

# LVM, Snapshot, Rsync, Bacula: Backup professionale a costo (quasi) zero

**Dr. Damiano Verzulli**

APM – Università “G' d'Annunzio”  
Chieti - Pescara

**damiano@verzulli.it**



Questo documento è rilasciato secondo i termini della licenza Creative Commons “Attribuzione - Non commerciale – Condividi allo stesso modo 2.5”. Secondo i termini di tale licenza, **questo documento può essere riprodotto, distribuito e modificato ma non per scopi commerciali.** Il testo integrale della licenza è disponibile all'indirizzo:

<http://creativecommons.org/licenses/by-nc-sa/2.5/it/legalcode>



# Il problema

- Implementare una soluzione di backup con:
  - (ragionevole) “backup in linea” per web e posta;
  - archiviazione storica dei LOG (firewall, flussi, sistemi);
  - possibilità di backup dei client (client/server)
  - garanzia assoluta di “controllo” e assenza di “vendor-lock-in” (no upgrade SW “forzati”)

**Soluzione  
necessariamente  
“Open Source”**

*Chi vi parla potrebbe essere  
scambiato per un  
“integralista F/OSS” ma  
(presuntuosamente?) NON  
ritiene di esserlo*



# Le criticità

- La soluzione doveva gestire volumi significativi:
  - **POSTA:** **+/- 400GB** x ca. 3.5 M file (mdir, in aumento)
  - **WEB:** **+/- 30GB** x ca. 300K file (in diminuzione)

**10 ore ca.  
di trasferimento  
su LAN a 100Mbps**

**+/- 430 GB per ogni  
“versione” da  
mantenere “on-line”**

***Ovviamente, a Chieti, esistono sia switch GETH sia HD da 1TB (...ma “entry”) ma... l'obiettivo era l'implementazione di una soluzione “per poveri, efficace, scalabile”!***



# L'idea

- Approcciare la notevole mole di dati con:
  - **RSYNC**: il formato MDIR si sposa benissimo con una soluzione “differenziale” visto che ad essere aggiornati sono principalmente i nuovi messaggi. Superato il primo RSYNC, gli altri vengono effettuati in tempi ragionevoli;
  - “**on-line**”: gestire sul file-system 7 alberi di sincronizzazione (lun/mar/mer/gio/ven/sab/dom) sui quali ruotare l'opportuno “rsync” giornaliero.

Ok ma... cosa fare per limitare i tempi di copia e l'uso dello spazio disco?



# LVM2 & snapshot

- Oltre alle caratteristiche classiche di Volume Management (*astrazione completa fra filesystem e storage sottostante attraverso i classici concetti di PV, VG, LV*), i **Logical Volume Manager** di Linux (LVM2 e eVMS) gestiscono pienamente gli **SNAPSHOT**;
- Lo SNAPSHOT è una “fotografia” fatta ad un certo filesystem (LV) in un certo istante temporale. A differenza di una copia classica offre due vantaggi fondamentali:
  - **SPAZIO**: lo snapshot occupa uno spazio che è dettato dalle differenze con l'originale;
  - **TEMPO**: a differenza della copia fisica, lo snapshot si crea in tempi dell'ordine del secondo.



# LVM2 & snapshot

...quindi:

- abbiamo recuperato un **PC entry-level** (PIV; 3.2Ghz; 2GB RAM);
- l'abbiamo dotato di **3 x 500 GB HD SATA**;
- abbiamo configurato l'LVM2 fornito da **UBUNTU LTS** con:
  - n. **1 V.G. unico da 1.5TB** - storageVG;
  - n. **1 L.V. da 600GB** - /dev/storageVG/backup
  - n. **7 snapshot da 50GB ciascuno** (R.O.)  
/dev/storageVG/Sun|Mon|Tue|Wed|Thu|Fri|Sat

....e di 1.5TB, sono “avanzati” 500GB ca.  
 $1500 - (600 + 7 * 50) = 550$



# LVM2 & snapshot

```
noc@nocbackup# lvsdisplay
```

## L.V. principale

```
--- Logical volume ---
LV Name                /dev/storageVG/backup
VG Name                storageVG
LV UUID                Z9nDyb-RZVD-hUfh-FFCd
LV Write Access        read/write
LV snapshot status    source of
                      /dev/storageVG/Sun [active]
                      /dev/storageVG/Mon [active]
                      /dev/storageVG/Tue [active]
                      /dev/storageVG/Wed [active]
                      /dev/storageVG/Thu [active]
                      /dev/storageVG/Fri [active]
                      /dev/storageVG/Sat [active]
LV Status              available
# open                 2
LV Size                600,00 GB
Current LE             153600
Segments               2
Allocation              inherit
Read ahead sectors     0
Block device           254:1
[...]
```

## Uno dei 7 Snapshot

```
noc@nocbackup# lvsdisplay
```

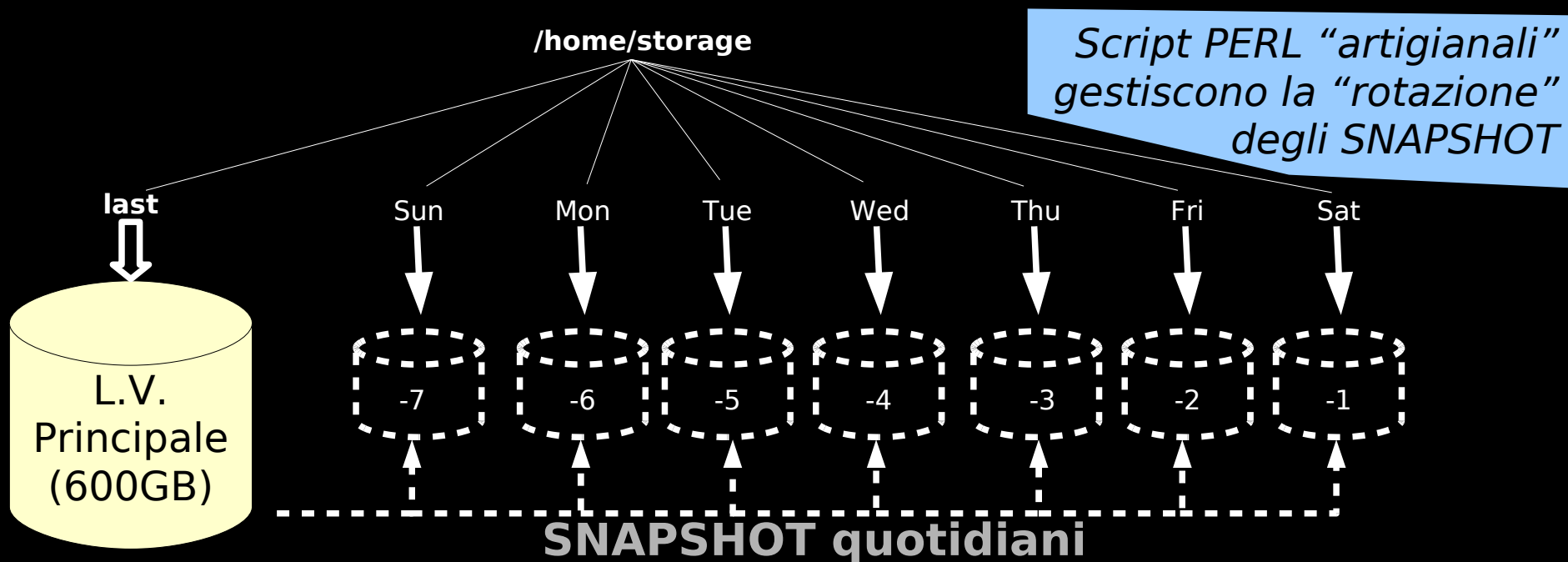
```
[...]
--- Logical volume ---
LV Name                /dev/storageVG/Sun
VG Name                storageVG
LV UUID                GcMdb1-HtnS-GFYD-p7GH
LV Write Access        read only
LV snapshot status    active destination for
                      /dev/storageVG/backup
LV Status              available
# open                 2
LV Size                600,00 GB
Current LE             153600
COW-table size         50,00 GB
COW-table LE           12800
Allocated to snapshot  46,18%
Snapshot chunk size    8,00 KB
Segments               1
Allocation              inherit
Read ahead sectors     0
Block device           254:2
[...]
```

**Lo snapshot più vecchio  
(7 gg) occupa il 46% dei  
50GB a lui riservati**



# L'accesso ai dati...

Ovviamente tutti ed 8 i L.V. possono essere “montati” ed acceduti senza problemi. Di conseguenza...



**Recuperare i dati vecchi di N giorni ( $N \leq 7$ ) equivale a fare “cd” + “scp”**





# La storicizzazione su nastro

- Ovviamente 7 giorni di “conservazione” del backup non sono sufficienti: bisogna aggiungere una gestione ottimale dei nastri
- Dopo un'attenta ricerca si è scelto di impostare una architettura basata su:

- **IBM TS/3100**: 2RU; 23+1 slot per nastri LTO4 (400/800GB); barcode reader; supporto per il doppio tape;



- **BACULA**: una soluzione di backup open-source, con il pieno supporto di diverse tape-library



# BACULA – Backup hi-end Open-Source

Parlare in dettaglio di BACULA è decisamente incompatibile con gli spazi ed i tempi di questa presentazione. In estrema sintesi...

- **BACULA**, in produzione da settembre 2008, non ha mai dato alcun problema, rivelandosi un SW all'altezza di Apache, MySQL, Samba per quanto riguarda stabilità, flessibilità, scalabilità.

D'altro canto è un software che “nasce” per essere high-end:

*“...if you are new to Unix systems or do not have offsetting experience with a sophisticated backup package, the Bacula project does not recommend using Bacula...”*

da [www.bacula.org](http://www.bacula.org)



# BACULA – Backup hi-end Open-Source

---

Nel nostro caso BACULA è configurato per:

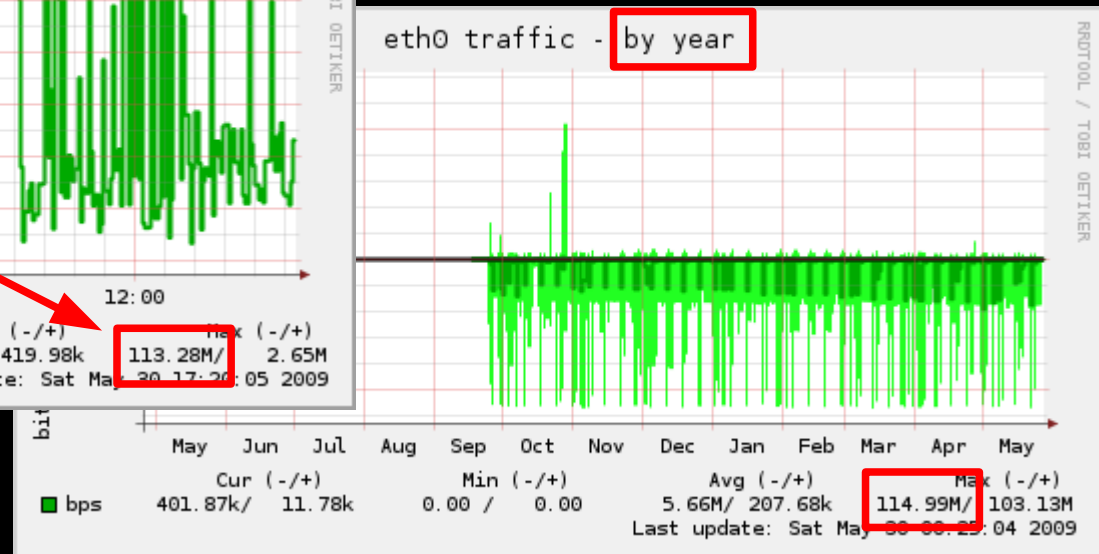
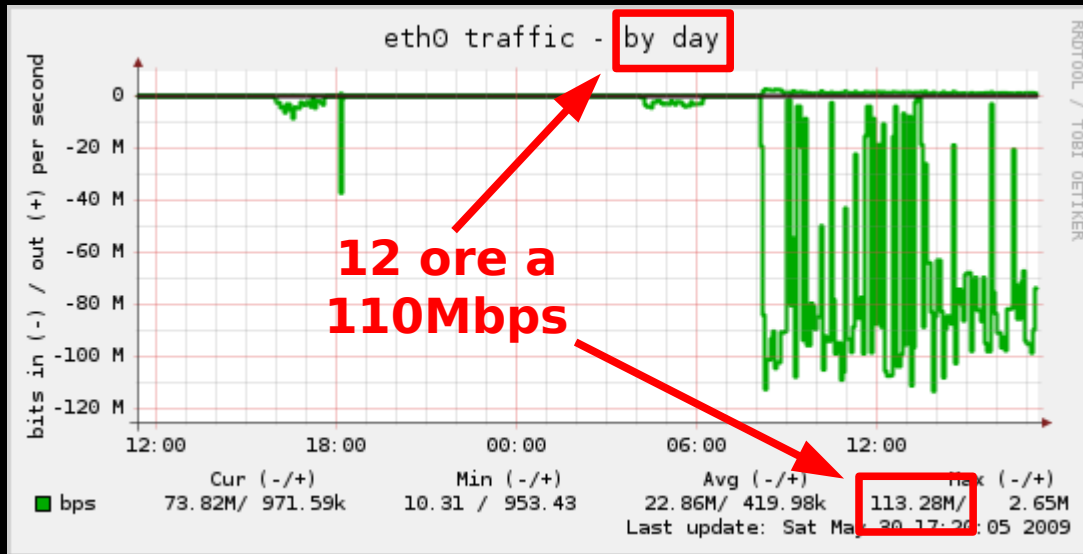
- Effettuare, giornalmente, settimanalmente e mensilmente una serie articolata di Backup totali, differenziali, incrementali;
- Mantenere:
  - alcuni di tali backup, a tempo indeterminato;
  - altri backup, per un mese;
  - altri backup, per una settimana,
- Distribuire tali backup su opportuni “set” di nastri

...tanto altro

# Bacula: considerazioni

Bacula, quando usato seriamente, può diventare... pericoloso:

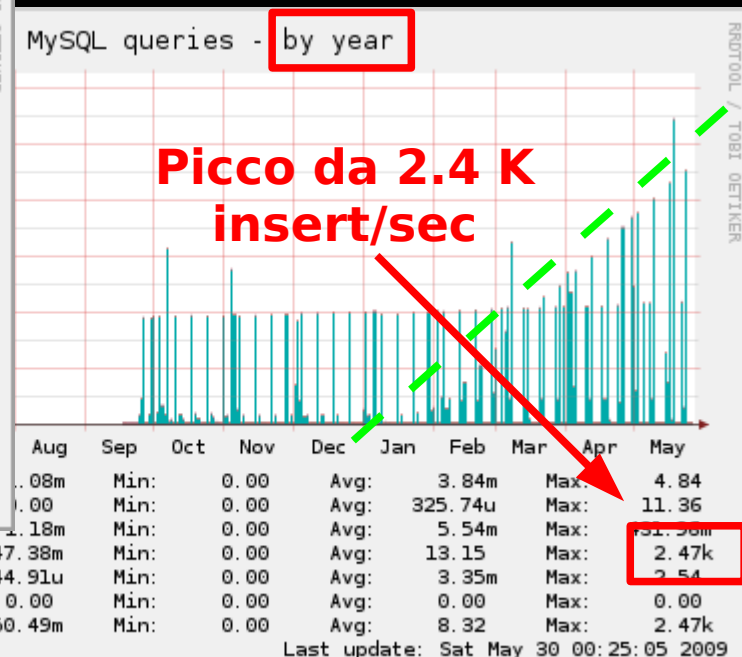
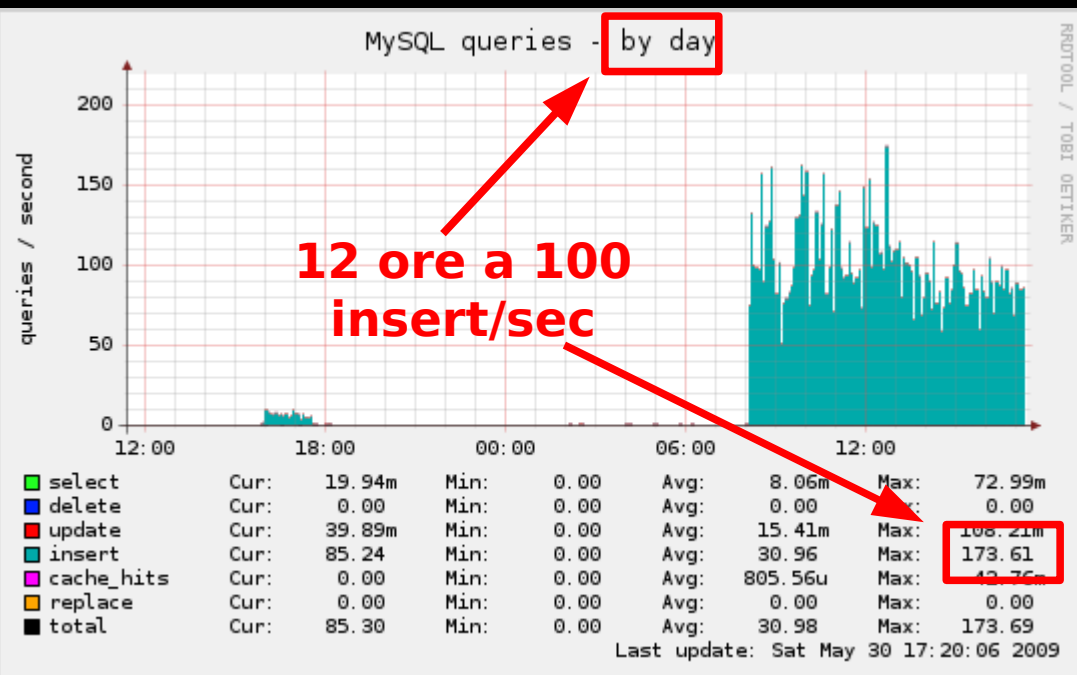
- **Traffico LAN:** rapida saturazione di segmenti FastETH



# Bacula: considerazioni

- **MySQL**: il nostro Bacula utilizza MySQL come repository informativo. Backuppare una massa rilevante di dati, equivale ad effettuare un DOS di "Insert" SQL:

**Dobbiamo ancora capire cosa è cambiato a gennaio 2009**



# Bacula: considerazioni

- **MySQL**: può rapidamente diventare un elemento **MOLTO** critico

```
mysql> select MediaId,VolumeName,LabelDate  
from Media order by 3 DESC;
```

MediaId	VolumeName	LabelDate
42	943CCCL3	2009-05-23 08:06:33
49	001MEDL3	2009-05-22 11:56:27
51	275JJJL3	2009-05-17 15:55:03
29	402CCCL3	2009-05-16 08:05:02
31	940CCCL3	2009-05-09 08:05:02
52	897UUUL3	2009-05-05 15:53:44
50	002MEDL3	2009-05-05 12:54:44
48	279JJJL3	2009-05-03 09:07:17
24	404CCCL3	2009-05-02 08:05:02
47	941CCCL3	2009-05-01 21:01:38
46	276JJJL3	2009-04-20 20:02:14
43	278JJJL3	2009-04-01 07:32:09
41	942CCCL3	2009-03-16 12:58:54
39	896UUUL3	2009-03-01 09:07:33
37	898UUUL3	2009-02-01 09:06:41
36	277JJJL3	2009-01-04 09:06:20
33	401CCCL3	2008-12-04 16:06:16
27	403CCCL3	2008-12-01 00:31:15
21	400CCCL3	2008-11-05 11:22:38
20	895UUUL3	2008-11-02 11:51:26
28	944CCCL3	2008-10-28 13:19:31

**21 nastri  
attualmente gestiti**

**67 milioni  
di record (file - 9 mesi)**

```
mysql> show tables;
```

Tables_in_bacula
BaseFiles
CDImages
Client
Counters
Device
File
FileSet
Filename
Job
JobMedia
Location
LocationLog
Log
Media
MediaType
Path
Pool
Status
Storage
UnsavedFiles
Version

```
21 rows in set (0.00 sec)
```

```
mysql> select count(*) from  
File;
```

count (*)
67326689

```
1 row in set (0.00 sec)
```



# Bacula: considerazioni

- Negli impieghi più pesanti, **volumi**, **tipologia** di backup e **velocità** del tape (=durata del backup) vanno attentamente valutati:

```
30-May 08:05 deimos.unich.it-dir JobId 1896: Start Backup JobId 1896, Job=nobackupMAILDaily.2009-05-30_08.05.33
[...]
30-May 08:08 deimos.unich.it-sd JobId 1896: Volume "943CCCL3" previously written, moving to end of data.
30-May 17:51 deimos.unich.it-sd JobId 1896: End of Volume "943CCCL3" at 565:14691 on device "AutochangerDrive" (/dev/nst0).
Write of 64512 bytes got -1.
30-May 17:51 deimos.unich.it-sd JobId 1896: Re-read of last block succeeded.
30-May 22:45 deimos.unich.it-dir JobId 1896:There are no more Jobs associated with Volume "404CCCL3". Marking it purged.
[...]
30-May 22:47 deimos.unich.it-sd JobId 1896: Recycled volume "404CCCL3" on device "AutochangerDrive" (/dev/nst0), all previous data lost.
30-May 22:47 deimos.unich.it-sd JobId 1896: New volume "404CCCL3" mounted on device "AutochangerDrive" (/dev/nst0) at 30-May-2009 22:47.
31-May 00:31 deimos.unich.it-sd JobId 1896: Job write elapsed time = 11:26:08, Transfer rate = 10.85 MB bytes/second
[...]
Job: nobackupMAILDaily.2009-05-30_08.05.33
Backup Level: Full
Client: nobackup-fd 2.2.8 (26Jan08)
FileSet: nobackupMAILSET 2008-09-27 08:05:00
Scheduled time: 30-May-2009 08:05:00
Start time: 30-May-2009 08:07:27
End time: 31-May-2009 02:05:29
Elapsed time: 17 hours 58 mins 2 secs
FD Files Written: 3,748,878
FD Bytes Written: 445,946,489,828 (445.9 GB)
Rate: 6894.4 KB/s
Volume name(s): 943CCCL3|404CCCL3
FD termination status: OK
SD termination status: OK
Termination: Backup OK
[...]
31-May 02:05 deimos.unich.it-dir JobId 1896: Begin pruning Jobs.
31-May 02:05 deimos.unich.it-dir JobId 1896: No Jobs found to prune.
31-May 02:05 deimos.unich.it-dir JobId 1896: Begin pruning Files.
31-May 02:05 deimos.unich.it-dir JobId 1896: No Files found to prune.
31-May 02:05 deimos.unich.it-dir JobId 1896: End auto prune.
```

```
Scheduled time: 30-May-2009 08:05:00
Start time: 30-May-2009 08:07:27
End time: 31-May-2009 02:05:29
Elapsed time: 17 hours 58 mins 2 secs
FD Files Written: 3,748,878
FD Bytes Written: 445,946,489,828 (445.9 GB)
Rate: 6894.4 KB/s
```

**18 ore di scrittura  
per 445 GB di dati  
(Posta/FULL)  
6.8 MB/sec**

# I risultati

- L'architettura di backup presentata, ci ha permesso di:
  - Recuperare, in 8/10 occasioni, il contenuto di mailbox che gli utenti hanno “accidentalmente” azzerato (causa download POP3 da nuovo client di posta, a seguito di crash di PC);
  - Recuperare, in due occasioni, il contenuto di mailbox a seguito di “buchi” nei client, riscontrati successivamente ai 7 giorni di “on-line”;
  - Estendere il backup:
    - Ad alcuni server “utente” in DMZ;
    - Al desktop di un utente “politicamente importante” (+/- 200GB)

Soddisfatti  
al  
**100%**





# In conclusione

## Vantaggi

- L'impiego di LVM+Snapshot/Rsync è risultato stabile e comodo;
- BACULA è più che perfetto nella gestione del backup, pilotando una tape-library IBM TS/3100;
- Il backup client/server è gestito anche su sistemi Win (XP, 2003, Vista)
- La documentazione disponibile è rilevante ed esaustiva

## Criticità

- È consigliabile ospitare MySQL su un server ad-hoc;
- L'impiego di una LAN ad-hoc, GigaEth, è fortemente consigliata;

## Criticità

- Noi non ne abbiamo trovate! :-)  
(...e non parlatemi di "supporto")



# È tutto....

---

# Domande?

Commenti, Suggestimenti e Critiche sono  
**\*più che bene accette\***

Dr. Damiano Verzulli  
damiano@verzulli.it