

LPI-Japan 主催 LPIC 304 技術解説セミナー



LPI-Japanアカデミック認定校
スキルブレイン株式会社 インストラクター
河原木 忠司



- 試験概要
- XenとKVMを利用した仮想化の概要
- ロードバランシング
- クラスタ管理



試験概要



初級

上級

LPICレベル1 LPIC-1

サーバの
構築、運用・保守

実務に必要なLinuxの基本操作とシステム管理が行えるエンジニアであることを証明できます。

LPICレベル2 LPIC-2

ネットワークを含む、
コンピュータシステムの構築、運用・保守

Linuxのシステムデザイン、ネットワーク構築において、企画、導入、維持、トラブルシューティングができるエンジニアであることを証明できます。

LPICレベル3 Core LPIC-3 Core

エンタープライズシステム
構築レベル

大規模システムの構築、運用・保守ができるエンジニアであることを証明できます。

LPICレベル3 Speciality LPIC-3 Speciality

各分野の最高レベルの技術力を持つ
専門家レベル

LPI-302 Mixed Environment:

Linux、Windows、Unixが混在するシステムの設計、構築、運用・保守ができるエキスパートエンジニアであることを証明できます。

LPI-303 Security:

セキュリティレベルの高いコンピュータシステムの設計、構築、運用・保守ができるエキスパートエンジニアであることを証明できます。

LPI-304 Virtualization & High Availability:

クラウドコンピューティングシステム(クラウド)の設計、構築、運用・保守ができるエキスパートエンジニアであることを証明できます。



- 約60問
- 90分
- 出題範囲
 - 仮想化
 - ロードバランシング(負荷分散)
 - クラスタ管理
 - クラスタストレージ



■ Xen

- CentOS 5.を利用
→ Debian 7に変更
- RH系ではver.6系では公式リポジトリから外されている

■ KVM

- Debian 6を利用

■ ロードバランシング

■ クラスタ管理

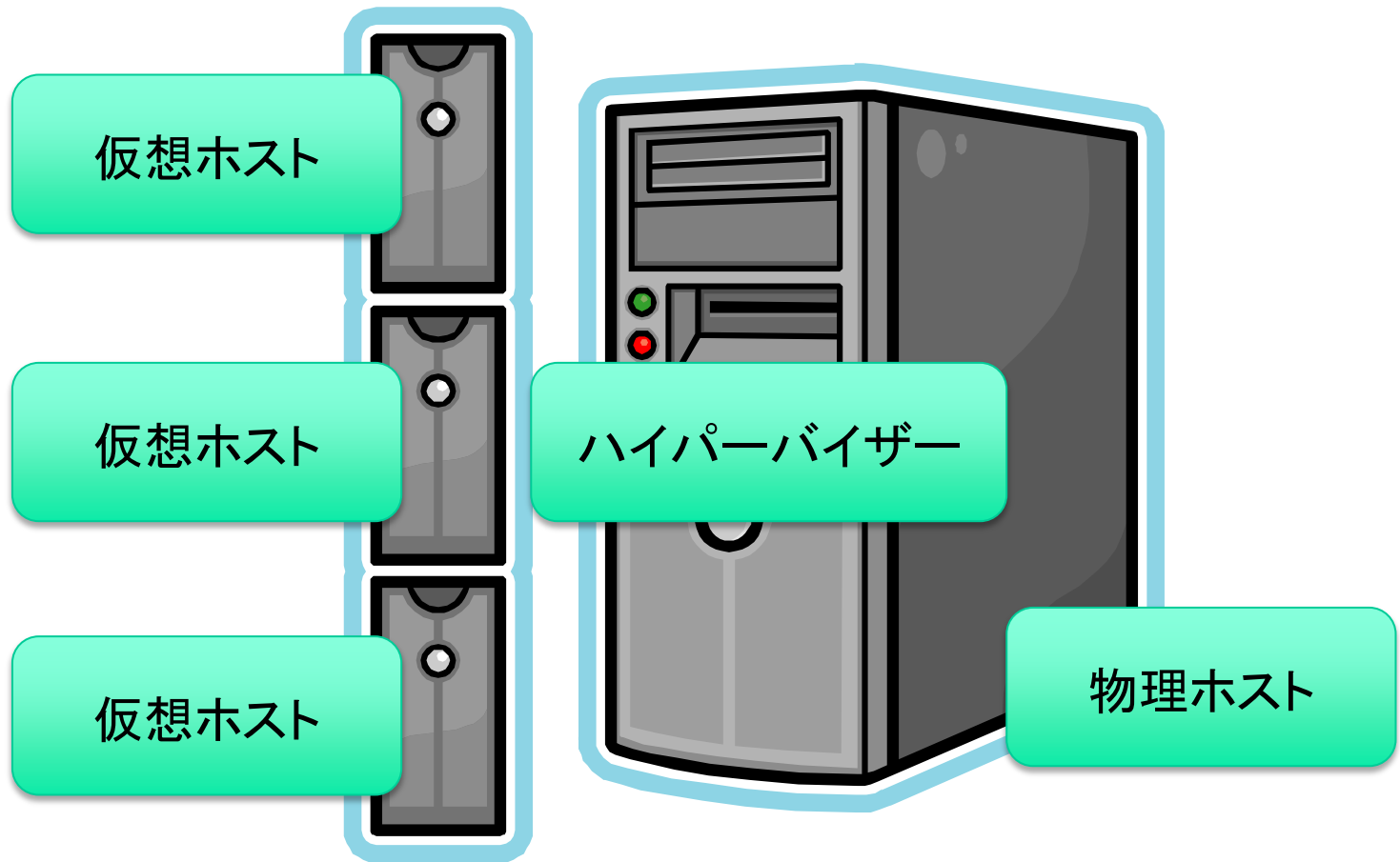
- VMWare Player上のCentOS5.5を利用



XenとKVMを利用した仮想化の概要



- 1台の物理ホスト上に、複数台の仮想ホストを起動し、リソースを利用
- メリット
 - エネルギー消費の効率化
 - システムリソースの有効利用
 - 古いシステムの継続利用





■ 準仮想化

- 仮想化をサポートする物理CPUを要しない
- Xenはこの形態の仮想化を主にサポート
- 環境が制限されるが、仮想化のオーバーヘッドを最小限にできる

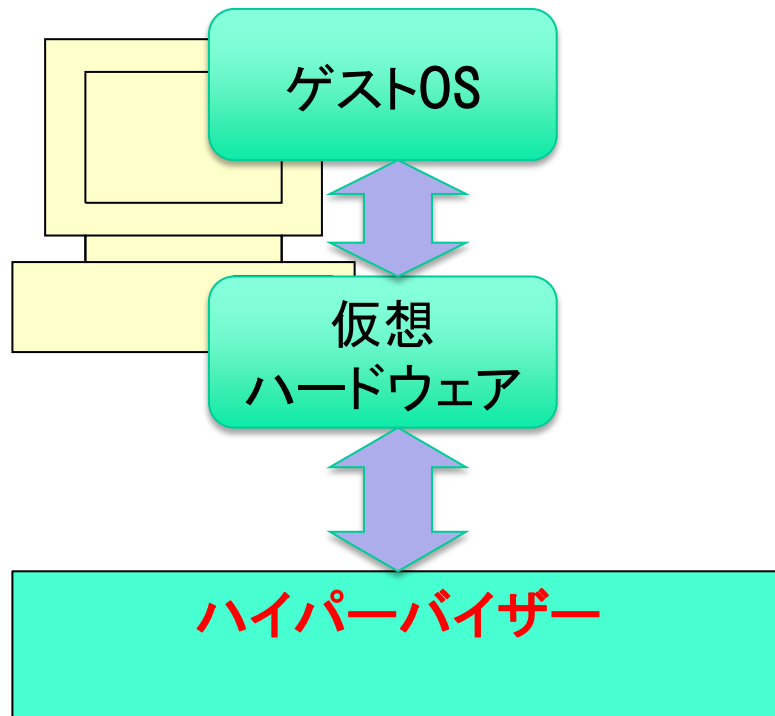
■ 完全仮想化

- 仮想化をサポートする物理CPUを要する
- KVMはこの形態の仮想化をサポート
- ホストOS上で、デバイスの動作をシミュレートするため、仮想化のオーバーヘッドが生じる

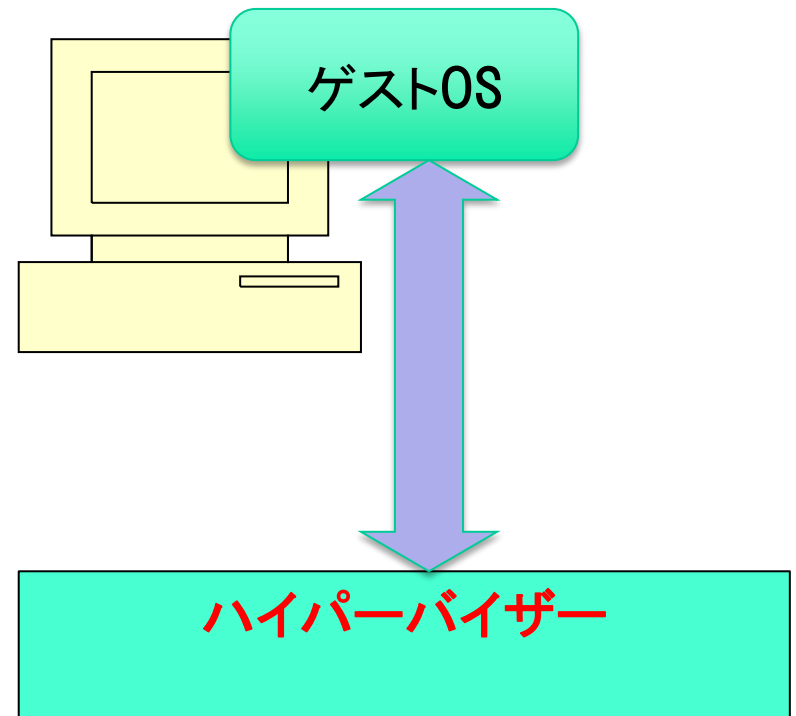


準仮想化と完全仮想化

■ 準仮想化



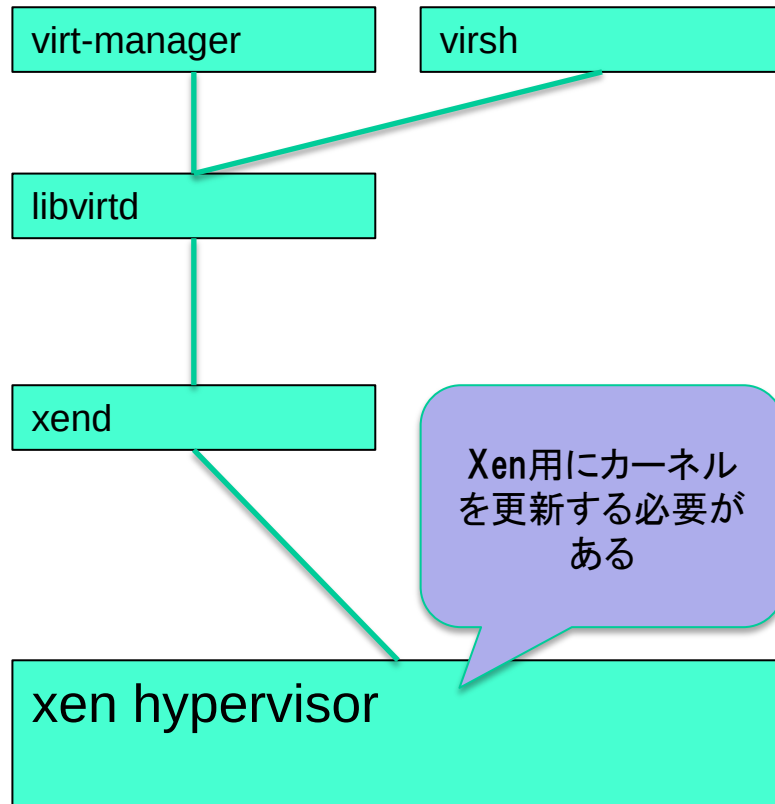
■ 完全仮想化



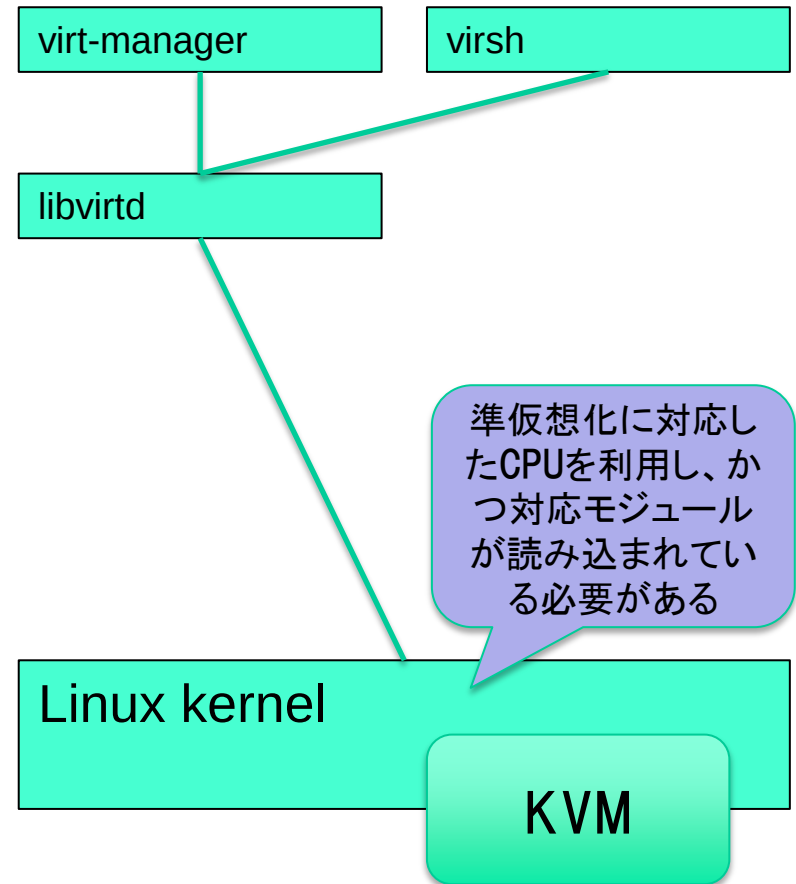


XenとKVMの構成要素

■ Xen



■ KVM





■インストール

```
yum install xen kernel-xen python-virtinst virt-viewer
```

■grubの設定

```
[root@localhost ~]# grep -v '^#' /boot/grub/menu.lst
```

```
default=0
```

```
timeout=5
```

```
splashimage=(hd0,1)/boot/grub/splash.xpm.gz
```

```
hiddenmenu
```

```
title CentOS (2.6.18-348.6.1.el5xen)
```

```
    root (hd0,1)
```

```
    kernel /boot/xen.gz-2.6.18-348.6.1.el5
```

```
    module /boot/vmlinuz-2.6.18-348.6.1.el5xen ro root=LABEL=/ rhgb quiet
```

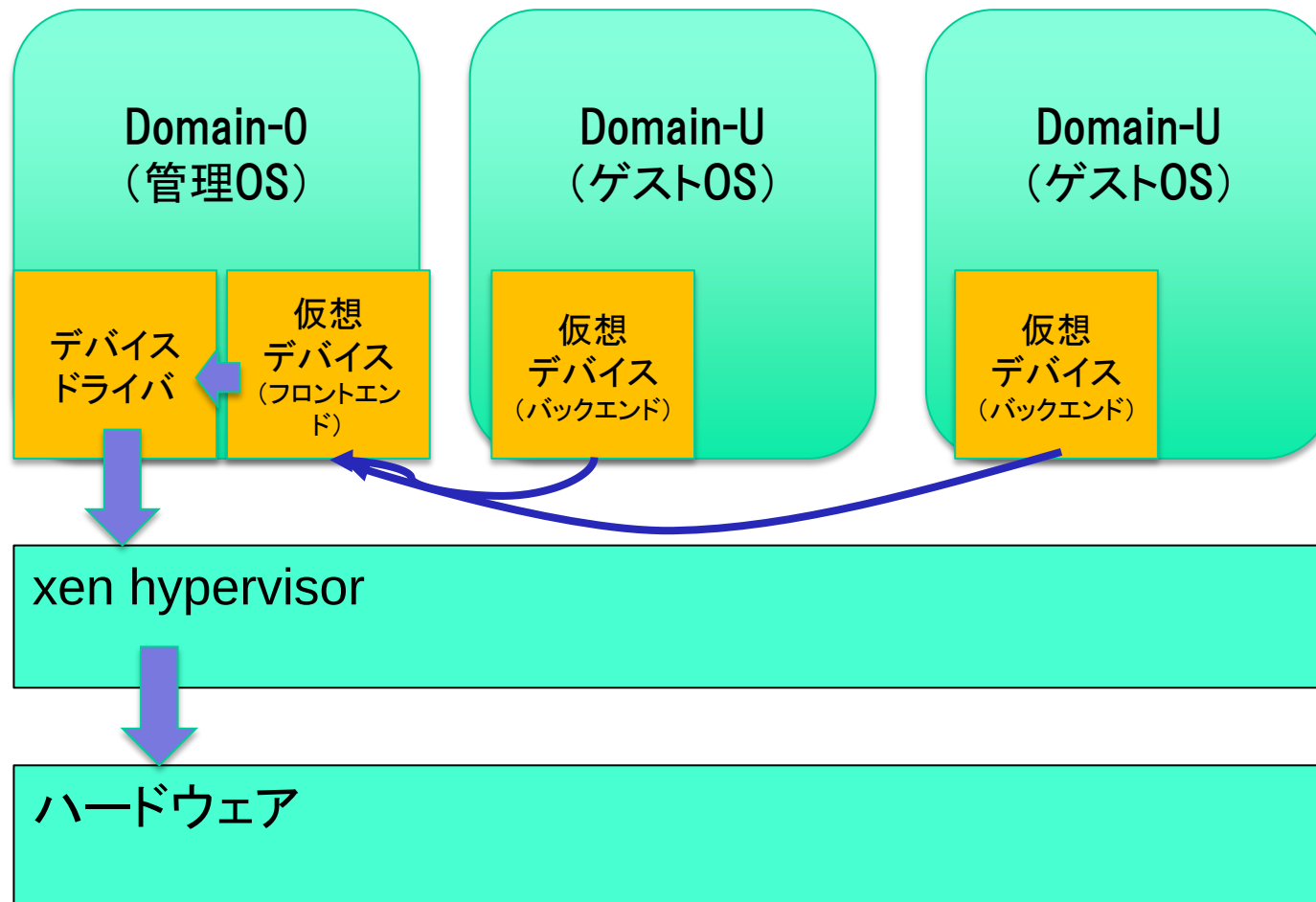
```
    module /boot/initrd-2.6.18-348.6.1.el5xen.img
```

```
title CentOS (2.6.18-238.el5PAE)
```

```
    root (hd0,1)
```

```
    kernel /boot/vmlinuz-2.6.18-238.el5PAE ro root=LABEL=/ rhgb quiet
```

```
    initrd /boot/initrd-2.6.18-238.el5PAE.img
```





```
[root@localhost ~]# virt-install --prompt
```

Would you like a fully virtualized guest (yes or no)? This will allow you to run unmodified operating systems. no

What is the name of your virtual machine? centos5

How much RAM should be allocated (in megabytes)? 512

What would you like to use as the disk (file path)?

/var/lib/xen/images/centos5.img

How large would you like the disk (/var/lib/xen/images/centos5.img) to be (in gigabytes)? 20

What is the install URL?

http://ftp.riken.jp/Linux/centos/5.8/os/i386/



- `xm create -c centos6`
- `virt-viewer centos6`
- `virt-manager`



- Xen環境において、仮想マシンの管理に利用
- 主なサブコマンド

サブコマンド名	用途
create	Domain-U(ゲストOS)の起動
shutdown	Domain-Uのシャットダウン
destroy	Domain-Uの終了(=電源OFF)
console	コンソール接続
list	Domainの一覧表示

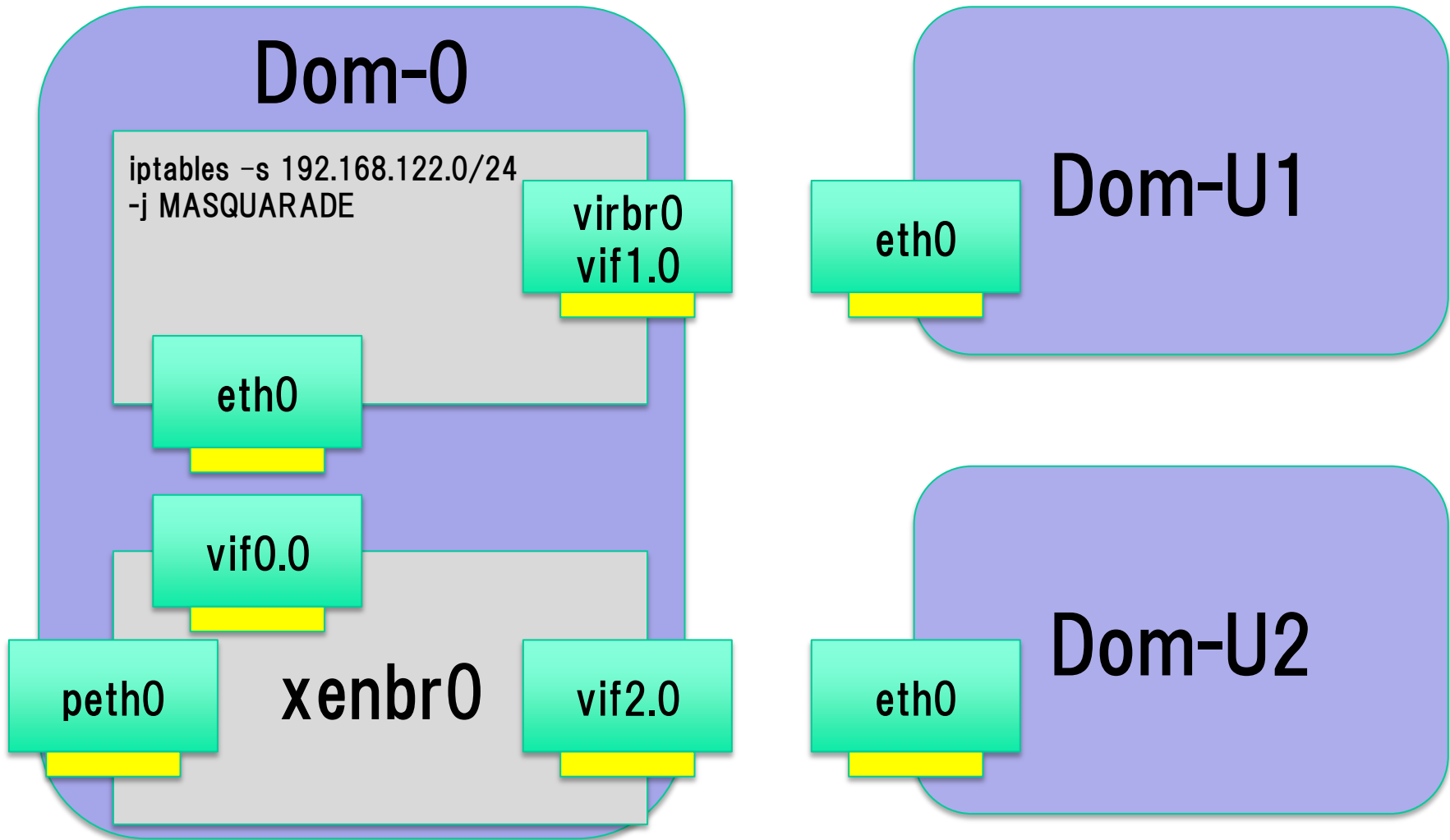


■ /etc/xen/

- 設定ファイル: xend-config.sxp

■ /etc/xen/auto/

- シンボリックリンクを設定すると、仮想マシンを自動起動





```
root@debian:~# lsmod | grep kvm
```

```
kvm_intel                32904  0
```

```
kvm                      185023  1 kvm_intel
```

```
root@debian:~# grep vmx /proc/cpuinfo
```

```
flags      : fpu vme de pse tsc msr pae mce cx8 apic sep mtrr pge  
mca cmov pat pse36 clflush dts acpi mmx fxsr sse sse2 ss ht tm pbe  
nx lm constant_tsc arch_perfmon pebs bts aperfmperf pni dtes64  
monitor ds_cpl vmx smx est tm2 ssse3 cx16 xtpr pdcm sse4_1 xsave  
lahf_lm ida tpr_shadow vnmi flexpriority
```

```
flags      : fpu vme de pse tsc msr pae mce cx8 apic sep mtrr pge  
mca cmov pat pse36 clflush dts acpi mmx fxsr sse sse2 ss ht tm pbe  
nx lm constant_tsc arch_perfmon pebs bts aperfmperf pni dtes64  
monitor ds_cpl vmx smx est tm2 ssse3 cx16 xtpr pdcm sse4_1 xsave  
lahf_lm ida tpr_shadow vnmi flexpriority
```

amd環境の場合、**kvm_amd**



```
root@debian:~# aptitude install kvm bridge-utils
```

以下の新規パッケージがインストールされます:

```
bridge-utils kvm libaio1{a} libvdeplug2{a} qemu-kvm
```

更新: 0 個、新規インストール: 5 個、削除: 0 個、保留: 0 個。

1,564 kB のアーカイブを取得する必要があります。展開後に 4,215 kB のディスク領域 が新たに消費されます。

先に進みますか? [Y/n/?] y



■ ディスクイメージの作成

```
root@debian:~# kvm-img create -f qcow2 centos.img 20GB
```

```
Formatting 'centos.img', fmt=qcow2 size=21474836480 encryption=off  
cluster_size=0
```

```
root@debian:~# ls -l *.img
```

```
-rw-r--r-- 1 root root 262144 2013-06-18 00:48 centos.img
```

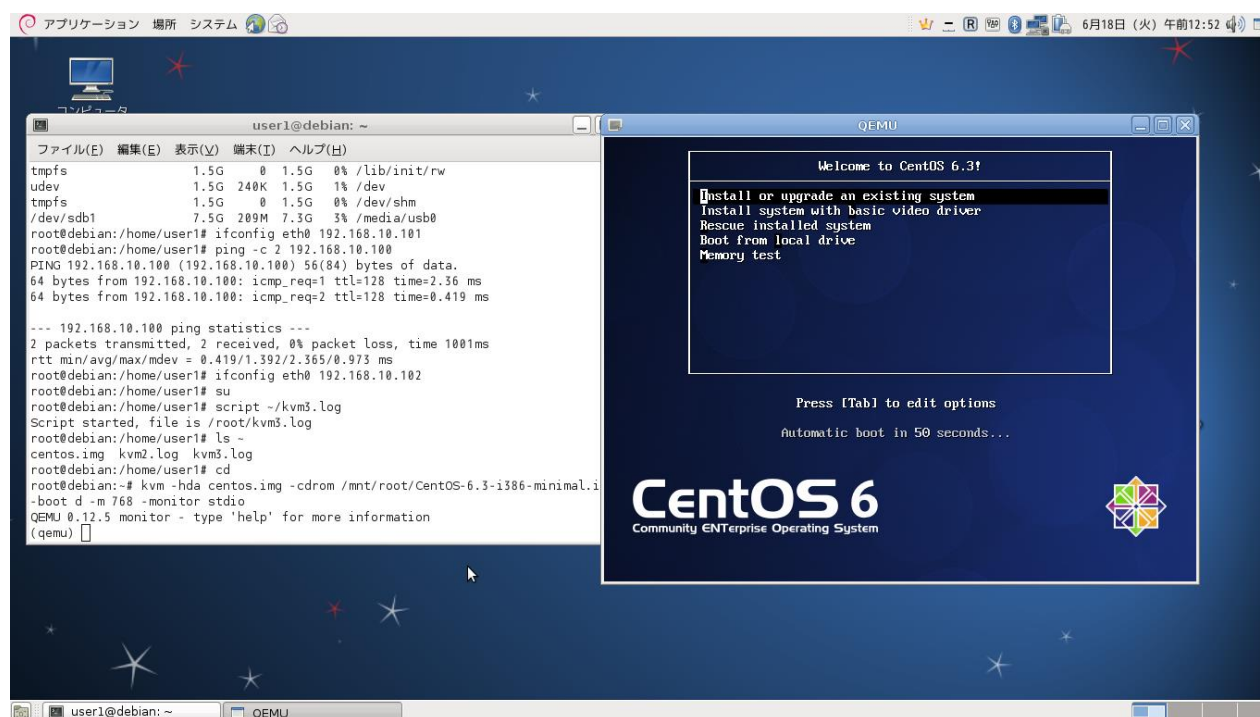
※ディストリビューションによっては、kvm-imgコマンドではなく、qemu-img
コマンドを利用



■ isoファイルを読み込んでブートし、インストール

```
root@debian:~# kvm -hda centos.img -cdrom CentOS-6.3-i386-  
minimal.iso -boot d -m 768 -monitor stdio
```

※ディストリビューションによっては、kvmコマンドではなく、qemu-kvmコマンドを利用





```
root@debian:~# aptitude install virtinst libvirt0 python-libvirt libvirt-bin
```

以下の新規パッケージがインストールされます:

```
etherboot-qemu{a} libgtk-vnc-1.0-0{a} libreadline5{a} libvde0{a} libvirt0  
libxenstore3.0{a} lvm2{a} openbios-ppc{a} openbios-sparc{a}  
openhackware{a} python-libvirt python-pycurl{a} python-urlgrabber{a}  
qemu{a} qemu-keymaps{a} qemu-system{a} qemu-user{a} qemu-utils{a}  
seabios{a} vde2{a} vgabios{a} virt-viewer{a} virtinst
```

更新: 0 個、新規インストール: 23 個、削除: 0 個、保留: 0 個。

20.5 MB のアーカイブを取得する必要があります。展開後に 57.5 MB のディスク領域が 新たに消費されます。

先に進みますか? [Y/n/?] y



```
root@debian:~# virt-install --prompt
```

仮想マシンの名前は何ですか? centos

どれだけの RAM を割り当てますか (メガバイト単位で)? 768

What would you like to use as the disk (file path)?

/imgfiles/centos.img

どの位の大きさのディスク (/imgfiles/centos.img) にしたいですか(ギガバイト単位で)? 20

What is the install CD-ROM/ISO or URL? /CentOS-6.3-i386-minimal.iso



■ユーザー権限

- virt-installを利用した場合、libvirt-qemuの権限でファイルが配置

■defaultネットワークの起動

- virsh net-start default



■ virsh

- 仮想マシンを制御
- libvirtにより提供

■ 主なサブコマンド

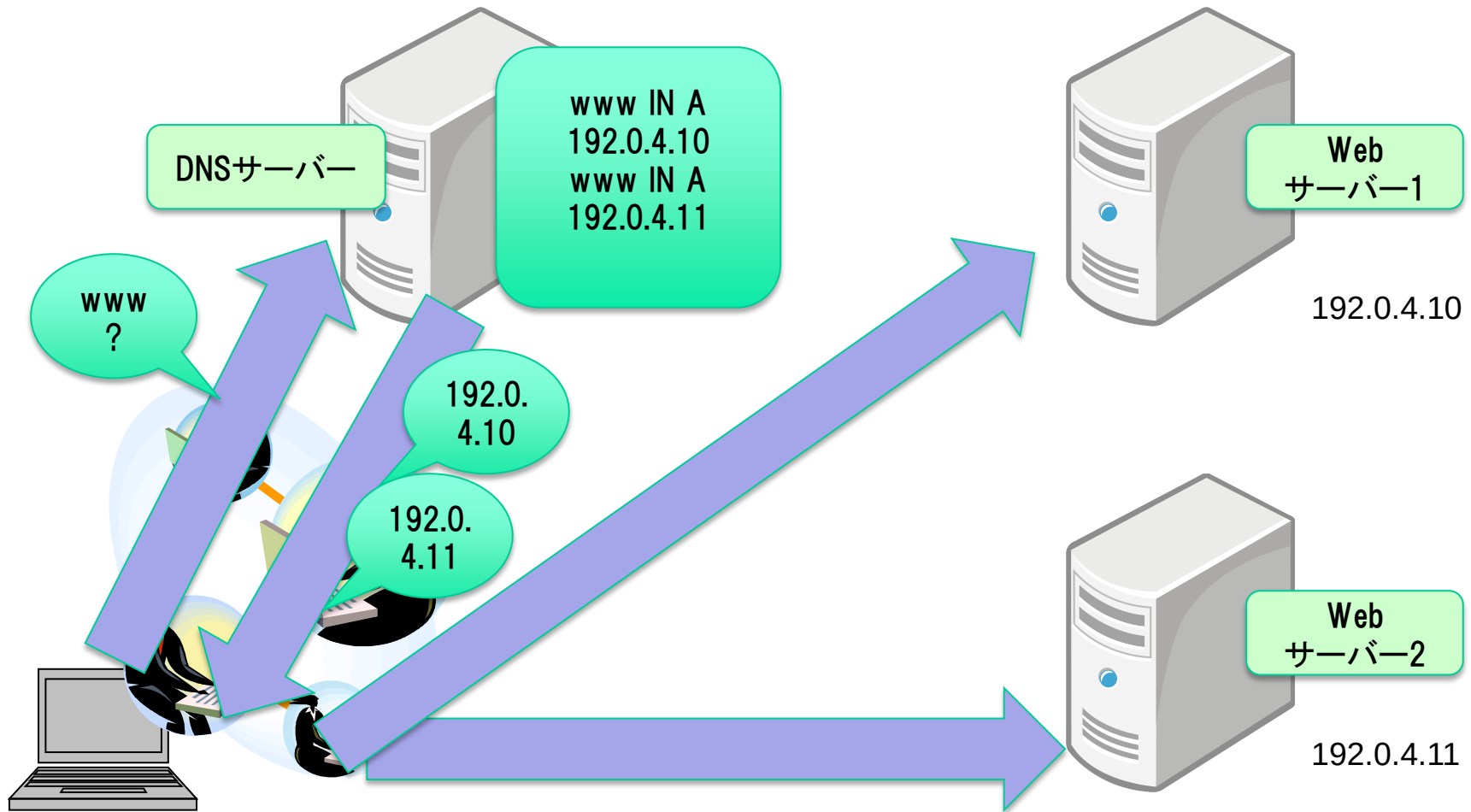
サブコマンド名	用途
start	Domain-U(ゲストOS)の起動
shutdown	Domain-Uのシャットダウン
destroy	Domain-Uの終了(=電源OFF)
console	コンソール接続
list	Domainの一覧表示
save	Domain-Uの状態の保存
setvcpus	仮想CPUの割り当て 例: <code>virsh setvcpus centos6 2</code>

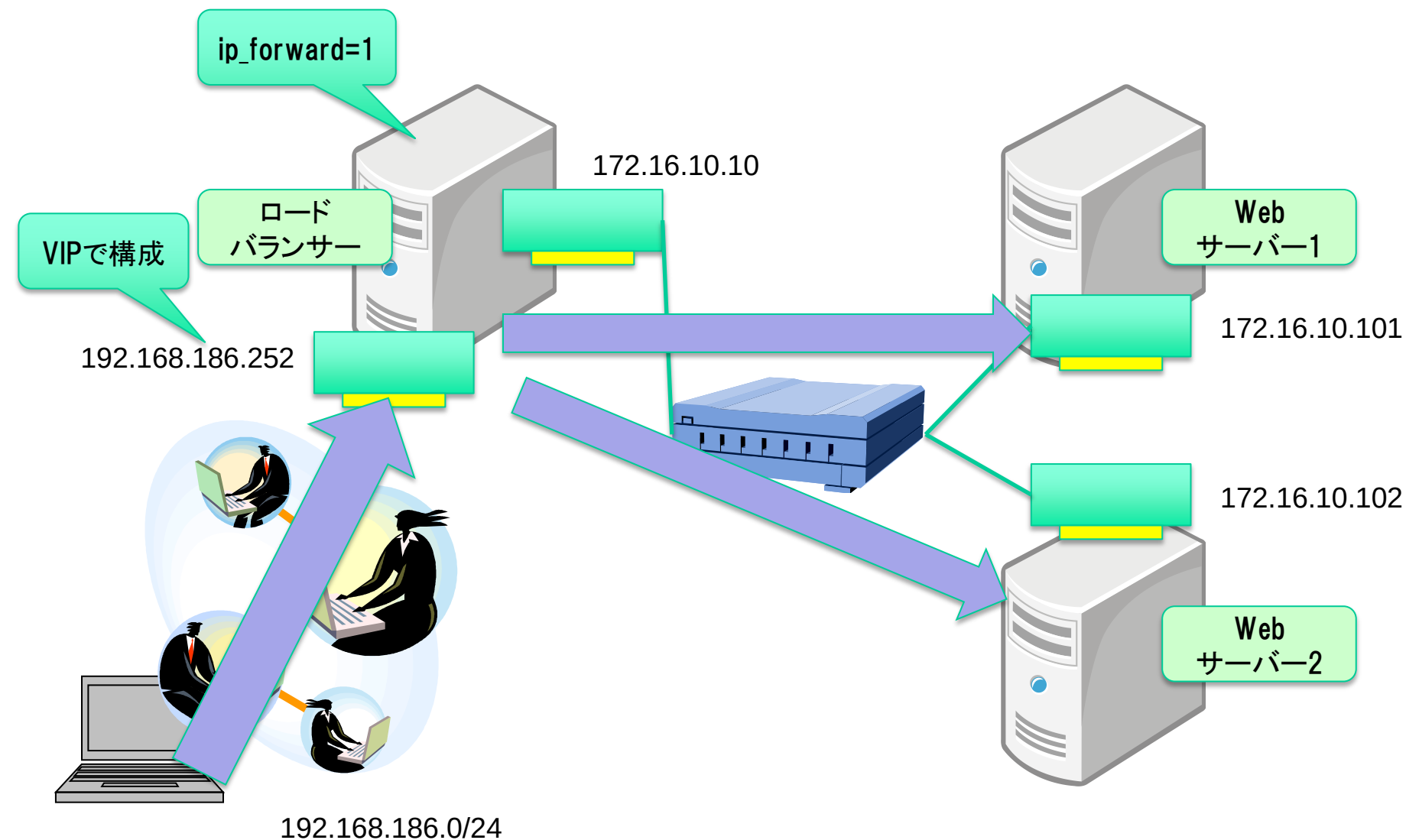


ロードバランシング



- 外部ネットワークからの接続を複数サーバーに転送し、負荷分散をする仕組み。
- ロードバランシングの手法
 - ラウンドロビン型
 - ランダム型
 - 送信元IPアドレス型
 - サーバー負荷連動型
 - 重み付け型







- Linuxシステムの負荷分散ソリューション
- ipvsadm: LVSの管理インターフェース



■オプション

- -A : 設定の追加
- -E : 設定の変更
- -D : 設定の削除

■プロトコル

- -t : TCP
- -u : UDP

■負荷分散の手法

- -s rr : ラウンドロビン
- -s wrr : ラウンドロビン + 重み付け
- -s lc : 負荷の少ないサーバーへ転送
- -s wlc : 負荷の少ないサーバーへ転送 + 重み付け



■オプション

- -a : 設定の追加
- -e : 設定の変更
- -d : 設定の削除

■プロトコル

- -t : TCP
- -u : UDP

■サーバーの指定: -r

■転送オプション

- -g : パケットに何もせず転送
- -m : マスカレードして転送



```
[root@lv ~]# ipvsadm -A -t 192.168.186.252:80 -s lc
[root@lv ~]# ipvsadm -a -t 192.168.186.252:80 -r 172.16.10.101 -m
[root@lv ~]# ipvsadm -a -t 192.168.186.252:80 -r 172.16.10.102 -m
[root@lv ~]# ipvsadm -ln
```

IP Virtual Server version 1.2.1 (size=4096)

Prot LocalAddress:Port Scheduler Flags

-> RemoteAddress:Port	Forward	Weight	ActiveConn	InActConn
-----------------------	---------	--------	------------	-----------

TCP 192.168.186.252:80 lc

-> 172.16.10.102:80	Masq	1	0	0
-> 172.16.10.101:80	Masq	1	0	0



```
[root@lv ~]# /etc/init.d/ipvsadm status
```

```
IP Virtual Server version 1.2.1 (size=4096)
```

```
Prot LocalAddress:Port Scheduler Flags
```

```
-> RemoteAddress:Port      Forward Weight ActiveConn  
InActConn
```

```
TCP 192.168.186.252:80 lc
```

```
-> 172.16.10.102:80      Masq    1      0      0
```

```
-> 172.16.10.101:80      Masq    1      0      0
```

```
[root@lv ~]# /etc/init.d/ipvsadm save
```

```
ipvsadm: Saving IPVS table to /etc/sysconfig/ipvsadm:      [ OK ]
```

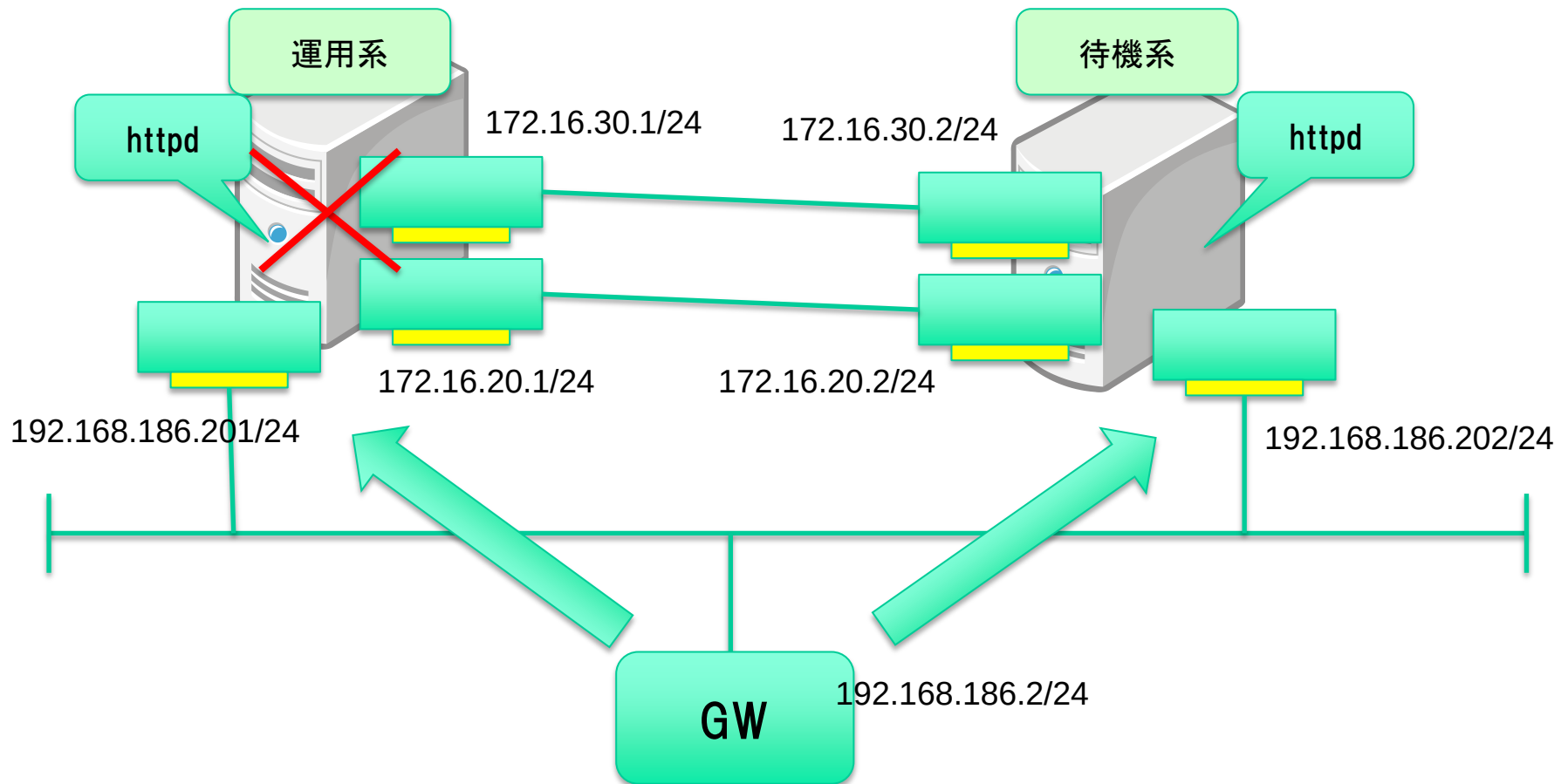
```
[root@lv ~]# chkconfig ipvsadm on
```



- システムダウン時を想定して、heartbeatにより、サーバーを監視



クラスタ管理





■ <http://linux-ha.sourceforge.jp>

■ リポジトリ情報とパッケージも配布

SourceForge.JP > ソフトウェアを探す > Linux-HA Japan > 概要

Linux-HA Japan

概要 ▾ ダウンロード ▾ ソースコード ▾ チケット ▾ 文書 ▾ コミュニケーション ▾ ニュース ▾

オープンソースのDAW「Ardour 3.2」リリース、動画の音声トラック編集をサポート [SF.JP Magazine]

プロジェクトの説明

Web



本ページはLinux-HA Japan 開発者向けサイトです。

プロジェクトのメインサイトはこちらです <http://linux-ha.sourceforge.jp/>

成果物のダウンロードはこちらです <http://sourceforge.jp/projects/linux-ha/releases/>

Linux-HA Japanプロジェクトは、Linux上で高可用クラスタシステムを構築するための部品として、オープンソースの、クラスタリソースマネージャ、クラスタ通信レイヤ、ブロックデバイス複製、その他、さまざまなアプリケーションに対応するための数多くのリソースエージェント、などを、日本国内向けに維持管理、支援等を行っているプロジェクトです。

主な製品として、Pacemaker、Heartbeat、Corosync、DRBD等を取り扱っています。

※現在、Linux-HA Japanの独自開発機能は個別にリリースしておりません。独自開発機能が必要な場合は、Pacemakerリポジトリに同梱されているRPMをご利用ください。

バグを

文書を

フォー

RSSを

インストール



■ 構成するコンポーネント

- pacemaker : フェールオーバー
- heartbeat : 監視
- pm_extra : pacemaker用拡張リソースエージェント・プラグイン

■ インストール手順

```
[root@pm2 ~]# tar xzvf pacemaker-1.0.13-1.1.el5.i386.repo.tar.gz
[root@pm2 ~]# cd pacemaker-1.0.13-1.1.el5.i386.repo
[root@pm2 pacemaker-1.0.13-1.1.el5.i386.repo]# yum -c
pacemaker.repo install pacemaker heartbeat pm_extras
```



```
[root@pm1 ~]# cat /etc/ha.d/ha.cf
```

```
pacemaker on
```

```
debug 0
```

```
udpport 694
```

```
keepalive 2
```

```
warntime 7
```

```
deadtime 10
```

```
initdead 10
```

```
logfacility local1
```

```
bcast eth1
```

```
bcast eth2
```

```
node pm1.example.net
```

```
node pm2.example.net
```

```
watchdog /dev/watchdog
```

```
respawn root /usr/lib/heartbeat/ifcheckd
```



```
[root@pm1 ~]# cat /etc/ha.d/authkeys
```

```
auth 1
```

```
1 sha1 secret
```

```
[root@pm1 ~]# chmod 0600 /etc/ha.d/authkeys
```



```
[root@pm1 ~]# vim /etc/syslog.conf
[root@pm1 ~]# grep local1 /etc/syslog.conf
*.info;mail.none;authpriv.none;cron.none;local1.none
/var/log/messages
local1.info                                /var/log/ha.log
[root@pm1 ~]# /etc/init.d/syslog restart
```



```
[root@pm1 ~]# /etc/init.d/heartbeat start
Starting High-Availability services:          [ OK ]
[root@pm1 ~]# crm_mon -A
Connection to the CIB terminated
Reconnecting...
```

=====

Last updated: Thu Jun 20 00:05:52 2013

Stack: Heartbeat

Current DC: pm1.example.net (cd0b09c4-5082-429a-8190-750a032f9988) - partition with quorum

Version: 1.0.13-30bb726

2 Nodes configured, unknown expected votes

0 Resources configured.

=====

Online: [pm1.example.net pm2.example.net]

Node Attributes:

* Node pm1.example.net:

+ pm2.example.net-eth1 : up

+ pm2.example.net-eth2 : up

* Node pm2.example.net:

+ pm1.example.net-eth1 : up

+ pm1.example.net-eth2 : up

**crm : pacemakerでリソースを
制御**



■ `[root@pm1 ~]# crm configure property no-quorum-policy="ignore"`
`stonith-enabled="false"`

- `no-quorum-policy`
3台以上の構成の場合に利用。
- `stonith-enabled`
STONITH機能を有効にするかどうか。

■ STONITH

制御不能になったサーバー(=スプリットブレイン)の電源を強制的に落とす



■ [root@pm1 ~]# crm configure rsc_defaults resource-stickiness="INFINITY" migration-threshold="1"

- resource-stickiness
自動フェイルバックの設定
- migration-threshold
故障回数指定



- `[root@pm1 ~]# crm configure primitive httpd ocf:heartbeat:apache
params configfile="/etc/httpd/conf/httpd.conf"
statusurl="http://localhost/" testregex="service" op monitor
interval="10s"`
- `[root@pm1 ~]# crm configure primitive vip ocf:heartbeat:IPaddr2
params ip="192.168.186.100" nic="eth0" cidr_netmask="24" op
monitor interval="10s"`
- `[root@pm1 ~]# crm configure group web vip httpd`



```
[root@pm1 ~]# crm_mon
```

```
=====
```

```
Last updated: Thu Jun 20 01:29:27 2013
```

```
Stack: Heartbeat
```

```
Current DC: pm2.example.net (9656fc25-eadc-415d-b84b-18523d75d16b) - partition w  
ith quorum
```

```
Version: 1.0.13-30bb726
```

```
2 Nodes configured, unknown expected votes
```

```
2 Resources configured.
```

```
=====
```

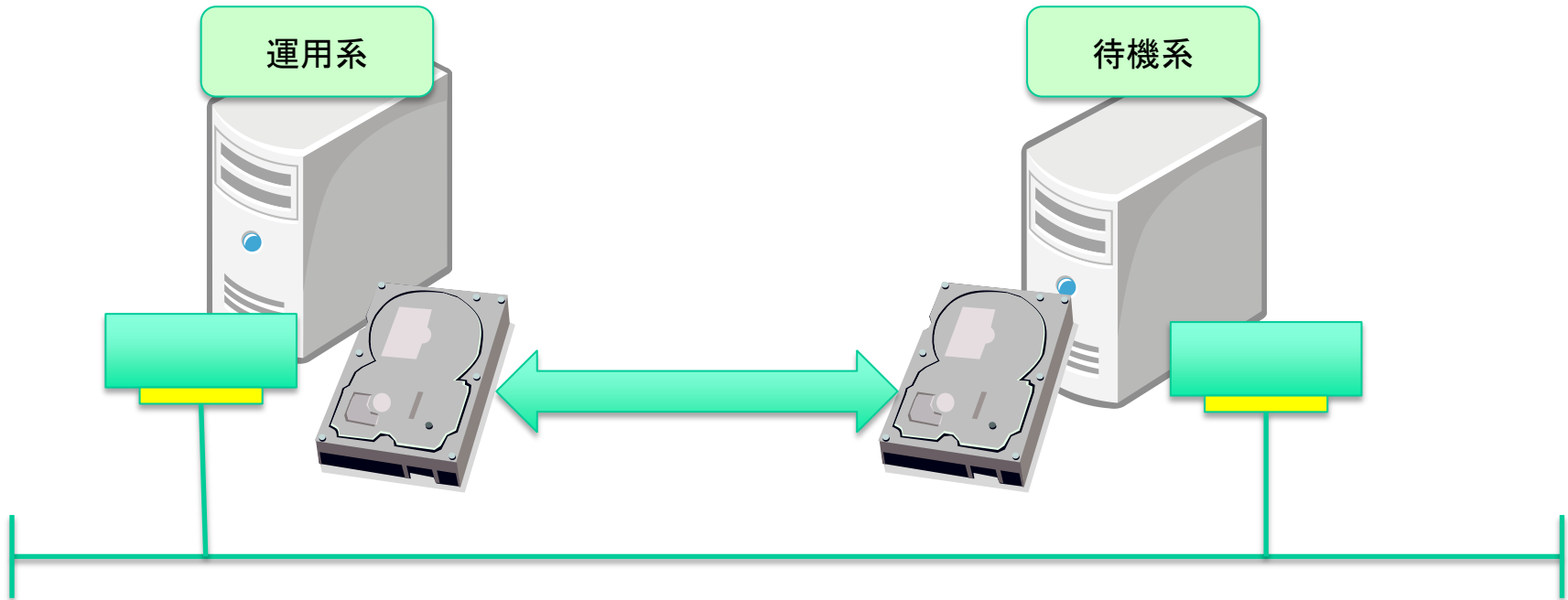
```
Online: [ pm1.example.net pm2.example.net ]
```

```
httpd (ocf::heartbeat:apache): Started pm2.example.net
```

```
vip (ocf::heartbeat:IPaddr2): Started pm2.example.net
```



- 同じコンテンツを提供できるようにdrdbでクラスタストレージを構成





質疑応答についてはお気軽にお声掛けください。

ご清聴ありがとうございました。

<http://www.skillbrain.co.jp>



info@skillbrain.co.jp