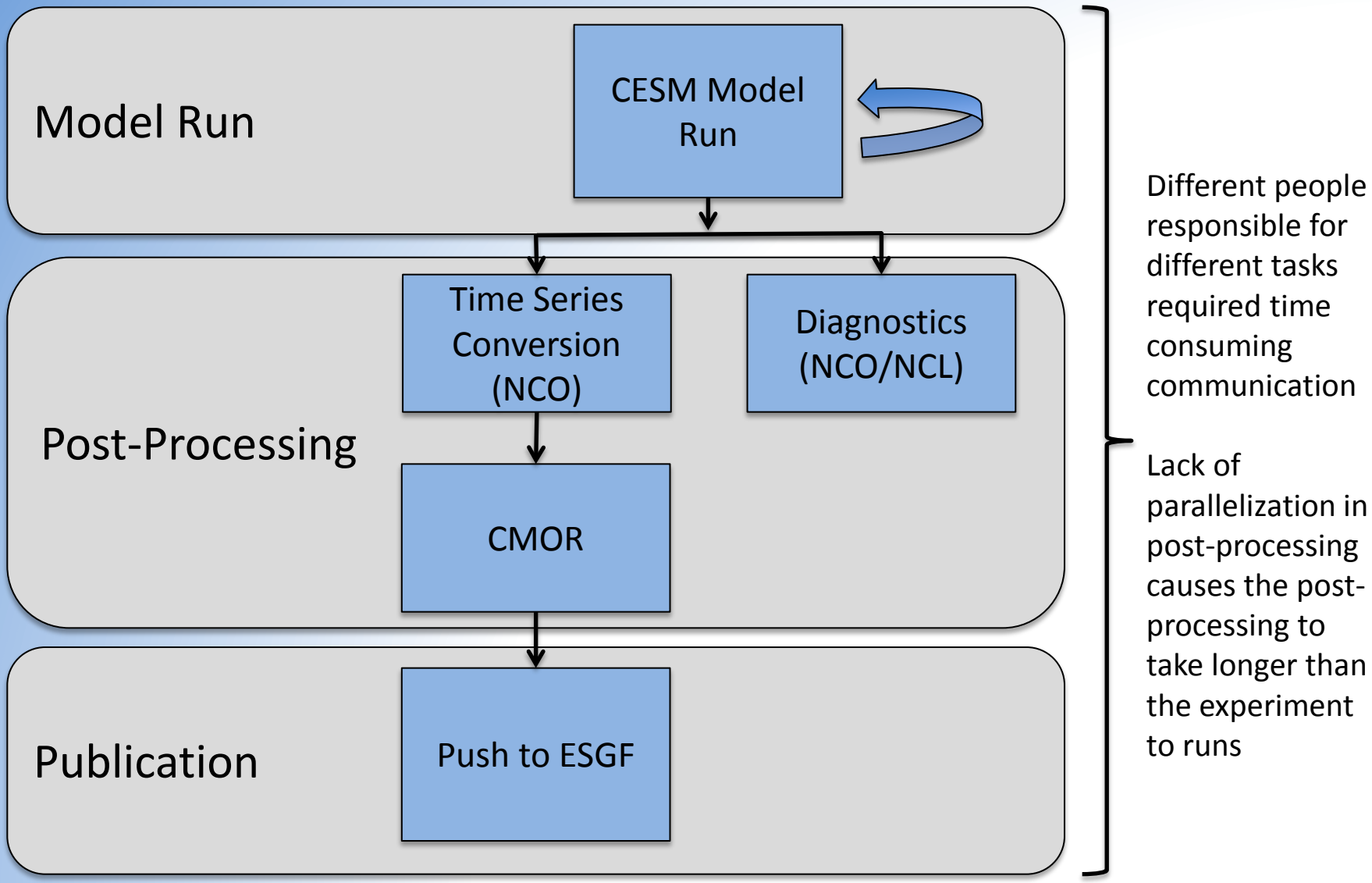
Optimizing Workflow for CESM

Sheri Mickelson
ASAP/TDD/CISL
CSEG/CGD

Thanks to all of the CESM workflow refactor team for their work and guidance

- Ben Andre
- Alice Bertini
- John Dennis
- Jim Edwards
- Mary Haley
- Jean-Francois Lamarque
- Michael Levy
- Sheri Mickelson
- Kevin Paul
- Sean Santos
- Jay Shollenberger
- Gary Strand
- Mariana Vertenstein

Current CESM Workflow



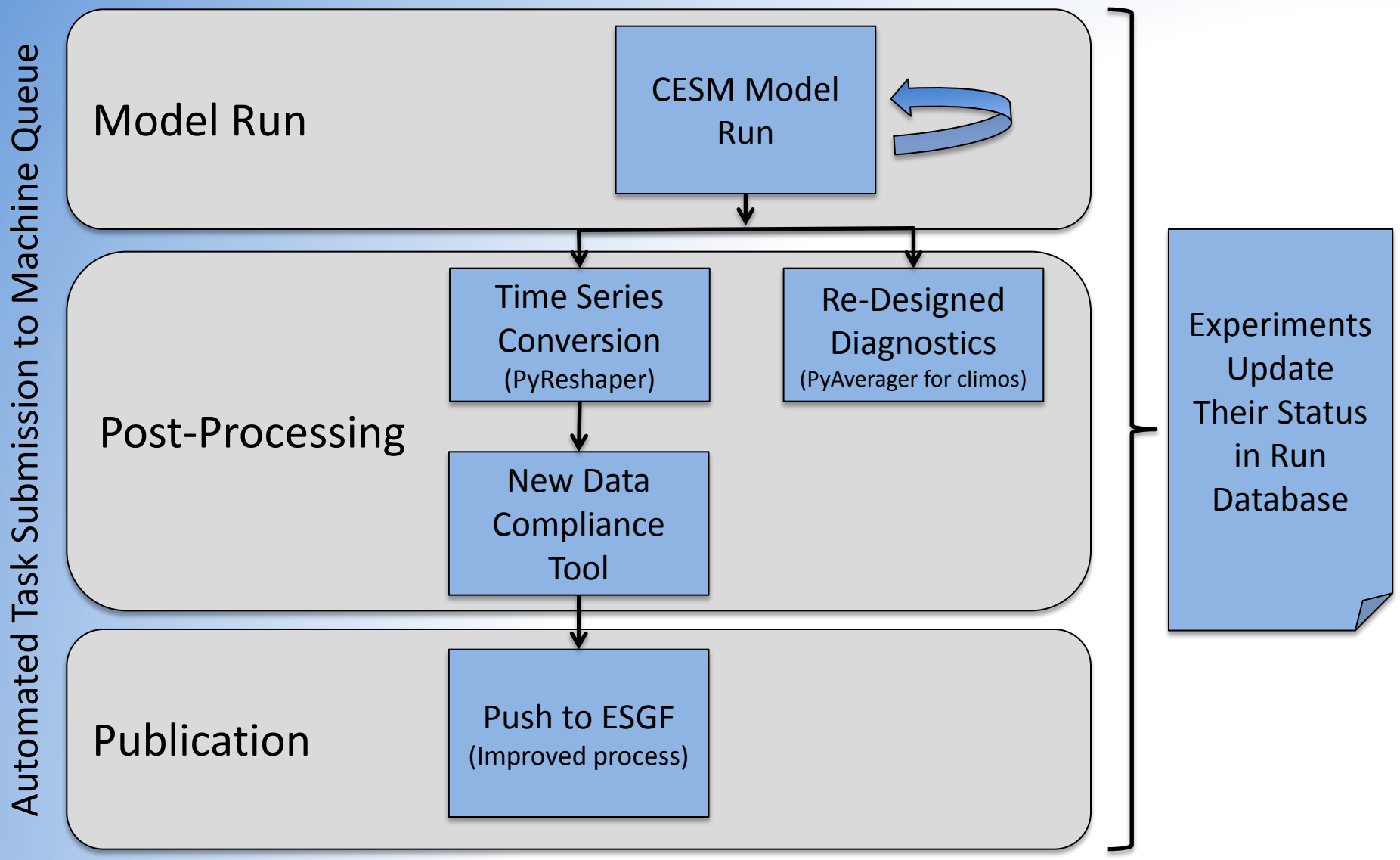
Different people responsible for different tasks required time consuming communication

Lack of parallelization in post-processing causes the post-processing to take longer than the experiment to runs

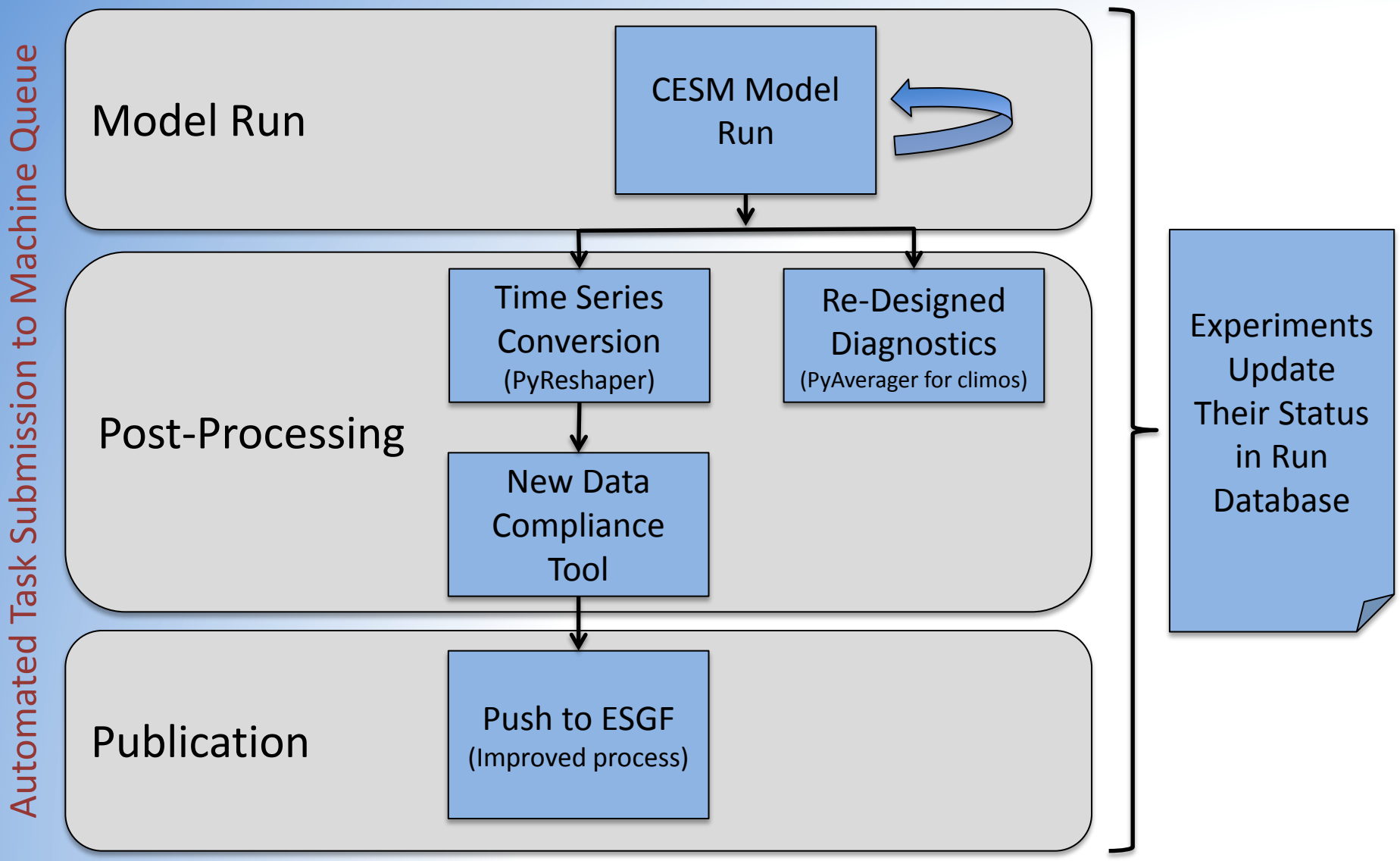
Goals For New CESM/CMIP6 Workflow

- Improve orchestration of the workflow and add in automation
- Examine the individual pieces of the workflow and improve where necessary
- Add parallelization into the workflow
- Create more formal project management tools

New CESM/CMIP6 Workflow



New CESM/CMIP6 Workflow

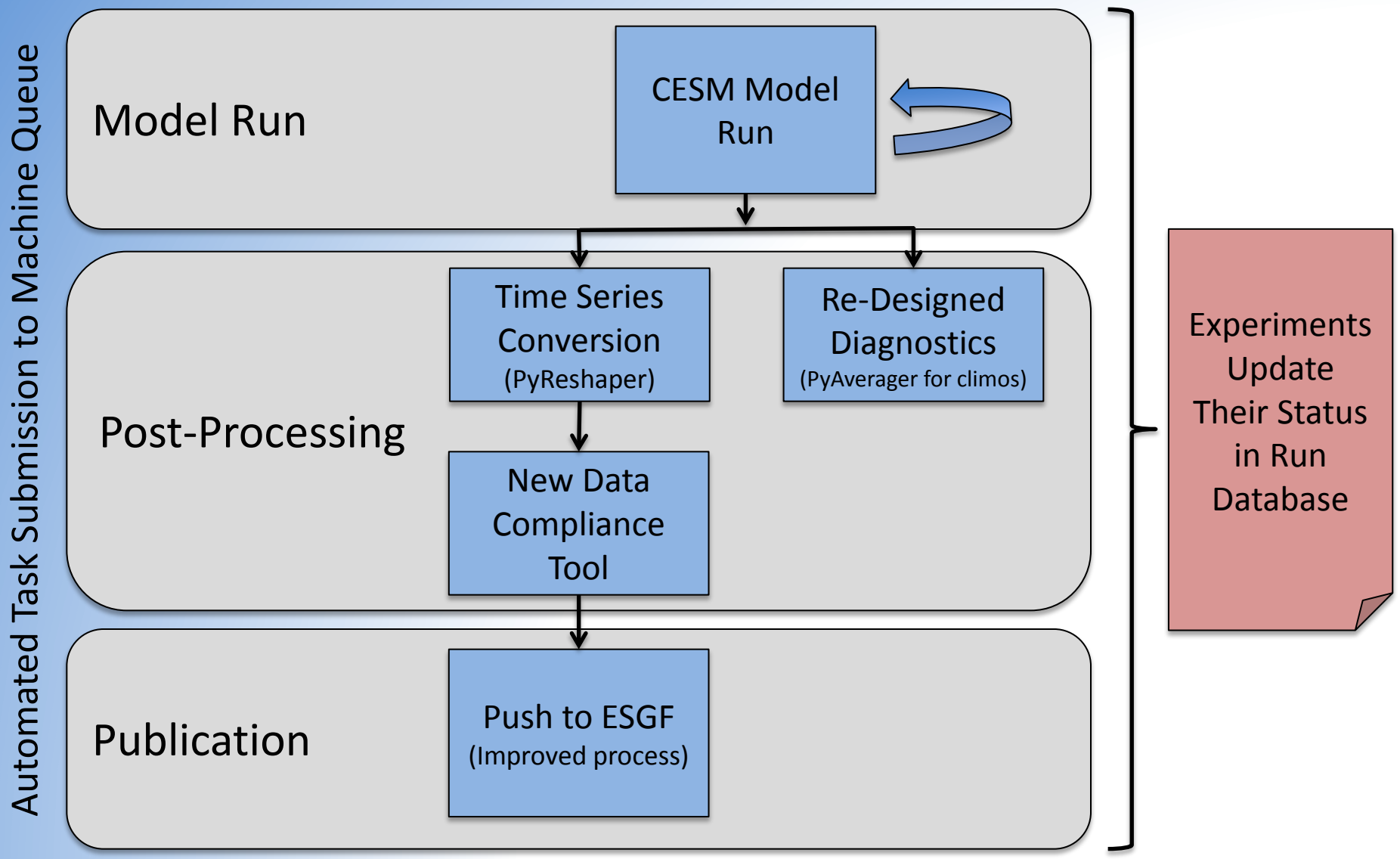


Simple diagram, but a lot going on behind the scenes …

Orchestration and automating the submission of CESM and post-processing tasks

- Users are able to turn post-processing tasks on/off and select date ranges to process
- Based on selections, post-processing jobs will automatically be submitted to the queuing system

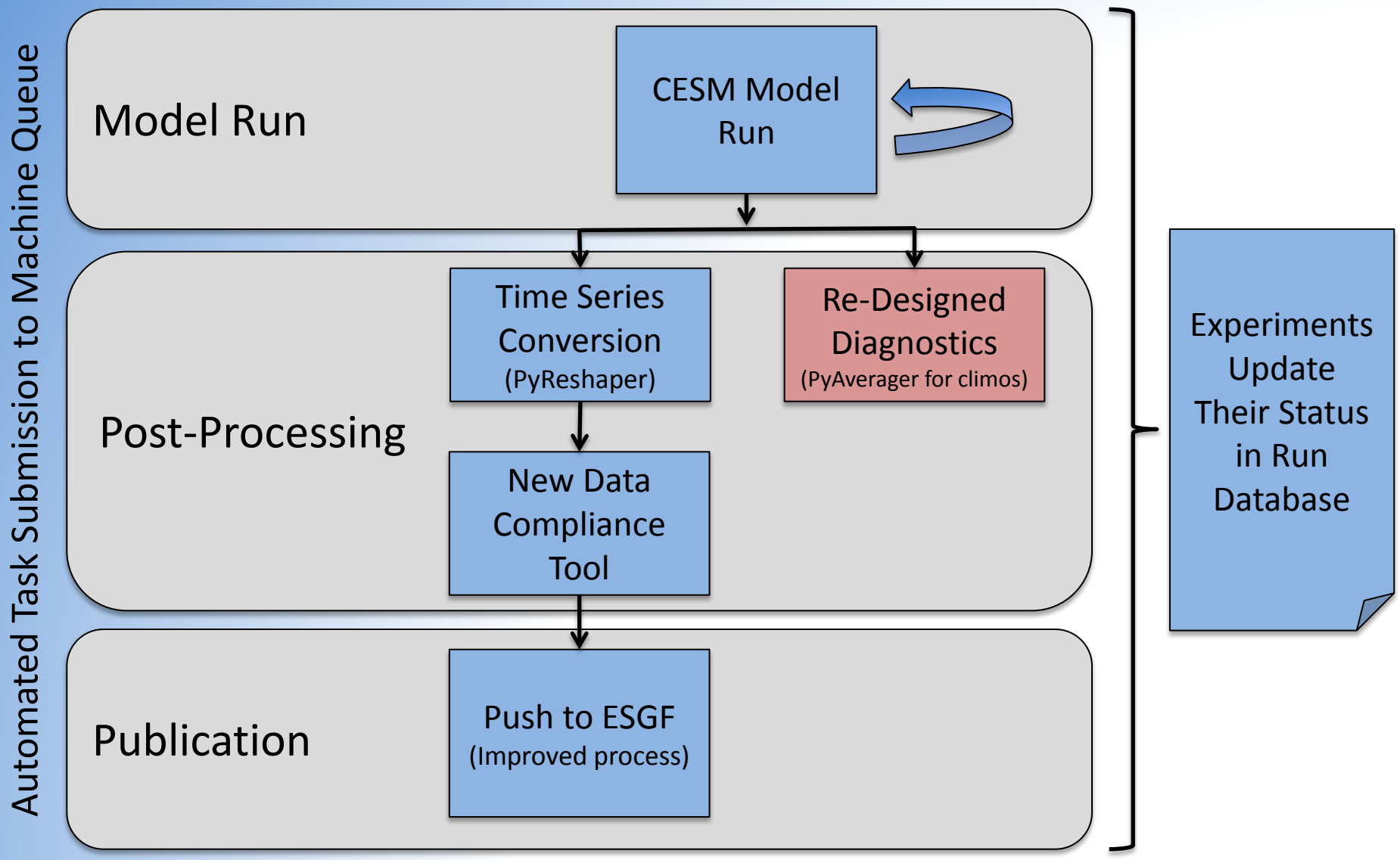
New CESM/CMIP6 Workflow



Enhancements to the Run Database

- There will be a separate section for CMIP6 experiments
- All experiments will update status to the database (simulation progress/color coded run status)
- CMIP6 timeline views
- Resource tracking (i.e. available disk space)
- Optional link to diagnostic web pages

New CESM/CMIP6 Workflow



Re-Design of Diagnostic Packages

- Can be automatically ran as part of a CESM run, but they still maintain the capability to ran standalone
  - Configured through XML Options
  - Sets up a Python Virtual Environment for users
- Instead of NCO, use the PyAverager to create the climatology files (from either time slice or time series files)
- Runs the same NCL plotting scripts in parallel

Climatology Files Created by the AMWG, OMWG, Land and Ice Diagnostic Packages

ATM - SE

b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_07_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_01_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_JJA_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_09_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_03_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_SON_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_11_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_12_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_ANN_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_DJF_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_08_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_02_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_MAM_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_10_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_04_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_05_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_06_climo.nc

ICE

ice_vol_b.e12.B1850C5CN.ne30_g16.init.ch.027_1-10.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cice.h_ANN_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cice.h_jfm_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cice.h_amj_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cice.h_jas_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cice.h_ond_climo.nc

LND

b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_ANN_ALL.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_SON_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_MAM_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_JJA_means.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_ANN_means.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_DJF_means.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_JJA_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_ANN_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_DJF_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_SON_means.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_MAM_means.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_MONS_means.nc

OCN

Arc_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Hud_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Atl_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Ind_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Lab_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Gin_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Sou_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Glo_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Pac_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0001.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0003.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0005.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0007.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0009.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_01_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_03_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_05_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_07_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_08_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_10_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_12_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_moc.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0002.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0004.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0006.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0008.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0010.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_02_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_04_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_06_climo.nc
mavg.0001-0010.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_09_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_11_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_mocm.nc
tavg.0001-0010.nc

57+ Climatology Files are Created

Climatology Files Created by the AMWG, OMWG, Land and Ice Diagnostic Packages

ATM - SE

b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_07_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_01_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_JJA_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_09_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_03_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_SON_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_11_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_12_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_ANN_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_DJF_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_08_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_02_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_MAM_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_10_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_04_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_05_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_06_climo.nc

ICE

ice_vol_b.e12.B1850C5CN.ne30_g16.init.ch.027_1-10.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cice.h_ANN_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cice.h_jfm_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cice.h_amj_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cice.h_jas_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cice.h_ond_climo.nc

LND

b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_ANN_ALL.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_SON_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_MAM_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_JJA_means.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_ANN_means.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_DJF_means.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_JJA_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_ANN_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_DJF_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_SON_means.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_MAM_means.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_MONS_means.nc

OCN

Arc_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Hud_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Atl_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Ind_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Lab_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Gin_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Sou_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Glo_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Pac_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0001.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0003.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0005.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0007.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0009.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_01_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_03_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_05_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_07_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_08_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_10_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_12_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_moc.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0002.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0004.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0006.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0008.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0010.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_02_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_04_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_06_climo.nc
mavg.0001-0010.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_09_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_11_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_mocm.nc
tavg.0001-0010.nc

	Original
ATM-SE	11 min
ICE	5 min
LND	18 min
OCN	8+ hours
	Total
	8-9 hours (1 degree resolution climatologies over 10 years)

Climatology Files Created by the AMWG, OMWG, Land and Ice Diagnostic Packages

ATM - SE
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_07_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_01_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_JJA_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_09_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_03_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_SON_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_11_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_12_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_ANN_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_DJF_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_08_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_02_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_MAM_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_10_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_04_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_05_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cam.h0_06_climo.nc

ICE
ice_vol_b.e12.B1850C5CN.ne30_g16.init.ch.027_1-10.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cice.h_ANN_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cice.h_amj_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.cice.h_jas_climo.nc

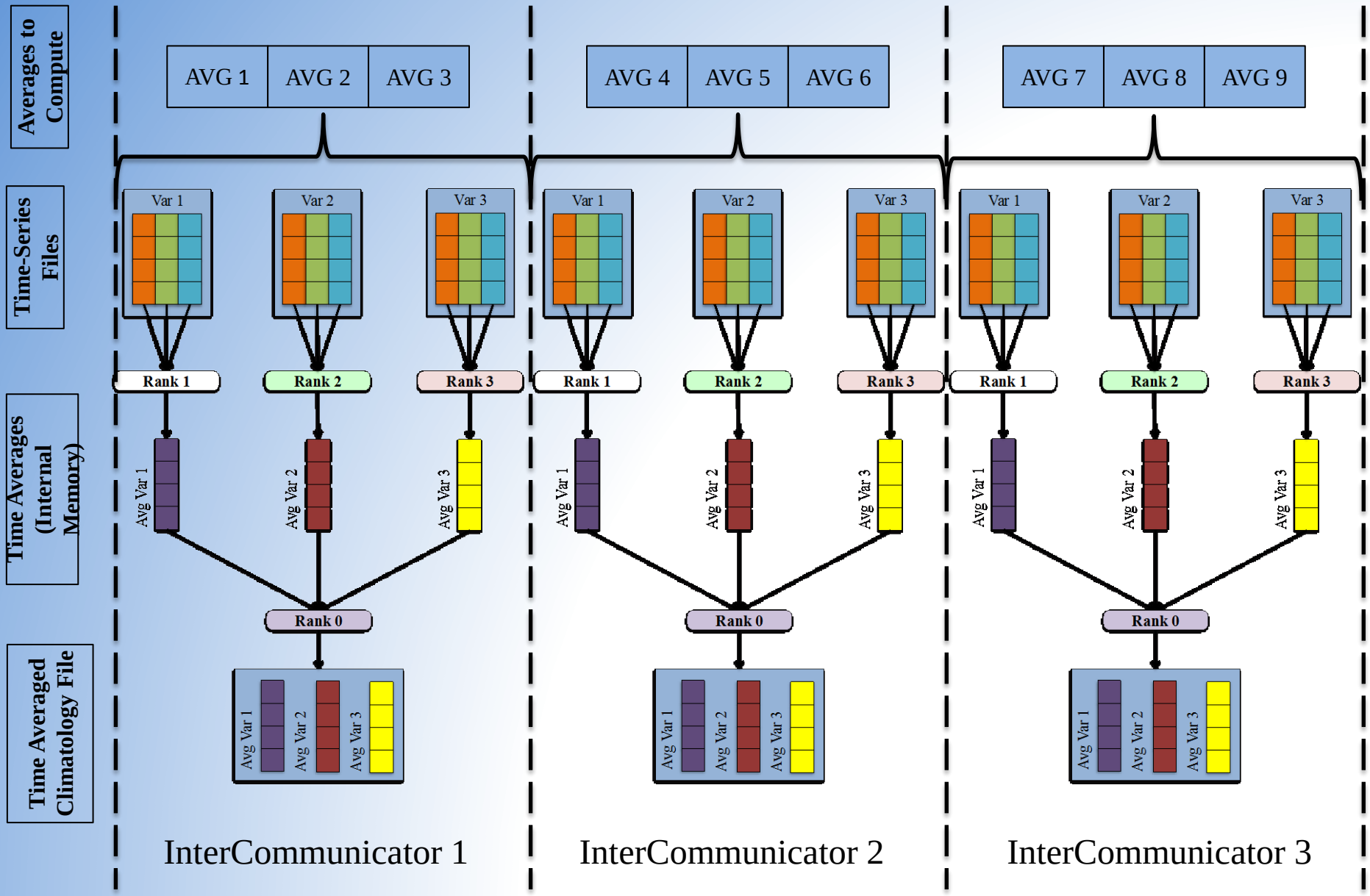
LND
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_JJA_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_SON_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_ANN_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_DJF_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_SON_means.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_MAM_means.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.clm2.h0_MONS_means.nc

OCN
Arc_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Hud_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Atl_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Ind_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Lab_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Gin_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Sou_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Glo_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
Pac_hor_mean_hor.meanConcat.b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_0001-0010.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0001.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0002.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0003.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0004.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0005.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0006.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0007.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0008.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0009.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h.0010.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_01_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_02_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_03_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_04_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_05_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_06_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_07_climo.nc
mavg.0001-0010.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_08_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_09_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_10_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_11_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_12_climo.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_mocm.nc
b.e12.B1850C5CN.ne30_g16.init.ch.027.pop.h_moc.nc
tavg.0001-0010.nc

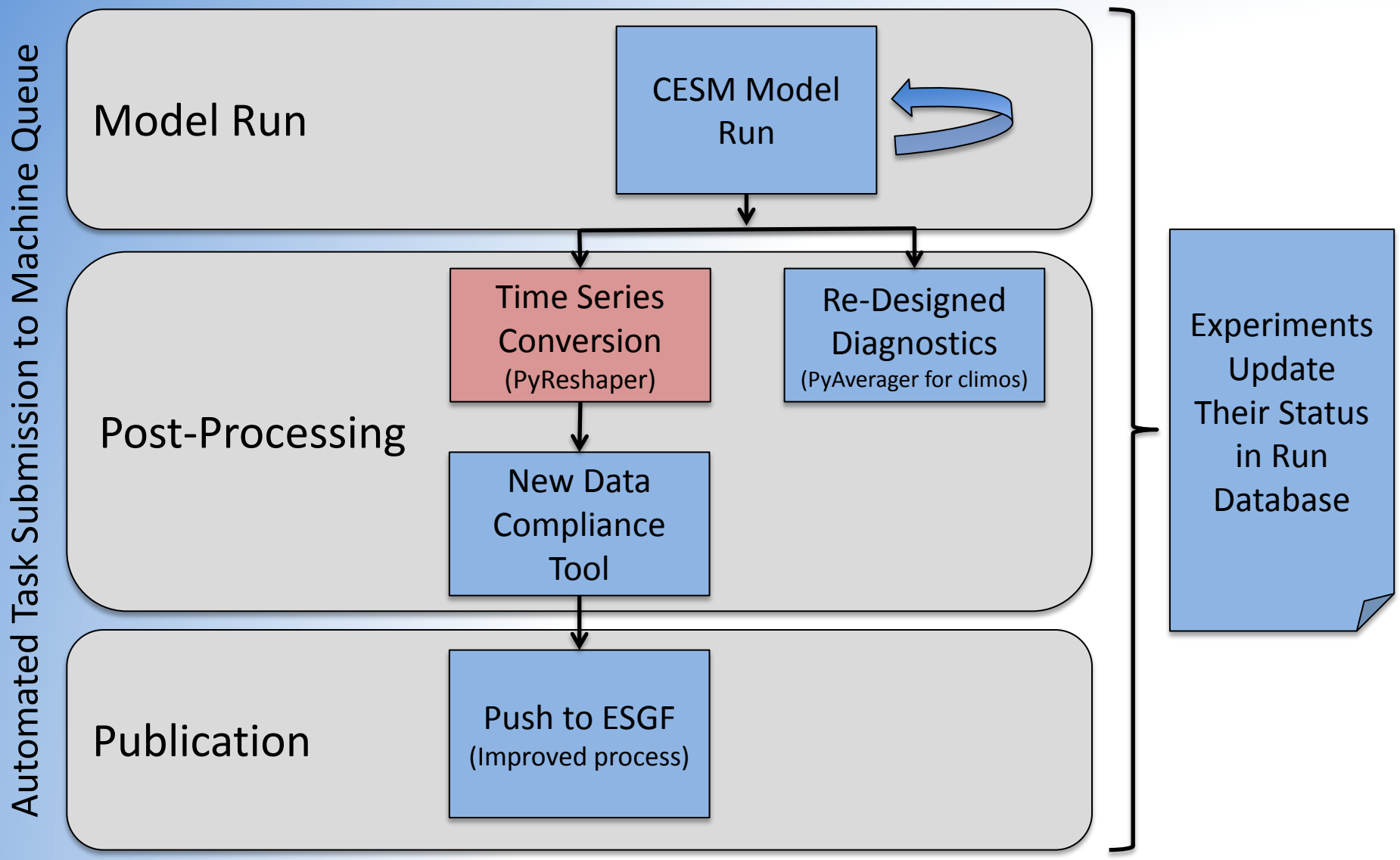
PyAverager increased performance by x100

	Original	PyAverager
ATM-SE	11 min	48 sec
ICE	5 min	11 sec
LND	18 min	30 sec
OCN	8+ hours	2 min
	Total	Total
	8-9 hours (1 degree resolution climatologies over 10 years)	Less than 4 min (1 degree resolution climatologies over 10 years)

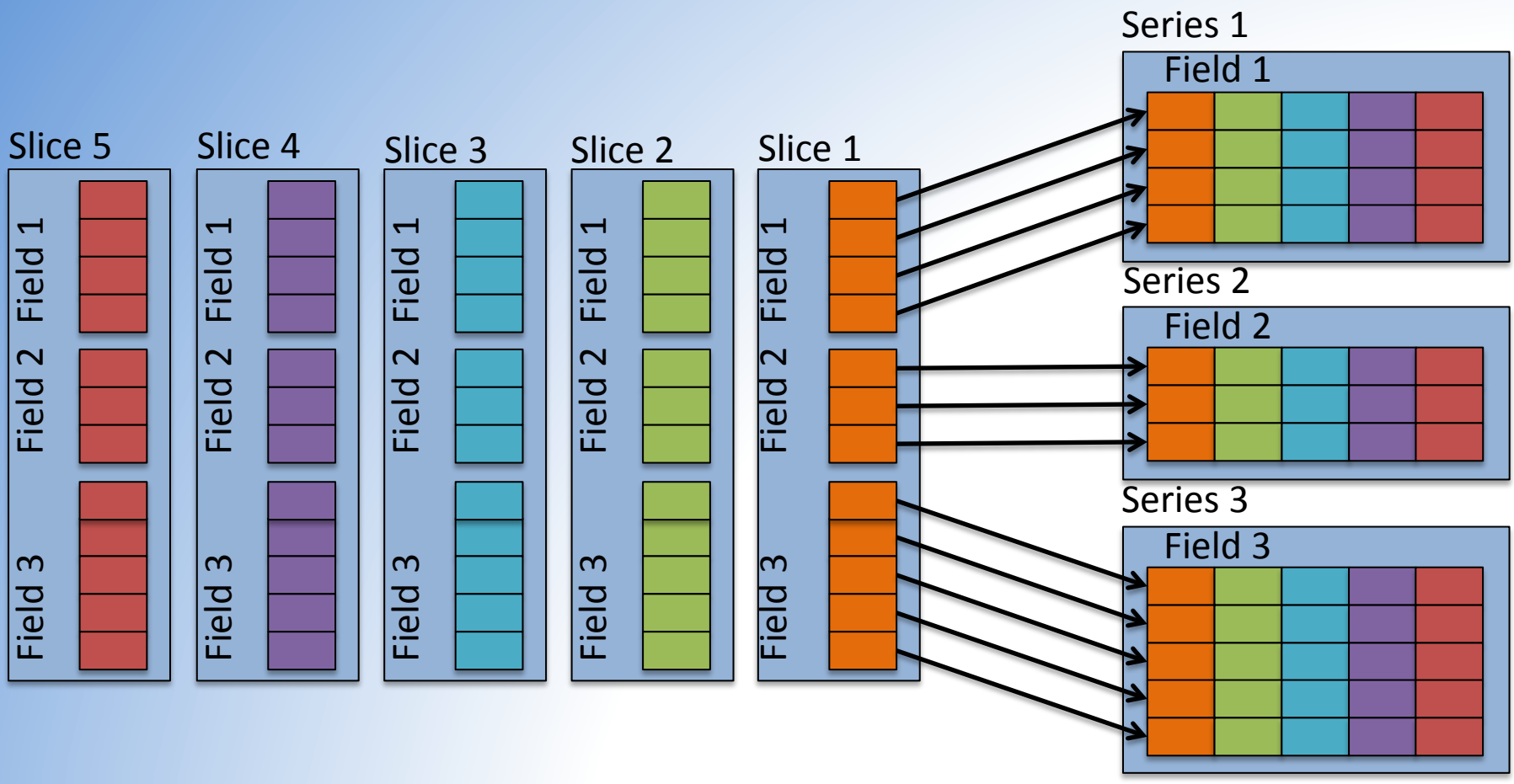
Partitioning of the PyAverager Tasks



New CESM/CMIP6 Workflow



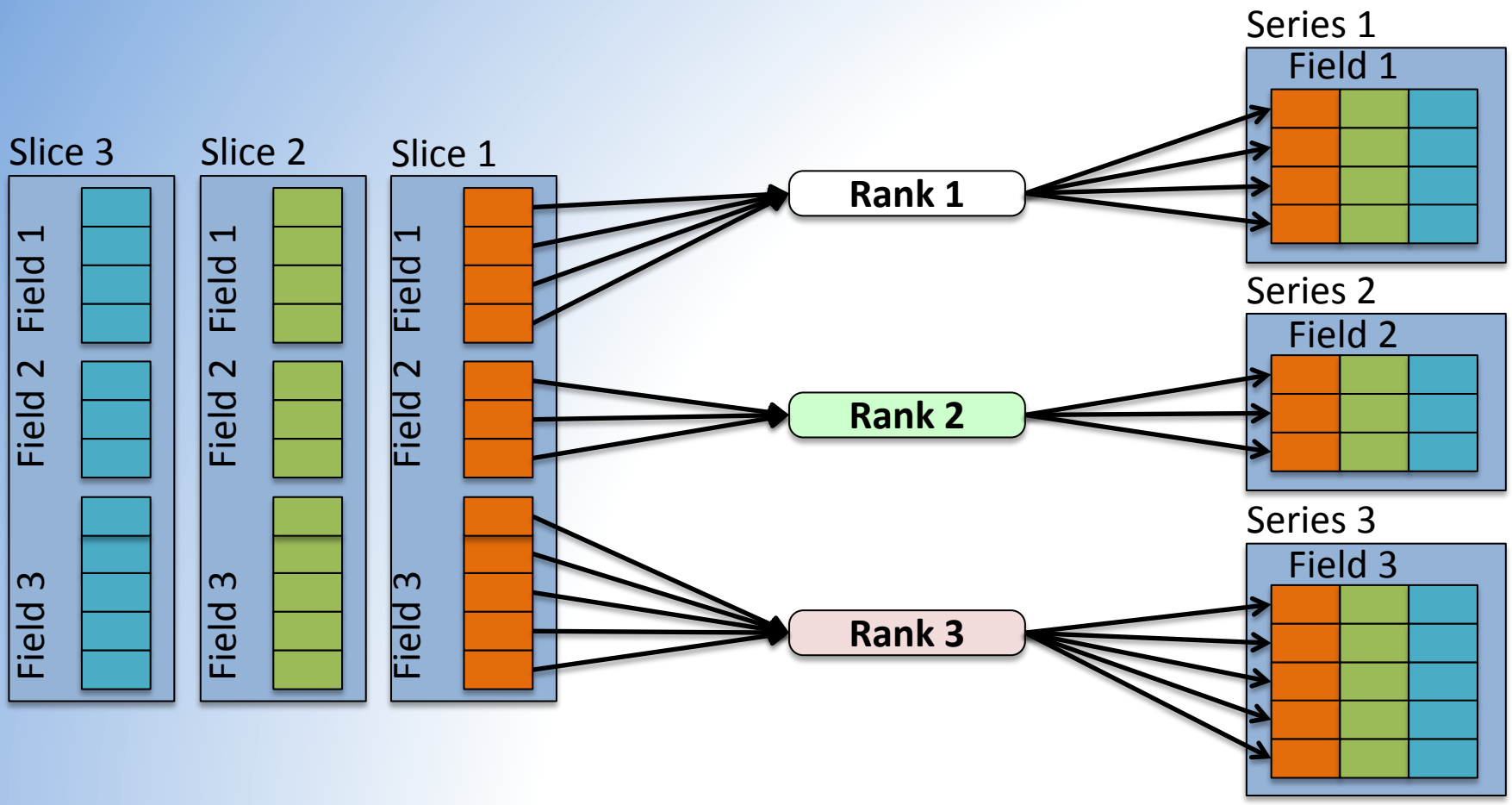
Time Slice to Time Series Conversion



- **This was one of the most expensive CMIP5 post-processing steps**
- The current post-processing suite works in serial using NCO

Task Parallelization Strategy
PyReshaper

Each rank is responsible for writing one (or more) time-series variables to a file



Time-Slice to Time-Series Conversion

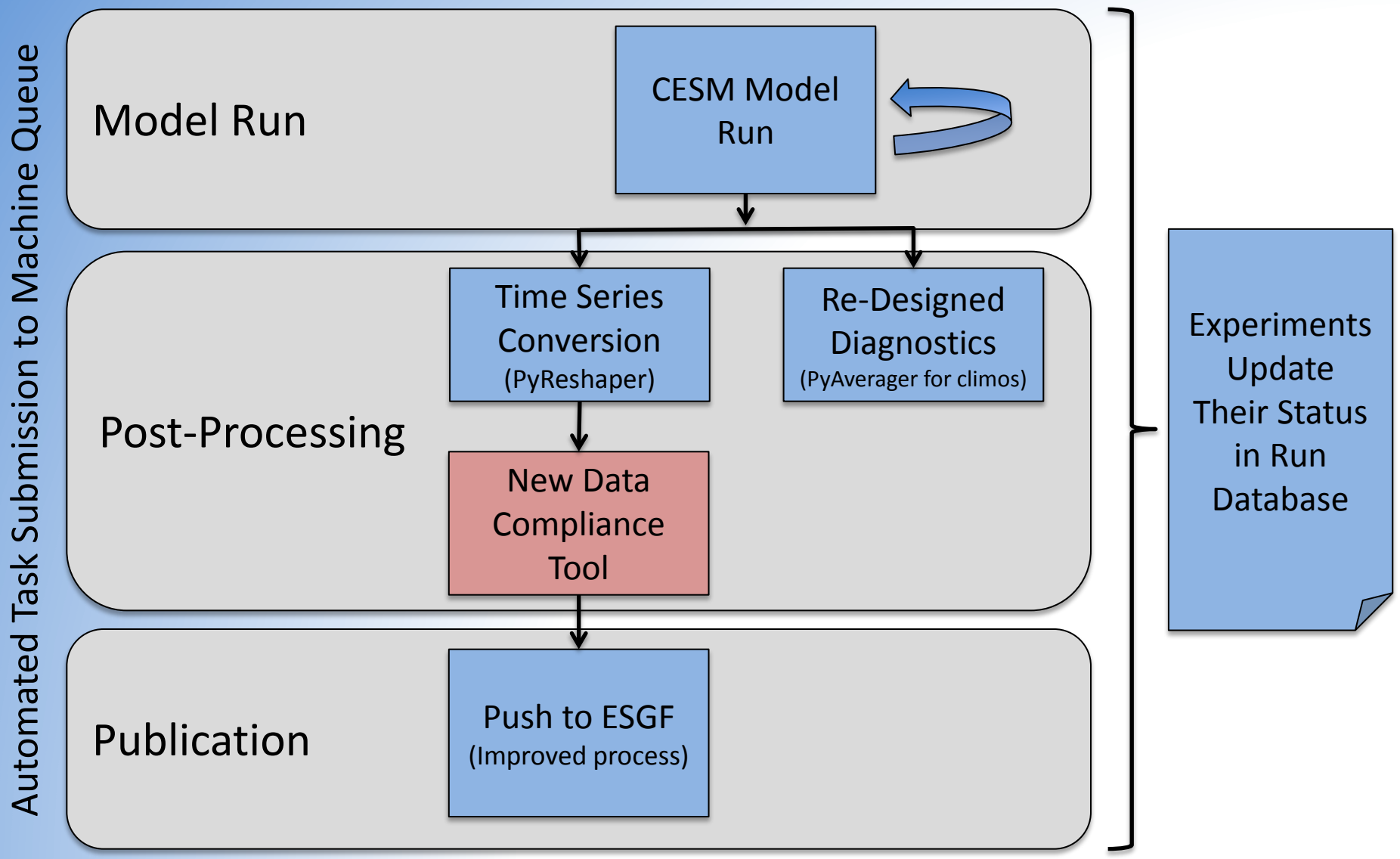
PyReshaper Timing Statistics

Existing Method (NCO)	Time (per MIP per Year)	Average Throughput (per run)
f09 x g16	225 minutes	1.85 MB/sec
ne120 x g16	478 minutes	4.85 MB/sec

New Method (PyReshaper)	Time (per MIP per Year)	Average Throughput (per run)
f09 x g16	4 minutes	104 MB/sec
ne120 x g16	8 minutes	290 MB/sec

- Times include the approximate full time to convert all component data to NetCDF4
- Conversions were ran on Yellowstone using 4 nodes/4 cores (16 cores total)
- We can expect a 2X increase in throughput if we double core counts for low-resolution data
- We can expect a 3X increase in throughput if we double core counts for high-resolution data

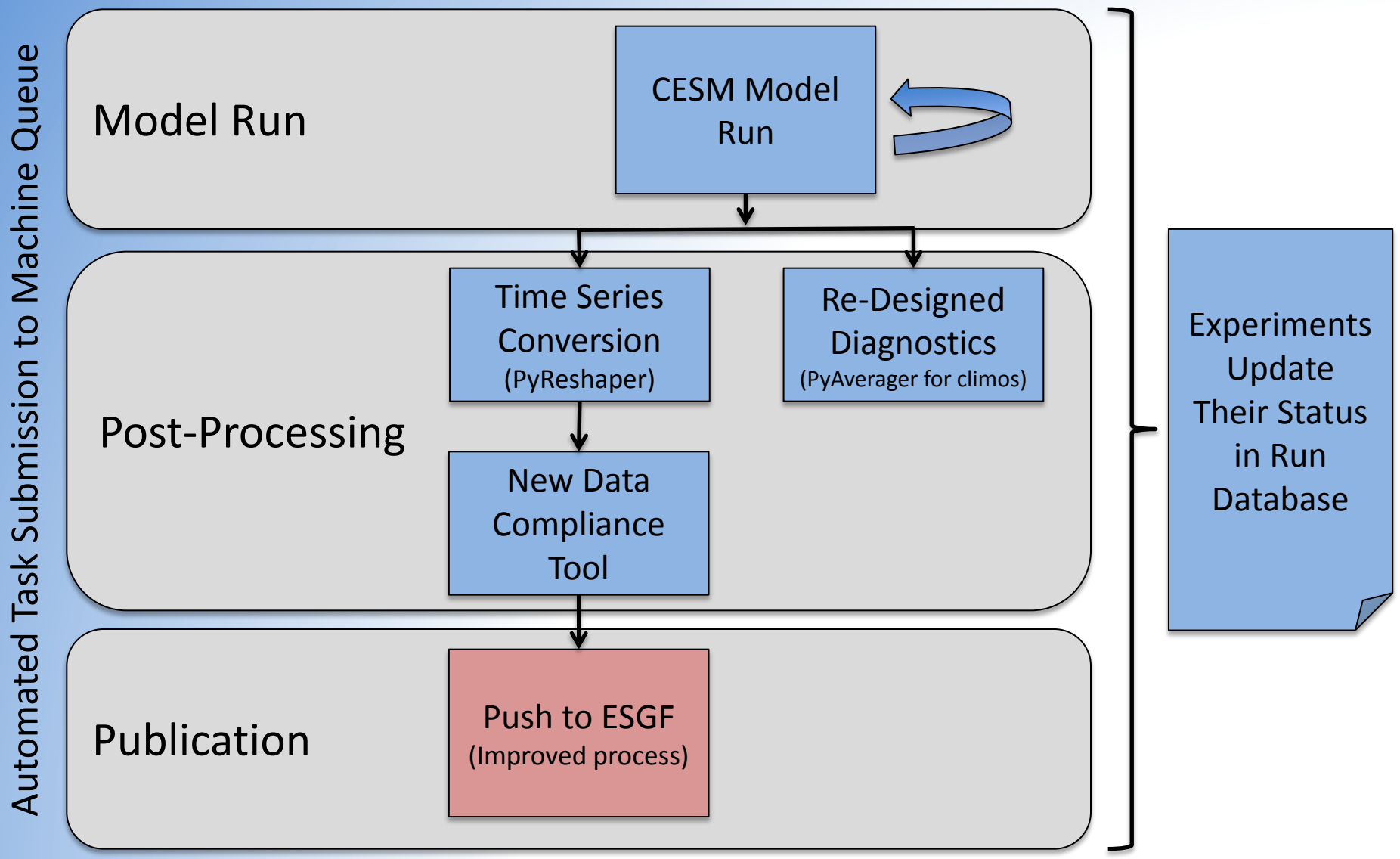
New CESM/CMIP6 Workflow



Data Compliance Tool

- Two Main Goals:
  - Simplify the use
  - Add parallelization to increase performance (this was another bottleneck in CMIP5)
- Use similar techniques that were used by the PyReshaper and PyAverager
- Prototyping work has been started with very promising results

New CESM/CMIP6 Workflow



ESGF Publication

- Move the data staging, directory structuring, and versioning responsibilities into the new compliance tool
- Streamline the submission process
- Test publication workflow for other ways we can improve the publication process

Tool Availability

- PyReshaper
  - https://github.com/NCAR-CISL-ASAP/PyReshaper
  - pip install PyReshaper
- PyAverager
  - https://github.com/NCAR-CISL-ASAP/pyAverager
  - pip install pyAverager
- ASAPPyTools
  - https://github.com/NCAR-CISL-ASAP/ASAPPyTools
  - pip install ASAPTools
- Automated Job Launch
  - cesm1_4_beta05/cime1.1.0
- Python Tools Website
  - https://www2.cisl.ucar.edu/tdd/asap/parallel-python-tools-post-processing-climate-data

Questions?

- PyReshaper
  - https://github.com/NCAR-CISL-ASAP/PyReshaper
  - pip install PyReshaper
- PyAverager
  - https://github.com/NCAR-CISL-ASAP/pyAverager
  - pip install pyAverager
- ASAPPyTools
  - https://github.com/NCAR-CISL-ASAP/ASAPPyTools
  - pip install ASAPTools
- Automated Job Launch
  - cesm1_4_beta05/cime1.1.0
- Python Tools Website
  - https://www2.cisl.ucar.edu/tdd/asap/parallel-python-tools-post-processing-climate-data