

Battling Growth and Staying Compliant Through Oracle® Application Archiving

By Sean Scott

Editor's Note: Data bloat can be a disastrous problem. Whether you're an archive administrator or a system user who'd like to see better performance, you'll find our first ORAtips article on data archiving to be very valuable indeed. A respected member of the Oracle® community, Sean Scott uses a holistic approach to detail the pros and cons for identifying a solution that best meets your current and future business needs including storage considerations and data compliance requirements.

Introduction

We live and work in an age where information is becoming more pervasive and valuable each day. As the custodians of that data, we are expected to maintain it, safeguard it, and make it increasingly available to more end users. Those end users have high expectations, too. No longer is it simply transaction processing and some reporting – data warehouses and data mining applications are placing greater demands on databases and hardware.

According to the Meta Group, Inc., data is growing at a rate of 125% annually. This growth is taxing the systems responsible for storing the data. As databases grow, performance drops, backup and recovery times increase, and storage and infrastructure costs rise.

But this same study reports that up to 80% of the data we store is inactive. So it would seem that addressing the problem of data growth is as simple as identifying the inactive data and removing it.

As databases grow, performance drops, backup and recovery times increase, and storage and infrastructure costs rise.

Unfortunately, in a competitive, global marketplace, every bit of data has value, and may hold the key to attracting or keeping customers. Thus, it is rare to find end users who would agree to have some or all of their data purged from the system in the interests of performance!

The Compliance Issue

Government regulations introduced in recent years complicate matters even further. The Sarbanes-Oxley Act, SEC Rule 17-a, HIPPA, and numerous foreign laws all require that records be kept for years and even decades. In some cases, this

data must be immediately searchable and available to auditors.

Often, the laws themselves are ambiguous regarding exactly what data should be kept, leaving businesses in the uneasy position of navigating multiple domestic and international regulations. The time and cost associated with choosing what to keep and what to discard is increasingly being weighed against the safer option of simply keeping everything.

Some legislation even mandates that the data be kept in its original form, frequently with a time-stamp, and its referential integrity intact. The implication is that data should be kept in the original application, and cannot be converted into XML format or dumped into a data warehouse for storage.

To put this in perspective, consider backup tapes taken by your business 10, 15, or 20 years ago. Does your company still possess the technology to restore those tapes? Will your database be able to accommodate the data? Will your application understand it? What changes will your business face in the next 10 to 20 years, and will your current data still fit?

Clearly, our existing databases and applications will not outlive the retention requirements of some of the data they presently manage!

This article will discuss how Data Archiving works to solve these problems, some things to consider when selecting a solution, and some resources you can use to get more information on the issue.”

Information Lifecycle Management and Oracle Data Archiving

Information Lifecycle Management (ILM) attempts to manage and control this data bloat. The premise behind ILM is that not all data is equal, and thus can be stored differently. Data is kept on the most appropriate storage device, based on its accessibility needs: Active data on high-cost, high-performing media; older, static data in an active archive instance and moderately priced storage; and finally – when the cost and need to maintain the data “actively” exceeds the value of the data – on a low cost storage device such as tape or CD.

Archiving Oracle data is emerging as one of the best solutions from just a handful of approaches to managing the information lifecycle. A host of vendors have appeared, offering a variety of products aimed at tackling the issues businesses face. But archiving is not simply a technology solution to a performance, financial, and/or legal problem. It is important that companies approach archiving holistically, and have buy-in from every affected department. And because each product currently available tackles archiving from its own unique perspective, it's important that businesses consider the needs of the entire enterprise when choosing a solution.

The Idea Is Simple...

Whether your objective is to improve your Oracle database performance, reduce backup and recovery times, manage data growth and control storage costs, or comply with legal requirements, the idea behind archiving is relatively simple.

An Oracle application that doesn't use archiving stores all of its data in a single table, stored on a single Oracle database instance, using expensive, redundant storage. Archiving utilizes

a database view instead of a table. The view, in turn, looks at two or more tables. “Active” or current data is still stored in a table on expensive, redundant storage, but the “inactive” or historical data is moved to less expensive storage.

Because the Oracle application sees the view and not the underlying complexity, the archiving solution is (in theory) invisible to the application, and transparent access to the data is achieved. There is no need to create or maintain multiple applications for querying active and historical data independently, and users aren't even aware that the data is stored in more than one place.

Because the historical data is static, it doesn't need to be backed up as frequently (each time new historical records are added to the history table is adequate – usually once a month). The archive tables can be kept in the production database, or moved to an archive instance (tuned for queries rather than OLTP). The archive database can even be located remotely, at a disaster recovery site, for example.

Conceivably, a 100GB production database could be reduced to 20GB in production and 80GB in archive (based on the Meta Group, Inc. figures for active and inactive data). The implications for performance and savings are obvious. With a smaller production footprint, it easier to maintain the most critical data and tune databases to satisfy user demands. Oracle upgrades will also run more quickly when there is less data to be updated.

...The Solution is Not

Unfortunately, it is not a simple matter to move 80% of your data! There are a multitude of factors that must be taken into account, and even different ways of moving the data. Fortunately, there are accepted best practices for archiving and ILM:

- Ensure that archiving is viewed as a business solution to a business problem, and not simply an application of technology to solve a technology problem. This cannot be overemphasized, and it is critical that the individuals responsible for making decisions within each business unit are a part of the process from the onset.
- Make a determination about what data can and cannot be archived. Some apparently static data should not be archived, such as configuration information or lookup data that is necessary for an application to work properly.
- Archive only closed transactions. Since it is presumed that archive data is static, it makes no sense to place data that may still be updated into an archive instance.
- Archive transactions at a cellular level, and only when all dependent data is eligible for archiving. For example, payments depend on purchase orders, and invoices depend on shipping and billing information. Until a transaction, as a whole, is complete, all of its components should remain part of the active instance.
- Include archived data in the backup and recovery plan, and be sure to perform backups of archive data any time it changes.
- Maintain an audit trail of archiving activities and guarantee that archived data cannot be altered.

Perhaps most important is to recognize that Oracle data archiving is not a solution to an existing problem, but a solution to a future problem. When the Oracle database has stopped performing well, when storage costs have become unmanage-

able, or when disaster recovery needs cannot be accomplished is not the time to address an archiving strategy. Instead, companies must have the vision to recognize that all of these problems are avoidable and make the investment in archiving before they impact the business. Often, such an initiative is driven from the “bottom up” by the user community and IT staff.

One of the most difficult challenges faced by archiving is the Oracle upgrade path.

Oracle® Upgrade Considerations

One of the most difficult challenges faced by archiving is the Oracle upgrade path. If Oracle Applications didn't change, it wouldn't be a concern, but realistically, nobody expects to be running the same version of a database or application three, five, or ten years down the road. Even an Oracle application patch can affect the integrity of archived data by altering the underlying database relationships.

Some regulations require that data must remain accessible and retain its integrity with other data in the database. But application changes sometimes require that the underlying table structure and its referential integrity change as well. Thus, it becomes important to store data relative to its original database structure, without changing the way in which the data itself is represented or the time stamp of individual rows. Obviously, the ability of an archiving solution to adapt to application changes and treat individual records according to their original version is an important factor in choosing a solution.

Different vendors approach this problem differently. Some detect and accommodate minor changes automatically while others require a script or patch be applied to revise data access. New Oracle application releases will almost always require a patch or upgrade to the archiving application. Fortunately, vendors of archiving solutions all work very closely with database and application vendors to keep their products tightly integrated with those they support.

Additional Features

Aside from simply moving and managing data, many archiving products provide additional features:

- Capability to easily “un-archive” data back into the active instance. This provides the peace of mind to take a somewhat more aggressive approach towards what records are archived, knowing that it isn't a permanent choice! Central to this ability is an intelligent application that recognizes all the dependent data that must be restored with a particular record in order to maintain its referential integrity.

- The ability to easily clone a database or restore subsets of a production database. This can be extremely useful when creating test or QA instances since an entire production database doesn't need to be restored. The storage costs of these non-production databases are reduced, and the manpower overhead is reduced as well.
- Automated, scheduled archiving based on established business rules, relieving IT application and database administrators from having to manually perform these activities.
- Business intelligence that can produce previews of different archiving strategies, providing management with an idea of how a particular approach would impact the business.
- Some archiving vendors focus on one or two specific products or databases, while others provide a broader reach.
- Most vendors offer consulting services to assist companies in deploying their archiving solutions. Some will even preview their products on-site and perform analysis on existing data in order to demonstrate the return on investment of their products.

As with any software choice, bells and whistles shouldn't determine which product is best for your business needs. Instead, they should only serve to sway a decision when two or more products address the critical needs. Remember, the critical features are the ability of the product to help your company accomplish its business goals for compliance, performance, and managing storage costs.

Community Support

Though the idea of archiving is still relatively new, a forest of confusing information and myth has grown in that short time. Fortunately, the Oracle Applications Users Group Archive and Purge Special Interest Group (OAUG A&P SIG) represents the end-user community to both Oracle and third-party vendors. Members of the group include vendors, users who have implemented archiving, and users who are considering an archiving solution. The OAUG A&P SIG is unique in that third-party vendors participate on the organizing committee; each allocated a single seat, to guarantee that all of the technologies and products are equally represented.

The OAUG Archive and Purge SIG Website is located at <http://archivesig.oaug.org/>. Their purpose is to provide “a forum for sharing information and experiences about installation, configuration, implementation, maintenance, and continued support of Oracle Applications (all modules and releases) archiving and purging solutions. The A&P SIG will collaborate with Oracle and third-party vendors for developing standards and benchmarks for data, images and reports, archive and purge functionality, as well as improving Applications performance, reliability, and return-on-investment. A&P SIG will also focus on best practices, statutory compliance, data management & storage, retrieval & access; patch management, and enhancements of archive/purge functionality.”


Companies that are using archiving, or are considering archiving in their enterprises, are encouraged to participate in the SIG and exchange ideas and experiences.

Conclusion

Archiving can provide a multitude of benefits to an enterprise. Reduc-

ing the amount of data held in an “active” Oracle database allows database administrators to focus their performance tuning efforts on the most frequently accessed, critical data. System availability is improved. Oracle upgrade paths are shortened. Storage costs are reduced. Regulatory compliance is met more easily, and decisions over whether or not to keep data are based on legal factors and not affected by performance and cost considerations.

Data growth is a fact of business, and our Oracle enterprise applications will continue to swell with information at increasing rates. Archiving is a means by which access to all of our data can be guaranteed over time, and by which our businesses can be protected during audits and legal discovery. But it is not a magic bullet and cannot be viewed simply as a reactionary fix to an existing performance or compliance problem. It must be a part of the overall information management strategy and be implemented as a means of preventing issues associated with the increasing amount of information we must store.

Sean Scott, Argentine, Inc. – Sean is CEO of Argentine, Inc., a provider of outsourced database administration, development, and design services. He has over fifteen years of experience with Oracle products, ten of them as a database administrator. He has presented at several Oracle Applications User Group meetings, IOUG Live!, and Oracle Open World, as well as a number of local and regional Oracle User Group meetings. He is currently vice-president of the San Diego Oracle Users Group. Sean may be contacted at Sean.Scott@ERPTips.com. 

ORAtips *Journal*

The information on our website and in our publications is the copyrighted work of Klee Associates, Inc. and is owned by Klee Associates, Inc. NO WARRANTY: This documentation is delivered as is, and Klee Associates, Inc. makes no warranty as to its accuracy or use. Any use of this documentation is at the risk of the user. Although we make every good faith effort to ensure accuracy, this document may include technical or other inaccuracies or typographical errors. Klee Associates, Inc. reserves the right to make changes without prior notice. NO AFFILIATION: Klee Associates, Inc. and this publication are not affiliated with or endorsed by Oracle Corporation. Oracle is a registered trademark of Oracle Corporation and/or its affiliates. Klee Associates, Inc. is a member of the Oracle Partner Network

This article was originally published by Klee Associates, Inc., publishers of JDEtips and SAPtips. For training, consulting, and articles on JD Edwards or SAP, please visit our websites: www.JDEtips.com and www.SAPtips.com.