Pacemaker 1.1

Clusters from Scratch

Creare cluster Active/Passive e Active/Active su Fedora

Andrew Beekhof

Pacemaker 1.1 Clusters from Scratch
Creare cluster Active/Passive e Active/Active su Fedora
Edizione 5

Autore	Andrew Beekhof	*andrew@beekhof.net*
Traduttore	Raoul Scarazzini	*rasca@miamammausalinux.org*
Traduttore	Dan Frîncu	*df.cluster@gmail.com*

Lo scopo di questo documento è di fornire una guida completa per costruire un cluster active/passive con Pacemaker e mostrare come può essere convertito in una configurazione active/active.

Il cluster userà:

1. Fedora 13 come sistema operativo

2. Corosync per fornire i servizi di messaging e membership,

3. Pacemaker per la gestione delle risorse

4. DRBD come alternativa prezzo/prestazioni allo storage condiviso,

5. GFS2 come cluster filesystem (nella modalità active/active)

6. La shell crm per la visualizzazione della configurazione e l'attuazione delle modifiche

Per via del processo grafico di installazione di Fedora, diversi screenshot sono inclusi. Ad ogni modo questa guida è composta primariamente dai comandi, dalle ragioni per cui questi vengono eseguiti e l'output da loro prodotto.

[1] An explanation of CC-BY-SA is available at *http://creativecommons.org/licenses/by-sa/3.0/*

Indice

Lista delle figure

Prefazione

Indice

1. Convenzioni del documento

Questo manuale utilizza numerose convenzioni per evidenziare parole e frasi, ponendo attenzione su informazioni specifiche.

Nelle edizioni PDF e cartacea questo manuale utilizza caratteri presenti nel set *Font Liberation*[1]. Il set Font Liberation viene anche utilizzato nelle edizioni HTML se il set stesso è stato installato sul vostro sistema. In caso contrario, verranno mostrati caratteri alternativi ma equivalenti. Da notare: Red Hat Enterprise Linux 5 e versioni più recenti, includono per default il set Font Liberation.

1.1. Convenzioni tipografiche

Vengono utilizzate quattro convenzioni tipografiche per richiamare l'attenzione su parole e frasi specifiche. Queste convenzioni, e le circostanze alle quali vengono applicate, sono le seguenti.

`Neretto monospazio`

Usato per evidenziare l'input del sistema, incluso i comandi della shell, i nomi dei file ed i percorsi. Utilizzato anche per evidenziare tasti e combinazione di tasti. Per esempio:

> Per visualizzare i contenuti del file **`my_next_bestselling_novel`** nella vostra directory di lavoro corrente, inserire il comando **`cat my_next_bestselling_novel`** al prompt della shell e premere **`Invio`** per eseguire il comando.

Quanto sopra riportato include il nome del file, un comando della shell ed un tasto, il tutto riportato in neretto monospazio e distinguibile grazie al contesto.

Le combinazioni si distinguono dai tasti singoli tramite l'uso del segno più, il quale viene usato per creare una combinazione di tasti. Per esempio:

> Premere **`Invio`** per eseguire il comando.
>
> Premere **`Ctrl`**+**`Alt`**+**`F2`** per usare un terminale virtuale.

Il primo esempio evidenzia il tasto specifico singolo da premere. Il secondo riporta una combinazione di tasti: un insieme di tre tasti premuti contemporaneamente.

Se si discute del codice sorgente, i nomi della classe, i metodi, le funzioni i nomi della variabile ed i valori ritornati indicati all'interno di un paragrafo, essi verranno indicati come sopra, e cioè in **`neretto monospazio`**. Per esempio:

[1] https://fedorahosted.org/liberation-fonts/

> Le classi relative ad un file includono **`filesystem`** per file system, **`file`** per file, e **`dir`** per directory. Ogni classe possiede il proprio set associato di permessi.

Proportional Bold

Ciò denota le parole e le frasi incontrate su di un sistema, incluso i nomi delle applicazioni; il testo delle caselle di dialogo; i pulsanti etichettati; le caselle e le etichette per pulsanti di selezione, titoli del menu e dei sottomenu. Per esempio:

> Selezionare **Sistema** → **Preferenze** → **Mouse** dalla barra del menu principale per lanciare **Preferenze del Mouse**. Nella scheda **Pulsanti**, fate clic sulla casella di dialogo **mouse per mancini**, e successivamente fate clic su **Chiudi** per cambiare il pulsante primario del mouse da sinistra a destra (rendendo così il mouse idoneo per un utilizzo con la mano sinistra).
>
> Per inserire un carattere speciale in un file **gedit** selezionare **Applicazioni** → **Accessori** → **Mappa del carattere** dalla barra del menu principale. Selezionare successivamente **Cerca** → **Trova…** dal menu **Mappa del carattere**, digitare il nome desiderato nel campo **Cerca** e selezionare **Successivo**. Il carattere desiderato sarà evidenziato nella **Tabella dei caratteri**. Eseguire un doppio clic sul carattere per poterlo posizionare nel campo **Testo da copiare** e successivamente fare clic sul pulsante **Copia**. Ritornare sul documento e selezionare **Modifica** → **Incolla** dalla barra del menu di **gedit**.

Il testo sopra riportato include i nomi delle applicazioni; nomi ed oggetti del menu per l'intero sistema; nomi del menu specifici alle applicazioni; e pulsanti e testo trovati all'interno di una interfaccia GUI, tutti presentati in neretto proporzionale e distinguibili dal contesto.

`Corsivo neretto monospazio` o ***`Corsivo neretto proporzionale`***

Sia se si tratta di neretto monospazio o neretto proporzionale, l'aggiunta del carattere corsivo indica un testo variabile o sostituibile . Il carattere corsivo denota un testo che non viene inserito letteralmente, o visualizzato che varia a seconda delle circostanze. Per esempio:

> Per collegarsi ad una macchina remota utilizzando ssh, digitare **`ssh`** ***`username@domain.name`*** al prompt della shell. Se la macchina remota è **`example.com`** ed il nome utente sulla macchina interessata è john, digitare **`ssh john@example.com`**.
>
> Il comando **`mount -o remount`** ***`file-system`*** rimonta il file system indicato. Per esempio, per rimontare il file system **`/home`**, il comando è **`mount -o remount /home`**.
>
> Per visualizzare la versione di un pacchetto attualmente installato, utilizzare il comando **`rpm -q`** ***`package`***. Esso ritornerà il seguente risultato: ***`package-version-release`***.

Da notare le parole in corsivo grassetto - username, domain.name, file-system, package, version e release. Ogni parola funge da segnaposto, sia esso un testo inserito per emettere un comando o mostrato dal sistema.

Oltre all'utilizzo normale per la presentazione di un titolo, il carattere Corsivo denota il primo utilizzo di un termine nuovo ed importante. Per esempio:

> Publican è un sistema di pubblicazione per *DocBook*.

1.2. Convenzioni del documento

Gli elenchi originati dal codice sorgente e l'output del terminale vengono evidenziati rispetto al testo circostante.

L'output inviato ad un terminale è impostato su **`tondo monospazio`** e così presentato:

```
books        Desktop   documentation  drafts  mss    photos   stuff  svn
books_tests  Desktop1  downloads      images  notes  scripts  svgs
```

Gli elenchi del codice sorgente sono impostati in **`tondo monospazio`** ma vengono presentati ed evidenziati nel modo seguente:

```
package org.jboss.book.jca.ex1;

import javax.naming.InitialContext;

public class ExClient
{
   public static void main(String args[])
       throws Exception
   {
      InitialContext iniCtx = new InitialContext();
      Object         ref    = iniCtx.lookup("EchoBean");
      EchoHome       home   = (EchoHome) ref;
      Echo           echo   = home.create();

      System.out.println("Created Echo");

      System.out.println("Echo.echo('Hello') = " + echo.echo("Hello"));
   }
}
```

1.3. Note ed avvertimenti

E per finire, tre stili vengono usati per richiamare l'attenzione su informazioni che in caso contrario potrebbero essere ignorate.

Nota

Una nota è un suggerimento o un approccio alternativo per il compito da svolgere. Non dovrebbe verificarsi alcuna conseguenza negativa se la nota viene ignorata, ma al tempo stesso potreste non usufruire di qualche trucco in grado di facilitarvi il compito.

Importante

Le caselle 'importante' riportano informazioni che potrebbero passare facilmente inosservate: modifiche alla configurazione applicabili solo alla sessione corrente, o servizi i quali necessitano di un riavvio prima di applicare un aggiornamento. Ignorare queste caselle non causa alcuna perdita di dati ma potrebbe causare irritazione e frustrazione da parte dell'utente.

Avvertimento

Un Avvertimento non dovrebbe essere ignorato. Se ignorato, potrebbe verificarsi una perdita di dati.

2. We Need Feedback!

If you find a typographical error in this manual, or if you have thought of a way to make this manual better, we would love to hear from you! Please submit a report in Bugzilla[2] against the product **Pacemaker.**

When submitting a bug report, be sure to mention the manual's identifier: *Clusters_from_Scratch*

If you have a suggestion for improving the documentation, try to be as specific as possible when describing it. If you have found an error, please include the section number and some of the surrounding text so we can find it easily.

[2] *http://bugs.clusterlabs.org*

Leggimi-Prima

Indice

1.1. Scopo di questo documento

Computer clusters can be used to provide highly available services or resources. The redundancy of multiple machines is used to guard against failures of many types.

This document will walk through the installation and setup of simple clusters using the Fedora distribution, version 14.

The clusters described here will use Pacemaker and Corosync to provide resource management and messaging. Required packages and modifications to their configuration files are described along with the use of the Pacemaker command line tool for generating the XML used for cluster control.

Pacemaker is a central component and provides the resource management required in these systems. This management includes detecting and recovering from the failure of various nodes, resources and services under its control.

When more in depth information is required and for real world usage, please refer to the *Pacemaker Explained*[1] manual.

1.2. Cos'è Pacemaker?

Pacemaker is a cluster resource manager. It achieves maximum availability for your cluster services (aka. resources) by detecting and recovering from node and resource-level failures by making use of the messaging and membership capabilities provided by your preferred cluster infrastructure (either Corosync or Heartbeat).

Pacemaker’s key features include:

- Rilevazione e ripristino di malfunzionamenti di nodi e servizi
- Storage agnostic, non richiede uno storage condiviso
- Resource agnostic, tutto quello che può essere scriptato può essere clusterizzato
- Supports STONITH for ensuring data integrity
- Supporto a cluster grandi e piccoli
- Supports both quorate and resource driven clusters

[1] http://www.clusterlabs.org/doc/

- Supports practically any redundancy configuration
- Configurazione replicata automaticamente che può essere aggiornata da qualsiasi nodo
- Capacità di specificare ordine, collocazione e anti-collocazione per i servizi lato cluster
- Support for advanced service types
  - Cloni: per servizi che necessitano di essere attivi su nodi multipli
  - Muliti-state: per servizi con modi multipli (ad esempio master/slave, primary/secondary/
- cluster shell unificata e scriptabile

1.3. Architettura di Pacemaker

Al livello più elevato il cluster è composto da tre componenti:

- Non-cluster aware components (illustrated in green). These pieces include the resources themselves, scripts that start, stop and monitor them, and also a local daemon that masks the differences between the different standards these scripts implement.
- Resource management Pacemaker provides the brain (illustrated in blue) that processes and reacts to events regarding the cluster. These events include nodes joining or leaving the cluster; resource events caused by failures, maintenance, scheduled activities; and other administrative actions. Pacemaker will compute the ideal state of the cluster and plot a path to achieve it after any of these events. This may include moving resources, stopping nodes and even forcing them offline with remote power switches.
- Low level infrastructure Corosync provides reliable messaging, membership and quorum information about the cluster (illustrated in red).

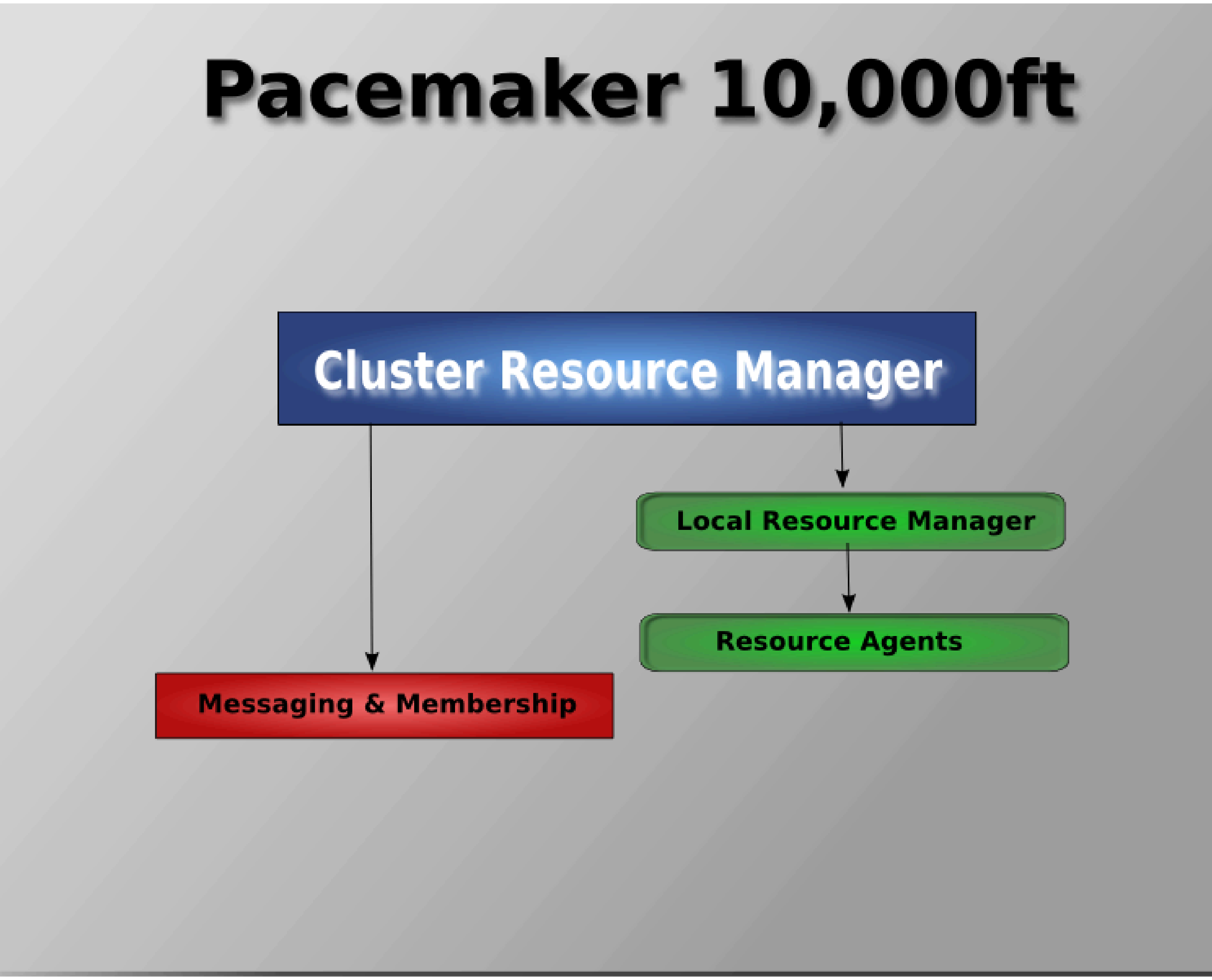


Figura 1.1. Panoramica concettuale dello Stack

When combined with Corosync, Pacemaker also supports popular open source cluster filesystems. [2]

Due to recent standardization within the cluster filesystem community, they make use of a common distributed lock manager which makes use of Corosync for its messaging capabilities and Pacemaker for its membership (which nodes are up/down) and fencing services.

[2] Even though Pacemaker also supports Heartbeat, the filesystems need to use the stack for messaging and membership and Corosync seems to be what they're standardizing on. Technically it would be possible for them to support Heartbeat as well, however there seems little interest in this.

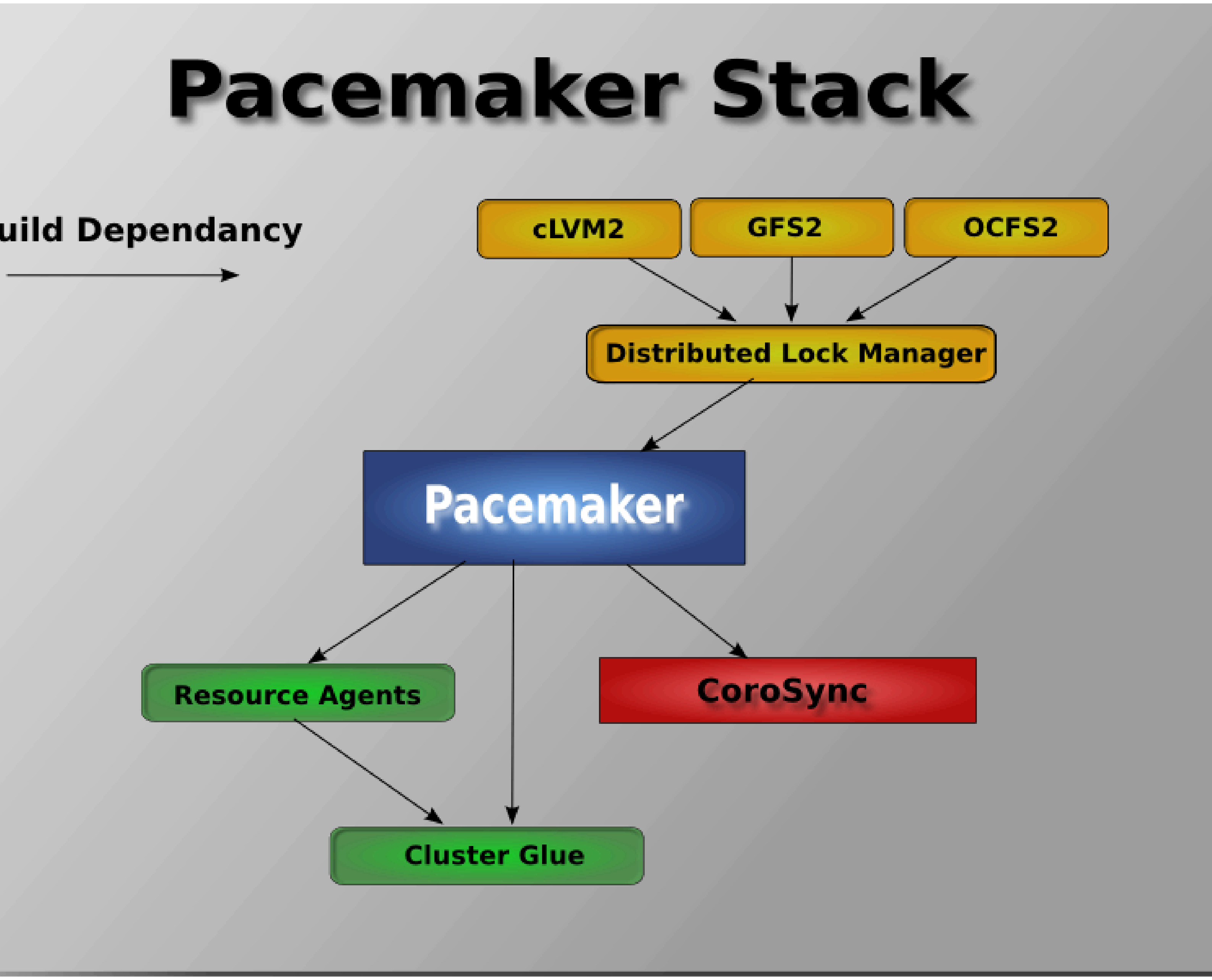


Figura 1.2. Lo stack Pacemaker

1.3.1. Componenti interni

Pacemaker stesso è composto da quatto componenti chiave (illustrati sotto nello stesso schema di colori del diagramma precedente):

- CIB (acronimo di come Cluster Information Base)
- CRMd (acronimo di Cluster Resource Management daemon)
- PEngine (acronimo di Policy Engine)
- STONITHd

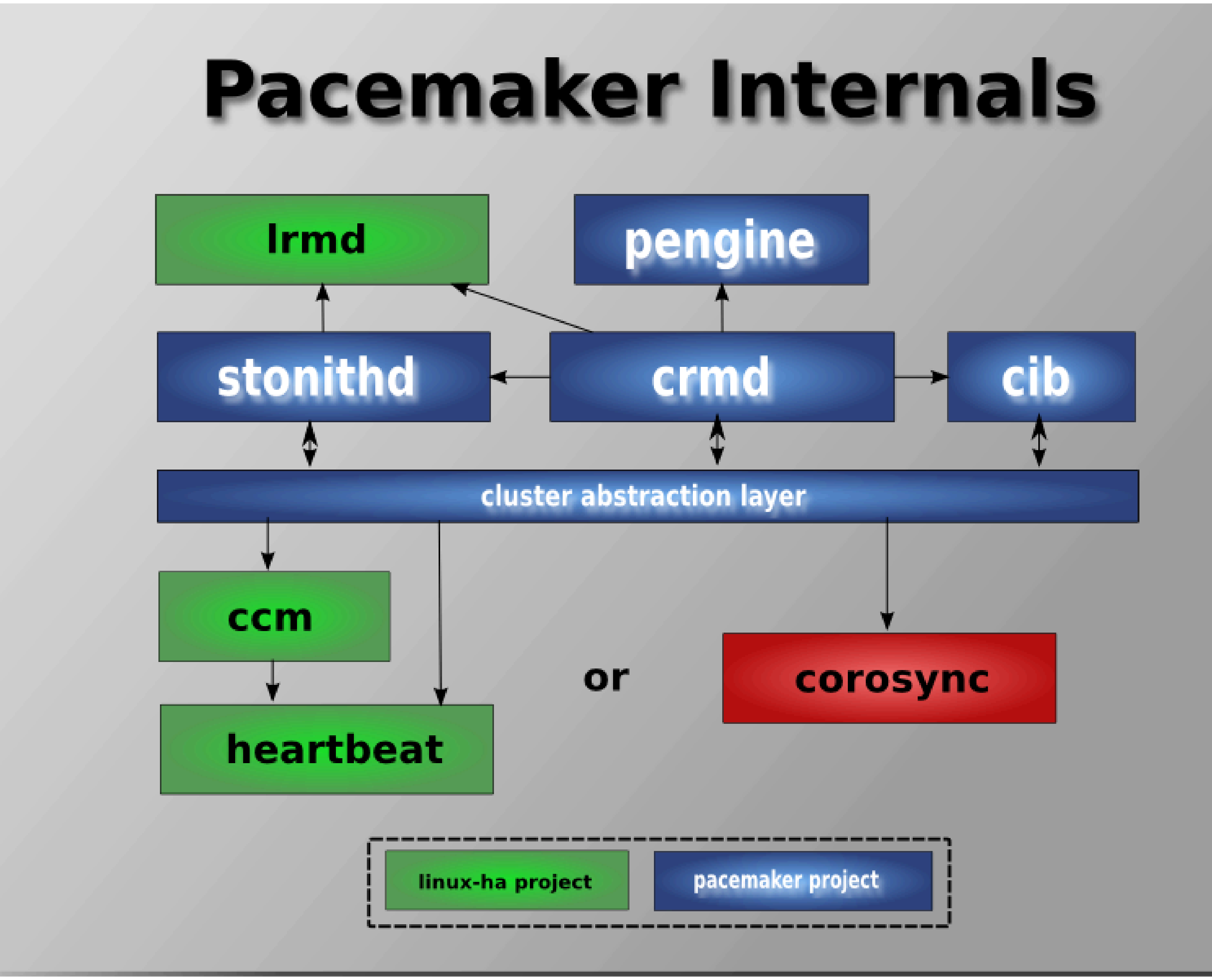


Figura 1.3. Componenti interni

The CIB uses XML to represent both the cluster’s configuration and current state of all resources in the cluster. The contents of the CIB are automatically kept in sync across the entire cluster and are used by the PEngine to compute the ideal state of the cluster and how it should be achieved.

This list of instructions is then fed to the DC (Designated Co-ordinator). Pacemaker centralizes all cluster decision making by electing one of the CRMd instances to act as a master. Should the elected CRMd process, or the node it is on, fail… a new one is quickly established.

The DC carries out the PEngine’s instructions in the required order by passing them to either the LRMd (Local Resource Management daemon) or CRMd peers on other nodes via the cluster messaging infrastructure (which in turn passes them on to their LRMd process).

Gli altri nodi riferiscono i risultati delle loro operazioni al DC. Attraverso l'analisi dei risultati aspettati e di quelli attuali, i nodi eseguiranno qualsiasi azione necessaria per attendere il completamento della precedente oppure interromperanno il processo, richiedendo al PEngine di calcolare nuovamente lo stato ideale del cluster basandosi sui risulati inaspettati.

In alcuni casi, potrebbe essere necessario spegnere i nodi per preservare dati condifivi o completare il ripristino di una risorsa. Per questo in Pacemaker esiste STONITHd. STONITH è un acronimo per Shoot-The-Other-Node-In-The-Head e viene implementato tipicamente con un switch di potenza remoto. In Pacemaker i dispositivi STONITH sono modellati come risorse (e configurati all'interno del CIB) per facilitare il monitoraggio delle anomalie. STONITHd si prende cura di capire la topologia STONITH così che i suoi client debbano unicamente richiedere unicamente la morte di un nodo ed esso si preoccupi del resto.

1.4. Tipologia dei cluster Pacemaker

Pacemaker non fa alcuna ipotesi in merito all'ambiente operativo, questo consente di supportare praticamente qualsiasi *configurazione ridondata*[3] come Active/Active, Active/Passive, N+1, N+M, N-to-1 e N-to-N.

In this document we will focus on the setup of a highly available Apache web server with an Active/ Passive cluster using DRBD and Ext4 to store data. Then, we will upgrade this cluster to Active/Active using GFS2.

[3] http://en.wikipedia.org/wiki/High-availability_cluster#Node_configurations

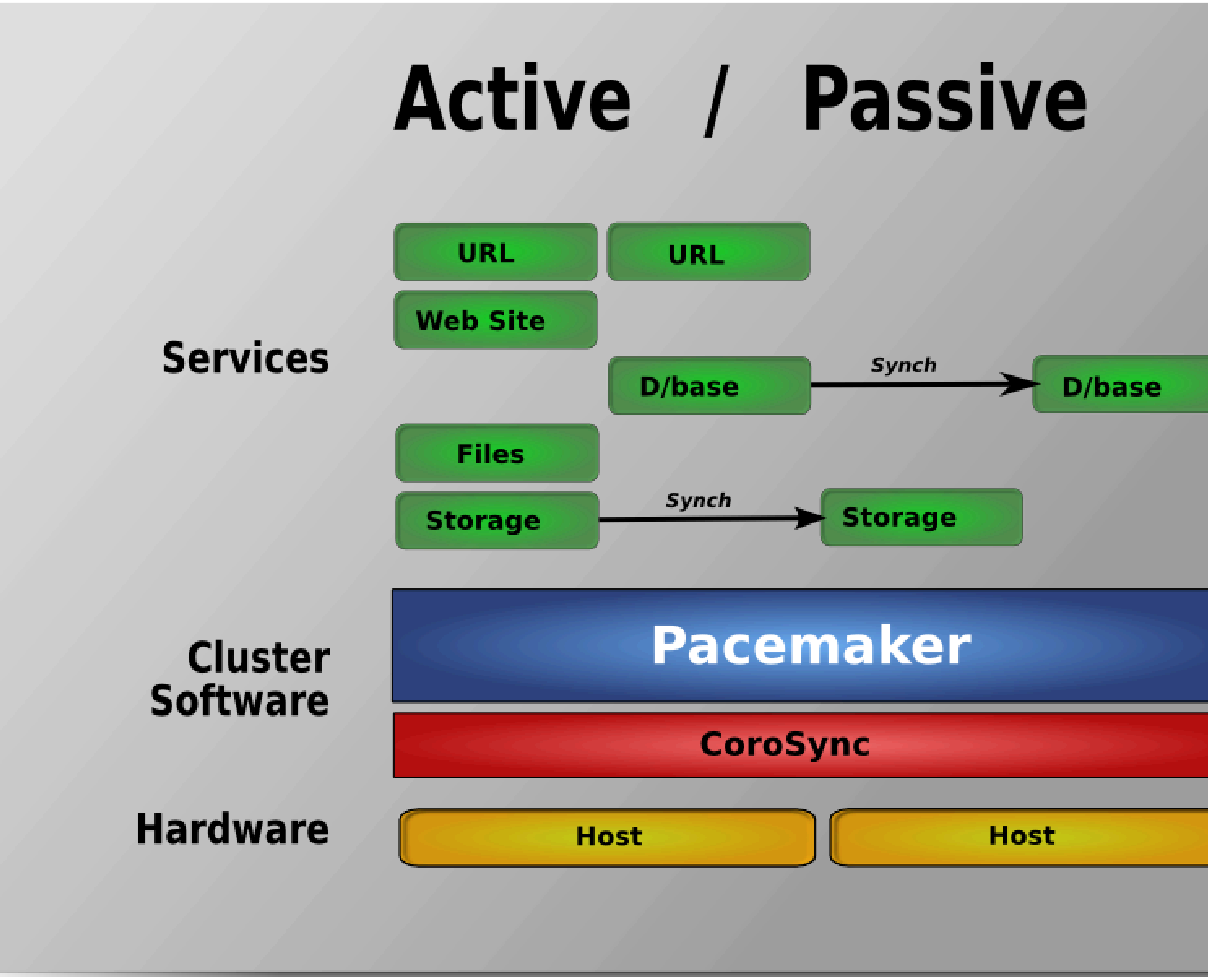


Figura 1.4. Ridondanza Active/Passive

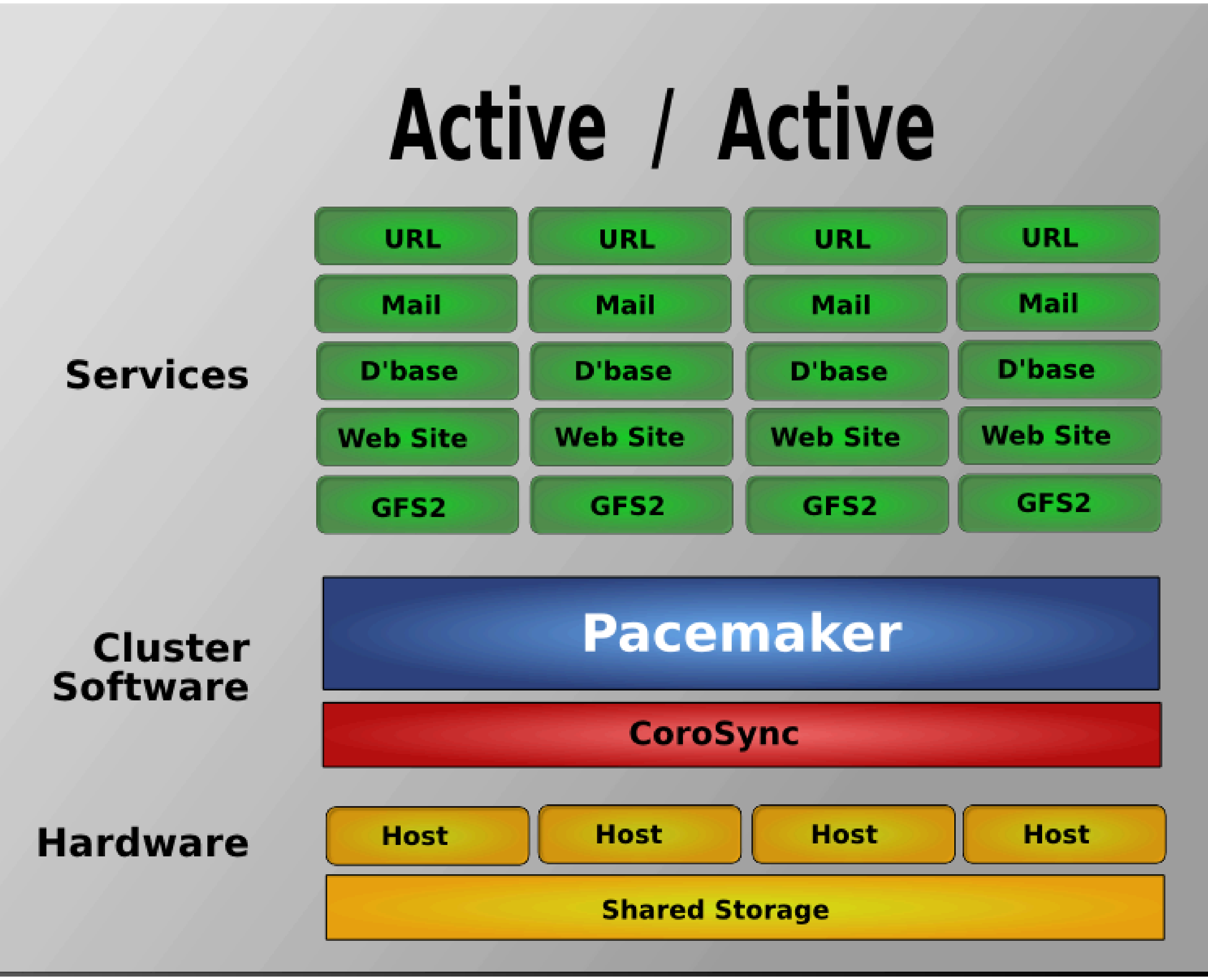


Figura 1.5. Ridondanza N a N

Installazione

Indice

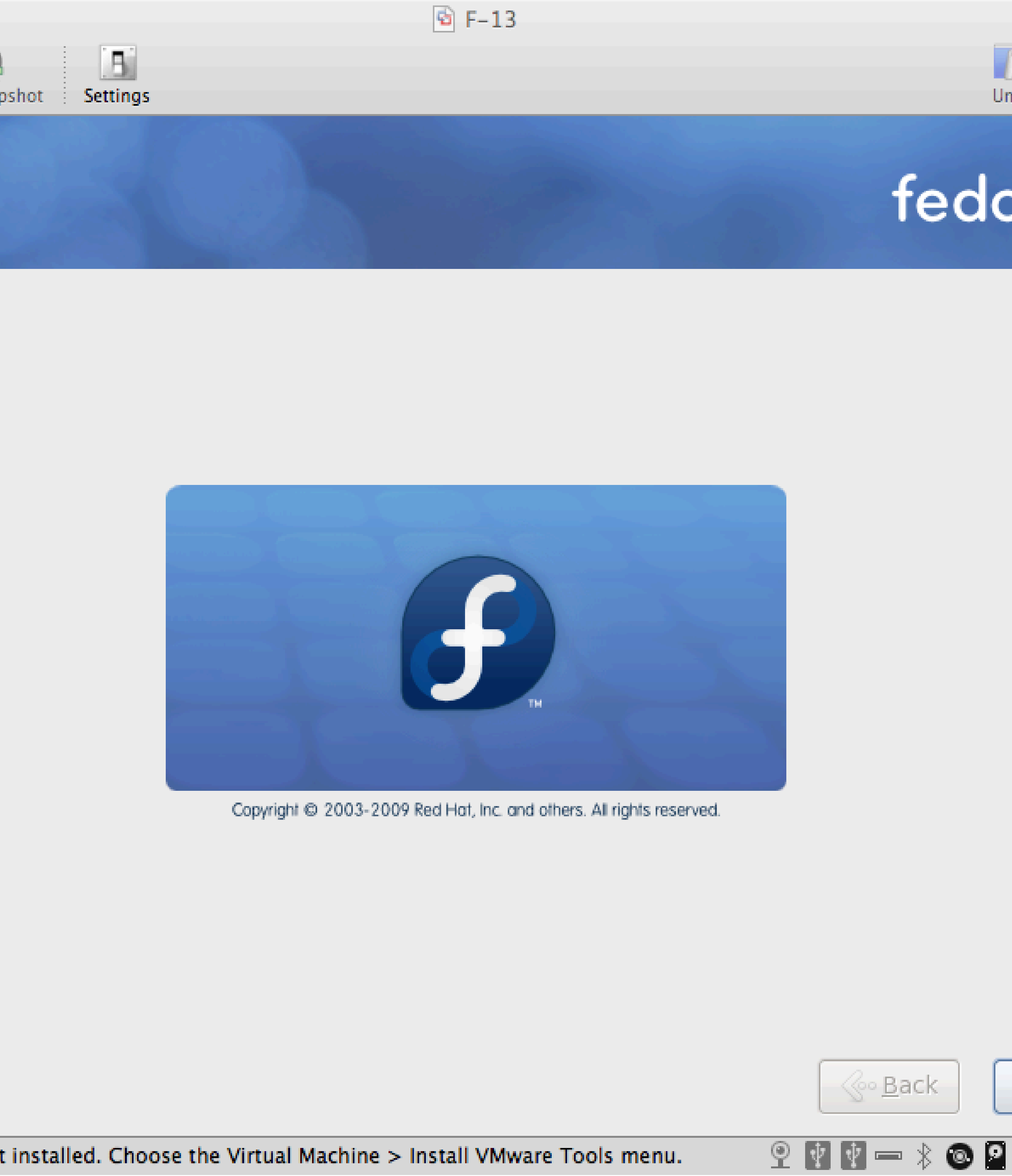
2.1. Installazione del sistema operativo

Detailed instructions for installing Fedora are available at *http://docs.fedoraproject.org/install-guide/f13/* in a number of languages. The abbreviated version is as follows…

Point your browser to *http://fedoraproject.org/en/get-fedora-all*, locate the Install Media section and download the install DVD that matches your hardware.

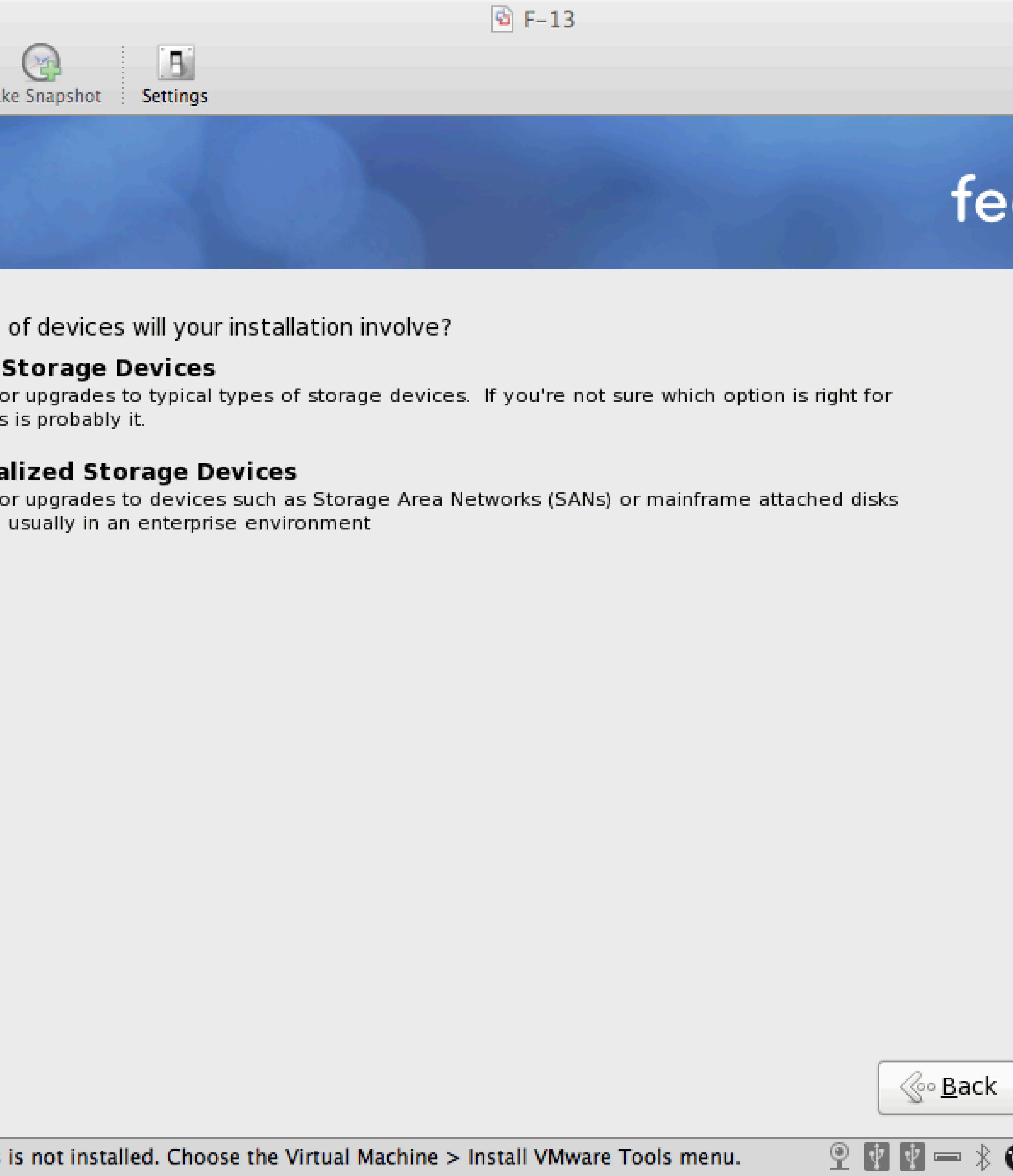
Burn the disk image to a DVD [1] and boot from it. Or use the image to boot a virtual machine as I have done here. After clicking through the welcome screen, select your language and keyboard layout [2]

[1] *http://docs.fedoraproject.org/readme-burning-isos/en-US.html*
[2] *http://docs.fedoraproject.org/install-guide/f13/en-US/html/s1-langselection-x86.html*


F-13
pshot
Settings
Un
fedo
TM
Copyright © 2003-2009 Red Hat, Inc. and others. All rights reserved.
Back

t installed. Choose the Virtual Machine > Install VMware Tools menu.

Figura 2.1. Installation: Good choice



of devices will your installation involve?

Storage Devices
or upgrades to typical types of storage devices. If you're not sure which option is right for
s is probably it.

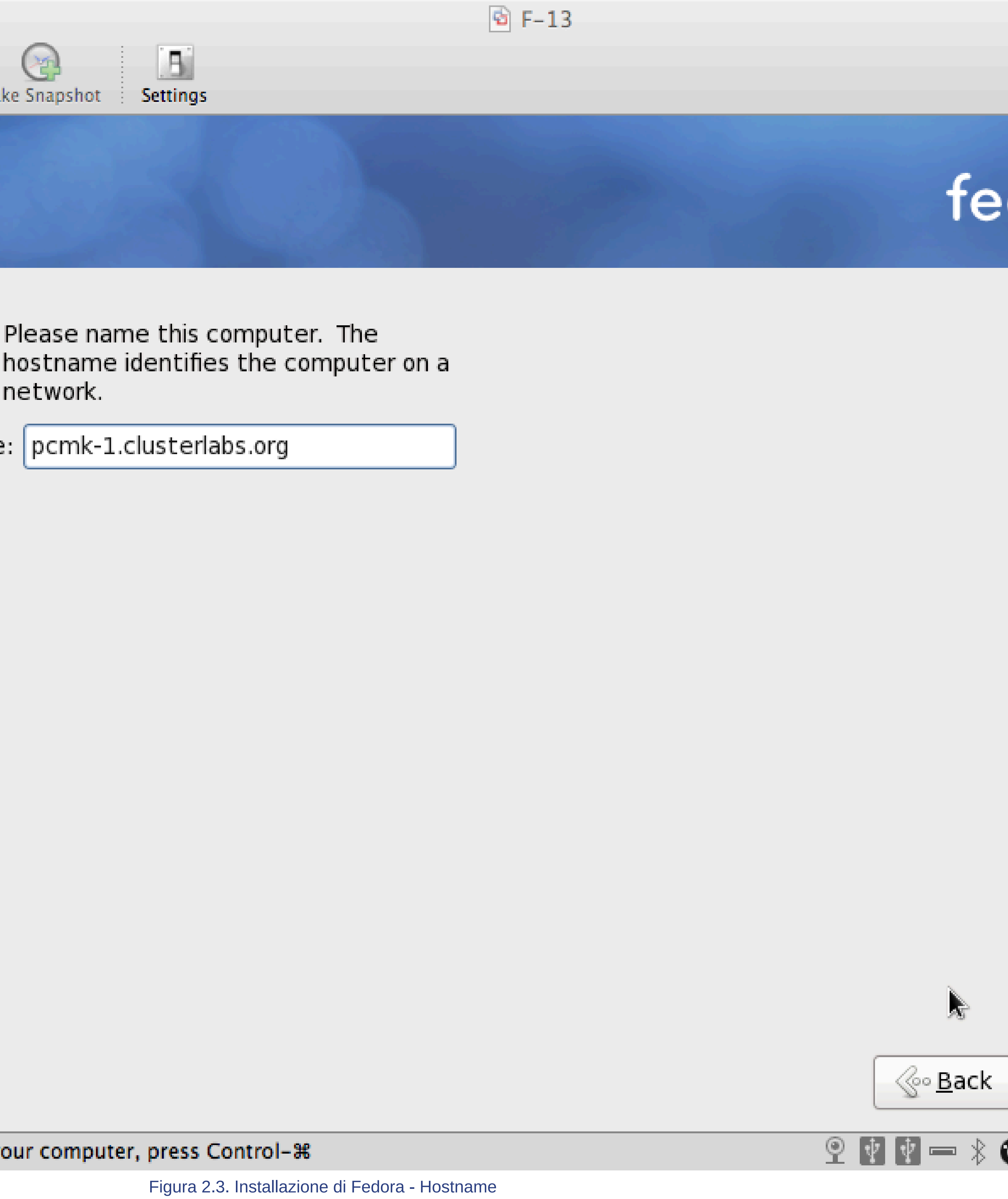
alized Storage Devices
or upgrades to devices such as Storage Area Networks (SANs) or mainframe attached disks
usually in an enterprise environment

is not installed. Choose the Virtual Machine > Install VMware Tools menu.

Figura 2.2. Installazione di Fedora - Storage Devices

Assign your machine a host name. [3] I happen to control the clusterlabs.org domain name, so I will use that here.

[3] *http://docs.fedoraproject.org/install-guide/f13/en-US/html/sn-networkconfig-fedora.html*



Please name this computer. The hostname identifies the computer on a network.

e: pcmk-1.clusterlabs.org

Figura 2.3. Installazione di Fedora - Hostname

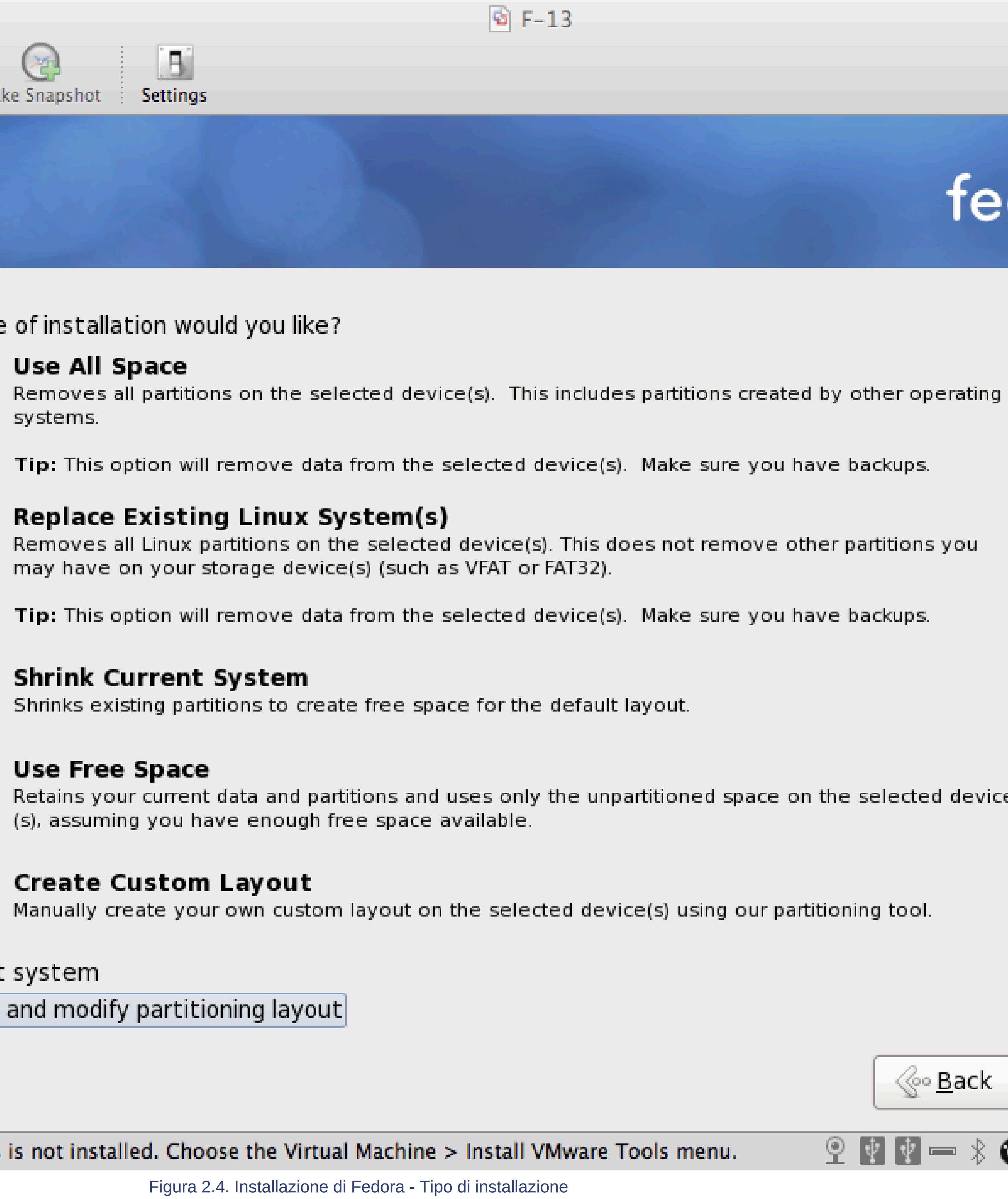
You will then be prompted to indicate the machine’s physical location and to supply a root password. [4]

Now select where you want Fedora installed. [5]

As I don’t care about any existing data, I will accept the default and allow Fedora to use the complete drive. However I want to reserve some space for DRBD, so I’ll check the Review and modify partitioning layout box.

[4] *http://docs.fedoraproject.org/install-guide/f13/en-US/html/sn-account_configuration.html*

[5] *http://docs.fedoraproject.org/install-guide/f13/en-US/html/s1-diskpartsetup-x86.html*



e of installation would you like?

Use All Space
Removes all partitions on the selected device(s). This includes partitions created by other operating systems.

Tip: This option will remove data from the selected device(s). Make sure you have backups.

Replace Existing Linux System(s)
Removes all Linux partitions on the selected device(s). This does not remove other partitions you may have on your storage device(s) (such as VFAT or FAT32).

Tip: This option will remove data from the selected device(s). Make sure you have backups.

Shrink Current System
Shrinks existing partitions to create free space for the default layout.

Use Free Space
Retains your current data and partitions and uses only the unpartitioned space on the selected device (s), assuming you have enough free space available.

Create Custom Layout
Manually create your own custom layout on the selected device(s) using our partitioning tool.

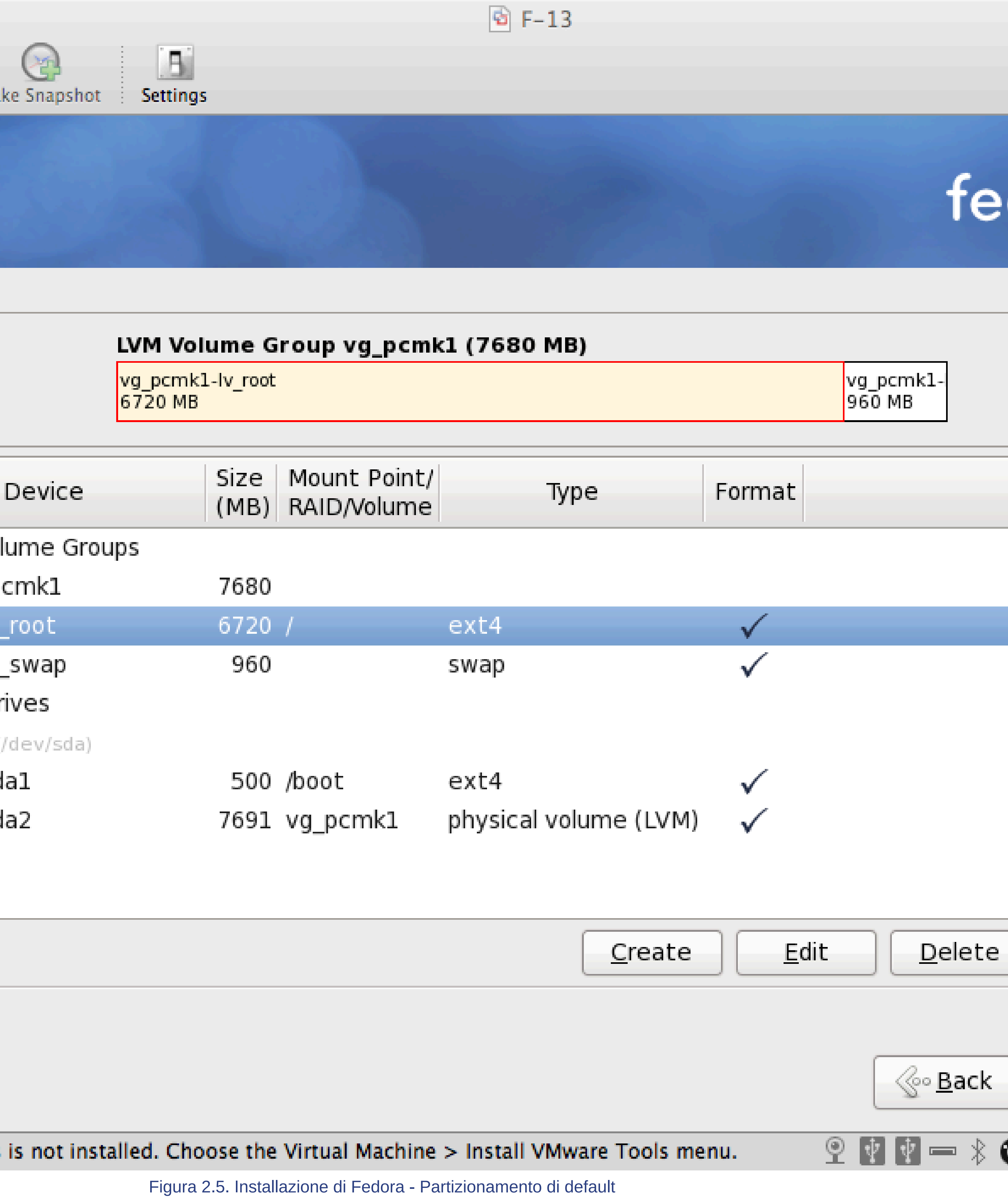
t system

and modify partitioning layout

is not installed. Choose the Virtual Machine > Install VMware Tools menu.

Figura 2.4. Installazione di Fedora - Tipo di installazione

By default, Fedora will give all the space to the / (aka. root) partition. Wel’ll take some back so we can use DRBD.



LVM Volume Group vg_pcmk1 (7680 MB)

vg_pcmk1-lv_root 6720 MB	vg_pcmk1- 960 MB

Device	Size (MB)	Mount Point/ RAID/Volume	Type	Format
lume Groups				
cmk1	7680			
_root	6720	/	ext4	✓
_swap	960		swap	✓
rives				
/dev/sda)				
la1	500	/boot	ext4	✓
la2	7691	vg_pcmk1	physical volume (LVM)	✓

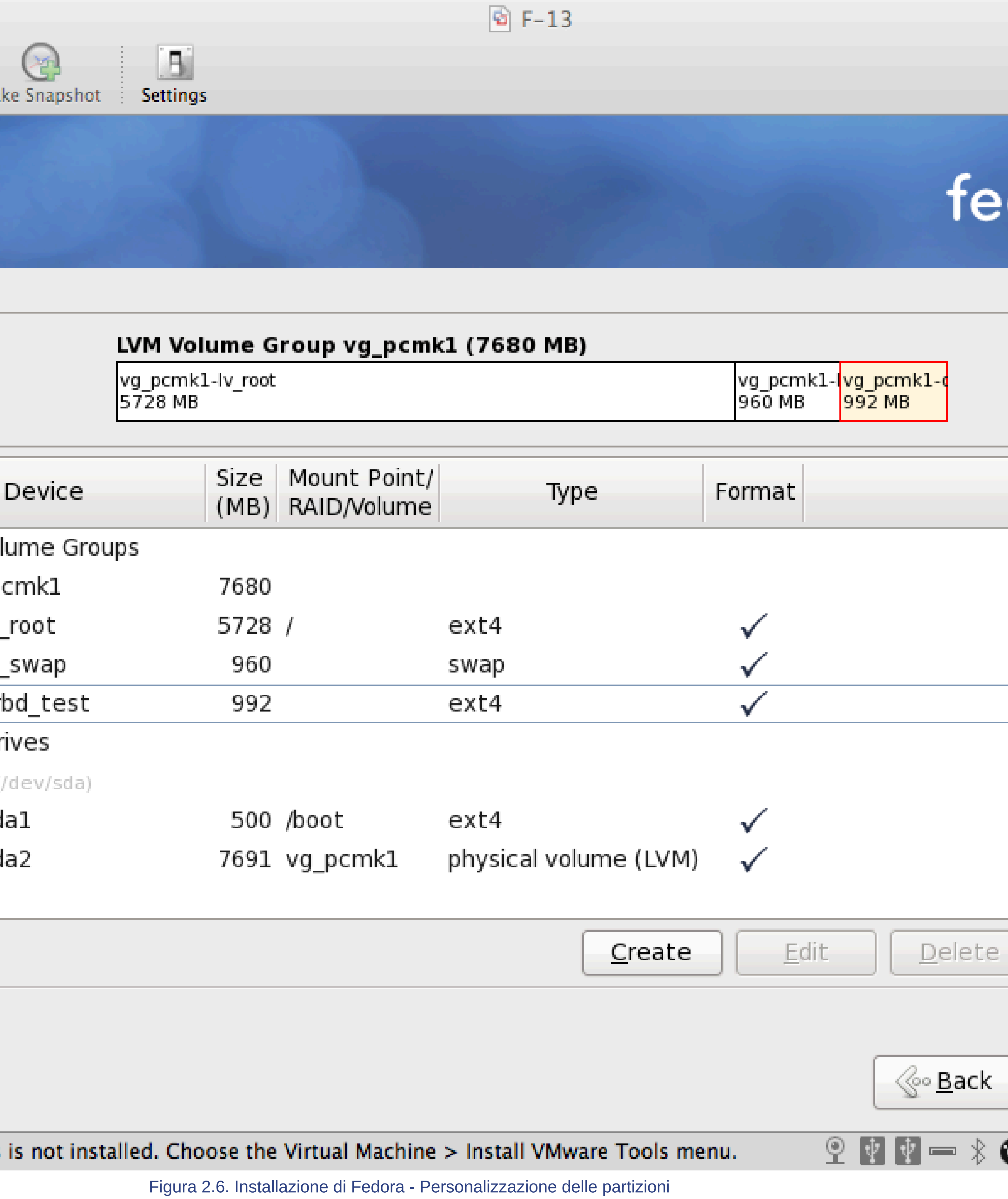
is not installed. Choose the Virtual Machine > Install VMware Tools menu.

Figura 2.5. Installazione di Fedora - Partizionamento di default

Lo schema di partizionamento finale dovrà somigliare a quanto illustrato nel diagramma sotto.

Importante

If you plan on following the DRBD or GFS2 portions of this guide, you should reserve at least 1Gb of space on each machine from which to create a shared volume. Fedora Installation - Customize PartitioningFedora Installation: Create a partition to use (later) for website data

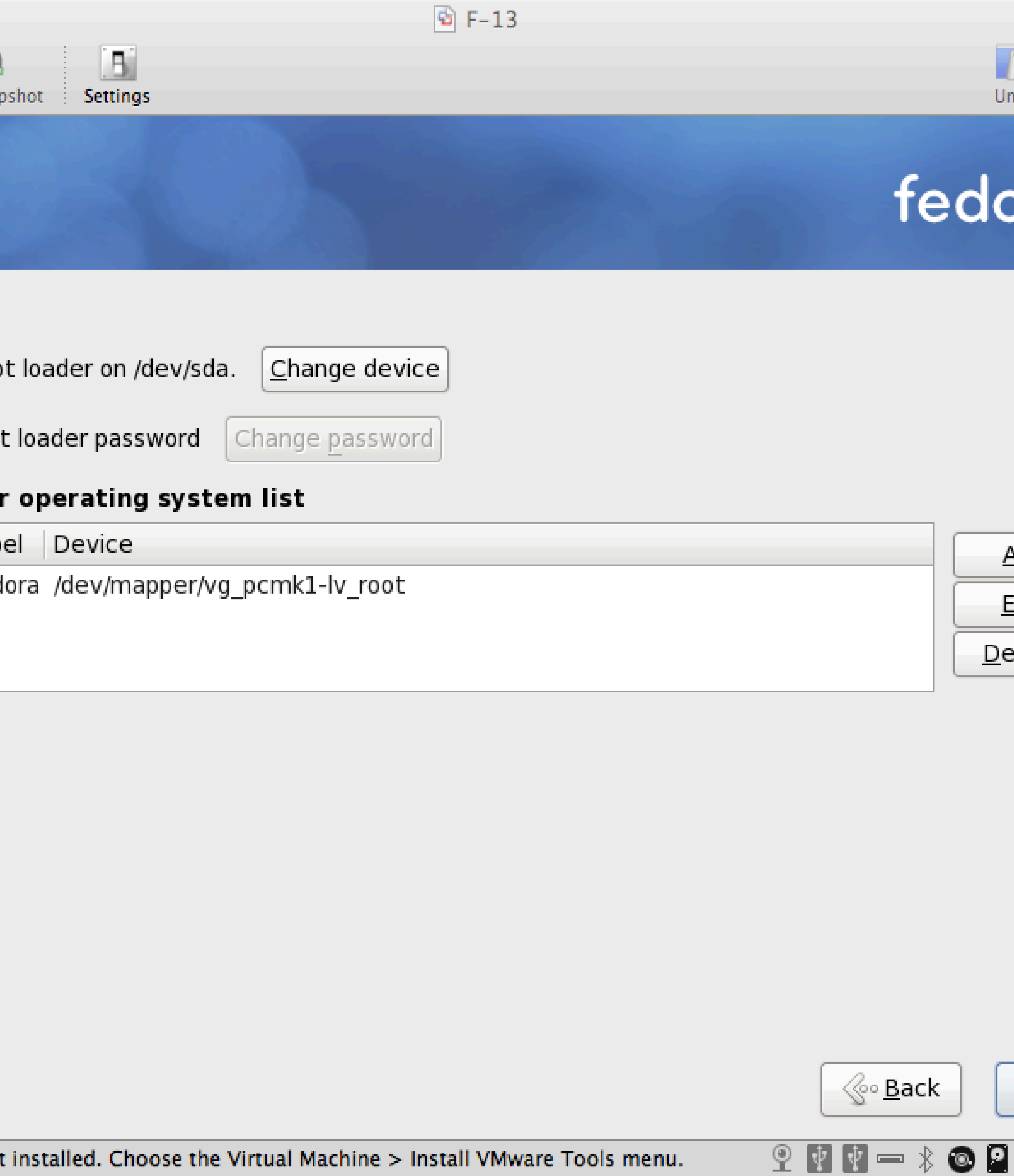


LVM Volume Group vg_pcmk1 (7680 MB)

Device	Size (MB)	Mount Point/ RAID/Volume	Type	Format
lume Groups				
cmk1	7680			
_root	5728	/	ext4	✓
_swap	960		swap	✓
bd_test	992		ext4	✓
rives				
/dev/sda)				
la1	500	/boot	ext4	✓
la2	7691	vg_pcmk1	physical volume (LVM)	✓

is not installed. Choose the Virtual Machine > Install VMware Tools menu.

Figura 2.6. Installazione di Fedora - Personalizzazione delle partizioni



ot loader on /dev/sda.

t loader password

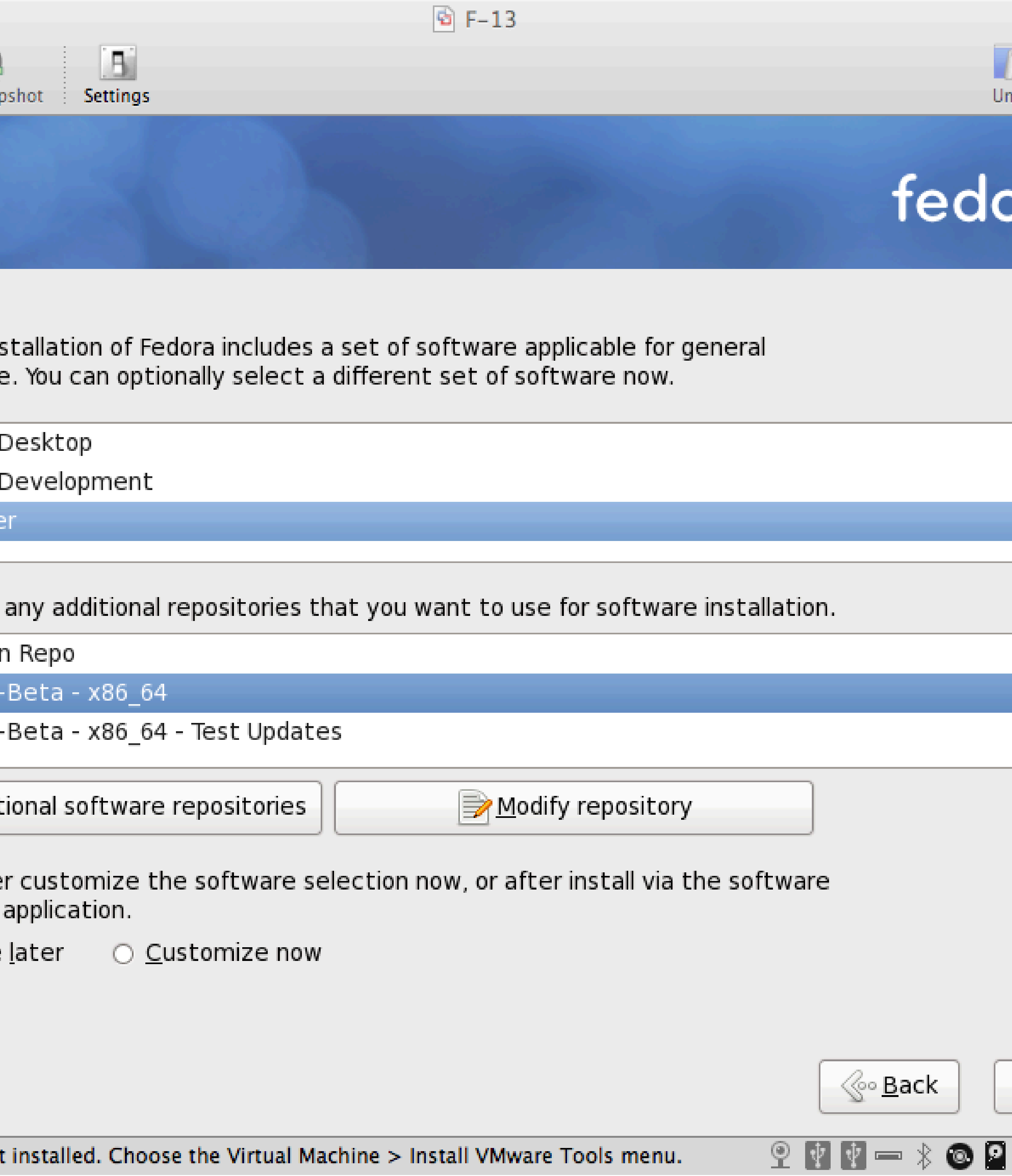
r operating system list

el	Device
lora	/dev/mapper/vg_pcmk1-lv_root

t installed. Choose the Virtual Machine > Install VMware Tools menu.

Figura 2.7. Installazione di Fedora - Bootloader

Next choose which software should be installed. Change the selection to Web Server since we plan on using Apache. Don’t enable updates yet, we’ll do that (and install any extra software we need) later. After you click next, Fedora will begin installing.



stallation of Fedora includes a set of software applicable for general
e. You can optionally select a different set of software now.

Desktop
Development
er

any additional repositories that you want to use for software installation.

n Repo
-Beta - x86_64
-Beta - x86_64 - Test Updates

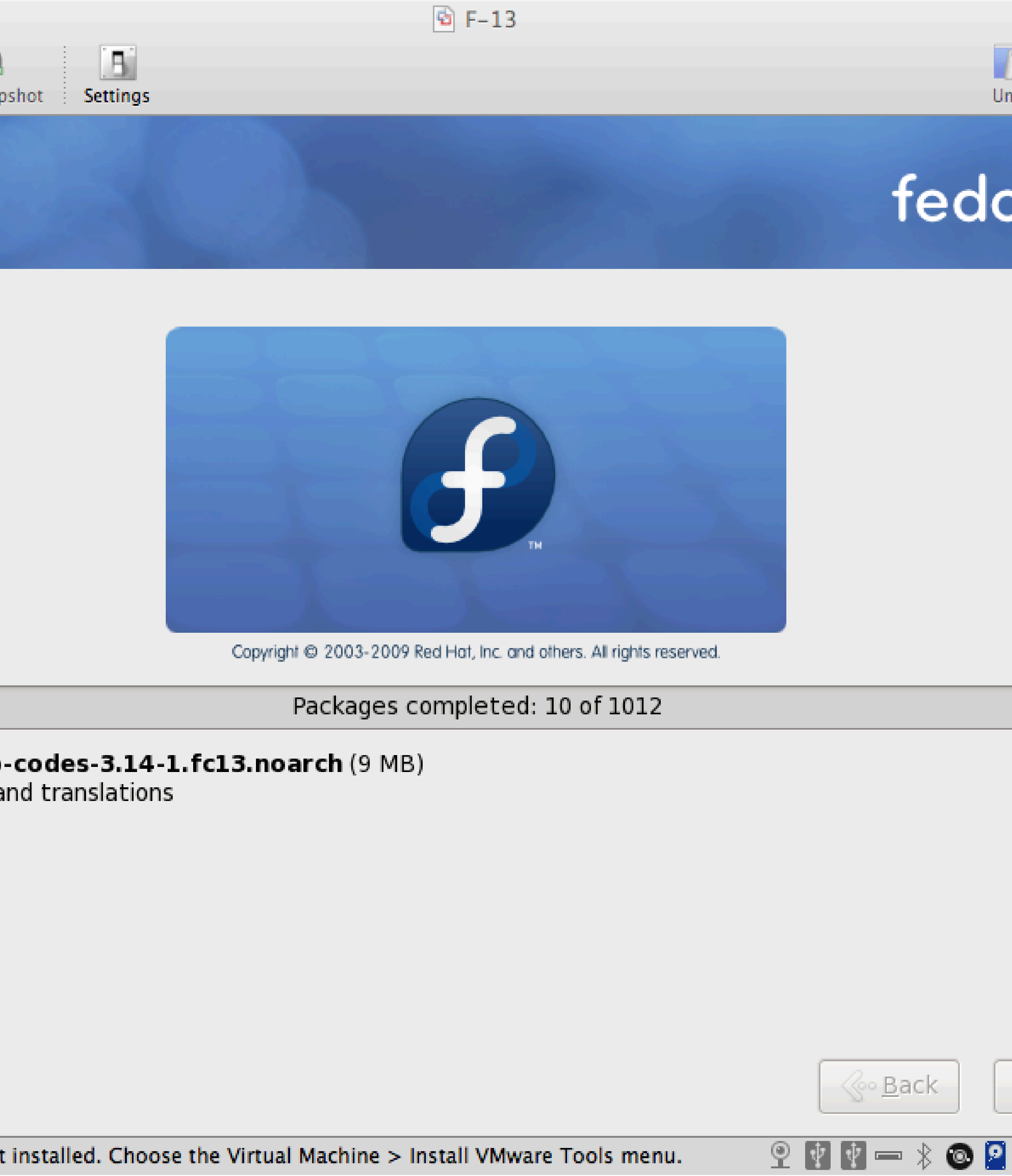
er customize the software selection now, or after install via the software
application.

e later ○ Customize now

t installed. Choose the Virtual Machine > Install VMware Tools menu.

Figura 2.8. Installazione di Fedora - Software

Go grab something to drink, this may take a while



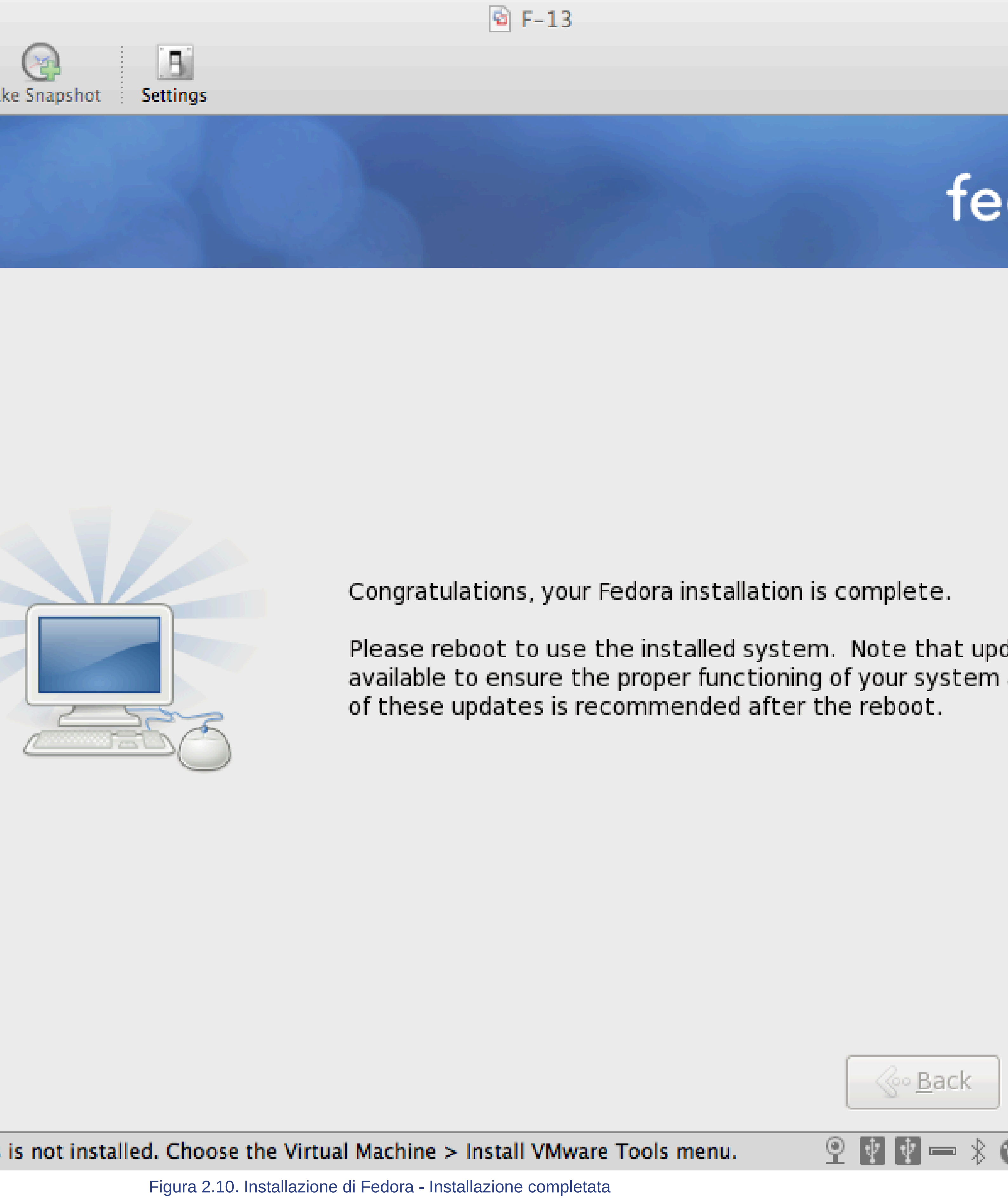
Copyright © 2003-2009 Red Hat, Inc. and others. All rights reserved.

Packages completed: 10 of 1012

-codes-3.14-1.fc13.noarch (9 MB)
and translations

t installed. Choose the Virtual Machine > Install VMware Tools menu.

Figura 2.9. Installazione di Fedora - Installazione



Congratulations, your Fedora installation is complete.

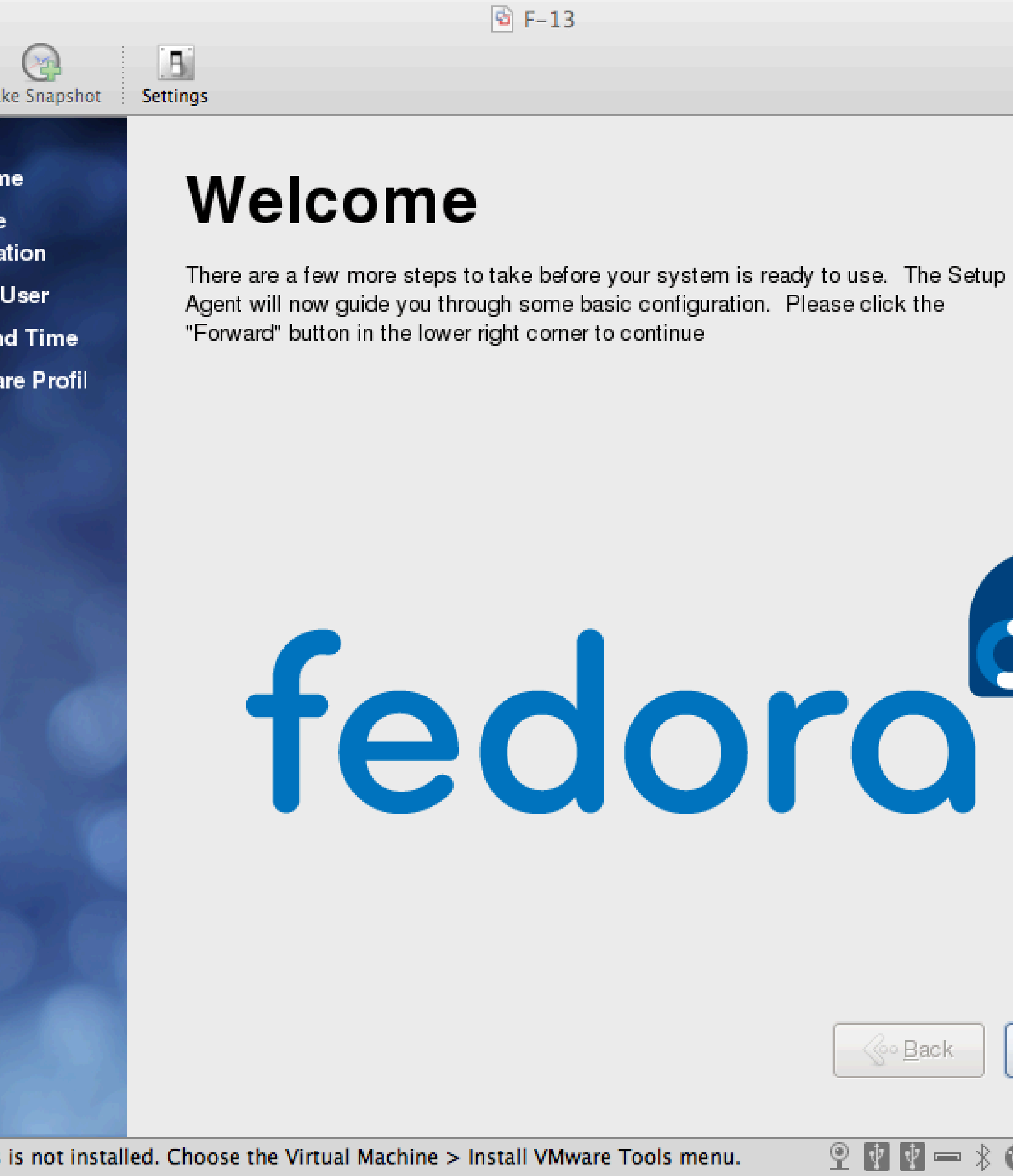
Please reboot to use the installed system. Note that upd
available to ensure the proper functioning of your system
of these updates is recommended after the reboot.

is not installed. Choose the Virtual Machine > Install VMware Tools menu.

Figura 2.10. Installazione di Fedora - Installazione completata

Once the node reboots, follow the on screen instructions [6] to create a system user and configure the time.

[6] *http://docs.fedoraproject.org/install-guide/f13/en-US/html/ch-firstboot.html*

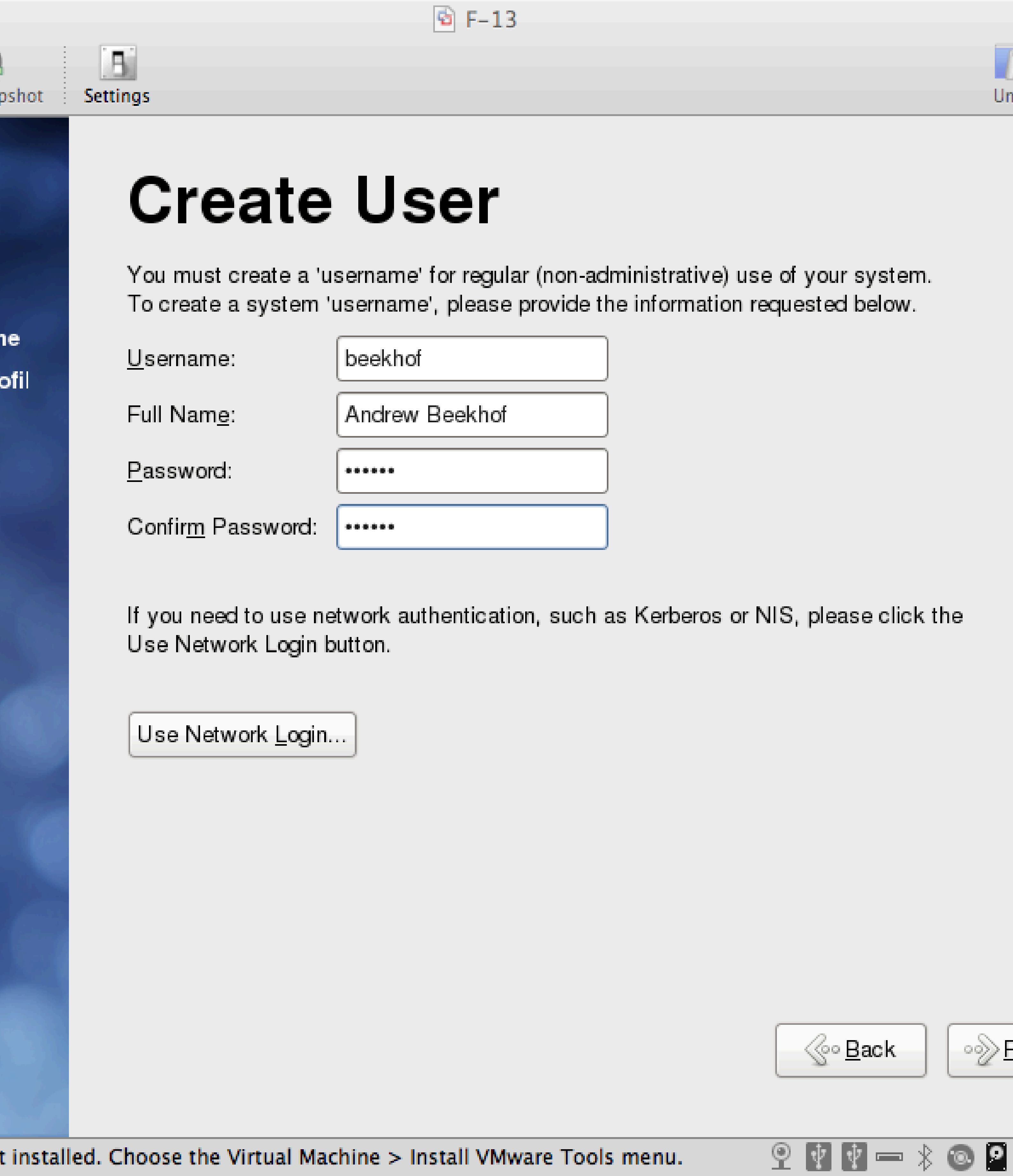


Welcome

There are a few more steps to take before your system is ready to use. The Setup Agent will now guide you through some basic configuration. Please click the "Forward" button in the lower right corner to continue

is not installed. Choose the Virtual Machine > Install VMware Tools menu.

Figura 2.11. Installazione di Fedora - Primo avvio



Create User

You must create a 'username' for regular (non-administrative) use of your system. To create a system 'username', please provide the information requested below.

Username: beekhof

Full Name: Andrew Beekhof

Password: ••••••

Confirm Password: ••••••

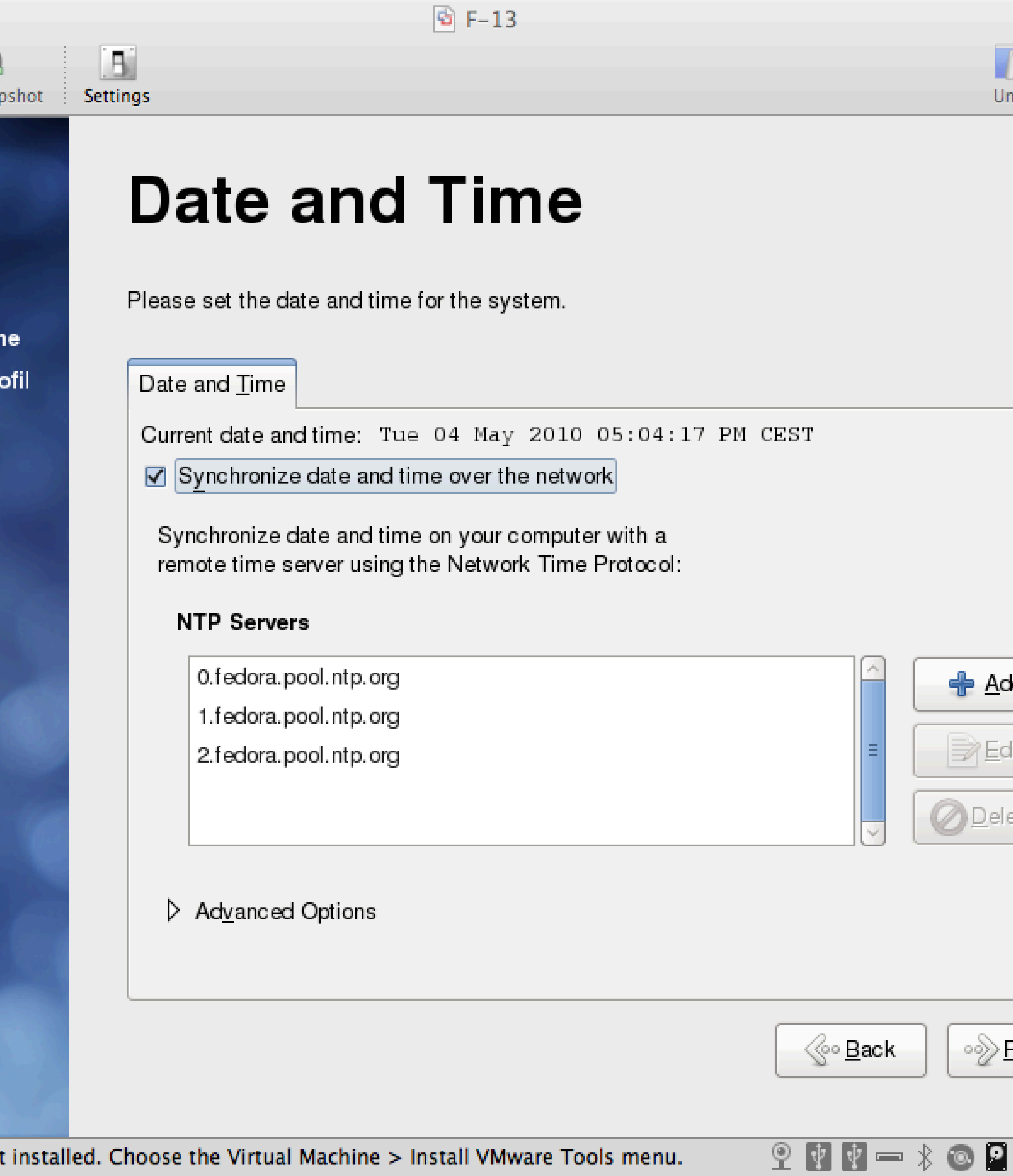
If you need to use network authentication, such as Kerberos or NIS, please click the Use Network Login button.

F-13
pshot
Settings
Un
ne
ofil
Use Network Login...
Back
F
t installed. Choose the Virtual Machine > Install VMware Tools menu.

Figura 2.12. Installazione di Fedora - Creazione utente non privilegiato

Nota

It is highly recommended to enable NTP on your cluster nodes. Doing so ensures all nodes agree on the current time and makes reading log files significantly easier. Fedora Installation - Date and TimeFedora Installation: Enable NTP to keep the times on all your nodes consistent



Date and Time

Please set the date and time for the system.

Current date and time: Tue 04 May 2010 05:04:17 PM CEST

☑ Synchronize date and time over the network

Synchronize date and time on your computer with a remote time server using the Network Time Protocol:

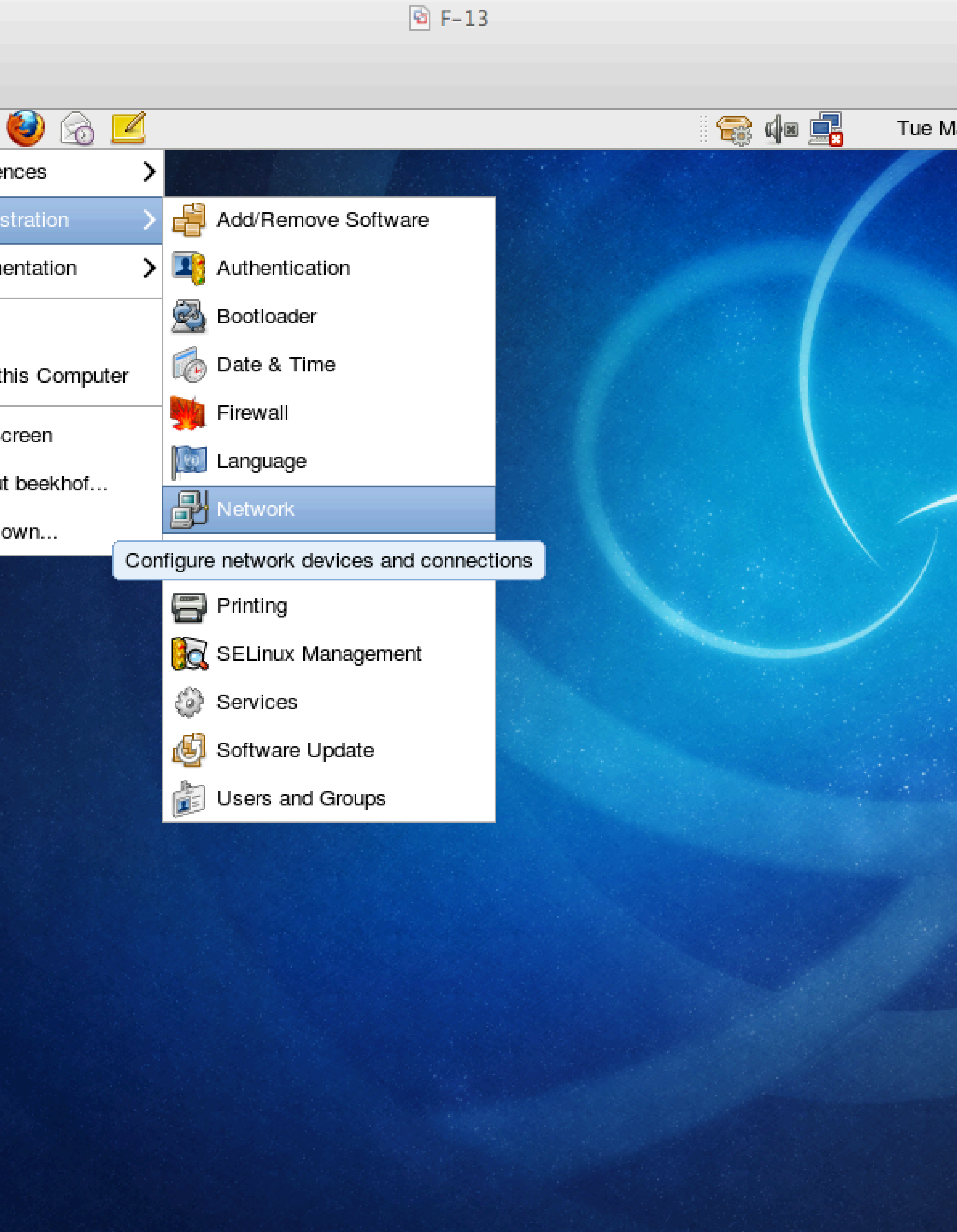
NTP Servers

0.fedora.pool.ntp.org
1.fedora.pool.ntp.org
2.fedora.pool.ntp.org

t installed. Choose the Virtual Machine > Install VMware Tools menu.

Figura 2.13. Installazione di Fedora - Data ed ora

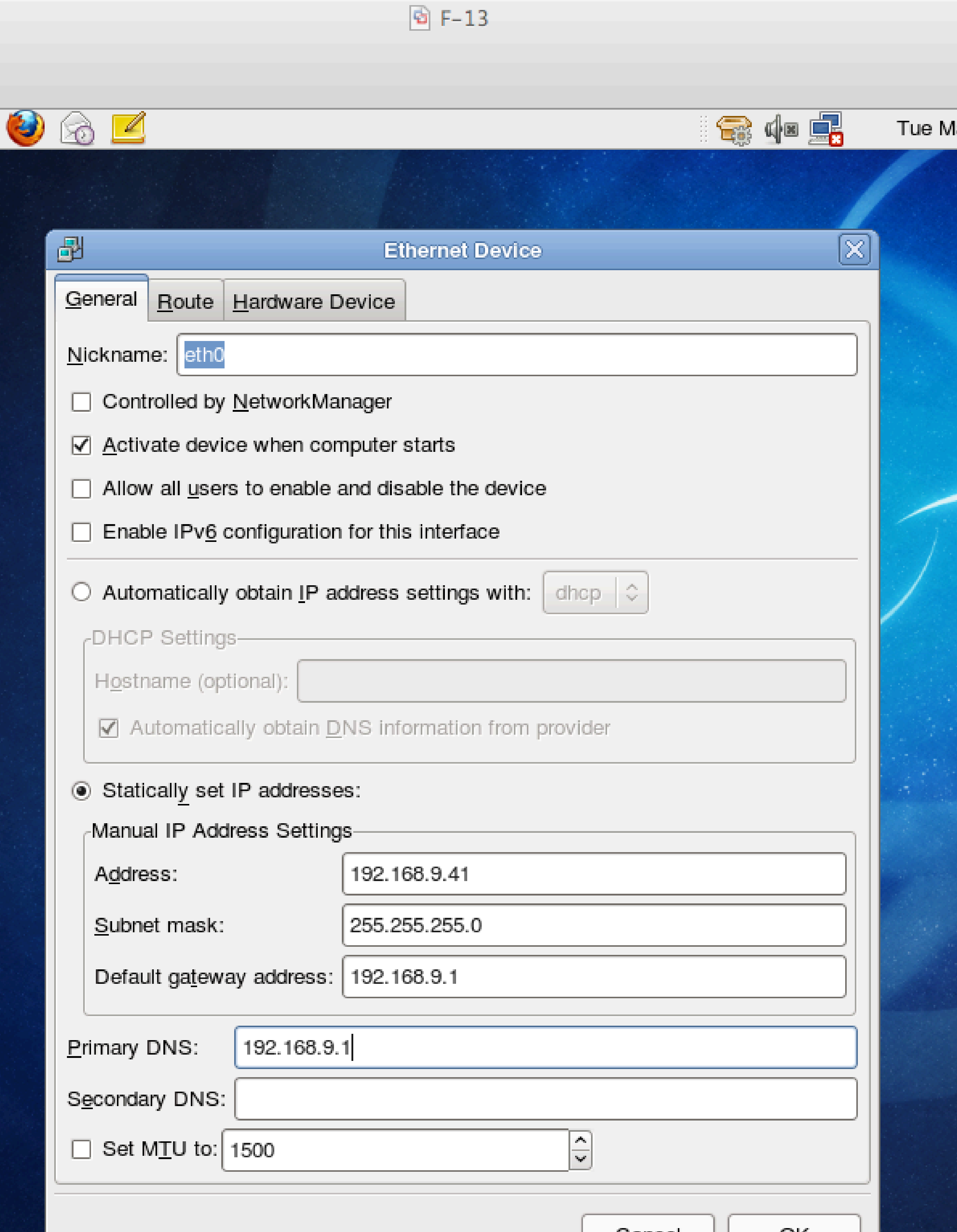
Cliccare su **`next`** fino alla comparsa della schermata di login. Cliccare sull'utente creato e fornire la password stabilita.



F-13
Tue M
ences
stration
entation
this Computer
creen
it beekhof...
own...
Add/Remove Software
Authentication
Bootloader
Date & Time
Firewall
Language
Network
Configure network devices and connections
Printing
SELinux Management
Services
Software Update
Users and Groups

Importante

Do not accept the default network settings. Cluster machines should never obtain an ip address via DHCP. Here I will use the internal addresses for the clusterlab.org network.



Nickname: eth0

□ Controlled by NetworkManager

☑ Activate device when computer starts

□ Allow all users to enable and disable the device

□ Enable IPv6 configuration for this interface

○ Automatically obtain IP address settings with: dhcp

DHCP Settings

Hostname (optional):

☑ Automatically obtain DNS information from provider

● Statically set IP addresses:

Manual IP Address Settings

Address: 192.168.9.41

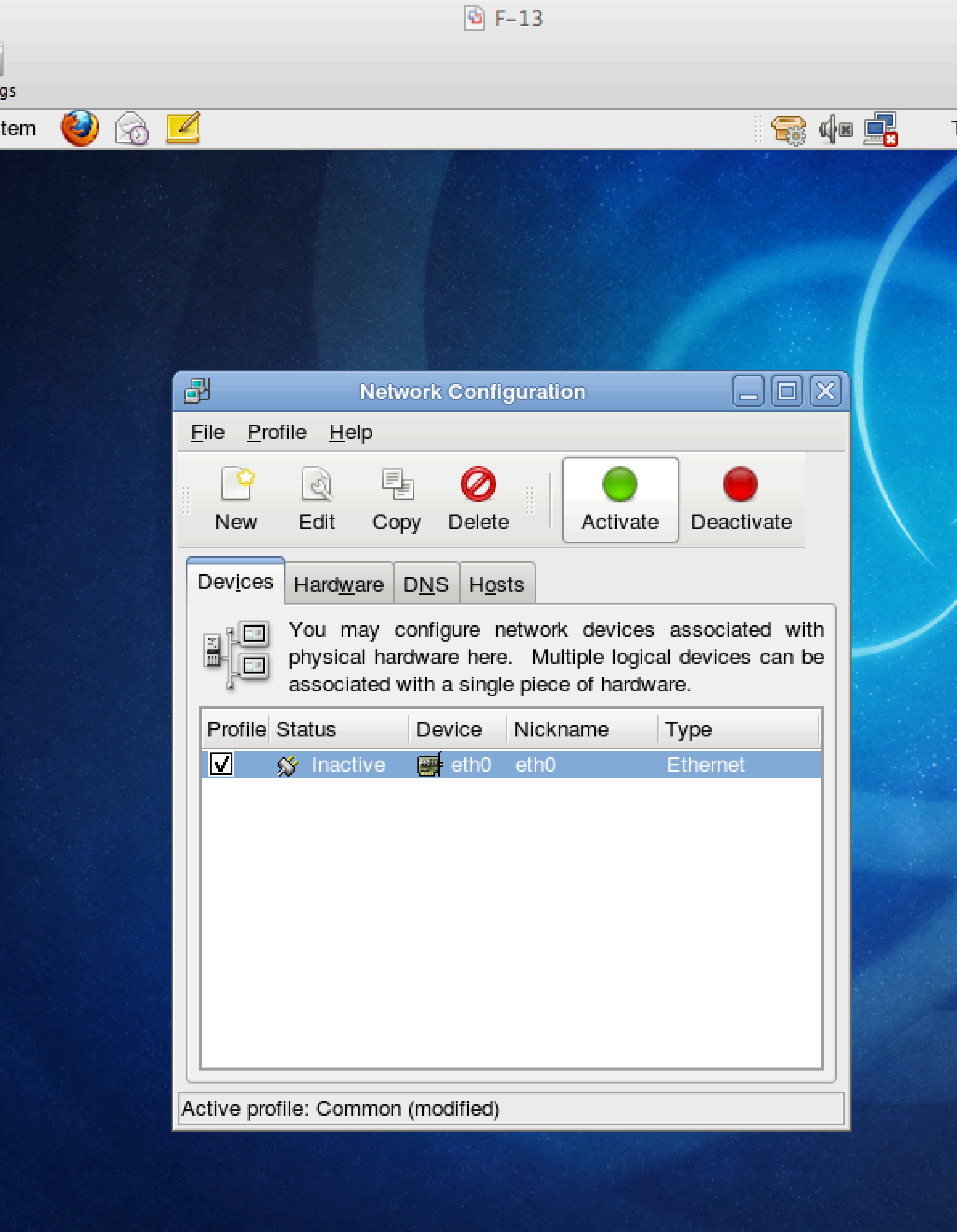
Subnet mask: 255.255.255.0

Default gateway address: 192.168.9.1

Primary DNS: 192.168.9.1

Secondary DNS:

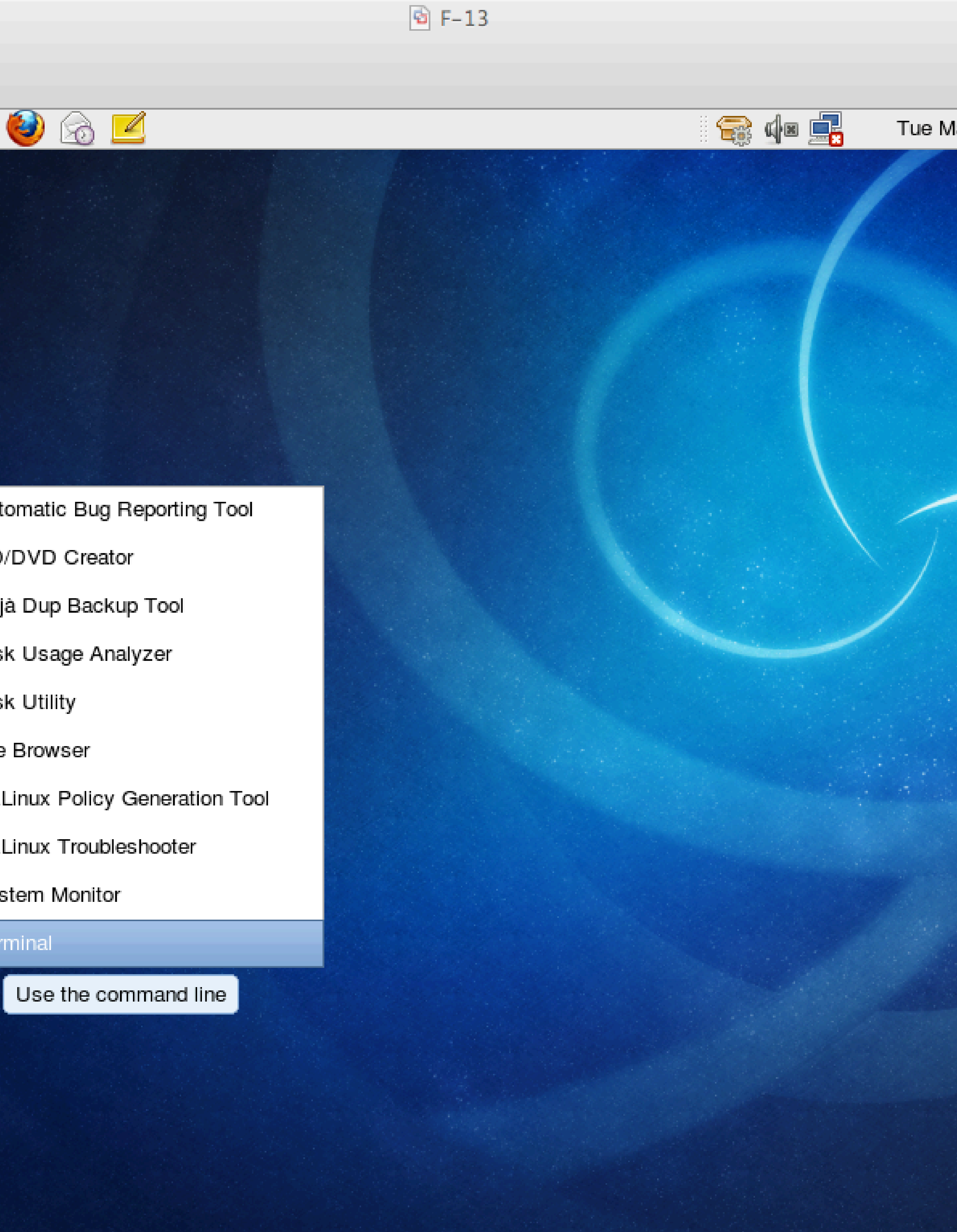
□ Set MTU to: 1500



You may configure network devices associated with physical hardware here. Multiple logical devices can be associated with a single piece of hardware.

Profile	Status	Device	Nickname	Type
☑	Inactive	eth0	eth0	Ethernet

Active profile: Common (modified)



F-13
Tue M
tomatic Bug Reporting Tool
)/DVD Creator
jà Dup Backup Tool
sk Usage Analyzer
sk Utility
e Browser
Linux Policy Generation Tool
Linux Troubleshooter
stem Monitor
rminal
Use the command line

Nota

Questa era l'ultima schermata, da qui in poi il lavoro verrà effettuato da terminale.

2.2. Installazione del software cluster

Go to the terminal window you just opened and switch to the super user (aka. "root") account with the su command. You will need to supply the password you entered earlier during the installation process.

```
[beekhof@pcmk-1 ~]$ su -
Password:
[root@pcmk-1 ~]#
```

Nota

Notare che lo username (il testo che precede il simbolo @) ora indica che l'utente utilizzato è "root".

```
# ip addr
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 16436 qdisc noqueue state UNKNOWN
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
    inet6 ::1/128 scope host
       valid_lft forever preferred_lft forever
2: eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UNKNOWN qlen 1000
    link/ether 00:0c:29:6f:e1:58 brd ff:ff:ff:ff:ff:ff
    inet 192.168.9.41/24 brd 192.168.9.255 scope global eth0
    inet6 ::20c:29ff:fe6f:e158/64 scope global dynamic
       valid_lft 2591667sec preferred_lft 604467sec
    inet6 2002:57ae:43fc:0:20c:29ff:fe6f:e158/64 scope global dynamic
       valid_lft 2591990sec preferred_lft 604790sec
    inet6 fe80::20c:29ff:fe6f:e158/64 scope link
       valid_lft forever preferred_lft forever
# ping -c 1 www.google.com
PING www.l.google.com (74.125.39.99) 56(84) bytes of data.
64 bytes from fx-in-f99.1e100.net (74.125.39.99): icmp_seq=1 ttl=56 time=16.7 ms

--- www.l.google.com ping statistics ---
1 packets transmitted, 1 received, 0% packet loss, time 20ms
rtt min/avg/max/mdev = 16.713/16.713/16.713/0.000 ms
# /sbin/chkconfig network on
#
```

2.2.1. Considerazioni sulla sicurezza

Per semplificare questa guida e focalizzare gli aspetti direttamente correlati al custering sarà necessario disabilitare il firewall della macchina e l'installazione di SELinux. Entrambe queste azioni creano limiti nella sicurezza e non dovrebbero essere applicate a macchine esposte al mondo.

Importante

```
TODO: Creare un Appendice che valuti (almeno) la riabilitazione del firewall.
```

```
# sed -i.bak "s/SELINUX=enforcing/SELINUX=permissive/g" /etc/selinux/config
# /sbin/chkconfig --del iptables
# service iptables stop
iptables: Flushing firewall rules:                         [  OK  ]
iptables: Setting chains to policy ACCEPT: filter          [  OK  ]
iptables: Unloading modules:                               [  OK  ]
```

Nota

Sarà necessario riavviare per rendere effettive le modifiche a SELinux. In alterniva qualcosa di simile a questo apparirà all'avvio di corosync:

```
May  4 19:30:54 pcmk-1 setroubleshoot: SELinux is preventing /usr/sbin/
corosync "getattr" access on /. For complete SELinux messages. run sealert -l
 6e0d4384-638e-4d55-9aaf-7dac011f29c1
May  4 19:30:54 pcmk-1 setroubleshoot: SELinux is preventing /usr/sbin/
corosync "getattr" access on /. For complete SELinux messages. run sealert -l
 6e0d4384-638e-4d55-9aaf-7dac011f29c1
```

2.2.2. Installazione del software cluster

Dalla versione 12, Fedora fornisce versioni recenti di tutto quanto si necessita, quindi avviando la shell lanciare:

```
# sed -i.bak "s/enabled=0/enabled=1/g"
/etc/yum.repos.d/fedora.repo
# sed -i.bak "s/enabled=0/enabled=1/g"
/etc/yum.repos.d/fedora-updates.repo
# yum install -y pacemaker corosync
Loaded plugins: presto, refresh-packagekit
fedora/metalink                                                    |  22 kB     00:00
fedora-debuginfo/metalink                                          |  16 kB     00:00
fedora-debuginfo                                                   | 3.2 kB     00:00
fedora-debuginfo/primary_db                                        | 1.4 MB     00:04
fedora-source/metalink                                             |  22 kB     00:00
fedora-source                                                      | 3.2 kB     00:00
fedora-source/primary_db                                           | 3.0 MB     00:05
updates/metalink                                                   |  26 kB     00:00
updates                                                            | 2.6 kB     00:00
updates/primary_db                                                 | 1.1 kB     00:00
updates-debuginfo/metalink                                         |  18 kB     00:00
updates-debuginfo                                                  | 2.6 kB     00:00
updates-debuginfo/primary_db                                       | 1.1 kB     00:00
updates-source/metalink                                            |  25 kB     00:00
updates-source                                                     | 2.6 kB     00:00
updates-source/primary_db                                          | 1.1 kB     00:00
Setting up Install Process
```

```
Resolving Dependencies
--> Running transaction check
---> Package corosync.x86_64 0:1.2.1-1.fc13 set to be updated
--> Processing Dependency: corosynclib = 1.2.1-1.fc13 for package:
 corosync-1.2.1-1.fc13.x86_64
--> Processing Dependency: libquorum.so.4(COROSYNC_QUORUM_1.0)(64bit) for package:
 corosync-1.2.1-1.fc13.x86_64
--> Processing Dependency: libvotequorum.so.4(COROSYNC_VOTEQUORUM_1.0)(64bit) for package:
 corosync-1.2.1-1.fc13.x86_64
--> Processing Dependency: libcpg.so.4(COROSYNC_CPG_1.0)(64bit) for package:
 corosync-1.2.1-1.fc13.x86_64
--> Processing Dependency: libconfdb.so.4(COROSYNC_CONFDB_1.0)(64bit) for package:
 corosync-1.2.1-1.fc13.x86_64
--> Processing Dependency: libcfg.so.4(COROSYNC_CFG_0.82)(64bit) for package:
 corosync-1.2.1-1.fc13.x86_64
--> Processing Dependency: libpload.so.4(COROSYNC_PLOAD_1.0)(64bit) for package:
 corosync-1.2.1-1.fc13.x86_64
--> Processing Dependency: liblogsys.so.4()(64bit) for package: corosync-1.2.1-1.fc13.x86_64
--> Processing Dependency: libconfdb.so.4()(64bit) for package: corosync-1.2.1-1.fc13.x86_64
--> Processing Dependency: libcoroipcc.so.4()(64bit) for package:
 corosync-1.2.1-1.fc13.x86_64
--> Processing Dependency: libcpg.so.4()(64bit) for package: corosync-1.2.1-1.fc13.x86_64
--> Processing Dependency: libquorum.so.4()(64bit) for package: corosync-1.2.1-1.fc13.x86_64
--> Processing Dependency: libcoroipcs.so.4()(64bit) for package:
 corosync-1.2.1-1.fc13.x86_64
--> Processing Dependency: libvotequorum.so.4()(64bit) for package:
 corosync-1.2.1-1.fc13.x86_64
--> Processing Dependency: libcfg.so.4()(64bit) for package: corosync-1.2.1-1.fc13.x86_64
--> Processing Dependency: libtotem_pg.so.4()(64bit) for package:
 corosync-1.2.1-1.fc13.x86_64
--> Processing Dependency: libpload.so.4()(64bit) for package: corosync-1.2.1-1.fc13.x86_64
---> Package pacemaker.x86_64 0:1.1.5-1.fc13 set to be updated
--> Processing Dependency: heartbeat >= 3.0.0 for package: pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: net-snmp >= 5.4 for package: pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: resource-agents for package: pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: cluster-glue for package: pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: libnetsnmp.so.20()(64bit) for package:
 pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: libcrmcluster.so.1()(64bit) for package:
 pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: libpengine.so.3()(64bit) for package:
 pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: libnetsnmpagent.so.20()(64bit) for package:
 pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: libesmtp.so.5()(64bit) for package: pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: libstonithd.so.1()(64bit) for package:
 pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: libhbclient.so.1()(64bit) for package:
 pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: libpils.so.2()(64bit) for package: pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: libpe_status.so.2()(64bit) for package:
 pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: libnetsnmpmibs.so.20()(64bit) for package:
 pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: libnetsnmphelpers.so.20()(64bit) for package:
 pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: libcib.so.1()(64bit) for package: pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: libccmclient.so.1()(64bit) for package:
 pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: libstonith.so.1()(64bit) for package:
 pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: liblrm.so.2()(64bit) for package: pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: libtransitioner.so.1()(64bit) for package:
 pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: libpe_rules.so.2()(64bit) for package:
 pacemaker-1.1.5-1.fc13.x86_64
```

```
--> Processing Dependency: libcrmcommon.so.2()(64bit) for package:
 pacemaker-1.1.5-1.fc13.x86_64
--> Processing Dependency: libplumb.so.2()(64bit) for package: pacemaker-1.1.5-1.fc13.x86_64
--> Running transaction check
---> Package cluster-glue.x86_64 0:1.0.2-1.fc13 set to be updated
--> Processing Dependency: perl-TimeDate for package: cluster-glue-1.0.2-1.fc13.x86_64
--> Processing Dependency: libOpenIPMIutils.so.0()(64bit) for package: cluster-
glue-1.0.2-1.fc13.x86_64
--> Processing Dependency: libOpenIPMIposix.so.0()(64bit) for package: cluster-
glue-1.0.2-1.fc13.x86_64
--> Processing Dependency: libopenhpi.so.2()(64bit) for package: cluster-
glue-1.0.2-1.fc13.x86_64
--> Processing Dependency: libOpenIPMI.so.0()(64bit) for package: cluster-
glue-1.0.2-1.fc13.x86_64
---> Package cluster-glue-libs.x86_64 0:1.0.2-1.fc13 set to be updated
---> Package corosynclib.x86_64 0:1.2.1-1.fc13 set to be updated
--> Processing Dependency: librdmacm.so.1(RDMACM_1.0)(64bit) for package:
 corosynclib-1.2.1-1.fc13.x86_64
--> Processing Dependency: libibverbs.so.1(IBVERBS_1.0)(64bit) for package:
 corosynclib-1.2.1-1.fc13.x86_64
--> Processing Dependency: libibverbs.so.1(IBVERBS_1.1)(64bit) for package:
 corosynclib-1.2.1-1.fc13.x86_64
--> Processing Dependency: libibverbs.so.1()(64bit) for package:
 corosynclib-1.2.1-1.fc13.x86_64
--> Processing Dependency: librdmacm.so.1()(64bit) for package:
 corosynclib-1.2.1-1.fc13.x86_64
---> Package heartbeat.x86_64 0:3.0.0-0.7.0daab7da36a8.hg.fc13 set to be updated
--> Processing Dependency: PyXML for package: heartbeat-3.0.0-0.7.0daab7da36a8.hg.fc13.x86_64
---> Package heartbeat-libs.x86_64 0:3.0.0-0.7.0daab7da36a8.hg.fc13 set to be updated
---> Package libesmtp.x86_64 0:1.0.4-12.fc12 set to be updated
---> Package net-snmp.x86_64 1:5.5-12.fc13 set to be updated
--> Processing Dependency: libsensors.so.4()(64bit) for package: 1:net-
snmp-5.5-12.fc13.x86_64
---> Package net-snmp-libs.x86_64 1:5.5-12.fc13 set to be updated
---> Package pacemaker-libs.x86_64 0:1.1.5-1.fc13 set to be updated
---> Package resource-agents.x86_64 0:3.0.10-1.fc13 set to be updated
--> Processing Dependency: libnet.so.1()(64bit) for package: resource-
agents-3.0.10-1.fc13.x86_64
--> Running transaction check
---> Package OpenIPMI-libs.x86_64 0:2.0.16-8.fc13 set to be updated
---> Package PyXML.x86_64 0:0.8.4-17.fc13 set to be updated
---> Package libibverbs.x86_64 0:1.1.3-4.fc13 set to be updated
--> Processing Dependency: libibverbs-driver for package: libibverbs-1.1.3-4.fc13.x86_64
---> Package libnet.x86_64 0:1.1.4-3.fc12 set to be updated
---> Package librdmacm.x86_64 0:1.0.10-2.fc13 set to be updated
---> Package lm_sensors-libs.x86_64 0:3.1.2-2.fc13 set to be updated
---> Package openhpi-libs.x86_64 0:2.14.1-3.fc13 set to be updated
---> Package perl-TimeDate.noarch 1:1.20-1.fc13 set to be updated
--> Running transaction check
---> Package libmlx4.x86_64 0:1.0.1-5.fc13 set to be updated
--> Finished Dependency Resolution

Dependencies Resolved


==========================================================================================
 Package               Arch     Version                             Repository      Size
==========================================================================================
Installing:
 corosync              x86_64   1.2.1-1.fc13                        fedora         136 k
 pacemaker             x86_64   1.1.5-1.fc13                        fedora         543 k
Installing for dependencies:
 OpenIPMI-libs         x86_64   2.0.16-8.fc13                       fedora         474 k
 PyXML                 x86_64   0.8.4-17.fc13                       fedora         906 k
 cluster-glue          x86_64   1.0.2-1.fc13                        fedora         230 k
 cluster-glue-libs     x86_64   1.0.2-1.fc13                        fedora         116 k
 corosynclib           x86_64   1.2.1-1.fc13                        fedora         145 k
 heartbeat              x86_64   3.0.0-0.7.0daab7da36a8.hg.fc13      updates        172 k
 heartbeat-libs         x86_64   3.0.0-0.7.0daab7da36a8.hg.fc13      updates        265 k
 libesmtp               x86_64   1.0.4-12.fc12                       fedora          54 k
 libibverbs             x86_64   1.1.3-4.fc13                        fedora          42 k
 libmlx4                x86_64   1.0.1-5.fc13                        fedora          27 k
 libnet                 x86_64   1.1.4-3.fc12                        fedora          49 k
 librdmacm              x86_64   1.0.10-2.fc13                       fedora          22 k
 lm_sensors-libs        x86_64   3.1.2-2.fc13                        fedora          37 k
 net-snmp               x86_64   1:5.5-12.fc13                       fedora         295 k
 net-snmp-libs          x86_64   1:5.5-12.fc13                       fedora         1.5 M
 openhpi-libs           x86_64   2.14.1-3.fc13                       fedora         135 k
 pacemaker-libs         x86_64   1.1.5-1.fc13                        fedora         264 k
 perl-TimeDate          noarch   1:1.20-1.fc13                       fedora          42 k
 resource-agents        x86_64   3.0.10-1.fc13                       fedora         357 k

Transaction Summary
=========================================================================================
Install      21 Package(s)
Upgrade       0 Package(s)

Total download size: 5.7 M
Installed size: 20 M
Downloading Packages:
Setting up and reading Presto delta metadata
updates-testing/prestodelta                                           | 164 kB     00:00
fedora/prestodelta                                                    |  150 B     00:00
Processing delta metadata
Package(s) data still to download: 5.7 M
(1/21): OpenIPMI-libs-2.0.16-8.fc13.x86_64.rpm                        | 474 kB     00:00
(2/21): PyXML-0.8.4-17.fc13.x86_64.rpm                                | 906 kB     00:01
(3/21): cluster-glue-1.0.2-1.fc13.x86_64.rpm                          | 230 kB     00:00
(4/21): cluster-glue-libs-1.0.2-1.fc13.x86_64.rpm                     | 116 kB     00:00
(5/21): corosync-1.2.1-1.fc13.x86_64.rpm                              | 136 kB     00:00
(6/21): corosynclib-1.2.1-1.fc13.x86_64.rpm                           | 145 kB     00:00
(7/21): heartbeat-3.0.0-0.7.0daab7da36a8.hg.fc13.x86_64.rpm          | 172 kB     00:00
(8/21): heartbeat-libs-3.0.0-0.7.0daab7da36a8.hg.fc13.x86_64.rpm     | 265 kB     00:00
(9/21): libesmtp-1.0.4-12.fc12.x86_64.rpm                             |  54 kB     00:00
(10/21): libibverbs-1.1.3-4.fc13.x86_64.rpm                           |  42 kB     00:00
(11/21): libmlx4-1.0.1-5.fc13.x86_64.rpm                              |  27 kB     00:00
(12/21): libnet-1.1.4-3.fc12.x86_64.rpm                               |  49 kB     00:00
(13/21): librdmacm-1.0.10-2.fc13.x86_64.rpm                           |  22 kB     00:00
(14/21): lm_sensors-libs-3.1.2-2.fc13.x86_64.rpm                      |  37 kB     00:00
(15/21): net-snmp-5.5-12.fc13.x86_64.rpm                              | 295 kB     00:00
(16/21): net-snmp-libs-5.5-12.fc13.x86_64.rpm                         | 1.5 MB     00:01
(17/21): openhpi-libs-2.14.1-3.fc13.x86_64.rpm                        | 135 kB     00:00
(18/21): pacemaker-1.1.5-1.fc13.x86_64.rpm                            | 543 kB     00:00
(19/21): pacemaker-libs-1.1.5-1.fc13.x86_64.rpm                       | 264 kB     00:00
(20/21): perl-TimeDate-1.20-1.fc13.noarch.rpm                         |  42 kB     00:00
(21/21): resource-agents-3.0.10-1.fc13.x86_64.rpm                     | 357 kB     00:00

Total                                                        539 kB/s | 5.7 MB     00:10
warning: rpmts_HdrFromFdno: Header V3 RSA/SHA256 Signature, key ID e8e40fde: NOKEY
fedora/gpgkey                                                         | 3.2 kB     00:00 ...
Importing GPG key 0xE8E40FDE "Fedora (13) <fedora@fedoraproject.org%gt;" from /etc/pki/rpm-
gpg/RPM-GPG-KEY-fedora-x86_64

Running rpm_check_debug
Running Transaction Test
Transaction Test Succeeded
Running Transaction
  Installing     : lm_sensors-libs-3.1.2-2.fc13.x86_64                           1/21
  Installing     : 1:net-snmp-libs-5.5-12.fc13.x86_64                            2/21
  Installing     : 1:net-snmp-5.5-12.fc13.x86_64                                 3/21
  Installing     : openhpi-libs-2.14.1-3.fc13.x86_64                             4/21
  Installing     : libibverbs-1.1.3-4.fc13.x86_64                                5/21
  Installing     : libmlx4-1.0.1-5.fc13.x86_64                                   6/21
  Installing     : librdmacm-1.0.10-2.fc13.x86_64                                7/21
  Installing     : corosync-1.2.1-1.fc13.x86_64                                   8/21
  Installing     : corosynclib-1.2.1-1.fc13.x86_64                                9/21
  Installing     : libesmtp-1.0.4-12.fc12.x86_64                                 10/21
  Installing     : OpenIPMI-libs-2.0.16-8.fc13.x86_64                            11/21
  Installing     : PyXML-0.8.4-17.fc13.x86_64                                    12/21
  Installing     : libnet-1.1.4-3.fc12.x86_64                                    13/21
  Installing     : 1:perl-TimeDate-1.20-1.fc13.noarch                            14/21
  Installing     : cluster-glue-1.0.2-1.fc13.x86_64                              15/21
  Installing     : cluster-glue-libs-1.0.2-1.fc13.x86_64                         16/21
  Installing     : resource-agents-3.0.10-1.fc13.x86_64                          17/21
  Installing     : heartbeat-libs-3.0.0-0.7.0daab7da36a8.hg.fc13.x86_64          18/21
  Installing     : heartbeat-3.0.0-0.7.0daab7da36a8.hg.fc13.x86_64               19/21
  Installing     : pacemaker-1.1.5-1.fc13.x86_64                                 20/21
  Installing     : pacemaker-libs-1.1.5-1.fc13.x86_64                            21/21

Installed:
  corosync.x86_64 0:1.2.1-1.fc13                    pacemaker.x86_64 0:1.1.5-1.fc13

Dependency Installed:
  OpenIPMI-libs.x86_64 0:2.0.16-8.fc13
  PyXML.x86_64 0:0.8.4-17.fc13
  cluster-glue.x86_64 0:1.0.2-1.fc13
  cluster-glue-libs.x86_64 0:1.0.2-1.fc13
  corosynclib.x86_64 0:1.2.1-1.fc13
  heartbeat.x86_64 0:3.0.0-0.7.0daab7da36a8.hg.fc13
  heartbeat-libs.x86_64 0:3.0.0-0.7.0daab7da36a8.hg.fc13
  libesmtp.x86_64 0:1.0.4-12.fc12
  libibverbs.x86_64 0:1.1.3-4.fc13
  libmlx4.x86_64 0:1.0.1-5.fc13
  libnet.x86_64 0:1.1.4-3.fc12
  librdmacm.x86_64 0:1.0.10-2.fc13
  lm_sensors-libs.x86_64 0:3.1.2-2.fc13
  net-snmp.x86_64 1:5.5-12.fc13
  net-snmp-libs.x86_64 1:5.5-12.fc13
  openhpi-libs.x86_64 0:2.14.1-3.fc13
  pacemaker-libs.x86_64 0:1.1.5-1.fc13
  perl-TimeDate.noarch 1:1.20-1.fc13
  resource-agents.x86_64 0:3.0.10-1.fc13

Complete!
#
```

2.3. Prima di continuare

Ripetere i passi dell'installazione così da avere due nodi Fedora con il software cluster installato.

For the purposes of this document, the additional node is called pcmk-2 with address 192.168.122.102.

2.4. Setup

2.4.1. Definire la rete

Verificare la reciproca comunicazione dei due nodi:

```
# ping -c 3 192.168.122.102
PING 192.168.122.102 (192.168.122.102) 56(84) bytes of data.
64 bytes from 192.168.122.102: icmp_seq=1 ttl=64 time=0.343 ms
64 bytes from 192.168.122.102: icmp_seq=2 ttl=64 time=0.402 ms
64 bytes from 192.168.122.102: icmp_seq=3 ttl=64 time=0.558 ms
```

```
--- 192.168.122.102 ping statistics ---3 packets transmitted, 3 received, 0% packet loss,
 time 2000ms
rtt min/avg/max/mdev = 0.343/0.434/0.558/0.092 ms
```

Figure 2.18. Verify Connectivity by IP address

Now we need to make sure we can communicate with the machines by their name. If you have a DNS server, add additional entries for the two machines. Otherwise, you'll need to add the machines to /etc/hosts . Below are the entries for my cluster nodes:

```
# grep pcmk /etc/hosts
192.168.122.101 pcmk-1.clusterlabs.org pcmk-1
192.168.122.102 pcmk-2.clusterlabs.org pcmk-2
```

Figure 2.19. Set up /etc/hosts entries

Ora è possibile verificare il setup usando nuovamente ping:

```
# ping -c 3 pcmk-2
PING pcmk-2.clusterlabs.org (192.168.122.101) 56(84) bytes of data.
64 bytes from pcmk-1.clusterlabs.org (192.168.122.101): icmp_seq=1 ttl=64 time=0.164 ms
64 bytes from pcmk-1.clusterlabs.org (192.168.122.101): icmp_seq=2 ttl=64 time=0.475 ms
64 bytes from pcmk-1.clusterlabs.org (192.168.122.101): icmp_seq=3 ttl=64 time=0.186 ms

--- pcmk-2.clusterlabs.org ping statistics ---3 packets transmitted, 3 received, 0% packet
 loss, time 2001ms
rtt min/avg/max/mdev = 0.164/0.275/0.475/0.141 ms
```

Figure 2.20. Verify Connectivity by Hostname

2.4.2. Configuarazione di SSH

SSH rappresenta una via conveniente e sicura per copiare file e lanciare comandi remotamente. Per raggiungere l'obiettivo di questa guida, verrà creata una chiave senza password (utilizzando l'opzione -N) in modo da poter eseguire comandi remoti senza che venga mostrato il prompt.

Avvertimento

Chiavi SSH non protette, senza una password, non sono raccomandate per server esposti al mondo.

Creazione di una nuova chiave ed abilitazione al proprietario della chiave di effettuare login:

Creazione ed attivazione di una nuova chiave SSH

```
# ssh-keygen -t dsa -f ~/.ssh/id_dsa -N ""
Generating public/private dsa key pair.
Your identification has been saved in /root/.ssh/id_dsa.
Your public key has been saved in /root/.ssh/id_dsa.pub.
The key fingerprint is:
91:09:5c:82:5a:6a:50:08:4e:b2:0c:62:de:cc:74:44 root@pcmk-1.clusterlabs.org
```

```
The key's randomart image is:
+--[ DSA 1024]----+
|==.ooEo..        |
|X O + .o o       |
| * A    +        |
|  +      .       |
| .      S        |
|                 |
|                 |
|                 |
|                 |
+-----------------+

# cp .ssh/id_dsa.pub .ssh/authorized_keys
```

Installare la chiave sull'altro nodo e verificare come sia possibile lanciare comandi remotamente, senza ricevere prompt

```
# scp -r .ssh pcmk-2:
The authenticity of host 'pcmk-2 (192.168.122.102)' can't be established.
RSA key fingerprint is b1:2b:55:93:f1:d9:52:2b:0f:f2:8a:4e:ae:c6:7c:9a.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'pcmk-2,192.168.122.102' (RSA) to the list of known
 hosts.root@pcmk-2's password:
id_dsa.pub                           100%  616     0.6KB/s   00:00
id_dsa                               100%  672     0.7KB/s   00:00
known_hosts                          100%  400     0.4KB/s   00:00
authorized_keys                      100%  616     0.6KB/s   00:00
# ssh pcmk-2 -- uname -npcmk-2
#
```

Figure 2.22. Installing the SSH Key on Another Host

2.4.3. Nomi di nodi abbreviati

During installation, we filled in the machine’s fully qualifier domain name (FQDN) which can be rather long when it appears in cluster logs and status output. See for yourself how the machine identifies itself:

```
# uname -n
pcmk-1.clusterlabs.org
# dnsdomainname clusterlabs.org
```

L'output del secondo comando è corretto, ma non è necessario il nome del dominio incluso nei dettagli essenziali dell'host. Per risolvere la situazione è necessario modificare il file /etc/sysconfig/netwirk. Ecco come dovrà apparire.

```
# cat /etc/sysconfig/network
NETWORKING=yes
HOSTNAME=pcmk-1.clusterlabs.org
GATEWAY=192.168.122.1
```

L'unica cosa da fare sarà di rimuovere la parte relativa al nome del dominio, che rimarrà comunque registrata altrove.

```
# sed -i.bak 's/\.[a-z].*//g' /etc/sysconfig/network
```

Per verificare l'efficacia dei cambiamenti apportati il contenuto del file dovrà essere simile a quanto riportato.

```
# cat /etc/sysconfig/network
NETWORKING=yes
HOSTNAME=pcmk-1
GATEWAY=192.168.122.1
```

Ma non è ancora finita. La macchina non sarà allineata con i nomi abbreviati finché non verrà riavviata, ma è possibile forzare l'aggiornamento.

```
# source /etc/sysconfig/network
# hostname $HOSTNAME
```

E' possibile quindi verificare che la macchina utilizzi il nome corretto

```
# uname -npcmk-1
# dnsdomainname clusterlabs.org
```

Le stesse operazioni andranno ripetute su pcmk-2

2.4.4. Configurare Corosync

Choose a port number and multi-cast [7] address. [8] Be sure that the values you chose do not conflict with any existing clusters you might have. For advice on choosing a multi-cast address, see *http://www.29west.com/docs/THPM/multicast-address-assignment.html* For this document, I have chosen port 4000 and used 226.94.1.1 as the multi-cast address.

Importante

The instructions below only apply for a machine with a single NIC. If you have a more complicated setup, you should edit the configuration manually.

```
# export ais_port=4000
# export ais_mcast=226.94.1.1
```

A questo punto è possibile determinare automaticamente l'indirizzo degli host. Non utilizzando l'indirizzo completo la configurazione è copiabile sugli altri nodi.

```
# export ais_addr=`ip addr | grep "inet " | tail -n 1 | awk '{print $4}' | sed s/255/0/`
```

Visualizzazione e verifica delle opzioni di configurazione

[7] *http://en.wikipedia.org/wiki/Multicast*

[8] *http://en.wikipedia.org/wiki/Multicast_address*

```
# env | grep ais_ais_mcast=226.94.1.1
ais_port=4000
ais_addr=192.168.122.0
```

Una volta soddisfatti dei valori scelti si potrà aggiornare la configurazione di Corosync

```
# cp /etc/corosync/corosync.conf.example /etc/corosync/corosync.conf
# sed -i.bak "s/.*mcastaddr:.*/mcastaddr:\ $ais_mcast/g" /etc/corosync/corosync.conf
# sed -i.bak "s/.*mcastport:.*/mcastport:\ $ais_port/g" /etc/corosync/corosync.conf
# sed -i.bak "s/.*bindnetaddr:.*/bindnetaddr:\ $ais_addr/g" /etc/corosync/corosync.conf
```

Finally, tell Corosync to load the Pacemaker plugin.

```
# cat <<-END >>/etc/corosync/service.d/pcmk
service {
        # Load the Pacemaker Cluster Resource Manager
        name: pacemaker
        ver:  1
}
END
```

The final configuration should look something like the sample in Appendix B, Sample Corosync Configuration.

Importante

When run in version 1 mode, the plugin does not start the Pacemaker daemons. Instead it just sets up the quorum and messaging interfaces needed by the rest of the stack. Starting the dameons occurs when the Pacemaker init script is invoked. This resolves two long standing issues:

a. Forking inside a multi-threaded process like Corosync causes all sorts of pain. This has been problematic for Pacemaker as it needs a number of daemons to be spawned.

b. Corosync was never designed for staggered shutdown - something previously needed in order to prevent the cluster from leaving before Pacemaker could stop all active resources.

2.4.5. Distribuire la configurazione

Ora è necessario copiare le modifiche effettuate sinora all'altro nodo:

```
# for f in /etc/corosync/corosync.conf /etc/corosync/service.d/pcmk /etc/hosts; do scp $f
 pcmk-2:$f ; done
corosync.conf                             100% 1528     1.5KB/s   00:00
hosts                                     100%  281     0.3KB/s   00:00
#
```

Verifica dell'installazione del cluster

Indice

3.1. Verifica dell'installazione di Corosync

Avvio di Corosync sul primo nodo

```
# /etc/init.d/corosync start
Starting Corosync Cluster Engine (corosync): [ OK ]
```

Controllo del corretto avvio del cluster e che è stata formata la membership iniziale

```
# grep -e "corosync.*network interface" -e "Corosync Cluster Engine" -e "Successfully read
 main configuration file" /var/log/messages
Aug 27 09:05:34 pcmk-1 corosync[1540]: [MAIN ] Corosync Cluster Engine ('1.1.0'): started and
 ready to provide service.
Aug 27 09:05:34 pcmk-1 corosync[1540]: [MAIN ] Successfully read main configuration file '/
etc/corosync/corosync.conf'.
# grep TOTEM /var/log/messages
Aug 27 09:05:34 pcmk-1 corosync[1540]: [TOTEM ] Initializing transport (UDP/IP).
Aug 27 09:05:34 pcmk-1 corosync[1540]: [TOTEM ] Initializing transmit/receive security:
 libtomcrypt SOBER128/SHA1HMAC (mode 0).
Aug 27 09:05:35 pcmk-1 corosync[1540]: [TOTEM ] The network interface [192.168.122.101] is
 now up.
Aug 27 09:05:35 pcmk-1 corosync[1540]: [TOTEM ] A processor joined or left the membership and
 a new membership was formed.
```

With one node functional, it’s now safe to start Corosync on the second node as well.

```
# ssh pcmk-2 -- /etc/init.d/corosync start
Starting Corosync Cluster Engine (corosync): [ OK ]
#
```

Controllo della corretta formazione del cluster

```
# grep TOTEM /var/log/messages
Aug 27 09:05:34 pcmk-1 corosync[1540]: [TOTEM ] Initializing transport (UDP/IP).
Aug 27 09:05:34 pcmk-1 corosync[1540]: [TOTEM ] Initializing transmit/receive security:
 libtomcrypt SOBER128/SHA1HMAC (mode 0).
Aug 27 09:05:35 pcmk-1 corosync[1540]: [TOTEM ] The network interface [192.168.122.101] is
 now up.
Aug 27 09:05:35 pcmk-1 corosync[1540]: [TOTEM ] A processor joined or left the membership and
 a new membership was formed.
Aug 27 09:12:11 pcmk-1 corosync[1540]: [TOTEM ] A processor joined or left the membership and
 a new membership was formed.
```

3.2. Verifica dell'installazione di Pacemaker

Ora che è stato verificato come Corosync sia funzionante è quindi possibile controllare il resto dello stack.

```
# grep pcmk_startup /var/log/messages
Aug 27 09:05:35 pcmk-1 corosync[1540]:  [pcmk ] info: pcmk_startup: CRM: InitializedAug 27
 09:05:35 pcmk-1 corosync[1540]:  [pcmk ] Logging: Initialized pcmk_startup
Aug 27 09:05:35 pcmk-1 corosync[1540]:  [pcmk ] info: pcmk_startup: Maximum core file size
 is: 18446744073709551615
Aug 27 09:05:35 pcmk-1 corosync[1540]:  [pcmk ] info: pcmk_startup: Service: 9Aug 27 09:05:35
 pcmk-1 corosync[1540]:  [pcmk ] info: pcmk_startup: Local hostname: pcmk-1
```

Now try starting Pacemaker and check the necessary processes have been started

```
# /etc/init.d/pacemaker start
Starting Pacemaker Cluster Manager: [ OK ]

# grep -e pacemakerd.*get_config_opt -e pacemakerd.*start_child -e "Starting Pacemaker" /var/
log/messages
Feb  8 13:31:24 pcmk-1 pacemakerd: [13155]: info: get_config_opt: Found 'pacemaker' for
 option: name
Feb  8 13:31:24 pcmk-1 pacemakerd: [13155]: info: get_config_opt: Found '1' for option: ver
Feb  8 13:31:24 pcmk-1 pacemakerd: [13155]: info: get_config_opt: Defaulting to 'no' for
 option: use_logd
Feb  8 13:31:24 pcmk-1 pacemakerd: [13155]: info: get_config_opt: Defaulting to 'no' for
 option: use_mgmtd
Feb  8 13:31:24 pcmk-1 pacemakerd: [13155]: info: get_config_opt: Found 'on' for option:
 debug
Feb  8 13:31:24 pcmk-1 pacemakerd: [13155]: info: get_config_opt: Found 'yes' for option:
 to_logfile
Feb  8 13:31:24 pcmk-1 pacemakerd: [13155]: info: get_config_opt: Found '/var/log/
corosync.log' for option: logfile
Feb  8 13:31:24 pcmk-1 pacemakerd: [13155]: info: get_config_opt: Found 'yes' for option:
 to_syslog
Feb  8 13:31:24 pcmk-1 pacemakerd: [13155]: info: get_config_opt: Found 'daemon' for option:
 syslog_facility
Feb  8 16:50:38 pcmk-1 pacemakerd: [13990]: info: main: Starting Pacemaker 1.1.5 (Build:
 31f088949239+):  docbook-manpages publican ncurses trace-logging cman cs-quorum heartbeat
 corosync snmp libesmtp
Feb  8 16:50:38 pcmk-1 pacemakerd: [13990]: info: start_child: Forked child 14022 for process
 stonith-ng
Feb  8 16:50:38 pcmk-1 pacemakerd: [13990]: info: start_child: Forked child 14023 for process
 cib
Feb  8 16:50:38 pcmk-1 pacemakerd: [13990]: info: start_child: Forked child 14024 for process
 lrmd
Feb  8 16:50:38 pcmk-1 pacemakerd: [13990]: info: start_child: Forked child 14025 for process
 attrd
Feb  8 16:50:38 pcmk-1 pacemakerd: [13990]: info: start_child: Forked child 14026 for process
 pengine
Feb  8 16:50:38 pcmk-1 pacemakerd: [13990]: info: start_child: Forked child 14027 for process
 crmd

# ps axf PID TTY   STAT  TIME COMMAND
  2 ?    S<   0:00 [kthreadd]
  3 ?    S<   0:00 \_ [migration/0]
... lots of processes ...
13990 ?  S    0:01 pacemakerd
14022 ?  Sa   0:00 \_ /usr/lib64/heartbeat/stonithd
14023 ?  Sa   0:00 \_ /usr/lib64/heartbeat/cib
14024 ?  Sa   0:00 \_ /usr/lib64/heartbeat/lrmd
14025 ?  Sa   0:00 \_ /usr/lib64/heartbeat/attrd
14026 ?  Sa   0:00 \_ /usr/lib64/heartbeat/pengine
14027 ?  Sa   0:00 \_ /usr/lib64/heartbeat/crmd
```

Next, check for any ERRORs during startup - there shouldn’t be any.

```
# grep ERROR: /var/log/messages | grep -v unpack_resources
#
```

Repeat on the other node and display the cluster’s status.

```
# ssh pcmk-2 -- /etc/init.d/pacemaker start
Starting Pacemaker Cluster Manager: [ OK ]
# crm_mon
============
Last updated: Thu Aug 27 16:54:55 2009Stack: openais
Current DC: pcmk-1 - partition with quorum
Version: 1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f
2 Nodes configured, 2 expected votes
0 Resources configured.
============
Online: [ pcmk-1 pcmk-2 ]
```

Capitolo 4.

Pacemaker Tools

Indice

4.1. Utilizzare i tool di Pacemaker

Nell'oscuro passato configurare Pacemaer richiedeva all'amministratore di leggere e scrivere XML. In vero stile UNIX c'erano inoltre parecchi comandi differendi, specializzati in aspetti differenti di interrogazione ed aggiornamento del cluster.

Con Pacemaker 1.0 tutto questo è cambiato ed ora esiste una cluster shell integrata e scriptabile che nasconde tutta la confusionaria impalcatura XML. Essa permette di accodare i diversi cambiamente in una sola volta ed applicarli automaticamente.

Va dedicato del tempo a familiarizzare con le peculiarità della shell.

```
# crm --help
```

```
usage:
    crm [-D display_type]
    crm [-D display_type] args
    crm [-D display_type] [-f file]

    Use crm without arguments for an interactive session.
    Supply one or more arguments for a "single-shot" use.
    Specify with -f a file which contains a script. Use '-' for
    standard input or use pipe/redirection.

    crm displays cli format configurations using a color scheme
    and/or in uppercase. Pick one of "color" or "uppercase", or
    use "-D color,uppercase" if you want colorful uppercase.
    Get plain output by "-D plain". The default may be set in
    user preferences (options).

Examples:

    # crm -f stopapp2.cli
    # crm < stopapp2.cli
    # crm resource stop global_www
    # crm status
```

The primary tool for monitoring the status of the cluster is crm_mon (also available as crm status). It can be run in a variety of modes and has a number of output options. To find out about any of the tools that come with Pacemaker, simply invoke them with the --help option or consult the included man pages. Both sets of output are created from the tool, and so will always be in sync with each other and the tool itself.

Additionally, the Pacemaker version and supported cluster stack(s) are available via the --feature option to pacemakerd.

```
# pacemakerd --features
```

```
Pacemaker 1.1.9-3.fc20.2 (Build: 781a388)
```

```
 Supporting v3.0.7:  generated-manpages agent-manpages ncurses libqb-logging libqb-ipc
 upstart systemd nagios  corosync-native
```

```
# pacemakerd --help
```

```
pacemakerd - Start/Stop Pacemaker

Usage: pacemakerd mode [options]
Options:
 -?, --help   This text
 -$, --version   Version information
 -V, --verbose   Increase debug output
 -S, --shutdown   Instruct Pacemaker to shutdown on this machine
 -F, --features   Display the full version and list of features Pacemaker was built with

Additional Options:
 -f, --foreground   (Ignored) Pacemaker always runs in the foreground
 -p, --pid-file=value  (Ignored) Daemon pid file location

Report bugs to pacemaker@oss.clusterlabs.org
```

```
# crm_mon --help
```

```
crm_mon - Provides a summary of cluster's current state.

Outputs varying levels of detail in a number of different formats.

Usage: crm_mon mode [options]
Options:
 -?, --help   This text
 -$, --version   Version information
 -V, --verbose   Increase debug output
 -Q, --quiet   Display only essential output

Modes:
 -h, --as-html=value Write cluster status to the named html file
 -X, --as-xml   Write cluster status as xml to stdout. This will enable one-shot mode.
 -w, --web-cgi   Web mode with output suitable for cgi
 -s, --simple-status  Display the cluster status once as a simple one line output (suitable
 for nagios)

Display Options:
 -n, --group-by-node   Group resources by node
 -r, --inactive   Display inactive resources
 -f, --failcounts   Display resource fail counts
 -o, --operations   Display resource operation history
 -t, --timing-details   Display resource operation history with timing details
 -c, --tickets    Display cluster tickets
 -W, --watch-fencing    Listen for fencing events. For use with --external-agent, --mail-to
 and/or --snmp-traps where supported
 -A, --show-node-attributes  Display node attributes

Additional Options:
 -i, --interval=value  Update frequency in seconds
 -1, --one-shot   Display the cluster status once on the console and exit
 -N, --disable-ncurses   Disable the use of ncurses
 -d, --daemonize   Run in the background as a daemon
 -p, --pid-file=value  (Advanced) Daemon pid file location
 -E, --external-agent=value A program to run when resource operations take place.
 -e, --external-recipient=value A recipient for your program (assuming you want the program
 to send something to someone).

Examples:

Display the cluster status on the console with updates as they occur:
 # crm_mon

Display the cluster status on the console just once then exit:

 # crm_mon -1

Display your cluster status, group resources by node, and include inactive resources in the
 list:

 # crm_mon --group-by-node --inactive

Start crm_mon as a background daemon and have it write the cluster status to an HTML file:

 # crm_mon --daemonize --as-html /path/to/docroot/filename.html

Start crm_mon and export the current cluster status as xml to stdout, then exit.:

 # crm_mon --as-xml


Report bugs to pacemaker@oss.clusterlabs.org
```

Nota

Se SNMP e/o le opzioni mail non sono listate allora Pacemaker non è stato compilato per supportare queste funzionalità. Potrebbe essere dovuto ad una scelta della distribuzione che si sta utilizzando oppure al fatto che le librerie necessarie non sono disponibili. In questo caso è bene contattare il fornitore dei pacchetti per avere maggiori dettagli.

Creare un cluster Active/Passive

Indice

5.1. Esplorare la configurazione esistente

Quando Pacemaker viene avviato automatica registra il numero ed i dettagli dei nodi nel cluster, così come lo stack è utilizzato e la versione di Pacemaker utilizzata.

Ecco come dovrebbe apparire la configurazione base.

```
# crm configure show
node pcmk-1
node pcmk-2
property $id="cib-bootstrap-options" \
    dc-version="1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f" \
    cluster-infrastructure="openais" \
    expected-quorum-votes="2"
```

For those that are not of afraid of XML, you can see the raw configuration by appending "xml" to the previous command.

Questo è l'ultimo XML ad essere utilizzato nel documento.

```
# crm configure show xml
<?xml version="1.0" ?>
<cib admin_epoch="0" crm_feature_set="3.0.1" dc-uuid="pcmk-1" epoch="13" have-quorum="1" num_updates="7" validate-with="pacemaker-1.0">
 <configuration>
  <crm_config>
   <cluster_property_set id="cib-bootstrap-options">
    <nvpair id="cib-bootstrap-options-dc-version" name="dc-version" value="1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f"/>
    <nvpair id="cib-bootstrap-options-cluster-infrastructure" name="cluster-infrastructure" value="openais"/>
    <nvpair id="cib-bootstrap-options-expected-quorum-votes" name="expected-quorum-votes" value="2"/>
   </cluster_property_set>
  </crm_config>
  <rsc_defaults/>
  <op_defaults/>
  <nodes>
   <node id="pcmk-1" type="normal" uname="pcmk-1"/>
   <node id="pcmk-2" type="normal" uname="pcmk-2"/>
  </nodes>
  <resources/>
  <constraints/>
 </configuration>
</cib>
```

Prima di effettuare qualsiasi cambiamento è buona norma controllare la validità della configurazione.

```
# crm_verify -L
crm_verify[2195]: 2009/08/27_16:57:12 ERROR: unpack_resources: Resource start-up disabled
 since no STONITH resources have been defined
crm_verify[2195]: 2009/08/27_16:57:12 ERROR: unpack_resources: Either configure some or
 disable STONITH with the stonith-enabled option
crm_verify[2195]: 2009/08/27_16:57:12 ERROR: unpack_resources: NOTE: Clusters with shared
 data need STONITH to ensure data integrity
Errors found during check: config not valid -V may provide more details
#
```

Come si può notare il tool ha trovato qualche errore.

In order to guarantee the safety of your data [1] , Pacemaker ships with STONITH [2] enabled. However it also knows when no STONITH configuration has been supplied and reports this as a problem (since the cluster would not be able to make progress if a situation requiring node fencing arose).

Per adesso la funzionalità verrà disabilitata e configurata in seguito nella sezione Configurare STONITH. E' importante notare che l'uso di STONITH è altamente consigliato, disabilitarlo indica al cluster di dare per scontato che i nodi falliti vengano spenti. Alcuni rivenditori potrebbero rifiutarsi di supportare cluster che hanno STONITH disabilitato.

Per disabilitare STONITH è necessario impostare l'opzione stonith-enabled a false.

```
# crm configure property stonith-enabled=false
# crm_verify -L
```

Con la nuova opzione impostata la configurazione del cluster è ora valida.

Avvertimento

The use of stonith-enabled=false is completely inappropriate for a production cluster. We use it here to defer the discussion of its configuration which can differ widely from one installation to the next. See *Sezione 9.1, «What Is STONITH»* for information on why STONITH is important and details on how to configure it.

5.2. Aggiungere una risorsa

La prima cosa da fare è configurare un indirizzo IP. Indipendentemente da dove i servizi cluster stanno funzionando è necessario un indirizzo per raggiungerli. Verrà scelto ed aggiunto 192.168.122.101 come indirizzo virtuale, con il nome di ClusterIP e verrà indicato al cluster di controllarlo ogni 30 secondi.

Importante

L'indirizzo scelto non dovrà essere già associato ad un nodo fisico

[1] If the data is corrupt, there is little point in continuing to make it available

[2] A common node fencing mechanism. Used to ensure data integrity by powering off "bad" nodes

```
# crm configure primitive ClusterIP ocf:heartbeat:IPaddr2 \
     params ip=192.168.122.101 cidr_netmask=32 \
     op monitor interval=30s
```

The other important piece of information here is ocf:heartbeat:IPaddr2.

This tells Pacemaker three things about the resource you want to add. The first field, ocf, is the standard to which the resource script conforms to and where to find it. The second field is specific to OCF resources and tells the cluster which namespace to find the resource script in, in this case heartbeat. The last field indicates the name of the resource script.

Per ottenere una lista delle classi di risorse disponibili, lanciare

```
# crm ra classesheartbeat
lsb ocf / heartbeat pacemakerstonith
```

Per poi trovare tutte i resource agent OCF disponibili con Pacemaker ed Heartbeat, lanciare

```
# crm ra list ocf pacemaker
ClusterMon   Dummy     Stateful    SysInfo    SystemHealth  controld
ping      pingd
# crm ra list ocf heartbeat
AoEtarget       AudibleAlarm      ClusterMon       Delay
Dummy         EvmsSCC        Evmsd         Filesystem
ICP          IPaddr         IPaddr2        IPsrcaddr
LVM          LinuxSCSI       MailTo         ManageRAID
ManageVE        Pure-FTPd       Raid1         Route
SAPDatabase      SAPInstance      SendArp        ServeRAID
SphinxSearchDaemon   Squid         Stateful        SysInfo
VIParip        VirtualDomain     WAS          WAS6
WinPopup        Xen          Xinetd         anything
apache         db2          drbd          eDir88
iSCSILogicalUnit    iSCSITarget      ids          iscsi
ldirectord       mysql         mysql-proxy      nfsserver
oracle         oralsnr        pgsql         pingd
portblock       rsyncd         scsi2reservation    sfex
tomcat         vmware
#
```

A questo punto va verificato come la risorsa IP sia stata aggiunta e visualizzato lo stato del cluster per vedere che ora è attiva.

```
# crm configure shownode pcmk-1
node pcmk-2primitive ClusterIP ocf:heartbeat:IPaddr2 \
    params ip="192.168.122.101" cidr_netmask="32" \
    op monitor interval="30s"
property $id="cib-bootstrap-options" \
    dc-version="1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f" \
    cluster-infrastructure="openais" \
    expected-quorum-votes="2" \
    stonith-enabled="false" \
# crm_mon
============
Last updated: Fri Aug 28 15:23:48 2009
Stack: openais
Current DC: pcmk-1 - partition with quorum
Version: 1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f
2 Nodes configured, 2 expected votes
1 Resources configured.
============
Online: [ pcmk-1 pcmk-2 ]
ClusterIP (ocf::heartbeat:IPaddr): Started pcmk-1
```

5.3. Effettuare un Failover

Trattandosi di un cluster ad alta-affidabilità, è necessario testare il failover della nostra nuova risorsa prima di proseguire.

Per prima cosa va identificato da quale nodo l'indirizzo IP è erogato

```
# crm resource status ClusterIP
resource ClusterIP is running on: pcmk-1
#
```

Shut down Pacemaker and Corosync on that machine.

```
# ssh pcmk-1 -- /etc/init.d/pacemaker stop
Signaling Pacemaker Cluster Manager to terminate: [ OK ]
Waiting for cluster services to unload:. [ OK ]
# ssh pcmk-1 -- /etc/init.d/corosync stop
Stopping Corosync Cluster Engine (corosync): [ OK ]
Waiting for services to unload: [ OK ]
#
```

Una volta che Corosync non sta più funzionando, è possibile verificare sull'altro nodo lo stato del cluster attraverso crm_mon.

```
# crm_mon
============
Last updated: Fri Aug 28 15:27:35 2009
Stack: openais
Current DC: pcmk-2 - partition WITHOUT quorum
Version: 1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f
2 Nodes configured, 2 expected votes
1 Resources configured.
============

Online: [ pcmk-2 ]OFFLINE: [ pcmk-1 ]
```

Ci sono tre cose da evidenziare in merito all'attuale stato del cluster. La prima è che, come aspettato, pcmk-1 è ora offline. Ad ogni modo è possibile anche notare come ClusterIP non sta funzionando da nessuna parte!

5.3.1. Quorum e Cluster a due nodi

This is because the cluster no longer has quorum, as can be seen by the text "partition WITHOUT quorum" (emphasised green) in the output above. In order to reduce the possibility of data corruption, Pacemaker's default behavior is to stop all resources if the cluster does not have quorum.

Un cluster viene definito con quorum quando più della metà dei nodi conosciuti o aspettati sono online o, attraverso la matematica, quando la seguente equazione è vera:

```
total_nodes < 2 * active_nodes
```

Therefore a two-node cluster only has quorum when both nodes are running, which is no longer the case for our cluster. This would normally make the creation of a two-node cluster pointless [3], however it is possible to control how Pacemaker behaves when quorum is lost. In particular, we can tell the cluster to simply ignore quorum altogether.

```
# crm configure property no-quorum-policy=ignore
# crm configure show
node pcmk-1
node pcmk-2
primitive ClusterIP ocf:heartbeat:IPaddr2 \
    params ip="192.168.122.101" cidr_netmask="32" \
    op monitor interval="30s"
property $id="cib-bootstrap-options" \
    dc-version="1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f" \
    cluster-infrastructure="openais" \
    expected-quorum-votes="2" \
    stonith-enabled="false" \
    no-quorum-policy="ignore"
```

Dopo alcuni istanti il cluster avvierà l'indirizzo IP sui nodi rimanenti. E' da notare che il cluster non ha comunque il quorum.

```
# crm_mon
============
Last updated: Fri Aug 28 15:30:18 2009
Stack: openais
Current DC: pcmk-2 - partition WITHOUT quorum
Version: 1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f
2 Nodes configured, 2 expected votes
1 Resources configured.
============
Online: [ pcmk-2 ]
OFFLINE: [ pcmk-1 ]
ClusterIP (ocf::heartbeat:IPaddr): Started pcmk-2
```

Viene quindi simulato il recovery del nodo attraverso il riavvio dello stack cluster su pcmk-1 ed il controllo dello stato del cluster.

```
# /etc/init.d/corosync start
Starting Corosync Cluster Engine (corosync): [ OK ]
# /etc/init.d/pacemaker start
Starting Pacemaker Cluster Manager: [ OK ]# crm_mon
============
Last updated: Fri Aug 28 15:32:13 2009
Stack: openais
Current DC: pcmk-2 - partition with quorum
Version: 1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f
2 Nodes configured, 2 expected votes
1 Resources configured.
============
Online: [ pcmk-1 pcmk-2 ]

ClusterIP    (ocf::heartbeat:IPaddr):    Started pcmk-1
```

Qui appare qualcosa che potrebbe sembrare sorprendente: l'IP è tornato attivo sul nodo originale!

[3] Actually some would argue that two-node clusters are always pointless, but that is an argument for another time

5.3.2. Evitare che le risorse si muovano dopo il recovery

In some circumstances, it is highly desirable to prevent healthy resources from being moved around the cluster. Moving resources almost always requires a period of downtime. For complex services like Oracle databases, this period can be quite long.

To address this, Pacemaker has the concept of resource stickiness which controls how much a service prefers to stay running where it is. You may like to think of it as the "cost" of any downtime. By default, Pacemaker assumes there is zero cost associated with moving resources and will do so to achieve "optimal" [4] resource placement. We can specify a different stickiness for every resource, but it is often sufficient to change the default.

```
# crm configure rsc_defaults resource-stickiness=100
# crm configure show
node pcmk-1
node pcmk-2
primitive ClusterIP ocf:heartbeat:IPaddr2 \
    params ip="192.168.122.101" cidr_netmask="32" \
    op monitor interval="30s"
property $id="cib-bootstrap-options" \
    dc-version="1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f" \
    cluster-infrastructure="openais" \
    expected-quorum-votes="2" \
    stonith-enabled="false" \
    no-quorum-policy="ignore"rsc_defaults $id="rsc-options" \
    resource-stickiness="100"
```

Se ora si effettua nuovamente il test di failover si osserverà che come da pronostico ClusterIP verrà spostata su pcmk-2 quando pcmk-1 viene messo offline.

```
# ssh pcmk-1 -- /etc/init.d/pacemaker stop
Signaling Pacemaker Cluster Manager to terminate:         [  OK  ]
Waiting for cluster services to unload:.                  [  OK  ]
# ssh pcmk-1 -- /etc/init.d/corosync stop
Stopping Corosync Cluster Engine (corosync):        [ OK ]
Waiting for services to unload:              [ OK ]
# ssh pcmk-2 -- crm_mon -1
============
Last updated: Fri Aug 28 15:39:38 2009
Stack: openais
Current DC: pcmk-2 - partition WITHOUT quorum
Version: 1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f
2 Nodes configured, 2 expected votes
1 Resources configured.
============

Online: [ pcmk-2 ]
OFFLINE: [ pcmk-1 ]
ClusterIP    (ocf::heartbeat:IPaddr):    Started pcmk-2
```

Quando però pcmk-1 torna online ClusterIP rimane su pcmk-2.

```
# /etc/init.d/corosync start
Starting Corosync Cluster Engine (corosync): [ OK ]
# /etc/init.d/pacemaker start
Starting Pacemaker Cluster Manager: [ OK ]
```

[4] It should be noted that Pacemaker's definition of optimal may not always agree with that of a human's. The order in which Pacemaker processes lists of resources and nodes creates implicit preferences in situations where the administrator has not explicitly specified them

```
# crm_mon
============
Last updated: Fri Aug 28 15:41:23 2009
Stack: openais
Current DC: pcmk-2 - partition with quorum
Version: 1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f
2 Nodes configured, 2 expected votes
1 Resources configured.
============

Online: [ pcmk-1 pcmk-2 ]

ClusterIP    (ocf::heartbeat:IPaddr):    Started pcmk-2
```

Apache - Aggiungere ulteriori servizi

Indice

6.1. Forward

Ora che è stato realizzato un semplice, ma funzionale cluster a due nodi active/passive, è tempo di aggiungere servizi reali. Si partirà con Apache poiché molti cluster lo includono ed è relativamente semplice da configurare.

6.2. Installazione

Before continuing, we need to make sure Apache is installed on both hosts.

```
# yum install -y httpdSetting up Install Process
Resolving Dependencies
--> Running transaction check
---> Package httpd.x86_64 0:2.2.13-2.fc12 set to be updated
--> Processing Dependency: httpd-tools = 2.2.13-2.fc12 for package:
 httpd-2.2.13-2.fc12.x86_64
--> Processing Dependency: apr-util-ldap for package: httpd-2.2.13-2.fc12.x86_64
--> Processing Dependency: /etc/mime.types for package: httpd-2.2.13-2.fc12.x86_64
--> Processing Dependency: libaprutil-1.so.0()(64bit) for package: httpd-2.2.13-2.fc12.x86_64
--> Processing Dependency: libapr-1.so.0()(64bit) for package: httpd-2.2.13-2.fc12.x86_64
--> Running transaction check
---> Package apr.x86_64 0:1.3.9-2.fc12 set to be updated
---> Package apr-util.x86_64 0:1.3.9-2.fc12 set to be updated
---> Package apr-util-ldap.x86_64 0:1.3.9-2.fc12 set to be updated
---> Package httpd-tools.x86_64 0:2.2.13-2.fc12 set to be updated
---> Package mailcap.noarch 0:2.1.30-1.fc12 set to be updated
--> Finished Dependency Resolution

Dependencies Resolved

=======================================================================================
Package        Arch       Version        Repository     Size
=======================================================================================
Installing:
httpd        x86_64      2.2.13-2.fc12      rawhide      735 k
Installing for dependencies:
apr         x86_64      1.3.9-2.fc12       rawhide      117 k
apr-util      x86_64      1.3.9-2.fc12       rawhide      84 k
apr-util-ldap    x86_64      1.3.9-2.fc12       rawhide      15 k
httpd-tools     x86_64      2.2.13-2.fc12      rawhide      63 k
mailcap       noarch      2.1.30-1.fc12      rawhide      25 k
```

```
Transaction Summary
=======================================================================================
Install    6 Package(s)
Upgrade    0 Package(s)

Total download size: 1.0 M
Downloading Packages:
(1/6): apr-1.3.9-2.fc12.x86_64.rpm                 | 117 kB   00:00
(2/6): apr-util-1.3.9-2.fc12.x86_64.rpm                | 84 kB   00:00
(3/6): apr-util-ldap-1.3.9-2.fc12.x86_64.rpm            | 15 kB   00:00
(4/6): httpd-2.2.13-2.fc12.x86_64.rpm                | 735 kB   00:00
(5/6): httpd-tools-2.2.13-2.fc12.x86_64.rpm              | 63 kB   00:00
(6/6): mailcap-2.1.30-1.fc12.noarch.rpm                | 25 kB   00:00
 ---------------------------------------------------------------------------------------
Total                           875 kB/s | 1.0 MB   00:01
Running rpm_check_debug
Running Transaction Test
Finished Transaction Test
Transaction Test Succeeded
Running Transaction
 Installing   : apr-1.3.9-2.fc12.x86_64                      1/6
 Installing   : apr-util-1.3.9-2.fc12.x86_64                  2/6
 Installing   : apr-util-ldap-1.3.9-2.fc12.x86_64                 3/6
 Installing   : httpd-tools-2.2.13-2.fc12.x86_64                4/6
 Installing   : mailcap-2.1.30-1.fc12.noarch                  5/6
 Installing   : httpd-2.2.13-2.fc12.x86_64                   6/6

Installed:
 httpd.x86_64 0:2.2.13-2.fc12

Dependency Installed:
 apr.x86_64 0:1.3.9-2.fc12      apr-util.x86_64 0:1.3.9-2.fc12
 apr-util-ldap.x86_64 0:1.3.9-2.fc12 httpd-tools.x86_64 0:2.2.13-2.fc12
 mailcap.noarch 0:2.1.30-1.fc12

Complete!
```

Inoltre il cluster necessita del tool wget per assicurarsi la capacità di controllare lo stato del server Apache.

```
# yum install -y wgetSetting up Install Process
Resolving Dependencies
--> Running transaction check
---> Package wget.x86_64 0:1.11.4-5.fc12 set to be updated
--> Finished Dependency Resolution

Dependencies Resolved

===========================================================================================
Package    Arch       Version           Repository        Size
===========================================================================================
Installing:
wget      x86_64     1.11.4-5.fc12         rawhide        393 k

Transaction Summary
===========================================================================================
Install    1 Package(s)
Upgrade    0 Package(s)

Total download size: 393 k
Downloading Packages:
wget-1.11.4-5.fc12.x86_64.rpm                      | 393 kB   00:00
Running rpm_check_debug
Running Transaction Test
Finished Transaction Test
Transaction Test Succeeded
```

```
Running Transaction
 Installing   : wget-1.11.4-5.fc12.x86_64                         1/1

Installed:
 wget.x86_64 0:1.11.4-5.fc12

Complete!
```

6.3. Preparazione

Prima di tutto è necessario creare la pagina che Apache servirà. Su Fedora la docroot di default di Apache è /var/www/html, qui verrà creato il file index.

```
[root@pcmk-1 ~]# cat <<-END >/var/www/html/index.html <html>
 <body>My Test Site - pcmk-1</body>
 </html>
 END
```

Per il momento l'architettura verrà semplificata in modo da servire unicamente un sito statico e sincronizzare manualmente i dati tra i due nodi. Quindi lo stesso comando andrà lanciato su pcmk-2.

```
[root@pcmk-2 ~]# cat <<-END >/var/www/html/index.html <html>
 <body>My Test Site - pcmk-2</body>
 </html>
 END
```

6.4. Enable the Apache status URL

In order to monitor the health of your Apache instance, and recover it if it fails, the resource agent used by Pacemaker assumes the server-status URL is available. Look for the following in /etc/httpd/conf/httpd.conf and make sure it is not disabled or commented out:

```
<Location /server-status>
   SetHandler server-status
   Order deny,allow
   Deny from all
   Allow from 127.0.0.1
</Location>
```

6.5. Aggiornare la configurazione

At this point, Apache is ready to go, all that needs to be done is to add it to the cluster. Lets call the resource WebSite. We need to use an OCF script called apache in the heartbeat namespace [1] , the only required parameter is the path to the main Apache configuration file and we'll tell the cluster to check once a minute that apache is still running.

```
# crm configure primitive WebSite ocf:heartbeat:apache params configfile=/etc/httpd/conf/
httpd.conf op monitor interval=1min
# crm configure show
node pcmk-1
node pcmk-2primitive WebSite ocf:heartbeat:apache \ params configfile="/etc/httpd/conf/
httpd.conf" \ op monitor interval="1min"primitive ClusterIP ocf:heartbeat:IPaddr2 \
    params ip="192.168.122.101" cidr_netmask="32" \
```

[1] Compare the key used here ocf:heartbeat:apache with the one we used earlier for the IP address: ocf:heartbeat:IPaddr2

```
    op monitor interval="30s"
property $id="cib-bootstrap-options" \
    dc-version="1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f" \
    cluster-infrastructure="openais" \
    expected-quorum-votes="2" \
    stonith-enabled="false" \
    no-quorum-policy="ignore"
rsc_defaults $id="rsc-options" \
    resource-stickiness="100"
```

Dopo una breve attesa il cluster dovrbbe avviare apache

```
# crm_mon
============
Last updated: Fri Aug 28 16:12:49 2009
Stack: openais
Current DC: pcmk-2 - partition with quorum
Version: 1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f
2 Nodes configured, 2 expected votes
2 Resources configured.
============

Online: [ pcmk-1 pcmk-2 ]

ClusterIP    (ocf::heartbeat:IPaddr):    Started pcmk-2
WebSite    (ocf::heartbeat:apache):    Started pcmk-1
```

Attenzione però, la risorsa WebSite non sta girando sullo stesso host dell'indirizzo IP!

6.6. Assicurarsi che le risorse funzionino sullo stesso host

To reduce the load on any one machine, Pacemaker will generally try to spread the configured resources across the cluster nodes. However we can tell the cluster that two resources are related and need to run on the same host (or not at all). Here we instruct the cluster that WebSite can only run on the host that ClusterIP is active on.

For the constraint, we need a name (choose something descriptive like website-with-ip), indicate that its mandatory (so that if ClusterIP is not active anywhere, WebSite will not be permitted to run anywhere either) by specifying a score of INFINITY and finally list the two resources.

Nota

If ClusterIP is not active anywhere, WebSite will not be permitted to run anywhere.

Importante

Colocation constraints are "directional", in that they imply certain things about the order in which the two resources will have a location chosen. In this case we’re saying **`WebSite`** needs to be placed on the same machine as **`ClusterIP`**, this implies that we must know the location of **`ClusterIP`** before choosing a location for **`WebSite`**.

```
# crm configure colocation website-with-ip INFINITY: WebSite ClusterIP
# crm configure show
node pcmk-1
node pcmk-2
primitive WebSite ocf:heartbeat:apache \
    params configfile="/etc/httpd/conf/httpd.conf" \
    op monitor interval="1min"
primitive ClusterIP ocf:heartbeat:IPaddr2 \
    params ip="192.168.122.101" cidr_netmask="32" \
    op monitor interval="30s"colocation website-with-ip inf: WebSite
 ClusterIPproperty $id="cib-bootstrap-options" \
    dc-version="1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f" \
    cluster-infrastructure="openais" \
    expected-quorum-votes="2" \
    stonith-enabled="false" \
    no-quorum-policy="ignore"
rsc_defaults $id="rsc-options" \
    resource-stickiness="100"
# crm_mon
============
Last updated: Fri Aug 28 16:14:34 2009
Stack: openais
Current DC: pcmk-2 - partition with quorum
Version: 1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f
2 Nodes configured, 2 expected votes
2 Resources configured.
============

Online: [ pcmk-1 pcmk-2 ]

ClusterIP    (ocf::heartbeat:IPaddr):    Started pcmk-2
WebSite    (ocf::heartbeat:apache):    Started pcmk-2
```

6.7. Controllare l'ordinamento di start e stop di una risorsa

When Apache starts, it binds to the available IP addresses. It doesn't know about any addresses we add afterwards, so not only do they need to run on the same node, but we need to make sure ClusterIP is already active before we start WebSite. We do this by adding an ordering constraint. We need to give it a name (choose something descriptive like apache-after-ip), indicate that its mandatory (so that any recovery for ClusterIP will also trigger recovery of WebSite) and list the two resources in the order we need them to start.

```
# crm configure order apache-after-ip mandatory: ClusterIP WebSite
# crm configure show
node pcmk-1
node pcmk-2
primitive WebSite ocf:heartbeat:apache \
    params configfile="/etc/httpd/conf/httpd.conf" \
    op monitor interval="1min"
primitive ClusterIP ocf:heartbeat:IPaddr2 \
    params ip="192.168.122.101" cidr_netmask="32" \
    op monitor interval="30s"
colocation website-with-ip inf: WebSite ClusterIPorder apache-after-ip inf: ClusterIP
 WebSiteproperty $id="cib-bootstrap-options" \
    dc-version="1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f" \
    cluster-infrastructure="openais" \
    expected-quorum-votes="2" \
    stonith-enabled="false" \
    no-quorum-policy="ignore"
rsc_defaults $id="rsc-options" \
    resource-stickiness="100"
```

6.8. Specificare una Location preferita

Pacemaker non si basa su una sorta di simmetria tra i nodi hardware, quindi si potrebbe prevedere di avere anche una macchina più potente di un'altra. In questi casi ha senso erogare le risorse su questa macchina, se è disponibile. Per fare ciò viene creata una location constraint. Anche in questo caso viene assegnato un nome descrittivo (prefer-pcmk-1), specificata la risorsa che girerà qui (WebSite), quanto peso avrà per girare qui (verrà utilizzato 50, ma in una situazione a due nodi ogni valore superiore allo 0 andrà bene) ed il nome dell'host.

```
# crm configure location prefer-pcmk-1 WebSite 50: pcmk-1
# crm configure show
node pcmk-1
node pcmk-2
primitive WebSite ocf:heartbeat:apache \
    params configfile="/etc/httpd/conf/httpd.conf" \
    op monitor interval="1min"
primitive ClusterIP ocf:heartbeat:IPaddr2 \
    params ip="192.168.122.101" cidr_netmask="32" \
    op monitor interval="30s"location prefer-pcmk-1 WebSite 50: pcmk-1colocation website-
with-ip inf: WebSite ClusterIP
property $id="cib-bootstrap-options" \
    dc-version="1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f" \
    cluster-infrastructure="openais" \
    expected-quorum-votes="2" \
    stonith-enabled="false" \
    no-quorum-policy="ignore"
rsc_defaults $id="rsc-options" \
    resource-stickiness="100"
# crm_mon
============
Last updated: Fri Aug 28 16:17:35 2009
Stack: openais
Current DC: pcmk-2 - partition with quorum
Version: 1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f
2 Nodes configured, 2 expected votes
2 Resources configured.
============

Online: [ pcmk-1 pcmk-2 ]

ClusterIP    (ocf::heartbeat:IPaddr):    Started pcmk-2WebSite    (ocf::heartbeat:apache):
  Started pcmk-2
```

Attenzione però, le risorse sono comunque su pcmk-2!

Anche se al momento viene preferito pcmk-1 a pcmk-2, questa preferenza è (intenzionalmente) minore dello stickiness della risorsa (quanto è stato preferito non avere un downtime superfluo).

Per osservare gli attuali punteggi di piazzamento è possibile usare un tool denominato ptest

```
ptest -sL
```

Nota

Include output There is a way to force them to move though…

6.9. Spostamento manuale delle risorse all'interno del cluster

C'è sempre la necessità da parte di un amministratore di soprassedere il cluster e forzare le risorse a spostarsi in una posizione specifica. E' possibile ignorare felicemente le location constraint create sopra, fornendo il nome della risorsa che si vuole spostare e la destinazione della stessa, il comando farà il resto.

```
# crm resource move WebSite pcmk-1
# crm_mon
============
Last updated: Fri Aug 28 16:19:24 2009
Stack: openais
Current DC: pcmk-2 - partition with quorum
Version: 1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f
2 Nodes configured, 2 expected votes
2 Resources configured.
============

Online: [ pcmk-1 pcmk-2 ]

ClusterIP    (ocf::heartbeat:IPaddr):    Started pcmk-1
WebSite    (ocf::heartbeat:apache):    Started pcmk-1
```

Notice how the colocation rule we created has ensured that ClusterIP was also moved to pcmk-1. For the curious, we can see the effect of this command by examining the configuration

```
# crm configure show
node pcmk-1
node pcmk-2
primitive WebSite ocf:heartbeat:apache \
    params configfile="/etc/httpd/conf/httpd.conf" \
    op monitor interval="1min"
primitive ClusterIP ocf:heartbeat:IPaddr2 \
    params ip="192.168.122.101" cidr_netmask="32" \
    op monitor interval="30s"
location cli-prefer-WebSite WebSite \
    rule $id="cli-prefer-rule-WebSite" inf: #uname eq pcmk-1
location prefer-pcmk-1 WebSite 50: pcmk-1
colocation website-with-ip inf: WebSite ClusterIP
property $id="cib-bootstrap-options" \
    dc-version="1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f" \
    cluster-infrastructure="openais" \
    expected-quorum-votes="2" \
    stonith-enabled="false" \
    no-quorum-policy="ignore"
rsc_defaults $id="rsc-options" \
    resource-stickiness="100"
```

Viene evidenziata la constraint automatizzata utilizzata per spostare la risorsa in pcmk-1

6.9.1. Restituire il controllo nuovamente al cluster

Una volta terminata una qualsiasi attività che richiede lo spostamento della risorsa in pcmk-1, in questo caso nessuna, è possibile consentire al cluster di riprendere le normali operazioni con il comando unmove. Dal momento che in precedenza è stata configurata una default stickiness, la risorsa rimarra in pcmk-1.

```
# crm resource unmove WebSite
# crm configure show
```

```
node pcmk-1
node pcmk-2
primitive WebSite ocf:heartbeat:apache \
    params configfile="/etc/httpd/conf/httpd.conf" \
    op monitor interval="1min"
primitive ClusterIP ocf:heartbeat:IPaddr2 \
    params ip="192.168.122.101" cidr_netmask="32" \
    op monitor interval="30s"
location prefer-pcmk-1 WebSite 50: pcmk-1
colocation website-with-ip inf: WebSite ClusterIP
property $id="cib-bootstrap-options" \
    dc-version="1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f" \
    cluster-infrastructure="openais" \
    expected-quorum-votes="2" \
    stonith-enabled="false" \
    no-quorum-policy="ignore"
rsc_defaults $id="rsc-options" \
    resource-stickiness="100"
```

E' da notare come la constraint automatizzata non esiste più. Se viene controllato lo stato del cluster è possibile osservare che, così come aspettato, le risorse sono ancora attiva su pcmk-1.

```
# crm_mon
============
Last updated: Fri Aug 28 16:20:53 2009
Stack: openais
Current DC: pcmk-2 - partition with quorum
Version: 1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f
2 Nodes configured, 2 expected votes
2 Resources configured.
============

Online: [ pcmk-1 pcmk-2 ]

 ClusterIP    (ocf::heartbeat:IPaddr):    Started pcmk-1
 WebSite    (ocf::heartbeat:apache):    Started pcmk-1
```

Capitolo 7.

Storage replicato con DRBD

Indice

7.1. Background

Even if you’re serving up static websites, having to manually synchronize the contents of that website to all the machines in the cluster is not ideal. For dynamic websites, such as a wiki, it’s not even an option. Not everyone care afford network-attached storage but somehow the data needs to be kept in sync. Enter DRBD which can be thought of as network based RAID-1. See *http://www.drbd.org/* for more details.

7.2. Installare i pacchetti DRBD

Since its inclusion in the upstream 2.6.33 kernel, everything needed to use DRBD ships with Fedora 13. All you need to do is install it:

```
# yum install -y drbd-pacemaker drbd-udev
Loaded plugins: presto, refresh-packagekit
Setting up Install Process
Resolving Dependencies
--> Running transaction check
---> Package drbd-pacemaker.x86_64 0:8.3.7-2.fc13 set to be updated
--> Processing Dependency: drbd-utils = 8.3.7-2.fc13 for package: drbd-
pacemaker-8.3.7-2.fc13.x86_64
--> Running transaction check
---> Package drbd-utils.x86_64 0:8.3.7-2.fc13 set to be updated
--> Finished Dependency Resolution

Dependencies Resolved

=================================================================================
 Package                Arch           Version              Repository      Size
=================================================================================
Installing:
 drbd-pacemaker         x86_64         8.3.7-2.fc13         fedora          19 k
Installing for dependencies:
 drbd-utils             x86_64         8.3.7-2.fc13         fedora         165 k

Transaction Summary
=================================================================================
Install       2 Package(s)
Upgrade       0 Package(s)

Total download size: 184 k
Installed size: 427 k
```

```
Downloading Packages:
Setting up and reading Presto delta metadata
fedora/prestodelta                                        | 1.7 kB     00:00
Processing delta metadata
Package(s) data still to download: 184 k
(1/2): drbd-pacemaker-8.3.7-2.fc13.x86_64.rpm             |  19 kB     00:01
(2/2): drbd-utils-8.3.7-2.fc13.x86_64.rpm                 | 165 kB     00:02
 ---------------------------------------------------------------------------------
Total                                             45 kB/s | 184 kB     00:04
Running rpm_check_debug
Running Transaction Test
Transaction Test Succeeded
Running Transaction
  Installing     : drbd-utils-8.3.7-2.fc13.x86_64                             1/2
  Installing     : drbd-pacemaker-8.3.7-2.fc13.x86_64                         2/2

Installed:
  drbd-pacemaker.x86_64 0:8.3.7-2.fc13

Dependency Installed:
  drbd-utils.x86_64 0:8.3.7-2.fc13

Complete!
```

7.3. Configurare DRBD

Prima di configurare DRBD è necessario definire a parte lo spazio che questo dovrà utilizzare.

7.3.1. Creare una partizione per DRBD

Se si dispone di più di 1Gb liberi è bene servirsene. Negli scopi di questa guida 1Gb è fin troppo spazio per un file html e sufficiente per ospitare più avanti i metadata di GFS2.

```
# lvcreate -n drbd-demo -L 1G VolGroup
Logical volume "drbd-demo" created
# lvs
LV    VG    Attr  LSize  Origin Snap% Move Log Copy% Convert
drbd-demo VolGroup -wi-a- 1.00G
lv_root  VolGroup -wi-ao  7.30G
lv_swap  VolGroup -wi-ao 500.00M
```

La medesima operazione va ripetuta sul secondo nodo, avendo cura di utilizzare la stessa grandezza per la partizione.

```
# ssh pcmk-2 -- lvs
LV   VG    Attr  LSize  Origin Snap% Move Log Copy% Convert
lv_root VolGroup -wi-ao  7.30G
lv_swap VolGroup -wi-ao 500.00M
# ssh pcmk-2 -- lvcreate -n drbd-demo -L 1G VolGroup
Logical volume "drbd-demo" created
# ssh pcmk-2 -- lvs
LV    VG    Attr  LSize  Origin Snap% Move Log Copy% Convert
drbd-demo VolGroup -wi-a- 1.00G
lv_root  VolGroup -wi-ao  7.30G
lv_swap  VolGroup -wi-ao 500.00M
```

7.3.2. Scrivere la configurazione DRBD

There is no series of commands for building a DRBD configuration, so simply copy the configuration below to /etc/drbd.conf

Informazioni dettagliate sulle direttive utilizzate in questa configurazione (e le alternative disponibili) è disponibile presso *http://www.drbd.org/users-guide/ch-configure.html*

Avvertimento

Be sure to use the names and addresses of your nodes if they differ from the ones used in this guide.

```
global {
 usage-count yes;
}
common {
 protocol C;
}
resource wwwdata {
 meta-disk internal;
 device  /dev/drbd1;
 syncer {
  verify-alg sha1;
 }
 net {
  allow-two-primaries;
 }
 on pcmk-1 {
  disk   /dev/mapper/VolGroup-drbd--demo;
  address  192.168.122.101:7789;
 }
 on pcmk-2 {
  disk   /dev/mapper/VolGroup-drbd--demo;
  address  192.168.122.102:7789;
 }
}
```

Nota

TODO: Spiegare la ragione dell'opzione allow-two-primaries

7.3.3. Inizializzare e caricare DRBD

Sistemata la configurazione, è possibile inizializzare DRBD

```
# drbdadm create-md wwwdata
md_offset 12578816
al_offset 12546048
bm_offset 12541952

Found some data
==> This might destroy existing data! <==

Do you want to proceed?
[need to type 'yes' to confirm] yes
Writing meta data...
initializing activity log
NOT initialized bitmap
```

```
New drbd meta data block successfully created.
success
```

Caricando il modulo DRBD del kernel e verificando che tutto sia a posto

```
# modprobe drbd# drbdadm up wwwdata# cat /proc/drbdversion: 8.3.6 (api:88/proto:86-90)
GIT-hash: f3606c47cc6fcf6b3f086e425cb34af8b7a81bbf build by root@pcmk-1, 2009-12-08 11:22:57
 1: cs:WFConnection ro:Secondary/Unknown ds:Inconsistent/DUnknown C r----
  ns:0 nr:0 dw:0 dr:0 al:0 bm:0 lo:0 pe:0 ua:0 ap:0 ep:1 wo:b oos:12248
```

Repeat on the second node

```
# ssh pcmk-2 -- drbdadm --force create-md wwwdata
Writing meta data...
initializing activity log
NOT initialized bitmap
New drbd meta data block successfully created.
success
# ssh pcmk-2 -- modprobe drbd
WARNING: Deprecated config file /etc/modprobe.conf, all config files belong into /etc/
modprobe.d/.
# ssh pcmk-2 -- drbdadm up wwwdata
# ssh pcmk-2 -- cat /proc/drbd
version: 8.3.6 (api:88/proto:86-90)
GIT-hash: f3606c47cc6fcf6b3f086e425cb34af8b7a81bbf build by root@pcmk-1, 2009-12-08 11:22:57
 1: cs:Connected ro:Secondary/Secondary ds:Inconsistent/Inconsistent C r----
    ns:0 nr:0 dw:0 dr:0 al:0 bm:0 lo:0 pe:0 ua:0 ap:0 ep:1 wo:b oos:12248
```

Ora va indicato a DRBD quale set di dati utilizzare. Dal momento che entrambe le parti contengono dati non necessari è possibile lanciare il seguente comando su pcmk-1:

```
# drbdadm -- --overwrite-data-of-peer primary wwwdata
# cat /proc/drbd
version: 8.3.6 (api:88/proto:86-90)
GIT-hash: f3606c47cc6fcf6b3f086e425cb34af8b7a81bbf build by root@pcmk-1, 2009-12-08 11:22:57
1: cs:SyncSource ro:Primary/Secondary ds:UpToDate/Inconsistent C r----
   ns:2184 nr:0 dw:0 dr:2472 al:0 bm:0 lo:0 pe:0 ua:0 ap:0 ep:1 wo:b oos:10064
    [=====>..............] sync'ed: 33.4% (10064/12248)K
    finish: 0:00:37 speed: 240 (240) K/sec
# cat /proc/drbd
version: 8.3.6 (api:88/proto:86-90)
GIT-hash: f3606c47cc6fcf6b3f086e425cb34af8b7a81bbf build by root@pcmk-1, 2009-12-08 11:22:57
1: cs:Connected ro:Primary/Secondary ds:UpToDate/UpToDate C r----
   ns:12248 nr:0 dw:0 dr:12536 al:0 bm:1 lo:0 pe:0 ua:0 ap:0 ep:1 wo:b oos:0
```

pcmk-1 is now in the Primary state which allows it to be written to. Which means it's a good point at which to create a filesystem and populate it with some data to serve up via our WebSite resource.

7.3.4. Popolare DRBD con i dati

```
# mkfs.ext4 /dev/drbd1
mke2fs 1.41.4 (27-Jan-2009)
Filesystem label=
OS type: Linux
Block size=1024 (log=0)
Fragment size=1024 (log=0)
3072 inodes, 12248 blocks
612 blocks (5.00%) reserved for the super user
First data block=1
Maximum filesystem blocks=12582912
2 block groups
```

```
8192 blocks per group, 8192 fragments per group
1536 inodes per group
Superblock backups stored on blocks:
    8193

Writing inode tables: done
Creating journal (1024 blocks): done
Writing superblocks and filesystem accounting information: done

This filesystem will be automatically checked every 26 mounts or
180 days, whichever comes first. Use tune2fs -c or -i to override.
```

Now mount the newly created filesystem so we can create our index file

```
# mount /dev/drbd1 /mnt/
# cat <<-END >/mnt/index.html
 <html>
  <body>My Test Site - drbd</body>
 </html>
 END
# umount /dev/drbd1
```

7.4. Configurare il cluster per DRBD

Un comoda funzione della shell crm è quella di rendere interattivi comandi per effettuare diversi cambiamenti automatici.

Inizialmente va lanciata la shell. Il prompt cambierà, indicando l'accesso alla modalità interattiva.

```
# crm cib
crm(live) #
```

Next we must create a working copy of the current configuration. This is where all our changes will go. The cluster will not see any of them until we say it’s ok. Notice again how the prompt changes, this time to indicate that we’re no longer looking at the live cluster.

```
cib crm(live) # cib new drbd
INFO: drbd shadow CIB created
crm(drbd) #
```

Ora è possibile creare il clone DRBD e visualizzare la configurazione revisionata.

```
crm(drbd) # configure primitive WebData ocf:linbit:drbd params drbd_resource=wwwdata \
    op monitor interval=60s
crm(drbd) # configure ms WebDataClone WebData meta master-max=1 master-node-max=1 \
    clone-max=2 clone-node-max=1 notify=truecrm(drbd) # configure shownode pcmk-1
node pcmk-2primitive WebData ocf:linbit:drbd \
 params drbd_resource="wwwdata" \
 op monitor interval="60s"primitive WebSite ocf:heartbeat:apache \
    params configfile="/etc/httpd/conf/httpd.conf" \
    op monitor interval="1min"
primitive ClusterIP ocf:heartbeat:IPaddr2 \
    params ip="192.168.122.101" cidr_netmask="32" \
    op monitor interval="30s"ms WebDataClone WebData \
 meta master-max="1" master-node-max="1" clone-max="2" clone-node-max="1" notify="true"
location prefer-pcmk-1 WebSite 50: pcmk-1
colocation website-with-ip inf: WebSite ClusterIP
order apache-after-ip inf: ClusterIP WebSite
property $id="cib-bootstrap-options" \
    dc-version="1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f" \
    cluster-infrastructure="openais" \
    expected-quorum-votes="2" \
    stonith-enabled="false" \
    no-quorum-policy="ignore"
rsc_defaults $id="rsc-options" \
    resource-stickiness="100"
```

Quando i cambiamenti saranno giudicati sufficienti sarà possibile indicare al cluster di iniziare ad usarli, ed utilizzare crm_mon per verificare che tutto stia funzionando.

```
crm(drbd) # cib commit drbdINFO: commited 'drbd' shadow CIB to the cluster
crm(drbd) # quitbye
# crm_mon
============
Last updated: Tue Sep 1 09:37:13 2009
Stack: openais
Current DC: pcmk-1 - partition with quorum
Version: 1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f
2 Nodes configured, 2 expected votes
3 Resources configured.
============

Online: [ pcmk-1 pcmk-2 ]

ClusterIP    (ocf::heartbeat:IPaddr):    Started pcmk-1
WebSite (ocf::heartbeat:apache):    Started pcmk-1Master/Slave Set: WebDataClone Masters: [
 pcmk-2 ] Slaves: [ pcmk-1 ]
```

Nota

Dettagli su come aggiungere una seconda risorsa DRBD

Ora che DRBD sta funzionando è possibile configurare una risorsa filesystem per utilizzarlo. In aggiunta alla definizione del filesystem si necessita inoltre di indicare al cluster dove questa sia localizzata (cioè solo sul nodo Primary di DRBD) e quando ad essa è consentito avviarsi (solo dopo che il nodo Primary è stato attivato).

Ancora una volta verrà utilizzata la shell in modalità interattiva

```
# crm
crm(live) # cib new fs
INFO: fs shadow CIB created
crm(fs) # configure primitive WebFS ocf:heartbeat:Filesystem \
    params device="/dev/drbd/by-res/wwwdata" directory="/var/www/html" fstype="ext4"
crm(fs) # configure colocation fs_on_drbd inf: WebFS WebDataClone:Master
crm(fs) # configure order WebFS-after-WebData inf: WebDataClone:promote WebFS:start

We also need to tell the cluster that Apache needs to run on the same
machine as the filesystem and that it must be active before Apache can
start.

crm(fs) # configure colocation WebSite-with-WebFS inf: WebSite WebFS
crm(fs) # configure order WebSite-after-WebFS inf: WebFS WebSite
```

E' tempo di revisionare la configurazione aggiornata:

```
crm(fs) # crm configure show
```

```
node pcmk-1
node pcmk-2
primitive WebData ocf:linbit:drbd \
    params drbd_resource="wwwdata" \
    op monitor interval="60s"
primitive WebFS ocf:heartbeat:Filesystem \
    params device="/dev/drbd/by-res/wwwdata" directory="/var/www/html" fstype="ext4"
primitive WebSite ocf:heartbeat:apache \
    params configfile="/etc/httpd/conf/httpd.conf" \
    op monitor interval="1min"
primitive ClusterIP ocf:heartbeat:IPaddr2 \
    params ip="192.168.122.101" cidr_netmask="32" \
    op monitor interval="30s"
ms WebDataClone WebData \
    meta master-max="1" master-node-max="1" clone-max="2" clone-node-max="1" notify="true"
location prefer-pcmk-1 WebSite 50: pcmk-1
colocation WebSite-with-WebFS inf: WebSite WebFS
colocation fs_on_drbd inf: WebFS WebDataClone:Master
colocation website-with-ip inf: WebSite ClusterIP
order WebFS-after-WebData inf: WebDataClone:promote WebFS:start
order WebSite-after-WebFS inf: WebFS WebSite
order apache-after-ip inf: ClusterIP WebSite
property $id="cib-bootstrap-options" \
    dc-version="1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f" \
    cluster-infrastructure="openais" \
    expected-quorum-votes="2" \
    stonith-enabled="false" \
    no-quorum-policy="ignore"
rsc_defaults $id="rsc-options" \
    resource-stickiness="100"
```

Dopo la revisione della nuova configurazione questa va caricata e si potrà osservarla in azione all'interno del cluster.

```
crm(fs) # cib commit fs
INFO: commited 'fs' shadow CIB to the cluster
crm(fs) # quit
bye
# crm_mon
============
Last updated: Tue Sep 1 10:08:44 2009
Stack: openais
Current DC: pcmk-1 - partition with quorum
Version: 1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f
2 Nodes configured, 2 expected votes
4 Resources configured.
============

Online: [ pcmk-1 pcmk-2 ]

ClusterIP    (ocf::heartbeat:IPaddr):    Started pcmk-1
WebSite (ocf::heartbeat:apache): Started pcmk-1
Master/Slave Set: WebDataClone
    Masters: [ pcmk-1 ]
    Slaves: [ pcmk-2 ]
WebFS (ocf::heartbeat:Filesystem): Started pcmk-1
```

7.4.1. Testare la migrazione

We could shut down the active node again, but another way to safely simulate recovery is to put the node into what is called "standby mode". Nodes in this state tell the cluster that they are not allowed to run resources. Any resources found active there will be moved elsewhere. This feature can be particularly useful when updating the resources' packages.

Mettendo il nodo locale in standby consentirà di osservare il cluster spostare tutte le risorse sull'altro nodo. Da notare inoltre come lo stato del nodo cambierà, indicando che non può più ospitare risorse.

```
# crm node standby
# crm_mon
============
Last updated: Tue Sep 1 10:09:57 2009
Stack: openais
Current DC: pcmk-1 - partition with quorum
Version: 1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f
2 Nodes configured, 2 expected votes
4 Resources configured.
============
Node pcmk-1: standbyOnline: [ pcmk-2 ]

ClusterIP    (ocf::heartbeat:IPaddr):    Started pcmk-2
WebSite (ocf::heartbeat:apache):    Started pcmk-2
Master/Slave Set: WebDataClone
    Masters: [ pcmk-2 ]    Stopped: [ WebData:1 ]
WebFS  (ocf::heartbeat:Filesystem):  Started pcmk-2
```

Fatto quanto necessario su pcmk-1 (in questo caso nulla, si è solo osservato le risorse spostarsi), è possibile consentire il nodo ad essere nuovamente parte integrante del cluster.

```
# crm node online
# crm_mon
============
Last updated: Tue Sep 1 10:13:25 2009
Stack: openais
Current DC: pcmk-1 - partition with quorum
Version: 1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f
2 Nodes configured, 2 expected votes
4 Resources configured.
============
Online: [ pcmk-1 pcmk-2 ]
ClusterIP    (ocf::heartbeat:IPaddr):    Started pcmk-2
WebSite (ocf::heartbeat:apache):    Started pcmk-2
Master/Slave Set: WebDataClone
    Masters: [ pcmk-2 ]
    Slaves: [ pcmk-1 ]
WebFS  (ocf::heartbeat:Filesystem):  Started pcmk-2
```

Notare come ora l'impostazione di stickiness delle risorse prevenga la migrazione dei servizi nuovamente verso pcmk-1.

Conversione in Active/Active

Indice

8.1. Requisiti

The primary requirement for an Active/Active cluster is that the data required for your services is available, simultaneously, on both machines. Pacemaker makes no requirement on how this is achieved, you could use a SAN if you had one available, however since DRBD supports multiple Primaries, we can also use that.

The only hitch is that we need to use a cluster-aware filesystem. The one we used earlier with DRBD, ext4, is not one of those. Both OCFS2 and GFS2 are supported, however here we will use GFS2 which comes with Fedora.

We'll also need to use CMAN for Cluster Membership and Quorum instead of our Corosync plugin.

8.2. Adding CMAN Support

CMAN v3[1] is a Corosync plugin that monitors the names and number of active cluster nodes in order to deliver membership and quorum information to clients (such as the Pacemaker daemons).

In a traditional Corosync-Pacemaker cluster, a Pacemaker plugin is loaded to provide membership and quorum information. The motivation for wanting to use CMAN for this instead, is to ensure all elements of the cluster stack are making decisions based on the same membership and quorum data. [2]

In the case of GFS2, the key pieces are the dlm_controld and gfs_controld helpers which act as the glue between the filesystem and the cluster software. Supporting CMAN enables us to use the versions already being shipped by most distributions (since CMAN has been around longer than Pacemaker and is part of the Red Hat cluster stack).

[1] http://docs.redhat.com/docs/en-US/Red_Hat_Enterprise_Linux/6/html-single/Cluster_Suite_Overview/index.html#s2-clumembership-overview-CSO

[2] A failure to do this can lead to what is called *internal split-brain* - a situation where different parts of the stack disagree about whether some nodes are alive or dead - which quickly leads to unnecessary down-time and/or data corruption.

Avvertimento

Ensure Corosync and Pacemaker are stopped on all nodes before continuing

Avvertimento

Be sure to disable the Pacemaker plugin before continuing with this section. In most cases, this can be achieved by removing /etc/corosync/service.d/pcmk and stopping Corosync.

8.2.1. Installing the required Software

```
# yum install -y cman gfs2-utils gfs2-cluster
Loaded plugins: auto-update-debuginfo
Setting up Install Process
Resolving Dependencies
--> Running transaction check
---> Package cman.x86_64 0:3.1.7-1.fc15 will be installed
--> Processing Dependency: modcluster >= 0.18.1-1 for package: cman-3.1.7-1.fc15.x86_64
--> Processing Dependency: fence-agents >= 3.1.5-1 for package: cman-3.1.7-1.fc15.x86_64
--> Processing Dependency: openais >= 1.1.4-1 for package: cman-3.1.7-1.fc15.x86_64
--> Processing Dependency: ricci >= 0.18.1-1 for package: cman-3.1.7-1.fc15.x86_64
--> Processing Dependency: libSaCkpt.so.3(OPENAIS_CKPT_B.01.01)(64bit) for package:
 cman-3.1.7-1.fc15.x86_64
--> Processing Dependency: libSaCkpt.so.3()(64bit) for package: cman-3.1.7-1.fc15.x86_64
---> Package gfs2-cluster.x86_64 0:3.1.1-2.fc15 will be installed
---> Package gfs2-utils.x86_64 0:3.1.1-2.fc15 will be installed
--> Running transaction check
---> Package fence-agents.x86_64 0:3.1.5-1.fc15 will be installed
--> Processing Dependency: /usr/bin/virsh for package: fence-agents-3.1.5-1.fc15.x86_64
--> Processing Dependency: net-snmp-utils for package: fence-agents-3.1.5-1.fc15.x86_64
--> Processing Dependency: sg3_utils for package: fence-agents-3.1.5-1.fc15.x86_64
--> Processing Dependency: perl(Net::Telnet) for package: fence-agents-3.1.5-1.fc15.x86_64
--> Processing Dependency: /usr/bin/ipmitool for package: fence-agents-3.1.5-1.fc15.x86_64
--> Processing Dependency: perl-Net-Telnet for package: fence-agents-3.1.5-1.fc15.x86_64
--> Processing Dependency: pexpect for package: fence-agents-3.1.5-1.fc15.x86_64
--> Processing Dependency: pyOpenSSL for package: fence-agents-3.1.5-1.fc15.x86_64
--> Processing Dependency: python-suds for package: fence-agents-3.1.5-1.fc15.x86_64
---> Package modcluster.x86_64 0:0.18.7-1.fc15 will be installed
--> Processing Dependency: oddjob for package: modcluster-0.18.7-1.fc15.x86_64
---> Package openais.x86_64 0:1.1.4-2.fc15 will be installed
---> Package openaislib.x86_64 0:1.1.4-2.fc15 will be installed
---> Package ricci.x86_64 0:0.18.7-1.fc15 will be installed
--> Processing Dependency: parted for package: ricci-0.18.7-1.fc15.x86_64
--> Processing Dependency: nss-tools for package: ricci-0.18.7-1.fc15.x86_64
--> Running transaction check
---> Package ipmitool.x86_64 0:1.8.11-6.fc15 will be installed
---> Package libvirt-client.x86_64 0:0.8.8-7.fc15 will be installed
--> Processing Dependency: libnetcf.so.1(NETCF_1.3.0)(64bit) for package: libvirt-
client-0.8.8-7.fc15.x86_64
--> Processing Dependency: cyrus-sasl-md5 for package: libvirt-client-0.8.8-7.fc15.x86_64
--> Processing Dependency: gettext for package: libvirt-client-0.8.8-7.fc15.x86_64
--> Processing Dependency: nc for package: libvirt-client-0.8.8-7.fc15.x86_64
--> Processing Dependency: libnuma.so.1(libnuma_1.1)(64bit) for package: libvirt-
client-0.8.8-7.fc15.x86_64
```

```
--> Processing Dependency: libnuma.so.1(libnuma_1.2)(64bit) for package: libvirt-
client-0.8.8-7.fc15.x86_64
--> Processing Dependency: libnetcf.so.1(NETCF_1.2.0)(64bit) for package: libvirt-
client-0.8.8-7.fc15.x86_64
--> Processing Dependency: gnutls-utils for package: libvirt-client-0.8.8-7.fc15.x86_64
--> Processing Dependency: libnetcf.so.1(NETCF_1.0.0)(64bit) for package: libvirt-
client-0.8.8-7.fc15.x86_64
--> Processing Dependency: libxenstore.so.3.0()(64bit) for package: libvirt-
client-0.8.8-7.fc15.x86_64
--> Processing Dependency: libyajl.so.1()(64bit) for package: libvirt-
client-0.8.8-7.fc15.x86_64
--> Processing Dependency: libnl.so.1()(64bit) for package: libvirt-
client-0.8.8-7.fc15.x86_64
--> Processing Dependency: libnuma.so.1()(64bit) for package: libvirt-
client-0.8.8-7.fc15.x86_64
--> Processing Dependency: libaugeas.so.0()(64bit) for package: libvirt-
client-0.8.8-7.fc15.x86_64
--> Processing Dependency: libnetcf.so.1()(64bit) for package: libvirt-
client-0.8.8-7.fc15.x86_64
---> Package net-snmp-utils.x86_64 1:5.6.1-7.fc15 will be installed
---> Package nss-tools.x86_64 0:3.12.10-6.fc15 will be installed
---> Package oddjob.x86_64 0:0.31-2.fc15 will be installed
---> Package parted.x86_64 0:2.3-10.fc15 will be installed
---> Package perl-Net-Telnet.noarch 0:3.03-12.fc15 will be installed
---> Package pexpect.noarch 0:2.3-6.fc15 will be installed
---> Package pyOpenSSL.x86_64 0:0.10-3.fc15 will be installed
---> Package python-suds.noarch 0:0.3.9-3.fc15 will be installed
---> Package sg3_utils.x86_64 0:1.29-3.fc15 will be installed
--> Processing Dependency: sg3_utils-libs = 1.29-3.fc15 for package:
 sg3_utils-1.29-3.fc15.x86_64
--> Processing Dependency: libsgutils2.so.2()(64bit) for package:
 sg3_utils-1.29-3.fc15.x86_64
--> Running transaction check
---> Package augeas-libs.x86_64 0:0.9.0-1.fc15 will be installed
---> Package cyrus-sasl-md5.x86_64 0:2.1.23-18.fc15 will be installed
---> Package gettext.x86_64 0:0.18.1.1-7.fc15 will be installed
--> Processing Dependency: libgomp.so.1(GOMP_1.0)(64bit) for
 package: gettext-0.18.1.1-7.fc15.x86_64
--> Processing Dependency: libgettextlib-0.18.1.so()(64bit) for
 package: gettext-0.18.1.1-7.fc15.x86_64
--> Processing Dependency: libgettextsrc-0.18.1.so()(64bit) for
 package: gettext-0.18.1.1-7.fc15.x86_64
--> Processing Dependency: libgomp.so.1()(64bit) for package: gettext-0.18.1.1-7.fc15.x86_64
---> Package gnutls-utils.x86_64 0:2.10.5-1.fc15 will be installed
---> Package libnl.x86_64 0:1.1-14.fc15 will be installed
---> Package nc.x86_64 0:1.100-3.fc15 will be installed
--> Processing Dependency: libbsd.so.0(LIBBSD_0.0)(64bit) for package: nc-1.100-3.fc15.x86_64
--> Processing Dependency: libbsd.so.0(LIBBSD_0.2)(64bit) for package: nc-1.100-3.fc15.x86_64
--> Processing Dependency: libbsd.so.0()(64bit) for package: nc-1.100-3.fc15.x86_64
---> Package netcf-libs.x86_64 0:0.1.9-1.fc15 will be installed
---> Package numactl.x86_64 0:2.0.7-1.fc15 will be installed
---> Package sg3_utils-libs.x86_64 0:1.29-3.fc15 will be installed
---> Package xen-libs.x86_64 0:4.1.1-3.fc15 will be installed
--> Processing Dependency: xen-licenses for package: xen-libs-4.1.1-3.fc15.x86_64
---> Package yajl.x86_64 0:1.0.11-1.fc15 will be installed
--> Running transaction check
---> Package gettext-libs.x86_64 0:0.18.1.1-7.fc15 will be installed
---> Package libbsd.x86_64 0:0.2.0-4.fc15 will be installed
---> Package libgomp.x86_64 0:4.6.1-9.fc15 will be installed
---> Package xen-licenses.x86_64 0:4.1.1-3.fc15 will be installed
--> Finished Dependency Resolution

Dependencies Resolved

=============================================================================
 Package              Arch        Version                 Repository    Size
=============================================================================
Installing:
 cman                 x86_64      3.1.7-1.fc15            updates      366 k
 gfs2-cluster         x86_64      3.1.1-2.fc15            fedora        69 k
 gfs2-utils           x86_64      3.1.1-2.fc15            fedora       222 k
Installing for dependencies:
 augeas-libs          x86_64      0.9.0-1.fc15            updates      311 k
 cyrus-sasl-md5       x86_64      2.1.23-18.fc15          updates       46 k
 fence-agents         x86_64      3.1.5-1.fc15            updates      186 k
 gettext              x86_64      0.18.1.1-7.fc15         fedora       1.0 M
 gettext-libs         x86_64      0.18.1.1-7.fc15         fedora       610 k
 gnutls-utils         x86_64      2.10.5-1.fc15           fedora       101 k
 ipmitool             x86_64      1.8.11-6.fc15           fedora       273 k
 libbsd               x86_64      0.2.0-4.fc15            fedora        37 k
 libgomp              x86_64      4.6.1-9.fc15            updates       95 k
 libnl                x86_64      1.1-14.fc15             fedora       118 k
 libvirt-client       x86_64      0.8.8-7.fc15            updates      2.4 M
 modcluster           x86_64      0.18.7-1.fc15           fedora       187 k
 nc                   x86_64      1.100-3.fc15            updates       24 k
 net-snmp-utils       x86_64      1:5.6.1-7.fc15          fedora       180 k
 netcf-libs           x86_64      0.1.9-1.fc15            updates       50 k
 nss-tools            x86_64      3.12.10-6.fc15          updates      723 k
 numactl              x86_64      2.0.7-1.fc15            updates       54 k
 oddjob               x86_64      0.31-2.fc15             fedora        61 k
 openais              x86_64      1.1.4-2.fc15            fedora       190 k
 openaislib           x86_64      1.1.4-2.fc15            fedora        88 k
 parted               x86_64      2.3-10.fc15             updates      618 k
 perl-Net-Telnet      noarch      3.03-12.fc15            fedora        55 k
 pexpect              noarch      2.3-6.fc15              fedora       141 k
 pyOpenSSL            x86_64      0.10-3.fc15             fedora       198 k
 python-suds          noarch      0.3.9-3.fc15            fedora       195 k
 ricci                x86_64      0.18.7-1.fc15           fedora       584 k
 sg3_utils            x86_64      1.29-3.fc15             fedora       465 k
 sg3_utils-libs       x86_64      1.29-3.fc15             fedora        54 k
 xen-libs             x86_64      4.1.1-3.fc15            updates      310 k
 xen-licenses         x86_64      4.1.1-3.fc15            updates       64 k
 yajl                 x86_64      1.0.11-1.fc15           fedora        27 k

Transaction Summary
=============================================================================
Install      34 Package(s)

Total download size: 10 M
Installed size: 38 M
Downloading Packages:
(1/34): augeas-libs-0.9.0-1.fc15.x86_64.rpm           | 311 kB     00:00
(2/34): cman-3.1.7-1.fc15.x86_64.rpm                  | 366 kB     00:00
(3/34): cyrus-sasl-md5-2.1.23-18.fc15.x86_64.rpm      |  46 kB     00:00
(4/34): fence-agents-3.1.5-1.fc15.x86_64.rpm          | 186 kB     00:00
(5/34): gettext-0.18.1.1-7.fc15.x86_64.rpm            | 1.0 MB     00:01
(6/34): gettext-libs-0.18.1.1-7.fc15.x86_64.rpm       | 610 kB     00:00
(7/34): gfs2-cluster-3.1.1-2.fc15.x86_64.rpm          |  69 kB     00:00
(8/34): gfs2-utils-3.1.1-2.fc15.x86_64.rpm            | 222 kB     00:00
(9/34): gnutls-utils-2.10.5-1.fc15.x86_64.rpm         | 101 kB     00:00
(10/34): ipmitool-1.8.11-6.fc15.x86_64.rpm            | 273 kB     00:00
(11/34): libbsd-0.2.0-4.fc15.x86_64.rpm               |  37 kB     00:00
(12/34): libgomp-4.6.1-9.fc15.x86_64.rpm              |  95 kB     00:00
(13/34): libnl-1.1-14.fc15.x86_64.rpm                 | 118 kB     00:00
(14/34): libvirt-client-0.8.8-7.fc15.x86_64.rpm       | 2.4 MB     00:01
(15/34): modcluster-0.18.7-1.fc15.x86_64.rpm          | 187 kB     00:00
(16/34): nc-1.100-3.fc15.x86_64.rpm                   |  24 kB     00:00
(17/34): net-snmp-utils-5.6.1-7.fc15.x86_64.rpm       | 180 kB     00:00
(18/34): netcf-libs-0.1.9-1.fc15.x86_64.rpm           |  50 kB     00:00
(19/34): nss-tools-3.12.10-6.fc15.x86_64.rpm          | 723 kB     00:00
(20/34): numactl-2.0.7-1.fc15.x86_64.rpm              |  54 kB     00:00
(21/34): oddjob-0.31-2.fc15.x86_64.rpm                |  61 kB     00:00
(22/34): openais-1.1.4-2.fc15.x86_64.rpm              | 190 kB     00:00
(23/34): openaislib-1.1.4-2.fc15.x86_64.rpm           |  88 kB     00:00
```

```
(24/34): parted-2.3-10.fc15.x86_64.rpm                | 618 kB     00:00
(25/34): perl-Net-Telnet-3.03-12.fc15.noarch.rpm      |  55 kB     00:00
(26/34): pexpect-2.3-6.fc15.noarch.rpm                | 141 kB     00:00
(27/34): pyOpenSSL-0.10-3.fc15.x86_64.rpm             | 198 kB     00:00
(28/34): python-suds-0.3.9-3.fc15.noarch.rpm          | 195 kB     00:00
(29/34): ricci-0.18.7-1.fc15.x86_64.rpm               | 584 kB     00:00
(30/34): sg3_utils-1.29-3.fc15.x86_64.rpm             | 465 kB     00:00
(31/34): sg3_utils-libs-1.29-3.fc15.x86_64.rpm        |  54 kB     00:00
(32/34): xen-libs-4.1.1-3.fc15.x86_64.rpm             | 310 kB     00:00
(33/34): xen-licenses-4.1.1-3.fc15.x86_64.rpm         |  64 kB     00:00
(34/34): yajl-1.0.11-1.fc15.x86_64.rpm                |  27 kB     00:00
 -----------------------------------------------------------------------------
Total                                        803 kB/s |  10 MB     00:12
Running rpm_check_debug
Running Transaction Test
Transaction Test Succeeded
Running Transaction
  Installing : openais-1.1.4-2.fc15.x86_64                              1/34
  Installing : openaislib-1.1.4-2.fc15.x86_64                           2/34
  Installing : libnl-1.1-14.fc15.x86_64                                 3/34
  Installing : augeas-libs-0.9.0-1.fc15.x86_64                          4/34
  Installing : oddjob-0.31-2.fc15.x86_64                                5/34
  Installing : modcluster-0.18.7-1.fc15.x86_64                          6/34
  Installing : netcf-libs-0.1.9-1.fc15.x86_64                           7/34
  Installing : 1:net-snmp-utils-5.6.1-7.fc15.x86_64                     8/34
  Installing : sg3_utils-libs-1.29-3.fc15.x86_64                        9/34
  Installing : sg3_utils-1.29-3.fc15.x86_64                            10/34
  Installing : libgomp-4.6.1-9.fc15.x86_64                             11/34
  Installing : gnutls-utils-2.10.5-1.fc15.x86_64                       12/34
  Installing : pyOpenSSL-0.10-3.fc15.x86_64                            13/34
  Installing : parted-2.3-10.fc15.x86_64                               14/34
  Installing : cyrus-sasl-md5-2.1.23-18.fc15.x86_64                    15/34
  Installing : python-suds-0.3.9-3.fc15.noarch                         16/34
  Installing : ipmitool-1.8.11-6.fc15.x86_64                           17/34
  Installing : perl-Net-Telnet-3.03-12.fc15.noarch                     18/34
  Installing : numactl-2.0.7-1.fc15.x86_64                             19/34
  Installing : yajl-1.0.11-1.fc15.x86_64                               20/34
  Installing : gettext-libs-0.18.1.1-7.fc15.x86_64                     21/34
  Installing : gettext-0.18.1.1-7.fc15.x86_64                          22/34
  Installing : libbsd-0.2.0-4.fc15.x86_64                              23/34
  Installing : nc-1.100-3.fc15.x86_64                                  24/34
  Installing : xen-licenses-4.1.1-3.fc15.x86_64                        25/34
  Installing : xen-libs-4.1.1-3.fc15.x86_64                            26/34
  Installing : libvirt-client-0.8.8-7.fc15.x86_64                      27/34

Note: This output shows SysV services only and does not include native
      systemd services. SysV configuration data might be overridden by native
      systemd configuration.

  Installing : nss-tools-3.12.10-6.fc15.x86_64                         28/34
  Installing : ricci-0.18.7-1.fc15.x86_64                              29/34
  Installing : pexpect-2.3-6.fc15.noarch                               30/34
  Installing : fence-agents-3.1.5-1.fc15.x86_64                        31/34
  Installing : cman-3.1.7-1.fc15.x86_64                                32/34
  Installing : gfs2-cluster-3.1.1-2.fc15.x86_64                        33/34
  Installing : gfs2-utils-3.1.1-2.fc15.x86_64                          34/34

Installed:
  cman.x86_64 0:3.1.7-1.fc15           gfs2-cluster.x86_64 0:3.1.1-2.fc15
  gfs2-utils.x86_64 0:3.1.1-2.fc15

Dependency Installed:
  augeas-libs.x86_64 0:0.9.0-1.fc15
  cyrus-sasl-md5.x86_64 0:2.1.23-18.fc15
  fence-agents.x86_64 0:3.1.5-1.fc15
  gettext.x86_64 0:0.18.1.1-7.fc15
  gettext-libs.x86_64 0:0.18.1.1-7.fc15
  gnutls-utils.x86_64 0:2.10.5-1.fc15
  ipmitool.x86_64 0:1.8.11-6.fc15
  libbsd.x86_64 0:0.2.0-4.fc15
  libgomp.x86_64 0:4.6.1-9.fc15
  libnl.x86_64 0:1.1-14.fc15
  libvirt-client.x86_64 0:0.8.8-7.fc15
  modcluster.x86_64 0:0.18.7-1.fc15
  nc.x86_64 0:1.100-3.fc15
  net-snmp-utils.x86_64 1:5.6.1-7.fc15
  netcf-libs.x86_64 0:0.1.9-1.fc15
  nss-tools.x86_64 0:3.12.10-6.fc15
  numactl.x86_64 0:2.0.7-1.fc15
  oddjob.x86_64 0:0.31-2.fc15
  openais.x86_64 0:1.1.4-2.fc15
  openaislib.x86_64 0:1.1.4-2.fc15
  parted.x86_64 0:2.3-10.fc15
  perl-Net-Telnet.noarch 0:3.03-12.fc15
  pexpect.noarch 0:2.3-6.fc15
  pyOpenSSL.x86_64 0:0.10-3.fc15
  python-suds.noarch 0:0.3.9-3.fc15
  ricci.x86_64 0:0.18.7-1.fc15
  sg3_utils.x86_64 0:1.29-3.fc15
  sg3_utils-libs.x86_64 0:1.29-3.fc15
  xen-libs.x86_64 0:4.1.1-3.fc15
  xen-licenses.x86_64 0:4.1.1-3.fc15
  yajl.x86_64 0:1.0.11-1.fc15

Complete!
```

8.2.2. Configuring CMAN

Nota

The standard Pacemaker config file will continue to be used for resource management even after we start using CMAN. There is no need to recreate all your resources and constraints to the *cluster.conf* syntax, we simply create a minimal version that lists the nodes.

The first thing we need to do, is tell CMAN complete starting up even without quorum. We can do this by changing the quorum timeout setting:

```
# sed -i.sed "s/.*CMAN_QUORUM_TIMEOUT=.*/CMAN_QUORUM_TIMEOUT=0/g" /etc/sysconfig/cman
```

Next we create a basic configuration file and place it in /etc/cluster/cluster.conf. The name used for each clusternode should correspond to that node’s uname -n, just as Pacemaker expects. The nodeid can be any positive mumber but must be unique.

Basic cluster.conf for a two-node cluster

```
<?xml version="1.0"?>
<cluster config_version="1" name="my_cluster_name">
  <logging debug="off"/>
  <clusternodes>
    <clusternode name="pcmk-1" nodeid="1"/>
    <clusternode name="pcmk-2" nodeid="2"/>
  </clusternodes>
```

```
</cluster>
```

8.2.3. Redundant Rings

For those wishing to use Corosync's multiple rings feature, simply define an alternate name for each node. For example:

```
    <clusternode name="pcmk-1" nodeid="1"/>
        <altname name="pcmk-1-internal"/>
    </clusternode>
```

8.2.4. Configuring CMAN Fencing

We configure the fence_pcmk agent (supplied with Pacemaker) to redirect any fencing requests from CMAN components (such as dlm_controld) to Pacemaker. Pacemaker's fencing subsystem lets other parts of the stack know that a node has been successfully fenced, thus avoiding the need for it to be fenced again when other subsystems notice the node is gone.

Avvertimento

Configuring real fencing devices in CMAN will result in nodes being fenced multiple times as different parts of the stack notice the node is missing or failed.

The definition should be placed in the fencedevices section and contain:

```
<fencedevice name="pcmk" agent="fence_pcmk"/>
```

Each clusternode must be configured to use this device by adding a fence method block that lists the node's name as the port.

```
<fence>
  <method name="pcmk-redirect">
    <device name="pcmk" port="node_name_here"/>
  </method>
</fence>
```

Putting everything together, we have:

cluster.conf for a two-node cluster with fencing

```
<?xml version="1.0"?>
<cluster config_version="1" name="mycluster">
  <logging debug="off"/>
  <clusternodes>
    <clusternode name="pcmk-1" nodeid="1">
      <fence>
        <method name="pcmk-redirect">
          <device name="pcmk" port="pcmk-1"/>
        </method>
      </fence>
    </clusternode>
    <clusternode name="pcmk-2" nodeid="2">
```

```
      <fence>
        <method name="pcmk-redirect">
          <device name="pcmk" port="pcmk-2"/>
        </method>
      </fence>
    </clusternode>
  </clusternodes>
  <fencedevices>
    <fencedevice name="pcmk" agent="fence_pcmk"/>
  </fencedevices>
</cluster>
```

8.2.5. Bringing the Cluster Online with CMAN

The first thing to do is check that the configuration is valid

```
# ccs_config_validate
Configuration validates
```

Now start CMAN

```
# service cman start
Starting cluster:
   Checking Network Manager...                            [  OK  ]
   Global setup...                                        [  OK  ]
   Loading kernel modules...                              [  OK  ]
   Mounting configfs...                                   [  OK  ]
   Starting cman...                                       [  OK  ]
   Waiting for quorum...                                  [  OK  ]
   Starting fenced...                                     [  OK  ]
   Starting dlm_controld...                               [  OK  ]
   Starting gfs_controld...                               [  OK  ]
   Unfencing self...                                      [  OK  ]
   Joining fence domain...                                [  OK  ]
```

Once you have confirmed that the first node is happily online, start the second node.

```
[root@pcmk-2 ~]# service cman start
Starting cluster:
   Checking Network Manager...                            [  OK  ]
   Global setup...                                        [  OK  ]
   Loading kernel modules...                              [  OK  ]
   Mounting configfs...                                   [  OK  ]
   Starting cman...                                       [  OK  ]
   Waiting for quorum...                                  [  OK  ]
   Starting fenced...                                     [  OK  ]
   Starting dlm_controld...                               [  OK  ]
   Starting gfs_controld...                               [  OK  ]
   Unfencing self...                                      [  OK  ]
   Joining fence domain...                                [  OK  ]
# cman_tool nodes
Node  Sts   Inc   Joined               Name
   1   M    548   2011-09-28 10:52:21  pcmk-1
   2   M    548   2011-09-28 10:52:21  pcmk-2
```

You should now see both nodes online. To begin managing resources, simply start Pacemaker.

```
# service pacemaker start
 Starting Pacemaker Cluster Manager: [  OK  ]
```

and again on the second node, after which point you can use crm_mon as you normally would.

```
[root@pcmk-2 ~]# service pacemaker start
 Starting Pacemaker Cluster Manager: [  OK  ]
# crm_mon -1
```

8.3. Creazione un Filesystem GFS2

8.3.1. Preparazione

Before we do anything to the existing partition, we need to make sure it is unmounted. We do this by telling the cluster to stop the WebFS resource. This will ensure that other resources (in our case, Apache) using WebFS are not only stopped, but stopped in the correct order.

```
# crm_resource --resource WebFS --set-parameter target-role --meta --parameter-value Stopped
# crm_mon
============
Last updated: Thu Sep 3 15:18:06 2009
Stack: openais
Current DC: pcmk-1 - partition with quorum
Version: 1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f
2 Nodes configured, 2 expected votes
6 Resources configured.
============

Online: [ pcmk-1 pcmk-2 ]

Master/Slave Set: WebDataClone
    Masters: [ pcmk-1 ]
    Slaves: [ pcmk-2 ]
ClusterIP    (ocf::heartbeat:IPaddr):    Started pcmk-1
```

Nota

Notare come sia Apache che WebFS sono stati fermati.

8.3.2. Creazione e popolamento della partizione GFS2

Ora che il cluster e le sue parti stanno funzionando senza problemi, è possibile creare la partizione GFS2.

Avvertimento

Questa operazione cancellerà qualsiasi contenuto sul device DRBD. Assicurarsi quindi di avere una copia di qualsiasi dato cruciale.

Nella creazione di una partizione GFS2 è necessario specificare diversi parametri aggiuntivi.

Per cominciare, va utilizzata l'opzione -p per specificare la volontà di utilizzare il DLM del Kernel. Proseguendo, l'opzione -j indica che la risorsa dovrà riservare abbastanza spazio per due journal (uno per ogni nodo che accede al filesystem).

Infine verrà utilizzato -t per specificare il nome della lock table. Il formato di questo campo è clustername:fsname. Per fsname è sufficiente scegliere un nome unico e descrittivo e dato che non è stato specificato ancora clustername è possibile usare il default (pcmk).

To specify an alternate name for the cluster, locate the service section containing **`name: pacemaker`** in corosync.conf and insert the following line anywhere inside the block:

```
clustername: myname
```

Questa operazione andrà eseguita in ciascun nodo, accertandosi di effettuare un restart dello stesso prima di continuare.

```
# mkfs.gfs2 -p lock_dlm -j 2 -t pcmk:web /dev/drbd1
This will destroy any data on /dev/drbd1.
It appears to contain: data

Are you sure you want to proceed? [y/n] y

Device:          /dev/drbd1
Blocksize:         4096
Device Size        1.00 GB (131072 blocks)
Filesystem Size:      1.00 GB (131070 blocks)
Journals:         2
Resource Groups:      2
Locking Protocol:     "lock_dlm"
Lock Table:        "pcmk:web"
UUID:           6B776F46-177B-BAF8-2C2B-292C0E078613
```

A questo punto è possibile ripopolare il nuovo filesystem con i dati (le pagine web). Per ora verrà creata una versione alternativa dell'home page.

```
# mount /dev/drbd1 /mnt/# cat <<-END >/mnt/index.html
<html>
<body>My Test Site - GFS2</body>
</html>
END
# umount /dev/drbd1
# drbdadm verify wwwdata#
```

8.4. Riconfigurare il cluster per GFS2

```
# crm
crm(live) # cib new GFS2
INFO: GFS2 shadow CIB created
crm(GFS2) # configure delete WebFS
crm(GFS2) # configure primitive WebFS ocf:heartbeat:Filesystem params device="/dev/drbd/by-
res/wwwdata" directory="/var/www/html" fstype="gfs2"
```

Ora che abbiamo creato nuovamente la risorsa è possibile ricreare tutte le constraint che questa utilizzava. Questo perché la shell rimuoverà automaticamente qualsiasi constraint che riferisca a WebFS.

```
crm(GFS2) # configure colocation WebSite-with-WebFS inf: WebSite WebFS
crm(GFS2) # configure colocation fs_on_drbd inf: WebFS WebDataClone:Master
crm(GFS2) # configure order WebFS-after-WebData inf: WebDataClone:promote WebFS:start
crm(GFS2) # configure order WebSite-after-WebFS inf: WebFS WebSite
crm(GFS2) # configure show
node pcmk-1
node pcmk-2
primitive WebData ocf:linbit:drbd \
    params drbd_resource="wwwdata" \
    op monitor interval="60s"
primitive WebFS ocf:heartbeat:Filesystem \
    params device="/dev/drbd/by-res/wwwdata" directory="/var/www/html" fstype="gfs2"
primitive WebSite ocf:heartbeat:apache \
    params configfile="/etc/httpd/conf/httpd.conf" \
    op monitor interval="1min"
primitive ClusterIP ocf:heartbeat:IPaddr2 \
    params ip="192.168.122.101" cidr_netmask="32" \
    op monitor interval="30s"
ms WebDataClone WebData \
    meta master-max="1" master-node-max="1" clone-max="2" clone-node-max="1" notify="true"
colocation WebSite-with-WebFS inf: WebSite WebFS
colocation fs_on_drbd inf: WebFS WebDataClone:Master
colocation website-with-ip inf: WebSite ClusterIP
order WebFS-after-WebData inf: WebDataClone:promote WebFS:start
order WebSite-after-WebFS inf: WebFS WebSite
order apache-after-ip inf: ClusterIP WebSite
property $id="cib-bootstrap-options" \
    dc-version="1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f" \
    cluster-infrastructure="openais" \
    expected-quorum-votes="2" \
    stonith-enabled="false" \
    no-quorum-policy="ignore"
rsc_defaults $id="rsc-options" \
    resource-stickiness="100"
```

Prima di caricare la configurazione nel cluster questa va revisionata, la shell andrà terminata e si dovrà osservare il responso del cluster

```
crm(GFS2) # cib commit GFS2
INFO: commited 'GFS2' shadow CIB to the cluster
crm(GFS2) # quit
bye
# crm_mon
============
Last updated: Thu Sep 3 20:49:54 2009
Stack: openais
Current DC: pcmk-2 - partition with quorum
Version: 1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f
2 Nodes configured, 2 expected votes
6 Resources configured.
============

Online: [ pcmk-1 pcmk-2 ]

WebSite (ocf::heartbeat:apache):    Started pcmk-2
Master/Slave Set: WebDataClone
    Masters: [ pcmk-1 ]
    Slaves: [ pcmk-2 ]
ClusterIP    (ocf::heartbeat:IPaddr):    Started pcmk-2WebFS (ocf::heartbeat:Filesystem):
 Started pcmk-1
```

8.5. Riconfigurare Pacemaker per l'Active/Active

Quasi tutto è a posto. Le versioni recenti di DRBD supportano l'operare in modalità Primary/Primary ed inoltre il filesystem utilizzato è cluster aware. Tutto ciò che rimane da fare è configurare nuovamente il cluster per sfruttare queste peculiarità.

Questo richiederà diversi cambiamenti, pertanto verrà utilizzata ancora una volta la modalità interattiva.

# crm # cib new active

There’s no point making the services active on both locations if we can’t reach them, so lets first clone the IP address. Cloned IPaddr2 resources use an iptables rule to ensure that each request only gets processed by one of the two clone instances. The additional meta options tell the cluster how many instances of the clone we want (one "request bucket" for each node) and that if all other nodes fail, then the remaining node should hold all of them. Otherwise the requests would be simply discarded.

```
# configure clone WebIP ClusterIP \
    meta globally-unique="true" clone-max="2" clone-node-max="2"
```

Ora bisogna indicare a ClusterIP come decidere quali richieste sono processate e da quali host. Per fare ciò è necessario definire il parametro clusterip_hash.

Aprire la risorsa ClusterIP

```
# configure edit ClusterIP
```

ed aggiungere il seguente testo alla linea dei parametri

```
clusterip_hash="sourceip"
```

In modo che la definizione completa sia simile alla seguente:

```
primitive ClusterIP ocf:heartbeat:IPaddr2 \
    params ip="192.168.122.101" cidr_netmask="32" clusterip_hash="sourceip" \
    op monitor interval="30s"
```

Questa è la trascrizione completa

```
# crm crm(live)
# cib new active
INFO: active shadow CIB created
crm(active) # configure clone WebIP ClusterIP \
    meta globally-unique="true" clone-max="2" clone-node-max="2"
crm(active) # configure shownode pcmk-1
node pcmk-2
primitive WebData ocf:linbit:drbd \
    params drbd_resource="wwwdata" \
    op monitor interval="60s"
primitive WebFS ocf:heartbeat:Filesystem \
    params device="/dev/drbd/by-res/wwwdata" directory="/var/www/html" fstype="gfs2"
primitive WebSite ocf:heartbeat:apache \
    params configfile="/etc/httpd/conf/httpd.conf" \
    op monitor interval="1min"
primitive ClusterIP ocf:heartbeat:IPaddr2 \
    params ip="192.168.122.101" cidr_netmask="32" clusterip_hash="sourceip" \
    op monitor interval="30s"
ms WebDataClone WebData \
    meta master-max="1" master-node-max="1" clone-max="2" clone-node-max="1" notify="true"
clone WebIP ClusterIP \
    meta globally-unique="true" clone-max="2" clone-node-max="2"
colocation WebSite-with-WebFS inf: WebSite WebFS
colocation fs_on_drbd inf: WebFS WebDataClone:Master
colocation website-with-ip inf: WebSite WebIPorder WebFS-after-WebData inf:
 WebDataClone:promote WebFS:start
order WebSite-after-WebFS inf: WebFS WebSiteorder apache-after-ip inf: WebIP WebSite
property $id="cib-bootstrap-options" \
    dc-version="1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f" \
    cluster-infrastructure="openais" \
    expected-quorum-votes="2" \
    stonith-enabled="false" \
    no-quorum-policy="ignore"
rsc_defaults $id="rsc-options" \
    resource-stickiness="100"
```

Notare come qualsiasi constraint riferita a ClusterIP sia stata aggiornata per utilizzare invece WebIP. Questo è un vantaggio aggiuntivo nell'utilizzo della shell crm.

A questo punto è necessario convertire le risorse filesystem ed Apache in cloni. Di nuovo, la shell aggiornerà automaticamente ogni constraint interessata.

```
crm(active) # configure clone WebFSClone WebFS
crm(active) # configure clone WebSiteClone WebSite
```

L'ultimo passo è quello di comunicare al cluster che ora è consentito promuovere entrambe le istanze a PRimary (o Master).

```
crm(active) # configure edit WebDataClone
```

Cambiare master-max a 2

```
crm(active) # configure show
node pcmk-1
node pcmk-2
primitive WebData ocf:linbit:drbd \
    params drbd_resource="wwwdata" \
    op monitor interval="60s"
primitive WebFS ocf:heartbeat:Filesystem \
    params device="/dev/drbd/by-res/wwwdata" directory="/var/www/html" fstype="gfs2"
primitive WebSite ocf:heartbeat:apache \
    params configfile="/etc/httpd/conf/httpd.conf" \
    op monitor interval="1min"
primitive ClusterIP ocf:heartbeat:IPaddr2 \
    params ip="192.168.122.101" cidr_netmask="32" clusterip_hash="sourceip" \
    op monitor interval="30s"
ms WebDataClone WebData \
    meta master-max="2" master-node-max="1" clone-max="2" clone-node-max="1" notify="true"
clone WebFSClone WebFSclone WebIP ClusterIP \
    meta globally-unique="true" clone-max="2" clone-node-max="2"
clone WebSiteClone WebSitecolocation WebSite-with-WebFS inf: WebSiteClone WebFSClone
colocation fs_on_drbd inf: WebFSClone WebDataClone:Master
colocation website-with-ip inf: WebSiteClone WebIP
order WebFS-after-WebData inf: WebDataClone:promote WebFSClone:start
order WebSite-after-WebFS inf: WebFSClone WebSiteClone
order apache-after-ip inf: WebIP WebSiteClone
property $id="cib-bootstrap-options" \
    dc-version="1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f" \
    cluster-infrastructure="openais" \
    expected-quorum-votes="2" \
    stonith-enabled="false" \
    no-quorum-policy="ignore"
rsc_defaults $id="rsc-options" \
    resource-stickiness="100"
```

Prima di caricare la configurazione nel cluster questa va revisionata, la shell andrà terminata e si dovrà osservare il responso del cluster

```
crm(active) # cib commit active
INFO: commited 'active' shadow CIB to the cluster
crm(active) # quit
```

```
bye
# crm_mon
============
Last updated: Thu Sep 3 21:37:27 2009
Stack: openais
Current DC: pcmk-2 - partition with quorum
Version: 1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f
2 Nodes configured, 2 expected votes
6 Resources configured.
============

Online: [ pcmk-1 pcmk-2 ]

Master/Slave Set: WebDataClone
    Masters: [ pcmk-1 pcmk-2 ]
Clone Set: WebIP Started: [ pcmk-1 pcmk-2 ]
Clone Set: WebFSClone Started: [ pcmk-1 pcmk-2 ]
Clone Set: WebSiteClone Started: [ pcmk-1 pcmk-2 ]
```

8.5.1. Testare il recovery

Nota

TODO: mettere un nodo in standby per dimostrare il failover

Configurare STONITH

Indice

9.1. What Is STONITH

STONITH is an acronym for Shoot-The-Other-Node-In-The-Head and it protects your data from being corrupted by rogue nodes or concurrent access.

Solo perché un nodo non è responsivo non significa che non stia accedendo ai dati. L'unica via per essere sicuri al 100% che i dati sono al sicuro è quella di utilizzare STONITH in modo da essere certi che il nodo sia effettivamente offline, prima di consentire ad altri nodi di accedere ai dati.

STONITH gioca un ruolo importante anche nel caso in cui un servizio clusterizzato non può essere stoppato. In questo caso il cluster utilizza STONITH per forzare l'intero nodo offline, rendendo di fatto sicuro l'avvio del servizio altrove.

9.2. Quale device STONITH andrebbe utilizzato

E' cruciale che il device STONITH sia in grado di differenziare un guasto del nodo da un guasto di rete.

The biggest mistake people make in choosing a STONITH device is to use remote power switch (such as many on-board IMPI controllers) that shares power with the node it controls. In such cases, the cluster cannot be sure if the node is really offline, or active and suffering from a network fault.

Likewise, any device that relies on the machine being active (such as SSH-based "devices" used during testing) are inappropriate.

9.3. Configurare STONITH

1. Find the correct driver: **`stonith_admin --list-installed`**

2. Since every device is different, the parameters needed to configure it will vary. To find out the parameters associated with the device, run: **`stonith_admin --metadata --agent type`**

   ```
   The output should be XML formatted text containing additional
   parameter descriptions. We will endevor to make the output more
   friendly in a later version.
   ```

3. Enter the shell crm Create an editable copy of the existing configuration cib new stonith Create a fencing resource containing a primitive resource with a class of stonith, a type of type and a parameter for each of the values returned in step 2: **`configure primitive …`**

4. If the device does not know how to fence nodes based on their uname, you may also need to set the special **`pcmk_host_map`** parameter. See **`man stonithd`** for details.

5. If the device does not support the list command, you may also need to set the special **`pcmk_host_list`** and/or **`pcmk_host_check`** parameters. See **`man stonithd`** for details.

6. If the device does not expect the victim to be specified with the port parameter, you may also need to set the special **`pcmk_host_argument`** parameter. See **`man stonithd`** for details.

7. Upload it into the CIB from the shell: **`cib commit stonith`**

8. Once the stonith resource is running, you can test it by executing: **`stonith_admin --reboot nodename`**. Although you might want to stop the cluster on that machine first.

9.4. Esempio

Assuming we have an chassis containing four nodes and an IPMI device active on 10.0.0.1, then we would chose the fence_ipmilan driver in step 2 and obtain the following list of parameters

Obtaining a list of STONITH Parameters

```
# stonith_admin --metadata -a fence_ipmilan
```

```
<?xml version="1.0" ?>
<resource-agent name="fence_ipmilan" shortdesc="Fence agent for IPMI over LAN">
<longdesc>
fence_ipmilan is an I/O Fencing agent which can be used with machines controlled by IPMI.
 This agent calls support software using ipmitool (http://ipmitool.sf.net/).

To use fence_ipmilan with HP iLO 3 you have to enable lanplus option (lanplus / -P) and
 increase wait after operation to 4 seconds (power_wait=4 / -T 4)</longdesc>
<parameters>
        <parameter name="auth" unique="1">
                <getopt mixed="-A" />
                <content type="string" />
                <shortdesc>IPMI Lan Auth type (md5, password, or none)</shortdesc>
        </parameter>
        <parameter name="ipaddr" unique="1">
                <getopt mixed="-a" />
                <content type="string" />
                <shortdesc>IPMI Lan IP to talk to</shortdesc>
        </parameter>
        <parameter name="passwd" unique="1">
                <getopt mixed="-p" />
                <content type="string" />
                <shortdesc>Password (if required) to control power on IPMI device</shortdesc>
        </parameter>
        <parameter name="passwd_script" unique="1">
                <getopt mixed="-S" />
                <content type="string" />
                <shortdesc>Script to retrieve password (if required)</shortdesc>
        </parameter>
        <parameter name="lanplus" unique="1">
                <getopt mixed="-P" />
                <content type="boolean" />
                <shortdesc>Use Lanplus</shortdesc>
        </parameter>
        <parameter name="login" unique="1">
                <getopt mixed="-l" />
                <content type="string" />
                <shortdesc>Username/Login (if required) to control power on IPMI device</
shortdesc>
        </parameter>
        <parameter name="action" unique="1">
```

```
                <getopt mixed="-o" />
                <content type="string" default="reboot"/>
                <shortdesc>Operation to perform. Valid operations: on, off, reboot, status,
 list, diag, monitor or metadata</shortdesc>
        </parameter>
        <parameter name="timeout" unique="1">
                <getopt mixed="-t" />
                <content type="string" />
                <shortdesc>Timeout (sec) for IPMI operation</shortdesc>
        </parameter>
        <parameter name="cipher" unique="1">
                <getopt mixed="-C" />
                <content type="string" />
                <shortdesc>Ciphersuite to use (same as ipmitool -C parameter)</shortdesc>
        </parameter>
        <parameter name="method" unique="1">
                <getopt mixed="-M" />
                <content type="string" default="onoff"/>
                <shortdesc>Method to fence (onoff or cycle)</shortdesc>
        </parameter>
        <parameter name="power_wait" unique="1">
                <getopt mixed="-T" />
                <content type="string" default="2"/>
                <shortdesc>Wait X seconds after on/off operation</shortdesc>
        </parameter>
        <parameter name="delay" unique="1">
                <getopt mixed="-f" />
                <content type="string" />
                <shortdesc>Wait X seconds before fencing is started</shortdesc>
        </parameter>
        <parameter name="verbose" unique="1">
                <getopt mixed="-v" />
                <content type="boolean" />
                <shortdesc>Verbose mode</shortdesc>
        </parameter>
</parameters>
<actions>
        <action name="on" />
        <action name="off" />
        <action name="reboot" />
        <action name="status" />
        <action name="diag" />
        <action name="list" />
        <action name="monitor" />
        <action name="metadata" />
</actions>
</resource-agent>
```

from which we would create a STONITH resource fragment that might look like this

Sample STONITH Resource

```
# crm crm(live)# cib new stonith
INFO: stonith shadow CIB created
crm(stonith)# configure primitive impi-fencing stonith::fence_ipmilan \
 params pcmk_host_list="pcmk-1 pcmk-2" ipaddr=10.0.0.1 login=testuser passwd=abc123 \
 op monitor interval="60s"
```

And finally, since we disabled it earlier, we need to re-enable STONITH. At this point we should have the following configuration.

```
crm(stonith)# configure property stonith-enabled="true"crm(stonith)# configure shownode
 pcmk-1
```

```
node pcmk-2
primitive WebData ocf:linbit:drbd \
    params drbd_resource="wwwdata" \
    op monitor interval="60s"
primitive WebFS ocf:heartbeat:Filesystem \
    params device="/dev/drbd/by-res/wwwdata" directory="/var/www/html" fstype="gfs2"
primitive WebSite ocf:heartbeat:apache \
    params configfile="/etc/httpd/conf/httpd.conf" \
    op monitor interval="1min"
primitive ClusterIP ocf:heartbeat:IPaddr2 \
    params ip="192.168.122.101" cidr_netmask="32" clusterip_hash="sourceip" \
    op monitor interval="30s"primitive ipmi-fencing
 stonith::fence_ipmilan \ params pcmk_host_list="pcmk-1
 pcmk-2" ipaddr=10.0.0.1 login=testuser passwd=abc123 \ op monitor interval="60s"ms
 WebDataClone WebData \
    meta master-max="2" master-node-max="1" clone-max="2" clone-node-max="1" notify="true"
clone WebFSClone WebFS
clone WebIP ClusterIP \
    meta globally-unique="true" clone-max="2" clone-node-max="2"
clone WebSiteClone WebSite
colocation WebSite-with-WebFS inf: WebSiteClone WebFSClone
colocation fs_on_drbd inf: WebFSClone WebDataClone:Master
colocation website-with-ip inf: WebSiteClone WebIP
order WebFS-after-WebData inf: WebDataClone:promote WebFSClone:start
order WebSite-after-WebFS inf: WebFSClone WebSiteClone
order apache-after-ip inf: WebIP WebSiteClone
property $id="cib-bootstrap-options" \
    dc-version="1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f" \
    cluster-infrastructure="openais" \
    expected-quorum-votes="2" \
    stonith-enabled="true" \
    no-quorum-policy="ignore"
rsc_defaults $id="rsc-options" \
    resource-stickiness="100"
crm(stonith)# cib commit stonithINFO: commited 'stonith' shadow CIB to the cluster
crm(stonith)# quit
bye
```

Appendice A. Riepilogo delle configurazioni

Indice

A.1. Configurazione finale del cluster

```
# crm configure show
node pcmk-1
node pcmk-2
primitive WebData ocf:linbit:drbd \
    params drbd_resource="wwwdata" \
    op monitor interval="60s"
primitive WebFS ocf:heartbeat:Filesystem \
    params device="/dev/drbd/by-res/wwwdata" directory="/var/www/html" fstype="gfs2"
primitive WebSite ocf:heartbeat:apache \
    params configfile="/etc/httpd/conf/httpd.conf" \
    op monitor interval="1min"
primitive ClusterIP ocf:heartbeat:IPaddr2 \
    params ip="192.168.122.101" cidr_netmask="32" clusterip_hash="sourceip" \
    op monitor interval="30s"
primitive ipmi-fencing stonith::fence_ipmilan \
    params pcmk_host_list="pcmk-1 pcmk-2" ipaddr=10.0.0.1 login=testuser passwd=abc123 \
    op monitor interval="60s"
ms WebDataClone WebData \
    meta master-max="2" master-node-max="1" clone-max="2" clone-node-max="1" notify="true"
clone WebFSClone WebFS
clone WebIP ClusterIP \
    meta globally-unique="true" clone-max="2" clone-node-max="2"
clone WebSiteClone WebSite
colocation WebSite-with-WebFS inf: WebSiteClone WebFSClone
colocation fs_on_drbd inf: WebFSClone WebDataClone:Master
colocation website-with-ip inf: WebSiteClone WebIP
order WebFS-after-WebData inf: WebDataClone:promote WebFSClone:start
order WebSite-after-WebFS inf: WebFSClone WebSiteClone
order apache-after-ip inf: WebIP WebSiteClone
property $id="cib-bootstrap-options" \
    dc-version="1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f" \
    cluster-infrastructure="openais" \
    expected-quorum-votes="2" \
    stonith-enabled="true" \
    no-quorum-policy="ignore"
rsc_defaults $id="rsc-options" \
    resource-stickiness="100"
```

A.2. Lista nodi

La lista dei nodi è popolata automaticamente dal cluster.

```
node pcmk-1
node pcmk-2
```

A.3. Opzioni del cluster

Qui è dove il cluster registra automaticamente alcune informazioni in merito al cluster

- dc-version - la versione (incluso l'hash del codice sorgente originale) di Pacemaker usata nel DC
- cluster-infrastructure - l'infrastruttura cluster utilizzata (heartbeat or openais)
- expected-quorum-votes - il numero massimo di nodi che ci si aspetta facciano parte del cluster

e dove l'amministratore può assegnare valori alle opzioni che controllano il modo in cui il cluster opera

- stonith-enabled=true - Fai uso di STONITH
- no-quorum-policy=ignore - Ignora la perdita di quorum e continua ad ospitare le risorse.

```
property $id="cib-bootstrap-options" \
    dc-version="1.1.5-bdd89e69ba545404d02445be1f3d72e6a203ba2f" \
    cluster-infrastructure="openais" \
    expected-quorum-votes="2" \
    stonith-enabled="true" \
    no-quorum-policy="ignore"
```

A.4. Risorse

A.4.1. Opzioni di default

Qui vengono configurate le opzioni del cluster che vanno applicati a tutte le risorse

- resource-stickiness - Specifica l'impossibilità o meno di muovere risorse ad altre macchine

```
rsc_defaults $id="rsc-options" \
    resource-stickiness="100"
```

A.4.2. Fencing

Nota

TODO: Aggiungi il testo qui

```
primitive ipmi-fencing stonith::fence_ipmilan \
    params pcmk_host_list="pcmk-1 pcmk-2" ipaddr=10.0.0.1 login=testuser passwd=abc123 \
    op monitor interval="60s"
clone Fencing rsa-fencing
```

A.4.3. Servizio Address

Users of the services provided by the cluster require an unchanging address with which to access it. Additionally, we cloned the address so it will be active on both nodes. An iptables rule (created as part of the resource agent) is used to ensure that each request only gets processed by one of the two clone instances. The additional meta options tell the cluster that we want two instances of the clone (one "request bucket" for each node) and that if one node fails, then the remaining node should hold both.

```
primitive ClusterIP ocf:heartbeat:IPaddr2 \
    params ip="192.168.122.101" cidr_netmask="32" clusterip_hash="sourceip" \
    op monitor interval="30s"
clone WebIP ClusterIP
    meta globally-unique="true" clone-max="2" clone-node-max="2"
```

Nota

TODO: Il RA quando clonato dovrebbe controllare l'opzione globally-unique=true

A.4.4. DRBD - Storage condiviso

Qui viene definito il servizio DRBD e specificata quale risorsa DRBD (da drbd.conf) questi debba controllare. La risorsa viene definita come master/slave e, per rispettare il setup active/active, entrambe le istanze vengono abilitate ad essere promosse specificando master-max=2. Viene valorizzata inoltre l'opzione notify, così che il cluster comunicherà all'agent DRBD quando il suo nodo cambierà stato.

```
primitive WebData ocf:linbit:drbd \
    params drbd_resource="wwwdata" \
    op monitor interval="60s"
ms WebDataClone WebData \
    meta master-max="2" master-node-max="1" clone-max="2" clone-node-max="1" notify="true"
```

A.4.5. Cluster Filesystem

Il Cluster Filesystem si assicura che i file vengano letti e scritti nella maniera corretta. E' necessario specificare il block device (fornito da DRBD), dove si vuole effettuare l'operazione di mount e che viene utilizzato GFS2. Di nuovo questo è un clone, perché è inteso essere attivo su entrambi i nodi. La constraint aggiuntiva assicura che la risorsa possa essere avviata su nodi con gfs-control attivo e istanze drbd.

```
primitive WebFS ocf:heartbeat:Filesystem \
    params device="/dev/drbd/by-res/wwwdata" directory="/var/www/html" fstype="gfs2"
clone WebFSClone WebFS
colocation WebFS-with-gfs-control inf: WebFSClone gfs-clone
colocation fs_on_drbd inf: WebFSClone WebDataClone:Master
order WebFS-after-WebData inf: WebDataClone:promote WebFSClone:start
order start-WebFS-after-gfs-control inf: gfs-clone WebFSClone
```

A.4.6. Apache

Infine viene definito il servizio Apache. E' necessario solamente specificare al cluster dove trovare il file di configurazione principale e costringere questo ad essere eseguito su nodi con il filesystem richiesto montato e l'indirizzo IP attivo.

```
primitive WebSite ocf:heartbeat:apache \
    params configfile="/etc/httpd/conf/httpd.conf" \
    op monitor interval="1min"
clone WebSiteClone WebSite
colocation WebSite-with-WebFS inf: WebSiteClone WebFSClone
colocation website-with-ip inf: WebSiteClone WebIP
order apache-after-ip inf: WebIP WebSiteClone
order WebSite-after-WebFS inf: WebFSClone WebSiteClone
```

Appendice B. Sample Corosync Configuration

Sample Corosync.conf for a two-node cluster

```
# Please read the Corosync.conf.5 manual page
compatibility: whitetank

totem {
    version: 2

    # How long before declaring a token lost (ms)
    token:     5000

    # How many token retransmits before forming a new configuration
    token_retransmits_before_loss_const: 10

    # How long to wait for join messages in the membership protocol (ms)
    join:      1000

    # How long to wait for consensus to be achieved before starting a new
    # round of membership configuration (ms)
    consensus:   6000

    # Turn off the virtual synchrony filter
    vsftype:    none

    # Number of messages that may be sent by one processor on receipt of the token
    max_messages:  20

    # Stagger sending the node join messages by 1..send_join ms
    send_join: 45

    # Limit generated nodeids to 31-bits (positive signed integers)
    clear_node_high_bit: yes

    # Disable encryption
    secauth:    off

    # How many threads to use for encryption/decryption
    threads:      0

    # Optionally assign a fixed node id (integer)
    # nodeid:     1234

    interface {
        ringnumber: 0

        # The following values need to be set based on your environment
        bindnetaddr: 192.168.122.0
        mcastaddr: 226.94.1.1
        mcastport: 4000
    }
}

logging {
    debug: off
    fileline: off
    to_syslog: yes
    to_stderr: off
    syslog_facility: daemon
    timestamp: on
}

amf {
    mode: disabled
}
```

Appendice C. Approfondimenti

- Project Website *http://www.clusterlabs.org*
- Cluster Commands A comprehensive guide to cluster commands has been written by Novell and can be found at: *http://www.novell.com/documentation/sles11/book_sleha/index.html?page=/documentation/sles11/book_sleha/data/book_sleha.html*
- Corosync *http://www.corosync.org*

Appendice D. Revision History

Revisione 1-1	**Mon May 17 2010**	**Andrew Beekhof** *andrew@beekhof.net*
Import from Pages.app		
Revisione 2-1	**Wed Sep 22 2010**	**Raoul Scarazzini** *rasca@miamammausalinux.org*
Italian translation		
Revisione 3-1	**Wed Feb 9 2011**	**Andrew Beekhof** *andrew@beekhof.net*
Updated for Fedora 13		
Revisione 4-1	**Wed Oct 5 2011**	**Andrew Beekhof** *andrew@beekhof.net*
Update the GFS2 section to use CMAN		
Revisione 5-1	**Fri Feb 10 2012**	**Andrew Beekhof** *andrew@beekhof.net*
Generate docbook content from asciidoc sources		

Indice analitico