

Séneca FX Server - Sistema de Archivos

El directorio de instalación : */opt/spica/seneca*

Séneca FX se instala en la carpeta */opt/spica/senecafx-DISTRO* siendo *DISTRO* el nombre de una distribución en particular, y que normalmente sera *SERVER*. En cualquier caso, durante la instalación se crea el enlace simbólico ***/opt/spica/seneca*** apuntando a la distribución que se haya instalado.

```
$ ls -l /opt/spica
lrwxrwxrwx 1 root root seneca -> ./senecafx-SERVER
drwxrwxr-x 15 seneca seneca senecafx-SERVER
```

Es aconsejable utilizar siempre este enlace para trabajar con Séneca FX. Así, para situarnos en el directorio de instalación de Séneca FX haremos lo siguiente,

```
$ cd /opt/spica/seneca
$ pwd -P

/opt/spica/senecafx-SERVER
```

También es muy recomendable al trabajar Séneca FX iniciar sesión con el usuario ***seneca*** (en lugar de *root*) para no alterar los permisos de algún archivo o directorio. Además, el script de inicio de este usuario ya nos sitúa en el directorio correcto,

```
$ su seneca
$ pwd

/opt/spica/seneca
```

Contenido del directorio */opt/spica/seneca*

El contenido de la carpeta */opt/spica/seneca* es el siguiente:

```
/opt
+-spica
+-seneca
+-- conf
+-- scripts
+-- seneca
+-- anneo
+-- lucio
```

```
+-- cli
+-- rakelib
+-- ruby
+-- rgloader
+-- bundle
+-- nginx
+-- redis
+-- solr
+-- pids -> /var/seneca/pids
+-- logs -> /var/seneca/logs
+-- temp -> /var/seneca/temp
+-- data -> /srv/data
```

donde,

- La carpeta "*conf*" contiene los archivos de configuración de Séneca FX.
- La carpeta "*scripts*" contiene una serie de scripts y utilidades Bash que dan soporte al entorno de ejecución del usuario *seneca*, la herramienta CLI, las actualizaciones online, etc. Esta carpeta se añade durante la instalación a la variable de entorno PATH.
- La carpeta "*seneca*" contiene la aplicación Web de Séneca FX, instanciada por el middleware de integración HTTP/Ruby (Passenger) cuando se inicia el servidor Nginx.
- La carpeta "*anneo*" contiene la aplicación del servicio de colas de Séneca FX. Es un demonio que se inicia/detiene independientemente del servidor Web.
- La carpeta "*lucio*" contiene la aplicación del 'crawling' para la ingestas. Es un demonio que se inicia/detiene independientemente del servidor Web.
- La carpeta "*cli*" contiene la aplicación del CLI de seneca FX.
- Las carpetas "*ruby*, *rakelib*, *rgloader* y *bundle*" contienen los archivos necesarios para ejecutar el interprete del language Ruby, con una configuración específica y librerías específicas para Séneca FX.
- La carpeta "*nginx*" contiene los binarios del servidor Nginx que hace de front-end HTTP de Séneca FX. También se incluye en esta carpeta el middleware Passenger que permite la integración de Séneca FX con el servidor Nginx.
- La carpeta "*redis*" contiene los binarios del servidor Redis que almacena datos NoSQL (colas y jobs para dar soporte al scheduler de Séneca FX).
- La carpeta "*solr*" contiene los binarios del servidor Solr que se utiliza para la indexación y búsquedas.
- Las carpetas "*temp*, *logs* y *pids*" contienen datos generados por la aplicación(es) tales como archivos temporales, logs, etc. y están enlazadas por defecto al directorio ***/var/seneca***.
- Finalmente, la carpeta "*data*" contiene los datos de usuario tales como archivos de video y audio, archivos de imagen, transcripciones, etc.. Por defecto se enlaza a la carpeta ***/srv/seneca*** y dado que

puede llegar a almacenar un número muy elevado de archivos, y de gran tamaño, suele montarse en un disco adicional o en una carpeta de red (NAS/SAN).

El directorio de archivos temporales, cache y logs : `/var/seneca`

Los archivos temporales, logs, cache del servidor Web, y demás archivos generados por la aplicación durante su funcionamiento pero que no es necesario preservar en caso de que sea necesario restaurar o migrar a otro equipo se encuentran en esta carpeta. Se puede utilizar cualquier otra carpeta -no se recomienda- pero será necesario cambiar los enlaces `./pids`, `./logs`, `./temp` en `opt/spica/seneca` para que apunten a la nueva carpeta.

Contenido del directorio `/var/seneca`

El contenido de la carpeta `/var/seneca` es el siguiente:

```
/var
+-seneca
  +-pids
  +-logs
  +-logrotate
  +-temp
```

donde,

- La carpeta "`pids`" contiene los archivos `.pid` generados por los distintos servicios que forman parte de Séneca FX y se ejecutan en segundo plano.
- La carpeta "`logs`" contiene los archivos `.log` de las distintas aplicaciones y los generados por tareas tales como actualizaciones, backups, etc.
- La carpeta "`logrotate`" contiene los archivos `.log` una vez han sido rotados.
- La carpeta "`temp`" contiene archivos temporales, organizados en varias carpetas según la aplicación que los genera.

Por simplicidad los permisos en esta carpeta, sub-carpeta y archivos que contienen son de lectura y escritura para todos los usuarios.

El directorio de archivos de usuario : `/srv/seneca`

Esta carpeta contiene todos los datos no temporales, los proporcionados por los usuarios - archivos de video y audio, transcripciones, etc.- así como los derivados de estos, y que en conjunto forman el archivo en disco de Séneca FX. Esta carpeta debe incluirse en las copias de seguridad.

Cuando el espacio en disco no sea suficiente, se recomienda montar esta carpeta en un disco físico diferente al de sistema o en una unidad de red, pero mantener el enlace `./data` en `opt/spica/seneca` a este directorio.

El propietario de esta carpeta, sub-carpetas y archivos que contienen debe ser el usuario *seneca* y el grupo *seneca*, con permisos de lectura y escritura únicamente para este usuario.

Contenido del directorio */srv/seneca*

El contenido de la carpeta */srv/seneca* es el siguiente:

```
/srv
+-seneca
  +-redis
  +-solr
  +-tenants
  +-backups
```

donde,

- La carpeta '*redis*' es utilizada por el servidor Redis para mantener una copia - actualizada cada minuto - del estado de las colas y tareas del planificador de tareas de Séneca FX.
- La carpeta '*solr*' es utilizada por el servidor Solr para mantener el índice de texto que se utiliza para las búsquedas.
- La carpeta '*legacy*' contiene archivos necesarios para la migración de versiones anteriores de Séneca.
- La carpeta '***tenants***' contiene el archivo en disco (audio, video, images, transcripciones, etc.) organizados en carpetas separadas para cada instancia o base de datos gestionada por Séneca FX.
- La carpeta '***backups***' contiene archivos de backup tanto de la base de datos SQL como de el índice de texto del Solr. Los backups diarios, que se realizan automáticamente, así como backups previos a las actualizaciones o aquellos realizados a discrección se encuentran en esta carpeta.

De las carpetas anteriores, la carpeta '*tenants*' es la más importante ya que contiene todos los archivos que es necesario preservar - junto con un copia de la base de datos - en caso de que sea necesario restaurar Seneca FX en otro equipo.

El primer nivel dentro de la carpeta *tenants* es un directorio por cada instancia de Séneca FX o '*tenant*', cuyo nombre es el identificador o 'prefix' del 'tenant'. Como ejemplo, si se han configurado dos tenants, *example1* y *example2*, tendremos lo siguiente

```
/srv
+-seneca
  +-tenants
    +-example1
      +-brand
      +-archive
      +-default
      +-images
      +-signatures
```

```
+-transcripts
+-scratch
+-uploads
+-legacy
+-example2
+-brand
+-archive
+-default
+-images
+-signatures
+-transcripts
+-scratch
+-uploads
+-legacy
```

donde,

- La carpeta '*brand*' contiene archivos de logo y hojas de estilo para personalizar las paginas html.
- La carpeta '**archive**' contiene todos los archivos relacionados con los '*items*' del catalogo: masters, proxys, transcripciones, speech-to-text, etc., organizados en diferentes carpetas:
 - '*default*', que contiene masters, proxys, fotos y documentos adjuntos de los items del catálogo
 - '*images*' contiene los posters del oradores, calificadoros, e items del catálogo que son proporcionados por el usuario (son se extraen de los videos).
 - '*transcripts*', que contiene los documentos de transcripción
 - '*signatures*', que contiene firmas digitales
- La carpeta '*scratch*' contiene archivos derivados tales como posters, cortes para descargas, etc.
- La carpeta '*uploads*' contiene archivos subidos por la web y que están a la espera de ser incorporados al catálogo.
- La carpeta '*legacy*' contiene archivos necesarios para la migración de versiones anteriores de Séneca.

[EOF]