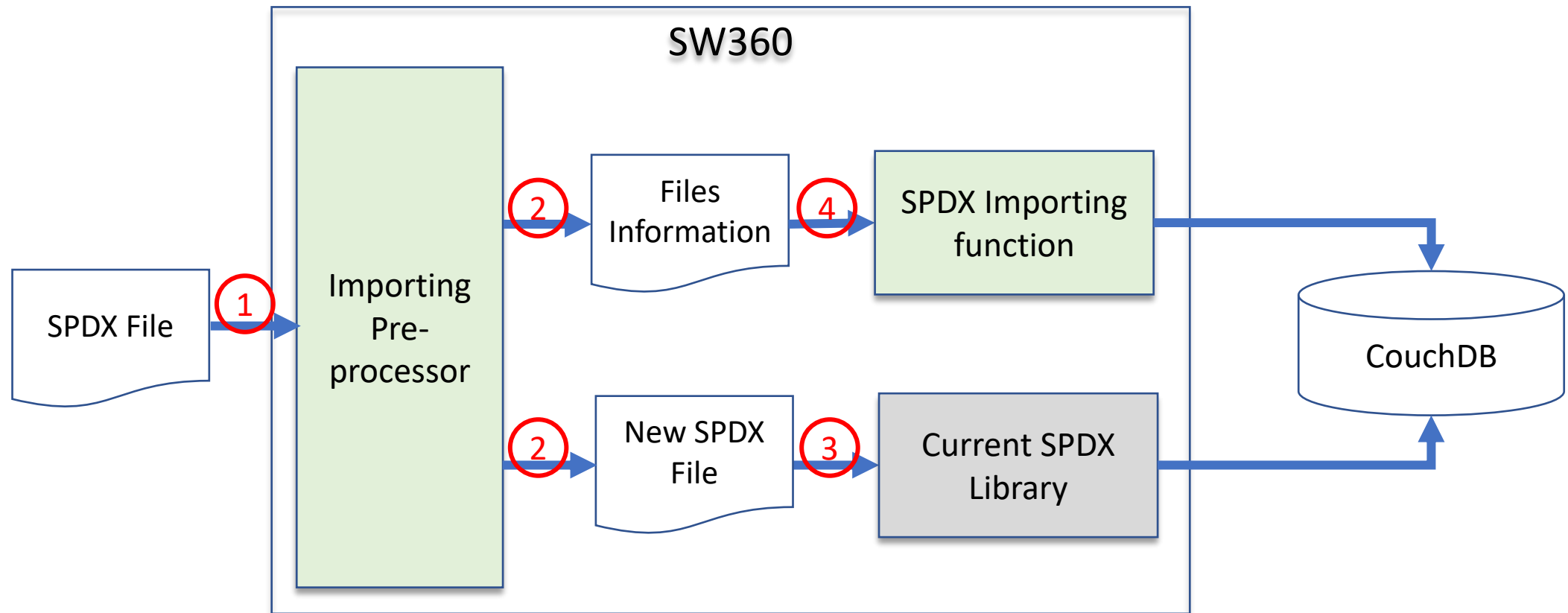


Importing



(1): When an SPDX file is imported into SW360, it will be handled by the Importing Pre-processor.

(2): Importing Pre-processor will split SPDX file into 2 smaller files:

- Files Information: contain File Information data

- New SPDX File: the content of this new SPDX file is the same as the original SPDX file, except one point that there is no File Information data. Therefore, the size of this file will be much smaller than the original one.

(3): Because the size of new SPDX file is very small now, it will be handled properly by the current SPDX library

(4): In the future, when we need to process Files Information, we will write a new function to handle the Files Information file. The size of this file will be large, but the structure is simple: just a small pattern is repeated multiple time. Low-level technique, like Shell script will be used. We have tested using command 'grep' to extract data, the speed is very fast. There is case that people use grep to process 80GB file: <https://stackoverflow.com/questions/13913014/grepping-a-huge-file-80gb-any-way-to-speed-it-up>