Linux KVM による仮想化環境の構築

○早川正人^{A)}、若松 進^{A)}、大下 弘^{A)}、野崎公隆^{A)}、

千代谷一幸^{A)}、雨宮尚範^{A)}、松岡 孝^{B)}
 ^{A)} 工学系技術支援室 情報通信技術系
 ^{B)} 共通基盤技術支援室 情報通信技術系

概要

仮想化技術の1つである KVM は「Kernel-based Virtual Machine」の略であり、Linux カーネルに仮想マシン の管理機能を統合したものである。完全仮想化(フル・バーチャライゼーション)による仮想マシン環境を 提供し、動作中の仮想マシンを別のホストに無停止で移動させる「ライブマイグレーション」もサポートさ れていて、仮想マシンが提供中のサービスを阻害することなく、物理マシンの保守作業等が可能になる。

今回は工学研究科技術部研修において、KVM ベースの仮想化サーバの構築を行い、ライブマイグレーションの動作確認を実行したので報告する。

1 KVM について

一般的に OS の仮想化は、ほかの OS の上で仮想マシンを動かすもの(ホスト OS 型)と、ハードウェアの 上で直接仮想マシンを動かすもの(ハイパーバイザ型)の大きく2種類に分かれる。KVM は、ハードウェア のエミュレーションやゲスト OS の管理用のフロントエンドとして「QEMU」を使い、Linux の上でゲスト OS を動かすので、一見前者のホスト OS 型に見えるが、Linux 自体をハイパーバイザにしてしまうことを考 慮すると、後者のハイパーバイザ型といえる。

KVM は Linux Kernel にマージされ、kernel 2.6.20 から標準機能として利用でき、Red Hat Enterprise Linux 6 では仮想化技術の中心に据えている。なお、CPU の仮想化支援技術を利用しているので, Inte VT や AMD-V といった仮想化支援機能に対応した CPU と,それらの機能を有効にできる BIOS が必要となる。



図 1. KVM のアーキテクチャ

通常の	通常の プロセス	通常の プロセス					
プロセス	CentOS 5.6	CentOS 5.6					
KVM Scientific Linux 6.1							
	ハードウェ	ア					

図2.構築した仮想化サーバの仕様

2 実験環境の構築

仮想化サーバ構築用には、表1に示すスペックのPCを2台用意し、NASにはThecus N7700 Proを使用した。また、.サーバ用のPCには、オンボードNIC(eth1)の他にIntel NIC(eth0)を追加、ネットワークアド レスとして10.10/16(NET2)、192.168.1/24(NET1)を想定し、ネットワークデバイスの設定を行った。 表 1. 仮想化サーバ用 PC の部品構成

PC パーツ	詳細	
CPU	AMD PhenomII (4Core, 2.5GHz)	
Memory	Transcend (PC3-10600, 8GB) $\times 2$	
HDD	Seagate (7200rpm, 1TB)	
Motherboard	Biostar (Graphics, GbE, PCI-E)	



図3.実験環境の構成

2.1 ホスト OS のインストール

ホスト OS となる Linux ディストリビューションは、Red Hat Enterprise Linux をベースとして高い互換性を 持ち、無償で配布されている Scientific Linux (サイエンティフィック・リナックス)を選択し、Scientific Linux の公式ページ (http://www.scientificlinux.org/)等から Scientific Linux 6.1 の 64bit 版の ISO イメージをダウンロ ードした。ISO イメージからインストール用ディスク (DVD2 枚)を作成した後、DVD から起動して OS の インストール作業を行った。なお、インストール時の各サーバの設定内容を表 2 に示す。

表 2. 仮想化サーバの設定内容

ホスト名		SV1	SV2		
IDアドレフ	NET1(eth0)	192.168.1.11	192.168.1.12		
	NET2(eth1)	10.10.1.11	10.10.1.12		
		/ (ルート):100GB			
パーティション	/構成	swap: 16GB			
		/guest:残り全部(最大容量まで使用)			
インストール		Desktop を選択し、ソフ	トウェアのカスタマイズ		
パッケージの調	選択	で仮想化関連のパッケー	ージを選択する		

2.2 ネットワークデバイスの設定

仮想化サーバのネットワークデバイスとして Intel NIC を eth0、オンボード NIC を eth1 に割り当て、IP ア ドレスを設定し、ホスト名でコンピューターを同定できるように「/etc/hosts」ファイルに記載した。また、 ゲストマシンをサーバとして利用するためには、ネットワークブリッジでゲストマシンと外部コンピュータ ーの相互アクセスを可能にする必要があるため、ブリッジインターフェース br0 を作成し、eth0 をブリッジ に参加させた。設定はネットワークスクリプト ifcfg-br0、ifcfg-eth0、ifcfg-eth1 を編集することで行った。

/etc/hosts への追記
 iscsi1 10.10.1.10

sv1 192.168.1.11 sv2 192.168.1.12 sv10 10.10.1.11 sv20 10.10.1.12 /etc/sysconfig/network-scripts/ifcfg-br0の作成 DEVICE=br0 ONBOOT=yes BOOTPROTO=none IPADDR=192.168.1.11 (仮想化サーバ2は192.168.1.12) NETMASK=255.255.255.0 GATEWAY=192.168.1.1 TYPE=Bridge /etc/sysconfig/network-scripts/ifcfg-eth0 への追記 BRIDGE=br0 /etc/sysconfig/network-scripts/ifcfg-eth1 への追記(同名の設定項目は置き換えた) . BOOTPROTO=none IPADDR=10.10.1.11 (仮想化サーバ2は10.10.1.12) NETMASK=255.255.255.0 GATEWAY=192.168.1.1

2.3 NAS の設定

設定は NAS をネットワークに接続した後、NAS の機能である Web 管理インターフェース(図 4)を用いて行い、iSCSI ターゲットとして 5TB の領域を作成した。なお、設定項目の詳細を表 3 に示す。

	-								
9	\$152/11/1	0	-						
	725 RAD	1D	RAD L-KIL	774452	10.00	使用 ディスク	NUM	データ 日間	15C5/ 有限
		RAD	6.	EXT3	Healthy	1234567	9305.7 GI	0.9 GB/4508.6 GB	4852.4+
AID	1-7-	5.850		LORAL ST					

表 3. NAS の設定詳細

	設定項目	設定値
Notwork 1	IP	10.10.1.10
INCLWOIK I	Netmask	255.255.0.0
	RAID Level	RAID 6
RAID	RAID ID	RAID
	Data Percentage	50%
	Allocation	50%
iSCSI Volume	iSCSI Target Volume	Enable
	Target Name	iscsi1

図 4. Web 管理インターフェース

2.4 iSCSI デバイスのマウント

iSCSI デバイスを利用できる様にするために下記の設定を行った。

1) iSCSI イニシエータ パッケージをインストールする

yum install iscsi-initiator-utils

2) サービスを起動する

service iscsi start

3) 自動起動の設定

chkconfig iscsid on

chkconfig iscsi on

- 4) iSCSI ターゲットに接続する(iSCSI ターゲットの IP は 10.10.1.10)
 # iscsiadm --mode discovery --type sendtargets --portal 10.10.1.10
 10.10.1.10:3260,1 iqn.2006-01.com.openfiler:tsn.b5ffc134d692
- 5) iSCSI ターゲットに接続できているかを確認する

iscsiadm -m node

10.10.1.10:3260,1 iqn.2006-01.com.openfiler:tsn.b5ffc134d692

6) ノードに接続する

iscsiadm --mode node --targetname iqn.2006-01.com.openfiler:tsn.b5ffc134d692 --login

- fdisk でデバイスとしての認識を確認 # fdisk -1
- 8) パーティションの作成 (/dev/sda を使用)

parted /dev/sda
(parted) mklabel gpt
(parted) mkpart

パーティションの名前? []?

ファイルシステムの種類? [ext2]? ext3

(parted) mkpartfs

パーティションの名前? []?

ファイルシステムの種類? [ext2]? ext3

(parted) quit

9) ファイルシステムの作成(ファイルシステムの容量により相当な時間を要する場合がある)

mkfs.ext3 /dev/sda1

```
10) マウントと確認
```

mkdir /iscsi1

mount -t ext3 /dev/sda1 /iscsi1

df -h /dev/sda1

	Filesystem	サイズ	使用	残り	使用%	マウント位置
	/dev/sda1	88G	184M	83G	1%	/iscsi1
11)	再起動後に自動的に	マウントさ	れるよう	fstab ?	を編集する(末尾に一行追加)

- # tail -1 /etc/fstab
- /dev/sda1 /iscsi1 ext3 _netdev 0.0
- 2.5 仮想化サーバの設定

仮想化サーバ用の OS (ゲスト OS) は CentOS5.6 の 32bit 版を使用し、ホストとなるマシンのローカルディ スク上にインストールし動作確認を実行した後、ライブマイグレーションの動作テストのために iSCSI 機能 を有する NAS のディスク領域にインストールした。

まず、仮想マシンの管理ツールである仮想マシンマネージャーを起動する(図 5)。この仮想マシンマネージャーには現在インストールされている仮想マシンが表示されており、稼働状況を確認することができる。

また、仮想マシンの起動、停止もここで行うことができる。新たに仮想マシンを追加する場合は、"新しい仮 想マシンの作成"ボタンを押して「新しい仮想マシン」ウインドウを表示し、必要事項を入力していく。

Pfreiber Mil Jahl 🕼 🖯 🗶 🖬	Togil &			
-		は 新しい仮想マシン ■●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●	*	■ 新しい仮想マシン 新しい仮想マシンを作成
2 전 2014년 2014		インストールメディアの場所 ○ CD-ROM または DVD を使用(<u>B</u>) ○		 ○ この長期マシンでストレージを有効に(E) ● コンピューターのハードディスク上にディスクイメージを存成(E) 20.0 ○ GB
		● ISO イメージを使用(): /guest/iso/CentOS-5.6-i386-bin-DVD.iso	✓ (@HI(W))	88.0 いわ avanabre in the default location ☑ 今すぐディスク全体を割り当てる(A) ① ○ 管理しているストレージか、他の既存のストレージを選択する(M)
		OS の種類とバージョンを選択してください OS の種類(工): Linux パージョン(X): Red Hat Enterprise Linux 5.4 or later	0 k	(#10(<u>8</u>))
Million-Mill		(++>ttA(C)) R&(B)	L AC(E)	年ャンセル(C) 展る(B) 激む(P)

図 5. 仮想マシンマネージャー

₩ 6	新] [い仮相マシンの作成
四 0.		V 「IX 泡が Y 🖌 🖌 Vノ I F IX

図 7. ローカルディスクへ作成

ステップ1で新規に作成する仮想マシンの名前を入力したあとインストール方法を選択する。今回は予め ダウンロードしておいた ISO イメージを使用するため、ローカルのインストールメディアを選択した。

ステップ2(図6)では、インストールメディアの場所として"ISOイメージを使用"を選び、参照ボタン を押して ISO メディアの検索画面からイメージファイルを指定する。次に OS の種類を"Linux"、バージョ ンを "Red Hat Enterprise Linux 5.4 or later" と選択した。

ステップ3で仮想マシンに割り当てるメモリ容量(512MB)と CPUの数(1)を指定する。

新規ストレージボリューム

名前(N): centos-nz2

ストレージボリュームのクォータ iscsi1's 使用可能な領域: 4543.87 GB

С МВ

k

最大容量(P): 20000 0 MB

フォーマット(<u>F</u>): raw

反想マシンカ

割り当て(A): 0

ステップ 4 (図 7) ではインストールするストレージを指定するが、ホストマシンのディスク上にインスト ールするには "コンピューターのハードディスク上にディスクイメージを作成"を選択して、必要なディス ク容量を指定する。また、NAS ヘインストールするには"管理しているストレージか、他の既存のストレー ジを選択する"を選択し、参照ボタンを押してストレージボリュームを指定する。「ストレージボリュームの 検索または作成」画面(図8)で iSCSI に対して新規ボリュームボタンを押して図9の作成画面に移り、ボリ ュームの名前と使用容量(20GB)を指定した。新しく作成したボリュームを選択して元の画面に戻る(図10)。

ら直接使用できるストレージユニットを作成します。

ージポリュームを追加

トレージユニット-Jing <u>Mame</u> 作成す。 ポリュームの手 日本 日本 日本 日本

Format 形式

Capacity: ボリュームの最大 サイブ

ポリュームに割り当てる

キャンセル(<u>C</u>) 完了(E)

Storage Pools	88 ~	917	フォーマット	(####
default	CereDS-5.6-(386-bin-DVD.ise	3.97.08	ise-	
Decentrations	central rut2 img	19 53 68	124	
DYTASTELETIONU-	centos5-clone-iscsi.img	19.53 GB	124	
interia.	centos5-hayakawa.img	20.00 GB	raw	centos5-hayakaw
SPERSONAL PRODUCT	centos5_on_iscs.img	20.00 08	13%	centos5_on_iscai
	critos-waka.img	9.77 GB	raw	critos-waka
				10
and the second se	and set and a set of the set of t		and the local damage of the	and a maintenance of

図 8. ストレージボリューム選択

	NUME:	190	
🔂 新しい仮	想マシンを	作成	
Ready to begin installat	tion of centos-nz		
OS: Red Ha	t Enterprise Linux	5.4 or later	
インストール、ローカル	CDROM/ISO		
XEV-: 512 M	P		
CPU: 1			
XFD-9/19.50	Discol/consta-412.1		
	and the second second	and the state of the second seco	Warden and American Street and
□ 1 2	ストールの形に設定	モカスタマイズす	6(<u>U</u>)
0.49	21-108032	モカスタマイズす	6(<u>U</u>)
 ロイン ア 詳細なオブション 	2)664682	8.0297429	6(2)
 マ 詳細なオブション ホストデバイス ethio 	2)	6	6(2)
 マ 詳細なオブション ホストデバイス ethic マ 留面 MAC アドレ; 	2)SOUCER	0	8(1)
 マ 詳細にオブション ホストデバイス etb0 ※ 開業 MAC アドレ3 52:54:00:a8:64:30 	21-LOBCER (Bridge 'br0') (BEREM)	0	6(<u>y</u>)
 イン 伊藤松オブション ホストデバイス eth0 田田 MAC アドレス 52:54:00:48:54:30 6354:00:48:54:30 	21-LOBCER (Bridge tor) (BREM) (ARE M) (km 0)	, () () () () () () () () () () () () ()	6(<u>y</u>)
 マンクロション アメントアバイス etbo 20 回波 MAC アドレン 52 54:00:06:05 32 の目的にの場所(2) アーキテクチャー(点) 	211-0001282 (Bridge 'br0') (Bridge 'br0') (Bridge 'br0') (Bridge 'br0') (Bridge 'br0')	¢ 5	6(<u>1</u>)

図 9. ストレージボリューム作成

2/r/AD NED ARD ARD ARD ARD ARD ARD ARD ARD ARD AR	
E Bach Barrow Crudes Anternational E contractor E cont	
Ell CIU dist	
Constant action References Constant for Constant for	
Eritas-ar Xr/o	
m tentact.tr	-
entasS-close 620	
Sco.	
Ren Ren	
eentes5.an.lacal	
Int contra	

▶ 新しい仮想マシンを作成

○ コンピューターのハードディスク上にディスクイメージを作成(B)

管理しているストレージか、他の既存のストレージを選択する(M)

図 10. NAS ヘインストール

キャンセル(C) 戻る(B) 進む(E)

9すぐディスク全体を割り当てる(A) **①**

⑦ この仮想マシンでストレージを有効に(E)

登録(店)... /iscsi1/centos-nz2.img

20.0 🗘 🖉 🛱

図 11. ネットワーク設定 図 12. 仮想マシンマネージャー 図 13. ホスト OS のデスクトップ

ステップ 5 (図 11) では接続するネットワークなどを指定する。ネットワークにはブリッジ接続を利用するため"ホストデバイス eth0 (Bridge 'br0')"を選択した。そして仮想化の種類として"kvm"、アーキテクチャとして"i686"を選択した。ここで"完了"ボタンを押すとゲスト OS のインストールが始まる。

インストールが終わると、仮想マシンマネージャー(図 12)に新たに仮想マシンが追加され、ホストマシン上で実行中であることが確認できる。(図 13)

3 ライブマイグレーション

ライブマイグレーションとは稼働中の仮想マシンを停止させずに別のホストマシンに移動する技術のこと であり、ホストマシンを停止させる計画メンテナンス時に、あらかじめ仮想マシンを別のホストマシンに移 行させておけば、仮想マシンで提供するサービスの無停止でのメンテナンス実施が可能になる。

ライブマイグレーションを行なうと、あるホス ト上で動作している仮想マシン上のメモリイメ ージが丸ごと別のホスト上の仮想マシンに移し 替えられ、稼働中の OS やアプリケーションソフ ト、ネットワーク接続などを一切停止・切断させ ることなく新しいホスト上で動作を継続するこ とができる。厳密には切り替えの瞬間にミリ秒単 位の瞬断が生じるが、ネットワークのセッション などが切断されることは無く、ユーザからは移動 が行われたことは分からない。



図 14. ライブマイグレーションとシステム構成

3.1 ライブマイグレーションの実行方法

ライブマイグレーションを実行するためには、ホスト1 (sv1) でゲスト OS を動作させ。仮想マシンマネ ージャーでゲストにカーソルを合わせ、右クリックするとマイグレーションを選択(図15) することができ る。その後でマイグレーション先のホスト (SV2) を選択(図16) して実行する。マイグレーションに要す る時間は、アプリケーションが動作している状態で約50秒程度、誰もログインしていない場合は36秒程度 であった。



図 15. ライブマイグレーションの実行1

P707-595 ## 5374 👹 🙆 🗾 🖩		- 4 ² 4 ² 4 ¹	D10 (A) 4# 1.52 PM
titlub woka 8.87	90 + 8 K	 ##V>>V# 	
PERMIT ARTIVIST AROU - TRADU	1	ファイル白 単単白 単月公 ヘルフ出	
PATHONN BE DATE ORIGIN	- 10 K Q2	● ●本 ト Ⅱ ⑥ ▼	
		4.8	· ON BRE
フォイル(E) 編集(E) 集号(E) 原類(E) ブックマーク(E)	V-A(1) =A(7(8)	W keighant (2004).	
4 4 · · · · · · · · · · · · · · · · · ·		- att	
MACRANUL* Rowth Discot*		In second to	
E LEANERIE		In contract class	
		E testest hejelese	
All Contents	C ##75567486-5x5		
In the second second	The second second second	interfact, or , bett	
APRIL IN CO.	-	A Distance in the	
CHANNERS AND A	All other water	and the second s	
antos webs #20	#1	1 10 10 10 10 10 10 10 10 10 10 10 10 10	
その(の) 単語ないの(語) 単分(の) ホーを得能(の)	#75-CVTF-U-babilly []	IIII sametar	
	5 manual friends	and a second class	
		and the second the second	
	Terrent Counters	and a second a	
		- age	
		-	
17.2 h (2000). 7	CORTA.		
		a alla	

図 16. ライブマイグレーションの実行 2

Programme and and both	• • 0 🗠 🖬		4 1 1	100200-00-948-138
	andre weeks \$12	1710 - T.K	· REVOUT	8-09 0.0
3-47回、 単品なたっ国、 単目の	1-1283	7+(h() ##() AN() ^67()		
7797-515 MR 5374	08698	- 15 m Q2	5 BK 18	
-		And a state of the	4.8	+ DV BEN
PARAMULANT AND AREA TATANE PART AND			W Acceleral (2008).	
A	(Select prove-are	HALL CANNEL OF		
44.49.04	2 In the Jack stack an	franking at he and the second second	inter control to	
#2128-1-5* @Cm	OS Dispost*		- #L*	
* IFFIELER			mit contest close	
	the second se		man control territory	
			- ete	
Contrasts		 ##75587750-535 	· IIII control it	
COLUENA			and control or beat	
		Higrating VM 'color-waka'	a	
A RELEASED	-	VM prim waka to will bill av2 ~7 c 20-5 mil	· III tons and	
A 19 19 19	41-11	er. Lus <mmusegst.< td=""><td>and a context</td><td></td></mmusegst.<>	and a context	
STATIALISEN N	Cathor and a R.S.	-	- +i+	
PART ANTION AND	1-1200	Migrating Banain	100 F00 F00	
		215 121 HE	ALL BALL	
		(advante)	and contest-class	
		(arrent	- +1+	
		# & a dimit # a figure a refer	BLD.	
			man contrad-to	
この範疇マシンのグラフィカルコンソームに重要や			- m + + + +	
			He was a second and second	
			and enter subs	
			-eets	

図 17. ライブマイグレーションの実行3

2 - (AR) 6823/238 8200 8-82880	2 F(AD 編集) 単20 AA7H
	L Max 1 H B v
	With Conception and Conception of Conception
	and the second s
	E state
	in trained down
ゲストロ戦争していません	E and a second despites a
	E second
	E tester, so had
	E the set
	E strantes
Color-webs EW73>	+ 2 4 + 12 00MA
57717-012 BR 1374 960089	= 13.00 QN
and the second se	
2+(A)) #5() 8() 8() 7-27-2() 7-6() -67()	e control topoteen
🏟 🗢 + 🔞 🔘 🏫 🐨 the Investment and Endproporties at	· (4
Backs-u. Ecents Dieper*	a and a set of the set
A TAPORTER .	
Contract of the local division of the local	a state
Contenta T # # ph	1. 丁学帮某些帮
(全学技術センタ	P- 工学系技统支援者
EXCLUSION OF THE PARTY	2
· ANYDOVA-Stat. · ctos usia SNVDO . · ctos usia BNVDO	

図 18. ライブマイグレーションの実行 4

3.2 ライブマイグレーションの動作確認

ライブマイグレーションの動作確認を次の方法で行った。ゲスト OS 上に DVDiso ファイル(約 4GB)を 置き、それを Windows 端末上の ftp クライアントで取得するのに 6 分程度を要する。このデータ転送中にラ イブマイグレーションを実行して、正常にファイルが受け取れるかどうかを確認した。その結果、受信ファ イルは元のものと一致した。したがって、ライブマイグレーション中もとぎれることなくサービスが行える ことが証明された。ただ、ライブマイグレーション中の転送速度は、数%程度に減少して(図 20)応答が極 端に悪くなった。次に、ゲスト OS のメモリ使用量を変更した場合とメモリ割り当て量を変更した場合で測 定を実行した。







図 20. クライアント PC のネットワーク使用率

4 まとめ

今回の実験では、無償版のソフトを利用して仮想化サーバを構築し、実際に運用して動作確認できた。また、ライブマイグレーションを実行して、その動作を確認した。その結果、ライブマイグレーションの実行時間はメモリの最大使用量に依存する事がわかり、なるべくサーバ負荷が少ない時に実行するのが最善であるとわかった。また、ゲスト OS に割り当てるメモリ量を調節する事で、ライブマイグレーションの実行時間を最適化できる事がわかった。

参考文献

- [1] 野崎ほか: "プライベートクラウド用サーバの構築",名古屋大学工学研究科・工学部「技報」Vol.14, 2012 年 3 月, PP. (2012 年 2 月 現時点で不明)
- [2] クラウド Watch Linux のカーネルに入った仮想化技術「KVM」,
 http://cloud.watch.impress.co.jp/docs/column/virtual/20110523_447267.html, Impress Group, 2012 年 2 月
- [3] クラウド Watch Linux 標準のサーバー仮想化機能「KVM」を試す、
 http://cloud.watch.impress.co.jp/docs/column/virtual/20110606_450502.html, Impress Group, 2012 年 2 月
- [4] COMPUTERWORLD 「KVM」—Linux 標準の仮想化機能の得意領域を知る, http://www.computerworld.jp/contents/111429/, IDG Interactive, 2012 年 2 月
- [5] @IT Linux 標準の仮想化技術「KVM」の仕組み, http://www.atmarkit.co.jp/flinux/rensai/kvm01/kvm01a.html, アイティメディア株式会社, 2012 年 2 月
- [6] IT 用語辞典, http://e-words.jp/, 株式会社インセプト, 2012 年 2 月