Astronomical Data Analysis Software and Systems VIII
ASP Conference Series, Vol. 172, 1999
D. M. Mehringer, R. L. Plante, and D. A. Roberts, eds.

DASH – Distributed Analysis System Hierarchy

Masafumi Yagi, Yoshihiko Mizumoto, Eiji Nishihara, Michitoshi Yoshida, Ryo Asai, Yoshihiro Chikada, George Kosugi[1], Tadafumi Takata[1], Ryusuke Ogasawara[1]

National Astronomical Observatory, Osawa, Mitaka, Tokyo 181-8588, Japan

Yasuhide Ishihara, Hiroshi Yanaka

Fujitsu Limited, Nakase, Mihama, Chiba 261-8588, Japan

Yasuhiro Morita, Hiroyuki Nakamoto

SEC Co. Ltd., 22-14 Sakuragaoka, Shibuya, Tokyo 150-0031, Japan

Abstract. DASH is a distributed data reduction and analysis system constructed on CORBA. It plays an important role in the SUBARU computer system not only as an offline analysis system but also as a test and development environment of online quick analysis. We abstracted an object from astronomical data reduction/analysis procedures; 'PROCube' (data reduction PROcedure Cube) and implemented it, so that we could share and reuse the reduction procedure. We are also trying to make it possible to retrieve sufficient original data for an analysis semi-automatically in cooperation with SUBARU data archive system (STARS).

1. Introduction

DASH is an object-oriented data reduction and analysis system for 8.2-m SUBARU telescope at Hawaii. In early 1999, SUBARU will obtain scientific first light image and soon it will begin to produce several tens TB data per year. DASH is designed as the SUBARU observatory system to manage, reduce and analyze such a huge amount of astronomical data produced by various instruments. In addition, reproducibility and portability of the reduction, seamless operability on heterogeneous computer system, open architecture, trial-and-error processing, and pipeline processing, and research target oriented analysis are required. To satisfy these requirements, we adopt CORBA as a distributed object environment and construct a distributed data reduction and analysis system (Takata et al. 1998). We have made two prototypes in 1997 (Mizumoto et al. 1998a, 1998b), and we have just tested Ptorotype-3 at SUBARU Observatory in Nov.

[1]Subaru Telescope, National Astronomical Observatory of Japan, Hilo, HI 96720, USA

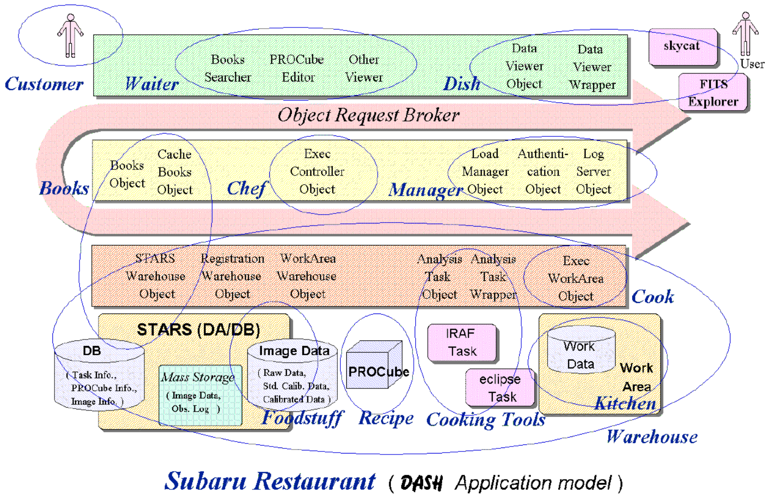

Figure 1. DASH application model.

1998. DASH project would be completed by early 2000, and we have achieved just as we had scheduled so far (Mizumoto et al. 1998b).

2. Components of DASH: SUBARU Restaurant

DASH system was developed with Object Oriented Methodology. When DASH was designed, the system was modeled as a restaurant, with three tiers: UIs, controls, and data. The distributed objects are designed based on the model and the role of each object is understood through this metaphor (Mizumoto et al. 1998b).

3. The trinity of SUBARU software system: the role of DASH

There are three subsystems in SUBARU software system; Subaru Observation Software System (SOSS; Sasaki et al. 1998) at the summit of Mauna Kea, Subaru Telescope data ARchive System (STARS; Takata et al. 1998) with quick-look image system for archive search (Baba et al. 1999) at Hilo base, and DASH.

Before one begins to analyze data in DASH, he or she will search in STARS and retrieve raw data into DASH. With an observation dataset, the retrieval of relevant calibration data can be done automatically. The data are reduced and analyzed in DASH to obtain scientific results. Then, reduced and standard calibration data are registered into archives.

The feedback from DASH to SOSS is two-fold. As we observe with SUBARU, the astronomical results of DASH are directly reflected to the next observation plan. For example, if some find that more DARK frames are needed for precise analysis, the information is reflected to the auto-scheduling system in the SOSS.

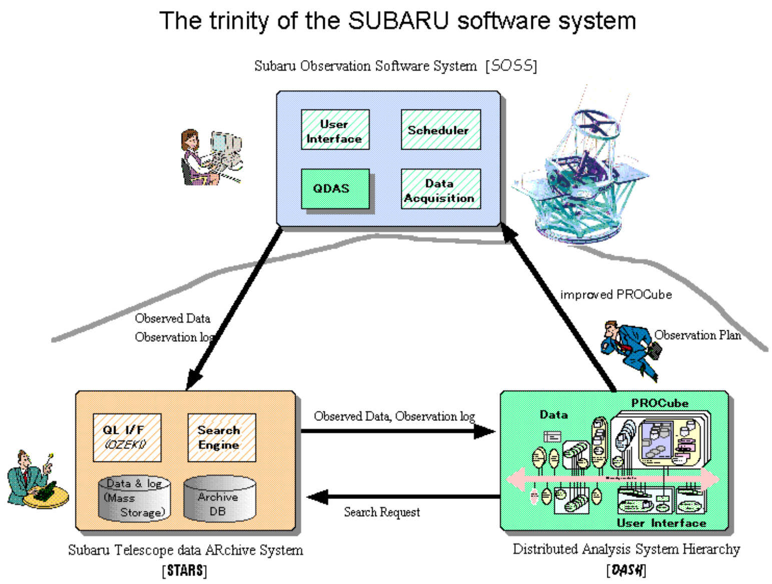

Figure 2. SUBARU software system overview.

The second feedback is improvement of quick data analysis system (QDAS) for quick look. We can use the same PROCube in the quick data analysis system in SOSS. There is a small subset of DASH in the observational system for data processing as a pipeline. An improvement of PROCube in DASH is reflected to SOSS and observer can get quick and precise result of the target at summit during observation. The three subsystems, SOSS, STARS, and DASH thus form a trinity.

4. PROCube: data reduction PROCedure Cube

PROCube is an object that holds all information about data processing (Mizumoto et al. 1998b). It contains data, engines, flow of reduction and analysis, tunable parameters for each engine, a log of reduction, and the current status.

Making a new PROCube is similar to drawing a flowchart of a reduction; each engine has several inputs, outputs and some parameters, which must be provided by the user. In PROCube Editor, this process is visualized; the user can cut & paste the names of data to connect the engines, set parameters for engines, and execute it. Intermediate files are automatically created and deleted when they are no longer needed. As PROCube holds all information about the reduction, the results are always reproducible.

Once a PROCube is created, it could be reused for reduction of another data by simply changing the original data. QDAS in SOSS already uses this PROCube pipeline, using data from DAQ system. In the next prototype, we try to fill the calibration data automatically with observational datasets (Kosugi et al. 1998).

5. SASH: Standalone Analysis System Hierarchy

DASH is designed as a part of SUBARU software system at SUBARU Observatory, Hawaii, however, people may desire to use DASH system at their institute (e.g., in Japan) for their analysis. In the future, we will be able to connect CORBA directly from Japan to Hawaii. Currently, because of rather poor internet connection, objects cannot be directly distributed. Our future plans is to extract some part of DASH as a stand-alone system without CORBA (SASH) so that the same interfaces and the same PROCubes as in DASH could be used in SASH.

Visiting astronomers to SUBARU will analyze their data in DASH at Hawaii for a few days after observation. When they leave, they can detach their environment, bring it along with their data, and continue the analysis on SASH at their home institute. In future, these satellite SASHs could be re-attached on DASH, or form satellite DASH systems connected to the original.

6. Results of Prototype-3 and the plan for the next prototype

With DASH prototype-3, we can reduce and analyze test data of some test instruments semi automatically with PROCubes. User authentication, data and engines sharing in the same proposal group are implemented and tested.

The main targets of next prototype are the stand-alone system (SASH), and the observational dataset. Another important targets are consideration for engines with GUI and development of algorithm for load dispensing. In Prototype-3, only non-interactive engines could be used in PROCubes, though such interactive engines are necessary for some analysis. Only data transfer load is considered in prototype-3, since the data transfer is one of the dominant factors of the system load. In next prototype we will monitor CPU load in some test system, and will change the algorithm accordingly. This DASH prototype-4 will be used and tested with the real data from SUBARU next year.

References

Baba, H., Yasuda, N., Ichikawa, S., Watanabe, M., Horaguchi, T., Hamabe, M., Aoki, W., & Ozawa, T. 1999, this volume, 413

Kosugi, G., et al. 1998, in SPIE Proc., Vol. 3349, Observatory Operations to Optimize Scientific Return, ed. P. J. Quinn (Bellingham: SPIE), 421

Mizumoto, Y., et al. 1998a, in ASP Conf. Ser., Vol. 145, Astronomical Data Analysis Software and Systems VII, ed. R. Albrecht, R. N. Hook, & H. A. Bushouse (San Francisco: ASP), 332

Mizumoto, Y., et al. 1998b, in SPIE Proc., Vol. 3349, Observatory Operations to Optimize Scientific Return, ed. P. J. Quinn (Bellingham: SPIE), 173

Sasaki, T., et al. 1998, in SPIE Proc., Vol. 3349, Observatory Operations to Optimize Scientific Return, ed. P. J. Quinn (Bellingham: SPIE), 427

Takata, T., et al. 1998, in SPIE Proc., Vol. 3349, Observatory Operations to Optimize Scientific Return, ed. P. J. Quinn (Bellingham: SPIE), 247