s3archive

1 What is s3archive ?

`s3archive` is a simple python script which can be used to archive the content of an Amazon S3[1] bucket. It is a quick and dirty wrapper around Tim Kay's aws[2]. It can be used to backup the files you store on S3.

2 Download

S3archive is available both as a Pacman[3] package (x86_64 only), or in source form. For source or Pacman installation, please refer to the Generic download and install instructions[4].

- Latest source tarball[5]

- PGP signature[6]

3 Basic concepts

The task of `s3archive` is very basic: given a prefix (or "path"), download all objects stored on S3 whose name begins with the prefix. Those objects will be stored in a regular local directory I call an archive.

If `s3archive` is run again with the same prefix to archive, it will keep or remove the previous archive depending on user requirements. More precisely, `s3archive` has to be given the number of archives to keep on the local filesystem.

`s3archive` does not implement compression, differential archives, or other fancy stuff. It is meant as a very basic and simple tool.

4 Usage

`s3archive` takes exactly three arguments

1. The prefix or S3 path to backup, formed as follows: `$bucket/$prefix`, for example if you want to backup all files with path `foo/bar` in the bucket `baz`, then the argument should be `baz/foo/bar`.

---

[1] `http://aws.amazon.com/s3/`
[2] `http://timkay.com/aws/`
[3] `http://www.archlinux.org/pacman/`
[4] `http://svasey.org/projects/download-install-doc_en.html`
[5] `http://svasey.org/public-repo.svasey.org/src/s3archive.tar.gz`
[6] `http://svasey.org/public-repo.svasey.org/src/s3archive.tar.gz.sig`

2. The directory where to store your archives.

3. The number of archives of the given path to keep on the local filesystem. Note that if there are more than that number, the oldest will be deleted to make room for the new one.

In addition, one command line option can be given: `--awsopts`, which can be used to specify additionnal, space-separated options to pass to `aws`.