

Question

[Michael Gosselin](#) · May 22, 2018

Faster finds in large Unix directories

We have this challenge at our site. When we first designed it many years ago, we decided that the best way to store files was with a unique identifier, which matched one of the fields in the corresponding record. For example, if the unique identifier was a nine-digit field (such as a SSN), we'd save a file as nnnnnnnnn.ext, where nnnnnnnnn is the nine-digit number, and ext is the file extension. If we needed a change to the file, we'd file as nnnnnnnnn_hdate.ext, where hdate is the horolog date. And for 18 years, this was just fine.

However, in year 19, we discovered one of the growing pains of success. We had always searched files in the directory with a call from \$ZF call from Cache. To find all the relevant files, we write something like `S X=$ZF(1S -1 /filepath/nnnnnnnnn* >output.txt"`) to list all the files in that directory into a single file we'd read and process. However, at some point, the searches started taking longer, and we'd get CSPs timing out after 60 seconds, and complaints from the customer service people that our searches were too slow.

To get around this, short-term, we moved every file from /filepath to /filepath_old, deleted the directory /filepath, then renamed /filepath_old to /filepath. This worked for about three months, then the searches started getting slower... again.

One idea we did have was to create sub-directories within the directory to sort the files. This would work for some cases, but then gets problematic when the search is looking for a date range. Not only that, we'd need to have an effective way to pass the unique identifier, and create the proper directory name for lookup to pass to the \$ZF call, or even to any %File calls.

So, I am asking for help and guidance on faster lookups, and better directory structures, for Cache with our Unix system. If you've had a similar problem, or can shed some light onto this, I would be most appreciative.

Thank you.

[#Caché](#) [#HP-UX](#)

Source URL: <https://community.intersystems.com/post/faster-finds-large-unix-directories>