HARVARD
UNIVERSITY

HUIT Infrastructure Governance

Software and OS Package Sourcing Guidelines

Authors: - Fanton, Joel	**Audience Level:** - Solution Architect and Project Manager - Application Developer and Designer - DevOps Staff
Version: 1.0 **Last Revised:** 11-Apr-2023 **Status:** Initial Release **Document Type:** Single Topic Guidance	**Distribution Scope:** - HUIT
Workgroup Members: - Conetta, Todd - Fanton, Joel - Laroche, Matt - Lazri, Cody - Pacheco, Al - Rota, Ben	**Reviewers:** - DeSilva, Indika - Moran, Keith - Rao, Spu - Sevier, Raoul

Purpose

The purpose of these guidelines, hereinafter “Guidelines”, is to provide guidance and governance for the sourcing of software and packages installed standalone or as part of an operating system on HUIT supported systems. Such software, hereinafter “Software,” includes components commonly referred to as application servers and “middleware” as well as packages and libraries required to run internally developed and off-the-shelf applications.

Software shall be installed via automation and infrastructure as code. Installation processes shall source approved software from designated sources as outlined in this document.

Assumptions

These guidelines assume that Software will be installed on HUIT supported operating systems which currently include:

- Red Hat Enterprise Linux (hereinafter “RHEL”)
- Amazon Linux
- Ubuntu
- Windows
- Containers whose image reflects one of the above operating systems

Guiding Principles

- Installed software shall be appropriately lifecycle-managed so that new versions become available as they are released and older versions are retired
- Installed software shall be patched or upgraded on a regular basis so that published security issues and bugs can be remediated in due course
- Artifactory is a preferred software repository for software not distributed with operating system distribution repositories given its capability to scan software for known security issues

Guidelines

These Sourcing Guidelines recognizes that there is no one-size fits all standard that will satisfy all software installation needs. Accordingly, these guidelines reflect both the type of software being installed as well as the operating system onto which it is being installed.

Special Note for Windows systems:

There is no concept of a package manager or standard distribution repository for Windows systems.

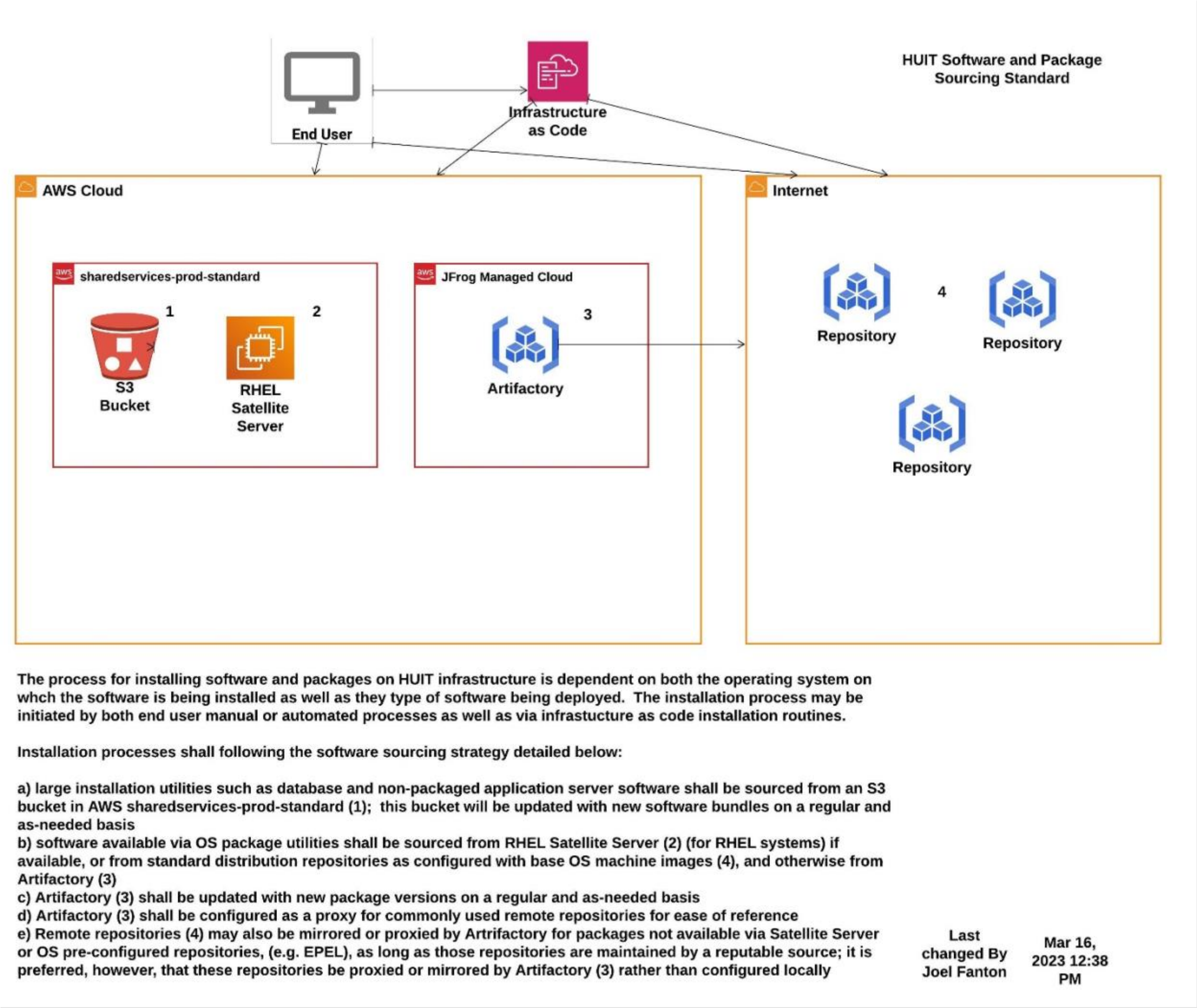
Updates to the Windows operating system are handled via the Windows update utility. In addition, SCCM is used on Windows systems for certain software installations and handled via group policy.

For software not installed by Windows updates and SCCM, software may be sourced from Artifactory or from S3.

Accordingly, references to RHEL Satellite server below do not apply to Windows systems. However, "standard distribution repositories", for Windows systems, may be interpreted as those software sources which Windows uses for processing updates to the Windows operating system.

Software sourcing architecture

Software supporting HUIT-developed and purchased systems shall be sourced from the repositories detailed in the following diagram:



The process for installing software and packages on HUIT infrastructure is dependent on both the operating system on whch the software is being installed as well as they type of software being deployed. The installation process may be initiated by both end user manual or automated processes as well as via infrastucture as code installation routines.

Installation processes shall following the software sourcing strategy detailed below:

a) large installation utilities such as database and non-packaged application server software shall be sourced from an S3 bucket in AWS sharedservices-prod-standard (1); this bucket will be updated with new software bundles on a regular and as-needed basis
b) software available via OS package utilities shall be sourced from RHEL Satellite Server (2) (for RHEL systems) if available, or from standard distribution repositories as configured with base OS machine images (4), and otherwise from Artifactory (3)
c) Artifactory (3) shall be updated with new package versions on a regular and as-needed basis
d) Artifactory (3) shall be configured as a proxy for commonly used remote repositories for ease of reference
e) Remote repositories (4) may also be mirrored or proxied by Artrifactory for packages not available via Satellite Server or OS pre-configured repositories, (e.g. EPEL), as long as those repositories are maintained by a reputable source; it is preferred, however, that these repositories be proxied or mirrored by Artifactory (3) rather than configured locally

Last changed By Joel Fanton   Mar 16, 2023 12:38 PM

Artifactory

Artifactory provides multiple repository types – RPM, DEB, Maven, Python, etc. It can also operate as a generic fileserver with a "generic" repository type.

Artifactory will provide repositories for custom RPM/DEB packages provided by Harvard (e.g. various SMVP tools that need to be installed on servers).

Artifactory will also serve as a mirror to ~~“authorized”~~ upstream repositories (e.g., EPEL) authorized by the IGC. These will be repositories that provide useful packages that are verified to be updated in a timely manner and to follow security best practices.

Red Hat Enterprise Satellite Server

Satellite server is systems management software provided by Red Hat to manage servers running RHEL. HUIT uses satellite server to deploy RHEL specific operating system patches and software packages.

For RHEL servers OS patches shall always be deployed from satellite server. In addition, required system software available in satellite server shall be sourced from satellite server as primary. If required system software is not available in satellite server then Artifactory should be consulted.

AWS S3

Certain large software installers (e.g. Oracle database software, Weblogic application server software) will be staged in AWS S3 and shall be sourced from S3 for download and install on servers.

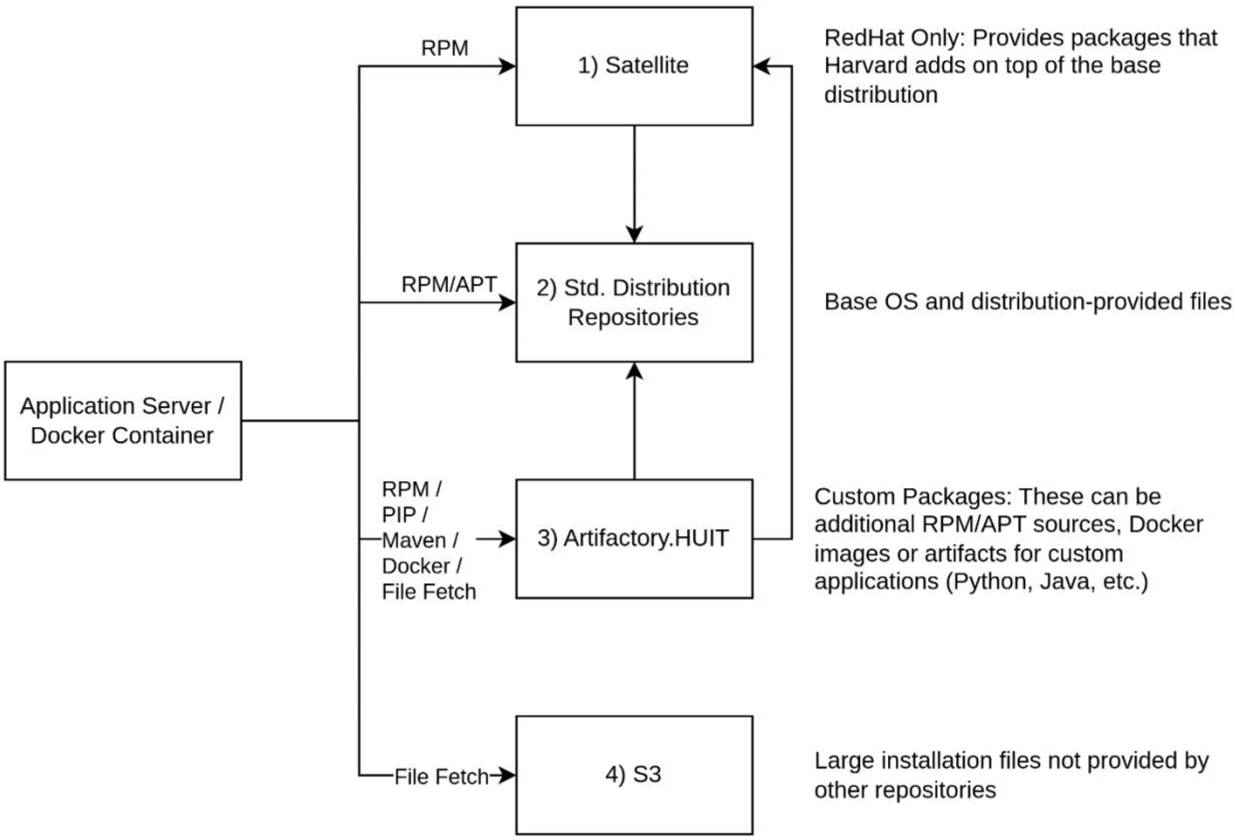
In practice

An application server will have the following options from which to fetch artifacts.

Standard OS Distribution Repositories / Satellite: Your first choice for finding packages and should be used for packages provided by the Linux distribution. This will be core OS files and often interpreters/compilers such as Java, Python, etc. that are provided by your distribution (Amazon Linux, RedHat, Ubuntu, etc.). This should be the preferred route for any dependencies as security updates are provided by the upstream providers in a timely manner and fetching new versions is easy and standardized.

Artifactory.HUIT: Your next choice for finding packages. Shall be used as an additional package repository for RPM or APT as well as a repository for application builds (e.g. custom python, Java, etc. applications). Generally speaking, application builds should be stored in Artifactory to be deployed to an application server.

S3: Your last choice for finding packages. Similar to Artifactory S3 shall be used to store large installation files for anything not found in either the standard package management or Artifactory sources. This may be things like Weblogic or other proprietary installation materials.


RPM
1) Satellite
RedHat Only: Provides packages that Harvard adds on top of the base distribution
RPM/APT
2) Std. Distribution Repositories
Base OS and distribution-provided files
Application Server / Docker Container
RPM / PIP / Maven / Docker / File Fetch
3) Artifactory.HUIT
Custom Packages: These can be additional RPM/APT sources, Docker images or artifacts for custom applications (Python, Java, etc.)
File Fetch
4) S3
Large installation files not provided by other repositories

Scenario 1: RedHat 9 Application server installing dependencies for an Apache Tomcat Java application*

The application server shall be built from an approved machine image (e.g. AMIBuilder AMI for RHEL9).

To complete the software build out for an Apache Tomcat Java application:

- the latest LTS version of Java is 17 and can be installed using the RedHat upstream repositories via Satellite server
- Apache Tomcat is not provided by the standard RedHat repositories and shall be staged in and fetched from Artifactory via mirror of upstream repository or via specific version staging for RHEL
- the application .WAR file shall be stored in and fetched from Artifactory.

Scenario: Windows 2022 Application server installing dependencies for an Apache Tomcat Java application*

The application server shall be built from an approved machine image (e.g. AMIBuilder AMI for Windows 2022).

To complete the software build out for an Apache Tomcat Java application:

- the latest LTS version of Java is 17; if it is not available as part of the standard Windows OS installation, it can be installed using Artifactory

- Apache Tomcat is not provided as part of the standard Windows OS installation and can be fetched from Artifactory via mirror of upstream repository or via specific version staging for Windows
- the application .WAR file shall be stored in and fetched from Artifactory.

*note that some software installation packages including .exe and .msi installer may include dependent software which will be installed as required and as such will not be subject to these software sourcing guidelines.

Software to be staged/sourced in Artifactory includes but is not limited to the following (either via upstream mirror or stage of specific version)

- OpenJDK
- Apache/Tomcat
- Jenkins
- HUIT Security MVP software (CloudAware, Crowdstrike, Splunk forwarder, Nessus)
- Ansible Bootstrap package
- Cloud-os-accounts (package to install DevOps support accounts)
- Oracle Instant Client
- Weblogic Apache plugins
- AWS SSM agent
- AWS Codedeploy agent

Tools links:

Artifactory: artifactory.huit.harvard.edu

AWS S3: sharedservices-prod- huit-devops-software-repository: https://s3.console.aws.amazon.com/s3/buckets/huit-devops-software-repository?region=us-east-1&tab=objects

Installation Decision Matrix:

For Java and Weblogic (others to be added):

Component	Vendor	Source	Lifecycle Management
Java	Oracle	S3 – via Oracle provided software bundle (e.e. EBusiness Suite,	Via Oracle quarterly CPU/PSU

		PeopleSoft) for Linux or Windows	
Java	OpenJDK	**Linux systems:** RHEL Satellite Server/standard OS distribution repository **Windows systems:** via Artifactory	Via OS patching schedules Via Artifactory lifecycle management
Weblogic	Oracle	S3 – via Oracle provided software bundle (e.e. EBusiness Suite, PeopleSoft) for Linux or Windows	Via Oracle quarterly CPU/PSU