Universal File System Extended Attributes Namespace

François Revol
HAIKU® Project
revol@free.fr

File Meta-data

- Attributes (POSIX or alike)
  - Name (yes, that's meta-data!)
  - Type (dir, file...)
  - Owner (uid, gid...), Permissions (rwxrwxrwx)
  - Timestamps (atime, ctime, mtime, crtime...)
- Extended Attributes
  - EA = xattrs
  - Resource forks
  - Named streams

Extended Attributes

- Generic storage method for meta-data
- Name-value pairs attached to files
- **Not** part of the file itself!
  - Does not require knowledge of the file format
  - But could be extracted from file content (ID3, EXIF)
- Semantics are OS or application defined
- Low-level (file-system) (predates XML & DC…)
- Operating system / file-system specific

Windows – NTFS

- Named streams
  - Alternative streams for files
  - Accessed by path: “foo.txt:somestream”
- Also supports Extended Attributes
  - Name, value pairs
  - Inherited from OS/2
- Usage patterns
  - Not much (and WinFS disappeared)

Linux – ext3/4

- Names (UTF-8?), value (binary / string) pairs
- Atomic access
- Namespace: restricted to “user.*” for applications
- Ext3/4:
  - Namespace prefix stored as 8bit integer
    - EXT4_XATTR_INDEX_USER=1 /* “user.*” */, …
      EXT4_XATTR_INDEX_SECURITY=6 /* “security.*” */
  - 1 block max storage per inode
- XFS, reiserfs: no practical storage limitation
- GNU libattr userland API

Linux – ext3/4

- Usage patterns
  - Beagle and other metadata indexing tools
    - Didn't take off yet
  - FreeDesktop.org-specified
    - “user.mime_type”
    - “user.xdg.origin.url”
    - “user.dublincore.title” & other DC properties…
  - Apache mod_mime_xattr sends type & charset
  - Nepomuk?
  - Slowly growing

Solaris

- Name, value pairs
- Stored as regular files in a tree
  - man fsattr(5)
- Accessible as file descriptors
  - `openat(fd, name, O_XATTR)`
  - `attropen(filename, name, oflag)`
- Custom support in `tar`/`cpio` or `star`

Mac OS X – HFS+

- Historical HFS: Resource fork (binary blob)
- HFS+ supports xattrs
  - Name, value (bin or string) pairs
- Namespace
  - Reverse DNS naming by convention
- Usage patterns
  - “com.apple.ResourceFork” maps HFS metadata
  - “com.apple.metadata:kMDItemWhereFroms” (urls)
  - “com.apple.quarantine” (Safari downloads) …

BeOS & Haiku – BFS

- Name (UTF-8), type (uint32), value (bin) tuplets
  - Type field adds semantics (int32, float, string…)
  - MIME database describes them more
  - Names can be indexed by the filesystem
- fs_attr.h syscalls
- High-level API (C++)
  - BNode::ReadAttr(), BNode::WriteAttr() …
- Live Queries
  - Notify applications of new matching files

BeOS & Haiku – BFS

- Usage patterns (Pervasive)
  - “BEOS:TYPE” (MIME type)
  - “BEOS:APP_SIG” (Application signature)
  - “BEOS:ICON” (HVIF binary icon)
  - “META:url” (Internet shortcut address)
  - “META:{name,email,phone, …}” (Contact infos)
  - “MAIL:{from, to, subject, …}”
  - “Music:{Artist, Album, Track, …}”
  - …

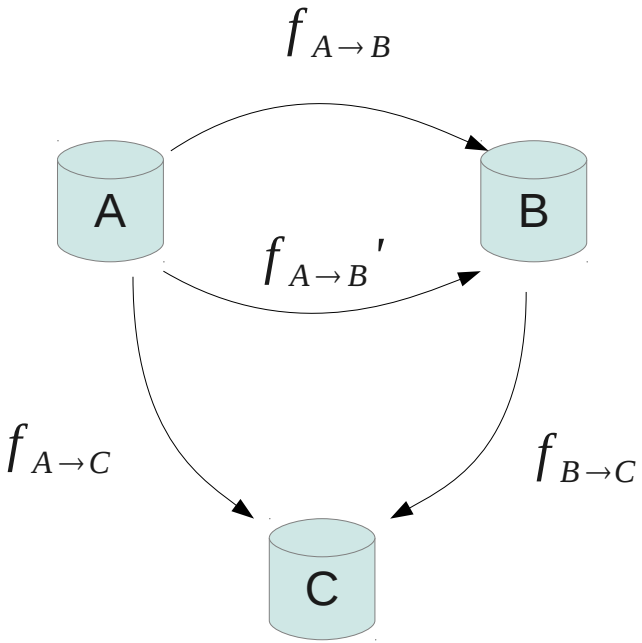
Problems using xattrs

- Often not considered in file transfers
- No support on some filesystems (FAT…)
  - Backing store schemes are also incompatible
- When a mapping exists it is
  - Unilaterally defined
  - Inconsistent
  - Not resilient to composition
- File preservation is thus incomplete
  - Backup, Archival, Digital preservation …

No standardized mapping

- Foreign to native mapping is vendor-specific
- Propositions only consider their OS or fs
- Sometimes several mapping exist
  - NFTS-3g and Samba do not agree
- Mapping composition is not idempotent

Mapping functions



$$f_{A\to B} \neq f_{A\to B}'$$

$$f_{A\to B} \circ f_{B\to C} \neq f_{A\to C}$$

Sample case

- In Haiku
  - People file on BFS
  - Copied to NTFS
- In Windows
  - Copied to a Samba share
- On Linux
  - Samba copies to ext3
- In Haiku
  - Copied from ext3 → **Unusable**

```
"BEOS:TYPE"    'MIMS' "application/x-person"
"META:email"   'CSTR' "revol@free.fr"
"IM:status"    'CSTR' "Offline"
"_trk/pinfo_le"'RAWT' 00 BA E3 EC A7 09...
```

```
"haiku.BEOS_TYPE_MIMS" "application/x-person"
"haiku.META_email_CSTR" "revol@free.fr"
"haiku.IM_status_CSTR" "Offline"
"haiku._trk_pinfo_le_RAWT" 00 BA E3 EC A7 09...
```

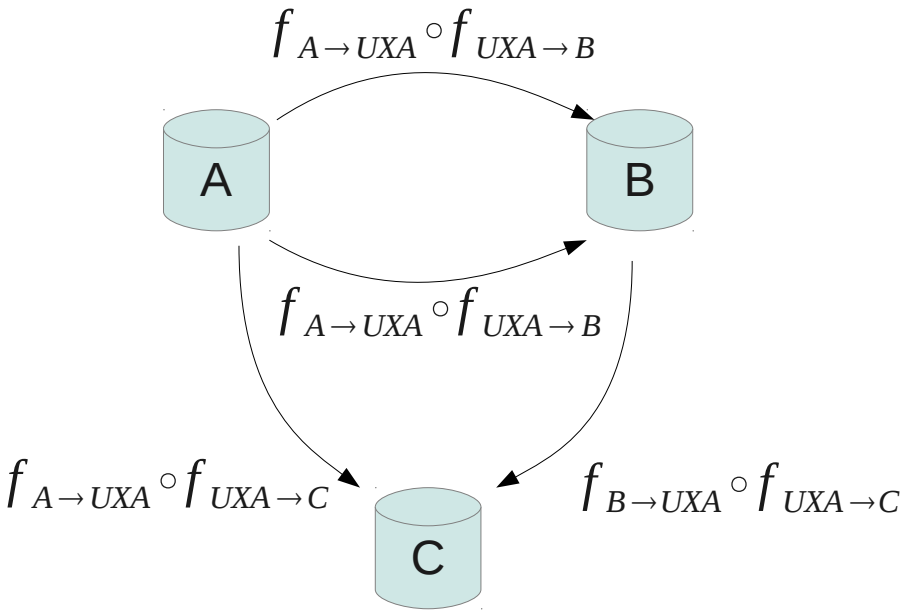
```
"user.DosStreams"
05 00 00 00 00 00 00...  '................'
00...-42 45 4f 53 5f...  '........BEOS_TYP'
45 00 53 4d 49 4d 61...  'E.SMIMapplicatio'...
```

```
"linux.user.DosStreams"
05 00 00 00 00 00 00...  '................'
00...-42 45 4f 53 5f...  '........BEOS_TYP'
45 00 53 4d 49 4d 61...  'E.SMIMapplicatio'...
```

Proposition: UXA

- Unified xattr namespace
- Each vendor defines its UXA mapping
- OS translates to UXA from foreign fs
- OS presents the UXA namespace in their own
- Separate Transport & Presentation layers
  - Transport layer only cares about preservation
  - Higher-level software could perform more complex remapping and add semantics

Mapping functions

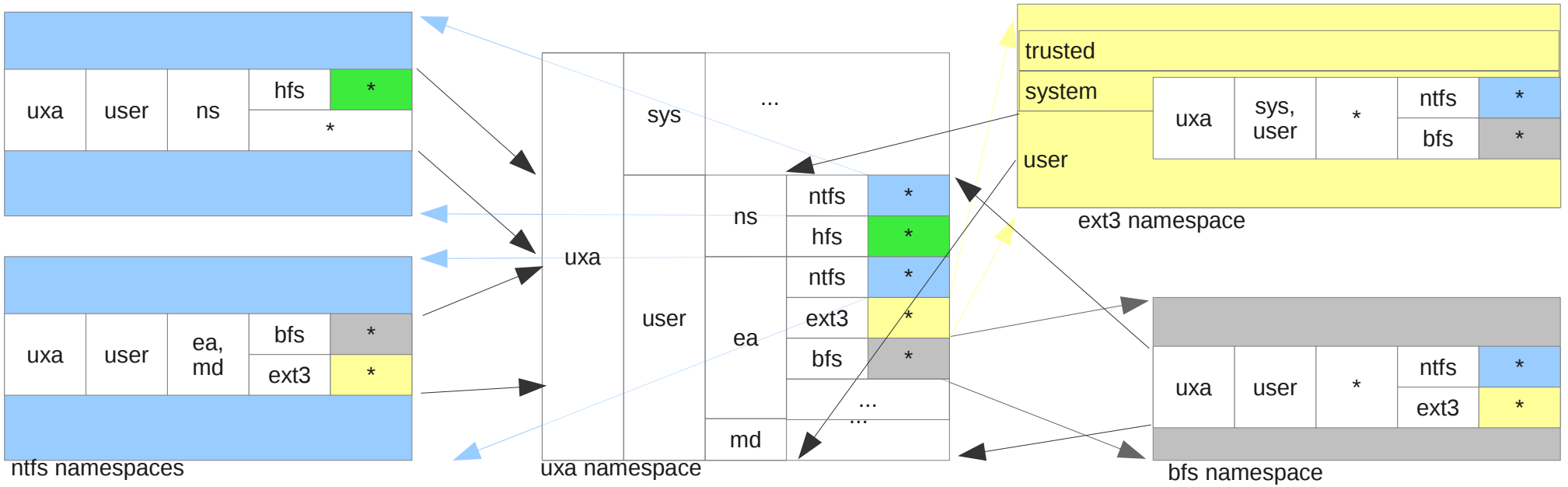


$$\forall A, B, C:$$
$$f = f_{A\to B}\circ f_{B\to C}$$
$$f = f_{A\to UXA}\circ f_{UXA\to B}\circ f_{B\to UXA}\circ f_{UXA\to C}$$
$$f = f_{A\to UXA}\circ f_{UXA\to C}$$
$$f_{A\to B}\circ f_{B\to C} = f_{A\to C}$$

UXA namespace

- Root level: “uxa” defines the root placeholder
- Access level: “user” editable vs. ”sys”tem
- Subtype level
  - “ea”: Extended Attribute
  - “ns”: Named Stream
  - “md”: (other) MetaData
- Vendor level
  - Defines the vendor namespace the EA belongs to

UXA namespace



Higher-level possibilities

- Modified libattr
  - Translates known attributes to native ones
    - uxa.user.ea.bfs.BEOS:TYPE → user.mime_type
- Samba filtering module
- Synchronization applications
- Migration assistants
- DC mapping
- RDF/XML/…-defined mappings

Shortcomings

- Limited storage space
  - Best effort
  - Backup servers should account for it
- No backing store
  - Best effort
  - Could be used as a canonical format in agreed-upon backing store file (or existing ones)
- ACLs to handle with care (might break security)
- Synchronization issues on converted data

So what now?

- Write an UXA RFC
- Forward proposal to interested parties
- Write mapping RFCs and register at IANA
- Fix existing software
  - Samba, NTFS-3g …
  - rsync, tar, cpio, zip, GNU coreutils …
    - Though most userland tools use libattr, so just fix libattr

Questions?

Extended attributes
Can't talk to each others
Covered by snow