PREVENT, BACKUP AND RESTORE PROCEDURES FOR A WEB SERVER

Cosmin Cătălin OLTEANU[1]

Abstract

In These days when everything is stored online on a server a real disaster recovery plan should be a must for every network administrator. From my experience I would say that the first step is to prevent and than to make a detailed backup plan and a restore one too. Custom cron scripts can make automated scheduled various commands just tools for trustworthy systems.

Key words: *informatic system back-up plan, mysql dump, prevent power failure*

1. INTRODUCTION

In the WEB integrated informatic systems era, reliability of servers is a big issue.

If we have a real disaster plan for prevention and recovery of data loss and hardware failure, we can say that we have a trust worthy system and we can rely on it.

Every user should trust the system and should work with confidence that his work will be also available tomorrow.

2. PREVENT, BACKUP AND RESTORE A REAL DISASTER RECOVERY PLAN

These days when everything is stored online on a server a real disaster recovery plan should be a must for every network administrator.

From my experience I would say that the first step is to prevent and than to make a detailed backup plan and a restore one too.

Let's assume that we have a web generated application that is used as an integrated informatic system for a university campus.

We'll go further and we'll presume that everything is based on Linux (any distribution available, like Fedora 13, Gentoo, Debian etc.).

Fig. 1.Gentoo Logo.

[1] Lectuer, PhD, "Nicolae Titulescu" University, Bucharest, e-mail: contact@olteanucosmin.ro

Fig. 2.Fedora 13 Logo.

Data losses can happen any time due to many reasons and we have to manage that paper plan is not always equal with real plan:

- We can have accidental data loss
- We can have intentional data loss
- We have to deal with small budgets for components and not ideal ones etc.

I have to say that we have a few causes and I have discovered from my previous experience that this causes can be:

- electrical / power problems;
- failure of devices;
- bad coding
- Database bug's etc.

a) electrical / power problems that can be avoided

I our country every issue about power supplier is a closed one because is just one supplier and we just do not have an alternative and usually when we encounter such problems the results are devastating.

Just a few years ago a mail server that I usually maintain just burn out because of a big overvoltage. The UPS and mainboard were fried.

After a few days (after I have replaced the components and the server was working again) I have tried to find an assurance company to have all the equipment assured but from 11 companies none of them could make an offer for electrical problems generated by supplier. The conclusion was just annoying: No one takes responsibility due to power failure. All must be done by lawyers, court law suites and time just could be extended for years until something is done.

All I could do in an environment where "time is money" was to separate the problem in two stages and deal with both of them:

1) problems generated through power cables (220v);

2) problems generated through small curents cables (UTP).

For the first problem the solution was to install a good automatic voltage regulator (Fig. 4) doubled by powerful Uninterruptible Power Supply (Fig.3).

For these I have chosen APC products like:

Fig. 3. APC Smart-UPS 1000VA.

Fig. 4. APC Line-R 1200VA Automatic voltage regulation

For the second problem I have to say that for UTP Ethernet stable connections at 1 Gbit/s I used also an APC product (Fig.5):

Fig. 5. APC APC ProtectNet standalone surge protector for 10/100/1000
ase-T Ethernet lines

b) failure of devices

For this type of failures I have encountered only problems with magnetic storage devices like hard drives. The others components were stable in time (remember that we have server components where the quality and control process is very reliable).

The first thing to do is to have a server managed with mirror RAID 1 enabled (Fig. 3), doubled by enterprise hard disk in SAS technology.

I have chose RAID 1 as a solution because if we have problems with one hard drive we have just to change it and reconstruct the RAID matrix before everything is back as it was in no time.

I recommend for such hard disks to use Seagate 7/24 latest technology of Barracuda® ES.2 with a rate of 1.2 M hrs. MTBF or Constellation™ ES with a rate of 1.8 M hrs. MTBF.

I have to say that these HDD's proved to be the most reliable on the market and the warranty of 5 years is all that we need for the moment.

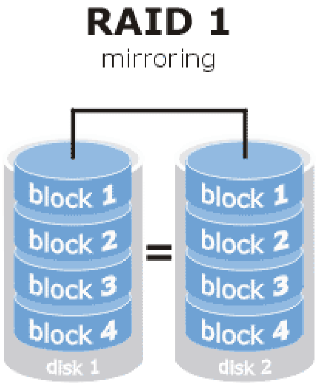


Fig. 6.RAID mirroring diagramme.

c) bad coding

The problems generated by any code of a certain application are hard to find and can be eliminated only in time. For this type of errors the solution is to have strict back-up of database and to update that application as quick as the resolving patch appear.

d) Database bug's

Database platforms are quite often improved by new general release or by small patches. A successful database on linux world is MySql that is quite secure and reliable.

The patches of the distribution must be installed on a daily bases to have a secure system.

Four layers of back-up

In a real time server environment based on Linux if we discuss about back-up we have to talk about a plan for every layer that is needed to be managed in order to have a full working server.

In real life I realized that the layers can be grouped as:

- ✓ Operating System of server with particular configuration files for every service/server system
- ✓ Database files
- ✓ Web Application files
- ✓ Log files

a) Operating System of server with particular configuration files for every service/server system

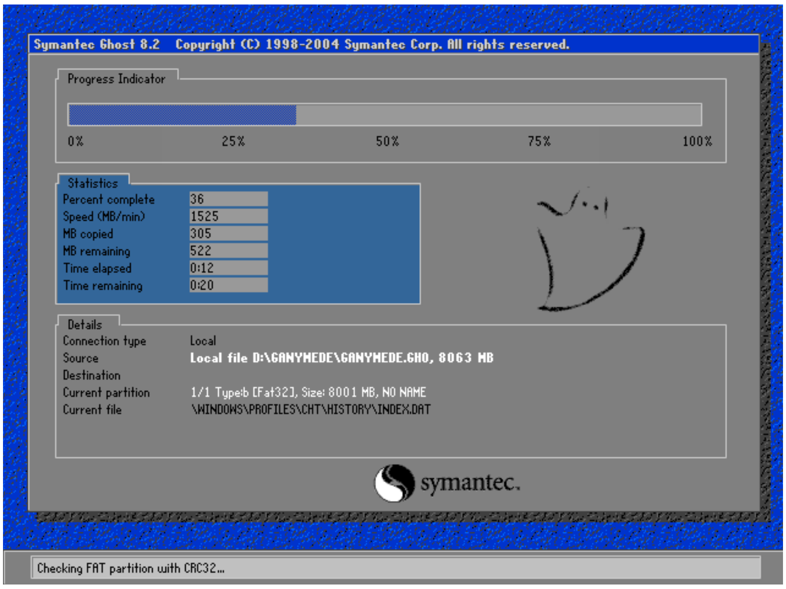
Usually the Operating System (OS) of a production server is kept frozen about new installations of additional software but open for security patches of existent services.

As a backup procedure for OS, I have implemented two methods for quick restore.

First method is to have a clone image for whole hard drive. I usually use Symantech Professional GHOST 11 (Fig. 7) or dd (1).

```
dd if=/dev/rdsk/dks0d1s0 bs=32k of=/dev/rdsk/dks0d2s0  (1)
```

(Example of dd backup)



```
Symantec Ghost 8.2   Copyright (C) 1998-2004 Symantec Corp. All rights reserved.

Progress Indicator
0%            25%            50%            75%            100%

Statistics
Percent complete      36
Speed (MB/min)        1525
MB copied             305
MB remaining          522
Time elapsed          0:12
Time remaining        0:20

Details
Connection type       Local
Source                Local file D:\GANYMEDE\GANYMEDE.GHO, 8063 MB
Destination
Current partition     1/1 Type:b [Fat32], Size: 8001 MB, NO NAME
Current file          \WINDOWS\PROFILES\CHT\HISTORY\INDEX.DAT

symantec.

Checking FAT partition with CRC32...
```

Fig. 7.Symantec GHOST clonning program

Let's say that this kind of back-up is done on a monthly basis.

Further more, I generate also an automated back-up on a cron schedule for every service personal config file by using tar (2).

```
tar –czpf /mnt/etc/etc.tar.gz /etc     (2)
```

After the archive is generated an automated script verify if the connection with another server is stable and active (4) and save the archive on a date folder (3) for future storage.

```
data_noua=`date +%w_%Y`(3)

if !(mount -t cifs //192.168.X.X/docum /mnt/profback -o username=nume_utilizator,
password=parola,rw)
then(4)
exit 1
```

b) Database files

In order to have a real back-up of the database, I had to make a couple of scripts that runs on a daily schedule, usually night a 4 o'clock.

There are two methods for a MySql database full back-up.

An sql dump (5) and full real archive of table files (6) of the database.

```
mysqldump -u ${mysql_user} -p${mysql_pass} --opt -A > ${date}_mysql_dump.sql
bzip2 -9 ${date}_mysql_dump.sql(5)

cd ${nt_mysql}
tar -czpf /tmp/site/bd_site_nt.tar.gz.(6)
```

The resulting files are automatically copied on a daily schedule basis on another server and from there are copied monthly on blue ray disks for archive and storage.

b) Web Application files

The web application files are the ones that interrogate the database and display the results. These files dynamically generates web pages for the real management of the integrated informatic system.

It must be mentioned that the updates are quite often available and the files are modified almost daily.

For that reason the back-up is done at night by a cron schedule. The script make an archive of all files and copy it on another server for safety on a date folder (7).

```
#!/bin/bash
mysql_user="nume_utilizator"
mysql_pass="parola"
http_base_dir="/var/www/localhost"
backup_dir="/mnt/backups" (7)
pwd=`pwd`
date=`date +%j_%d_%m_%Y`
random=`echo $RANDOM`
mkdir /tmp/${random}
cd /tmp/${random}
cd ${http_base_dir}
tar cvfp /tmp/${random}/${date}_http_dump.tar.
cd /tmp/${random}
bzip2 -9 ${date}_http_dump.tar
/opt/bin/rar a ${date}_backup.rar.
mv ${date}_backup.rar /mnt/backups/
chown -R coc:users /mnt/backups
rm -rf /tmp/${random}
```

d) Log files

With the log files we can discover if we have bugs in php files or in sql interrogation scripts etc. Also we can see if some hardware is not working properly.

The logs are kept only for a month older and for daily basis (8).

```
data_noua=`date +%w_%Y`
tar –czpf /mnt/log/log.tar.gz /var/log  (8)
```

The date folders are available for code developers in order to solve code bugs and errors.

If we have all the back-up’s for the restore procedure we have just to take in reverse order the back-up plan.

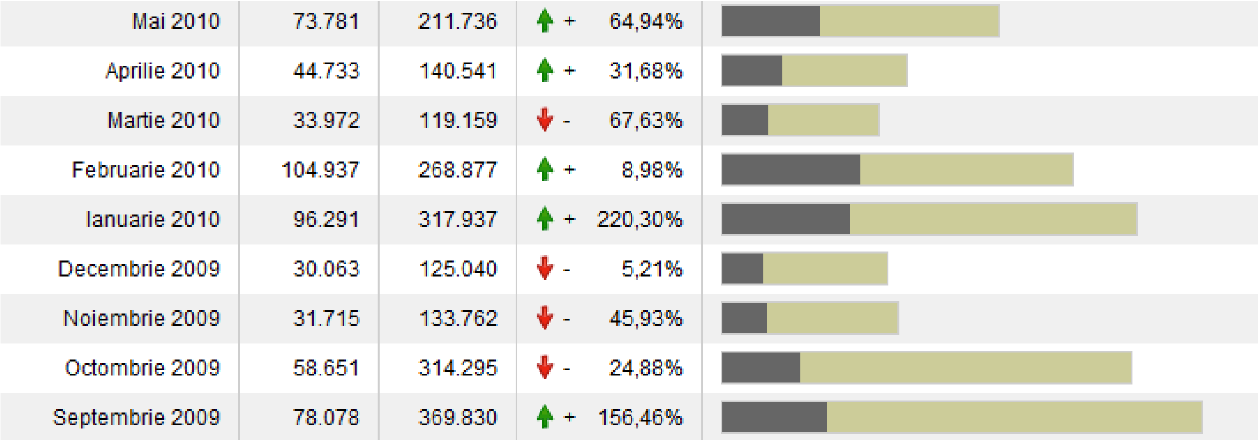
6. CONCLUSIONS

Prevention is the most important thing that can be done in an server environment. If a disaster appear only the back-ups can save the work of hundreds of people.

From my experience I have developed some schedule cron scripts that helps me in order to have all organized and in good conditions.

The results are quite encouraging because the production server of the integrated informatic system is customized by me from 2004 and everything works fine and data can be restored quickly.

This is why the users trust is growing and in February 2010 I had 104937 visits than January 2009 when I had only 22459 (Fig. 8).



Mai 2010	73.781	211.736	↑ +	64,94%
Aprilie 2010	44.733	140.541	↑ +	31,68%
Martie 2010	33.972	119.159	↓ -	67,63%
Februarie 2010	104.937	268.877	↑ +	8,98%
Ianuarie 2010	96.291	317.937	↑ +	220,30%
Decembrie 2009	30.063	125.040	↓ -	5,21%
Noiembrie 2009	31.715	133.762	↓ -	45,93%
Octombrie 2009	58.651	314.295	↓ -	24,88%
Septembrie 2009	78.078	369.830	↑ +	156,46%

Fig. 8. Statistics over new visitors.