
Date-stamped Backups Implemented with Hard Links and Rsync

Raymond E. Marcil
marcilr@gmail.com

Revision 25 (November 19, 2012)

Abstract

This document contains the design of a script to generate date-stamped backups using hard links and rsync. The script will generate backups for every day for the past week, every month for the past year, and every prior year.

Contents

Contents	2
List of Figures	3
List of Tables	3
1 Introduction	5
2 Hard Links and Rsync	6
2.1 Hard Links	6
2.2 Rsync	6
3 Examples	7
3.1 Escaping < and > Symbols	7
3.2 Enumerate	7
3.3 Comments	7
3.4 Footnotes	7
3.5 Simple Table Examples	8
3.6 Another Simple Table Example	8
3.7 Verbatim	9

List of Figures

1	Create Hard Link	6
2	File and Directory Structure	9

List of Tables

1	EASEMENTS_17B Table	8
2	USS XML index elements	8
3	XREF_MTR_QMQ Table	8

Definitions and Abbreviations

- **hard link** - The term is used in file systems which allow multiple hard links to be created for the same file. This has the effect of creating multiple names for the same file, causing an aliasing effect: e.g. if the file is opened by one of its names, and changes are made to its content, then these changes will also be visible when the file is opened by an alternative name. By contrast, a soft link on such file systems is not a link to a file itself, but to a file name. This also creates aliasing, but in a different way. Directories are files, so multiple hard links to directories are possible; however, their unrestricted creation is usually avoided, because of the cyclic structures this may create.

http://en.wikipedia.org/wiki/Hard_link

- **rsync** - A fast, versatile, remote (and local) file-copying tool.

1 Introduction

The goal is to create a script that will generate backups for every day for the past week, every month for the past year, and every prior year.

```
cp -al
```

Figure 1: Create Hard Link

2 Hard Links and Rsync

The backup script uses hard links and rsync for easy backups and deduplication of data.

2.1 Hard Links

With “hard links it is possible to associate multiple directory entries with a single inode.”¹ This facilitates easy deduplication for date-stamped backups created using hard links. Inodes are only deleted when all hard links to the inode are deleted.²

“The standard GNU coreutils `cp` command comes with a `-l` flag that causes it to create (hard) links instead of copies (it doesn’t hard-link directories, though, which is good; you might want to think about why that is). Another handy switch for the `cp` command is `-a` (archive), which causes it to recurse through directories and preserve file owners, timestamps, and access permissions.

Together, the combination `cp -al` makes what appears to be a full copy of a directory tree, but is really just an illusion that takes almost no space. If we restrict operations on the copy to adding or removing (unlinking) files—i.e., never changing one in place—then the illusion of a full copy is complete. To the end-user, the only differences are that the illusion-copy takes almost no disk space and almost no time to generate.”³

2.2 Rsync

Rsync is a fast, versatile, remote (and local) file-copying tool. It can be used to easily create a date-stamped backup from the original files.

¹[Understanding UNIX / Linux symbolic \(soft\) and hard links](#)

²[How to find and delete all hard links to a file](#)

³[Easy Automated Snapshot-Style Backups with Linux and Rsync](#) by Mike Rubel

3 Examples

Series of useful L^AT_EX markup. Need to break out to separate examples.tex file.

3.1 Escaping < and > Symbols

To get \$<\$ or \$>\$ just wrap the symbols in \$ for math mode.

3.2 Enumerate

1. DNR - Alaska State Department of Natural Resources
 - HI - Historical Index, not maintained since 1982
 - LE - Land Estate, maintained by SGU
 - ME - Mineral Estate, maintaind by SGU
2. Alaska State Surveys
 - ASBLT - As-Built Survey
 - ASCS - Cadastral Survey

3.3 Comments

COMMENTS Comment — *Sean Weems, Spring 2003*

We should get the **COMMENTS** column searchable via the landrecords application before we do much anything else – shouldn't be too hard.

Errata: Plats spanning multiple sections

A few anomalies can be observed in the **AKPLATS** table. Specifically plats exist that span multiple sections. Since the table only has a single column, **SCODE**, that accepts a single section code, SGU (Status Graphics Unit) has handled this problem by entering multiple rows in the table, each with a different section that point to the same plat or file. Multiple section plats are indicated by setting the **TCODE** column to the value 37, and making an appropriate notation like *Section 24-25-26-27* in the **REMARKS** column.

[FIXME: Perhaps the **SCODE** column should accept an array of sections?]

3.4 Footnotes

Some footnotes here ?? for an example. Yet another 2 example.

3.5 Simple Table Examples

Column Name	Type	Description
EQS	VARCHAR2(1)	!NULL map shows village selections
ITM_COL	VARCHAR2(1)	USGS ITM column: 1-6
ITM_ROW	VARCHAR2(1)	USGS ITM row: A-E
QMQ_ABBR_DNR	VARCHAR2(3)	Three character DNR abbreviation for the QMQ
RASTER_FILENAME	VARCHAR2(50)	Physical path to file
RASTER_PATHNAME	VARCHAR2(50)	URL path to PDF of map
SCODE	VARCHAR2(2)	Supplement map code: 1,2,3,...
COMMENTS	VARCHAR2(256)	Plat comments

Table 1: EASEMENTS_17B Table

XML element	Description
FNUM	US Survey file number
MERIDIAN	BLM meridian code 12 = Copper River 13 = Fairbanks 28 = Seward 44 = Kateel 45 = Umiat
TOWNSHIP	Five character Township code
RANGE	Five character Range code
PAGE	Survey page number 1,2,3,...
FILENAME	Relative path to file in direcorey

Table 2: USS XML index elements

3.6 Another Simple Table Example

Column Name	Type	Description
MTR	VARCHAR2(9)	Meridian, Township, Range, example: <i>C026S054E</i>
QMQ	VARCHAR2(3)	Quarter Million Quadrangle code, example: <i>DIL</i> (Dillingham quadrangle)

Table 3: XREF_MTR_QMQ Table


```

gis/raster/
  dnr/
    map_library/
    plats/
      SP/YYYYMMDD/*.pdf          # indexed
      HI/YYYYMMDD/*.pdf          # Indexed
      ASLS/YYYYMMDD/*.pdf        # Indexed
    recorded-plats/
      YYYYMMDD/*.pdf
  blm/
    easements_17b/YYYYMMDD/*.pdf # indexed
    mtp/YYYYMMDD/*.pdf            # non-indexed
    usrs/YYYYMMDD/*.pdf           # indexed
    usrs-notes/YYYYMMDD/*.pdf     # indexed
    uss/YYYYMMDD/*.pdf            # indexed
    uss-notes/YYYYMMDD/*.pdf      # indexed
    usms/YYYYMMDD/*.pdf           # indexed
    usms-notes/YYYYMMDD/*.pdf     # indexed
  usgs/
    drg/
      collared/
        250K/
        63K/
        25K/
        24/
      decollared/
      tools/
      missing\_data/
    dem/
    doq/
    topo/

```

Figure 2: File and Directory Structure

3.7 Verbatim

Appendix

Links

Build a Home Terabyte Backup System Using Linux

by Duncan Napier - Nov 29, 2005

<http://www.linuxjournal.com/article/8590>

Easy Automated Snapshot-Style Backups with Linux and Rsync by [Mike Rubel](#)

http://www.mikerubel.org/computers/rsync_snapshots/