Astronomy & Iris at Edinburgh

Joe Zuntz

Team

- James Perry

- Joe Zuntz

- George Beckett

- Terry Sloan

- Bob Mann

Surveys

- Photometric surveys stack multiple wide-field images of the sky for broad astro uses - solar system to cosmology

- Ongoing programme of surveys

    - Current generation:
      DES, KiDS, HSC

    - Next generation:
      Rubin, Euclid, Roman

- Big Data Challenges:
  approaching exabyte size

- Simulation Challenge

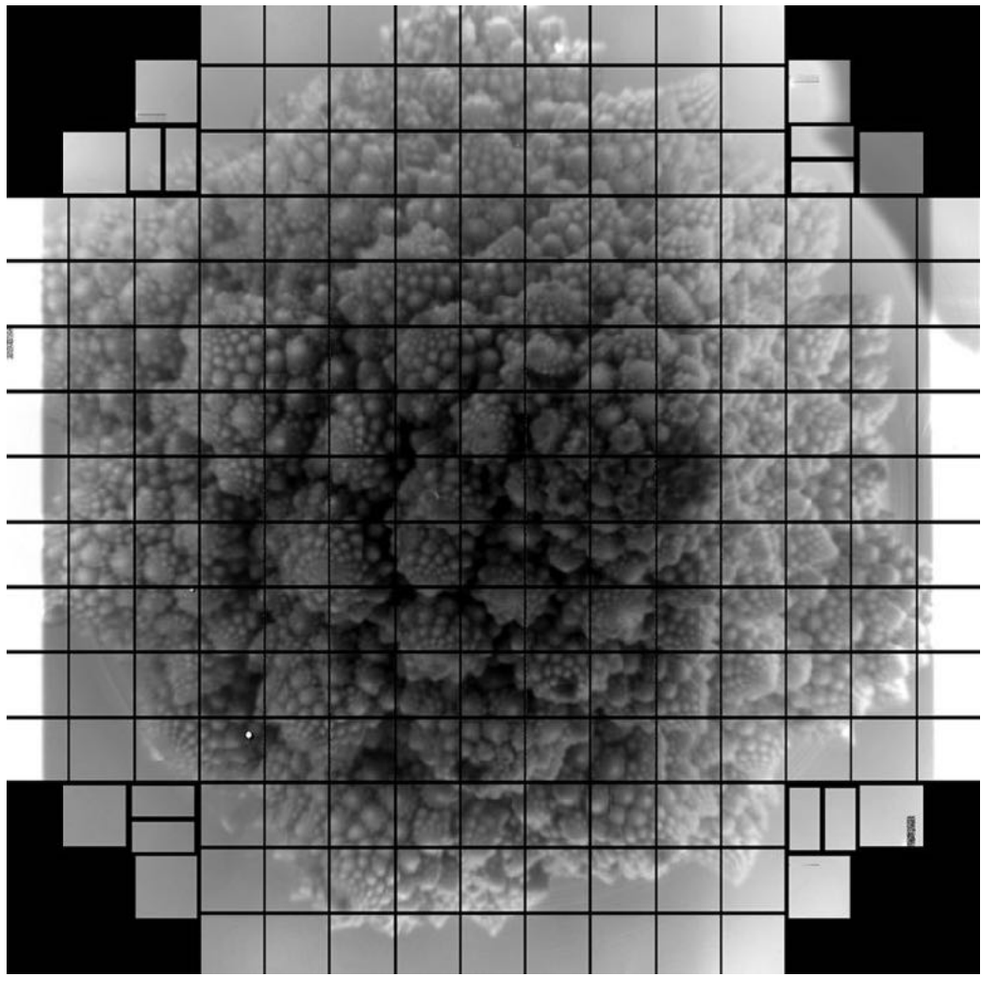
Rubin Observatory

- 10 year main survey
  37B sources

- Could image the entire sky
  every 3 days

- Camera complete and used
  to image some broccoli

- Strong UK involvement
  100 PIs
  400 PDRAs/PhDs



- PPRP Infrastructure

- First science light 2023?

Rubin & IRIS

- Many potential HPC/HTC science activities from Rubin

- Focus here on activities in LSST:UK Work Package 3.10

    - Operational support for LSST Cosmology (DESC)

- Current phase: supporting simulation, processing, and analysis of *Data Challenge 2* images

    - Designed to test LSST pipelines & science analysis

Flow

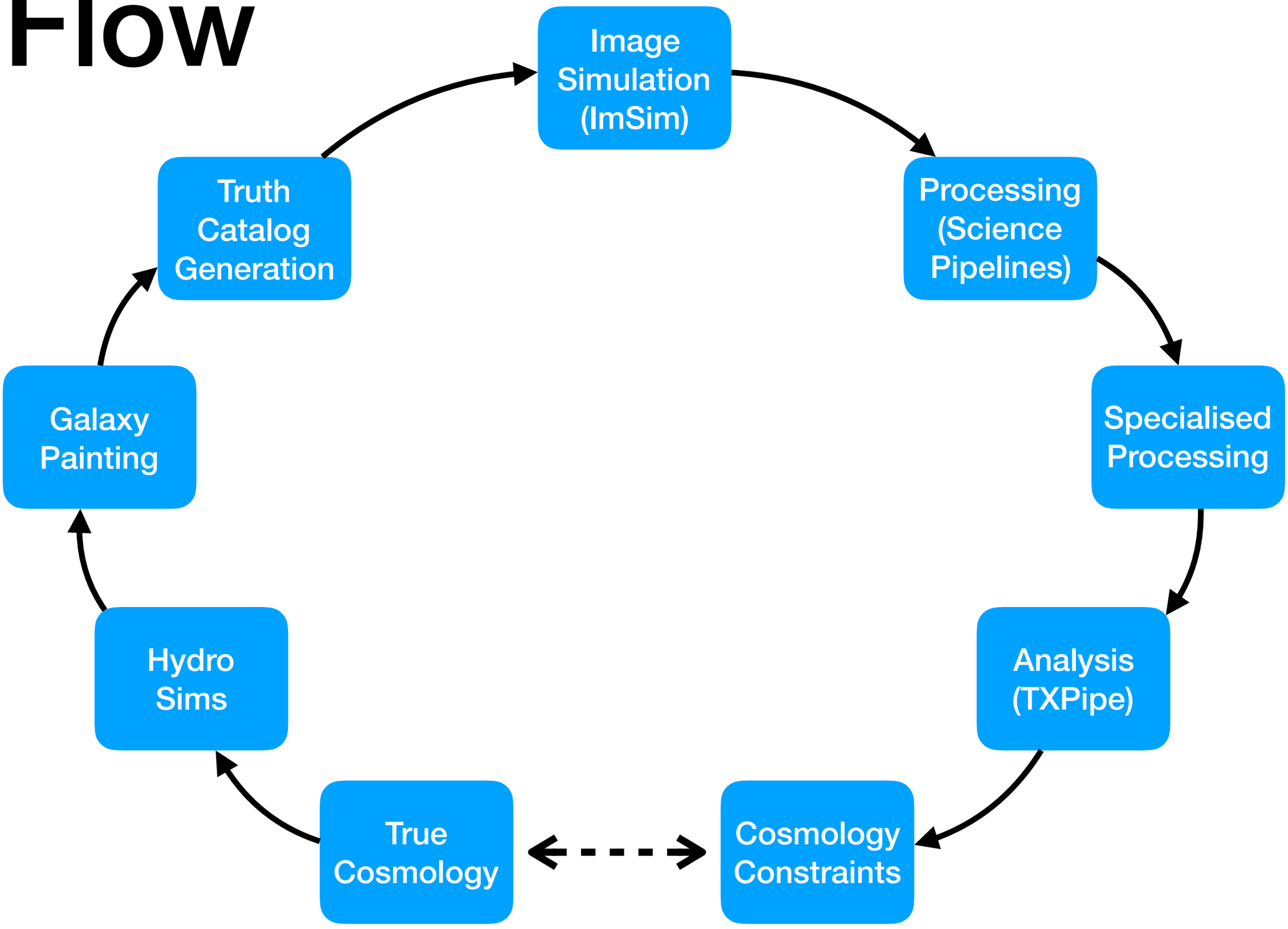


Image Simulation (ImSim)
Truth Catalog Generation
Processing (Science Pipelines)
Galaxy Painting
Specialised Processing
Hydro Sims
Analysis (TXPipe)
True Cosmology
Cosmology Constraints

Flow

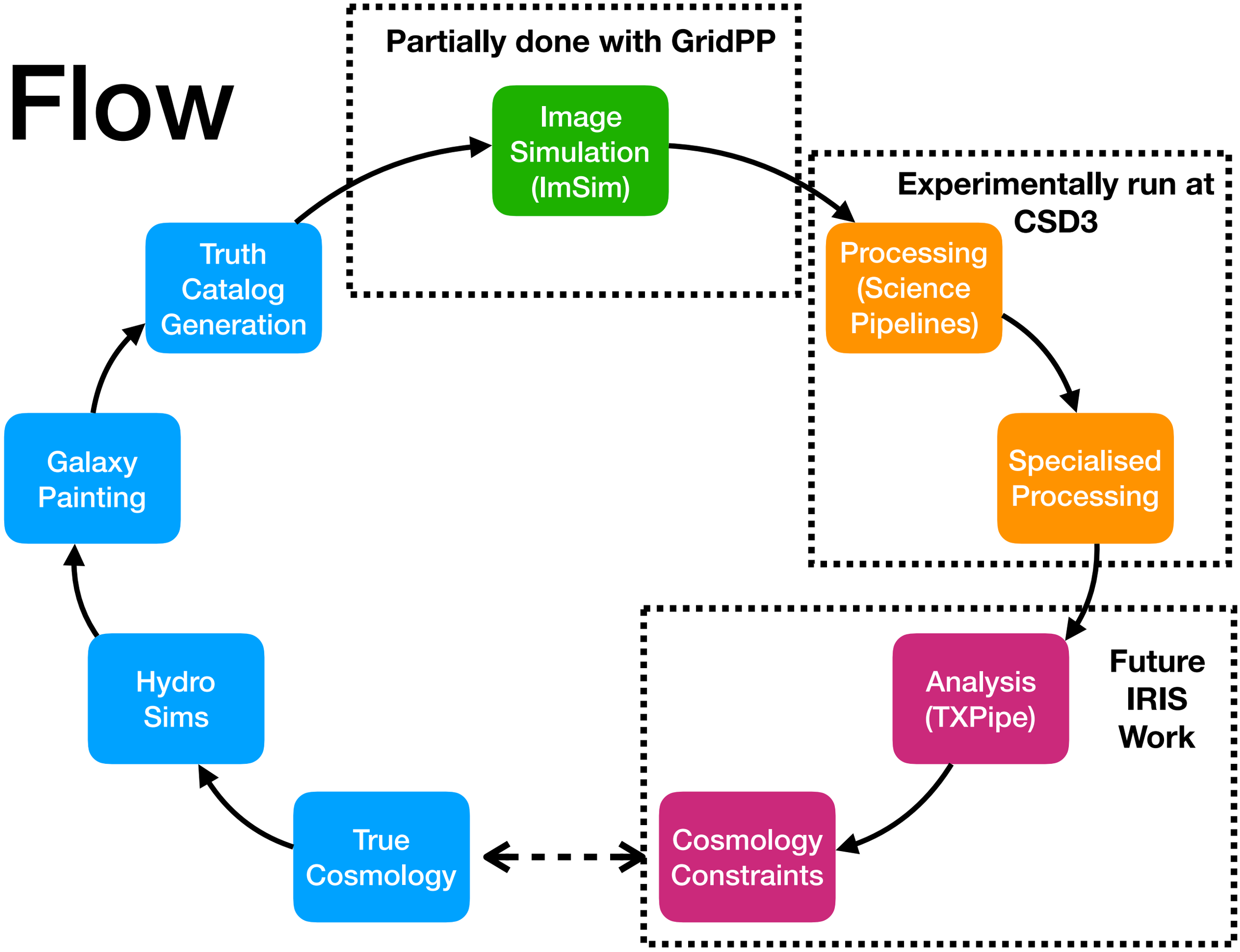

Partially done with GridPP
Image Simulation (ImSim)
Truth Catalog Generation
Galaxy Painting
Hydro Sims
True Cosmology
Experimentally run at CSD3
Processing (Science Pipelines)
Specialised Processing
Analysis (TXPipe)
Cosmology Constraints
Future IRIS Work

Image Simulation

- Used ImSim code to turn input truth catalogs to noisy images
- Divided tasks by time and sky region
- Ran 2/7 years of simulations on GridPP
- 650k individual jobs, took several months total
- Several iterations due to errors in both runs and inputs

Image Simulation Challenges

- Great deal of compute and support available: powerful resource

- Inhomogeneity
    - Couldn’t rely on data accessibility, s/w usability, OS, cores/node

- Unsuitable assumptions
    - Watchdogs took CPU-light periods as sign of dead job

- Extensive hand-holding
    - JP had to monitor this daily

Image Processing

- LSST Project builds the Science Pipelines
    - Image calibration, processing, combination, and object detection and modelling
- A more standard HPC job
    - Run on CSD3 - seemed to work nicely
- Parsl workflow tool - useful for connecting stages.
    - Monitoring occasionally killed by login node watchdog
- Docker/singularity
    - Some issues with MPI mismatches

Future: Science Analysis

- Running bespoke code on catalog output of previous stages
- Generate summary stats and science constraints
- Experimenting with *Ceci* - interfacing science to workflow tools, so we can mix and match
- Adding provenance, standardised I/O, parallelism tools
- Currently run at NERSC, expect easy port to CSD3 and highly useful resources in future