

Gina Steiner und Rainer Weigle

Back it up, Scotty!

Comparing extensions backup and synchronisation

No one wants to lose data with TYPO3, but mistakes do happen. If you have set up a backup in time you should feel much more secure. But, as anyone affected will soon realise, the backup is only as good as the restoration. So which systems offer resilient backup or synchronisation options, and which details must be taken into account?

Everyone who regularly works with computers should take an interest in the subject of backup. User, software and hardware errors are all too common. While high quality hardware, RAID systems and an uninterrupted power source can improve the reliability of hardware, they do not prevent software or user errors, which are much more often the cause of failures.

It is important to distinguish between the individual domains. A backup secures all files at a fixed time (usually at night). Ideally securing into a consistent state takes place in an external medium or an external server. In a synchronisation two TYPO3 systems are regularly aligned with one another. It should encompass the database and the files. However, version control is often not possible. If a failure is not detected by the automatic synchronisation, it will be silently adopted by the second entity.

Thus a synchronisation does not replace a backup in which several versions from the last days or weeks were stored. Normally nothing is archived with backups: to keep the amount of space needed small the backup is cyclically overwritten after days or weeks.

With all methods it is important to test whether a complete restoration can be easily and reliably effected, for, as already mentioned, a backup is only as good as its restoration. A reliable backup can be easily achieved with the help of a shell program. If you don't have shell access, you will either have to resort to extensions for backup and synchronisation, or leave the backup to the host.

Search criteria and uncovered external extensions

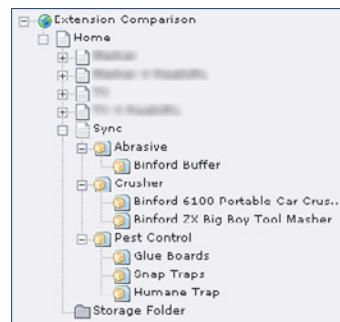
A search for extensions in the TER with the terms „backup“, „import“, „export“ and „sync“ resulted in more than a dozen hits. According to their authors, the extensions „ed_lpesync“ and „df_t3sync“ are out of date. According to its author, „www_t3sync“ is too complex in the absence of documentation. The extensions „backup“, „w4_backup“, „zor_backup“ and „agprjmgm“ would not run properly on the Debian Sarge test system: often they displayed the same problems. As an alternative, „w4x_backup“ was tested without success on four different systems, including two TYPO3 standard installations at hosts.

„treesync_base“ and „treesync“ were amalgamated, as were „dbsync“ and „doc_dbsync_fr“. Thus the three extensions „ml_dbsync“, „treesync_base“ and „dbsync“ were left over for testing. The tests were conducted with PHP Version 4.3.10-20, MySQL Version 4.0.24 and Apache Version 2.0.54. A fundamental problem with TYPO3/PHP-based backup solutions are the restrictive PHP settings of many host packages. Thus, for example, „upload_max_filesize“ and „max_execution_time“ are often not sufficient to fully re-integrate a larger backup.

Database Sync (ml_dbsync)

With the aid of this extension it is possible to create, in the TYPO3 DB, TYPO3 pages and page contents from the contents of external

databases which are foreign to TYPO3. In the backend this extension becomes visible in Modul Tools > DB Sync.



The pages created from an external database can be recognised by a special icon.

Pages that are created out of the external database can be immediately recognised by a special page icon. If, in addition, new content elements were introduced via TYPO3, these are not overwritten or deleted by a new synchronisation. If, however, you change the content elements created by „ml_dbsync“ with TYPO3, they will be overwritten with the external DB values with a new synchronisation. The configuration for the interpretation of the external database content occurs with XML, which, though it might take some getting used to, is very flexible.

The synchronisation can only go from the external database to the TYPO3 database and not vice versa, and no files are aligned. Thus, while a backup cannot be made possible in this way, the possibilities for database synchronisation are very powerful and flexible. The importation of files from external databases is only possible when these allow access via the network and the database user is approved for external hosts. This should be taken into account when considering safety matters.

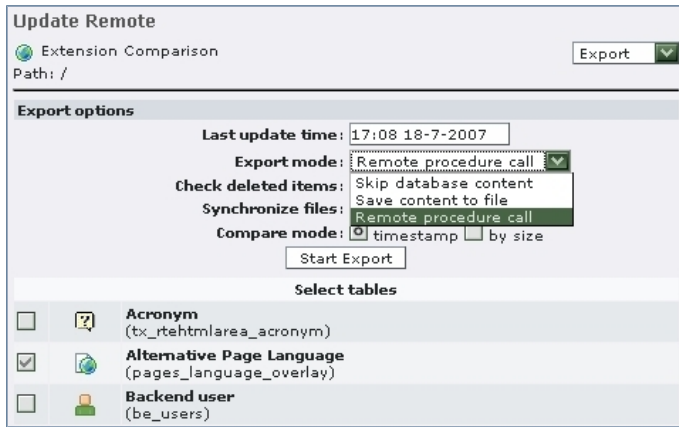
The synchronisation can only go from the external database to the TYPO3 database and not vice versa, and no files are aligned. Thus, while a backup cannot be made possible in this way, the possibilities for database synchronisation are very powerful and flexible. The importation of files from external databases is only possible when these allow access via the network and the database user is approved for external hosts. This should be taken into account when considering safety matters.

Automatization is possible with the help of the extension „gabriel“ [1]. „Gabriel“ is selected at defined intervals by Cron-job. Instructions for the configuration of the extension can be found in the TER under „ml_tut_gabriel“ [2].

The not yet released Version 0.10.0 was available for the test, recently it has become available in the TER. „ml_dbsync“ gives you the sensible option of deactivating all users during synchronisation. No documentation was available, but the T3N article „10 000 Pages in 15 Minutes“ (in volume 6) as well the author himself can offer help. On the whole the solution, especially for the importation of TYPO3-foreign database contents, proved itself powerful and flexible.

Remote synchronisation toolkit and synchronisation control module (treesync_base + treesync)

These extensions make possible the synchronisation of two identical TYPO3 entities, for example a development system and a live system. For this the extension „treesync_base“ is necessary on both systems and, additionally, „treesync“ has to be installed on the development system. Not only is an FTP server necessary in the live system, but both systems must be present right from the start as exact copies.



'treesync_base' and 'treesync' offer comfortable setting options for the alignment of the systems.

Nothing now stands in the way of synchronisation in both directions. This is effected in the backend module Tools > Rsync by a drop-down following the choice of export or import. The tables for synchronisation can be selected with a click, and this allows a partial synchronisation too. So that all files may be adjusted, all subordinate folders can be individually selected in „uploads“ and „fileadmin“. The selection of „skip database content“ starts the exclusive adjustment of the files, the selection of „save content to file“ from the drop down starts the backup of contents as a file. When files are named the date and time are used so that a backup can be produced with these extensions. When synchronisation is effective all database operations are logged in the backend. The documentation explains which settings the extension requires, although a little more information about the FTP server and rights would have been helpful. The encryption of the FTP traffic with TLS is supported.

In short: (partial) synchronisation in both directions including files in „uploads“ and „fileadmin“ is a success.

DB Sync (dbsync + doc_dbsync_fr)

These extensions facilitate the importation not only of complete databases, but also of individual tables. An importation within the same database, for example an importation of „fe_users“ into „be_users“, is possible too. This extension does not offer an alignment of files, the alignment of files is only provided for in one direction.

Following its installation, which is dependent on „api_made“, the extension can be reached through the backend module Tools > DB Sync. The tables of other databases may now be selected for importation, or internal table-entries may be imported. In doing so you can decide between the options „Add records to local tables“ and „Empty local tables before synchronisation“. When importing internal table-entries, not only may these be restricted to particular records, it is also possible to choose the page to which the records are assigned.

In the case of imports of „fe_users“ into „be_users“ the source and target fields must be assigned to one another. The extension offers sensible default settings, but the fields may also be freely assigned via the drop-down. A clamp in „Do not include deleted records“ excludes deleted records. The concise documentation is sufficient and is also available in French through „doc_dbsync_fr“.

With these extensions, too, importation from an external database is only possible when the database allows access through the network and the database user has been released for external hosts. In summary: no alignment of files, but on the one hand a quick database (complete) importation, and on the other

hand the possibility of the alignment of two different internal tables – a well thought-out system.

Solutions without external extensions

If you don't have quick access, you can use a system extension or the host's backups. If you do have quick access, these diverse external tools can be used for your own flexible backup.

TYPO3's internal means

With the system extension „impexp“ TYPO3 already brings with it the option of a backup. The selection of „Import from .t3d“ or „Export to t3d“ from the click menu under „More options...“ can effect importation or exportation. Although extensions cannot be aligned here either, there is the possibility of noting which extensions need to be present in the target system. When importing, a notice to that effect will be displayed.

It is possible to export (partial) trees and to include files in the export. Each individual content element and every link can be flexibly excluded from the export, and relations can be automatically dissolved and exported as well or be exported as static references. Once the update button has been engaged, red notices allow you to recognise whether and where the settings lead to errors.

Frequently used export settings can be saved or loaded via an individual save or load button as default settings. These default settings can be tied to persons and, in that case, cannot be used by any other backend user, even with respect to an Admin matter. So that export files can be distinguished from one another, the indication of diverse meta-information is possible. The files can either be saved in „fileadmin“ or on the local disc and, when no individual specifications are made, obtain from the system a file name in the form of „T3D_19-07-10-27-09-z.t3d“. Likewise the files can be imported from fileadmin or from the local disc. However, the fact that the import button becomes the update button following importation according to the optional definition of the settings and the activation of the preview button can lead to confusion. In short: (partial) export and import, powerful setting options, inclusion of all necessary files (but not of the extensions), backup possible by saving in a file – hardly any wishes remain unfulfilled.

Hosts

Where the TYPO3 package was rented from a TYPO3 host, the backup is often included in the package price. Some hosts have told us about their backup strategies. All result in a daily dumping of the database into the file system and all carry out a file system backup.

elios.de	4 days' full external backup, longer by request and redundant with an external data centre.
jweiland.net	7 days' full backup locally, the backup script can be extended by the user. An additional 3 days' full backup externally, but here the restoration is charged for.
mittwald.de	10 days' full backup externally, the restoration is charged for. On request all the files in the client web space can be secured.
punkt.de	14 days' incremental backup externally, longer or on a different system by request.

External tools

If access to the shell of a web server is possible it is possible to resort to classical backup methods. As a rule a dumping of the (MySQL) database should be filed in the file system and, subsequently, a backup of the file system should be carried out. Comprehensive information about this may be found, for example, in O'Reilly's recently published book Backup & Recovery.



So that the dumping of the database can be as consistent as possible, the tables should not be recordable during the dump („--lock-all-tables“ for MyISAM tables, or „--single-transaction“ for InnoDB tables). This can, however, lead to short-term timeouts during recordable database access by TYPO3 entities.

As an alternative to dumping, the MySQL database files can be saved directly. Here too it must be ensured that no recordable access follows and that all database indices are recorded on the hard drive. For MyISAM tables the perlscript „mysqlhotcopy“ automates all the necessary steps [3].

The backup of the files can be achieved either with classical backup software (for example „amanda“ or „bacula“ in the open source domain) or with a simple disc to disc backup with „rsync“ on a second disc or backup server. „dirvish“ [4] is a very helpful and easy to configure tool for rsync-based backups inclusive of version control. For its version control this tool uses hard links, with the result that unmodified files are not saved more than once, although a complete representation of the file tree is present at all times during the backup. The restoration is easily accomplished by „rsync“ or „cp“.

If an external FTP backup domain is provided for in the hosting package, the c't tool „ftplicity“ [5] can comfortably effect backups, which are divided among handy files and saved in encrypted form in the FTP server.

All steps in the backup are automatically taken by Cron-jobs and should be regularly checked for completeness and functionality (for example through a restoration in a test system). Database dumps can also be carried out with php-MyAdmin (as integrated in TYPO3 for example), but here the PHP limitations mentioned above should be taken into account.

Destination table	Source table
 Website user (fe_users)	 Backend user (be_users)
Username: Text input	Username: (Text input) <input type="text"/>
Password: Text input	Password: (Text input) <input type="password"/>
Groups: Selector box	Group: (Selector box) <input type="text"/>
Lock to domain: Text input	Lock to domain: (Text input) <input type="text"/>
Name: Text input	Name: (Text input) <input type="text"/>
Address: Text area	None <input type="text"/>

With importation with DB Sync the fields are allocated manually to the source and target table-fields.

Conclusion

Unfortunately none of the extensions offer a complete backup for the TYPO3 installation (including extensions and sources). Similarly, no extension comes with automatization. The external extensions are more likely to cover special cases. The system extension can be helpful for one-off backups or synchronisations, especially when partial trees are to be exported. Thus, for a reliable, complete and automated backups, either hosts or individual shell scripts are required. If archiving data is important for you, you will not be able to get around an individual shell solution. In future, the integrated automatization (perhaps even without Cron-job) and the sending of an error notification via email might be considered with respect to extensions.

Call up

The extension comparison team [6] has set itself the task of making the selections of extensions easier and of supporting users in the search for the extension which suits them best. The call up in the last T3N magazine brought in more team members, but we would like to find some more people to actively help us with our work so

that, in the long term, we will be able to provide systematic extension comparisons, not only in the magazine but elsewhere too. We would love to hear feedback, suggestions and criticisms with respect to our previous articles. Please reply by email to: extcomp@googlemail.com.

Links and Literature

 [Softlink 1864](#)

- [1] gabriel: <http://typo3.org/documentation/document-library/extension-manuals/gabriel/1.0.0/view/>
- [2] ml_tut_gabriel: http://typo3.org/documentation/document-library/extension-manuals/ml_tut_gabriel/
- [3] mysqlhotcopy: <http://dev.mysql.com/doc/refman/5.0/en/disaster-prevention.html>
- [4] dirvish: <http://www.dirvish.org/>
- [5] ftplicity: <http://www.heise.de/ct/06/13/links/216.shtml>
- [6] Extension Comparison Team: http://wiki.typo3.org/index.php/Extension_Comparison

THE AUTHOR



Gina Steiner lives in Hamburg, where she studied meteorology and was a technical webmaster at the MPI for Meteorology from 2000 to 2004. Since 2003 she has been working with TYPO3, since 2004 she has been working with it almost exclusively as a freelancer. She is an active participant in the TYPO3 user group HH, a member of the TYPO3 Association, and is organising the TYPO3 Developer Days in Hamburg/Elmshorn.

THE AUTHOR



Rainer Weigle lives in Hamburg, where he studied oceanography. Since 2000 he has been working at the Max Planck Institute for Meteorology, most recently as head of the IT department. This department is not only responsible for Linux clusters, networks, servers and 500 desktops, it also runs approximately ten TYPO3 entities for various departments of Hamburg University.