Advanced Jenkins

Core Practices to Managing and Scaling Jenkins for the Enterprise

DARIN POPE
DEVELOPER ADVOCATE, CLOUDBEES

CONTENTS

Jenkins is arguably one of the most popular development tools on the planet, with some reports estimating that over 70% of all continuous integration pipelines run on Jenkins. It's great at helping small, agile development teams build, test, and deploy multiple times a day. Some estimates say Jenkins supports more than 180,000 sites, running 20 million jobs across more than 750,000 build agents.

However, as teams, environments, projects, and market pressures increase, so do the administrative tasks associated with maintaining and administering Jenkins. Scaling Jenkins usage across a growing enterprise, without affecting the security or stability of its software development lifecycle (SDLC) is challenging, particularly for highly regulated industries.

CORE PRACTICES

In this Refcard, we outline specific best practices for setting up efficient, secure Jenkins pipelines.

USE JUST ENOUGH PIPELINE

Jenkins Pipeline (or simply Pipeline with a capital "P") is a suite of plugins that supports implementing and integrating continuous delivery pipelines into Jenkins. This allows you to automate the SDLC and deliver important changes more efficiently to your users and customers.

Pipeline code works beautifully for its intended role of automating build, test, deploy, and administration tasks. But, as it is pressed into more complex roles and unexpected uses, some users have run into snags. Using best practices — and avoiding common mistakes — can help you design a pipeline that is more robust, scalable, and high-performing.

Often, development teams make basic mistakes that can sabotage their pipeline. (Yes, you ***can*** sabotage yourself when you're creating a pipeline.) In fact, it's easy to spot someone who is going down this dangerous path — and it's usually because they don't understand some key technical concepts about Pipeline. This invariably leads to scalability mistakes that you'll pay dearly for down the line.

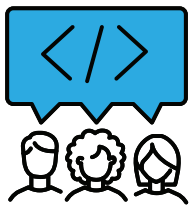
JUST SAY NO TO PIPELINES IN PROGRAMMING LANGUAGES

Perhaps the biggest misstep people make is deciding that they need to write their entire pipeline in a programming language. After all, Pipeline is a domain specific language (DSL). However, that does not mean that it is a general-purpose programming language.

CloudBees

Need Help Managing Your Jenkins at Scale?

CloudBees Continuous Integration can help you:

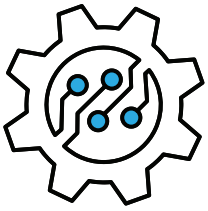


Reduce Administrative Overhead

Simplify management across controllers, onboard easily with RBAC, and configure it all as code

Build More Safely

Built-in security and compliance for hardened Jenkins workflows.



Automate at Scale

Eliminate endless re-scripting. Repeatability every time.

Learn More

If you treat the DSL as a general-purpose programming language, you are making a serious architectural blunder by doing the wrong work in the wrong place. Remember that the core of Pipeline code runs on the controller. So, you should be mindful that everything you express in the Pipeline domain specific language (DSL) will compete with every other Jenkins job running on the controller.

For example, it's easy to include a lot of conditionals, flow control logic, and requests using scripted syntax in the pipeline job. Experience tells us this is not a good idea and can result in serious damage to pipeline performance. Organizations with poorly written Pipeline jobs bring a controller to its knees, while only running a few concurrent builds.

Wait a minute, you might ask: "Isn't the controller supposed to handle code?" Yes, the controller certainly is there to execute pipelines. But it's much better to assign individual steps of the pipeline to command-line calls that execute on an agent. So, instead of running a lot of conditionals inside the pipeline DSL, it's better to put those conditionals inside a shell script or batch file and call that script from the pipeline.

However, this begs another question: "What if I don't have any agents connected to my controller?" If this is the case, then you've just made another bad mistake in scaling Jenkins pipelines. Why? Because the first rule of building an effective pipeline is to make sure you use agents. If you're using a Jenkins controller and haven't defined any agents, then your first step should be to define at least one agent and use that agent instead of executing on the controller.

For the sake of maintaining scalability in your pipeline, the general rule is to avoid processing any workload on your controller. If you're running Jenkins jobs on the controller, you are sacrificing controller performance. So, try to avoid using Jenkins controller capacity for things that should be passed off to an agent. Then, as you grow and develop, all your work should be running agents. This is why we always recommend setting the number of executors on the controller to zero.

USE JUST ENOUGH PIPELINE TO KEEP YOUR PIPELINE SCALABLE

All of this serves to highlight the overarching theme of "using just enough pipeline." Simply put, you want to use enough code to connect the pipeline steps and integrate tools — but no more than that. Limit the amount of complex logic embedded in the Pipeline itself (similarly to a shell script), and avoid treating it as a general-purpose programming language. This makes the pipeline easier to maintain, protects against bugs, and reduces the load on controllers.

Another best practice for keeping your pipeline lean, fast, and scalable is to use declarative syntax instead of scripted syntax for your Pipeline. Declarative naturally leads you away from the kinds of mistakes just described. It is a simpler expression of code and an easier way to define your job. It's computed at the startup of the pipeline instead of executing continually during the pipeline.

Therefore, when creating a pipeline, start with declarative, and keep it simple for as long as possible. Anytime a script block shows up inside of a declarative pipeline, you should extract that block and put it in a shared library step. That way, the declarative pipeline is still clean. Joining together the declarative and the shared library will take care of the vast majority of use cases you'll experience.

That being said, you cannot assume that declarative plus a shared library will solve every problem. There are cases where scripted is the right solution. However, declarative is a great starting point until you discover that you absolutely must use scripted. Just remember, at the end of the day, you'll do well to follow the adage: "Use just enough pipeline and no more."

USE DECLARATIVE PIPELINE

Maintaining multiple configurations of larger, more complex pipelines is another pain point commonly experienced by enterprise Jenkins users. Declarative Pipelines provide a more modern, opinionated approach to building software. The resulting Jenkinsfile can then be committed to a Git repo in a "pipeline as code" fashion. It is the recommended way of building Pipelines at scale as it makes Pipeline easier to manage.

A declarative Pipeline provides a structured hierarchical syntax to simplify the creation of Pipelines and the associated Jenkinsfiles. In its simplest form, a Pipeline runs on an agent and contains stages, while each stage contains steps that define specific actions. Here is an example:

```
pipeline {
    agent any
    stages {
        stage('Build') {
            steps {
                sh 'mvn install'
            }
        }
    }
}
```

Let's define a few terms before we go any further:

- **Agent**: An `agent` section specifies where tasks in a Pipeline run. An `agent` must be defined at the top level inside the `pipeline` block to define the default agent for all stages in the Pipeline. An `agent` section may optionally be specified at the stage level, overriding the default for this stage and any child stages. In this example, the agent is specified with `any`, meaning this Pipeline will run on any available agent.

- **Stages**: A `stage` represents a logical grouping of tasks in the Pipeline. Each `stage` may contain a `steps` section with steps to be executed, or stages to be executed sequentially, in parallel, or expanded into a parallel matrix.
- It contains commands that run processes like Build, Deploy, or Tests. These commands are included in the `steps` section in the stage. It is advisable to have descriptive names for a `stage` as these names are displayed in the UI and logs.
- There can be one or more `stage` sections inside a `stages` section. At least one `stage` must be defined in the `stages` section.
- **Steps**: The `` `steps` `` section contains a set of commands. Each command performs a specific action and is executed one-by-one:

```
pipeline {
    agent any
    stages {
        stage('Example') {
            steps {
                sh 'mvn compile'
            }
        }
    }
}
```

AVOID CREATING A MONOLITHIC CONTROLLER

The number one question Jenkins administrators have when broaching this subject is: "How many jobs *can* I run on my controller?" Unfortunately, there is not a simple answer. The number of variables is too great. The pipeline complexity variable alone is enough to sway a guestimate into too many unknowns. We do, however, recommend using the 5,000 jobs figure as a high-water mark to help you understand when it is time to horizontally scale.

Employing a monitoring solution to monitor macro metrics (CPU/RAM/DISK I/O) and creating baselines will allow you to properly plan for scaling. Micro-metrics such as garbage collection logs, object creation rate, and thread counts can be analyzed and baselined so that you understand the needs of your jobs and properly plan for additional controllers as needed in your installation.

DON'T LET YOUR PLUGINS MANAGE YOU

The large ecosystem of plugins that extend the functionality of Jenkins is a critical factor in its popularity and success. Despite all the benefits that plugins bring to Jenkins, they also bring with them a host of problems. The management of plugins within a given Jenkins controller — and especially across a fleet of Jenkins controllers — has security, stability, and other implications that are compounded at scale.

CONSIDER YOUR PLUGIN MANAGEMENT STRATEGY

Jenkins administrators can go about managing their plugins in many different ways. To determine the best plugin management strategy for your organization, consider the following questions:

- What is your operational model? For example, do you…
  - Have centralized management for all plugin-related changes on all controllers?
  - Default to one-time plugin deployment for all controllers, then self-managed afterwards?
  - Allow team/project administrators of each controller to manage and install plugins?
- Will controllers be restricted on the list of plugins available to them?
- Do you prefer the definition of plugin list to be in code or via UI?
- Is the installation of plugins per controller optional or mandatory?
- What are the network restrictions and/or corporate policy restrictions on plugins?
- Do you have any internally developed plugins?
- What type of controllers will be provisioned?

The comparison chart below demonstrates plugin management methods available natively in Jenkins.

NATIVE PLUGIN MANAGEMENT METHODS FOR JENKINS

OPTION	DESCRIPTION
Manage Plugins	• Available to all Jenkins administrators. • Plugins can be installed, uninstalled, and updated via the UI. The Advanced tab lets administrators install custom plugins, or plugins that are not available from the update center or the plugin catalog. • This is the simplest and most common way of installing plugins.
Commands With Jenkins CLI	• Use existing commands that can perform plugin installations, enable, disable, and list plugins. • This CLI command can be performed by a script and will have to be triggered on each controller. • From the CLI, plugins can be installed by name (via update center) and by a link to a plugin (via URL or file).
Jenkins Configuration as Code (JCasC) for Controllers	• Simplifies the management of a Jenkins installation by capturing the configuration of Jenkins controllers in human-readable declarative configuration files that can then be applied to a controller in a reproducible way. • Plugins are managed using two configuration files, plugins.yaml and plugin-catalog.yaml. Plugins are installed automatically upon boot or restart of a controller, and upon a hot reload on the controller. • Plugins listed in the JCasC bundle aren't optional; they are installed automatically every time.

For smaller organizations, the above methods may cover their plugin management needs. Larger organizations, particularly those in highly regulated industries such as banking, healthcare, or insurance, require more robust team management capabilities that tie-in to their RBAC strategy. They are also more likely to require ongoing performance and security testing of plugins used. This might entail testing plugins, plugin versions, and plugin dependencies to determine their stability. For those organizations, more advanced plugin management methods are available by using a commercial, Jenkins-based CI solution.

As software organizations grow, so do the number of plugins that are required to support Jenkins pipelines and application development teams. A Jenkins administrator can end up with hundreds of plugins to manage. Without visibility into which plugins are being used and not used, unused plugins are oftentimes left installed to prevent disruption to the entire system. Over time, this leads to plugin bloat, which can negatively impact system stability and whether or not upgrades are performed.

USE CONTAINERS FOR BUILD AGENTS

A continuous integration environment is a mixed bag of machines, platforms, build toolchains, and operating systems. You need the utmost flexibility to manage these machines and build them to be interchangeable. In general, you don't want to tie builds to a specific build machine.

Make use of container images and Dockerfiles. You won't have to worry about configuring and managing tool installers or setting up build agent images. With Declarative Pipelines, you can specify that the build runs in a specific container image, or you can have your build environment expressed as a Dockerfile that is stored in source control. You can also use the same behavior at a per-stage level rather than across the entire pipeline.

DEVELOPERS CAN HAVE COMPLETE CONTROL OVER THEIR BUILD ENVIRONMENT

If you need a new version of a tool, update the Dockerfile in source control as part of the pull request to upgrade the tool. If you need to build an old branch, the old branch will have the old build environment.

WHAT IF I'M USING WINDOWS AND LINUX CONTAINERS?

If you are working in an environment with a mixture of Windows and Linux containers, then you will probably want to use labels to differentiate the build agents configured to run Windows containers from the agents configured to run Linux containers.

KEY TAKEAWAYS

While Jenkins is the leading automation platform for continuous integration (CI), enterprises have unique needs. They must be able to scale Jenkins across multiple teams without increasing the administrative burden. Enterprises must ensure security and compliance without inhibiting productivity. At the same time, a growing number of enterprises are embracing a hybrid cloud strategy. Supporting a vast range of infrastructure types and optimizing software delivery across multiple tools and teams present complex challenges.

Organizations facing these challenges should consider looking for solutions that help them with the burden of managing and scaling Jenkins. Commercial solutions can extend Jenkins with enterprise functionality that embeds best practices, rapid onboarding, security, and compliance.

WRITTEN BY DARIN POPE,
DEVELOPER ADVOCATE, CLOUDBEES

Darin is a developer advocate for CloudBees, where he produces technical digital and written content for users. He is also the co-host of a weekly podcast named DevOps Paradox.

DZone

DZone, a Devada Media Property, is the resource software developers, engineers, and architects turn to time and again to learn new skills, solve software development problems, and share their expertise. Every day, hundreds of thousands of developers come to DZone to read about the latest technologies, methodologies, and best practices. That makes DZone the ideal place for developer marketers to build product and brand awareness and drive sales. DZone clients include some of the most innovative technology and tech-enabled companies in the world including Red Hat, Cloud Elements, Sensu, and Sauce Labs.

Devada, Inc.
600 Park Offices Drive
Suite 150
Research Triangle Park, NC 27709

888.678.0399 | 919.678.0300