

Chapter 4

Dependant Data Mart Environment

1. Introduction

In chapter 3 we described an organisation wide data warehouse. Although these types of data warehouses were the first to be introduced, they can sometimes fail to provide information to the end user effectively and within an acceptable time frame (Bersin, 1996). To address these weaknesses, an organisation may choose both an organisation wide data warehouse that addresses all operations and a simplified data structure for selected business units or subjects.

The following are primary reasons why such an environment is more favourable than a single organisation wide data warehouse (D2K, 1996):

- Business departments which implement a data mart as an extension of the data warehouse are able to customise the flow of data introduced into the data warehouse.
- The amount of historical data needed by the business function will be far less. Accordingly, the business unit will have access to historical data relevant only to their operations. This in turn will improve access and query response times.
- The business function will be able to select data extraction and analysis tools specific to their needs instead of being required to use an organisation wide package which might not cater for their exact needs.
- The unit cost of processing and storage on the size machine that is appropriate to the data mart is significantly less than the unit cost of processing and storage for the data warehouse environment.

The key advantage of such a data structure is improved response times to access data and process queries (Bersin, 1996).

2. Aim

In this chapter we provide the internal auditor with an understanding of how the dependant data mart relies on the existing data warehouse. We also aim to show what control risks exist within this environment. We will however only identify the outstanding internal control risks which are unique to the data mart environment in this chapter. This is because certain internal control risks and considerations identified in chapter 2 and 3 of this study also apply to the data mart environment.

This chapter concludes by providing suitable internal control considerations which can be applied in assessing these unique internal control risks.

3. Understanding the dependant data mart environment

3.1 Background

A data mart is generally defined as a database or collection of databases designed to help managers make strategic decisions about their business. A data warehouse however combines databases across an entire organisation. Data marts are usually smaller and focus on a particular subject or department (AOL, 1996).

3.2 Development considerations

The development of the data mart environment should comply with a development methodology similar to that outlined in Chapter 2. Since the scope of the development project will be restricted to a single business unit or subject, the project will need to focus more aggressively on efficiency and effectiveness. This will ensure that project deliverables are attained within a suitable timeframe and within budget constraints.

The data mart's key advantage over a single organisation wide data warehouse is improved response time. Therefore increased focus should be applied in ensuring rapid response to queries raised by users. To assist in achieving the goal of improved response time, the development team should adopt a more rigid development strategy

than that adopted for the organisation wide data warehouse. Accordingly, this strategy should consider the following specifications (Bersin, 1996):

- The most effective extraction and interrogation tools should be selected to ensure that queries are processed in the shortest possible time frame.
- Data marts should be able to address an increase in capacity over time. Studies show that data marts grow in size by 30 to 100% per annum if the full functionality of the data mart is realised by the user.
- Common data elements should be used if the organisation intends on implementing a number of dependant data marts for certain business units or subjects. This will ensure that data is consistent in quality.
- Ongoing monitoring of the data mart environment is in place where the size of the environment dictates.

The transfer of data from the existing data warehouse to the data mart should be controlled. This interface between the two environments should consider (D2K, 1996):

- The interface should execute on a periodic basis.
- The interface may update the environment only with new data or may refresh all data resident within the environment.
- The interface should link the metadata resident within the data mart to that resident within the originating data warehouse. The interface needs to describe how the data between the two environments is linked so that if the user drills down between the two environments, the data marts meta data can easily find the heritage of the data within the data warehouse.

4. Internal control risks and considerations within the data mart environment

In this section two unique internal control risks are identified which may exist within the data mart environment. Under each of the risks identified a brief explanation of the risk is provided. We will also indicate which of COBIT's information criteria, viz. effectiveness, integrity, availability, efficiency, confidentiality are affected.

The internal auditor is provided with suitable internal control considerations which can be applied in assessing each of the internal control risks.

4.1 A lack of sufficient response time monitoring on a periodic basis

4.1.1 Risk explanation

The inconsistent or total lack of ongoing response time monitoring within the data mart environment may result in information not being provided to the user on time. In such instances the user will either stop utilising the data mart environment or be unable to make informed management decisions (Bersin, 1996).

According to COBIT's information criteria identified in chapter 1, the risk identified affects the efficiency, effectiveness and availability aspects of information.

4.1.2 Internal control considerations

The following internal control considerations are applicable (Bersin, 1996):

- As part of the monitoring procedures, the management team should monitor what data is being accessed, what response times are being achieved and how much data is requested. They will also ascertain what the busiest times of the day, week and month are.
- Escalation procedures should be in place to ensure that unsatisfactory trends are reported upon and actioned.
- Users should be trained on what is considered acceptable response times and what to do if the data mart environment does not meet minimum standards.
- Controls should be in place to limit the number of ad-hoc queries which are not catered for as part of the pre-prepared reports. They should also identify queries during the data mart's initial project development.
- The network on which the data mart relies upon should be monitored on a periodic basis to ensure that no related network hardware or software problems exist.

4.2 *Transfer of data from the organisation wide data warehouse to the data mart is not controlled*

4.2.1 Risk explanation

The data mart may become inefficient and ineffective if the project team fails to ensure that only the most necessary data is transferred from the organisation wide data warehouse to the data mart (Bersin, 1996).

The upload of data from the data warehouse to the data mart may also create data integrity and availability problems if not controlled. The uncontrolled uploading of data may arise if:

- The frequency of updating data is not in agreement with the end user's needs.
- The data mart is not updated with data changes made in the data warehouse environment.
- The inability of the data mart environment to notify the users of subsequent changes made to data already relied upon, may result in incorrect management decisions being made. This is applicable in instances where a total refresh of data occurs.

In all three instances, the significant loss of decision making and the reliance on incorrect data can cause significant financial losses to the organisation.

According to COBIT's information criteria identified in chapter 1, the risks identified affect the availability, integrity and reliability aspects of information.

4.2.2 Internal control considerations

The following internal control considerations are applicable (Bersin, 1996):

- Procedures should be in place to ensure that the data transferred to the data mart is based upon assessments performed which determine which data types are most frequently accessed by the end users within the originating data warehouse. This

- will guarantee that the data mart's response times are efficient and that only the most needed data is retained within the data mart environment.
- Procedures should be in place to ensure that all data transferred from the operational system is transferred at the most appropriate time. The project team in conjunction with the end user should identify the time when the upload of data should take place so as to reflect the correct data characteristics.
 - Users in conjunction with the project team should have agreed upon whether the uploading of data to the data mart should be on a total or partial refresh basis.
 - The project team should have taken steps to ensure that the procedures which will be followed in updating data already resident within the data mart complies with an approved methodology.
 - Suitable checks should be in place to ensure that any significant changes made to operational data already relied upon by the user are effectively followed-up and communicated.

5. Summary

The dependant data mart is an extension of the data warehouse. This extension allows for improved access to data by segregating data types according to business functions or subjects.

In addition to the internal control risks identified in chapters 2 and 3, we identified an additional two internal control risks relating to the dependant data mart. These are:

- A lack of sufficient response time monitoring on a periodic basis.
- The transfer of data from the organisation wide data warehouse to the data mart is not controlled.

Suitable internal control considerations were also provided as a means of assessing the extent of these internal control risks.

6. Conclusion

For organisations seeking improved response times from their organisation data warehouses, data marts are fast and inexpensive to implement and are able to show a

fast return on investment. To achieve this however, it is important to realise that the data mart should be developed according to a rigid development methodology, similar to that introduced in chapter 2 of this study.

To ensure that the data mart environment also provides ongoing benefit to the organisation, the environment should be monitored on an ongoing basis to ensure that it does not become unmanageable over time.

2. Aim

In this chapter we provide the internal auditor with an understanding of the internal data warehouse environment. We also identify what internal controls should be in place