JCSDA Core Team

Claude Gibert
Rick Grubin

EWOK (Experiments and Workflow Orchestration Kit)
R2D2 (Research Repository for Data and Diagnostics)

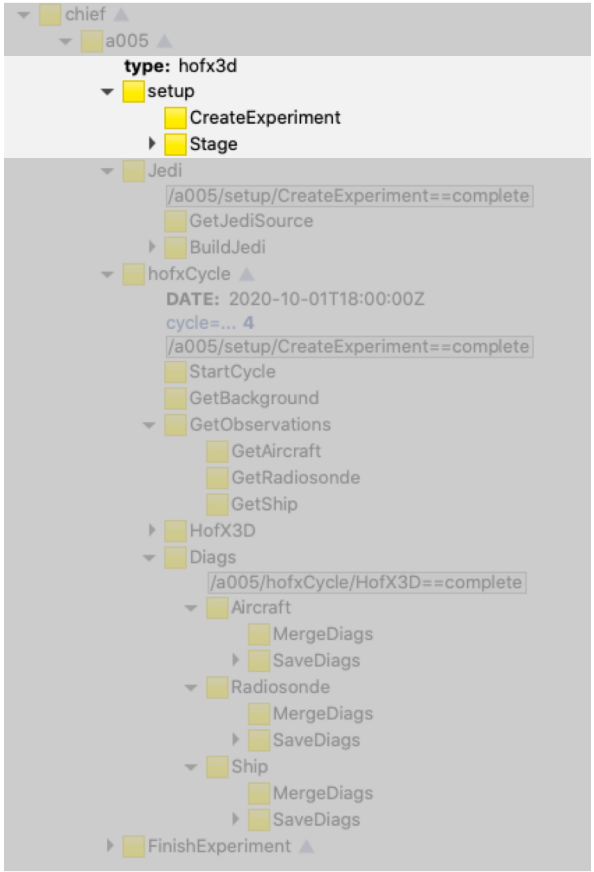
Today

- Workflow: ecflow
- How ewok attempts to leverage generic experiments
- How r2d2 abstracts meteorological data access
- How to parameterize experiments:
  - Setting up experiments
  - Specifying branches and tags to build a bundle
  - Machine specific information

ECFlow and EWOK

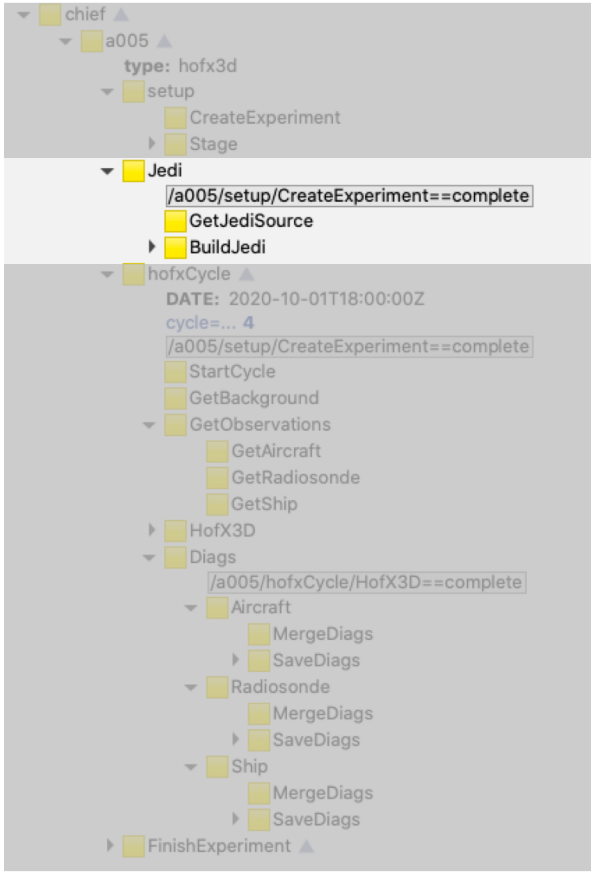
Live demo ☺

Prepare the suite



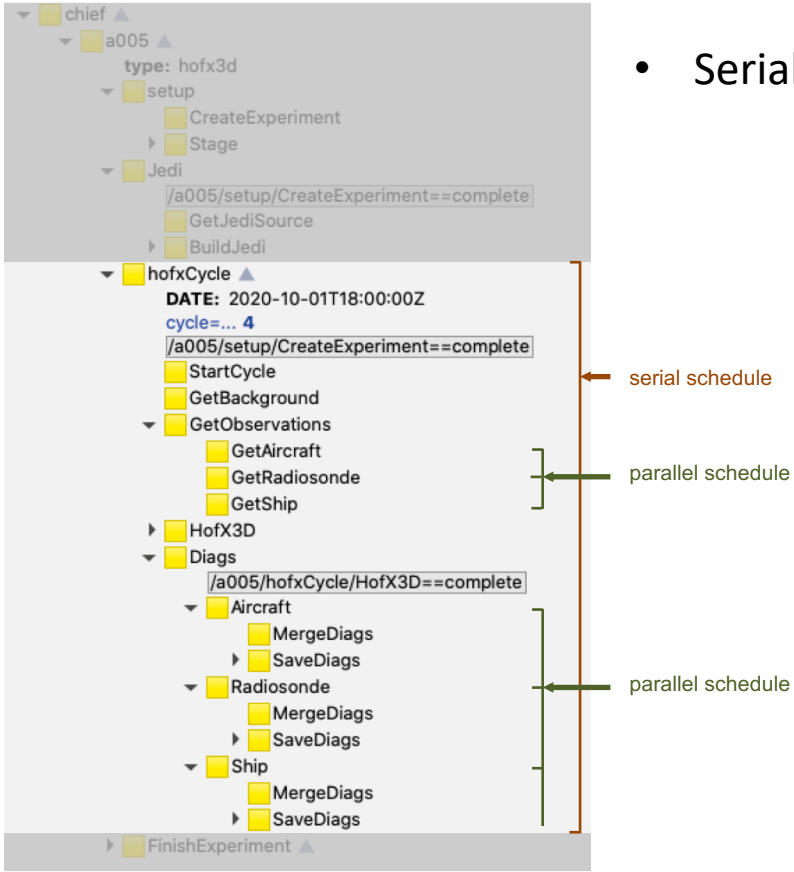
- Create directories
- Copy static model files (vertical levels, geometry, field sets, etc..)

Prepare the code



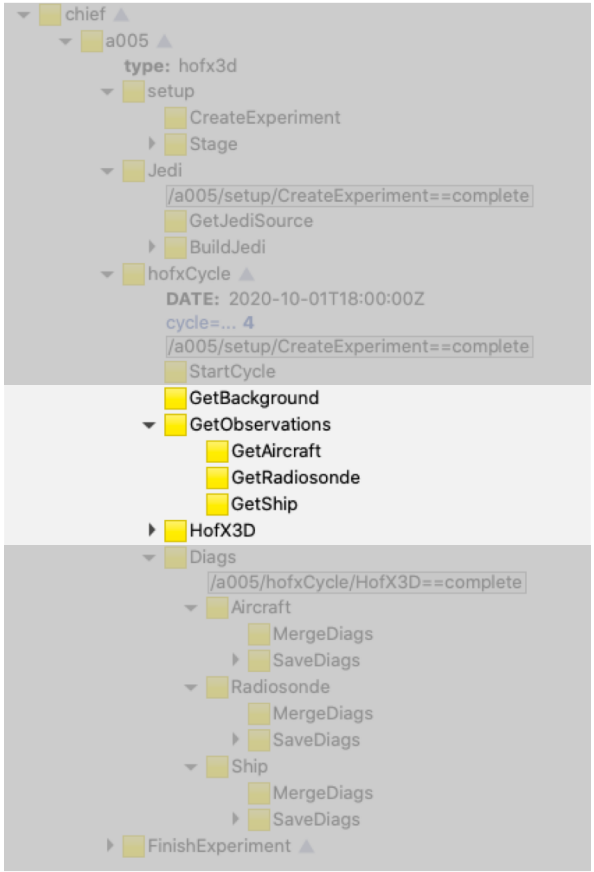
- Pull git repositories, branches, tags, commits
- Build the bundle (ecbuild, make)

Cycle



- Serial and parallel scheduling

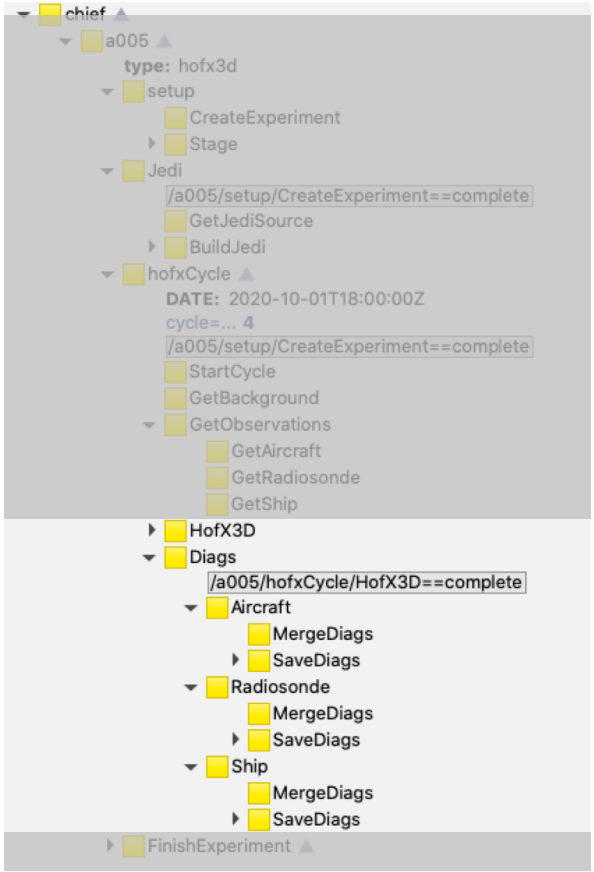
Get model input data



- Get or produce the background
- Get observations

- When invoking the binary, ewok build the yaml necessary to run the model using templates sitting in an ewok directory in the oops repository.
- The templates refer to yaml files residing in the chosen model repository, for example, in fv3-jedi/ewok/gfs or fv3-jedi/ewok/geos.

Postprocess and save data



- Merge files
- Prepare diagnostics
- Save data

- The output data, available in r2d2, can be visualized anywhere. Comparisons between experiments can be made.

Complete

- Save the experiment configuration, including git information to be able to reproduce the experiment.
- Cleanup files, experiment.

Building block

- Prepare the suite
- Prepare the code
- Cycle
  - Get input data
  - Preprocess
  - Run model
  - Postprocess and save output data

Those are similar for most types of experiments. This enables us to have a small set of generics suites.

Task Genericity

- Calling the model binaries with the right configuration files. Templates are used

- Moving data around. In order to deal with data in a generic way, R2D2 provides a high level interface. Meteorological data (fields, observations) are described with a set of key value pairs mapping to a unique piece of data (a file).

R2D2 - fetch

```
fetch(
    type='ob',
    provider='ncdiag',
    experiment='oper',
    obs_type=['aircraft', 'radiosonde'],
    date=date_sequence('2020-10-01T00:00:00Z', '2020-10-01T18:00:00Z' 'PT6H'),
    time_window='PT6H'
)
```

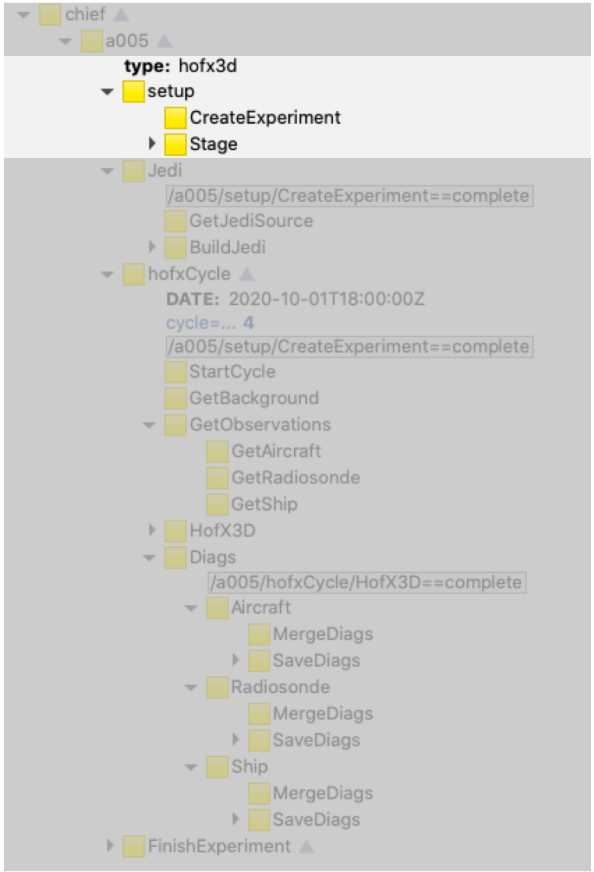
R2D2 - fetch

```
fetch(
    type='fc',
    model='gfs',
    experiment='oper',
    resolution='c768',
    step='PT6H',
    date=date_sequence('2020-10-01T00:00:00Z', '2020-10-01T18:00:00Z' 'PT6H'),
    file_type=['fv_core.res', 'fv_tracer.res'],
    tile=[1, 2, 3, 4, 5, 6],
    target_file='$(type).$(step).$(date).nc'
)
```

Today

- Workflow: ecflow
- How ewok attempts to leverage generic experiments
- How r2d2 abstracts meteorological data access
- How to parameterize experiments:
  - Setting up experiments
  - Specifying branches and tags to build a bundle
  - Machine specific information

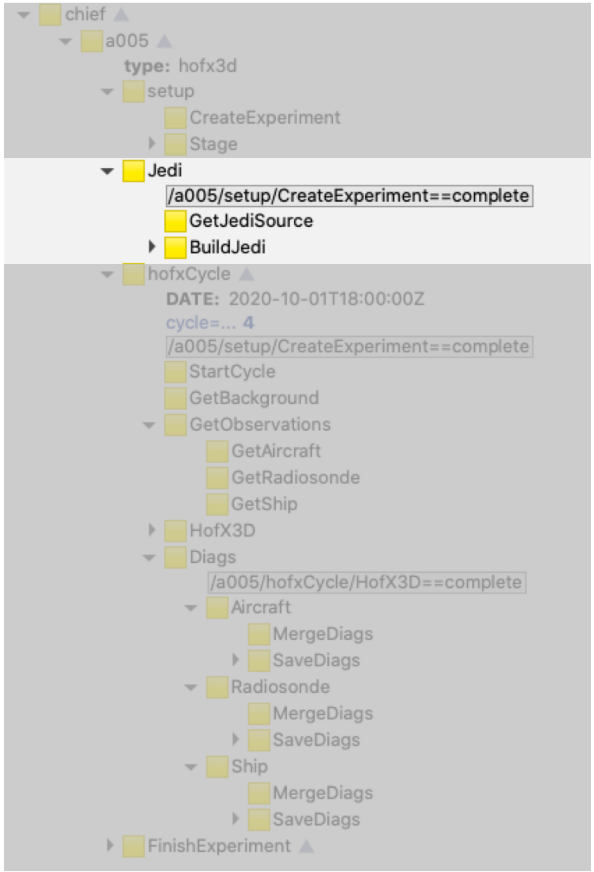
Prepare the suite



- Create directories
- Copy static model files (vertical levels, geometry, field sets, etc..)

```
# Templates
STAGE:
  - git://fv3-jedi/ewok/gfs/stage.yaml
GEOMETRY: git://fv3-jedi/ewok/gfs/geometry.yaml
BACKGROUND: git://fv3-jedi/ewok/gfs/bg.yaml
AN_TEMPLATE: git://fv3-jedi/ewok/gfs/an.yaml
OB_TEMPLATE: git://fv3-jedi/ewok/gfs/ob.yaml
R2D2: git://fv3-jedi/ewok/gfs/r2d2.yaml
step_cycle: PT6H
```

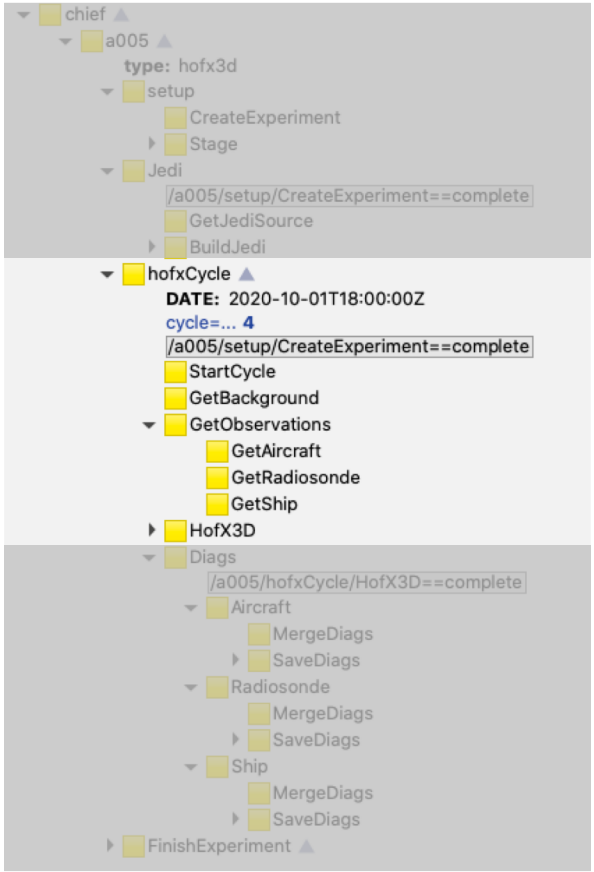
Prepare the code



- Pull git repositories, branches, tags, commits
- Build the bundle (ecbuild, make)

```yaml
# Repos
model_repos:
  - url: $(github)/FMS.git
    branch: release-stable
    project: fms
  - url: $(github)/GFDL_atmos_cubed_sphere.git
    branch: release-stable
    project: fv3
  - url: $(github)/femps.git
    branch: develop
  - url: $(github)/fv3-jedi-linearmodel.git
    branch: develop
    project: fv3-jedi-lm
  - url: $(github)/fv3-jedi.git
    branch: develop
```

Get model input data

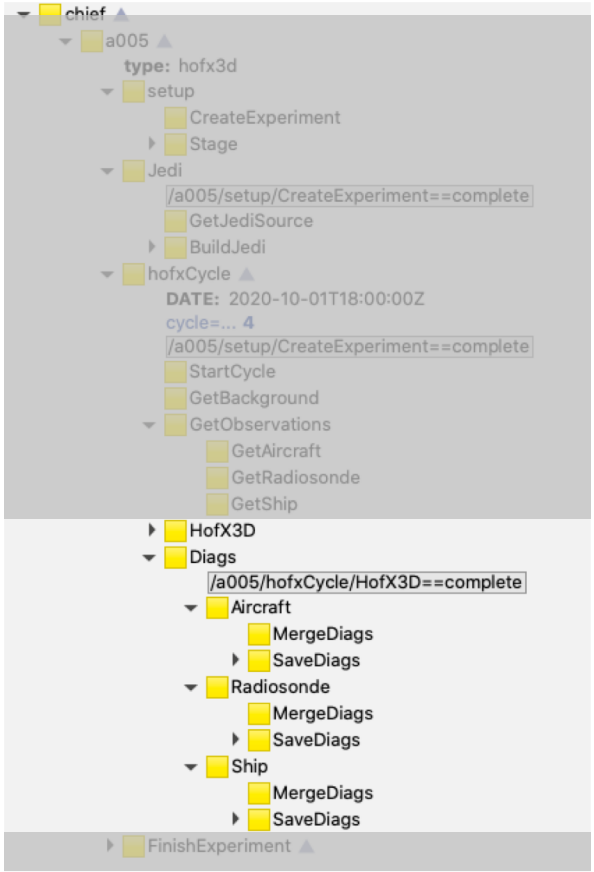


- Get or produce the background
- Get observations

```yaml
# Templates
OBSERVATIONS:
  - git://ufo/ewok/aircraft.yaml
  - git://ufo/ewok/radiosonde.yaml
  - git://ufo/ewok/ship.yaml

# Parameters observation experiment (usually oper)
obs_experiment: oper
bg_init: oper
# needed in input.yaml for hofx
forecast_length: PT6H
```

Run the model, postprocess and save data



- Merge files
- Prepare diagnostics
- Save data

```yaml
# resolution
horizontal_resolution: c12
vertical_resolution: 64

# executables
hofx:
  exec: fv3jedi_hofx_nomodel.x
  args: ''
  ntasks: 6

# host
HOST: hosts/standard_mpi.yaml
```

ECFlow and EWOK

Questions? ☺