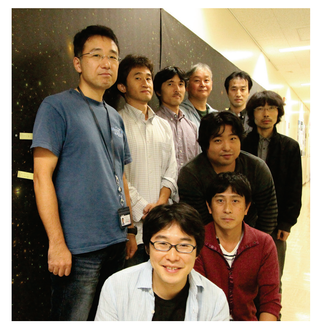
Data Processing in Subaru Strategic Survey Program with Hyper Suprime-Cam

* Email: furusawa.hisanori@nao.ac.jp

Hisanori Furusawa*, Michitaro Koike, Sogo Mineo, Yoshihiko Yamada, Hiroyuki Ikeda, Satoshi Takita, Yuki Okura, Masayuki Tanaka, Tadafumi Takata (NAOJ), & HSC DR Team



HSC Subaru Strategic Program (HSC-SSP)

The Subaru Strategic Program (SSP) with Hyper Suprime-Cam (HSC) is an ongoing wide-field imaging survey project at Subaru Telescope, devoting 300 nights to cover 1400 sq.deg fields around the equator in 5 optical bands (and a few NBs in 4 selected deep fields). The survey is motivated by various scientific interests including weak-lensing cosmology and galaxy sciences. We have made 2 public data releases to date. The products are available at: **https://hsc-release.mtk.nao.ac.jp/**

Concept of Data Production

The project aims at providing well-calibrated science-ready products to the astronomical community. In order to attain uniform products across the entire survey footprint, a workflow of observation and data production has been designed so that exposure planning and data production are performed based on measured quality of raw data (Figs. 1 & 2). In this scheme, the production cycle involves a quick quality assurance at the observing site in Hawaii, followed by full production run conducted in Japan. The onsite QA system (Furusawa et al. 2018) extracts raw data quality and ingest it to a QA database, which is used to determine exposures to be combined into the final products. A single production (1-2 times/year) generates >500TB data per release. The products are transferred to the archive system and loaded to a database for external access. The pipelines are run across PC clusters so that the whole processing could be done within a ~2 months time scale for every data release.

HSC Data Production Flow

Fig. 1 The onsite QA is performed soon after an exposure is obtained. The production process refers to the QA database.

Processing & Archive Systems

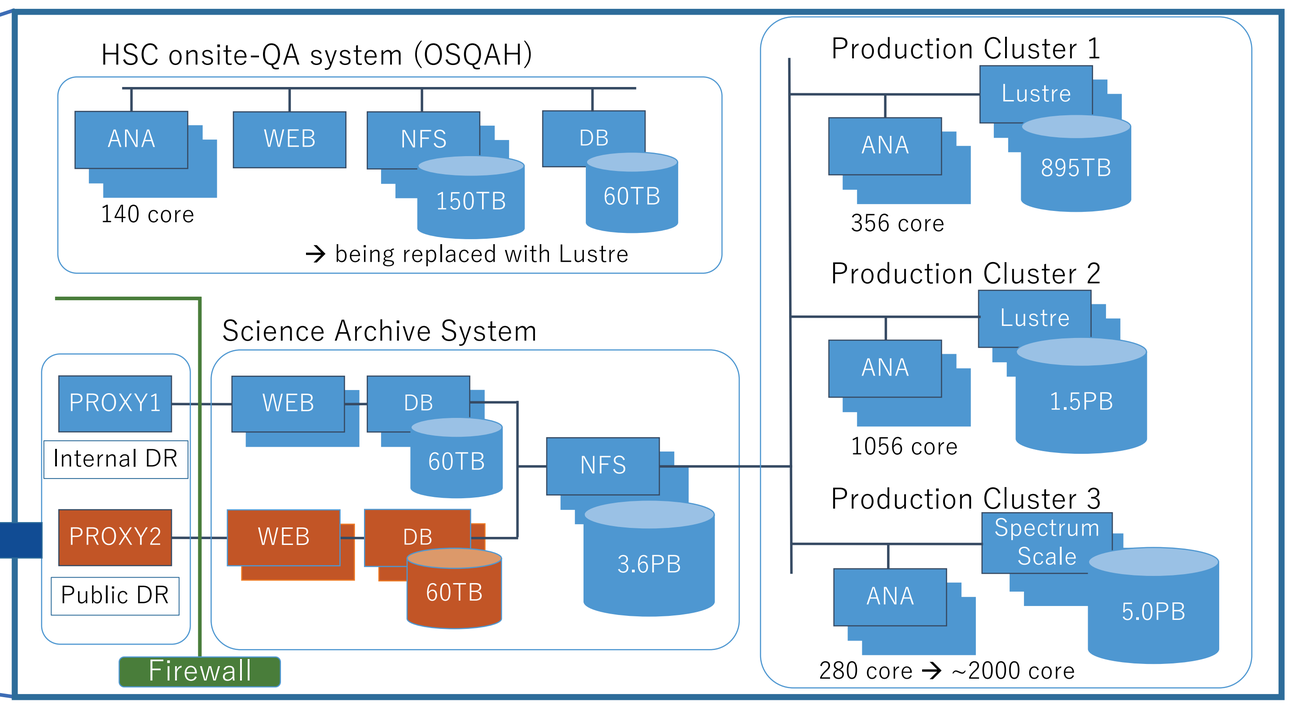


Fig. 2 The system components constitute OSQAH, archive, and clusters.

HSC Data Archive Services

HSC-SSP PDR: https://hsc-release.mtk.nao.ac.jp (Fig. 3)

- Catalog Archive Server (CAS)
  - Ruby on Rails for SQL execution and job scheduler
  - PostgreSQL 10.9 DBMS + distributed DB extension
    - + new ongoing development (see Takata+ P11.11)
  - Apache Python module modified for access control
  - Schema browser
  - Python APIs for connecting DB services
- Data Archive Server
  - Image search, download, and cutout services
- Interactive Viewer (hscMap)
  - Web-client application for viewing pseudo-color images and overlaying catalog sources and their measurements

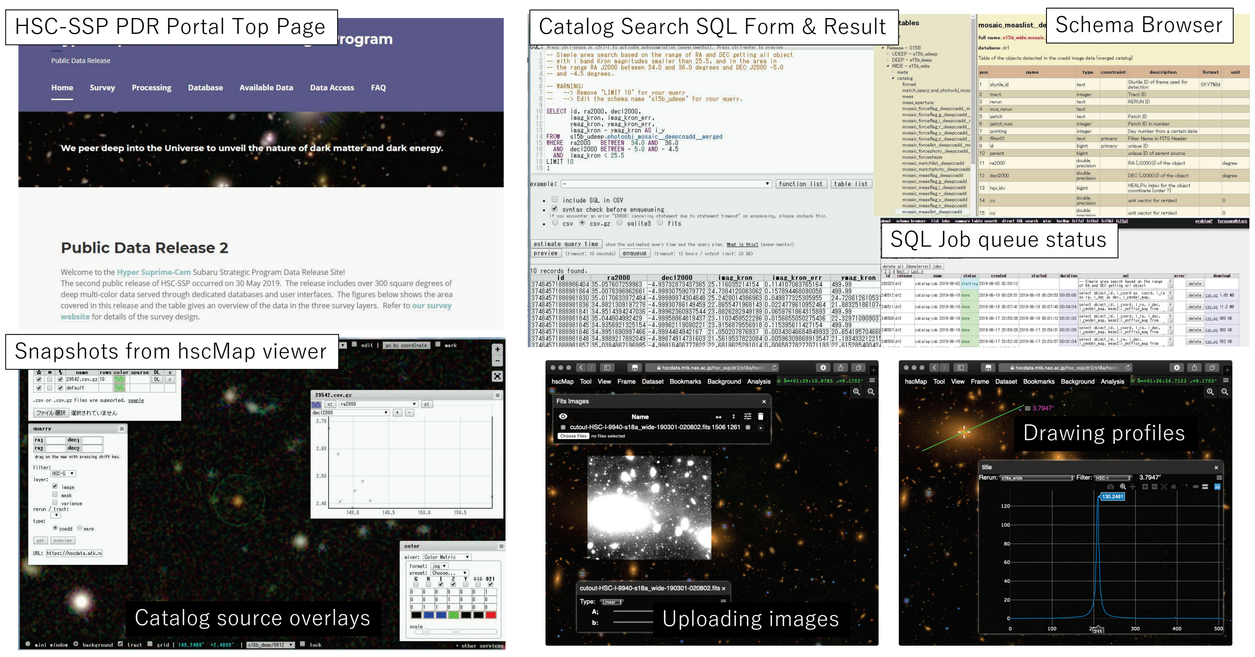


Fig. 3 Snapshots from the HSC data release service.

HSC Data Analysis Pipeline

We use hscPipe (a compatible version of the LSST stack) software for processing HSC-SSP data (Bosch et al. 2018). The pipeline features are being updated and enhanced in the collaboration framework, including source detection with deblender, sky subtraction, and chromatic photometric calibration. Fig. 4 shows a trial of removing ghosts and satellite trails on coadd images, in which ghost positions are predicted by the optical design, and Hough transform is used to detect trail pixels. We plan to update the pipeline to incorporate those new functions, and improve the products towards constructing a final state of the data release.

References: Furusawa, H., et al. 2018, PASJ, 70, S3; Takata, T. et al. 2019 P11.11; Bosch, J., et al. 2018, PASJ, 70, S5

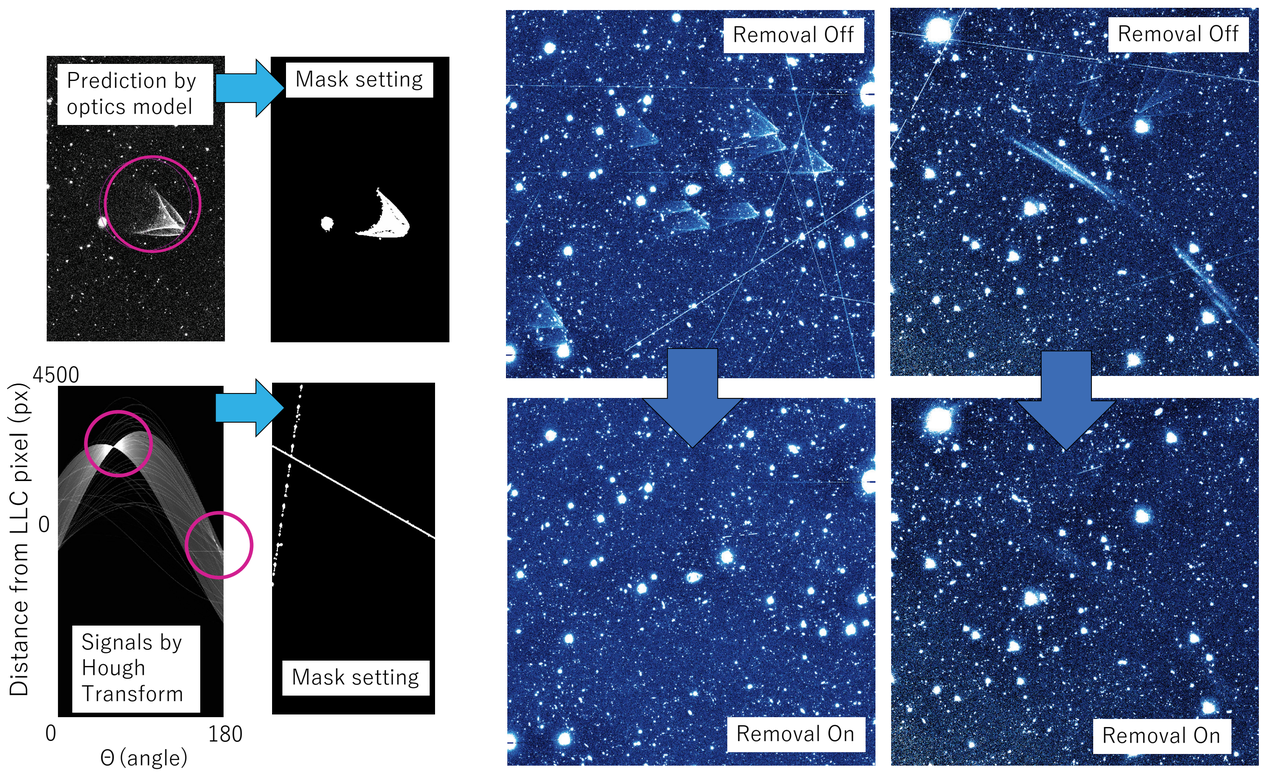


Fig. 4 A preliminary result of ghost & satellite trails algorithms.