

Re: Sort files by filename

Source: <http://linux.derkeiler.com/Mailing-Lists/Fedora/2007-08/msg00259.html>

- *From:* Rodolfo Alcazar Portillo <rodolfo.alcazar@xxxxxxxxxxxxx>
 - *Date:* Wed, 01 Aug 2007 17:42:26 -0400
-

Am Dienstag, den 31.07.2007, 14:50 -0400 schrieb Mark Haney:

Rodolfo Alcazar Portillo wrote:

On 7/31/07, Mark Haney <mhaney@xxxxxxxxxxxxxxxxxxxxx> wrote:

Rodolfo Alcazar Portillo wrote:

Am Dienstag, den 31.07.2007, 11:00 -0400
schrieb Miner, Jonathan W (CSC)

...

```
[rodolfoap] /home/rodolfoap > for n in $(seq
10 40); do touch
XXXX20070515_112011_942_${n}.bz2;
done
[rodolfoap] /home/rodolfoap > rm
XXXX20070515_112011_942_11.bz2
[rodolfoap] /home/rodolfoap > rm
XXXX20070515_112011_942_22.bz2
[rodolfoap] /home/rodolfoap > rm
XXXX20070515_112011_942_33.bz2
[rodolfoap] /home/rodolfoap > for n in $(seq
10 40); do if [ ! -e *_${n}.bz2 ]; then echo
NOT FOUND: $n; fi; done
NOT FOUND: 11
NOT FOUND: 22
NOT FOUND: 33
[rodolfoap] /home/rodolfoap >
```

Good luck.

I think this would probably work except I really need to know which ones are out of sequence as is, changing the filename would eliminate that capability and the missing files are the ones I need. See, I'm receiving the files from a separate server and they are sent to

Re: Sort files by filename

me in
this format. I need to know which ones I /don't/ receive so I
know
which ones I need to have resent. Make sense?

Yes. But you must have a pattern with which we can compare your received files. Which is the exact pattern? If you have this clear, we can probably solve the problem with bash.

Waiting...

Okay, here goes, the XXXX part denotes a 4 character symbol for a NWS radar site, the 8 digits after that is the date followed by '_HHMMSS_' of time the data was collected. The next part I'm not certain, but believe it's the scan number of that particular radar scan, then the last 2 digits are the part number of a complete scan. In other words, I have a series of weather radar scans that get broken up into 'radials' (the .bz2 files) and numbered. Our customers get fed this data through our servers, but a couple have reported missing files. I'm trying to track down the problem, but can't go through 2GB of volume scans every day for possibly one missing file.

The files are stored in directories in this format: XXXX/date/ so I have multiple volume scans in each day's directory.

Does this help?

Barely. Mark, I do agree with the rsync solution mentioned early. But, forcing it with bash, what I would do is:

1. Obtain the daily radar sites list

```
[rodolfoap] /home/rodolfoap > ls -1 *.bz2|sed -e "s/^(...)\(.*\)^\1/"|sort|uniq > radarsites.$(/bin/date "+%Y%m%d")
```

...the four points are the four first characters...

...you can have a predetermined file, so you can know if you have a server missing:

```
[rodolfoap] /home/rodolfoap > cat radarsites.$(/bin/date "+%Y%m%d")
```

XXXX

XYXX

XYYX

XYYY

YYYY

```
[rodolfoap] /home/rodolfoap > cat radarsites.official
```

XXXX

XYXX

XYYX

Re: Sort files by filename

XYYY
YYYY
ZZZZ

```
[rodolfoap] /home/rodolfoap > diff radarsites.official radarsites.$(/bin/date "+%Y%m%d")
6d5
< ZZZZ
```

...That means: in the left file (official radar list) stays a value not found on the right file. You can play with the result.

2. Find the minimum and maximum values of each set. That is what you dont say. Is there an obligated minimum? maximum? Anyway, this is a list of your set ranges:

```
[rodolfoap] /home/rodolfoap > for a in $(cat radarsites.$(/bin/date "+%Y%m%d")); do MIN=$(ls -l
${a}*/sort -n|head -1); MAX=$(ls -l ${a}*/sort -n|tail -1);echo $MIN -- $MAX; done
XXXX20070515_112011_942_10.bz2 -- XXXX20070515_112011_942_40.bz2
XYXX20070515_112011_942_10.bz2 -- XYXX20070515_112011_942_40.bz2
XYYX20070515_112011_942_10.bz2 -- XYYX20070515_112011_942_40.bz2
XYYY20070515_112011_942_10.bz2 -- XYYY20070515_112011_942_40.bz2
YYYY20070515_112011_942_10.bz2 -- YYYY20070515_112011_942_40.bz2
```

Now you have all elements to find your missing files. If there is anymore we can help with, please be clear.

Good luck!

Rodolfo Alcazar – rodolfo.alcazar@xxxxxxxxxxxxx
otbits.blogspot.com / counter.li.org: #367962

– Now all of us can talk to the NSA – just by dialing any number.
David Letterman, on National Security Agency's eavesdropping program

—
fedora–list mailing list
fedora–list@xxxxxxxxxxxxx
To unsubscribe: <https://www.redhat.com/mailman/listinfo/fedora–list>