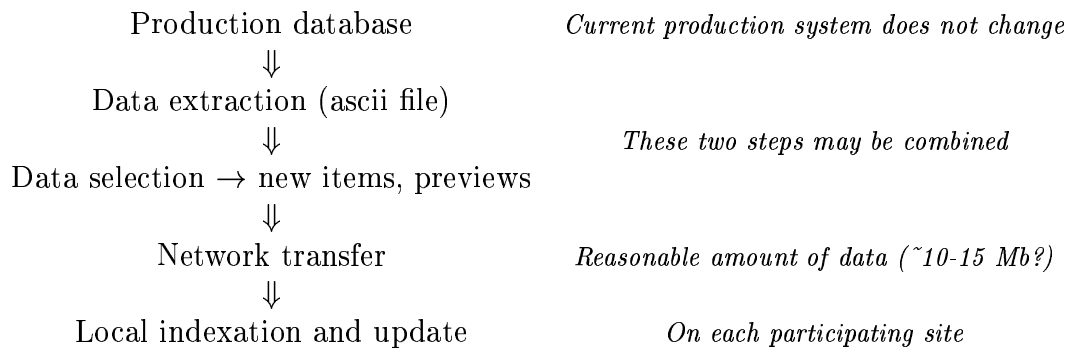


Updating servers

Cellule MathDoc

May 3, 2000

The proposed updating process (which will apply to the servers of the international access net, as well to national or regional access nodes,...) is given by the following flow-chart :



Notes

1. We assume that the current production system remains unchanged.
2. This scheme applies to the loading of new data (newly reviewed items and preview items). They are identified by their Zentralblatt-MATH volume number. The updating of existing data may be done in a similar way, provided the extraction step is able to identify modified items. This can be done by a simple comparison of data. (the current production system apparently does not allow easy identification of modified data).
3. Network transfers should not be too big (update frequency should be high enough). Some measurements done on recent updates (second half of 1999 and january-march 2000) show that the amount of data to be transfered can be kept to well under 10 Mb on a monthly basis. By transferring only new items, modified items, and a list of deleted items (previews), the typical size of update data is less than 8 Mb (gzip compressed) and even less than 5 Mb (bzip2 compressed). This can be reduced even further if we transmit only modified fields for changed items (less than 4 Mb).
4. The current structure of data files does not allow updating them in place. The local updating process is actually a merging process. (and thus needs more disk space).
5. Software requirements include: the development of a tool for data extraction, selection and sorting, the development of an indexer (compatible with existing indexing rules) and merging/updating procedures.

[*Draft 0.2*]