

An introduction to



Tracker,

SPARQL,

... and whatnot

Index

- **Introduction**

- What is Tracker?
- A bit of history...
- Tracker in GNOME

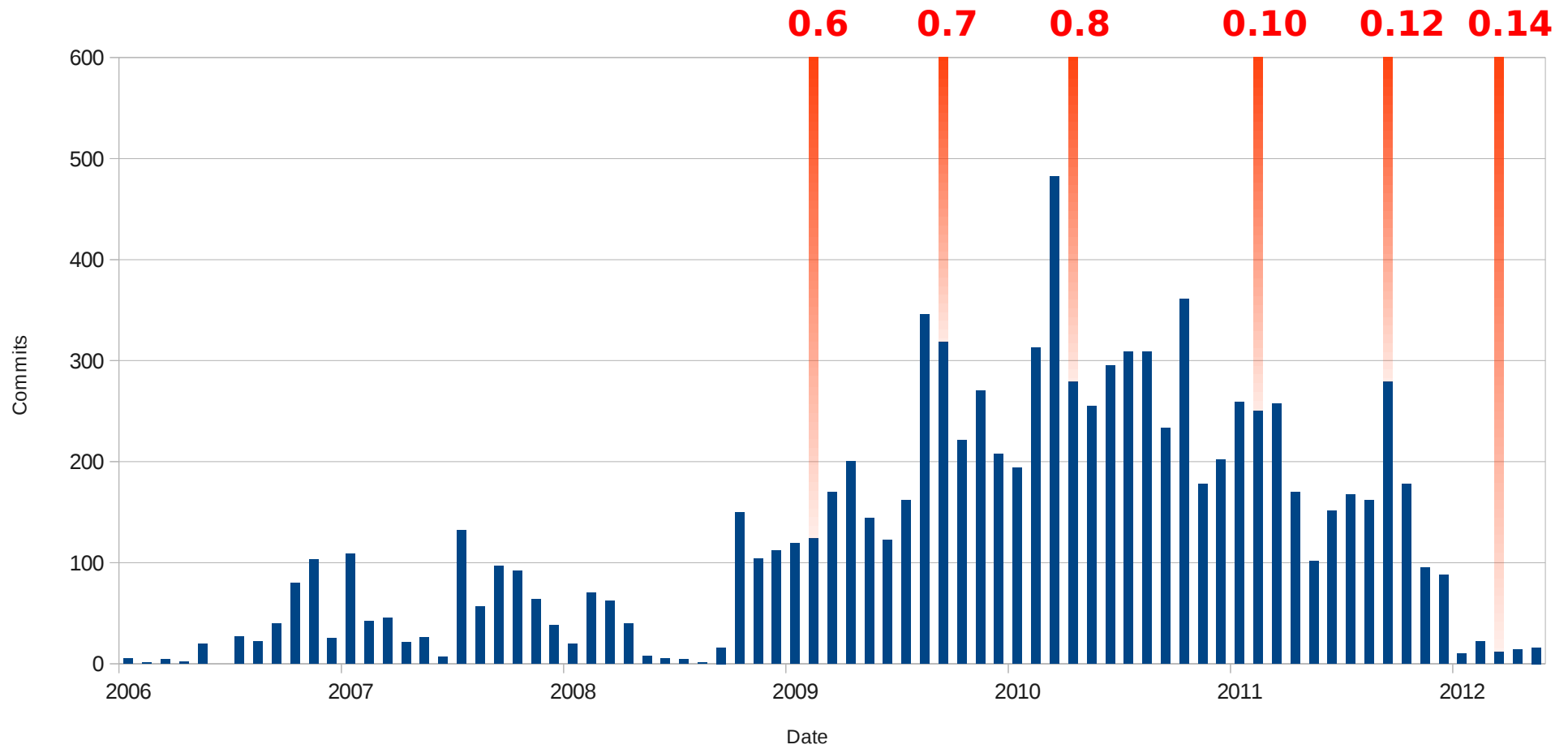
- **Tracker internals**

- Architecture
- Example flow
- SPARQL
- FTS
- Tools

Introduction: What is Tracker?

- Tracker is...
 - ✓ ... a semantic data storage.
 - ✓ SPARQL, sqlite, Nepomuk ontologies
 - ✓ ... a filesystem crawler
 - ✓ GFileMonitor (inotify in Linux)
 - ... a file contents/metadata extractor
 - ✓ poppler, gstreamer...
 - ✓ ... a full text search (FTS) engine.
 - ✓ sqlite's fts3/fts4 plus “fts:” SPARQL extensions

Introduction: a bit of history...



Introduction: Tracker in GNOME

Love or Hate?

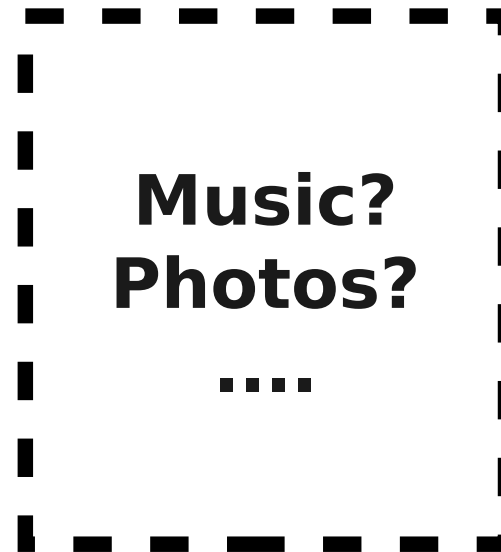


Picture: **GUADec 2010 Hate/Love wall**
© Sense Hofstede, (CC BY-SA)

Introduction: Tracker in GNOME



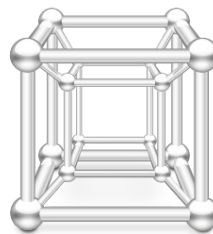
Documents



Needle



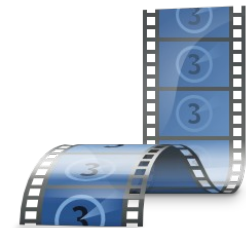
**Nautilus
(Tags)**



Boxes

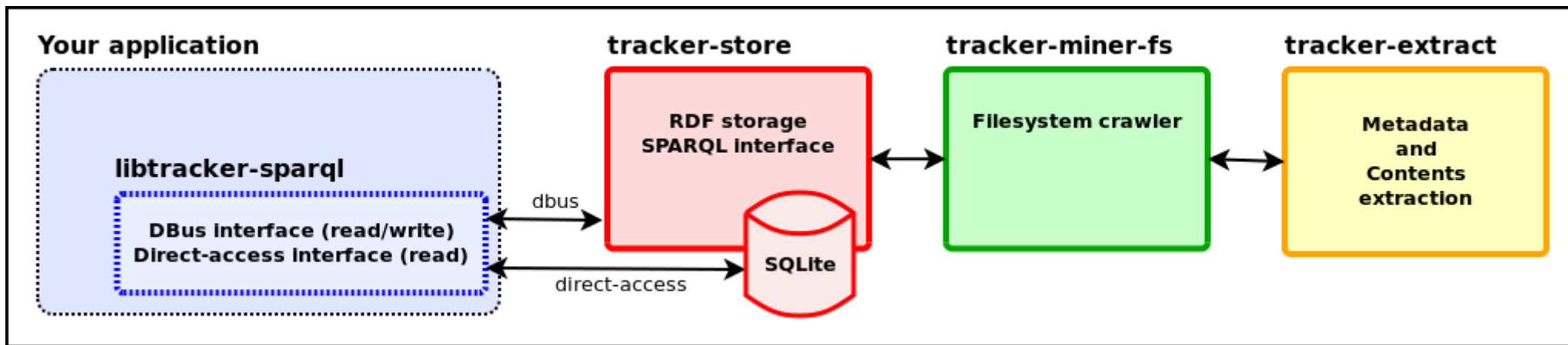


Evolution

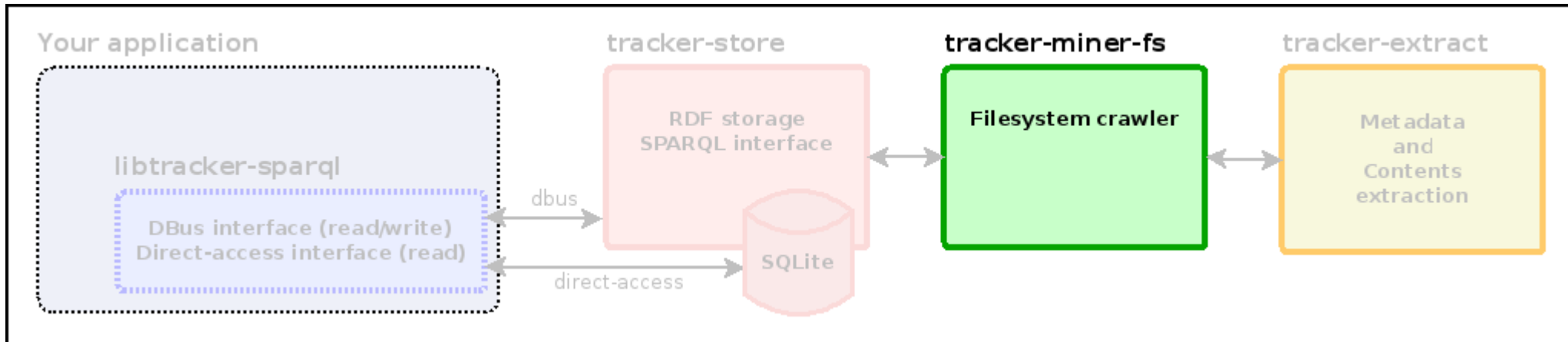


Totem

Tracker internals: Architecture

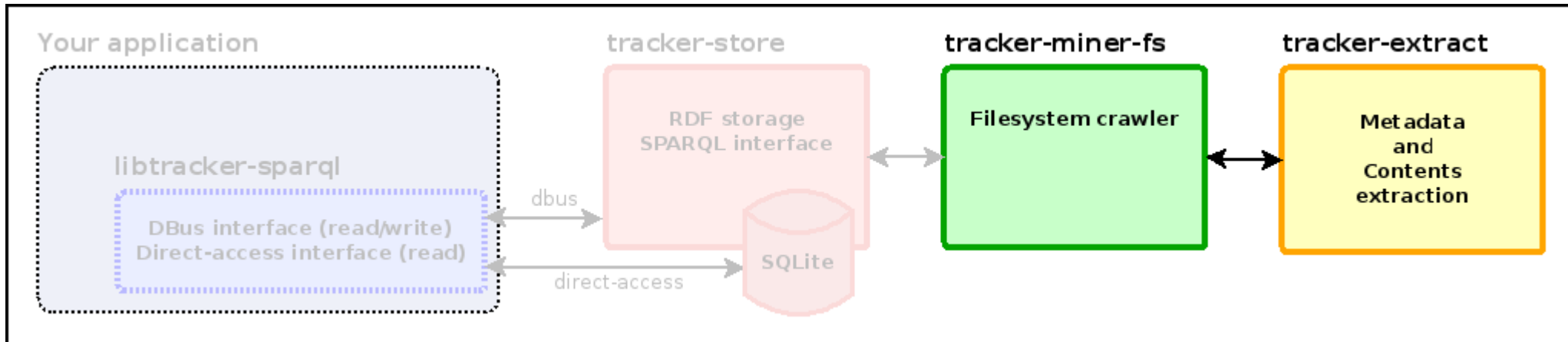


Tracker internals: Example flow



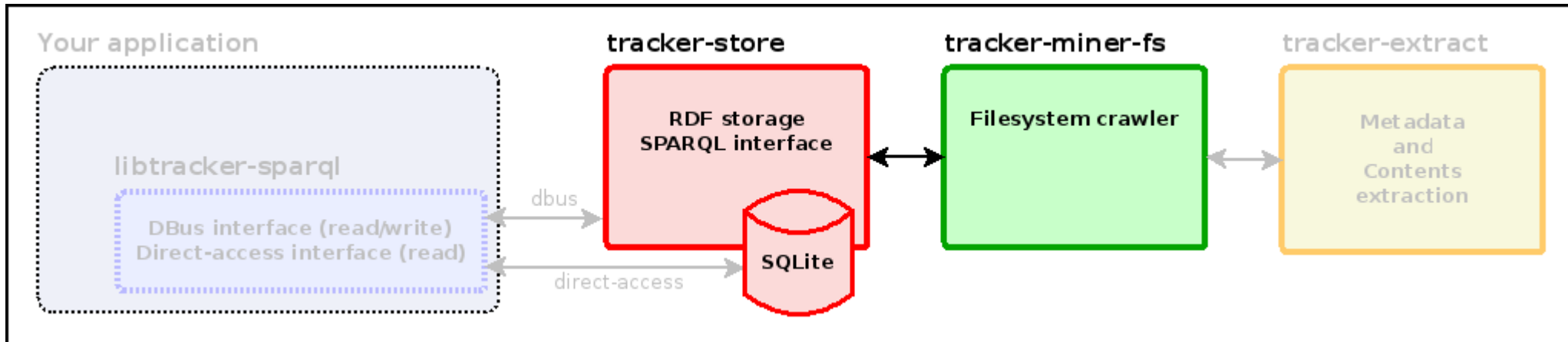
- (1) User downloads a new file,
`tracker-miner-fs` detects it automatically.

Tracker internals: Example flow



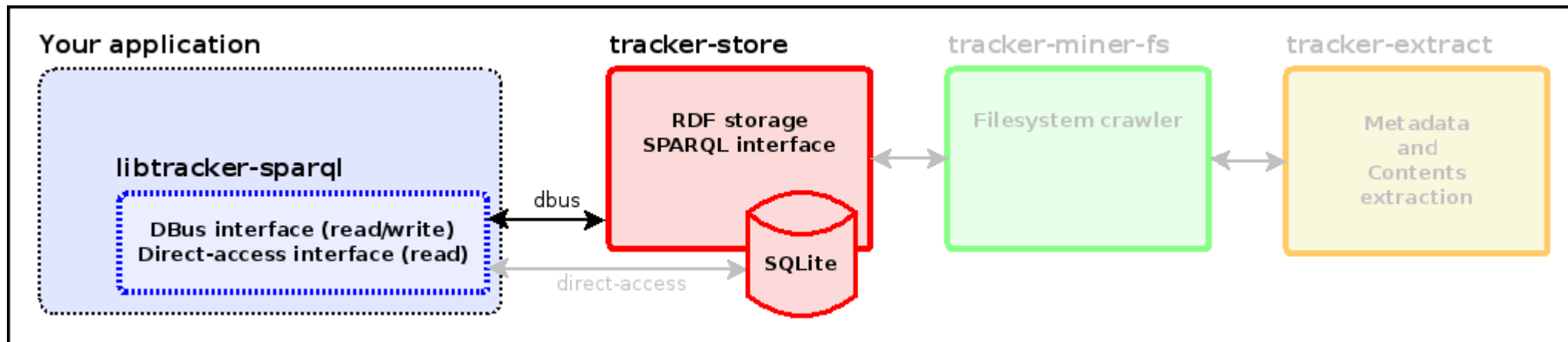
(2) **tracker-miner-fs** tells **tracker-extract** to extract metadata & contents from it.

Tracker internals: Example flow



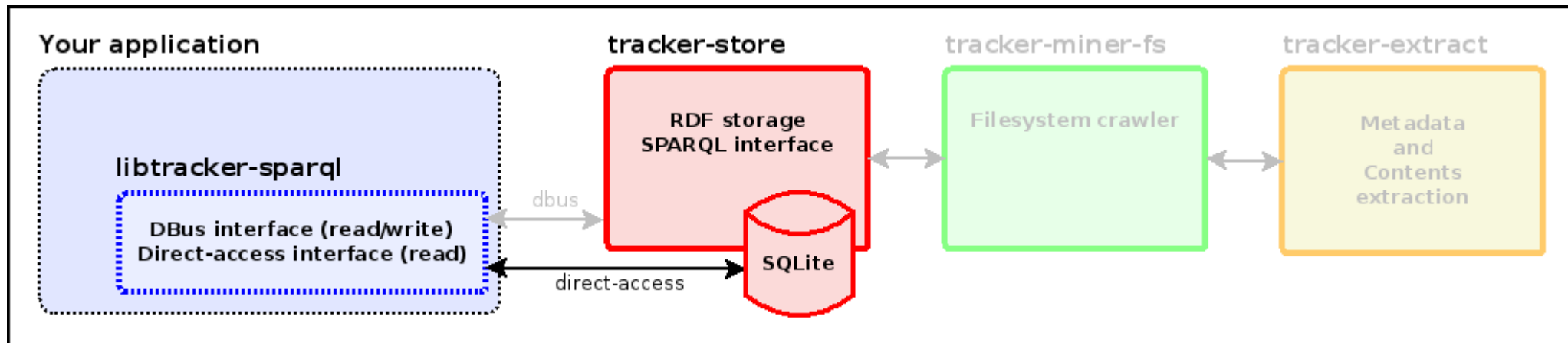
(3) **tracker-miner-fs** collects all extracted info and pushes it to **tracker-store**.

Tracker internals: Example flow



(4) **tracker-store** stores the info in the **SQLite** database, and notifies about the new item via **DBus** to listener **applications**.

Tracker internals: Example flow



(5) The **application** gets the notification about the new file and makes read-only queries to the **SQLite** database directly.

Tracker internals: SPARQL

- Tracker stores **Resources** uniquely identified by URNs
 - urn:uuid:11f22ad4-2e7c-a288-f10a-0b155ddbdfdde
- Each Resource is of a (set of) given **Classes(s)**.
 - nfo:TextDocument, nfo:Document, nfo:FileDataObject, nie:InformationElement, nie:DataObject
- Each Resource has different **Properties**.
 - 'nie:mimeType' = 'application/pdf'
 - The Class itself is specified as “rdf:type” properties
- Classes and Properties are defined by **Ontologies**.
 - 'nie', 'nfo', 'tracker'...
- <https://live.gnome.org/Tracker/Documentation/Examples>

Tracker internals: SPARQL

- **tracker-info** NOFX/TheBrews.ogg

```
'urn:uuid:b679e353-1375-74c0-4305-e7a659b2ee41'
```

Results:

```
'rdf:type' = 'nmm#MusicPiece'
```

```
'nie:mimeType' = 'audio/x-vorbis+ogg'
```

```
'nie:title' = 'The Brews'
```

```
'nfo:genre' = 'Alternative & Punk'
```

```
'nmm:musicAlbum' = 'urn:album:Punk%20in%20Drublic'
```

```
'nmm:performer' = 'urn:artist:NOFX'
```

...

Tracker internals: SPARQL

- Get the location for all music files:

```
$ tracker-sparql -q "
```

```
SELECT ?url
```

```
WHERE {
```

```
  ?song a      nmm:MusicPiece .
```

```
  ?song nie:url ?url .
```

```
}"
```

Tracker internals: SPARQL

```
file:///home/aleksander/Music/NOFX/Punk%20in%20Drublic/10%20-%20Dying%20Degree.ogg
file:///home/aleksander/Music/NOFX/Punk%20in%20Drublic/17%20-%20Scavenger%20Type.ogg
file:///home/aleksander/Music/NOFX/Punk%20in%20Drublic/16%20-%20Reeko.ogg
file:///home/aleksander/Music/NOFX/Punk%20in%20Drublic/02%20-%20Leave%20it%20Alone.ogg
file:///home/aleksander/Music/NOFX/Punk%20in%20Drublic/05%20-%20Don't%20Call%20Me%20White.ogg
file:///home/aleksander/Music/NOFX/Punk%20in%20Drublic/07%20-%20Perfect%20Government.ogg
file:///home/aleksander/Music/NOFX/Punk%20in%20Drublic/12%20-%20Lori%20Meyers.ogg
file:///home/aleksander/Music/NOFX/Punk%20in%20Drublic/15%20-%20Happy%20Guy.ogg
file:///home/aleksander/Music/NOFX/Punk%20in%20Drublic/09%20-%20The%20Quass.ogg
file:///home/aleksander/Music/NOFX/Punk%20in%20Drublic/01%20-%20Linoleum.ogg
file:///home/aleksander/Music/NOFX/Punk%20in%20Drublic/08%20-%20The%20Brews.ogg
file:///home/aleksander/Music/NOFX/Punk%20in%20Drublic/13%20-%20Jeff%20Wears%20Birkenstocks.ogg
file:///home/aleksander/Music/NOFX/Punk%20in%20Drublic/11%20-%20Fleas.ogg
file:///home/aleksander/Music/NOFX/Punk%20in%20Drublic/04%20-%20The%20Cause.ogg
file:///home/aleksander/Music/NOFX/Punk%20in%20Drublic/14%20-%20Punk%20Guy.ogg
```

...

Tracker internals: SPARQL

- Get the song and album titles for all music files of a given artist:

```
$ tracker-sparql -q "
```

```
SELECT ?title ?albumtitle
```

```
WHERE {
```

```
    ?song nmm:performer [ nmm:artistName 'NOFX' ] ;
```

```
    nie:title ?title ;
```

```
    nmm:musicAlbum ?album .
```

```
    ?album nmm:albumTitle ?albumtitle
```

```
}"
```

Tracker internals: SPARQL

Dying Degree, Punk in Drublic
Scavenger Type, Punk in Drublic
Reeko, Punk in Drublic
Leave it Alone, Punk in Drublic
Don't Call Me White, Punk in Drublic
Perfect Government, Punk in Drublic
Lori Meyers, Punk in Drublic
Happy Guy, Punk in Drublic
The Quass, Punk in Drublic
Linoleum, Punk in Drublic
The Brews, Punk in Drublic
Jeff Wears Birkenstocks, Punk in Drublic
Fleas, Punk in Drublic
The Cause, Punk in Drublic
Punk Guy, Punk in Drublic
Dig, Punk in Drublic
My Heart is Yearning, Punk in Drublic

...

Tracker internals: FTS

- Full Text Search is **a side feature**
- Get all documents which contain words starting with the 'gnom*' prefix but which do not contain the 'tracker' word:

```
$ tracker-sparql -q "  
SELECT nie:url(?u)  
WHERE { ?u a          nfo:Document ;  
         fts:match 'gnom* -tracker'  
}"
```

Tracker internals: FTS

Results:

```
file:///home/aleksander/Documents/ModemManager%20talks/FOSDEM2012%20-%20LTE%20and%20ModemManager.pdf
```

```
file:///home/aleksander/Documents/ModemManager%20talks/FOSDEM2012%20-%20LTE%20and%20ModemManager.odp
```

```
file:///home/aleksander/Documents/Repository/notes/dotemacs.el
```

```
file:///home/aleksander/Documents/Repository/notes/gnulinux-ivan.txt
```

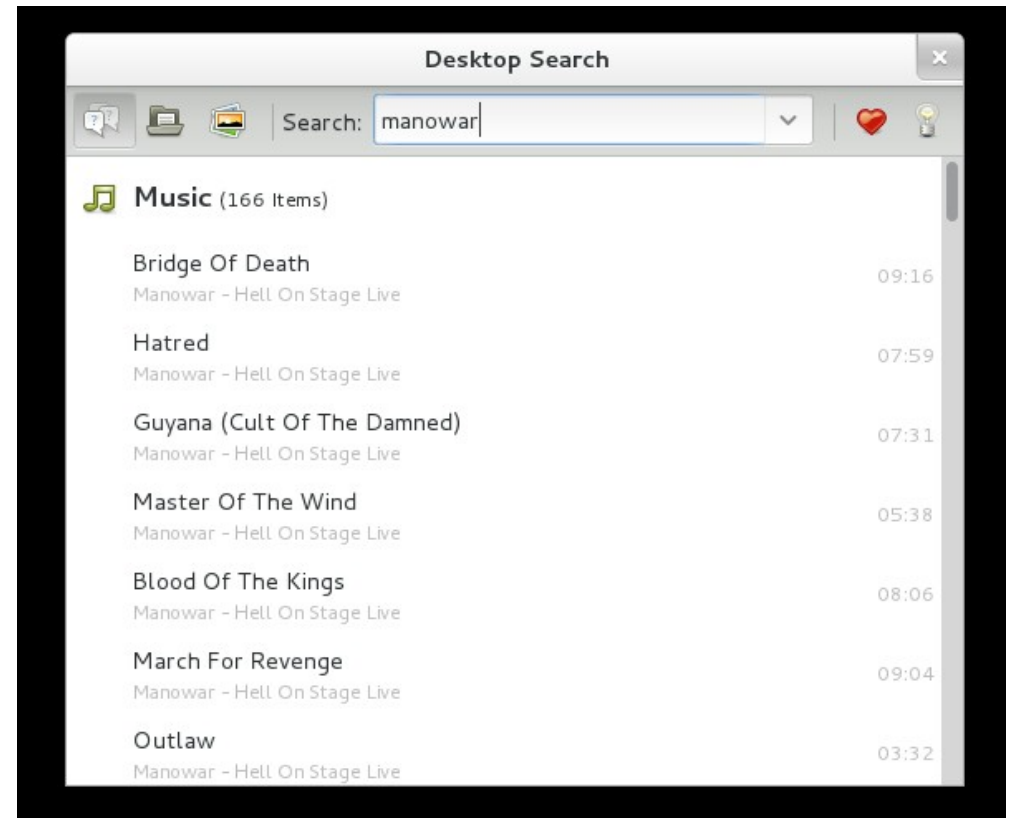
```
file:///home/aleksander/Documents/Repository/articles/filesystem-monitoring/filesystem-monitoring.tex
```

...

- `$ tracker-search "gnom* -tracker"`

Tracker internals: Tools

- Command line interface:
 - tracker-control
 - tracker-info
 - tracker-search
 - tracker-sparql
 - tracker-stats
 - tracker-tag
- Graphical user interface:
 - tracker-preferences
 - tracker-needle
 - Tracker plugin for gnome-shell



Thanks!



- **Website:**

<http://projects.gnome.org/tracker>

- **Mailing list:**

[tracker-list \(at\) gnome.org](mailto:tracker-list@gnome.org)

- **Bug reporting:**

<https://bugzilla.gnome.org/browse.cgi?product=tracker>

- **Repository:**

<git://git.gnome.org/tracker>

- **IRC:**

#tracker in GimpNet