Maximizing Storage Efficiency with Oracle and SQL Server Databases

James Green, vExpert
David M. Davis, vExpert
ActualTech Media
September 2016
Table of Contents

Table of Contents ............................................................................................................. 2

Introduction .................................................................................................................... 3
The Downside to Copy Data ............................................................................................. 3

What is Copy Data Management? .................................................................................... 4

Challenges in Mission Critical Database Storage ......................................................... 5
Excessive Copies .............................................................................................................. 5
Copy Data Security ......................................................................................................... 5
Copy Timeline .................................................................................................................. 5
Stale Data ....................................................................................................................... 5
Storage Integrations ....................................................................................................... 5

Solutions in Mission Critical Database Storage ........................................................... 6
Automate Copy Processes ............................................................................................... 6
Eliminate Copy Delays .................................................................................................... 6
Refresh Data Frequently ................................................................................................. 6
Utilize Copy Data Masking .............................................................................................. 7
Flexible Tool Selection ................................................................................................. 7

Recommendations .......................................................................................................... 8
What is Catalogic ECX? .................................................................................................. 8
Learn More ...................................................................................................................... 8
About ActualTech Media ............................................................................................... 8

ActualTech Media © 2016. All rights reserved.
Under no circumstances is this document be sold, copied, or reproduced in any way except with written permission.
The information contained with the document is given in good faith and is believed to be accurate, appropriate, and reliable at the time it is given, but is provided without any warranty of accuracy, appropriateness or reliability.
The author does not accept any liability or responsibility for any loss suffered from the reader’s use of the advice, recommendation, information, assistance or service, to the extent available by law.

All trademarks are the property of their respective owners.
Introduction

Organizations require copies of production data for use in development, QA, analysis, and reporting. For various reasons such as performance impact, security implications, and data integrity, performing anything considered a dev or test function on production data is rarely a good idea. Instead, organizations create frequent copies of production data sets and continually refresh them to allow secondary functions to be performed with real data.

This practice isn’t without its drawbacks, however. According to IDC, over 80% of organizations maintain more than ten copies of every Oracle database instance at any given time.

IT now live in a world with a serious copy data problem.

The Downside to Copy Data

Whether your company’s critical applications use Oracle or SQL Server as their backend database, it’s likely that you’re using more storage than necessary for your mission critical data thanks to the need to create dev/test and other copies. You’re also probably spending quite a bit of time manually copying and refreshing copies of that data.

The end result is that you’re paying more for your mission critical storage than you need to. On top of that, you’re spending much more time than you should managing all of that copy data.

With many companies moving their mission critical storage to all-flash storage and with the price of all-flash storage still being exceptionally higher than traditional storage, it’s time to learn how you can maximize your storage efficiency using copy data management, or CDM.
What is Copy Data Management?

*Copy data* is the collective set of all data not currently being used in production (e.g., a snapshot, backup, vault, or replica of a data set made for various IT or business functions—data recovery, test/dev, analytics or operations functions). *Copy data* is distinct from *backup data* in that backup data is for the sole purpose of recovery whereas copy data is used in ongoing development and test processes and is often manipulated as a part of those processes.

Creating data copies results in duplicate data floating around an organization. These data copies are frequently not well understood, are not tracked, and can be difficult to provision with any level of promptness. Solving these problems is the job of Copy Data Management solutions.

Copy Data Management (CDM) is a relatively new IT category of solutions intended to centralize the creation and management of copies of production data. Figure 2 shows some management features that one might look for in a CDM solution.

A Copy Data Management solution assumes responsibility for the operational process of making copies and delivering them to data consumers. This can include scheduling copy creation, managing copy retention, driving copy distribution, tracking copies, reporting on copies, and integrating copy creation with application or hypervisor stacks.
Challenges in Mission Critical Database Storage

There are a number of challenges associated with copying production data. Left unaddressed, these challenges can have a negative impact on an IT organization’s bottom line.

**Excessive Copies**

First and foremost, a lack of organization around copy data frequently leads to a quantity of duplicates that is disproportionately large in relation to the number of copies that are actually needed. The immediately obvious result of this challenge is increased storage usage, both in terms of performance (IOPS consumed) and in terms of capacity used.

The not-so-obvious implications include confusion around which copy a developer or analyst should use, accidental use of stale data, and even security risk due to the fact that the excessive copies probably aren’t tracked or inventoried.

**Copy Data Security**

There’s more to the security story, too. Because creating copies is commonly a manual process, organizations often also rely on manual or homegrown systems to mask sensitive data in the copies, if they even mask the data at all. This manual masking process takes time and is potentially unreliable.

**Copy Timeline**

Due to the nature of the tools at the disposal of many IT operations personnel, the timeline to deliver a new copy of production data once it’s been requested is frequently longer than desirable. The process to provision these copies is clunky and the level of automation of the task in most organizations is quite low. While not incredibly complicated, it does take an administrator some time to complete.

**Stale Data**

Because of the effort involved in creating new copies and the amount of time it takes to deliver them, the copies of production data that are in use tend to be older than what developers and analysts might like. In some cases, the copy data is weeks behind production, making it difficult to address the current state of the business today.

**Storage Integrations**

Lastly, the manual copy that many organizations perform doesn’t take advantage of native all-flash array functions like the ability to create a zero copy clone of the volume which has no additional capacity footprint. Letting these array integrations go to waste is a costly mistake in the long run.
Solutions in Mission Critical Database Storage

To solve the challenges associated with copy data management, the most reasonable course of action for most organizations is to procure a tool that brings together all the copy data and keeps track of the age, and relevance of the data, while also securing it. The following are some of the outcomes that you might expect from such a tool.

**Automate Copy Processes**
Anyone who has spent much time in IT operations knows that humans are error prone. Automation is not only important in the modern era because of its utility in accelerating time to value, but it also reduces risk by ensuring that the intended outcome is realized.

Further, by leveraging automation of copy data, much more granular management of copies is possible. This keeps excess duplication in check and makes sure that resources aren’t being wasted.

**Eliminate Copy Delays**
By leveraging automation as well as integration with the storage array—accomplished by combining the use of robust API integration with mature snapshot technology like that built-in to many modern all-flash arrays—the metadata snapshot requires no actual copying of data and, as a result, the snapshot completes almost immediately.

The expedient nature of this new delivery paradigm also allows you to experience previously unheard of possibilities like a self-service portal where developers and DBAs can request fresh copies of production data to be delivered on their own timeframe. The ability to offer this removes IT operations as a bottleneck and keeps development teams moving forward.

**Refresh Data Frequently**
Automating the copy process allows data consumers to rely on fresh data, as well, without creating administrative overhead. For example, the tool of choice might be configured to deliver a new copy of the production data at 6:00 AM each day.

To accomplish this by hand would wear on the IT operations, and while it’s something that has been scripted in the past, relying on cobbled together scripts for such an important business function is downtime waiting to happen. A tool with this automation baked in can make relevant data available on a schedule in a supportable, reliable way.
Utilize Copy Data Masking

Weighing heavy on the minds of most CIOs and other business leaders today is the possibility of a security breach resulting in the loss of sensitive data. Copy data is one of the more susceptible vectors by which an organization could be attacked since it is all-too-often not tracked, sometimes not scrubbed for sensitive data, and not always secured like it should be.

One measure to strengthen an organization’s security posture with regard to copy data is to employ a mechanism that masks sensitive data columns from production databases. Not everyone who will develop against or report on a data set should be privy to information like social security numbers or pay rates.

A robust CDM tool uses policy and role-based filtering to keep sensitive data away from prying eyes. Figure 4 shows a simplified example of what masked copy data might look like. The column with the sensitive data is not available to this particular user.

<table>
<thead>
<tr>
<th>First Name</th>
<th>Last Name</th>
<th>Social Security Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>Doe</td>
<td>XXX-XX-XXXX</td>
</tr>
<tr>
<td>Jane</td>
<td>White</td>
<td>XXX-XX-XXXX</td>
</tr>
<tr>
<td>Jim</td>
<td>Smith</td>
<td>XXX-XX-XXXX</td>
</tr>
<tr>
<td>George</td>
<td>Brown</td>
<td>XXX-XX-XXXX</td>
</tr>
</tbody>
</table>

Flexible Tool Selection

In the quickly evolving world of IT operations, there’s an ever-present challenge with regard to tooling: new, helpful tools are developed and available, but legacy systems aren’t able to interface with those tools. As new tools are implemented, the evaluators should always be looking at how one tool interfaces with others.

A robust Copy Data Management tool should expose controls via a standard interface so that other tools can easily interact with it. As a specific example, in the modern enterprise, many teams would like the ability to automate procedures with tools like Chef, Puppet, and Ansible.
Recommendations

As you may have guessed, the tool described in the previous section does exist, and really can help organizations solve this copy data problem. It's in the interest of all stakeholders to look at tools like this that can reduce friction across departmental silos, ease capacity and performance burdens on the infrastructure, and add controls to increase security and reduce risk.

One such copy data management tool is Catalogic's ECX solution.

**What is Catalogic ECX?**
With this software tool, IT organizations can address all of the previously described storage woes. ECX offers an extensive list of benefits, like:

- Automate the creation and use of copy data—snapshots, vaults, clones, and replicas—on existing enterprise storage infrastructure.
- Agentless model simplifies management while still providing application awareness.
- Integrated data masking keeps information secure.
- APIs allow integration with popular DevOps tools.
- Supports cataloging to Oracle RMAN.
- Automate test and development infrastructure provisioning, reducing management time as much as 99%.
- Database log capture provides point-in-time recovery.
- Simple licensing based on storage controllers.
- Supports a multi-vendor selection of storage arrays, including systems from EMC Unity, IBM, Pure Storage and NetApp.

**Learn More**
To learn more about Catalogic Software, Inc. and about the ECX software that's been discussed, visit [www.catalogicsoftware.com](http://www.catalogicsoftware.com) to request a free trial. It only takes about 15 minutes to deploy, so there's no excuse to not experience the operational and utilization advantages that can be gained by using a tool like ECX!

**About ActualTech Media**
ActualTech Media delivers authoritative content services and assets for top IT companies across the globe. Leading IT industry influencers Scott D. Lowe, David M. Davis, James Green, and partners develop trusted, third-party content designed to educate, convince, and convert IT buyers. ActualTech Media helps its clients reach the right technical and business audiences with content that resonates and leads to results.