CORPUS STRUCTURE

ceph.git/ceph-object-corpus is a submodule.:

```
bin/    # misc scripts
archive/$version/objects/$type/$hash  # a sample of encoded objects from a specific version
```

You can also mark known or deliberate incompatibilities between versions with:

```
archive/$version/forward_incompat/$type
```

The presence of a file indicates that new versions of code cannot decode old objects across that $version (this is normally the case).

HOW TO GENERATE AN OBJECT CORPUS

We can generate an object corpus for a particular version of ceph like so.

1. Checkout a clean repo (best not to do this where you normally work):

   ```
   git clone ceph.git
   cd ceph
   git submodule update --init --recursive
   ```

2. Build with flag to dump objects to /tmp/foo:

   ```
   rm -rf /tmp/foo ; mkdir /tmp/foo
   ./do_autogen.sh -e /tmp/foo
   make
   ```

3. Start via vstart:

   ```
   cd src
   MON=3 OSD=3 MDS=3 RGW=1 ./vstart.sh -n -x
   ```

4. Use as much functionality of the cluster as you can, to exercise as many object encoder methods as possible:

   ```
   ./rados -p rbd bench 10 write -b 123
   ./ceph osd out 0
   ./init-ceph stop osd.1
   for f in ../qa/workunits/cls/*.sh ; do PATH=".:$PATH" $f ; done
   ../qa/workunits/rados/test.sh
   ./ceph_test_librbd
   ./ceph_test_libcephfs
   ./init-ceph restart mds.a
   ```

Do some more stuff with rgw if you know how.

1. Stop:

   ```
   ./stop.sh
   ```

2. Import the corpus (this will take a few minutes):

   ```
   test/encoding/import.sh /tmp/foo `./ceph-dencoder version` ../ceph-object-corpus/archive
   test/encoding/import-generated.sh ../ceph-object-corpus/archive
   ```

3. Prune it! There will be a bazillion copies of various objects, and we only want a representative sample.:

```
pushd ../ceph-object-corpus
bin/prune-archive.sh
popd
```

4. Verify the tests pass:

```
make check-local
```

5. Commit it to the corpus repo and push:

```
pushd ../ceph-object-corpus
git checkout -b wip-new
git add archive/`../src/ceph-dencoder version`
git commit -m `../src/ceph-dencoder version`
git remote add cc ceph.com:/git/ceph-object-corpus.git
git push cc wip-new
popd
```

6. Go test it out:

```
cd my/regular/tree
cd ceph-object-corpus
git fetch origin
git checkout wip-new
cd ../src
make check-local
```

7. If everything looks good, update the submodule master branch, and commit the submodule in ceph.git.