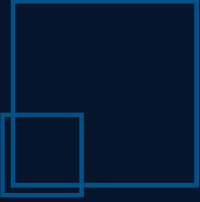


Massive Solutions

Massive Solutions

ОС Clustrx®: организация сервисов

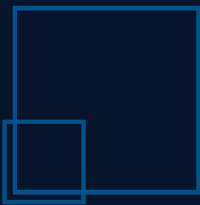
Евгений Остапец, Massive Solutions Ltd.



Загрузка

- аппаратное обеспечение
- системный BIOS
- Clustrx: kernel+initrd
- Clustrx: CBIOS (middleware)
- Clustrx: CB-Manager в контейнере
- Clustrx: сервисы в контейнерах
- прикладное ПО в контейнерах

hardware

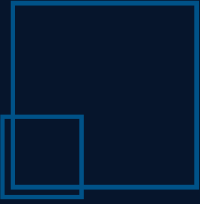


Загрузка

- аппаратное обеспечение
- системный BIOS
- Clustrx: kernel+initrd
- Clustrx: CBIOS (middleware)
- Clustrx: CB-Manager в контейнере
- Clustrx: сервисы в контейнерах
- прикладное ПО в контейнерах

BIOS

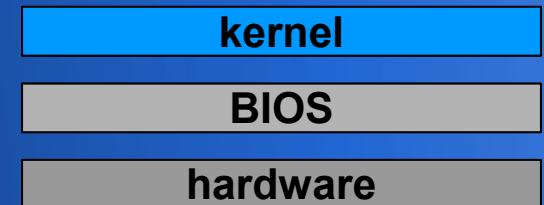
hardware

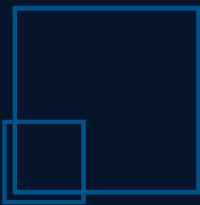


Загрузка

Massive Solutions

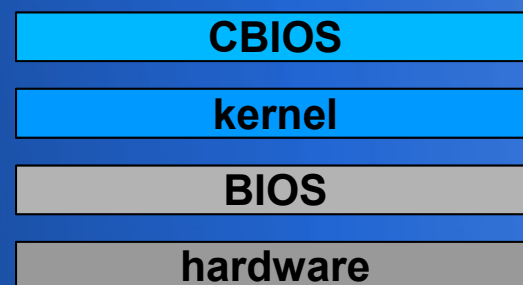
- аппаратное обеспечение
- системный BIOS
- Clustrx: kernel+initrd
- Clustrx: CBIOS (middleware)
- Clustrx: CB-Manager в контейнере
- Clustrx: сервисы в контейнерах
- прикладное ПО в контейнерах

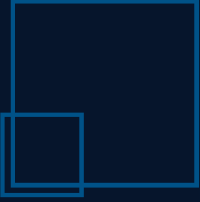




Загрузка

- аппаратное обеспечение
- системный BIOS
- Clustrx: kernel+initrd
- Clustrx: CBIOS (middleware)
- Clustrx: CB-Manager в контейнере
- Clustrx: сервисы в контейнерах
- прикладное ПО в контейнерах

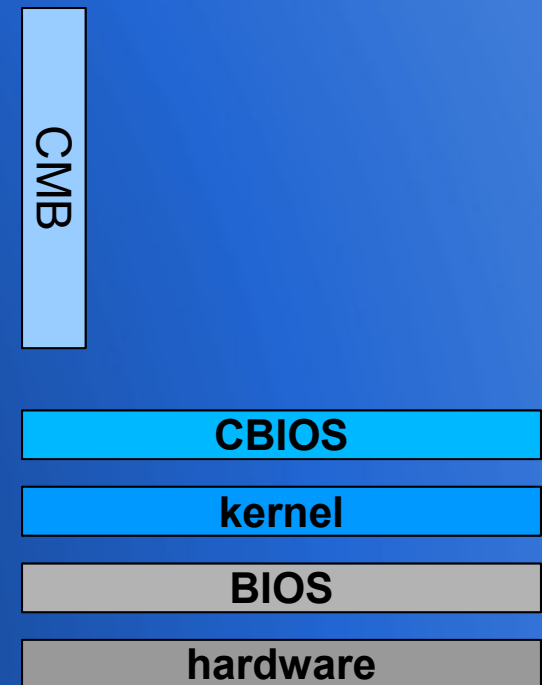


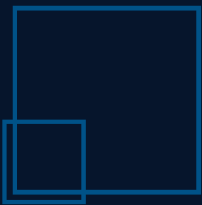


Загрузка

Massive Solutions

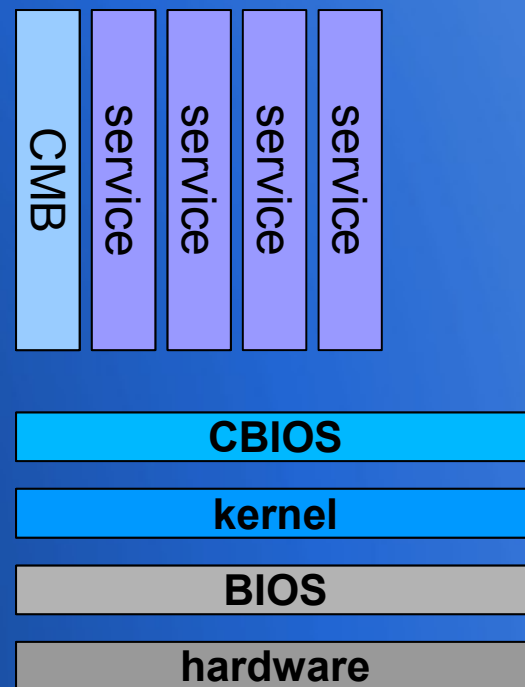
- аппаратное обеспечение
- системный BIOS
- Clustrx: kernel+initrd
- Clustrx: CBIOS (middleware)
- Clustrx: CB-Manager в контейнере
- Clustrx: сервисы в контейнерах
- прикладное ПО в контейнерах

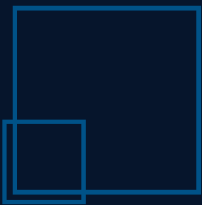




Загрузка

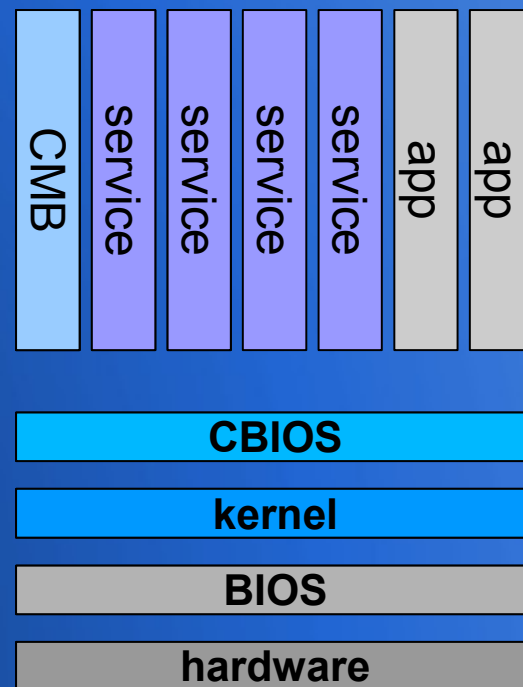
- аппаратное обеспечение
- системный BIOS
- Clustrx: kernel+initrd
- Clustrx: CBIOS (middleware)
- Clustrx: CB-Manager в контейнере
- Clustrx: сервисы в контейнерах
- прикладное ПО в контейнерах

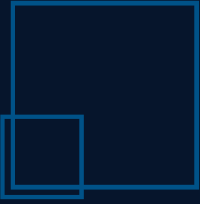




Загрузка

- аппаратное обеспечение
- системный BIOS
- Clustrx: kernel+initrd
- Clustrx: CBIOS (middleware)
- Clustrx: CB-Manager в контейнере
- Clustrx: сервисы в контейнерах
- прикладное ПО в контейнерах

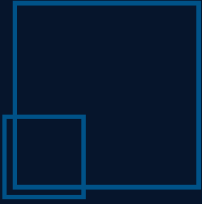




ClusterBIOS

Massive Solutions

- СВІОS – загружаемый образ ОС («прошивка»)
- цель: упрощение/обобщение нижнего уровня
 - гетерогенная среда, realtime-мониторинг
 - гибкое управление
 - свободная масштабируемость
 - состав:
 - kernel+initrd
 - СВ-Executor (+libraries)
 - СВ-Manager
 - сервисы разнесены по контейнерам

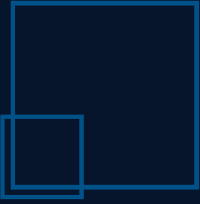


CB-Manager

Massive Solutions

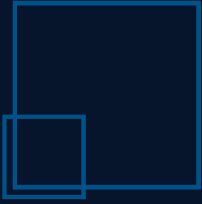
CB-Manager – встроенный гипервизор CBIOS

- функциональность управления узлом
- запуск/останов слотов и сервисов
- список и состояние доступных слотов
- установка/удаление слотов
- работа с образами вычислительных узлов
- конфигурационные интерфейсы
- обновление
- запускается как сервис-в-контейнере



ФС узла

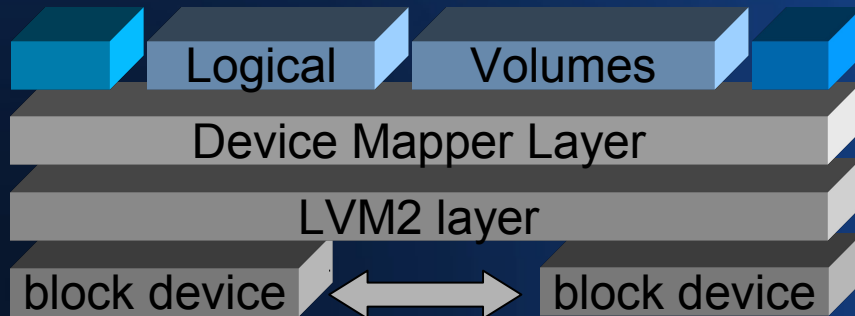
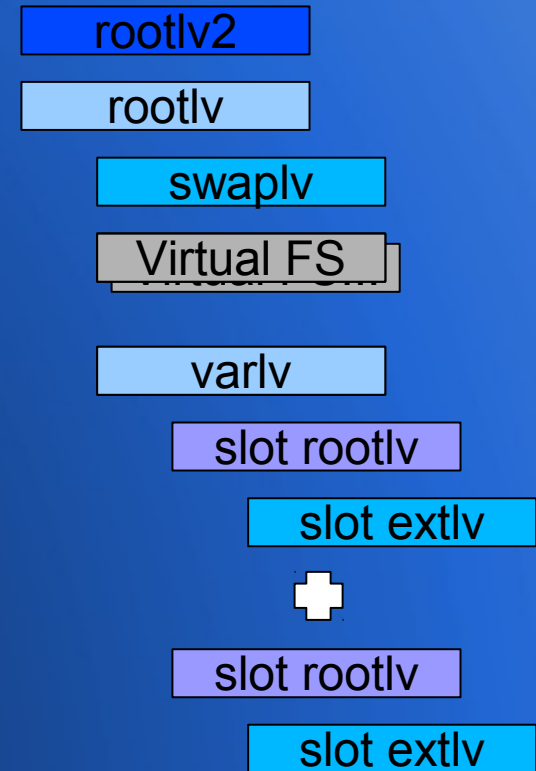
- ФС расположены на логических томах LVM
 - CBIOS: / (основной+запасной), /var, swap
 - отдельные LV (≥ 1) для каждого слота
- обобщенное именование томов для слотов:
/dev/mapper/lvm-slot<eid>-<iid>
 - eid: идентификатор контейнера на узле
 - iid: идентификатор раздела в контейнере

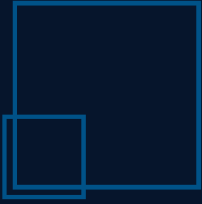


ФС узла

Massive Solutions

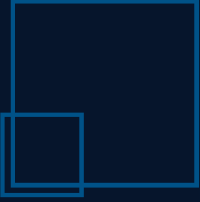
Filesystem	Size	Used	Avail	Use%	Mounted on
/dev/dm-0	2.0G	904M	1010M	48%	/
udevfs	5.0M	200K	4.9M	4%	/dev
shmfs	12G	0	12G	0%	/dev/shm
tmpfs	12G	37M	12G	1%	/tmp
/dev/dm-1	7.9G	328M	7.2G	5%	/var





Слоты

- цель:
 - изоляция сервисов друг от друга и C/BIOS
 - портативность сервисов по узлам и 64/32-bit
- слот может быть реализован как:
 - LXC-сервисный
 - LXC-системный
 - KVM-системный
- слот – это:
 - мера гранулярности сервисов в Clustrx
 - единица установки/обновления/удаления



СЛОТЫ

Massive Solutions

Управление из CLI

slot image list:

Shows available images

slot list:

Show slots list for all of management nodes

slot list <ObjectID>:

Show slot list for selected management node

slot show-runtime <ObjectID>:

Show detailed runtime slot information

slot show:

Show detailed slot list for all of management nodes

slot show <ObjectID>:

Show detailed slot list for selected management node

slot start <ObjectID>:

Start selected slot

slot startall <ObjectID>:

Start all slots at selected server

slot stop <ObjectID>:

Stop selected slot

slot stopall <ObjectID>:

Stop all slots at selected server

slot restart <ObjectID>:

Restarts selected slot

help slot:

Shows detailed information about command "slot"

СЛОТЫ

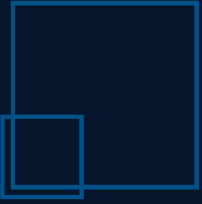
Massive Solutions

```
> slot show 31000040010
```

ParentID	Location	Type	Version	IP	Hostname	Status	Description
31000040010	31000040175	nfs				ON	NFS Slot
31000040010	31000040275	dhcp	0.901.25.28 2010-09-10	10.0.249.52	dhcpd52	ON	Sysmon Slot (DHCP)
31000040010	31000040375	dhcprelay	0.901.15.16 2010-07-14			ON	DHCP Relay Slot
31000040010	31000040475	ldap	0.901.36.27 2010-09-27	10.0.249.6	ldap01	ON	LDAP Slot
31000040010	31000040575	service	0.901.16.15 2010-07-19			ON	Service Slot
31000040010	31000040675	dns_master		10.0.249.1	dns-master	ON	Master DNS Slot
31000040010	31000040875	sysmon		10.0.249.17	sysmon17	ON	Sysmon Slot

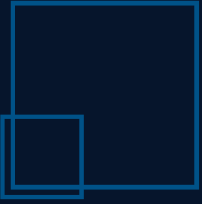
```
> slot show --by-type=sysmon
```

ParentID	Location	Type	Version	IP	Hostname	Status	Description
30100010010	30100010175	sysmon	0.901.15.16 2010-07-16	10.0.249.35	sysmon35	ON	Sysmon Slot
30100010010	30100010275	sysmon	0.901.15.16 2010-07-16	10.0.249.41	sysmon41	ON	Sysmon Slot
30100010010	30100010375	sysmon	0.901.16.20 2010-08-09	10.0.249.61	sysmon61	ON	Sysmon Slot
30100020010	30100020175	sysmon		10.0.249.42	sysmon42	ON	Sysmon Slot
30100020010	30100020275	sysmon	0.901.16.20 2010-08-09	10.0.249.62	sysmon62	ON	Sysmon Slot
30100030010	30100030175	sysmon		10.0.249.43	sysmon43	ON	Sysmon Slot
30100030010	30100030275	sysmon	0.901.16.20 2010-08-09	10.0.249.63	sysmon63	ON	Sysmon Slot
30100030010	30100030575	sysmon	0.901.32.32 2010-09-22	10.0.249.143	sysmon143	ON	Common Sysmon slot
30100040010	30100040175	sysmon	0.901.15.16 2010-07-16	10.0.249.44	sysmon44	ON	Sysmon Slot
30100040010	30100040275	sysmon	0.901.16.20 2010-08-09	10.0.249.64	sysmon64	ON	Sysmon Slot
31000030010	31000030475	sysmon		10.0.249.25	sysmon25	ON	Sysmon Slot
31000040010	31000040875	sysmon		10.0.249.17	sysmon17	ON	Sysmon Slot
41700030010	41700030375	sysmon		10.0.249.14	sysmon14	ON	Sysmon Slot
41700070010	41700070275	sysmon		10.0.249.30	sysmon30	ON	Sysmon Slot
41700070010	41700070375	sysmon		10.0.249.32	sysmon32	ON	Map-Proxy Sysmon Slot
41700070010	41700070575	sysmon	0.901.16.20 2010-08-09	10.0.249.33	sysmon33	ON	Sysmon Slot
41700070010	41700070775	sysmon	0.901.32.32 2010-09-22	10.0.249.34	sysmon34	ON	Common Sysmon slot
41700080010	41700080175	sysmon		10.0.249.28	sysmon28	ON	Sysmon Slot
41700080010	41700080275	sysmon	0.901.15.13 2010-07-07	10.0.249.70	sysmon70	ON	Sysmon Slot (cslog)



Слоты

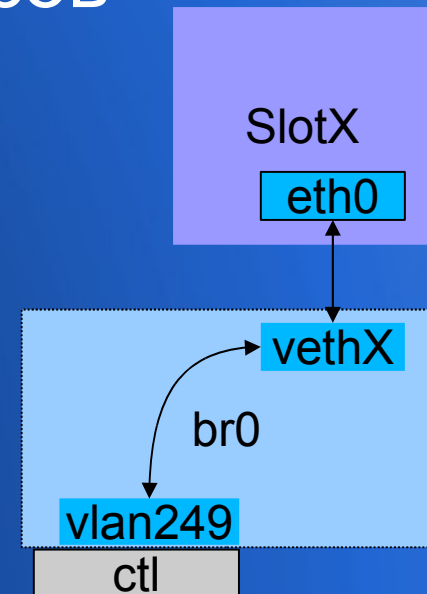
- LXC (Linux Containers):
 - виртуализация операционных окружений
 - собственные пространства pid, uid, ...
 - сеть: собственная или хостовая
 - общее ядро C/BIOS для всех окружений
- KVM (Kernel Virtual Machine):
 - полная виртуализация (Intel VT, AMD V)
 - каждый слот имеет свои ядро, диск, сеть, ...



Сеть узла

Massive Solutions

- несколько физических интерфейсов
 - ctl, svn, ib0...
 - VLAN управления (vlan249)
- один или более слотов
 - vethX на узле
 - eth0 в слоте со своей сетью
- один или более бриджей (brN)



Сеть узла

Massive Solutions

```
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 16436 qdisc noqueue state UNKNOWN
2: ctl: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP qlen 1000
3: svn: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP qlen 1000
6: vlan249@ctl: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state UP
7: br0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state UNKNOWN
15: veth5: <BROADCAST,MULTICAST,PROMISC,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP qlen
18: veth7: <BROADCAST,MULTICAST,PROMISC,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP qlen
24: veth12: <BROADCAST,MULTICAST,PROMISC,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP qlen
36: veth15: <BROADCAST,MULTICAST,PROMISC,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP qlen 1000
```

bridge name	bridge id	STP enabled	interfaces
br0	8000.0002c90450a6	yes	veth12 veth15 veth5 veth7 vlan249

