# Data Imaging for Data Governance and Regulatory Compliance

Enterprise applications often rely on more than one database, but there are no accepted methods today for versioning and managing complex data environments. Windocks introduced the use of SQL Server containers combined with database cloning, enabling delivery of complex Terabyte class data environments in seconds. This combination also provides unique support for versioned, auditable data images that enhance data governance and compliance.

Data Imaging for Enterprise Data Environments

Data images are built using Full or Differential SQL Server backups, snapshots, and SQL Server incremental log shipping, combined with SQL Server scripts to implement data masking during the image build. The resulting image is a full byte copy of the data in the form of Virtual Disks that can span multiple physical (or virtual) disks and large data sets.

The Virtual Disk, in turn, supports creation of Windows "differencing disks" which are writable clones. Clones are delivered in seconds, and only require 40 MB or less of storage.
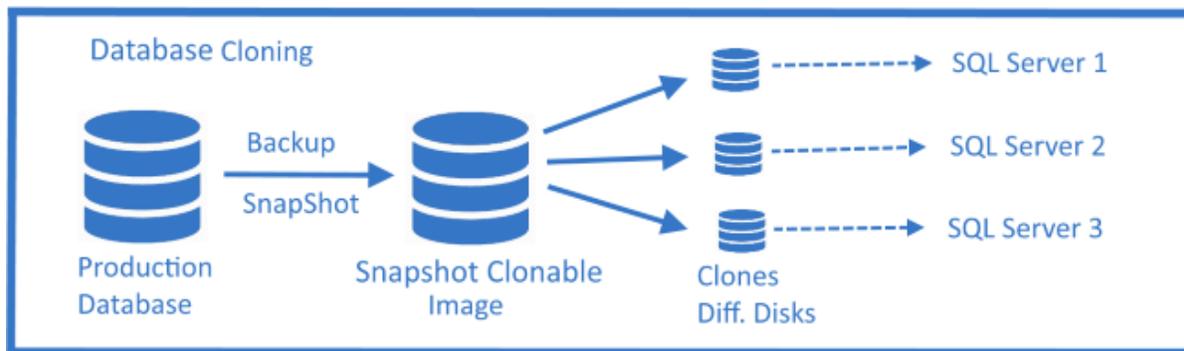
Images are built with a Dockerfile that specifies the location of backups, snapshots, or log shipping updates, and SQL Server scripts. Windocks images support multiple databases, with source files located on the Windocks host or a network attached file. In the example below the Dockerfile specifies two databases, located on network attached file shares, and SQL Server scripts for data masking.

```
                              dockerfile - Notepad
File  Edit  Format  View  Help
FROM mssql-2012
SETUPCLONING FULL customers \\test1\dbbackups\customerdatafull.bak
SETUPCLONING FULL operations \\test1\dbbackups\operationsfull.bak
COPY cleanseData.sql .
RUN cleanseData.sql
```

The resulting image is versioned and auditable, and supports delivery of multi-terabyte environments in seconds for Development and Test, and for reporting and BI. These data environments can now be delivered automatically, or provisioned by users, for use with any SQL Server container (both Windocks and Microsoft's), as well as with conventional SQL Server instances.
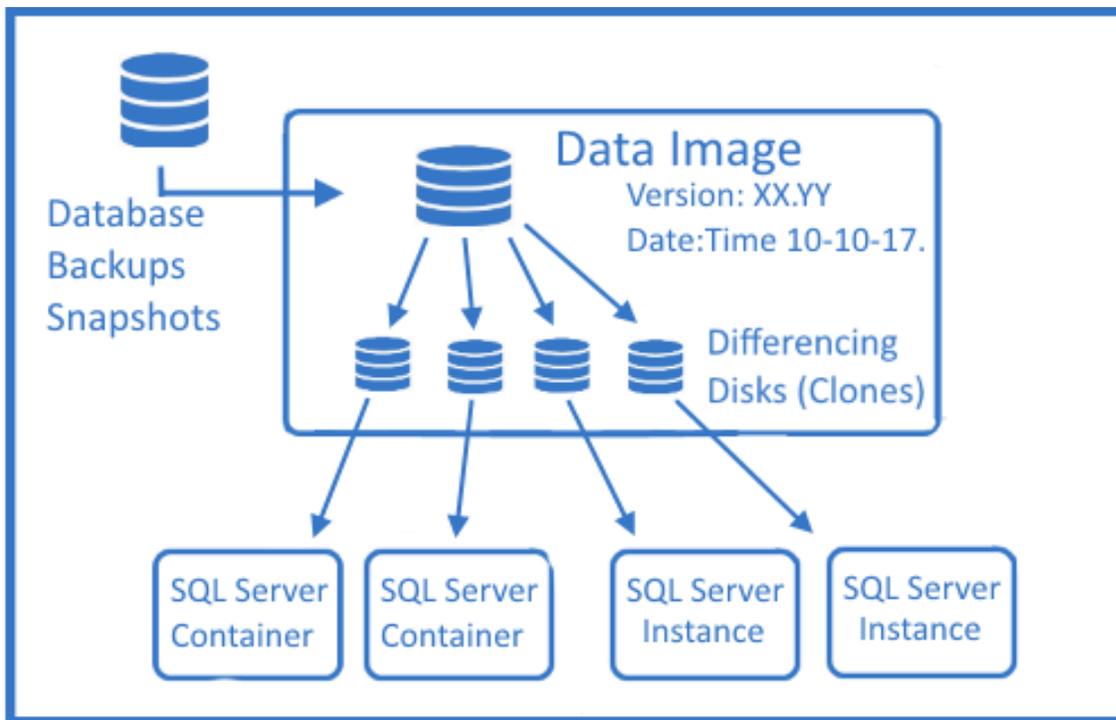
Data Imaging enhances Governance and Compliance

The design as described addresses the needs for data environments for Dev/Test and Reporting needs, but the new Data Image repository is ideally suited for expanding Data Governance and Regulatory Compliance needs.

Privacy/Security: security is improved with data delivered through a structured container process. Adhoc access to enterprise data can be curtailed, and approved and auditable images used to support dev and test, as well as reporting and BI needs. Privacy is enhanced as data masking that is implemented during the image build. National boundaries are respected with image registries hosted in the appropriate country, as the Windocks solution runs wherever Windows servers are supported (on-premise, private, or public cloud).

Quality: the container workflow enhances data quality and consistency by supporting use of production databases as the authoritative source of data. The workflow outlined above supports native Jenkins or Team City servers for Continuous Integration, making this approach unique as the first full stack Jenkins Continuous Integration solution that incorporates production database support.

Access and Use: Docker containers are emerging as the defacto standard for software development and test. Containers play a prominent role in Microsoft's strategies for Windows Server 2016 and SQL Server 2017. The approach outlined here provides organizations with an on-ramp to Docker based workflows on Windows Server 2012 and Server 2016, with support of all editions of SQL Server 2008 onward. Not only does this workflow improve access to data (on-demand and in seconds), with the latest preferred dev and test tooling, but it also is uniquely useful for SQL Server reporting and BI purposes. This workflow integrates with existing backup and DR system infrastructure, making it uniquely easy to add to existing systems and infrastructure.

Open: as a result of customer feedback Windocks is also expanding support for delivery of data environments from any Storage Area Network (SAN), from NetApp, EqualLogic, and others. Copy Data Management systems will also be supported, such as from Cohesity and Rubrik. Finally, support for MySQL, DB2, and other environments will also be added as requested by customers.

Conclusions

Data governance should not be an after-thought for modern software development and delivery strategies. Windocks combination of SQL Server containers with database cloning delivers benefits for development and test, reporting and BI, and enhances data governance and policy compliance. The solution installs with existing systems and delivers Terabyte class data environments in seconds, while creating immutable, versioned and auditable images that address many data governance needs.

Explore how Windocks can enhance your data governance and delivery with a free Windocks Community Edition. Download your copy here.