

# Fedora's MirrorManager - now and the future

Adrian Reber <adrian@lissas.de>

August 02, 2016

Overview

MirrorManager2

Usage

Known Problems

Discussion

# Definitions

- What is MirrorManager
  - Tool to redirect the users to the "best" mirror
- What is the "best" mirror
  - The closest mirror which is up to date
- What is the closest mirror
  - Same continent
  - Maybe Internet2 or other research and education networks
  - Same country
  - Same ASN (maybe peer ASN)
  - Same network

# MirrorManager Features

- Supports partial mirror
- Supports multiple categories (Fedora Linux, EPEL)
- Supports private mirrors (probably unique)
- Self service web interface (probably unique)

# MirrorManager Users

- The Fedora Project
- RPM Fusion

## Other Mirror Management Projects

- Mirrorbits (VideoLAN, SuperRepo)
- MirrorBrain (GNOME, LibreOffice, openSUSE, XFCE)
- Debian's HTTP redirector (reads debian mirror trace files)
- MirMon (only checks a single timestamp file)
- Many unpublished mirror management frameworks

## MirrorManager History

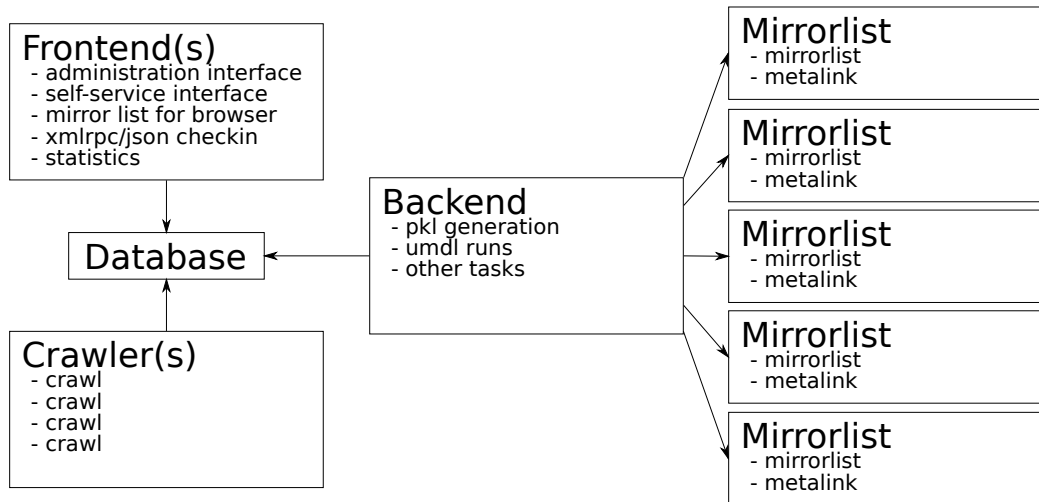
- Development started in January 2007 (git log)
  - TurboGears 1.x based application
- Fedora seems to use it since May 2007 (puppet git log)
- Rewrite required to run on RHEL7
- Pierre-Yves ported MirrorManager to Flask 2014 (git log)
- MirrorManager2 in production since spring 2015

## MirrorManager2 Changes

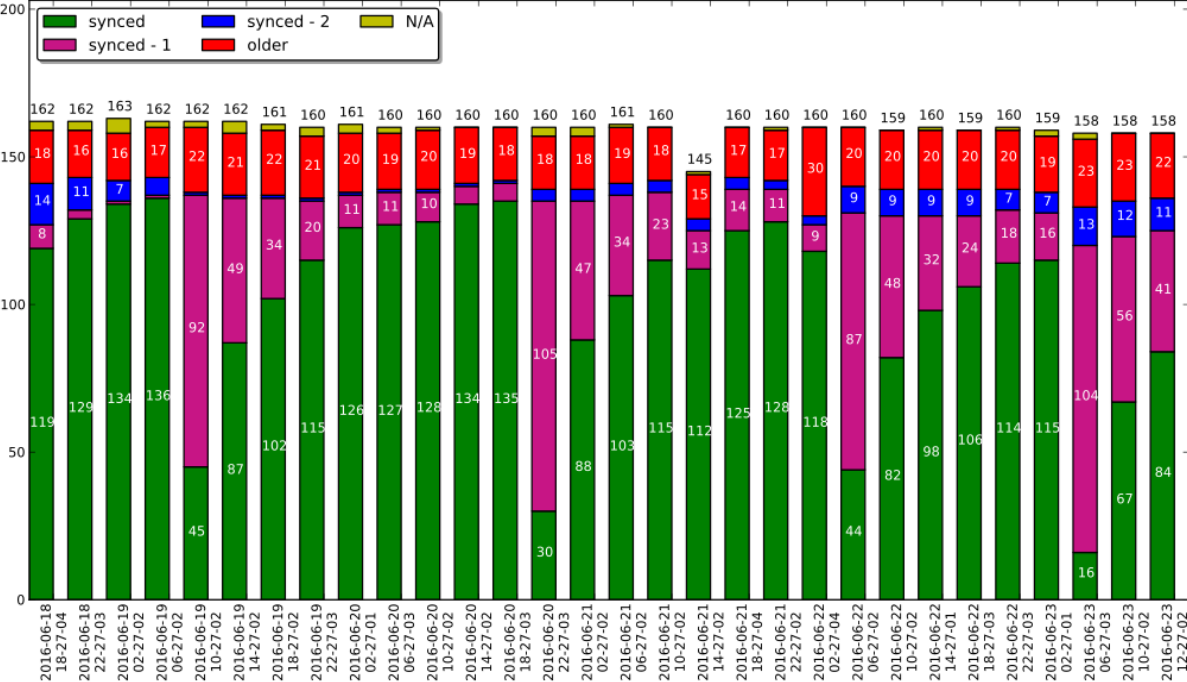
- Flask and SQLAlchemy based
- Crawler automatically disables mirrors
- The mirrorlist and metalink backend uses same data as MM1
- Disabled and removed any FTP occurrence
- Added possibility to query only for HTTPS mirrors
- Moving from cron based to being fedmsg based



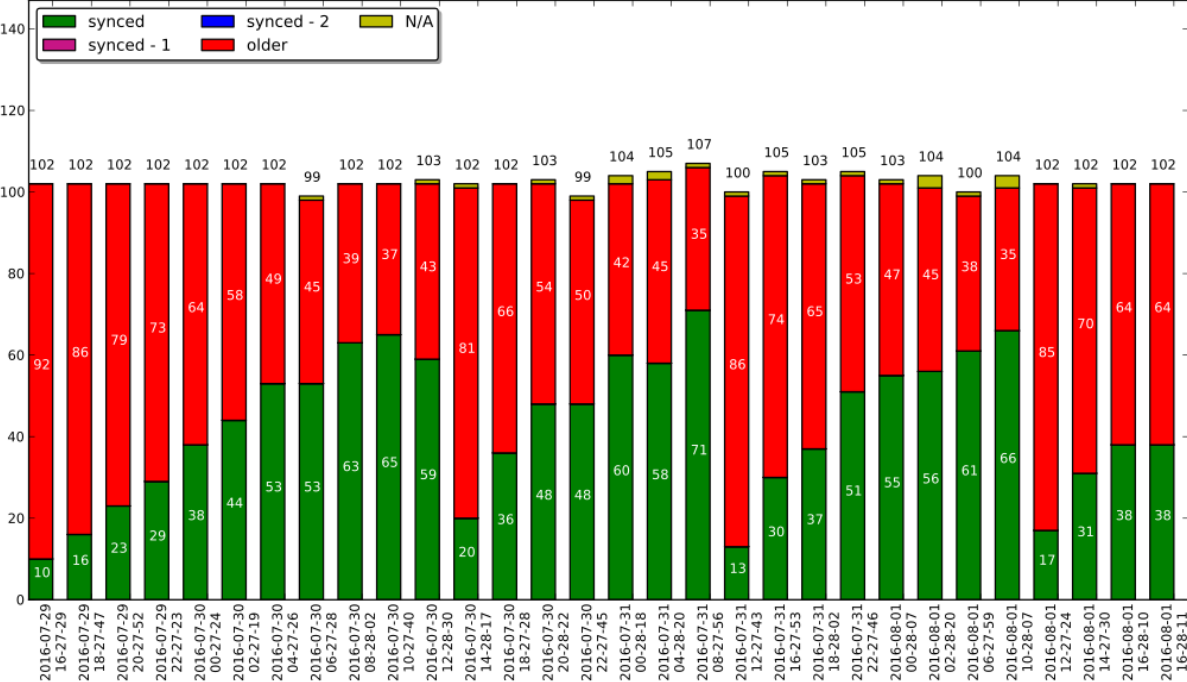
# MirrorManager2 Architecture



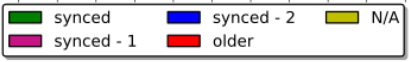
repomd.xml propagation for updates/22/x86\_64/repdata



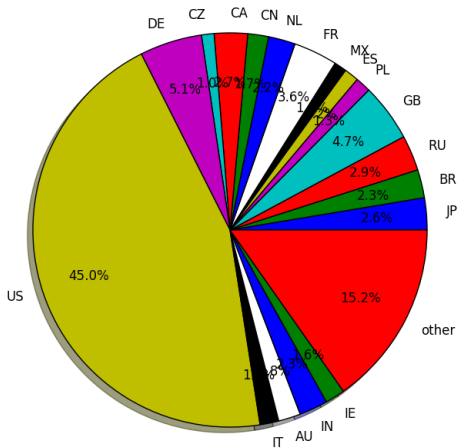
repomd.xml propagation for development/rawhide/Everything/x86\_64/os/repodata



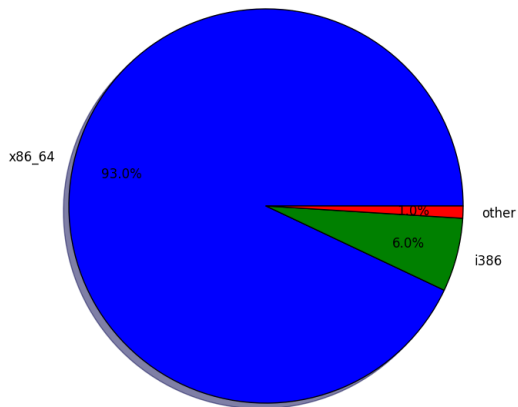
repomd.xml propagation for updates/24/x86\_64/repdata



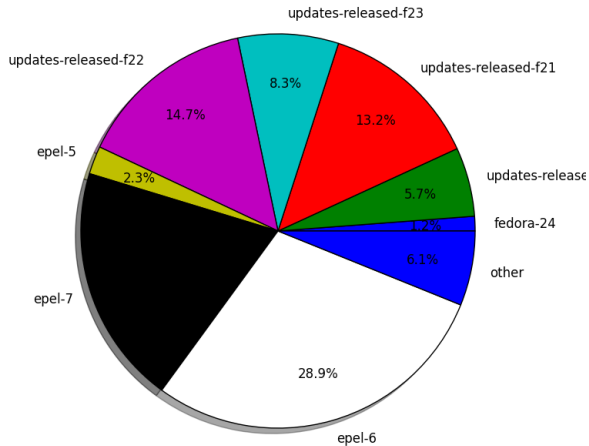
# Countries



# Architectures



# Repositories



## Adding a Mirror

- Site
- Site
  - Host
  - Host
    - Categories Carried (Fedora Linux)
    - Categories Carried (EPEL)
      - URL (RSYNC)
      - URL (HTTP)
      - URL (HTTPS)



## mirrorlist and metalink options

- `https://mirrors.fedoraproject.org/mirrorlist?repo=rawhide&arch=x86_64`
- `https://mirrors.fedoraproject.org/metalink?repo=rawhide&arch=x86_64`
- Additional Parameters
  - repo
  - arch
  - path
  - country
  - netblock
  - version (for CentOS)
  - cc (for CentOS)
  - protocol

# mirrorlist

```
1 $ curl -s "https://mirrors.fedoraproject.org/mirrorlist?repo=rawhide&arch=x86_64&country=pl"
2 # repo = rawhide arch = x86_64 country = PL
3 http://ftp.icm.edu.pl/pub/Linux/fedora/linux/development/rawhide/Everything/x86_64/os/
4 http://ftp.ps.pl/pub/Linux/fedora-linux/development/rawhide/Everything/x86_64/os/
```

```
1 $ curl -s "https://mirrors.fedoraproject.org/mirrorlist?repo=rawhide&arch=x86_64&country=global&protocol=https"
2 # repo = rawhide arch = x86_64 country = global protocol = https
3 https://muug.ca/mirror/fedora/linux/development/rawhide/Everything/x86_64/os/
4 https://ftp-stud.hs-esslingen.de/pub/fedora/linux/development/rawhide/Everything/x86_64/os/
5 https://mirrors.nic.cz/fedora/linux/development/rawhide/Everything/x86_64/os/
6 https://www.mirrorservice.org/sites/dl.fedoraproject.org/pub/fedora/linux/development/rawhide/Everything/x86_64/os/
```

# metalink

```
1 $ curl -s "https://mirrors.fedoraproject.org/metalink?repo=rawhide&arch=x86_64&country=pl"
2 <?xml version="1.0" encoding="utf-8"?>
3 <metalink version="3.0" xmlns="http://www.metalinker.org/" type="dynamic" pubdate="Mon, 01 Aug 2016
   16:38:00 GMT" generator="mirrormanager" xmlns:mm0="http://fedorahosted.org/mirrormanager">
4   <files>
5     <file name="repomd.xml">
6       <mm0:timestamp>1470037417</mm0:timestamp>
7       <size>4385</size>
8       <verification>
9         <hash type="md5">9426210e4ed8de0a41dc4336136e579a</hash>
10        <hash type="sha1">f6e5d61f4dbd5af59b441907eab8d76d6f2330bb</hash>
11        <hash type="sha256">acbc28024aa425da85.....c20dbfecc62f1301b87f829</hash>
12        <hash type="sha512">7d83e.....cc1d285ca9d3ff75d870c1d1ab7785d</hash>
13      </verification>
14      <mm0:alternates>
15        <mm0:alternate>
16          <mm0:timestamp>1469951484</mm0:timestamp>
17          <size>4385</size>
18        <verification>
```

## report\_mirror

- Tool required for private mirrors
- Also used by public mirrors
- Scans the local mirror and transmits the result to MM
- Only scans the directories
  - To minimize local I/O (from 2007)
- If running with different UID/GID than httpd or rsync it can lead to flapping mirrors

# Anycast Mirrors

- MirrorManager cannot handle anycast mirrors
- The crawler will always scan the same server
- GeoIP also gets confused
- Does not work right now

## Metalink Alternates

- Up to three different `repomd.xml` checksums in a metalink
- Checksums are removed from metalink if
  - older than 3 days (`max_stale_days`)
  - second most-recent entry is older than `max_propagation_days` [sic]

## update-master-directory-listing (UMDL)

- Crawls the NFS mounted Fedora tree
- Checks directory CTIMES
- Tries to find new or changed repositories in a second run
- Highly inefficient
- Lots of I/O operations (mainly stats) over 11TB ( $9 \times 10^6$  files)
- Rewrite exists from Pierre-Yves
- Not in production
- The information about the files probably exists somewhere else
- Also needed by `quick-fedora-mirror`

# The end.

Thanks for listening.