

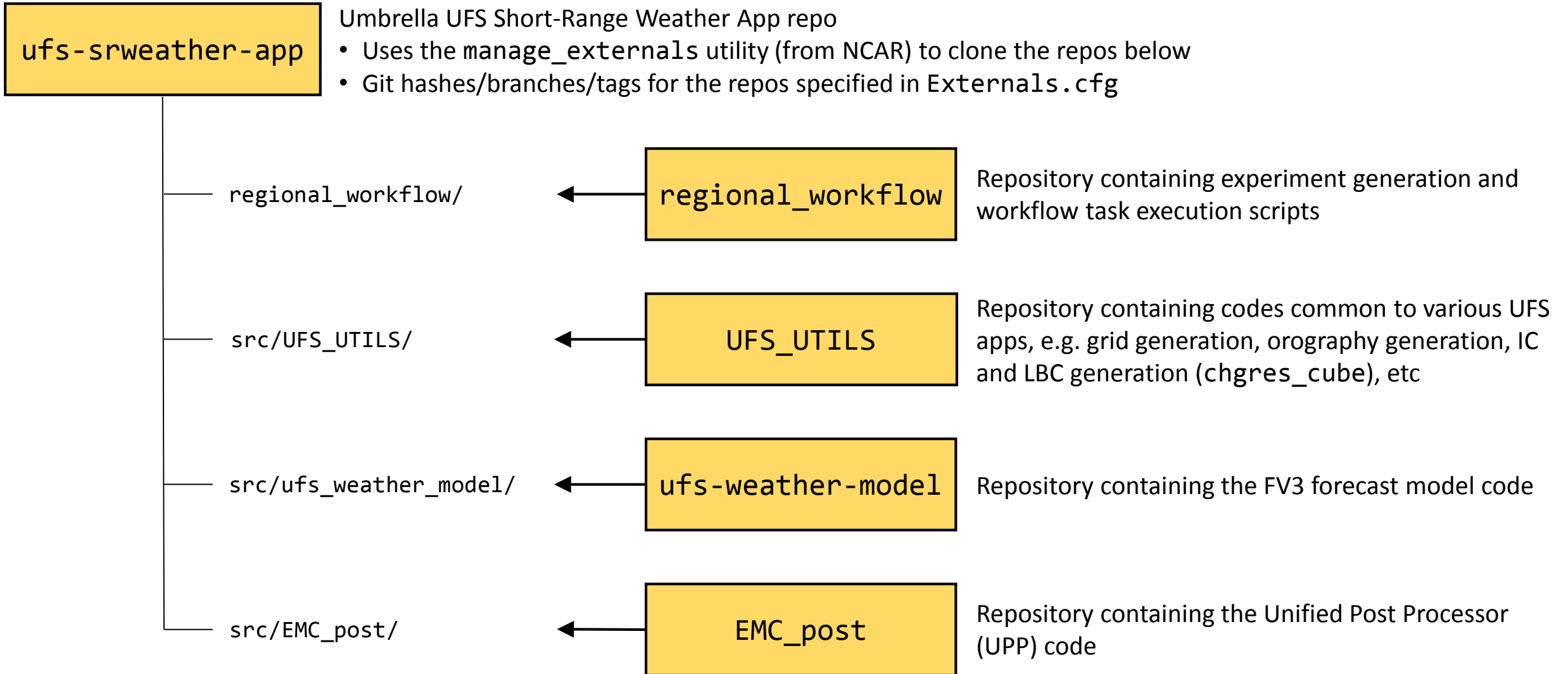
Configuring the UFS SRW App Experiment/Workflow

Gerard Ketefian

CIRES/NOAA/ESRL/GSL/DTC

September 20, 2021

UFS SRW App Repository Structure



Obtaining and Running the UFS SRW App

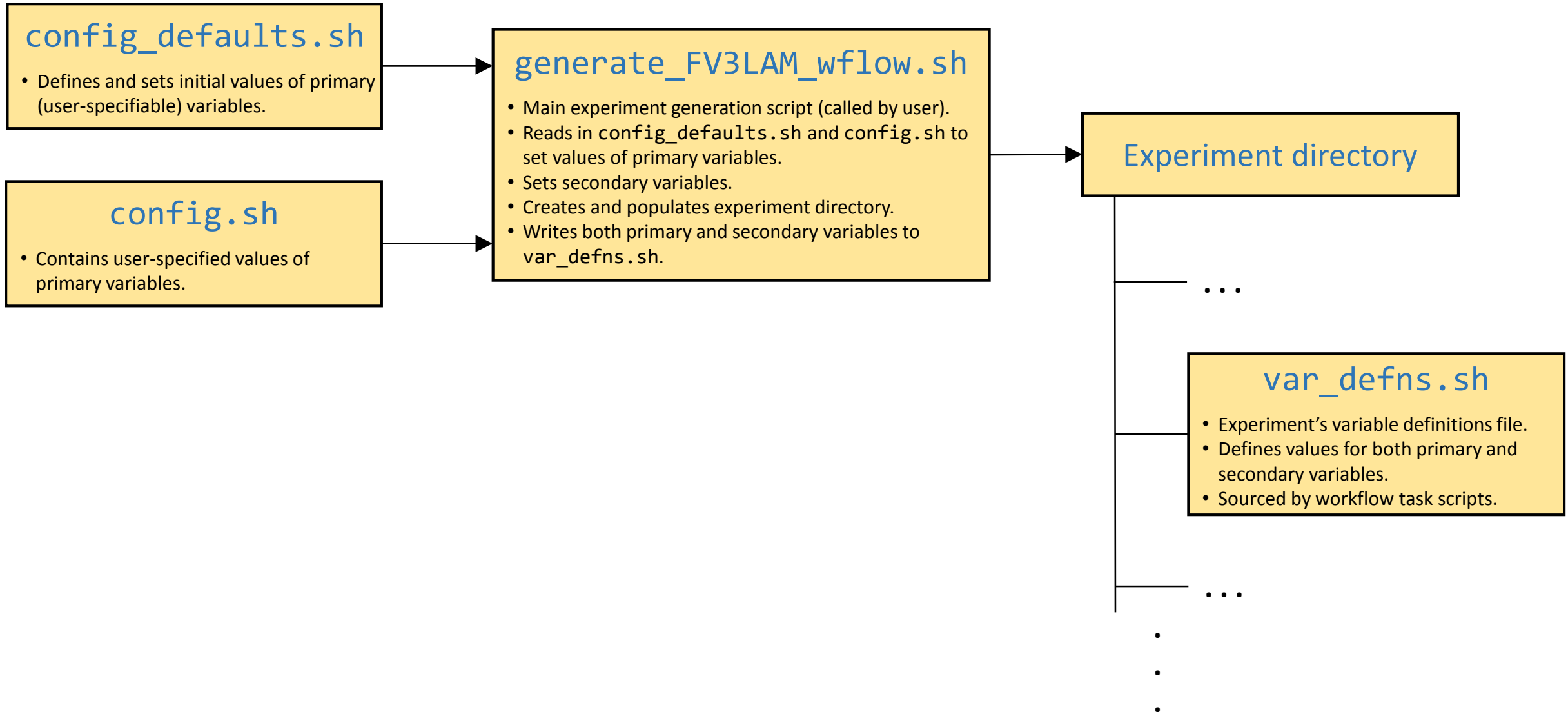
Do only once

- Clone the umbrella `ufs-srweather-app` repository from <https://github.com/ufs-community/ufs-srweather-app>.
- Use the `manage_externals` utility to clone the external repos:
 - `regional_workflow`
 - `UFS_UTILS`
 - `ufs-weather-model`
 - `EMC_post`
- Build the codes.

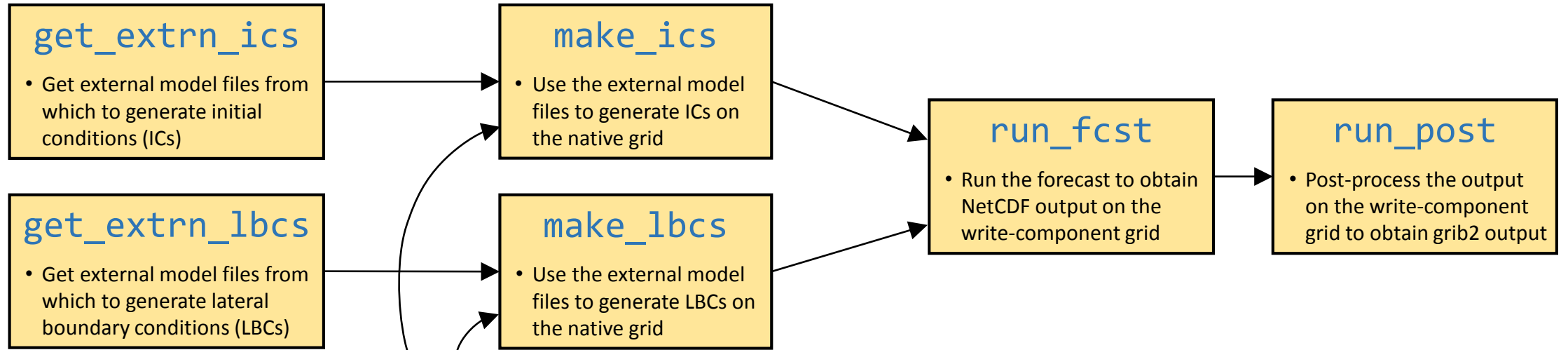
Do once per
experiment

- **Create/edit the configuration file (`config.sh`) that specifies experiment configuration variables.**
 - Examples provided in `regional_workflow` repo (`config.community.sh` and `config.nco.sh`)
- Generate an experiment directory and rocoto workflow XML by running the script `generate_FV3LAM_wflow.sh`.
- (Re)launch the experiment using `rocotorun` and monitor its status using `rocotostat` (or, alternatively, use the `launch_FV3LAM_wflow.sh` script in the experiment directory) until all workflow tasks are complete or one fails.
 - Automatic relaunch capability using `cron` also available via the `USE_CRONTAB_TO_RELAUNCH` experiment variable.

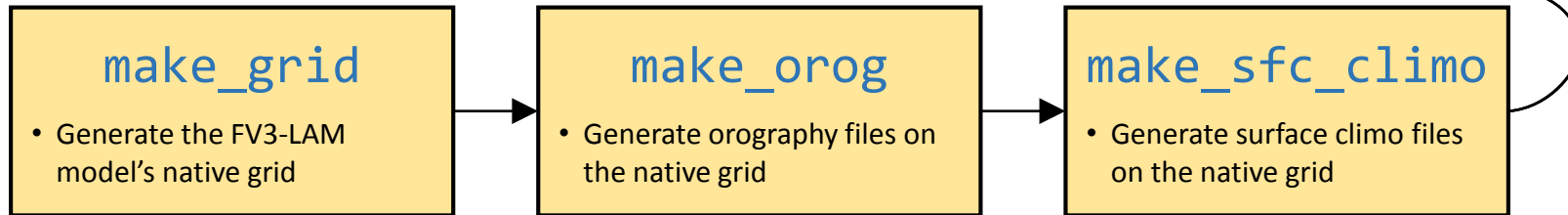
Experiment Generation Overview



Workflow Tasks



Run once per cycle



Pre-processing steps; run only once per experiment

OR

skip and instead use pregenerated grid, orography, and surface climo files

Experiment Variables (1 of 2)

- Categorized into two groups:
 - Primary variables:
 - User-specifiable.
 - Defined and assigned initial values in `config_defaults.sh`.
 - User specifies (a subset of) these in `config.sh`.
 - If not explicitly set by the user in `config.sh`, may get reset during experiment generation to a value that depends on other primary and/or secondary variables. Example:
If not specified in `config.sh`, `QUEUE_DEFAULT` gets reset from its initial value of a null string in `config_defaults.sh` to a value that depends on the primary variable `MACHINE`.
 - Secondary variables:
 - Cannot be specified by the user.
 - Set during experiment generation in a way that may depend on value of other primary and/or secondary variables. Example:
The location of the experiment directory (`EXPTDIR`; a secondary variable) is obtained from the primary variables `EXPT_BASEDIR` and `EXPT_SUBDIR` as follows:
`EXPTDIR="${EXPT_BASEDIR}/${EXPT_SUBDIR}"`

Experiment Variables (2 of 2)

- Both primary and secondary variables are recorded in the experiment's variable definitions file (`var_defns.sh`; located at the top level of the experiment directory, `$EXPTDIR`) during the experiment generation step.
- All workflow task scripts source `var_defns.sh` in order to have access to both primary and secondary experiment variables.

Experiment Configuration Files

- `config_defaults.sh`:
 - Defines the full set of primary (i.e. user-specifiable) experiment variables and sets their default values.
 - Located in `regional_workflow/ush`.
 - Read in by `generate_FV3LAM_workflow.sh`.
 - Contains description of each primary variable.
- **`config.sh`**: **This is the file users modify to customize the experiment.**
 - Used to customize the experiment; edit this file to specify values for the primary experiment variables.
 - Located in `regional_workflow/ush`.
 - Read in by `generate_FV3LAM_workflow.sh` (after `config_defaults.sh`).
 - Cannot be used to specify any arbitrary variable, only those defined in `config_defaults.sh`.

Configuration Variables:

Experiment Mode (1 of 2)

- Experiment can be run in one of two modes: research and operational
- Mode determines the directory structure of the experiment.
 - Research mode: All experiment components, including all workflow, input, and output files, are contained in a single directory specified by `$EXPTDIR`.
 - Operational mode: Experiment components are distributed among three different directories.
 - `$EXPTDIR`: Directory containing variable definitions, rocoto XML, and log files as well as cycle-independent input files to the forecast model.
 - `$STMP`: Directory containing “intermediate” files, e.g. cycle-dependent ICs and LBCs; cycle-dependent forecast input files; and raw forecast output files.
 - `$PTMP`: Directory containing post-processed files generated by UPP.

Configuration Variables:

Experiment Mode (2 of 2)

- **RUN_ENVIR**

- Determines mode in which the experiment will run: research or operational
- Valid values: "community" "nco"
- Set to "community" to run in research mode and to "nco" to run in operational mode (NCO = NCEP Central Operations, where NCEP = National Centers for Environmental Prediction).
- Operational mode not officially supported.
- During this training, will only run in research mode.

Configuration Variables: System Information

Using CHEYENNE for this training



- **MACHINE**

- Name of machine/platform on which the workflow is running.
- Valid values:
"WCOSS_CRAY" "WCOSS_DELL_P3" "HERA" "ORION" "JET" "ODIN" "CHEYENNE"
"STAMPEDE" "GAEA" "MACOS" "LINUX"
- Default value: "BIG_COMPUTER"

- **ACCOUNT**

- Account on the specified machine under which to submit jobs to the queue (and to which to charge core-hours).
- Default value: "an_account"

- **WORKFLOW_MANAGER**

- Name of the workflow manager, if any.
- Valid values: "rocoto" "none"
- Default value: "rocoto"
- If set to "none", the XML file that defines the rocoto workflow (FV3LAM_workflow.xml) is not generated.

Configuration Variables:

Experiment Directory Location

- **EXPT_BASEDIR**

- Base directory in which the experiment directory will be created.
- Default value: `${HOME}rfs/./expt_dirs`
where `${HOME}rfs` is the full path to your `ufs-srweather-app` directory.
⇒ `expt_dirs` is on same level in directory tree as `ufs-srweather-app`.

- **EXPT_SUBDIR**

- Name of subdirectory under `${EXPT_BASEDIR}` in which the experiment will be located.
- Full path to experiment directory is thus:
`EXPTDIR="${EXPT_BASEDIR}/${EXPT_SUBDIR}"`
- Must be specified by user in `config.sh`.
- Also considered the experiment's name.

Configuration Variables:

Forecast Dates, Cycle Hours, Forecast Length (1 of 2)

- **DATE_FIRST_CYCL, DATE_LAST_CYCL**
 - Starting days (not including the hour-of-day) of the first and last forecasts that the experiment will run.
 - Each must be an 8-digit string of the form YYYYMMDD, where YYYY is a 4-digit year, MM is a 2-digit month, and DD is a 2-digit day-of-the-month.
- **CYCL_HRS**
 - Array containing the hours of the day at which forecasts are launched.
 - Each element must be a 2-digit integer between "00" and "23".
- **FCST_LEN_HRS**
 - Length of each forecast, in hours.

Configuration Variables:

Forecast Dates, Cycle Hours, Forecast Length (2 of 2)

Example:

The following code snippet in `config.sh` will instruct the workflow to run 30 forecasts, each 36 hours long, that start on 2021060100, 2021060112, 2021060200, 2021060212, ..., 2021061500, and 2021061512:

```
DATE_FIRST_CYCL="20210601"  
DATE_LAST_CYCL="20210615"  
CYCL_HRS=( "00" "12" )  
FCST_LEN_HRS="36"
```

Configuration Variables: Predefined Grids & Physics Suites

- **PREDEF_GRID_NAME**

- Name of the predefined grid to use in all forecasts run by the experiment.

- Valid values:

- "RRFS_CONUS_25km" "RRFS_CONUS_13km" "RRFS_CONUS_3km"

- Must be specified by user in `config.sh` (unless using a custom grid; creation of custom grids will be covered on Day 2 of this training).

- **CCPP_PHYS_SUITE**

- Name of physics suite to use in all forecasts run by the experiment.

- Valid values: "FV3_GFS_v15p2" "FV3_RRFS_v1a1pha"

- Default value: "FV3_GFS_v15p2"

Configuration Variables:

External Model Specification

- **EXTRN_MDL_NAME_ICS, EXTRN_MDL_NAME_LBCS**
 - Names of the external models from which to generate initial conditions (ICs) and lateral boundary conditions (LBCs).
 - Valid values: "GSMGFS" "FV3GFS" "RAP" "HRRR" "NAM"
 - Default value: "FV3GFS"
- **LBC_SPEC_INTVL_HRS**
 - Time interval (in integer hours) between LBC specification forecast hours.
 - Default value: "6"
 - `get_extrn_lbc` task will attempt to get an external model forecast file for each LBC specification hour, e.g. if `LBC_SPEC_INTVL_HRS=3`, then files will be obtained for forecast hours 3, 6, 9, ...
- **FV3GFS_FILE_FMT_ICS, FV3GFS_FILE_FMT_LBCS**
 - If using the FV3GFS global model for generating ICs (i.e. if `EXTRN_MDL_NAME_ICS` is set to "FV3GFS") or LBCs (i.e. if `EXTRN_MDL_NAME_LBCS` is set to "FV3GFS"), the assumed format of the FV3GFS files.
 - Valid values: "nemsio" "grib2"
 - Default value: "nemsio"

Configuration Variables:

User-Staged External Model Data (1 of 2)

- **USE_USER_STAGED_EXTRN_MDL_FILES**
 - Flag that specifies whether to look for the external model files in user-specified directories.
 - Valid values: "TRUE" "FALSE"
 - Default value: "FALSE"
- **EXTRN_MDL_SOURCE_BASEDIR_ICS, EXTRN_MDL_SOURCE_BASEDIR_LBCS**
 - Base directories in which to look for external model files from which to generate ICs and LBCs.
 - Full directory obtained by appending the cycle's date and time to these variables. Example:
For a cycle that starts on 2021060100, the workflow scripts will look for the external model IC files in
`${EXTRN_MDL_SOURCE_BASEDIR_ICS}/2021060100/`
and for the LBC files in
`${EXTRN_MDL_SOURCE_BASEDIR_LBCS}/2021060100/`
- **EXTRN_MDL_FILES_ICS, EXTRN_MDL_FILES_LBCS**
 - Arrays containing names of the external model files from which ICs and LBCs will be generated.

Configuration Variables: User-Staged External Model Data (2 of 2)

Example:

The following code snippet in `config.sh` will instruct the workflow to use grib2-formatted files from the FV3GFS to generate the ICs and LBCs for 48-hour forecasts that use LBCs specified every 6 hours:

```
FCST_LEN_HRS="48"  
LBC_SPEC_INTVL_HRS="6"
```

```
EXTRN_MDL_NAME_ICS="FV3GFS"  
EXTRN_MDL_NAME_LBCS="FV3GFS"
```

```
FV3GFS_FILE_FMT_ICS="grib2"  
FV3GFS_FILE_FMT_LBCS="grib2"
```

```
USE_USER_STAGED_EXTRN_FILES="TRUE"
```

```
EXTRN_MDL_SOURCE_BASEDIR_ICS="/glade/p/ral/jntp/UFS_SRW_app/staged_extrn_md1_files/FV3GFS"
```

```
EXTRN_MDL_FILES_ICS=( "gfs.pgrb2.0p25.f000" )
```

```
EXTRN_MDL_SOURCE_BASEDIR_LBCS="/glade/p/ral/jntp/UFS_SRW_app/staged_extrn_md1_files/FV3GFS"
```

```
EXTRN_MDL_FILES_LBCS=( "gfs.pgrb2.0p25.f006" "gfs.pgrb2.0p25.f012" "gfs.pgrb2.0p25.f018" \  
"gfs.pgrb2.0p25.f024" "gfs.pgrb2.0p25.f030" "gfs.pgrb2.0p25.f036" \  
"gfs.pgrb2.0p25.f042" "gfs.pgrb2.0p25.f048" )
```

Since $FCST_LEN_HRS/LBC_SPEC_INTVL_HRS = 48/6 = 8$,
must specify 8 LBC files in `EXTRN_MDL_FILES_LBCS`

Configuration Variables: Pregenerated Fixed Files

The grid, orography, and surface climatology files depend only on the grid (thus called “fixed” files).

- No need to regenerate them for experiments that use the same grid.
- Can skip the `make_grid`, `make_orog`, and `make_sfc_climo` tasks and instead point to pregenerated files (generated by a previous experiment on the same grid).

- **RUN_TASK_MAKE_GRID**
 - Flag specifying whether or not to run the `make_grid` task as part of the workflow.
 - Valid values: "TRUE" "FALSE"
 - Default value: "TRUE"

- **GRID_DIR**
 - The directory in which to look for pregenerated grid files.
 - If `RUN_TASK_MAKE_GRID` is set to "FALSE", this is set automatically.

- **RUN_TASK_MAKE_OROG, OROG_DIR**
 - Same as `RUN_TASK_MAKE_GRID` and `GRID_DIR` but for the `make_orog` task.

- **RUN_TASK_MAKE_SFC_CLIMO, SFC_CLIMO_DIR**
 - Same as `RUN_TASK_MAKE_GRID` and `GRID_DIR` but for the `make_sfc_climo` task.

Configuration Variables: Post-processor (UPP)

- **USE_CUSTOM_POST_CONFIG_FILE**
 - Flag specifying whether or not to use a custom (i.e. user-specified) configuration file for the post-processor (UPP).
 - Valid values: "TRUE" "FALSE"
 - Default value: "FALSE"
 - If "FALSE", a default configuration file included in the `regional_workflow` repo is used.
- **CUSTOM_POST_CONFIG_FP**
 - Full path to the custom post-processor configuration file.
 - Used only if `USE_CUSTOM_CONFIG_FP` is set to "TRUE".

Configuration Variables:

Miscellaneous

- **PREEXISTING_DIR_METHOD**

- Flag that specifies how to deal with a preexisting experiment directory and most preexisting subdirectories (e.g. if running a workflow task a 2nd, 3rd, etc time).
- Valid values: "delete" "rename" "quit"
 - "delete": Delete existing directories.
 - "rename": Rename existing directories by appending the strings `_old001`, `_old002`, etc to the name.
 - "quit": Quit if an existing directory is encountered.
- Default value: "delete"

- **USE_CRON_TO_RELAUNCH**

- Flag specifying whether or not to add a line in the user's cron table that calls the experiment (re)launch script (`launch_FV3LAM_wflow.sh`).
- Valid values: "TRUE" "FALSE"
- Default value: "FALSE"

- **CRON_RELAUNCH_INTVL_MNTS**

- Interval (in minutes) between successive cron calls to `launch_FV3LAM_wflow.sh`.
- Default value: "03"
- Used only if `USE_CRON_TO_RELAUNCH` is set to "TRUE".

Questions?