



Technical FAQ

## JFrog Artifactory on NetApp for DevOps

### JFrog Artifactory Integration with ONTAP and StorageGRID Webscale

Bikash Roy Choudhury, NetApp  
November 2017

In partnership with



#### Abstract

JFrog Artifactory is the first, and only, universal Artifact Repository Manager on the market, and fully supports software packages created by any language or technology. JFrog Artifactory is the only enterprise-ready repository manager available that supports secure, clustered, High Availability Docker registries; and integrates with all major CI/CD and DevOps tools to provide an end-to-end, automated and bullet-proof solution for tracking artifacts from development to production.

NetApp® ONTAP® software provides a standard data management platform during the CI/CD workflow to store and protect data for JFrog Artifactory and other development tools that are used during the application development and delivery process. ONTAP also provides data availability, high performance, and scalability throughout the SDLC.

NetApp StorageGRID® Webscale is an object store offering that helps accelerate application development. It enables users at remote sites to collaborate for distributed file sharing of artifacts and binary packages that are managed by JFrog Artifactory.

## TABLE OF CONTENTS

<b>1 JFrog Artifactory and NetApp ONTAP</b> .....	<b>3</b>
1.1 Setup and Configuration .....	4
Which JFrog Artifactory components are used during setup? .....	4
Can JFrog Artifactory be installed on a virtual machine (VM) or a container? .....	4
Can JFrog Artifactory be configured in high-availability (HA) mode? .....	4
Where does ONTAP integrate into in the JFrog Artifactory setup? .....	5
Why is ONTAP important for JFrog Artifactory? .....	5
1.2 Data Protection and Recovery .....	5
Why is data protection important for JFrog Artifactory? .....	5
Does JFrog Artifactory have any native data backup and restore mechanism? .....	5
Why should I use NetApp ONTAP for backups with JFrog Artifactory? .....	5
How would JFrog Artifactory recover data if a file is accidentally deleted or lost? .....	5
How does ONTAP help with storage efficiency during data protection? .....	6
How does JFrog Artifactory perform data protection operations on ONTAP? .....	6
Can SnapCenter create consistent Snapshot copies for JFrog Artifactory database and repository volumes on ONTAP? .....	6
Do developers who use JFrog Artifactory need to know about the SnapCenter tool from NetApp? .....	6
What databases does SnapCenter support? .....	6
<b>2 JFrog Artifactory and NetApp StorageGRID Webscale</b> .....	<b>6</b>
2.1 Collaboration .....	7
How does StorageGRID Webscale enable data collaboration? .....	7
Does JFrog Artifactory have native data collaboration? .....	7
How does StorageGRID Webscale complement JFrog Artifactory? .....	7
How does JFrog Artifactory connect to two different endpoints—ONTAP in the CI/CD workflow and StorageGRID Webscale for data replication? .....	7
Where can I find the configuration file in JFrog Artifactory that manages the storage endpoints, such as ONTAP and StorageGRID Webscale, that access data over NFS and S3? .....	8
Does StorageGRID Webscale support identity and access management (IAM)? .....	8
Are policies that are configured in JFrog Artifactory enforced on the objects in StorageGRID Webscale? .....	8

## LIST OF FIGURES

Figure 1) JFrog Artifactory in CI/CD workflow on ONTAP .....	3
Figure 2) Anatomy of a binary artifact. ....	3
Figure 3) Replicate and share artifacts and binaries across remote sites with Artifactory and StorageGRID Webscale .....	7

# 1 JFrog Artifactory and NetApp ONTAP

JFrog Artifactory is a universal binary and artifact manager that is used in the continuous integration (CI) / continuous delivery (CD) workflow in the DevOps process. JFrog Artifactory expedites application delivery and enables faster software releases.

Figure 1 illustrates the CI/CD workflow that defines the requirements before the application is released into production. The entire workflow runs on NetApp ONTAP software, which provides a standard data management platform for a persistent datastore to tools and workloads that are leveraged throughout the software development lifecycle (SDLC). JFrog Artifactory plays an integral role to store and manage the prebuilt artifacts and binary packages that are essential in the application development and delivery process. For more information, see [TR-4547](#), [TR-4596](#), and [TR-4607](#).

Figure 1) JFrog Artifactory in CI/CD workflow on ONTAP.

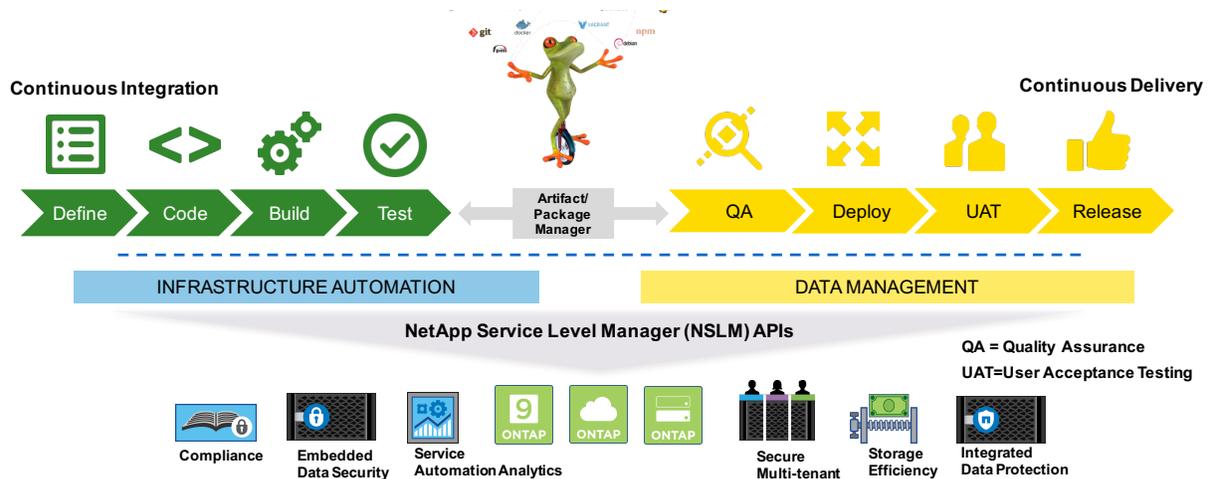
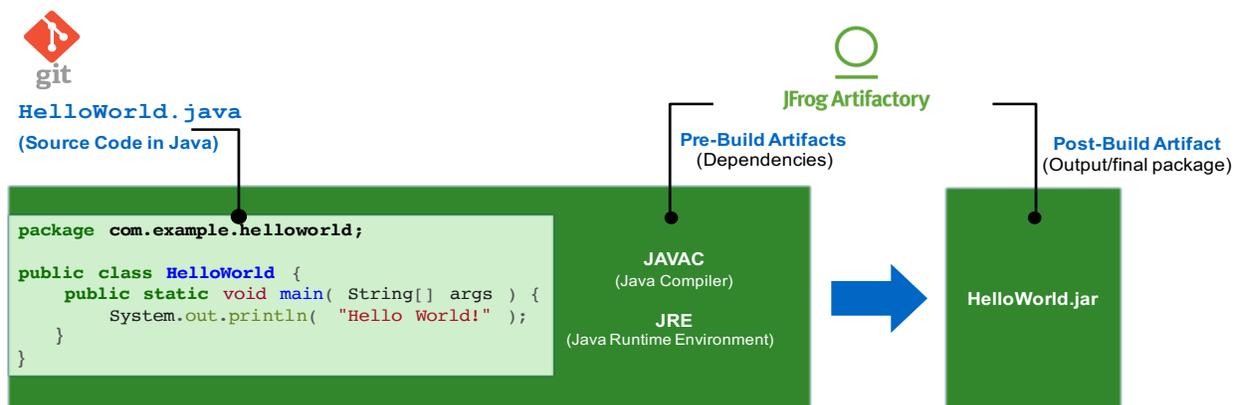


Figure 2 distinguishes Java source code from the binary repository and also provides the anatomy of a binary artifact. During a Java application development process, the different versions of source code (shown in Figure 2) are managed by a source-control management tool such as Git. Source code requires different prebuilt artifacts, such as the Java compiler, software development kits, and other library files.

Figure 2) Anatomy of a binary artifact.



The Java source code uses these prebuilt artifacts to compile the code and generates an output file such as `Hello.jar`. The `.jar` file, also known as Java Archive (JAR), is packaged with other dependencies and bundled as a `.war` file, also known as Web Archive (WAR). This `Hello.war` file is known as a binary package. JFrog Artifactory manages the different versions of the prebuilt artifacts and binary packages for various programming languages and repositories.

## 1.1 Setup and Configuration

### Which JFrog Artifactory components are used during setup?

Answer: JFrog Artifactory has two components in its setup. One component is a database that stores the metadata information about artifacts, builds, and binary packages, and the other component is a repository that stores the files as checksums.

### Can JFrog Artifactory be installed on a virtual machine (VM) or a container?

Answer: Yes, JFrog Artifactory can be installed on a VM and can run as a container. The `docker-compose.yml` file that can be used to install JFrog Artifactory in Docker containers is available in [TR-4607](#). The containers mount persistent data storage from ONTAP over NFSv3 by using NetApp Docker Volume Plugin to manage the provisioning of storage, which is then used to set up and configure Artifactory.

### Can JFrog Artifactory be configured in high-availability (HA) mode?

Answer: Yes, JFrog Artifactory can be configured in HA mode. In the following example, JFrog Artifactory is configured by using HA mode in a container setup with NGINX to load-balance the two Artifactory instances.

```
c736bc40931b      jfrog-docker-reg2.bintray.io/jfrog/artifactory-ha-
primary:5.1.4    "/bin/sh -c /runAr..." 4 days ago      Up 4 days
10042/tcp, 0.0.0.0:8082->8081/tcp, 0.0.0.0:10018->10017/tcp
root_artifactory-node2_1

5ed3815dba14      jfrog-docker-reg2.bintray.io/jfrog/artifactory-ha-
primary:5.1.4    "/bin/sh -c /runAr..." 4 days ago      Up 4 days
0.0.0.0:8081->8081/tcp, 0.0.0.0:10017->10017/tcp, 10042/tcp
root_artifactory-node1_1

c0a5c31fb6e1      jfrog-docker-reg2.bintray.io/jfrog/nginx-art:5.1.4
"/bin/tini -g -- /..." 4 days ago      Up 4 days      0.0.0.0:80-
>80/tcp, 0.0.0.0:443->443/tcp, 0.0.0.0:5000-5010->5000-5010/tcp
root_nginx_1
```

Starting with Artifactory 5.1.x, JFrog has introduced *sharding*, which allows Artifactory in HA mode to write to two local storage disks. JFrog also modified the database to store the details of the shard in both locations. As node 1 in an HA configuration writes the copy of the checksum in the first local disk under `shared-fs-1`, an internal remote process within Artifactory triggers to write to the `shared-fs-1` in the second local node. Finally, the internal process notifies the second Artifactory node in the cluster. This process generates redundant copies of the same checksum on two local storage disks. This JFrog design accounts for the fact that not all checksums are used all the time, and, based on the policy, these artifacts or checksums can be moved to Amazon S3.

When Artifactory identifies the checksums that are not accessed for a long time, it creates a `tmp` folder called `EVENTUAL`. It then deletes the bucket or folder under `shard-fs-1` in both of the storage local disks and triggers a move of the checksum into the S3 endpoint. Artifactory takes this approach because S3 does not understand sharding. The database updates the information about the checksum in its new location. When the checksum is requested, it is pulled from the S3 endpoint into `shared-fs-1`, and the process is handled internally by Artifactory.

When we configure this setup on ONTAP, where the home and data directories are configured over NFSv3, only one copy is in the `shard-fs-1` directory, and the second Artifactory node is notified about the copy. This approach provides high storage efficiency along with compaction and compression benefits for checksums in the cache and in the repository directories, respectively.

### **Where does ONTAP integrate into in the JFrog Artifactory setup?**

Answer: ONTAP provides persistent data storage for the database and for the repository components of JFrog Artifactory. The Artifactory database and repository are installed on a shared file system such as NFSv3 that mounts a NetApp FlexVol® volume to a VM, a physical machine, or a container.

### **Why is ONTAP important for JFrog Artifactory?**

Answer: ONTAP provides a standard data management platform in the continuous integration (CI) and continuous delivery (CD) workflow in which JFrog Artifactory plays a central role between the two processes. JFrog Artifactory manages multiple versions of prebuilt artifacts, builds, and binary packages for multiple repositories.

ONTAP complements JFrog Artifactory and benefits customers in the following ways:

- Automated data management for the CI/CD workflow for both private and public cloud
- Data protection for instant file, database, or repository recovery
- Cost-efficient backups that allow more binary packages and a smaller data footprint
- Faster application delivery through integrations with Git and Jenkins for CI on ONTAP

## **1.2 Data Protection and Recovery**

### **Why is data protection important for JFrog Artifactory?**

Answer: JFrog Artifactory manages various versions of prebuilt artifacts, builds, and binary packages. Data protection allows faster and reliable recovery from any data loss or any corruption from disasters. Data can be backed up to tape or to the cloud over S3. NetApp ONTAP complements JFrog Artifactory by providing better, faster, and cost-efficient methods for protecting data.

### **Does JFrog Artifactory have any native data backup and restore mechanism?**

Answer: Yes, JFrog Artifactory provides cron-based backup management services, along with a system-level import/export feature to back up and recover data.

### **Why should I use NetApp ONTAP for backups with JFrog Artifactory?**

Answer: Ideally, the prebuilt artifacts and binary packages are stored as checksums in the different Artifactory repositories. There could be more than one repository pointing to one checksum in the production environment. When any backup service is initiated on local storage, the artifact stores multiple copies in all the repositories that referred to it. This approach does not provide storage efficiency, thus increasing the storage costs.

Validation of a sample JFrog Artifactory dataset in a test environment indicated that if the production size of Artifactory is 10GB, the backup on local storage amounts to approximately 18GB. Thus, more space is consumed in every backup operation. JFrog Artifactory's native backup operations are also time-consuming.

### **How would JFrog Artifactory recover data if a file is accidentally deleted or lost?**

Answer: The file can be recovered by using the import feature in JFrog Artifactory. However, while recovering a single file, Artifactory also copies multiple other dependent files during the recovery process from the backup folder, thus increasing the data footprint in the production environment.

### **How does ONTAP help with storage efficiency during data protection?**

Answer: JFrog Artifactory is configured over a shared file system such as NFS on ONTAP. ONTAP uses native technologies such as NetApp FlexVol volumes and Snapshot™ copies to store and manage Artifactory databases and repositories in the CI/CD workflow. ONTAP provides a high degree of storage efficiency from data compaction and inline compression for the Artifactory data volumes in the production environment.

ONTAP Snapshot copies are thin-provisioned and pointer-based, which makes them space-efficient. JFrog Artifactory uses ONTAP Snapshot copies for data protection for the database and the repository, with no additional storage space overhead.

### **How does JFrog Artifactory perform data protection operations on ONTAP?**

Answer: JFrog Artifactory uses a Windows-based tool, NetApp SnapCenter® software, which can perform application-consistent backups on ONTAP. As part of the initial configuration, SnapCenter discovers the storage and identifies the resource groups to perform data protection. It also installs the custom agent in the database instance that is responsible for creating a consistent Snapshot copy without any disruption or manual process.

### **Can SnapCenter create consistent Snapshot copies for JFrog Artifactory database and repository volumes on ONTAP?**

Answer: Yes, SnapCenter creates a consistency group for JFrog Artifactory database and repository volumes on ONTAP. With the help of the database agent, SnapCenter can quiesce the database, flush all the buffers and logs, and then create a consistent Snapshot copy of both the database and the repository volumes.

JFrog Artifactory updates its database (metadata for the repositories) only after the update to the repository has been fully committed to persistent storage. JFrog Artifactory and ONTAP complement each other to provide consistency at the file system and the data layers.

### **Do developers who use JFrog Artifactory need to know about the SnapCenter tool from NetApp?**

Answer: Developers have more autonomy to create application-consistent Snapshot copies for the changes in JFrog Artifactory by using a plug-in. Developers can execute this plug-in from the CLI to call SnapCenter RESTful APIs to protect the changes in the Artifactory data volumes. Developers do not need to have any knowledge of the SnapCenter tool.

### **What databases does SnapCenter support?**

Answer: SnapCenter supports Oracle, SQL Server, MySQL, MongoDB, and DB2. You can use any of these choices as the pluggable database to connect to JFrog Artifactory.

## **2 JFrog Artifactory and NetApp StorageGRID Webscale**

JFrog Artifactory enables faster application delivery as a universal binary package manager throughout the CI/CD workflow. Business owners want to share and collaborate on artifacts and binaries that are generated by developers across their remote sites, globally. Therefore, a change to an artifact or a binary package at one site should replicate instantly to remote sites so that users have access to the latest builds and binary packages.

NetApp StorageGRID Webscale is a hyperscale global object store and a software-defined platform for a content repository of mission-critical applications. StorageGRID Webscale consists of a single namespace with a set of gateway, storage, and archive nodes at different sites in different geographic locations. For example, an organization might have three major data centers across the world, in the United States, Europe, and Asia.

## 2.1 Collaboration

### How does StorageGRID Webscale enable data collaboration?

Answer: StorageGRID Webscale connects multiple geographically dispersed sites into a single namespace that can be managed by a single admin node. Every remote site has a gateway and storage nodes. StorageGRID Webscale uses a content management system to asynchronously replicate data to multiple locations that are connected in a single grid. Data ingest or modification in one location triggers an instant copy of the change to all the locations on the grid. The speed of replication depends on the WAN speed.

### Does JFrog Artifactory have native data collaboration?

Answer: Yes, JFrog Artifactory uses the proxy feature for remote repositories. Remote users can search and list information about artifacts, builds, and binary packages from remote sites, but the data is pulled on demand from the central repository. Data is not prefetched to the remote sites for local users. Fetching data on demand to remote sites impairs performance, availability, and the flexibility to share files in a distributed manner.

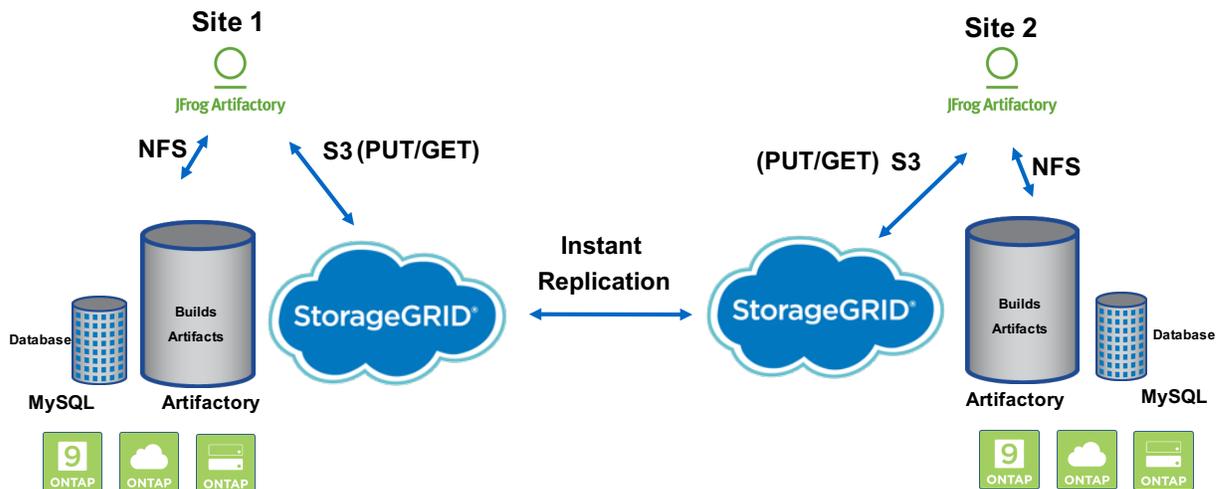
### How does StorageGRID Webscale complement JFrog Artifactory?

Answer: JFrog Artifactory and StorageGRID Webscale provide application-aware data replication to multiple remote sites. JFrog Artifactory synchronizes metadata across sites instantly. When a file is created or modified on one site, its metadata is propagated across other sites instantly by JFrog Artifactory. The data is asynchronously replicated by StorageGRID Webscale to all the remote locations. The replication speed depends on the WAN latency. Remote users can now access the binary packages with better performance.

### How does JFrog Artifactory connect to two different endpoints—ONTAP in the CI/CD workflow and StorageGRID Webscale for data replication?

Answer: The JFrog Artifactory Enterprise Edition supports NFSv3 and Amazon S3. JFrog Artifactory reads the files over NFSv3 from ONTAP and ingests them over S3 to StorageGRID Webscale. Because developers are transparent to files or objects, this approach provides data continuity. JFrog Artifactory gives developers data access and control throughout the CI/CD workflow.

Figure 3) Replicate and share artifacts and binaries across remote sites with Artifactory and StorageGRID Webscale.



**Where can I find the configuration file in JFrog Artifactory that manages the storage endpoints, such as ONTAP and StorageGRID Webscale, that access data over NFS and S3?**

Answer: The recommendation is to set up the JFrog Artifactory home, `/var/opt/jfrog/artifactory`, and the data, `/var/opt/jfrog/artifactory/data`, over NFSv3.

JFrog Artifactory reads files over NFSv3 by default. JFrog Artifactory has a `binarystore.xml` file in the `ART_HOME/etc` directory that stores the S3 configuration for StorageGRID Webscale.

**Does StorageGRID Webscale support identity and access management (IAM)?**

Answer: Yes, StorageGRID Webscale supports IAM at the object level.

**Are policies that are configured in JFrog Artifactory enforced on the objects in StorageGRID Webscale?**

Answer: Yes, StorageGRID Webscale follows S3 semantics. Any policy that is configured on JFrog Artifactory applies to objects in StorageGRID Webscale. For example, if the policy to “sync on deletes” is configured in JFrog Artifactory at one site, then deleting any object from JFrog Artifactory syncs the changes to all the sites in the namespace.

Refer to the [Interoperability Matrix Tool \(IMT\)](#) on the NetApp Support site to validate that the exact product and feature versions described in this document are supported for your specific environment. The NetApp IMT defines the product components and versions that can be used to construct configurations that are supported by NetApp. Specific results depend on each customer's installation in accordance with published specifications.

### **Copyright Information**

Copyright © 1994–2017 NetApp, Inc. All rights reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP "AS IS" AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

RESTRICTED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.277-7103 (October 1988) and FAR 52-227-19 (June 1987).

### **Trademark Information**

NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com/TM> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.