

SQL Server Database Provisioning Report

Survey on database provisioning requirements
among SQL Server Professionals

Foreword

Waiting until the point of release to see how our software behaves is a key cause of chaotic or even catastrophic software releases. Establishing production-like environments with realistic load and data sets for use in development and testing has long been a challenge for organizations.

The need to ‘bake quality in’, by ensuring realistic configurations and data sets are ‘shifted left’, has never been higher, and the DevOps movement recognizes that.

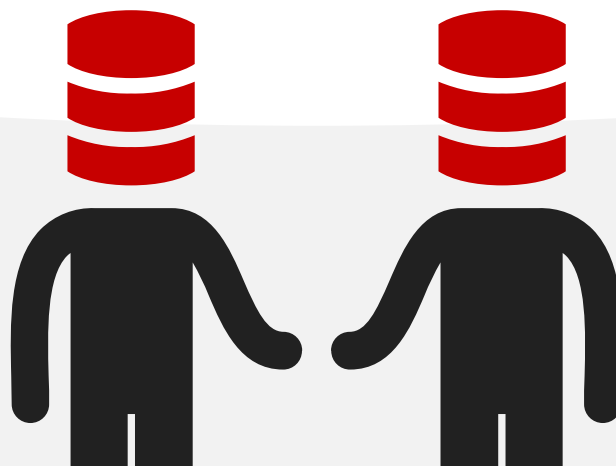
“Our goal is to ensure that Development and QA are routinely integrating the code with production-like environments at increasingly frequent intervals...

“...In other words, we will only accept development work as done when it can be successfully built, deployed and confirmed that it runs as expected in a production-like environment”

DevOps Handbook **Kim et al.**

Database provisioning carries its own challenges, with its own characteristics which have not commonly been addressed by provisioning toolsets. Some of these include concerns around data security, regulatory compliance, and internal infrastructure such as storage space.

This report seeks to identify how organizations currently provision databases, and to assess the importance of provisioning requirements.



Key Findings

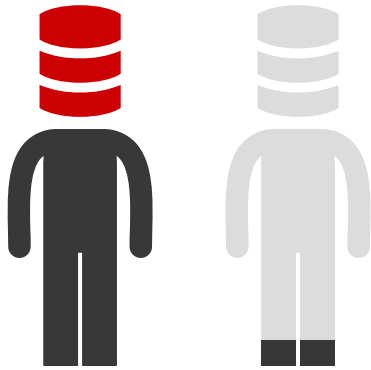
Over 460 SQL Server professionals that showed interest in Redgate's new provisioning tool SQL Clone participated in the survey. Respondents represented a wide range of job roles from a variety of industries. Among the key findings are:

- 72% of respondents provision a copy of production that is used in a non-production environment (i.e. Dev / Test / QA).
Only 28% of respondents always modify the data in some way before it reaches a non-production environment.
- 63% of organizations have provisioning controlled by the DBA. Only 29% allow individuals to self-serve.
- 54% of organizations provision weekly or more often with 28% provisioning daily.
- 83% of respondents agreed or strongly agreed that it is desirable to use production data in development and testing.
- The main factors restricting use of production data in non-production environments are **data security**, **regulatory compliance**, and **storage concerns**.

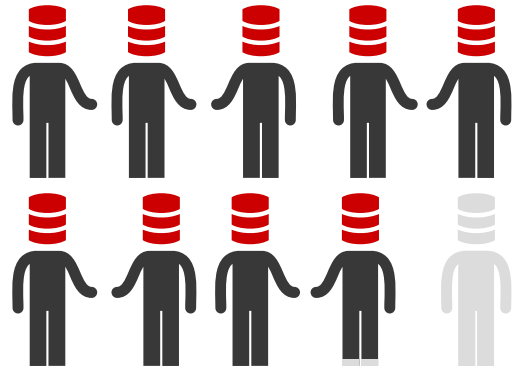


1

For databases in your day-to-day development environment (excluding QA, staging, etc.) do you? (multiple choice)

**53%**

Have a database that only you make use of

**89%**

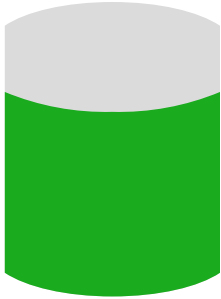
Use a database shared with other developers

As expected, more respondents work with databases in shared environments. However, many respondents also have databases that only they make use of. This varies slightly depending on job role. 59% of Application Developers have databases that only they use as opposed to DBAs (51%) or BI Developers (45%). Likewise DBAs and BI developers use shared databases more frequently (91% and 93%) than Application Developers 82%.

2

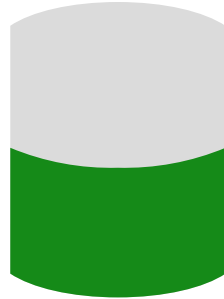
Where does the data in your development/test/QA environment come from? (multiple choice)

A copy of production



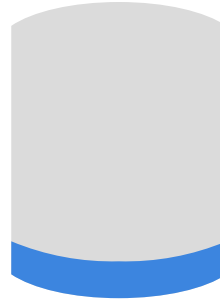
72%

A modified copy of production
(masked, subsetting, etc.)



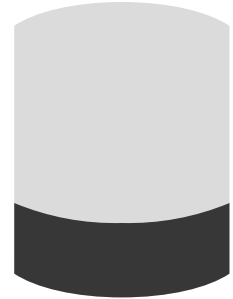
50%

Synthesized



13%

Manually created



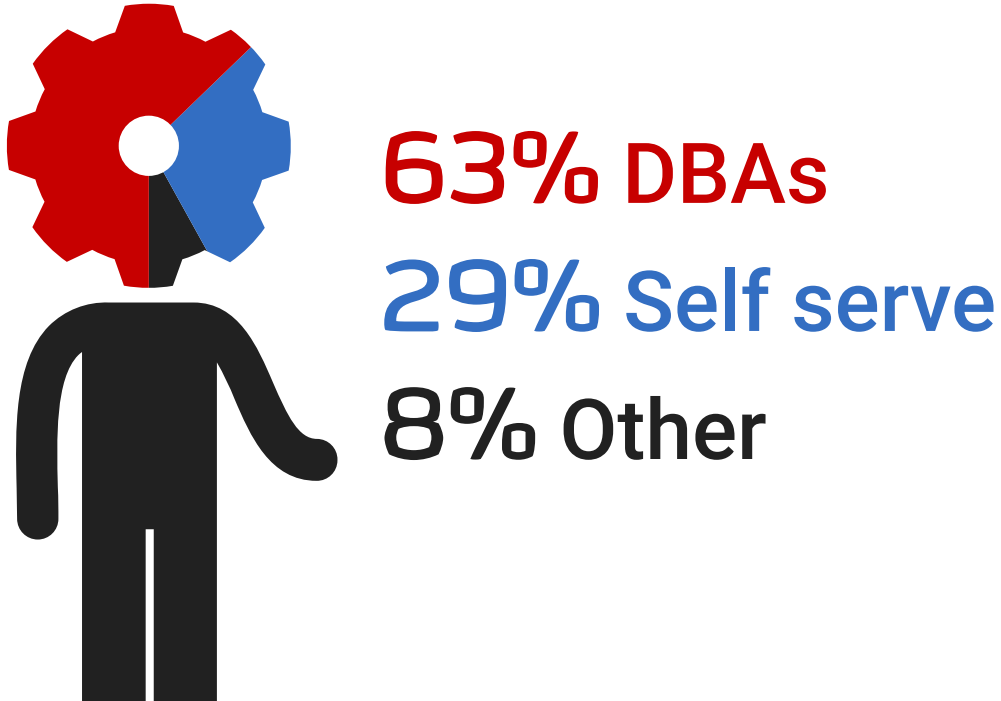
28%

Surprisingly, 72% of respondents work on data that comes directly from production in some instances in non-production environments. Astonishingly, 36% of respondents always use production data in non-production environments. This is most common for those that work in the media (50%), consulting (44%), and insurance (44%) industries. Surprisingly, 30% of healthcare organizations and 30% of government organizations that responded also always use data from production for development/test/QA purposes.

Only 28% of respondents never use a direct copy of production. This was most common in healthcare organizations (44%), followed by finance (29%) and insurance (26%).

3

Who controls the provisioning of databases?



The DBA controls the provisioning of databases in 63% of organizations, yet 29% of organizations allow teams to self-serve. We suspect that all those who can allow self-service do. Removing the need for ticketing, and thereby friction and dead time at the start of a development cycle, is self-evidently desirable.

On the other hand, the specifics of database provisioning (file moves, swapping permission sets, managing disk space) may mean that the barriers to enabling self-service are high.

For those that selected 'other', the provisioning was often owned by the DBA and a senior developer.

4

What is the highest frequency at which non-production databases are refreshed?



Daily

28%



Weekly

26%



Monthly

19%



Quarterly

12%



Other

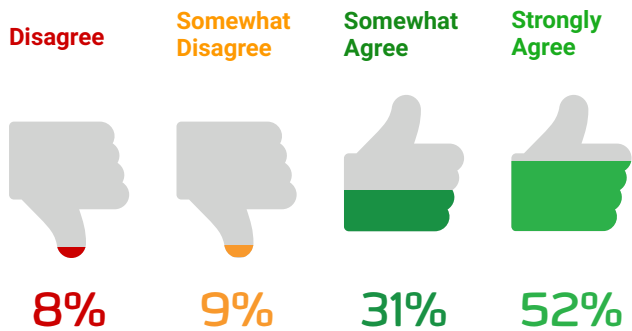
15%

A lot of provisioning of non-production databases is done as requested or on demand, however 28% of respondents provision as frequently as daily while 54% provision weekly or more. Surprisingly, those who self-serve provision at a similar frequency to those who have provisioning managed by a DBA. Among those who provision daily, 28% provision databases 1TB or larger.

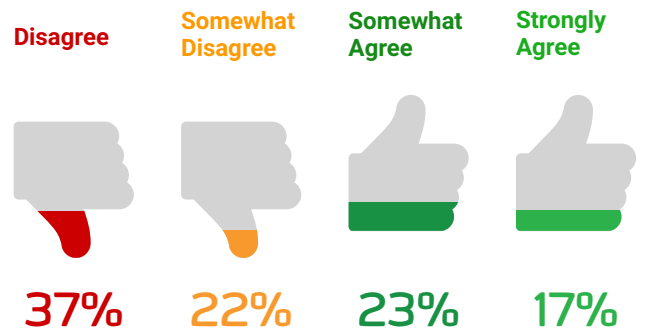
5

Do you agree or disagree with the following?

It is desirable to use Production data in Dev and Test



I am restricted from using Production data in dev and test

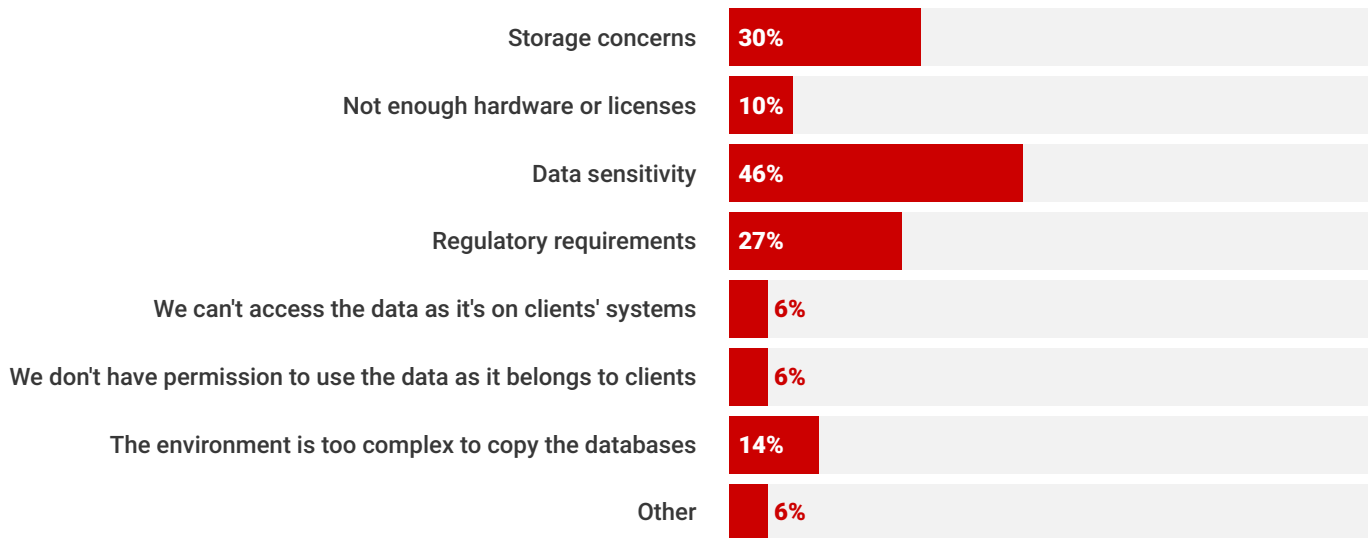


83% of respondents agree or strongly agree that it is desirable to use production data in development and testing. 76% of DBAs that responded agree with this as do 87% of Application Developers. This suggests Application Developers find having production data in development and test more desirable than DBAs.

40% of respondents are restricted from using production data in dev and test. 46% of those restricted do not use data in development/test/QA environments that has come directly from production.

6

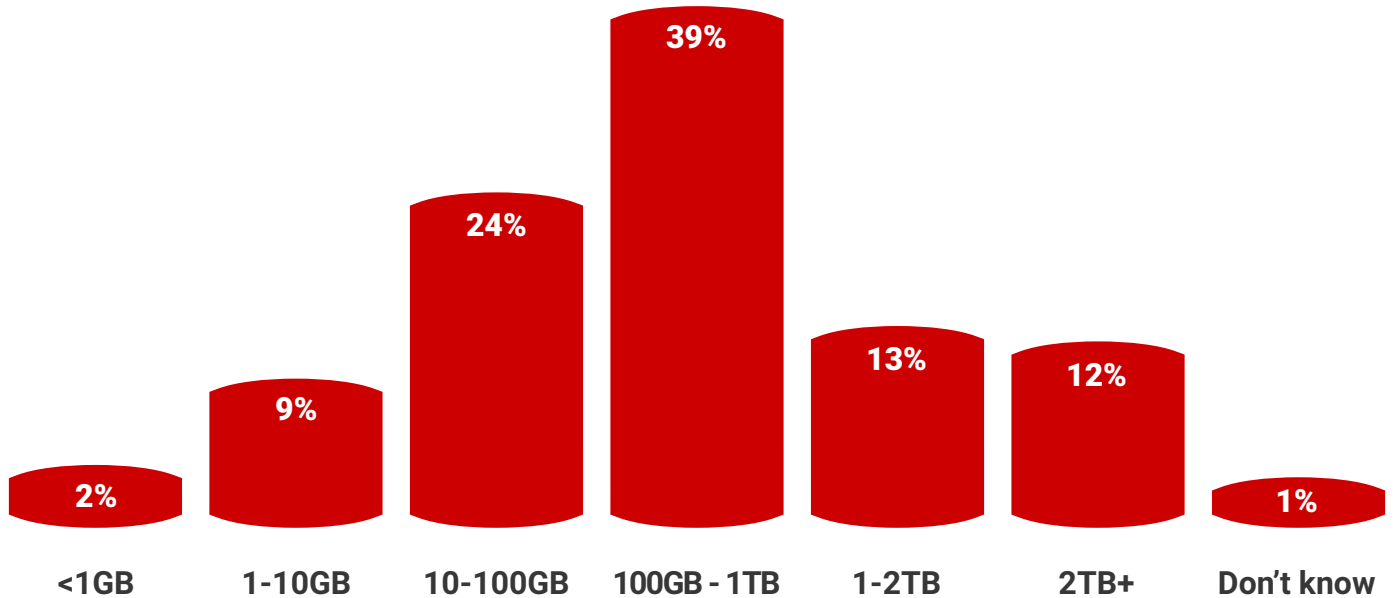
If you're restricted from using production data, why is that? (multiple choice)



Unsurprisingly, the main reason for restricting use of production data is due to data sensitivity and regulatory requirements. This reasoning was most frequent among finance (63%), government (62%), and healthcare (61%) organizations. With the GDPR, HIPAA, and other industry regulations, organizations are having to take more caution with how data is used and who has access to it.

7

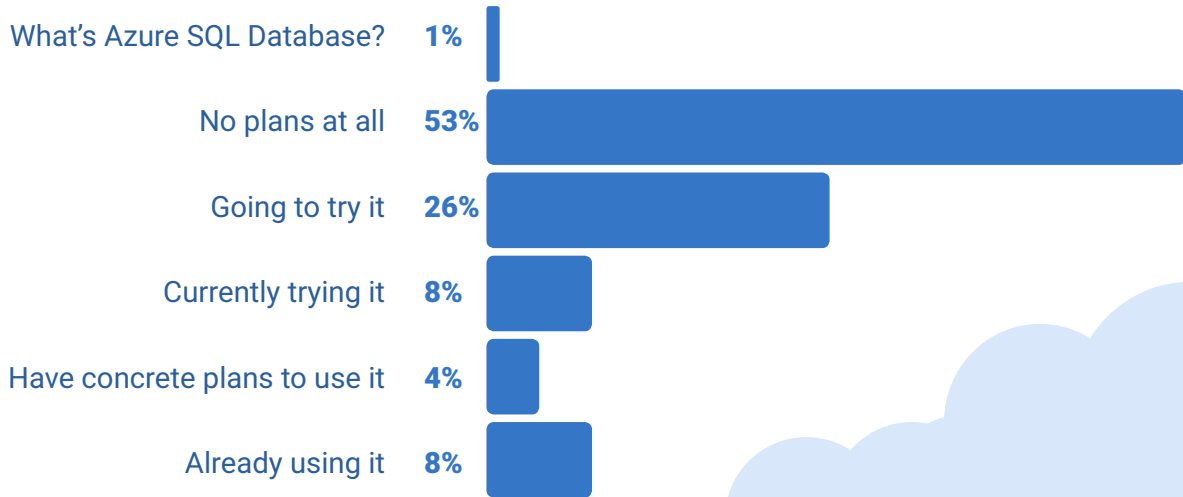
What size is the biggest database in your dev or test environments?



Storage concerns were also a key reason that use of production data is restricted. There are multiple reasons why this may be, including the cost of storage, the time it takes to manage storage, processes around requisitions, or not having enough storage. As expected, the majority of organizations have databases in development or test between 100GB and 1TB. Only 26% have databases 1TB or larger in development or test. Among those restricted from using production data due to storage concerns only, 41% have databases in development or test 1TB or larger and 36% of those provision daily.

8

What are your plans for Azure SQL Database in the next 12 months?

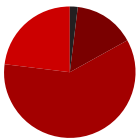


The adoption of Azure SQL Database is relatively low with only 8% of respondents using it and 4% with concrete plans to use it in the next 12 months. A large proportion are still looking at trying it. However, over 52% have no plans for Azure in the next 12 months. Of those organizations with storage concerns, 39% of respondents are currently using or trying Azure or have concrete plans to use it in the next 12 months. For those whose only concern was storage, 56% still had no plans to try out Azure SQL Database.

9

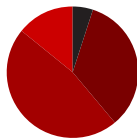
How important are the following?

I can quickly refresh databases in non-Production environments



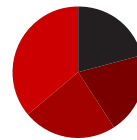
Not important 2%
Somewhat important 15%
Very important 60%
Critical 23%

The data is up-to-date in non-Production environments



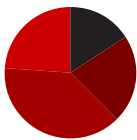
Not important 5%
Somewhat important 34%
Very important 47%
Critical 14%

Sensitive data is masked or removed from Production copies



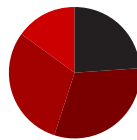
Not important 21%
Somewhat important 20%
Very important 23%
Critical 36%

Permissions are replaced in the Production copy



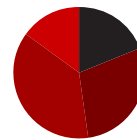
Not important 16%
Somewhat important 22%
Very important 39%
Critical 24%

Operational features (e.g. replication) are removed in copies



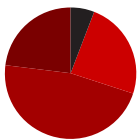
Not important 24%
Somewhat important 31%
Very important 30%
Critical 15%

Developers (and others) can self-serve copies of Production databases



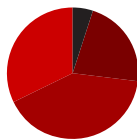
Not important 19%
Somewhat important 29%
Very important 37%
Critical 15%

Production databases can be copied on a schedule



Not important 6%
Somewhat important 24%
Very important 47%
Critical 23%

Access to provision and refresh databases copies is controlled

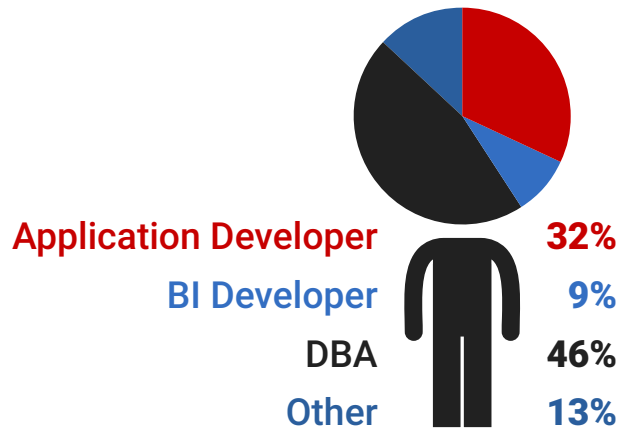


Not important 5%
Somewhat important 22%
Very important 41%
Critical 32%

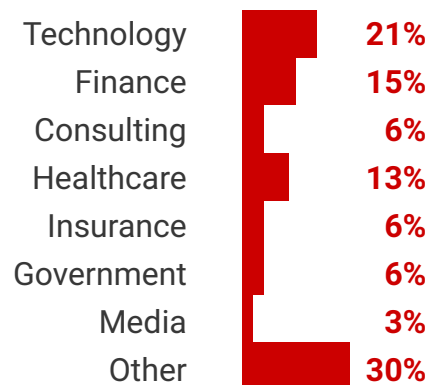
The most important provisioning operation to respondents was the ability to quickly refresh databases in non-production environments, with 83% rating this as very important or critical. Although the speed at which databases can be refreshed came out on top, surprisingly the ability for developers to self-serve is much lower with only 52% rating this as very important or critical. 19% felt this was not important at all. This indicates many organizations want some control over who can provision databases. This is reinforced by 73% of respondents rating controlling access to provision and refresh databases as very important or critical.

Survey Demographics

Job title



Industry



To learn more about how Redgate can help your organization overcome some of the challenges of provisioning SQL Server databases visit

www.red-gate.com/sql-clone

