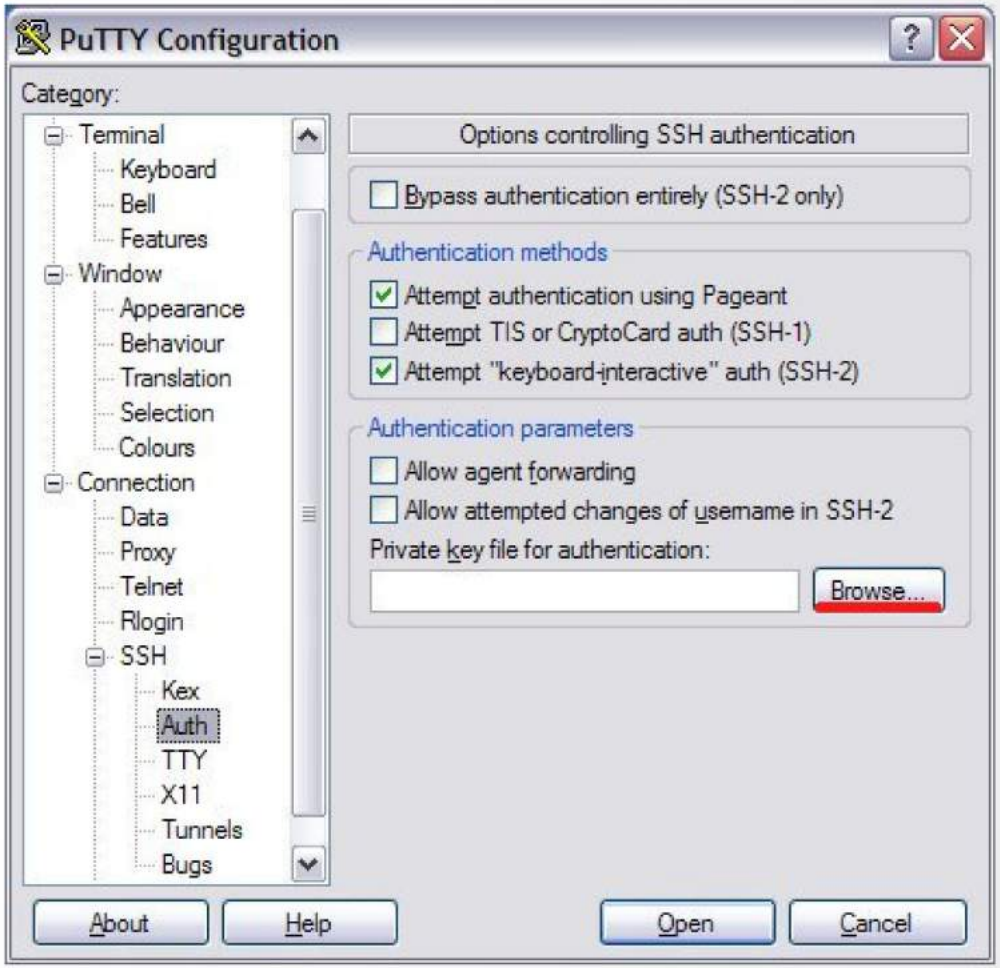
.ssh/authorized_keys — открытые ключи

OpenSSH 6

~?	**список команд**
~~	**~**
~^Z	**в фон**
~.	**завершение**

Putty

- Putty / puttygen



Удалённый доступ: sftp/scp

- Linux: scp / mc / Dolphin / Nautilus / Filezilla / …
- +sshfs
- Windows: FAR / WinSCP / Filezilla / …
- ssh-agent / pagent

Учётные записи

- LDAP
- NIS+
- Passwd + rsync

Учётные записи: LDAP

- phpLDAPadmin
- LDAP Account Manager
- OpenLDAP / Fedora Directory Server / Apache Directory Project / Mandriva Directory Server

Учётные записи: NIS+

- Не делайте так :)

Учётные записи: passwd

passwd, shadow, hosts, tcb, group, …

Rsync — только с узлов!

$ rsync -a server:/etc/passwd /etc/passwd

Управление узлами

- ssh
- ssh + bash
- pdsh
- **IPMI / ILO**
- **iKVM**

PDSH

```
$ pdsh -w 'node-[01-10]' 'w| grep loadaverage'
```

-a	выполнить команду на всех машинах, перечисленных в hostfile (см. ниже)
-w	выполнить команду на перечисленных узлах. Список задаётся через запятую (без пробелов), можно использовать диапазоны чисел, например node-[10-20]. Если в качестве списка указан '-', то список читается со стандартного ввода.
-x	исключить перечисленные узлы
-g	запустить команду на узлах перечисленных групп, список групп задаётся через запятую

PDSH

-f	задать число параллельных потоков исполнения
-u	задать таймаут в секундах выполнения команды на узле (по умолчанию таймаута нет)
-l	Выполнять команду от имени указанного пользователя (аналогично ssh)
-b	Прекращать выполнение по нажатию Ctrl-C (по умолчанию по нажатию Ctrl-C выдаётся текущий статус выполнения, а по второму выполнение прекращается)

PDSH

/etc/pdsh/machines

/etc/**dsh**/group/{g1,g2,g3...}
~/**.dsh**/group/{g1,g2,g3...}

WCOLL = имя файла со списком узлов

Screen (tmux)

 # screen -DRR
 # screen -list
 # screen -S имя_сессии / -r имя_сессии

Ctrl-A

c = create
k = kill
n = next
p = previous
d = detach
a = Ctrl-A

Screen

Ctrl-A

0-9	=	в окно с номером...
ESC	=	режим поиска/выделения
>	=	записать буфер в файл
h/H	=	hardcopy/start log
C-x	=	lock
t	=	время, хост, загрузка

IPMI

ipmitool

modprobe ipmi_devintf
ipmi_si
ipmi_msghandler

ipmitool COMMAND

ipmitool -I lan/lanplus -H host -U user
-P password COMMAND

IPMI

Команды:

lan	настройка сети
power	on/off/reset/cycle
mc	управление контроллером
sensor	печать датчиков
sol	activate
user	управление пользователями
channel	настройка каналов
session	информация о сессии
shell	ввод команд интерактивно

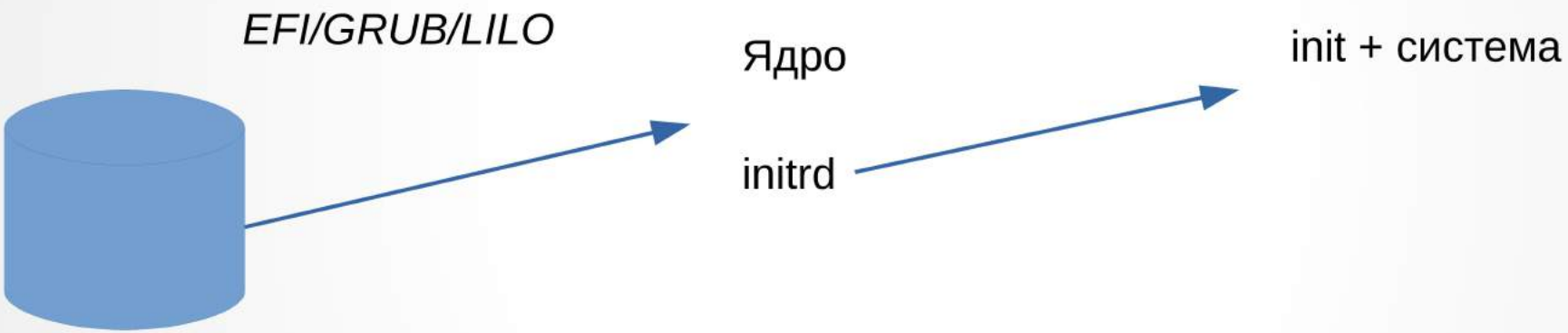
IPMI

```
# ipmitool lan print
Set in Progress         : Set Complete
Auth Type Support       : NONE MD2 MD5 PASSWORD
Auth Type Enable        : Callback : NONE MD2 MD5 PASSWORD
                        …
                        : OEM      : NONE MD2 MD5 PASSWORD
IP Address Source       : Static Address
IP Address              : 172.21.30.111
Subnet Mask             : 255.255.255.0
MAC Address             : 00:e0:81:46:0b:7d
Default Gateway IP      : 0.0.0.0
Default Gateway MAC     : 00:00:00:00:00:00
Backup Gateway IP       : 0.0.0.0
Backup Gateway MAC      : 00:00:00:00:00:00
802.1q VLAN ID          : Disabled
802.1q VLAN Priority    : 0
```

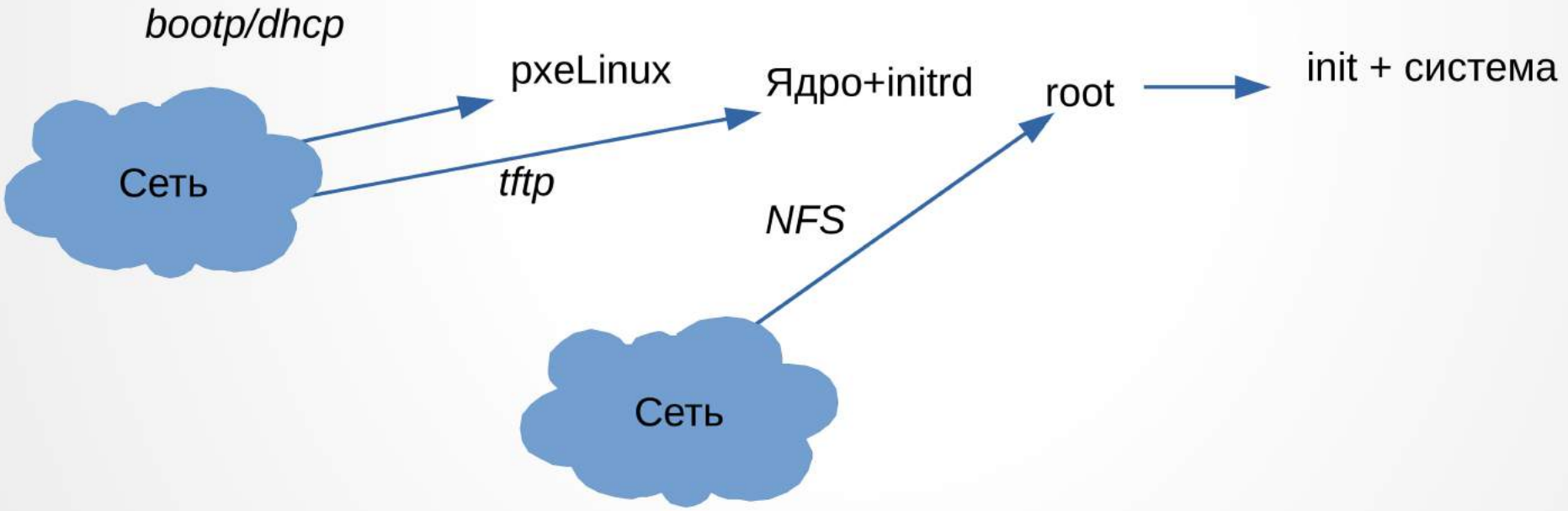
IPMI

```
# ipmitool lan set

usage: lan set <channel> <command> [option]

LAN set commands:
  ipaddr <x.x.x.x>                Set channel IP address
  netmask <x.x.x.x>               Set channel IP netmask
  macaddr <x:x:x:x:x:x>           Set channel MAC address
  defgw ipaddr <x.x.x.x>          Set default gateway IP address
  defgw macaddr <x:x:x:x:x:x>     Set default gateway MAC address
  bakgw ipaddr <x.x.x.x>          Set backup gateway IP address
  bakgw macaddr <x:x:x:x:x:x>     Set backup gateway MAC address

  password <password>     Set session password for this channel
```

Загрузка



Сетевая загрузка



Сетевая загрузка

- Одновременная загрузка
- Каталоги /var, /tmp, /etc
- Раздел подкачки
- Имя узла
- Syslog

DHCP

DHCP (Dynamic Host Configuration Protocol)
BOOTP

Client	Server
DHCPDISCOVER	DHCPOFFER
DHCPREQUEST	DHCPACK

DHCP (ISC)

```
max-lease-time 1200;
default-lease-time 600;
ddns-update-style none; ddns-updates off;

subnet 10.0.0.0 netmask 255.0.0.0 {
        range 10.0.5.0 10.0.5.100;
        allow unknown-clients;

        allow bootp;
        filename "pxelinux.0";
        next-server 10.0.0.2;
        option routers 10.0.0.2;
        option ntp-servers 10.2.0.1;
        option domain-name-servers 10.0.0.1,10.0.0.2;
        host monitor-1 {
                fixed-address 10.0.0.11;
                hardware ethernet 00:30:48:7E:00:42;
                option host-name monitor-1;
        }
}
```

PXE+TFTP

Syslinux → /var/lib/tftpboot

pxelinux.cfg/default - файл конфигурации:

```
default myimage
prompt 1
timeout 5

label myimage
  kernel img1/vmlinuz
  append initrd=img1/initrd ip=dhcp root=/dev/nfs \
      nfsroot=10.0.0.2:/noderoot.new \
      console=tty0 console=ttyS1,115200n8
```

Хранение данных

- Файловый сервер
- Сервер архивирования

✔ Дисковый массив - NAS / распределённая ФС

✗ iSCSI / ATA over Ethernet / ...

Сетевые файловые системы

- NFS
- PanFS
- Lustre
- GPFS
- GlusterFS
- GFS
- Hadoop/GoogleFS/...

NFS

RPC, portmap

portmap, nfsd, mountd

Число nfsd-процессов ~ число активных клиентов

Файловые системы

NFS

/etc/exports: список экспортируемых файловых систем

/path/to/fs кому_можно(опции) [кому_можно(опции) ...]

кому_можно: *, 1.2.3.4/24

опции:
(no_)root_squash
sync/async
no_subtree_check
ro/rw

exportfs -v / -r

NFS

/etc/exports

/export/dir host1,host2,host3(rw,root_squash)

all_squash	все пользователи клиента будет иметь права nobody (см. ниже) на сервере
anonuid	пользователь, права которого будут даны клиентам при операциях root_squash или all_squash (по умолчанию — nobody)
anonguid	аналогично anonuid, но для группы
(no_)subtree_check	не производить проверки прав пользователей в каталогах выше точки монтирования
async	позволить серверу подтверждать операции до их реального выполнения
sync	подтверждать операции только после их выполнения.

NFS

/etc/fstab

1.2.3.4/export/dir /dir nfs rw,sync 0 0

soft/hard	hard (по умолчанию) заставляет клиента повторять запрос до тех пор, пока не будет получен ответ. soft прекращает посылки после retrans посылок
retrans=n	число перепосылок запроса серверу
rsize/wsize=n	максимальный размер пакета для операций чтения/записи в байтах
ac/noac	клиент может/нет кешировать атрибуты файлов.
proto=udp/tcp	какой протокол использовать для соединения
intr/nointr	можно/нельзя прерывать файловые операции
acl/noacl	использовать/нет вспомогательный протокол NFSACL
nfsvers=N	использовать версию NFS N
sync/async	отсылать данные серверу до выхода из системного вызова
lock/nolock	разрешить/нет использование дополнительного протокола, позволяющего делать вызов flock для файлов на NFS

NFS

$ exportfs -r

$ exportfs -a = обновить экспорт

$ showmount -a = показать клиентов

$ showmount -e = показать экспорт

Управление: блокировки

- chage -E 1 (-E -1)
- passwd -L
- chmod 0 .ssh/authorized_keys
- pam_listfile
- pam_access

Не работают для ssh. Можно включить и настроить pam, тогда заработают.

Управление: блокировки

pam_access:

/etc/security/access.conf

+/- : кто : откуда

+ : root @sudo : crond :0 tty1 tty2 tty3

- : root : ALL

+ : @sudo : 192.168.

- : @sudo : ALL

Управление: квоты

Дисковые: quota/quotacheck/setquota

setquota -u|-g NAME -a|/dev/sda

repquota …

quotaon/quotaoff

quotacheck - проверить/создать базы

Управление: квоты

Процессор, память, и т. п.: ulimit

/etc/secure/limits.conf

ulimit -l = locked memory

ulimit -s = stack size

ulimit -u = user processes

ulimit -t = cpu time

Управление: ulimits

/etc/security/limits:

```
*               soft    cpu             100
@users          hard    nproc           50
```

Спасибо!

Вопросы?...

serg@parallel.ru