SOLIX®

Empowering the Data-driven Enterprise

MITIGATING RISK FROM DATABASE VOLUME GROWTH

A Lateral Approach to an Inevitable Problem

Data volume grows over time. This is an inescapable fact, without industry prejudice or database software affinity. When this happens, the database will no longer operate with peak efficiency. This fact is also inescapable and yet many organizations seem caught by surprise by the expense of maintaining consistent database performance over time. While database vendors love seeing performance decrease as it usually leads to more licensing fees, organizations and specifically application end users, express a different emotion.

Beyond additional hardware and software costs, inefficient databases also reduce productivity, a much harder number to calculate and likely much more expensive. A slower database equates to fewer transactions being processed per day. It could mean longer periods of database unavailability, as required maintenance tasks take longer to complete. Most software and hardware vendors will tell you that the answer is more software and hardware. Solix believes the answer is simple, keep the database from growing in size in the first place.

Database Active Archiving is a fundamental data management process. It is an aspect of data governance called information lifecycle management (ILM). Experts estimate that 80% of all data under management is either redundant, obsolete or trivial, how we handle this data bloat will factor directly into a company's bottom line.

Conventional Wisdom for Dealing with Data Volume Growth

Databases, simplistically speaking, do two things, they store data and retrieve data from that store. How long it takes to do either of those tasks impacts how productive the end user of an application is. If that end user is a customer, poor performance could send them to a competitor's website. Slow performance impacts the total number of business transactions that can be completed in a day. Both can potentially impact an organization's bottom line profitability.

Over time, the volume of data being managed by the database will grow. When a user requests information, the database must complete an increasing number of read operations to satisfy the end user request. Imagine looking for a person's name and address in a single phone book. Not a big deal. Now find the same information when you only have the person's name. A search through all the phone books in the country would take a great deal of time and effort. This is similar to what a database must process as the volume grows, only the database could contain, not only the current phone book data, but past phone books as well. All of this takes time and resources. Database software has continuously improved in sophistication, adding features and options to overcome many of the issues brought on by the growth in data volume. Indexing strategies, data partitioning, increased use of in-memory data stores. None of this comes cheap. Hardware vendors sell faster storage arrays, improved network protocols with greater throughput, faster processor cores and larger memory. Also, not cheap. Conventional wisdom when dealing with data volume growth in the database is additional hardware capacity and/or additional database options, each comes with a rather large price tag.

CASE STUDY:
WELL-KNOWN ONLINE RETAILER

This well-known online retailer has more than one million products on its site. They experienced explosive growth in its first decade of operations. Revenues increased 1000% in the decade. They are completely reliant on their IT infrastructure to satisfy both its shoppers and suppliers.

This intense growth proved too difficult for the Oracle eBusiness Suite (OEBS) applications to handle. Not only was the production application slowing but regular batch jobs to maintain clones for testing and take backups were simply taking too long. Other applications dependent on OEBS were also being negatively impacted.

"Thanksgiving to Christmas is our busiest season. We were watching our systems closely and trying to prepare for those busiest times. In '07 the system was basically brought to its knees, and we were very concerned about it crashing the system."

DEVELOPMENT MANAGER, ERP SYSTEMS

CONNECT WITH US

community.solix.com/blog
facebook.com/SolixTechnologies
twitter.com/SolixBigdata
info@Solix.com

When faced with conventional wisdom, it is often the simplest and least expensive answer that is overlooked. You might remember the story of the tractor trailer truck that was wedged into a tunnel. Its height, just barely over the restriction. The stuck tractor trailer had backed up traffic for more than a mile and rush hour hadn't even begun. Rescue crews contemplated cutting off the top of the trailer, an operation which could take hours. A little kid observing the scene made a simple suggestion. "Why not let the air out of the tires?" she said.

Lateral Thinking to Solve the Data Volume Issue

Solix employs the same lateral thinking that the little girl used to solve the problem mentioned above. The simple solution to ever increasing database size is to eliminate its growth in the first place.

Active database archiving employs simple timeframe techniques combined with, potentially complex business logic, to determine which data can safely be removed from the production database.

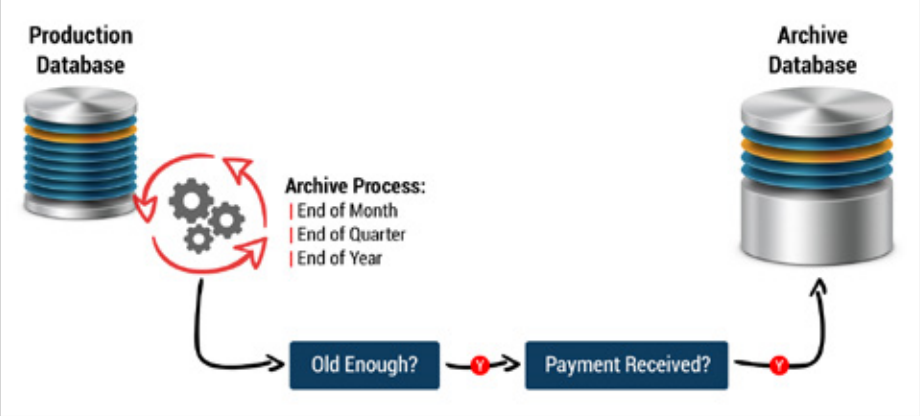

Figure 1: The active archive process ensures that only rows that meet all criteria for archive and purge are met. This process keeps the production database lean and efficient while the archive grows with the older less frequently accessed data.

While the process is simple enough to explain, correctly implementing an active archive process can be a challenge. This is true because every application has different logic. Every application implements referential integrity in its own way. Organizations differ in how they want to access archived data or even on what platform they want to archive data. Here are just some of the considerations that have to be addressed:

- **Archive Platform:** The production database is not limited to only archiving to another database. You might choose HDFS or an object store and that archive store might be on premises or in the cloud.
- **Archive Database:** If you choose to archive to another database, you may not be limited by the production database type. For example, the production database could be Oracle but the archive could be MySQL to save licensing costs.

Answers were not simple or cheap, however. The ERP team considered options ranging from bigger and faster hardware, application architecture changes, database tuning and purging and archiving. They refreshed their hardware and began working on the application architecture and database tuning. The results were less than impressive, While the online retailer made it through its 2008 fourth quarter, they realized the measures implemented could not be sustained over time. Better solutions were critical for the company.

They partnered with Solix and began examining the critical issues facing the retailer.

"The Solix team were really good at coming in and helping us analyze our data and the growth patterns we were having. They made some really good recommendations around the areas we should target first."

Fourth Quarter performance in 2009 improved after the Solix installation, and the need for hardware refreshes due to capacity or performance concerns abated.

CONNECT WITH US

community.solix.com/blog
facebook.com/SolixTechnologies
twitter.com/SolixBigdata
info@Solix.com

- **Access to Archive:** Do you need access to both the current and the archive data through the native application UI or can the archive data be accessed using other third party tools?

- **De-Archive, Legal Holds and Retention Policies:** Do you foresee a need to de-archive a transaction based on a change? Will you need the ability to place legal holds on some of the archived data due to pending litigation? Finally, do you need to implement strict retention policies for regulatory compliance?

It takes an experienced team to ask the correct questions and implement an active archiving solution. In many organizations, the knowledge of the application and its infrastructure, along with the answers to some of the questions posed above, do not exist within a single team or individual. The DBA might be very well versed in the database infrastructure and how it is maintained, but have little knowledge of how the application actually works. Application developers may understand how the application works but have only a passing understanding of the database infrastructure, and a limited understanding of the business user's requirements. Business users might be clear on their archived data access needs and retention policies but don't understand clearly their choices and trade-offs for implementing their requirements.

How to Justify Active Database Archiving

IT infrastructure teams are generally the first to realize that dealing with database growth using conventional wisdom is not sustainable. However, the forces aligned against a simple active archiving solution are powerful. This would include both database software and hardware vendors that often command the attention of decision makers due to the vendor's dominance of huge portions of the IT annual budget.

Most often an archiving initiative can be easily justified by looking at three main areas of savings.

- **End User Productivity:** The better performing the application is, the more work can be completed in a shorter amount of time. In addition, database maintenance tasks such as backups, reorganizations, cloning etc. get done faster making the database more available even for users of non-production copies of the database. Application upgrades and patches are completed in less time, meaning the application can have the latest security fixes and functional improvements with less downtime. The average initial improvement in performance after archive is 30%. A thirty percent boost in productivity could easily equal hundreds of thousands in savings per year for each application you maintain via active archiving.

  IT Staff Efficiency: When database performance is increased, time spent monitoring and tuning is reduced. This impacts system admins, database admins and storage admins. Admittedly, a much smaller universe of personnel when compared to end users, but often highly paid positions. Imagine if you could shave a modest 10 man hours per month, per database?

The primary database stood at 3.4 terabytes in 2009. The Solix database archiving solution has allowed the retailer to archive 1 terabyte of data, leaving the live database at 2.4 terabytes, despite continued growth. With the expertise of the Solix team, they targeted another 700 GB of data in their order and shipping data suitable for archiving.

The retailer doubled its product selection, increased its vendors and offered even richer product attributes on each SKU. **The Solix database archiving solution** has ensured the retailer's systems are prepared for the growth.

"Solix's commitment to working with the Oracle EBS, as well as interfacing with all of our systems, has been impressive. Patches and updates have been simple."

"The relationship is very important. Find a partner. Someone who will work with you. Solix really listened to what we wanted, what we needed and helped us accomplish those goals."

DEVELOPMENT MANAGER, ERP SYSTEMS
WELL-KNOWN ONLINE RETAILER

CONNECT WITH US

community.solix.com/blog
facebook.com/SolixTechnologies
twitter.com/SolixBigdata
info@Solix.com

- **Hardware and Software Upgrades:** Imagine being able to amortize hardware over a five-year refresh cycle instead of three? What if you could eliminate the need for extra cores on your database server? The savings could be in the tens to hundreds of thousands each year.

The best part is that the investment made in active archiving is relatively minimal when compared with the conventional wisdom alternatives. Single projects are generally completed in a less than a quarter and the benefits are immediately apparent. All at a fraction of what a conventional approach would cost.

Solix Database Archiving Solution

The Solix database archiving solution handles both the active archiving of data for applications still in production, but also the archiving of legacy applications that need to be retired gracefully with full regulatory compliance.

The Solix archiving solution is database and platform agnostic. Solix has developed specific knowledge bases for many well-known packaged applications such as Baan, JD Edwards, Siebel and SAP. Solix offers over 50 pre-packaged knowledge base modules in support of Oracle eBusiness Suite alone. These significantly reduce the time and effort to successfully complete an active archiving project. With Solix, you have several options for how and where you may want to archive your data. The Solix database archiving solution seamlessly plugs into the Solix Common Data Platform (CDP) to extend your archiving capabilities beyond simple structured data to include semi-structured and unstructured data.

Archiving Options:

- Database to Database (the same database platform or a different one) – Solix Database Archiving
- Database to Hadoop(HDFS) on premises or in the cloud– Supports full SQL access along with text search – Solix CDP
- Database to Object store on premises or in the cloud – Solix CDP

No matter how; or where you choose to archive, Solix fully supports the ability to implement legal holds, automatic retention policies and the ability to de-archive transactions back to the production database.

In addition to some of the previously mentioned packaged applications, the Solix database archiving solution also supports virtually any application on any type of database, both relational and non-relational (or pre-relational). Solix data management experts work with customers to help determine the best, most cost effective and sustainable approach for your organization's needs. With the widest variety of pricing and implementation options in the industry, Solix will work with you to tailor the correct solution for your organization.

Gartner

Gartner MQ Leader*

- Ranked Highest by Buyers
- Cost-Effective Solution
- Hadoop Support

CONNECT WITH US

community.solix.com/blog
facebook.com/SolixTechnologies
twitter.com/SolixBigdata
info@Solix.com

Why Solix

Solix is a recognized leader in the structured data archiving market[1]. Specifically recognized for high customer satisfaction and the widest array of pricing and deployment options. While this paper has focused on database active archiving, Solix was also singled out for its support of both semi-structured and unstructured data archiving as well.

With almost twenty years of experience, Solix has developed a deep center of excellence in the area of active and legacy data archiving. Supporting virtually any platform and database, the Solix team of experts in 2018 received awards for industry best practices[2]. Having also been singled out for ease of deployment, best array of pricing options, cloud support and high customer satisfaction, Solix is ready to partner with your organization, think laterally and avoid following expensive conventional wisdom.

CONNECT WITH US

community.solix.com/blog
facebook.com/SolixTechnologies
twitter.com/SolixBigdata
info@Solix.com

[1] Gartner Magic Quadrant for Structured Data Archiving and Application Retirement 2016 (most recent report)

[2] Frost & Sullivan Global Product Leadership Award 2018

SOLIX®
Empowering the Data-driven Enterprise

Copyright ©2018, Solix Technologies and/or its affiliates. All rights reserved.

This document is provided for information purposes only and the contents hereof are subject to change without notice.

This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchant- ability or fitness for a particular purpose.

We specially disclaim any liability with respect to this document and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.

Solix is a registered trademark of Solix Technologies and/or its affiliates. Other names may be trademarks of their respectively.

SOLIX TECHNOLOGIES, INC.

4701 Patrick Henry Dr., Bldg 20, Santa Clara, CA 95054

Toll Free: +1.888.GO.SOLIX (+1.888.467.6549)

Telephone: +1.408.654.6400

Fax: +1.408.562.0048

URL: https://www.solix.com

CONNECT WITH US

community.solix.com/blog

facebook.com/SolixTechnologies

twitter.com/SolixBigdata

info@Solix.com