

THE ESSENTIAL GUIDE TO

Storage for DevOps



Tintri

Contents

- 1 In Brief
- 2 What is DevOps?
- 3 When Storage Selection Goes Wrong...
- 4 What Matters?
- 5 Copy Data Management
- 7 How Copy Data Management Helps DevOps
- 8 Data Protection and Disaster Recovery
- 9 Should Your DR Environment Also Be Your Development Environment?
- 10 Quality of Service and Guaranteed Performance
- 11 Should Development and Production Share the Same Storage?
- 12 Monitoring and Troubleshooting
- 13 Automation
- 14 When Storage Selection Goes Right...
- 15 Create a Successful DevOps Strategy

In Brief

Many companies today are adopting a DevOps model to accelerate development efforts and deliver new applications and services. Choosing the right enterprise cloud storage provides the foundation to support your growing DevOps practice.

Organizations of all types and sizes are undergoing a digital transformation to better engage with customers. This transformation often means accelerating development, entering new markets, adding to product differentiation, and more. Increasingly, companies are adopting a DevOps methodology to achieve these goals. A recent survey suggests that DevOps can help you deploy code faster while reducing failure rates and outages.

When building your DevOps environment, you need to create a enterprise cloud infrastructure that is simple, flexible, and automated. Traditional storage systems have often proven to be the wrong choice for DevOps because they are not able to deliver these benefits.

Recently, a number of storage innovations have emerged that better suit the requirements of DevOps. By paying attention to a few selection criteria, you can identify storage that will not only meet your needs and budget, but simplify your DevOps environment.

This guide helps you understand the storage features that are most beneficial to your DevOps practice and provides specific guidelines on what to look for.

What is DevOps? Before you start thinking about enterprise cloud storage, it may be helpful to define what is meant by “DevOps” as it is used in this book.



Here are a couple of popular definitions:

“DevOps is a culture, movement or practice that emphasizes the collaboration and communication of both software developers and other information-technology (IT) professionals while automating the process of software delivery and infrastructure changes.”

–Wikipedia

“DevOps is a development methodology with a set of practices aimed at bridging the gap between Development and Operations, emphasizing communication and collaboration, continuous integration, quality assurance, and delivery with automated deployment.”

–Jabbari, et al¹

The first definition emphasizes the need to automate infrastructure changes in addition to software delivery. The ability to easily automate storage functions can have profound benefits for DevOps.

“Bridging the gap” that exists in traditional IT is prominent in the second definition. Development and IT operations people may have very different ideas about infrastructure. Reaching a consensus on what’s important in storage selection requires some education on both sides.

DEVELOPER PERSPECTIVE ON STORAGE

The typical developer doesn’t spend a lot of time thinking about storage. Developer priorities facilitated by storage include:

- Keeping test data up to date with production.
- Easy access to capacity as needed.
- Programmability.

OPERATIONS PERSPECTIVE ON STORAGE

IT operations teams have significant experience with storage. Priorities include:

- Accelerating release cycles.
- Availability and reliability.
- Features that streamline data management.

¹ Jabbari et. al., What is DevOps?: A Systematic Mapping Study on Definitions and Practices, 2016.

When Storage Selection Goes Wrong... The wrong storage fails to deliver the necessary performance, scaling, and copy management, increases overall costs, and quickly becomes an impediment to DevOps processes.

Traditional Storage Hobbles Cybersecurity Efforts

A government cybersecurity laboratory tests capabilities and trains personnel to prevent and defend against network intrusions.

The traditional approach to provisioning storage with LUNs and volumes did not scale quickly enough. The laboratory needed the ability to provision hundreds or even thousands of VMs in hours rather than weeks and to tear them down just as fast when training exercises completed.

Because of the dynamic environment, projects sometimes had to scale on short notice. Rolling out additional resources took days and overall performance deteriorated under the additional load.



PROBLEMS:

Existing storage was unable to scale quickly enough or guarantee performance for important projects and training exercises.

Financial Services Company Struggles with Storage CapEx, OpEx, and Data Copies

A Canadian financial services company relied on conventional storage for all its operations including development. Upfront capital costs and annual support costs were extremely high, while daily management tasks drove up operating costs.

Like many financial services companies, this firm is moving to a DevOps model to support its digital transformation. As a result, they needed to be able to provision capacity and performance quickly to support development efforts.

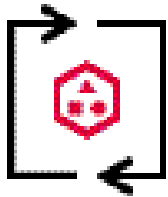
The DevOps team also needed multiple up-to-date copies of production data to support development and testing efforts. However, making a single copy took five hours and maintaining many copies was consuming a lot of expensive capacity.



PROBLEMS:

Existing storage had high CapEx and OpEx and was unable to provide up-to-date data copies to support development and testing efforts.

What matters? Storage can make a huge difference to the success of your DevOps efforts. If you have a chance to influence your storage buying decisions, there are five capabilities you should specify:



COPY DATA MANAGEMENT

Provisioning and maintaining up-to-date data copies results in higher quality software.



DATA PROTECTION AND DR

Both the production and development sides of your DevOps environment must be protected.



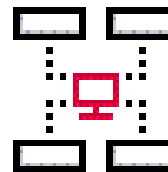
QUALITY OF SERVICE AND PERFORMANCE GUARANTEES

I/O performance for critical VMs has to be assured without overprovisioning.



MONITORING AND TROUBLESHOOTING

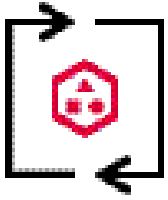
Problems are best avoided, but when a problem occurs the right tools are needed to find and fix it fast.



AUTOMATION

Storage that's easily automated simplifies the workflows needed in a DevOps environment.

Each of these topics is explored in more detail on the pages that follow →



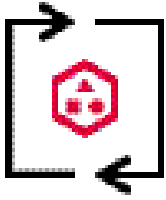
Copy Data Management As DevOps teams work to accelerate the delivery of new software features, they need up-to-date copies of production data to support development, integration testing, QA, and other essential tasks.

When development and test engineers roll out new versions of software, they need test data sets that have the same size and complexity as the production data set used by the application. In practice, synthetic test data is almost never an adequate substitute. Most teams make do with too few copies of production data, and those are often weeks or months out of date.

This is because the traditional approach to provisioning data copies is cumbersome, time-consuming, and difficult to manage. Physical copies have a big impact on storage systems. Full copies consume a large amount of expensive storage space and take hours to create. Because of the potential performance impact to production storage systems supporting live applications, it's rarely possible to create a new copy when it's needed, and provisioning a new set of copies can consume many man-hours of effort.

What Is Copy Data Management?

Copy data management allows a DevOps team to have the up-to-date data copies they need, when they need them. Instead of relying on multiple physical copies, copy data management methods create “virtual” copies, eliminating the data duplication and time needed for physical copies while removing the load on production storage and protecting the performance of production applications.



Copy Data Management



How Does Storage Enable Copy Data Management?

Modern storage systems provide capabilities such as space-efficient snapshots, cloning, and replication that can facilitate the creation of data copies. While there are differences between various vendors, these technologies typically work in the following ways:

- Space-efficient snapshots result in point-in-time data copies that only consume additional storage space as changes are made (aside from incremental storage needed for metadata).
- Cloning allows you to create writable “virtual” copies of a snapshot. You can create as many clones as you need and once again additional storage space is only consumed as changes are made.
- Replication is based on periodic snapshots. Once an initial copy is created, only the changes that have occurred since the last snapshot need to be replicated, saving time and bandwidth. This also significantly lowers the burden on production storage systems versus full data copies.
- Some storage companies and storage partners use these core technologies to automate the process of copy creation and copy data management.

Choosing Storage for Copy Data Management

- Assess your environment to understand the total number of copies you will require and the frequency with which they need to be refreshed.
- If new storage is needed, at a minimum select storage with snapshot, cloning, and replication features as described above combined with ease of automation.
- Ideally, look for storage vendors that understand DevOps requirements and that further facilitate the process of copy data management either directly or in conjunction with a partner.
- Choose storage that allows for visibility and action on individual VMs if possible. LUN-based storage architectures that predate virtualization may not be able to provide the right level of granularity and are therefore much less efficient.
- Seek advanced capabilities such as the ability to access any recovery point and move between recovery points without data loss.

How Copy Data Management Helps DevOps

TIMELINES

AGILITY

RECOVERY



Developer requirements

Build apps with latest production data to avoid rework

Distribute updates immediately across Dev/Test and QA

Loss of file, VM, or vDisk does not cause a shutdown



Operations requirements

Avoid environment rebuilds and accelerate release cycles

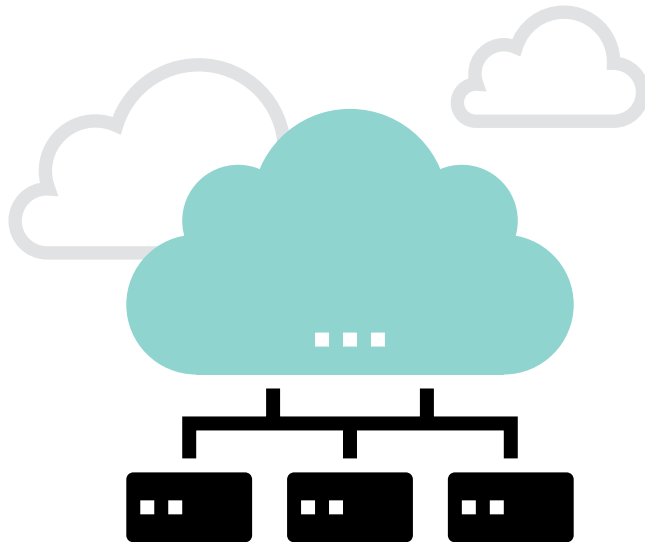
Automate rapid test cycles across hundreds of servers

Single click point-in-time restore



Data Protection and Disaster Recovery As DevOps becomes central to your business success, you have to rethink the level of protection it requires. Fortunately, the same storage tools that facilitate copy data management can protect both the production and development sides of your DevOps environment and facilitate a multiplicity of tasks.

For companies undergoing a digital transformation, DevOps can be a major contributor to your competitive advantage and business success. Of course you're going to protect your production environment, but now ensuring the productivity of your DevOps team, even in the face of disaster, also becomes a significant concern. In the past, if a development or test environment became corrupted, it was typical to just recreate it from scratch. Software builds could just be rerun. But DevOps requires greater immediacy. It may be time to rethink policies carried over from your company's earlier development practices.



Applying Storage Tools to Protect and Manage Your DevOps Environment

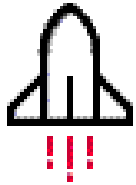
Fortunately, the same storage capabilities that support copy data management (described in the preceding section) can also serve as the foundation of a data protection strategy for all aspects of your DevOps environment, while satisfying other data management requirements as well.

- ♦ **Snapshots.** Regular snapshots of production VMs, developer and tester workspaces, and other important data provide a first line of defense against failures, mistakes, and data corruption.
- ♦ **Replication.** Periodic replication can be used to replicate your entire environment—or just the most critical parts—to a secondary location.
- ♦ **Cloning.** Cloning is extremely useful after a failure. A failed VM can be quickly restored, in either a primary or secondary location, by cloning a backup snapshot. Cloning makes a variety of DevOps functions simpler, such as creating multiple copies of test data sets, creating identical work environments for developers, or making a test environment identical to your production environment.

Choosing the right storage should not only make these tools available, it should make them programmable so that your DevOps team can incorporate them directly into their tools and daily workflows.

Should Your DR Environment Also Be Your Development Environment?

Midsized and larger organizations already tend to have a significant investment in DR facilities. Increasingly, these facilities are being co-opted for DevOps functions. The production environment is replicated regularly to protect it, providing up-to-date data that can be cloned as often as needed.



Quality of Service and Guaranteed Performance All aspects of your DevOps environment—production, development, and testing—can generate significant I/O. You have to be able to guarantee performance for critical VMs while ensuring that other workloads aren't starved for resources.

Quality of Service (QoS) and the ability to guarantee performance for particular workloads is a significant concern to organizations adopting DevOps. In a typical DevOps environment, there's often a production database that is critical to the completion of customer transactions. If that database is starved for I/O, everything that depends on it is going to slow down and customers may abandon transactions or leave the site before completion.

On the development side, you might also need to guarantee performance of critical functions. Software builds can take a long time to complete and you don't want restricted I/O to extend the process. Testing environments need to generate I/O loads and transaction rates similar to that seen in production.



What Is Storage QoS?

Storage QoS provides a mechanism to let you control how the performance of a storage system is allocated to different workloads. There are three factors to consider:

- **Ceiling.** Most QoS implementations allow you to set a ceiling on the number of IOPS or the maximum bandwidth that can be consumed.
- **Floor.** Much less commonly, QoS implementations let you guarantee a minimum number of IOPS or bandwidth that will be available. This guarantees a minimum level of performance is reserved for a particular workload.
- **Granularity.** Most QoS implementations work at the granularity of the LUN or volume. This is not always ideal. Setting QoS on a LUN containing 100 VMs might not be that meaningful; you may have to organize your storage carefully to get a real advantage from QoS under these circumstances. Ability to configure QoS at the VM or vDisk level is preferred.

Choosing Storage for QoS

The baseline requirement for storage QoS is the ability to set a ceiling. With that you can effectively guarantee performance for your most important workloads by limiting other workloads. The ability to set a floor or minimum guaranteed performance, and to be able to do so at the VM level, will greatly simplify QoS use.

Should Development and Production Share the Same Storage?

In the past, it was typical for production to have its own dedicated storage to ensure maximum application performance. Today, with the power of all-flash storage systems potentially coupled with QoS settings on individual VMs, it may be time to give that practice a second look. There are obvious advantages to consolidating DevOps on the same storage, including easier access to production data sets for testing and a smaller storage footprint.



Monitoring and Troubleshooting When you move to a DevOps environment with continuous integration and continuous deployment, you need the ability to monitor your infrastructure and correct problems and misconfigurations quickly.

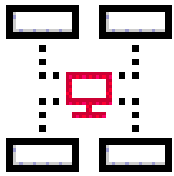
Visibility is one of the keys to DevOps success. Monitoring is necessary to make sure there aren't any incipient problems with applications or infrastructure. Because things can change quickly, misconfigurations are bound to occur and must be easily identified and corrected. And you should be able to see resource usage and plan for additions before the situation becomes critical. The right monitoring and management tools can enable your DevOps team to manage more storage with less time spent and less hassle, reducing the need for dedicated storage expertise.



Choosing Monitoring and Troubleshooting Features

Look for storage that provides the following monitoring features:

- ♦ **Basic monitoring.** At a minimum, the storage you choose should provide monitoring and alerting for all its core capabilities so you can easily verify that it is operating correctly.
- ♦ **Advanced monitoring.** More advanced capabilities, such as the ability to monitor storage performance in relation to networks and servers, can greatly simplify troubleshooting.
- ♦ **Predictive analytics.** The ability to forecast your future capacity and performance needs so you know exactly when to add resources can be extremely helpful.
- ♦ **Integration.** Storage should integrate with other monitoring tools in your environment through SNMP and/or API calls. The ability to access monitoring information through API calls is particularly desirable for DevOps.
- ♦ **VM-level monitoring.** Most storage monitoring tools operate at the LUN or volume level. Visibility at the VM level can simplify interpretation and speed up troubleshooting.



Automation A key requirement for DevOps is the ability to control all infrastructure through programming. Storage for DevOps must enable a high degree of automation.

As your organization moves from accommodating occasional software releases to a continuous integration model where software releases happen weekly, daily, or even more frequently, the number of environments needed for testing and validation grows rapidly. It's no longer practical, or even possible, to do this work manually. To be successful, everything must be automated.

The term “infrastructure-as-code” describes the ability to provision and manage IT infrastructure programmatically rather than physically or with interactive tools. Storage automation has a number of advantages:

- It lets you easily automate the process of creating and breaking down many environments as necessary.
- It lets you incorporate advanced storage features such as snapshots and cloning as part of your daily workflows.
- It eliminates the potential for error that results from manual or interactive configuration.

How Automation Benefits DevOps

DEVELOPERS

- The combination of automation and copy data management makes routine tasks simpler and faster, saving time
- You don't have to wait for operations to get resources
- You can incorporate storage capabilities as part of applications, reducing development effort

OPERATIONS

- Automation enables infrastructure self-service for:
 - Automated workflows
 - Ad hoc requirements
- Automation eliminates repetitive tasks, freeing you to add more business value

What to Look for in Storage Automation

The key to automation in the DevOps and cloud world is REST APIs. At a minimum, look for a full set of REST APIs to allow you to automate basic storage provisioning functions. Additional APIs may address advanced functions mentioned in this eBook: snapshots, clones, replication, QoS, monitoring, etc.

Storage infrastructure that's easy to orchestrate can be a big advantage. VM-level operations will simplify the automation process.

When Storage Selection Goes Right... The right storage can help your DevOps initiative succeed. By making informed storage choices, the organizations described on page 6 overcame their storage problems.

Cybersecurity Laboratory Overcomes Scaling Limitations, Guarantees Performance

When its existing storage failed to meet scalability and performance needs, this government cybersecurity facility went looking for storage that could keep up.

Where it used to take up to two weeks to deploy the hundreds of VMs needed for a customer event, the task can now be completed in minutes.

The new storage provides much better performance so it doesn't bog down when hundreds of VMs are deployed, and it provides QoS to guarantee performance at the level of the individual VM.

Additional benefits include a 50% reduction in footprint, reduced power and cooling loads, and reduced training requirements.



SUCCESS FACTORS:

Rapid spin-up of hundreds of VMs, guaranteed performance, ease of management.

Financial Services Firm Implements Data Copy Management, Reduces CapEx and OpEx

This IT team went looking for better storage alternatives because its expenses were too high and it was unable to easily provision data copies.

Where it used to take five hours to copy production data, the task can now be completed in five minutes, giving the team its weekends back. They can now create as many up-to-date copies as they need to support DevOps efforts.

The team chose an all-flash array for development. Advanced data reduction decreases the total storage capacity required. Space savings went from 1.7x to 4.3x; average latencies dropped to 0.3ms, and third-party backups ran 30% faster. Despite all these improvements, overall cost was less than half that of the previous storage solution.



SUCCESS FACTORS:

Copy data management, superior performance and data reduction. Lower upfront CapEx. Ease of management for lower OpEx.

Create a Successful DevOps Strategy The advice in this book should help you narrow the field of potential storage vendors. As you compare the options in terms of functionality and cost, keep the following guidelines in mind:



COPY DATA MANAGEMENT

The right storage will provide or support tools that enable you to streamline the process of copying production data for use by your developers and testers. No other single storage feature will return greater benefit in terms of streamlining your operations and increasing the quality of your code.



DATA PROTECTION AND DR

The same storage tools that facilitate copy data management can also help you integrate data protection for both production and development as part of your DevOps environment. Through automation, data protection can become part of the environment.



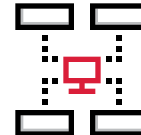
QoS

Quality of service functionality will let you ensure that critical VMs—both on the production and development side—get the I/O performance they need.



MONITORING AND TROUBLESHOOTING

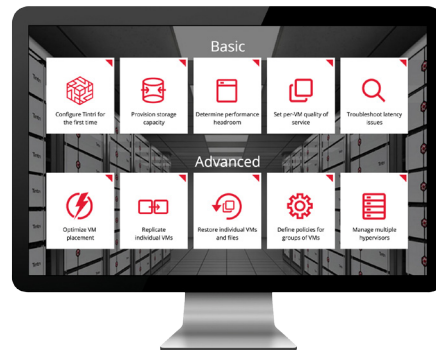
Storage expertise on your DevOps team may be limited. Storage should provide tools that simplify monitoring and management and accelerate tools. Predictive analytics tools can help with capacity planning.



AUTOMATION

DevOps moves too quickly for manual processes and handoffs from one team to another. You must be able to automate storage and other infrastructure to support self-service and to allow the creation of working and test environments as part of your routine workflow.

Thanks for reading!



We hope the Essential Guide to Storage for DevOps got you thinking.

Want to know exactly how Tintri can save you time and money? Then schedule a 30-minute demo to get your questions answered and see us in action. Just complete the form at:

tintri.com

Tintri VM-aware storage is the simplest for virtualized applications and cloud. Organizations including United Healthcare, NASA, Toyota and 7 of the Fortune 15 have said “No to LUNs”. With Tintri they manage only virtual machines, in a fraction of the footprint and at far lower cost than conventional storage.



For more information, visit www.tintri.com and follow us on Twitter @Tintri