

Becoming a CTAN mirror

Jim Hefferon

Have you downloaded something from CTAN that you found useful? Would you like to be useful in return?

Sites on the Comprehensive T_EX Archive Network come in two types. There are three core sites, where every day the archivists place new materials in the file tree. The other sites are mirrors—every night they sync their holdings with a core site and then offer those materials for download by others.

A mirror site, once set up, mostly runs itself.

So, if you are able to run a mirror site, that would be a big help to CTAN and to T_EX users near you, without causing too much trouble for you. This note describes the technical requirements for doing this, with the specific steps to become an official CTAN mirror.

1 Requirements

You need a computer with 20–30 gigabytes of available disk space and an Internet connection with a permanent IP address. If you have a firewall then you must make sure external users can get through to your site. And, you also need some experience running a Unix-like system, such as GNU/Linux or Mac OS X (a Windows-based system could work but the directions below don't cover this case).

2 You are only three steps away

To set up a mirror you need to (1) give visitors access to files by running either a web or FTP daemon, or both, (2) set up your system to sync your holdings with a core site every day, and (3) sign up to be an official mirror.

2.1 Give your visitors access

The two most popular way to offer the files to your visitors are over HTTP and over FTP. The examples below assume that you keep the archive materials in your `/var/ftp/pub/tex-archive` directory.

To offer the materials over HTTP you need a web server. We use Apache.¹ Setting up a server is beyond our scope but here are tips.

- One way to make the archive available is to put a link inside your document root. For example, this command (adjust paths as needed) makes `http://your site/tex-archive` show the top of the archive.

```
ln -s /var/ftp/pub/tex-archive \
    /var/www/html/tex-archive
```

¹ <http://httpd.apache.org>

- You don't want files named `index.html` to be served as the index of its directory. See the directives in Figure 1.

To offer materials over FTP, you must have an FTP daemon running. We use ProFTPD² but there are many others. Setting up this daemon is also beyond our scope but if your documentation does not cover how to allow anonymous access then we suggest choosing different software.

2.2 Sync with a core site

To get the materials, and to keep them up to date, run `rsync`.³ This program does the transfers efficiently, saving both you and us network usage.

You must mirror from one of the core sites. Pick the one nearest you. Use `rsync://rsync.cam.ctan.org/CTAN` for the one in England, or `rsync://rsync.dante.ctan.org/CTAN` for the German site, or `rsync://rsync.tug.ctan.org/CTAN` for the one in the northeast USA. (The examples below use the first.)

The following command will get everything on CTAN and put it on your hard drive.

```
rsync -av --delete \
    rsync://rsync.cam.ctan.org/CTAN \
    /var/ftp/pub/tex-archive
```

The first time that you run it, this command can take quite a long time (an hour or more, perhaps, depending on the connection speed). Before you commit to it, you can try using the `-n` option, as in `rsync -avn --delete ...`, which will say what would be done without doing it.

You must run the `rsync` command every day. To set this up, run `crontab -e` at the command prompt. In the editor that appears enter a line like this (the line breaks are only for *TUGboat*; it must be one line in the editor):

```
31 2 * * * rsync -a --delete
    rsync://rsync.cam.ctan.org/CTAN
    /var/ftp/pub/tex-archive
```

This makes your system run the `rsync` command at 31 minutes past the hour, of the second hour of the day, on every day of the month, and during every month of the year, and every day of the week, that is, Sunday thru Saturday.

Please change these numbers when you set your system up, so that not everyone in the world hits CTAN at the same instant. Pick a time that is in the middle of the night at the location of the core site that you are mirroring.

² <http://www.proftpd.org/>

³ <http://rsync.samba.org>

```

<Directory /> # web visitors can't see outside web tree
  Order Deny, Allow
  Deny from all
</Directory>

<Directory /var/www/html> # web visitors can see the web tree
  Order Allow, Deny
  Allow from all
  Options +FollowSymLinks
</Directory>

<Directory /var/www/html/tex-archive> # soft link to CTAN tree
  Order Allow, Deny
  Allow from all
  Options -ExecCGI, +FollowSymLinks, -Includes, -IncludesNOEXEC, +Indexes
  DirectoryIndex      # no value, so 'index.html' is not used
</Directory>

```

Figure 1: Part of your web configuration file

2.3 Become official

When you have the files, and the cron job, and you are offering public access, visit http://www.ctan.org/mirror_signup to sign up to become an official mirror.

After a day or two, you will be added to a mailing list that is very low volume, perhaps a half dozen messages a year. And, you will be put into the CTAN system, so that you will start seeing traffic.

3 Thanks

If you can run a mirror, then we at CTAN would appreciate it. It is a much needed service — we could not continue without our mirrors.

◇ Jim Hefferon
 Saint Michael's College
 Colchester, Vermont USA
 ftpmaint (at) tug dot ctan dot org

