

Cam's Unix Notes, 1 — URI and address syntaxes

Consider a web address, such as '<http://muralab.org>'. This 'string' (literally, the characters between the quotes) is known as a *uniform resource locator* (URL), which is part of a broader syntactic structure that is common in computer networking and has become popularized because of the world-wide web (the `http://` addresses you used to see everywhere). Try the following in your web-browser on one of our struc* workstation (type **it** into the location bar of Firefox, Chrome, or whatever internet browser software you prefer):

file:///u1

NB: This really is 'file' instead of the usual 'http' prefix. You may have run across 'ftp://' or 'smb://' prefixes in the past... In principle, this 'file' address is no different from any of those – These strings are all part of a general set of URI *schemes*; see the "Official IANA-registered schemes" table at http://en.wikipedia.org/wiki/URI_scheme for an exhaustive list of types of prefixes.

What do you see when you type the above 'file'?? Navigate and play around a bit... In doing so, you see that your web browser is now basically being used as a fancy graphical *file system browser*. (Most modern web browsers were intentionally designed this way, so that they can *handle* [that's the technical term] many types of network protocols, including http, ftp, samba/cifs, tel [for telephone numbers], dav [for WebDAV], and so on.)

Some starting points for further information on these concepts:

- http://en.wikipedia.org/wiki/URI_scheme
- http://en.wikipedia.org/wiki/Uniform_resource_locator
- http://en.wikipedia.org/wiki/File_URI_scheme (for filesystems)

So, what does all this mean in terms of your research and data exchange / communication in our lab??

Consider the following sample *uniform resource identifier* (URI; technically, URLs are a subset of URIs):

cmura@struc3:/u1/cmura/some/path/to/a/dir/.

This URI is of a form that you'll probably see me use over and over again in email communications with you, so please learn it (also, it's standard in the community). This URI is composed of just a few distinct elements:

- a **username** (aka 'login name'; may be the same or different on different machines; almost always the same within a single network, like our StrucBio)
- a **hostname** = '*the machine*' = **struc3**(.chem.virginia.edu, if beyond network, outside world)
- **/an/(absolute)/path/to/local/resource** – Two important pieces of terminology:
 - '*local*' means local to *the machine*, versus a remote server or some other machine (*e.g.*, a hard-disk stored at Dropbox's facilities, or your computer at home).
 - '*absolute path*' Homework question — *What's the difference between an "absolute path" versus a "relative path" in a Unix filesystem?* This is a straightforward concept (not as fancy as it sounds), and one that is vital to understand; make sure to look-up this concept and familiarize yourself with it, asking questions when you have any.
 - *What is meant by the vague-sounding 'resource'?* In this lingo, a '*resource*' is either a *directory* (usually called a '*folder*' in Windows) or simply a *regular file* – described as '*regular*' because everything in a Unix filesystem is, technically, a type of file. (A directory is just a special type of file – namely, one that *contains* other files... So a directory is just a special file that is a file container... recursion on this theme leads to the concept of ((sub)-sub-)directories [often referred to as 'subdirs'].)