

ESO scalable architecture for operational databases

I. Vera, A. Dobrzycki, A. M. Chavan, P. Nass, J. S. Lockhart

European Southern Observatory,
Karl-Schwarzschild-Str 2,
D-85748 Garching, Germany

ABSTRACT

The European Organisation for Astronomical Research in the Southern Hemisphere (ESO), headquartered in Garching, Germany, operates different state-of-the-art observing sites in Chile. To manage observatory operations and observation transfer, ESO developed an end-to-end Data Flow System, from Phase I proposal preparation to the final archiving of quality-controlled science, calibration and engineering data. All information pertinent to the data flow is stored in the central databases at ESO headquarters and replicated to and from the observatory database servers.

In the ESO's data flow model one can distinguish two groups of databases; the front-end databases, which are replicated from the ESO headquarters to the observing sites, and the back-end databases, where replication is directed from the observations to the headquarters.

A part of the front-end database contains the Observation Blocks (OBs), which are sequences of operations necessary to perform an observation, such as instrument setting, target, filter and/or grism ID, exposure time, etc. Observatory operations rely on fast access to the OB database and quick recovery strategies in case of a database outage.

After several years of operations, those databases have grown considerably. There was a necessity in reviewing the database architecture to find a solution that support scalability of the operational databases.

We present the newly developed concept of distributing the OBs between two databases, containing operational and historical information. We present the architectural design in which OBs in operational databases will be archived periodically at ESO headquarters. This will remedy the scalability problems and keep the size of the operational databases small. The historical databases will only exist in the headquarters, for archiving purposes.

Keywords: Observation preparation, database scalability, database replication, European Southern Observatory

1. INTRODUCTION

The ESO Data Flow System (DFS) [4] is an advanced computer system which provides a unified frame for observation data transfer and processing, from Phase I proposal preparation to the final archiving of quality-controlled science and engineering data. The system allows tracking of all the observations made at ESO telescopes from the proposal submission to the final data product.

Due to the necessity at the observing sites to have quick and reliable access to the data, observing sites are provided with their own local database servers and data are replicated from and to the observing sites. Databases at ESO are divided in two groups depending in the direction databases are replicated:

- The front-end databases [3], which contain information from Phase I proposal Submission and Phase 2 proposal preparation. They are replicated from the headquarters to the observing sites.
- The back-end databases which contains the information of all observations carried out by ESO telescopes. They are replicated from the observing sites to the headquarters.

A critical part of the front-end databases is the observing blocks (OBs) repository. An OB is the fundamental unit that can be scheduled or executed in an ESO telescope. It contains the sequence of operations necessary to perform an observation. They are created using the Phase 2 Proposal Preparation (P2PP) system developed at ESO [1]. The system

allows to create OBs and submit them into ESO OB repositories.

Telescopes at ESO are operated in two modes [2]: Visitor Mode (VM), where the investigator is personally present at the telescope and directing the observations; and Service Mode (SM), where observations are performed by ESO staff. Service and Visitor Mode observations follow a predetermined sequence of operations expressed in terms of Observation Blocks.

The data flow for OBs in SM differs from VM [6]. The different data flow for the different operation modes is shown in Figure 1. SM observations are initially scheduled from Garching and therefore they need to be submitted to ESO headquarters. SM OBs are transferred to the observing sites for execution using database replication. VM observations enjoy greater flexibility as they don't need to be scheduled and they are supported directly by the staff at the observing sites.

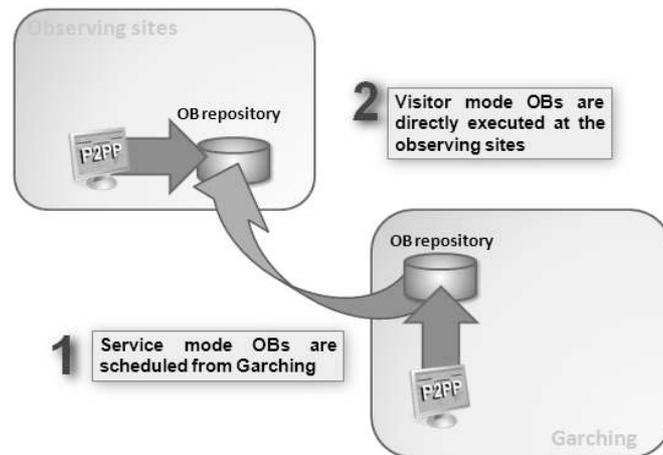


Figure 1. Old data flow architecture for observations in Service mode and Visitor Mode.

This database architecture approach has been taken for many years of operations, but there was a necessity in reviewing it as the repositories were growing considerably and operational databases were becoming more difficult to maintain and manipulate. Most of the OBs in the repositories were only kept for archiving purpose.

In this paper, we present the strategy adopted recently by ESO for keeping the operational databases at the observing sites small with all the information generated by ESO operations archived at the headquarters.

2. DATABASE ARCHITECTURE

The new OB repository architecture developed at ESO intends to keep operational databases small and with a more or less constant size. Having control of the database size improves the performance of the databases and the recovery time in case of a database outage. The new data architecture for the OB repositories is shown in Figure 2.

At the headquarters in Garching, the OB repository has been divided into the operational and the historical repository. The operational repository containing SM OBs is replicated to the observing sites. On the other hand the historical repository keeps the SM OBs which are not necessary any longer for the operations of the telescopes, as well as a copy of any OB created at the observing sites.

One of the challenges for the architecture was to transfer the OBs generated at the observing sites to the historical repository. This was achieved using the replication mechanism. Database replication can be selective in which objects to replicate, and OBs created at the observing sites can be easily distinguished from those created at Garching by means of the primary key. As it is shown in figure 2, the design includes replication of the OBs generated in the mountain to the historical database. Therefore such objects can be deleted from the observing sites whenever it is considered there are not

necessary any longer.

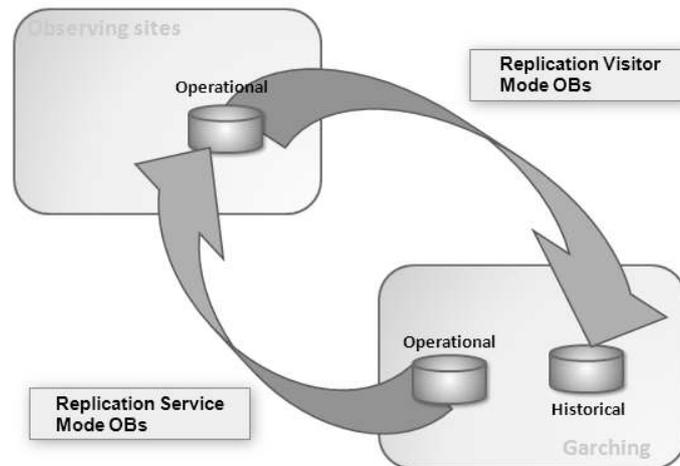


Figure 2. Database architecture with operational and historical repositories.

ESO has developed different mechanisms to control the data flow for OBs between operational and historical repositories. This mechanisms have been developed using database stored procedures, which provide a good performance and keep the logic encapsulated in the database. Using stored procedures allows to use the replication of stored procedure invocation, instead of replicating data-changing operations for replicated databases. This strategy is more efficient and avoids the possibility of overloading the replication servers.

In the next sections we explain the data flow and archiving mechanism for the different operational modes, the strategy to recover archived OBs, and the approach taken for monitoring the consistency of the databases.

3. SERVICE MODE OPERATIONS

3.-1 Standard Service Mode operation

Service mode observations are usually prepared before the start of the observing period. Investigators design their OBs using the P2PP system and submit them into the OBs repository at Garching. Once they are submitted to Garching they are replicated to the observing sites.

The User Support Department (USD) at Garching is responsible for reviewing the correctness of the OB. Once an OB is accepted, it is is “medium-term scheduled” (MTS). MTSs are a collection of instrument dependent queues containing the OBs and a series of reports describing the technical characteristics. MTS are stored into the database and replicated to the observing sites.

When an OB is in the MTS, becomes available for execution by the staff at the observatory. The execution of a SM OB is registered in the nightlog, which is used by the USD for review and resolution. Once all the OBs belonging to an observation run are successfully executed, they are removed from the MTS and the observing run is marked as “completed”.

Observation runs which have not been executed and will not be continued after the end of an observing period are marked as “terminated”.

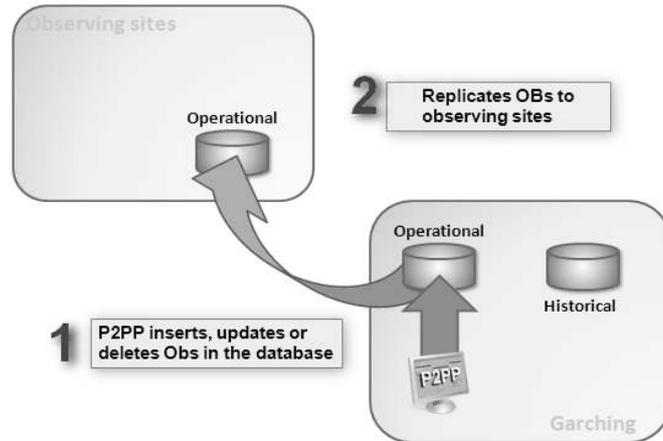


Figure 3. Standard data flow for a service mode run.

3.-2 Archiving procedure for Service Mode

Moving OBs to the historical database is the responsibility of the USD at Garching. In general, Service Mode observations are executed during their assigned observing period, typically six months long. There are special situations when OBs are carried over the next period, for example when there was a mistake executing the OB or the OB was not executed for some reason.

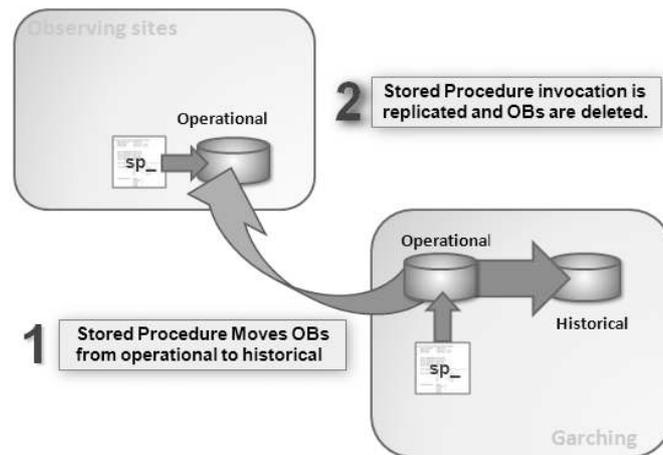


Figure 4. Service mode archiving procedure.

An archiving procedure has been developed at ESO which is run every six months (ESO's standard observing period). This stored procedure moves the OBs from the operational to the historical databases with the following criteria defined by the USD:

- OB must be three observing period old and it has not been modified in the last 540 days.
- The Observation Run for the OB must have a “completed” or “terminated” status.

The stored procedure invocation is replicated and an equivalent stored procedure is run over the operational database at the observing sites. This procedure deletes at the observing sites the same OBs which have been archived at the

headquarters.

4. VISITOR MODE OPERATIONS

4.1.- Standard Visitor Mode operation

Visitor mode operations have by definition more freedom, as the investigator is on-site when performing the observation. OBs are generally prepared with support of the ESO staff at the observatory and they are not checked into the repository. However a copy of each Visitor Mode OB is written to the operational repository at the observatory after execution for archiving purpose.

OBs written into the operational database at the observatory are replicated to the historical database at the headquarters. There is always a copy of the OBs created in the mountain at the historical database at the headquarters.

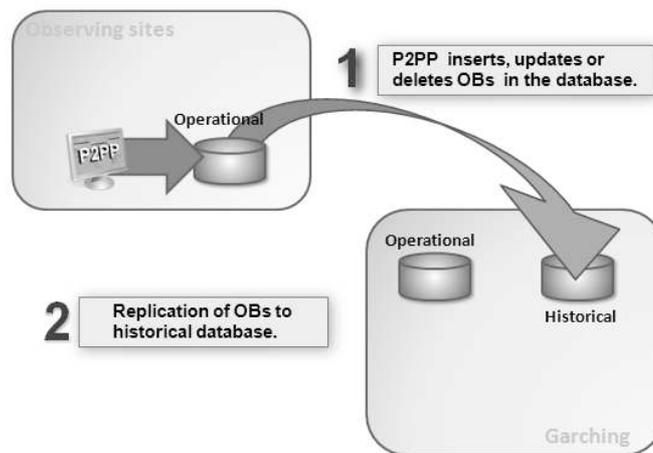


Figure 5. Standard data flow for a visitor mode run

4.-2 Archiving procedure for Visitor Mode

It is responsibility for the ESO staff at the observatory to remove Visitor Mode OBs from their repository. Generally Visitor Mode OBs are never re-executed and they could in principle be deleted as soon as the visitor leaves the mountain. However for safety reasons Visitor Mode OBs are removed only when they are at least 30 days old.

The archiving procedure developed for ESO observing sites runs every month. This procedure deletes all the Visitor Mode OBs that are older than 30 days. The stored procedure invocation is replicated to a dummy procedure in the historical repository. The dummy procedure does nothing as it is intended to leave a copy of the OBs in the historical repository.

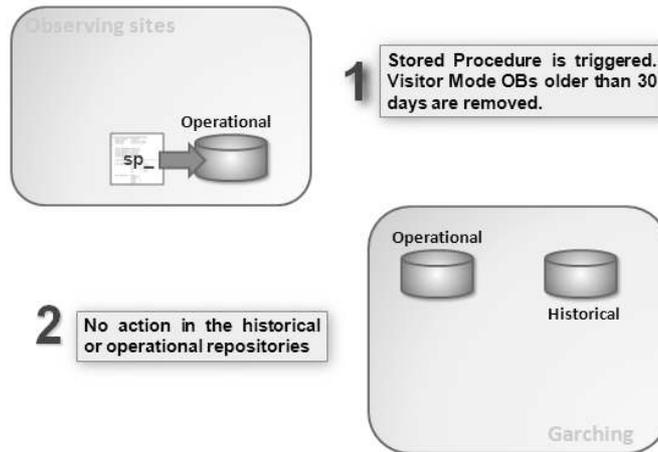


Figure 6. Visitor mode archiving procedure.

5. OB RECOVERY STRATEGY

It is possible that already archived OBs will have to be executed again, requiring OBs to be moved from the historical repository to the operational repository. This can be achieved with the use of a stored procedure in the historical repository. This stored procedure is triggered with the observing run identifier for the OBs to be re-executed as the input parameter. The procedure copies the OBs belonging to that observing run into the operational repository at Garching and deletes them from the historical repository.

OBs moved to the operational database at Garching are replicated to the operational database at the observing sites by the same mechanism used for standard service mode OBs. Therefore OBs are in the operational repositories and can be scheduled and executed again.

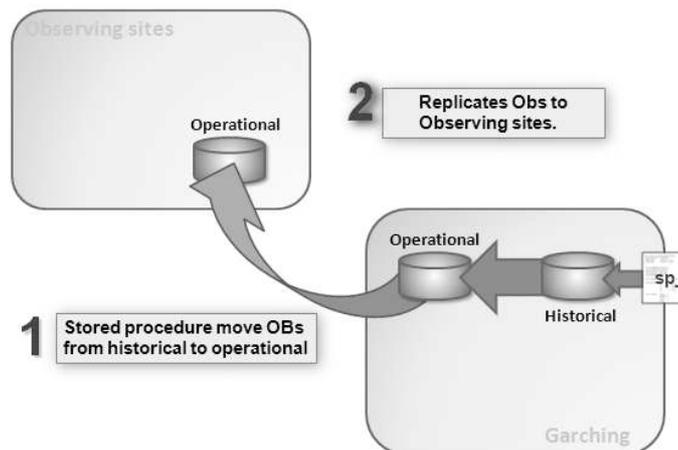


Figure 7. OBs recovery strategy

6. CONTENT MANAGMENT

Operational database at the observing sites contain both SM and VM OBs. That necessitated the creation of a mechanism to easily distinguish the content for different mode runs to be able to check the database consistency. The mechanism to interact with the operational repository as a SM or VM repository has been achieved using database views. For each table in the repository, SM and VM views have been created. The VM views return data created at the observatory, while the SM views return data which have been replicated from the operational repository at Garching.

Therefore, data consistency check is performed following two rules:

- All data in the operational repository at Garching must be returned from the SM views at the observing sites
- All data returned by the VM views at the observatory must exist in the historical repository at Garching.

These rules have been implemented in a script that is run daily, making sure no errors have occurred.

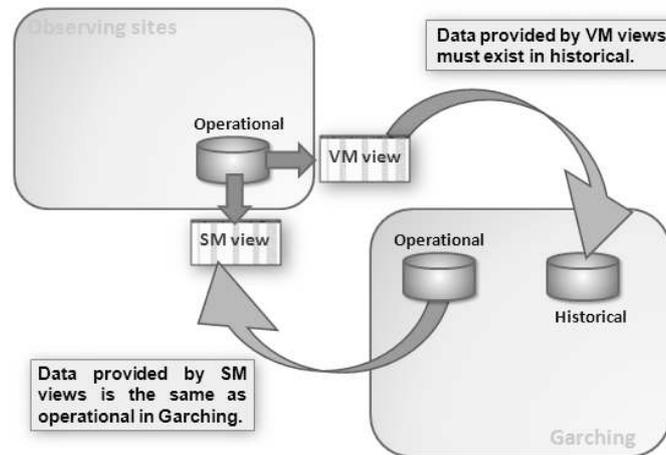


Figure 8. Content management of Service and Visitor mode runs using database views

7. CONCLUSIONS

We have presented the new database architecture developed at ESO for managing observing blocks. Observing blocks repositories are replicated between ESO headquarters and the observing sites. When telescopes are operated in the Service Mode, observing blocks are manipulated by the User Support Department at the headquarters and replicated to the observing sites. On the other hand, when operated in Visitor Mode, responsibility for those observing blocks is transferred to the ESO staff at the observatory.

The new approach involved creating a new repository at the headquarters to hold historical observing blocks. The User Support Department can remove Service Mode Observing Blocks which are not necessary anymore in the operational repositories and keep a copy in the new repository. In the same way the ESO staff at the observing sites can remove Visitor Mode Observing Blocks from the repositories at the mountain and a copy of those observing blocks is kept in the new historical repository. For situations when OBs have already been archived and have to be executed again, a strategy to recover them has been implemented. Finally, database content management has been improved.

All the new procedures have been developed using database technology, namely stored procedures and views. Database synchronization is achieved using database replication. Having the logic encapsulated in the database system is preferred to a more distributed system as it improves maintainability.

Having small databases will improve the efficiency of the queries, it make databases easy to recover in case of database outage and improves the efficiency of the content management.

REFERENCES

- [1] Chavan, A. M., & Albrecht, M. A., "A Distributed System for "Phase II" Proposal Preparation", ASP Conf. Ser. Vol. 125, Astronomical Data Analysis Software and Systems VI, 367-370 (1997).
- [2] Chavan, A. M. , Giannone, G., Silva, D., Krueger, T., & Miller, G. E., "Support tools for the VLT operations: the

NTT prototyping experience", Proc. SPIE 3349, 97 (1998).

[3] Chavan, A.M., et al., "A Front-End System for the VLT's Data Flow System", in Observatory Operations to Optimize Scientific Return", Proc. SPIE 4010 (2000)

[4] Grosbøl, P. & Peron, M., "The VLT Data Flow Concept", in ASP Conf. Ser., Vol. 125, Astronomical Data Analysis Software and Systems VI, ed. G. Hunt & H. E. Payne (San Francisco: ASP) (1997)

[5] Quinn, P.J., et al., "VLT Data Flow System: From Concepts to Operations", in Observatory Operations to Optimize Scientific Return, Proc. SPIE 3349 (1998)

[6] Quinn, P.J., et al., "The ESO Data Flow System in Operations: Closing the Data Loop", in Observatory Operations to Optimize Scientific Return, Proc. SPIE 4010 (2000)