3. Data Backup

Regular and redundant data backup is a vital necessity of every digital information system. When defining a backup solution for PECE, we followed the general guidelines of the Drupal community and the practical policies of RDA. In a nutshell, the overall goal of our backup policy to ensure PECE instances have, at all times, three backup copies in distinct machines.

The first backup level is the PECE backup, which is performed automatically on a regular basis by the Drupal framework. **The second level** is, generally, performed by the hosting company – which must provide regular, automated backups on the system level, that is, generating regular snapshots of a virtual machine where PECE is running. This is beyond the reach of automation of our platform and has to be set-up with the hosting company directly. We describe the tech requirements for hosting companies in the section on "Storage Costs" of this document. **The third and last but not least important form of redundant backup** is to generate an offline copy in a safe environment (in addition to the other two forms of automated online backup).

Regular backups are generated through the extension "Backup and Migrate" which performs full backups of the database and the PECE directory structure on the file system. The generated tarball files are useful for quickly restoring of the system in case of data or system failure. The backup functionally provides full Integration with `drush` (Drupal Shell) for facilitating the administrative tasks of more experienced system admins as well as a GUI for new PECE administrators who are not used to command-line interfaces. For users of the PECE VM distribution, we provide both options out-of-the-box.

Please observe that, given the key importance and sensitivity of this data management task, only administrators (users with the "administrator role" on the system) are allowed by default to generate and access backup files and system configurations. Administrative backup functions include:

- Database backup;
- File system backup;
- AES 256 encryption of backup files;
- Export and import previously generated backup files;
- Setup backup schedules (to run on cron jobs);
- Setup `$PATH` for backup files;
- Use `sftp` to send backup files to other machines.

Backups are generated with timestamp, AES encryption (given the sensitivity of the data they include in `.tar.gz` files) and then replicated to a different machine. Two options, thus, are offered to PECE administrators: to either use the GUI or the command-line interface (both offering automated backup solutions). Command-line tools facilitate the process of automation.

```
# Perform a new backup using PECE's backup profile
$ drush bam-backup pece_bkp
```

```
# Lists all the backups already generated (outputs backup ID numbers)
$ drush bam-backups

# Restore a particular backup, using its ID number
$ drush bam-restore $BACKUP_ID
```

These commands are based on `drush` to generate, list, and restore backups. Shell scripts can be additionally used – added as `cron` job – to 1) put the server in maintenance mode for backup purposes; 2) dump the contents of the database to a file; 3) generate a tarball of the Drupal directory structure; 4) assemble the DB dump and the tarball into another .tar.gz file; 5) use AES 256 to encrypt the package file; and 6) finally, upload the encrypted file to a different server via `sftp` (or, alternatively, synced with `rsync`). The option of scheduling and running a shell script for automated backup will shipped with the PECE distribution, thus offering an alternative for experienced system administrators running off of the PECE VM distribution or their own server infrastructure.

Instructions for Developers: Install modules **"backup_migrate"** v. 7-x-3.1, **"backup_migrate_sftp"** v. 7.x-1.0 and **"aes"** v. 7.x-2.0 (these modules depends on **php5-mcrypt** and **libssh2-php**, if you are using the PECE VM distribution, this is already configured out-of-the-box). The following configuration files for "Backup and Migrate" extension module are provided in the appendix section of this document, which can be imported to Drupal: `settings_profile.txt`; `schedule_profile.txt`;