Bacula
The Network Backup Solution

Presented by Kern Sibbald at FOSDEM
25 February 2007 in Brussels

Bacula – the Network Backup Tool for Linux, Mac, Unix and Windows

*It comes by night and sucks the vital essence from your computers.*
Introduction

Do you do backups?

- No
- Yes, I did one last month
- Yes, tarballs every week
- Sometimes I rsync ...
- Yes, CDs every week

Problems with the above that Bacula solves by:

- Finding the files you need to restore (GUI)
- Restoring to a point in time
- Knowing what was backed up when
- Scaling to handle 2000 machines
- Providing a bare-metal recovery (non-trivial)
**Introduction**

Bacula is a network backup solution, designed for Linux, Mac OS, Unix and Windows systems.

Project goals are to:

- backup any client from a Palm to a mainframe computer
- provide “Enterprise” features similar to the largest commercial applications
- assure data compatibility for 30 years (providing you have the appropriate hardware)
- use a Free and Open Source (GPL v2 + clarifications) license
The Network Backup Solution

Project History and Statistics

- **Bacula = Backup + Dracula**
- January 2000 – Project started
- 14 April 2002 – First release to Source Forge (version 1.16)
- 29 June 2006 – Release 1.38.11
- January 2007 – Release 2.0.0 (current 2.0.2)
- 41 Source Forge project members
- 13 (35) developers with subversion write access
- 656 bugs.bacula.org
- 1,444 bacula-users@lists.sourceforge.net
- 712 bacula-announce@lists.sourceforge.net

Downloads
- 74 version 1.16
- 10,919 version 1.36
- 12,408 version 1.38.11

*It comes by night and sucks the vital essence from your computers.*
The Six Bacula Components

1. Director (DIR)

- Control and administration for everything is centralized
- Basic unit is a Job (one client, one set of files, ...)
- Schedules, initiates and supervises all Jobs
- Maintains the catalog
- Typically one Director except in very large shops
The Six Bacula Components

2. File daemon or Client (FD)

- Does file backup, restore and verification as requested by Director
- Installed on each machine as a service (daemon)
- Communicates over network with Director and Storage daemon
- Needs access to all files to be backed up (root)
- Common code but adapted specifically to each OS
- Typically multiple File daemons per Director; one for each machine
The Six Bacula Components

3. Storage daemon (SD)
- Reads and writes data to the physical medium
  - Disk
  - Tape
  - DVD
  - USB
- Accepts orders and authorization from the Director
- Accepts and returns data to/from File daemons (FD)
- Sends file storage location to Director -> Catalog
- Typically one per Director
4. Console

- Allows user or administrator to control Bacula
- Communicates with Director via network
- Start jobs, review Job output, query/modify catalog
- Consoles available
  - TTY (bconsole)
  - wxWidgets (GUI) – Linux, Unix, Win32
  - Gnome (GUI)
  - Several web interfaces
  - Comprehensive Qt 4 console being developed (bat)
- Restricted consoles (ACL) for security
The Six Bacula Components

5. Catalog database

- Only component not written by Bacula team
- SQL database (MySQL, PostgreSQL, or SQLite) - unique
- Tracks Jobs run, Volumes used, File locations, ...
- Permits rapid restores
- Allows inquiry of when and where files were backed up
- Old data automatically pruned by Director
- Supports multiple databases of same vendor - scaling
The Six Bacula Components

6. Tray Monitor

- Gnome/KDE/Win32 GUI tray applet
- Monitors Director, File daemons, Storage daemon
- Near real-time display of activity
Main Components

The Network Backup Solution

- Console
  - User Commands
  - Commands
    - Authorization
      - File Attributes
      - Storage Location
  - File Attributes
  - Storage Location

- Director
  - File Attributes
  - Storage Location

- Catalog
  - SQL DBMS
  - MySQL
  - PostGreSQL
  - SQLite

- File Daemon
  - File Attributes + Data
  - File Attributes + Data

- Storage Daemon
  - File Attributes + Data
  - Physical Media

It comes by night and sucks the vital essence from your computers.
The Network Backup Solution

Features

- A central server and catalog with distributed backup
- All components communicate via the network and are deployed separately.
- Internal scheduler for automatic and simultaneous job execution with priorities.
- Interactive restore of one or more files from:
  - current backup
  - prior backup of time and date
  - list of files/directories to restore
  - restore by JobId
  - ...
- Simple administration with consoles (command line, GUI, and web)
Features (cont.)

- Labeled Volumes, to prevent accidental overwriting
- Support for ANSI / IBM labels
- Machine independent Volume data format - extensible
- Support for Unicode on Win32; UTF-8 on Unix
- Python interpreter for user “event” (job start, end, ...) scripting
- Rescue CDROM for “bare metal” recovery.
Bacula – Hardware Features

- Backups can span multiple volumes
- Multiple backups (jobs, clients, OSes) per volume
- Supports most tape drives with configurable Device resources
- Support for multiple drive autochangers (libraries)
- Supports tape barcode readers
- Extensive Pool and Volume library management
- Rapid restoration of individual files (one user reported 4 to 6 hours with tar and 3 to 4 minutes with Bacula!).
Bacula – Security Features

- Daemon authorization with CRAM-MD5
- Director and Storage daemon can be run non-root
- MD5, SHA1, ... signatures for each file
- CRC checksum for each Volume block
- Restricted consoles and tray-monitors
- Communications (TLS) encryption
- Data (PKI) encryption
- Tripwire like intrusion detection (Verify)
Technical Highlights

- OS support: Linux (all versions including the zSeries), Win32, Solaris, *BSD, Mac OS X, Irix, Tru64, AIX, HP-UX
- Backup has disk spooling capability to avoid “shoe-shine” on tapes
- Backup/restore of POSIX Access Control Lists (ACL), Mac resource forks, Win32 permissions
- Support for large files (>2GB) and 64 bit architectures
- Multi-thread implementation
- Originally written in C, now converted to a subset of C++
Bacula – Director Configuration File

Director {
  Name = bacula-dir
  Query File = “/usr/local/etc/query.sql”
  Working Directory = “/var/bacula”
  PID Directory = “/var/run”
  Maximum Concurrent Jobs = 20
  Password = “secret”
  Messages = Standard
}

Console {
  Name = Monitor
  Password = “monitor-secret”
  CommandACL = status, .status
  CatalogACL = BackupDB
}

It comes by night and sucks the vital essence from your computers.
Bacula Configuration - Job

Jobs are the basic unifying structure

- Name – unique name
- Type – what to do: backup, Backup, Migrate, Admin, Restore
- Level – level of detail of type: Full, Differential, Incremental
- FileSet – what to files to backup
- Client – where to get the files (machine name)
- Storage – where to put the files (which hardware)
- Pool – which set of Volumes (tapes, disk) to use
- Schedule – when to do it
Bacula – Director Configuration File

Job {
    Name = “Laptop”
    Type = Backup
    Client = laptop-fd
    FileSet = “Full Set”
    Schedule = “Weekly Cycle”
    Storage = File
    Messages = Standard
    Pool = Standard
    Write Bootstrap = “/var/bacula/laptop.bsr”
    Priority = 10
}

It comes by night and sucks the vital essence from your computers.
Client {
    Name = laptop-fd
    Address = laptop.example.org
    Catalog = MyCatalog
    Password = “secret-fd”
    File Retention = 30 days
    Job Retention = 6 months
    AutoPrune = yes
    Maximum Concurrent Jobs = 20
}
Bacula Configuration – FileSet

- Include/Exclude files and/or directories
- Regex or wildcard for file/directory name selection
- Compression using similar selection criteria
- Which filesystem types to backup
- ACL support
- Sparse file handling
- Signature (MD5, SHA1, ...)

*It comes by night and sucks the vital essence from your computers.*
FileSet {
  Name = "Full Set"
  Include {
    Options {
      signature=SHA1;    sparse = yes
      regex = ".*/c$";    wild = "*.txt"
      exclude = yes
    }
    File = /
    File = /usr
    File = /var
  }
  Exclude {
    File = /proc;    File = /tmp;    File = /sys;    File = /.journal
  }
}
Bacula – Director Configuration File (cont)

Schedule {
    Name = “Weekly Cycle”
    Run = Level=Full 1st sun at 2:05
    Run = Level=Differential 2nd-5th sun at 2:05
    Run = Level=Incremental mon-sat at 2:05
}
Bacula – File daemon Configuration File

FileDaemon {
    Name = laptop-fd
    Working Directory = /var/bacula
    PID Directory = /var/run
}

Director {
    Name = bacula-dir
    Password = “secret-fd”
}
Bacula – Storage Configuration File

Storage {
    Name = bacula-sd
    Working Directory = /var/bacula
    PID Directory = /var/run
}

Director {
    Name = bacula-dir
    Password = “secret-sd”
}
Bacula – Storage Configuration File (cont)

Device {
    Name = File
    Archive Device = /var/bacula/backups
    Device Type = File # DVD, FIFO, Tape
    Media Type = File
    Label Media = yes
    Random Access = yes
    ...
}
Bacula – Storage Configuration File (cont)

AutoChanger {
    Name = LTO-Changer
    Device = Drive-0, Drive-1
    Changer Device = /dev/sg0
    ...
}

Device {
    Name = Drive-0
    Archive Device = /dev/nst0
    Device Type = Tape  # DVD, File, FIFO
    Media Type = LTO-2
    Autochanger = yes
    ...
}

It comes by night and sucks the vital essence from your computers.
Developers

Most developers believe that backup programs are not very “sexy”. Perhaps, but they are extremely complicated:

- Database – SQL
- GUI
- Web
- Networking
- OS (every bit must be restored)
- Restore
  - Full, Diff, Inc
  - Find Volumes
  - Directory tree (100 million records, 5 million files)
  - where is the file?
- Job resource allocation (drives, Volumes)
Project development

Site: http://www.bacula.org/

Development style:
- SourceForge project
- Developer's guide with code style guidelines
- Developer SVN access. Currently 13 developers may commit
- Patches and commits reviewed by K. Sibbald
- All code tested using a regression test suite
- Email list for developers (bacula-devel)

License:
- GPL 2 (+ clarifications) copyright assigned to FSFE.
- Freedom Task Force (FTF)
Resources

For users and system administrators

- OS and Hardware compatibility lists (in manual)
- Bugs reports: [http://bugs.bacula.org/](http://bugs.bacula.org/)
- Email support list: bacula-users@lists.sourceforge.net

For developers

- Email list: bacula-devel@lists.sourceforge.net,
  bacula-commits@lists.sourceforge.net
- SVN at Source Forge
The Network Backup Solution

Credits

Thanks

- Dan Langille who created the original presentation
- Karl Cunningham who updated it
- This presentation draws heavily on their work

A .pdf copy of this presentation can be found at:

http://www.bacula.org -> Presentations -> ...
Live Demo

The following slides are not part of this slide show but are screenshots of the live demo that will be presented after the slide show.
The Network Backup Solution

It comes by night and sucks the vital essence from your computers.
The Network Backup Solution

It comes by night and sucks the vital essence from your computers.
The Network Backup Solution

Information about job

<table>
<thead>
<tr>
<th>JobId</th>
<th>Client</th>
<th>Job Name</th>
<th>FileSet</th>
<th>Level</th>
<th>StartTime</th>
<th>Duration</th>
<th>JobFiles</th>
<th>JobBytes</th>
<th>Errors</th>
<th>Pool</th>
<th>Volume Name</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>1433</td>
<td>Rufus</td>
<td>Rufus</td>
<td>RufusAll</td>
<td>I</td>
<td>2007-01-10 03:08:50</td>
<td>00:04:45</td>
<td>3695</td>
<td>657.2 Mb</td>
<td>0</td>
<td>Default</td>
<td>LTO-004</td>
<td>✓</td>
</tr>
</tbody>
</table>

Log: Rufus on Rufus (1433)

```
| roxie-dir: Start Backup JobId 1433, Job=Rufus.2007-01-10 03:05:01 |
| roxie-sd: Spooling data ... |
| rufus-fd: /boot is a different filesystem. Will not descend from / into /boot |
| rufus-fd: /home is a different filesystem. Will not descend from / into /home |
| rufus-fd: /tmp is a different filesystem. Will not descend from / into /tmp |
| rufus-fd: /usr is a different filesystem. Will not descend from / into /usr |
| rufus-fd: /sys is a different filesystem. Will not descend from / into /sys |
| rufus-fd: /dev is a different filesystem. Will not descend from / into /dev |
| rufus-fd: /mnt is a different filesystem. Will not descend from / into /mnt |
| rufus-fd: /net is a different filesystem. Will not descend from / into /net |
| rufus-fd: /smb is a different filesystem. Will not descend from / into /smb |
| rufus-fd: /misc is a different filesystem. Will not descend from / into /misc |
| roxie-sd: Job write elapsed time = 00:04:26, Transfer rate = 2.592 M bytes/second |
| roxie-sd: Committing spooled data to Volume "LTO-004". Despooling 690,353,872 bytes ... |
| roxie-sd: Despooling elapsed time = 00:00:11, Transfer rate = 62.75 M bytes/second |
| roxie-sd: Sending spooled attrs to the Director. Despooling 1,025,113 bytes ... |
| roxie-dir: Bacula 2.0.0 (28Dec06): 10-Jan-2007 03:13:35 |
```

JobId: 1433
Job: Rufus.2007-01-10 03:05:01
Backup Level: Incremental, since=2007-01-09 03:08:20
Client: "Rufus" 1.39.34 (28Dec06) i686-pc-linux-gnu,suse,10.2
FileSet: "RufusAll" 2005-12-26 19:31:20
Pool: "Default" (From Job resource)
Storage: "LTO-changer" (From Job resource)
Scheduled time: 10-Jan-2007 03:05:00
Start time: 10-Jan-2007 03:08:50
End time: 10-Jan-2007 03:13:35
Elapsed time: 4 mins 45 secs
Priority: 10
FD Files Written: 3,695
It comes by night and sucks the vital essence from your computers.
The Network Backup Solution

It comes by night and sucks the vital essence from your computers.

FileSet RufusAll

What is included:

/  
/boot  
/home  
/usr

What is excluded:

/home/kern/bacula/bin/working  
/home/kern/bacula/working  
.journal  
.autofsck  
/proc  
/lost+found  
/var/tmp  
/var/lock  
/var/spool/cups/tmp  
/usr/tmp

Tips: Warning, this is the current filesystem, it could have changed...
## Media

<table>
<thead>
<tr>
<th>Volume Name</th>
<th>Online</th>
<th>Vol Bytes</th>
<th>Vol Usage</th>
<th>Vol Status</th>
<th>Pool</th>
<th>Media Type</th>
<th>Last Written</th>
<th>When expire?</th>
<th>Select</th>
</tr>
</thead>
<tbody>
<tr>
<td>LTO-004</td>
<td></td>
<td>277.3 Gb</td>
<td></td>
<td>Append</td>
<td>Default</td>
<td>LTO-2</td>
<td>2007-01-10 05:04:44</td>
<td>2008-01-10 05:04:44</td>
<td></td>
</tr>
</tbody>
</table>
The Network Backup Solution

It comes by night and sucks the vital essence from your computers.
# The Network Backup Solution

It comes by night and sucks the vital essence from your computers.

## Clients

<table>
<thead>
<tr>
<th>Name</th>
<th>Select</th>
<th>Desc</th>
<th>Auto Prune</th>
<th>File Retention</th>
<th>Job Retention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matou</td>
<td></td>
<td>1.39.29 (04Dec06) i686-pc-linux-gnu,redhat,(Stentz)</td>
<td>1</td>
<td>30 days</td>
<td>365 days</td>
</tr>
<tr>
<td>MatouVerify</td>
<td></td>
<td></td>
<td></td>
<td>30 days</td>
<td>30 days</td>
</tr>
<tr>
<td>Minou</td>
<td></td>
<td>Windows NT 4.0,MVS,NT 4.0.1381</td>
<td>1</td>
<td>30 days</td>
<td>365 days</td>
</tr>
<tr>
<td>Polymatou</td>
<td></td>
<td>i686-pc-linux-gnu,redhat,(Stentz)</td>
<td>1</td>
<td>30 days</td>
<td>365 days</td>
</tr>
<tr>
<td>Roxie</td>
<td></td>
<td>2.0.0 (28Dec06) i686-pc-linux-gnu,suse,10.2</td>
<td>1</td>
<td>30 days</td>
<td>365 days</td>
</tr>
<tr>
<td>Rufus</td>
<td></td>
<td>1.39.34 (28Dec06) i686-pc-linux-gnu,suse,10.2</td>
<td>1</td>
<td>30 days</td>
<td>365 days</td>
</tr>
<tr>
<td>RufusVerify</td>
<td></td>
<td>i686-pc-linux-gnu,redhat,(Bordeaux)</td>
<td>1</td>
<td>30 days</td>
<td>30 days</td>
</tr>
<tr>
<td>Tibs</td>
<td></td>
<td>2.0.0 (04Jan07) Linux,Cross-compile,Win32</td>
<td>1</td>
<td>30 days</td>
<td>365 days</td>
</tr>
<tr>
<td>Timmy</td>
<td></td>
<td>i686-pc-linux-gnu,redhat,(Stentz)</td>
<td>1</td>
<td>30 days</td>
<td>365 days</td>
</tr>
<tr>
<td>timmy-tld</td>
<td></td>
<td>2.0.0 (28Dec06) i686-pc-linux-gnu,suse,10.2</td>
<td>1</td>
<td>30 days</td>
<td>365 days</td>
</tr>
<tr>
<td>Watchdog</td>
<td></td>
<td></td>
<td>1</td>
<td>24 hours</td>
<td>30 days</td>
</tr>
</tbody>
</table>

## Actions

- Last jobs
- Current jobs
- Status
- Stats
The Network Backup Solution

It comes by night and sucks the vital essence from your computers.

Client: Rufus (last 7 days)

<table>
<thead>
<tr>
<th>Name</th>
<th>Nb Jobs</th>
<th>Nb Bytes</th>
<th>Nb Files</th>
<th>Nb Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rufus</td>
<td>7</td>
<td>5.3 Gb</td>
<td>44540</td>
<td>0</td>
</tr>
</tbody>
</table>

Job Duration: 'Rufus'/all

Job Rate: 'Rufus'/all

Job Size: 'Rufus'/all