

Summary:

The pace of business is accelerating, with increasing demand for agile application development, and business analytics. The problem is, enterprise data is stuck in legacy storage systems. We lack cloud native infrastructure to streamline access, while simultaneously addressing needs for Data Governance and Regulatory Compliance.

Windocks is helping with a modern solution that includes security, management, and real time data delivery.

- Combines Docker's container technology with database cloning for modern approach, aligned with Microsoft and industry standards
- Immutable, auditable images for Data Governance and Regulatory Compliance
- Integrate with existing CI and storage systems, combined with Windows based data cloning
- Runs on public cloud or on premises
- Dramatically simpler and more affordable than competitive solutions.

"Windocks provisions SQL Server containers with a 3.5 TB NetApp hosted snapshot, on demand in 30 seconds. It's been a game changer for us!"

Enterprise Data Strategies

Today the largest personal transportation company owns no cars, and the largest hospitality provider owns no real estate. These businesses represent a new model of software defined business, with agile application development being center to their success.

The need for Digital Transformation is real, but CIOs and Enterprise Architects face the challenge of doing so without expanding IT budgets. As one CIO recently shared:

"I need to cut our effective spending, by getting more done with what we already own."

Enabling responsive application development to business and customer needs is imperative, but is also hindered by poor access to relational data. Gartner Group projects 70% of new projects will include relational data through 2020 (Gartner Operational Data Management Magic Quadrant, 2017). Unfortunately, little progress has been made in making enterprise data more accessible.

In fact, there is a growing chasm between the rapid application development achieved with Docker containers for .NET and Java front-end applications, and the glacial flow of enterprise data.

Dell shared industry surveys that show that over 80% of enterprises refresh databases for Dev/Test twice monthly or less. This stands in sharp contrast to the life-cycle of Docker containers, which is often measured in hours.

Strategic Questions

The consensus among IT leaders is that new initiatives should focus on cloud-native architectures. The question becomes how to implement strategies that address enterprise data management and delivery:

- 1) Will data services rely on storage systems or software defined solutions?
- 2) Can Data Governance and Regulatory Compliance be “shifted left” to meet GDPR and other requirements?
- 3) Should data ops be bi-modal or can cloudnative and legacy ops be combined with a software based solution?

Storage Arrays or Software Based Data Services?

The move toward Software Defined Data Centers is a continuation of the industry trend toward virtualization and cloud native architectures. The objective is to provide a framework for automating use of infrastructure, enabling higher utilization, reduced management costs, and improve responsiveness to users.

The problem is that we’ve seen poor progress in software defined storage.

Storage systems continue to rely on the Solaris UNIX operating system, with the ZFS based file system.

Investments today are being made in upgrading older Storage Arrays with all flash storage. These systems dramatically improve throughput, but offer little relief for data delivery for application development, and aren’t usable in public clouds. The systems are also under-utilized due to complex scripting involved in provisioning LUNS, Volumes, and mount points.

A new generation of storage appliance shows promise for greater data agility, with secondary storage solutions from Cohesity, Rubrik, and others. These systems simplify data access with restful API support, but still present increased costs and ongoing storage administration. Public cloud use is possible with some, using a UNIX software appliance VM.

Fortunately, software based options are now available, that deliver database cloning on Windows, and simplify integration and use of SAN based snapshots. Windows servers with database cloning support solutions that are infrastructure agnostic (public cloud or on premise), and provide DBAs and Developers with self-service tools, for a solution that supports both cloud-native and legacy projects. The new generation of software based data delivery also offers promise of simplifying Data Governance and Compliance needs.

Integrating Data Governance and Regulatory Compliance into Data Delivery

Compliance is a complex topic, ranging from government regulations such as Sarbanes-Oxley and the European Union Data Protection Act, to industry regulations such as PCI DSS

for payment cards and HIPAA for health data. Moving to a cloud platform can involve reduced control over certain aspects of security.

A primary concern is data location. The EU Data Protection Act (GDPR), strives to keep personal information inside of the European Union. This requirement can be challenging for organizations with geographically distributed development and test teams, who push for local access to data for their work.

Encryption plays an important role in varied regulatory environments. HIPAA, for example, requires that all user data, both moving and at rest, be encrypted.

To complicate matters, standards are evolving. Bringing visibility to users is a major goal of the **Cloud Security Alliance**, a three-year-old organization that appeals to auditors and service providers. A major goal of the CSA is development of standardized auditing frameworks to facilitate communications between users and cloud vendors.

Efforts of the CSA and other alliances, plus those of industry groups and government agencies, should enhance standards in the next several years. The CSA has formal alliances with ISO, ITU and NIST.

As the needs for enhanced Data Governance and Regulatory Compliance

grows, so will pressure to “shift left” and incorporate these standards into enterprise business practices. One opportunity is to evaluate how a Data strategy can support the creation of versioned, auditable data environments used within the organization.

Bi-Modal or a Unified Data Operations?

Gartner Group has brought the conundrum facing IT leaders into sharp contrast with the bimodal approach to IT management. Bi-modal approach makes intuitive sense to address conflicting needs for reliability and accuracy needed for legacy systems, contrasted against projects that call for agility and speed. The underlying premise is that conflicting needs cannot be efficiently addressed with existing tooling.

The bi-modal approach describes many operations today, with public clouds often supporting one set of infrastructure and processes and a different approach used on premise.

Enterprise data management is naturally biased in favor of bi-modal approaches, as primary Storage Arrays support on premise use only. Public clouds provide separate storage and delivery systems. Some SAN vendors repackage appliance to run as “cloud native” VMs. This repackaging, however, does little to address the costs of ownership and administration required for these systems.

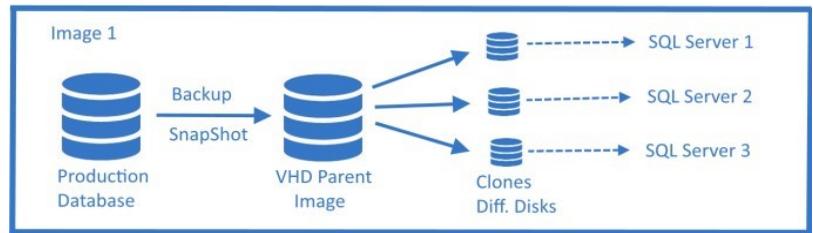
The question is whether a new generation of software based systems can provide a unifying layer of data management and cloning. Such a solution could unify cloud and on premise data management and delivery strategies.

The Case for Data Virtualization (database cloning)

Enterprises need improved access to data, in a form that spans both on premise and cloud use, with Data Governance and Regulatory Compliance that is “built in.” The current best practice for data delivery, involves the use of database virtualization (database cloning).

Database virtualization is supported in all Storage Array Network systems, and has been widely use for the past decade. Fast snapshots deliver clones that provide full read/write operation, with minimal storage, and are delivered in seconds.

Database clones or snapshots involve the creation of a snapshot, or virtual hard drive, which can includes a source database or can also be created through the use of database backups. The Virtual Hard Drive (VHD or snapshot) is a full byte copy of the environment, and supports fast delivery of clones which support full read/write operations for use by application development, reporting and BI.



As mentioned earlier Storage Array Network systems have supported fast snapshots for years. Use of the snapshots is complicated with vendor specific commands, and complex scripting involved in dealing with block level storage primitives (LUNs, volumes, mount points). The industry has added Docker volume plugin support, but the underlying complexity remains and appears to block broad usage.

Another problem associated with storage systems is the lack of support for data management, as these systems are managing blocks on varied storage media. Meta data is not available with these systems, so there is no apparent approach to address Data Governance or Regulatory Compliance with Storage Appliance focused strategies.

A new generation of storage appliances is coming to market from Cohesity, Rubrik, and others that simplify data access and support application native metadata. These systems provide restful APIs for data access, and can support application native data snapshots, allowing for richer metadata and the potential for improved Data Governance. These also typically support use of software appliances for use in public clouds.

In addition to a new generation of storage appliance, software based solutions for database cloning and delivery are now available that run on industry standard operating systems. Windocks combines use of Docker containers and database cloning that delivers versioned, auditable images stored in the image repository.

Legacy and Modern Data Delivery

Enterprise architects are challenged by a fragmented storage industry that lacks standards for interoperability. Where progress has been made with software defined networking standards, relatively little progress has been made in corresponding storage standards.

- Storage systems are dedicated appliances running various UNIX, and require dedicated storage admins.
- Re-packaged storage appliances into cloud-ready VMs does little to solve cost of ownership and administration challenges.
- Storage systems consist of LUNs and volumes, and lack application and user metadata. The result is poor audit support for Data Governance and Regulatory Compliance needs.
- Interoperability between storage systems is poor, and snapshots are underutilized due to storage complexity and ongoing management.

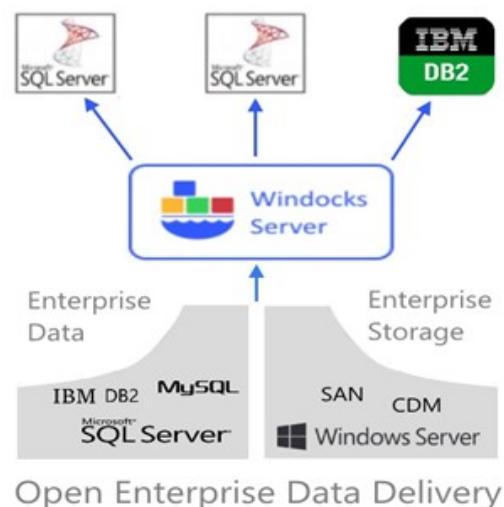
In contrast, a modern approach to storage is software based and runs on industry standard operating systems, delivering an open, affordable, easily managed system.

Finally, a modern solution provides interoperability with existing enterprise infrastructure, including Storage Arrays and Secondary Copy systems, and supports emerging standards like Docker containers. A modern software solution that runs on Windows or Linux should simplify the use of SAN hosted snapshots, enabling these systems to be used to support modern application initiatives.

Open

A modern, open software solution integrates with existing SAN infrastructure, simplifies use of snapshots, and support delivery of database clones throughout an enterprise. Organizations that have grown through acquisition and merger, can extend the useful life of mixed storage systems with the Windocks San-ready containers.

An open solution also supports emerging infrastructure standards, such as Docker containers, and should support delivery of database clones into multi-tiered application environments for Continuous Integration and Delivery.



Finally, an open and modern solution is designed to deliver end-user self-service of data environments, for Developers, Testers, and automated processes involved in Continuous Integration systems. Solutions should also support a full range of data environments, including SQL Server, MySQL, and other enterprise databases.

Security and Data Governance

Enterprise data strategies are hindered today by incomplete support for data versioning and security. Organizations need versioning support for complete environments, not just for individual databases. Images need to incorporate data masking during the image build, not through an ad hoc steps.

Windocks delivers on these needs, with a single image build step that delivers versioned environments that can include one or multiple databases, with data masking and user permissions applied during the image build. The result is a versioned, immutable, auditable data image that is stored in an enterprise data repository.

Such a system effectively “shifts left” Security and Data Governance into the default enterprise workflow for defining,

managing, and delivering data environments.

The result supports the most demanding audit requirements throughout the life cycle of enterprise data.

Data images support real time usage reporting in dashboards, and granular reporting on users and groups and images accessed. Data Images are immutable, stored in industry standard artifact repositories.

Finally, data images can also be prepared specifically for third party access, supporting secure data delivery and access for trading partners and others outside of the enterprise.

Scalability and Total Cost of Ownership

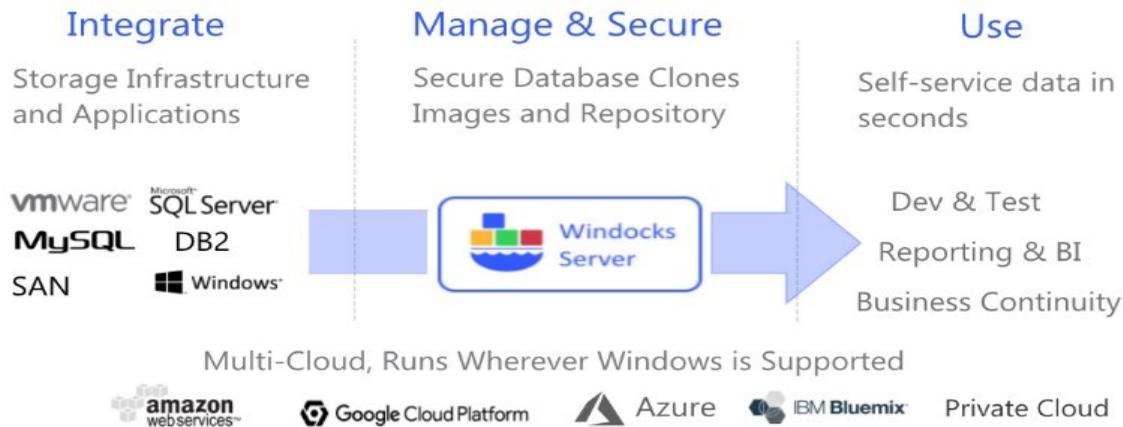
Enterprise scalability is commonly measured in terms of “scale out” or “scale up” system design, but should also include consideration for economical support of small teams within an enterprise. Cost of ownership is a function of Administration, integration, and licensing and pricing.

Windocks: a modern, open, data delivery platform

Windocks is a modern, open enterprise data delivery platform, that combines the use of containers and database clones for a cost effective, scalable solution for a enterprise relational database use-cases.

Windocks works wherever Windows Servers are supported, including any public cloud or on premise infrastructure. Uses include dynamic support for application development and test, Data Governance and

Enterprise Data Management, Security, and Delivery



Regulatory Compliance, DR testing and Recovery, Reporting and BI, Continuous Integration, Datacenter Consolidation, and Hybrid Cloud support.

Windocks delivers a rich set of benefits:

- Delivery of Terabyte databases in seconds
- End user self-service of environments
- DBA driven workflow does away with storage admin bottlenecks
- Instantly available SQL Server clones for efficient development, and production system support
- Simplified dev and test operations with an average 5:1 reduction in VMs used
- Flexibility to use SAN and CDM snapshots, or Windows clones

- Deliver data to Microsoft's SQL containers, and other environments
- Multi-tier support with .NET or Java, with Jenkins, Team City, and other CI pipelines
- Unmatched affordability with per user support at just over \$15/month

Development, Test, and QA

Windocks was first released in April of 2016, and is popular for support of application development, test, and QA. Hundreds of organizations rely on database cloning for delivery of database environments for application development and test.

Most use production databases (with data masking and user permissions applied) in the form of full and differential backups, or database snapshots. Backups are restored to a Virtual Hard Drive, with SQL Server scripts applied for data masking or other preparation.

The VHD becomes a full byte copy of the prepared environment. Clones utilize a Copy on Write method to

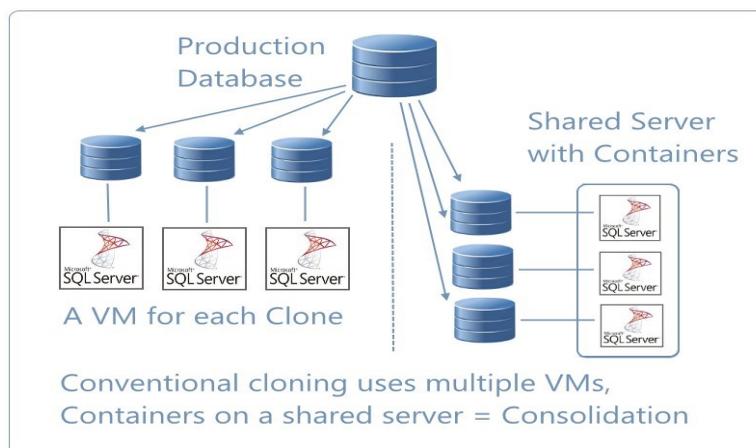
capture changes, and are functionally equivalent to snapshots provided by SANs and storage appliances. One advantage of Windows based cloning is meta data captured during image build, which is key to Data Governance and Regulatory Compliance strategies.

Database clones are served to any SQL Server instance, including Docker containers or conventional instances.

Docker containers are popular for orchestration of multi-tier environments combining .NET with SQL Server, for Jenkins, Team City and other Continuous Integration pipelines.

Windocks is also infrastructure agnostic, running wherever Windows Servers are supported, including on premise or public cloud. Windocks also creates new options for data delivery by offering SAN ready integration, simplifying the use of snapshots on older NetApp, EqualLogic, and other SAN systems. Windocks provides out of the box integration solutions, and provides automated cleanup of mount points and volumes.

Teams simplify development and test operations by working with isolated containers on a shared server, reducing the number of VMs used. Teams reduce use of VMs by 5:1 when working with containers.



Data Governance and Regulatory Compliance

A compelling aspects of Windocks is how Data Governance and Regulatory Compliance are built-in to the default workflow. Data images are immutable, versioned, and auditable artifacts, with rich metadata to support reporting and audibility of data images and their use.

Data images bring order to ad hoc use of enterprise data. Today, standard tooling supports database level versioning, and lacks support for integrated life cycle management. Windocks creates versioned data images that are complete, and can include scores of databases, with support for data masking scripts applied during the image build. The result is an industry-first complex data image, and auditable image repository.

Windocks delivers GDPR compliance by building customer privacy into data images with data masking and user/group permissions built into the default workflow, satisfying the requirement of “privacy by default, and by design.”

Organizations use data images to implement geographically defined repositories, with versioned and auditable images. With Windocks organizations can fulfill important GDPR requirements, while enhancing the efficiency and delivery of data within an organization and to external partners.

Windocks is also a modern, open design supporting enterprise metadata management systems for data quality, information catalogs, and Regulatory Compliance systems.

DR Testing and Recovery

Whether applications are hosted on a cloud or on premise Windocks supports native database backups and snapshots to create instantly available data environments. Windocks also supports automated incremental image updates through SQL Server log shipping updates, providing effective DR support in the event of a production system failure.

In the event of a production system failure, the latest image can be delivered and mounted to a replacement system within 1 minute. This provides many organizations with a new approach to High Availability and Disaster Recovery.

Windocks is also being used for DR system testing. Database replication can be

paused and an image created and used to test the viability of the replica. This can be automated, and accomplished in minutes, and is particularly popular for inter-cloud system replication and failover designs.

Hybrid Cloud Agility

Windocks simplifies the use of data on the cloud for Dev and Test, while ensuring user and group level access control to complete images, and improving overall release cycles.

Windocks data images are ultimately portable, and solve the long-standing problems of system configuration differences and “DLL hell” that have prevented developers from easily sharing work with larger teams. Windocks images will run reliably wherever the appropriate Windows Servers are located.

The process begins with a backup or snapshot taken from a production system, along with incremental updates, which is used to build a Windocks data image on an adjacent nonproduction system. SQL Server scripts are applied during the image build to ensure the data is properly prepared, including user permissions and data masking.

Cloud based systems located adjacent to production simplifies data transport, and is particularly popular for support of geographically distributed teams. On demand, teams selfprovision prepared environments in seconds. By working with isolated containers on a shared server teams simplify their operations, and reduce their use of VMs by 5:1 or more.

Reporting and BI

Reporting and Business Intelligence is another major use case for database clones, for immediate and timely access

to data for business insights and decision making.

The process of utilizing database clones for Reporting and BI follows the same workflow as described earlier for Dev and Test support. Images are built using native backups, and scripts are applied to the data to address user and group permissions, and data masking. Once built, the images support self-service provisioning of environments in seconds.

In the case of reporting and BI, however, the use of containers provides additional benefits with scale out support for SQL Server Report Server and SQL Server Analysis Server (SSRS and SSAS). In addition to provisioning large databases in seconds, Windocks can also automate the delivery of scale-out SSRS and SSAS instances. The combination of fast database delivery combined with elastic support for SSRS and SSAS is a particular benefit for SQL Server based Reporting and BI initiatives.

Continuous Integration

For many organizations, the ultimate goal of fast provisioning of databases for application development, is to incorporate the back-end into a Jenkins, TeamCity, or other CI pipeline.

Multitier application environments are delivered on demand, in seconds, on a shared server. CI pipelines are simplified with smaller, more affordable clusters, with dramatic improvement in speed and efficiency.

Datacenter Consolidation

While Datacenter Consolidation is not a use case per se, Windocks provides a uniquely powerful form of automated infrastructure that delivers significant benefits for increased data center utilization and consolidation.

Windocks is an implementation of Docker's source for Windows, and as such provides full stack support for .NET, Java, SQL Server, and other application stacks. Windocks is a high performance, lightweight container implementation that delivers increased server utilization. Windocks customers have regularly support 30 or more SQL Server containers on a single 8 CPU core host, realizing a corresponding 30:1 reduction in the number of VMs used.

In addition to increased server utilization, Windocks delivers dramatic savings in both network bandwidth and storage consumption. Data images are built and used in place, and Windocks database clones require only 40 MB of storage on delivery, reducing storage consumption by 99% or more.

About Windocks:

Windocks delivers virtualized database environments for organizations around the globe, enabling organizations to access, manage, and protect data faster, more efficiently, and simply than existing data systems. Database cloning combined with Docker based containers enables modern software development and delivery, and reporting. Windocks is the first open, modern platform for enterprise data delivery.

For additional information, visit Windocks at www.windocks.com, or contact Windocks at info@windocks.com.

Explore the use of SQL Server containers and database clones with a free Windocks Community Edition.

<https://www.windocks.com/community-dockerwindows>



References:

The Gartner Group Magic Quadrant for Operational Database Management Systems is available courtesy of Microsoft at:

<https://info.microsoft.com/gartner-odbms-magic-quadrantregister.html>

Dell's research is available here:

<https://www.quest.com/whitepaper/the-current-state-andadoption-of-devops8116848/>

<https://cloudsecurityalliance.org/>

<https://windocks.com/files/DataImagingEnhancedDataGovernanceCompliance.pdf>

<https://windocks.com/files/GetStartedWithSQLServerContainersAndDatabaseClones.pdf>