# 1

# Apache Tomcat



If you've written any Java servlets or JavaServer Pages (JSPs), chances are good that you've downloaded Tomcat. That is because Tomcat is a free, feature-complete *Servlet container* that developers of servlets and JSPs can use to run their code. Tomcat is used in Sun's reference implementation of the Servlet Container, which means that Tomcat's first goal is to be 100 percent compliant with the versions of the Servlet and JSP API specifications that it supports.

However, Tomcat is more than just a test server. Many corporations are using Tomcat in production environments because it has proven to be quite stable. These corporations range from Fortune 500 companies such as WalMart and General Motors to ISPs hosting multiple small-business Web sites. Tomcat is used in the real world to run everything from online photo albums (Webshots) to high performance financial Web applications (ETrade).

*A list of Tomcat-powered Web sites is at* `http://wiki.apache.org/tomcat/PoweredBy`.

Despite Tomcat's popularity, it suffers from a common shortcoming among open source projects: lack of complete documentation. Some documentation is distributed with Tomcat (mirrored at `http://tomcat.apache.org`), and there's an open source effort to write a Tomcat book (`http://tomcatbook.sourceforge.net/`). Even with these resources, however, there is a great need for additional material.

This book has been created not just to fill in some of the documentation holes, but to use the combined experience of the authors to help Java developers and system administrators make the most of the Tomcat product. Whether you're trying to learn enough to just get started developing Web applications or want to understand the more arcane aspects of Tomcat configuration, you should find what you're looking for within these pages.

The first two chapters are designed to provide newcomers with some basic background information that is prerequisite learning for subsequent chapters. If you're a system administrator with no previous Java experience, we advise you to read these first two chapters, and likewise if you're a Java developer who is new to Tomcat. If you're well informed about Tomcat and Java, you'll

probably want to jump straight ahead to Chapter 3, although skimming this chapter and its successor is likely to add to your present understanding.

The following topics are discussed in this chapter:

- The origins of the Tomcat server
- The terms of Tomcat's license and how it compares to other open source licenses
- How Tomcat fits into the Java "big picture"
- An overview of integrating Tomcat with Apache and other Web servers

# Humble Beginnings: The Apache Project

One of the earliest Web servers was developed by Rob McCool at the National Center for Supercomputer Applications (NCSA), University of Illinois, Urbana-Champaign. This Web server was referred to colloquially as the NCSA project, or NCSA for short. By 1995, the NCSA server was quite popular, but its future was uncertain because the primary developer, McCool, had left NCSA the previous year. A group of developers got together and compiled all the NCSA bug fixes and enhancements they had found, and patched them into the NCSA code base. The developers released this new version in April 1995, and called it Apache, which was somewhat of an acronym for "A PAtCHy Web Server."

Apache was readily accepted by the developer community from its earliest days, and less than a year after its release, it unseated NCSA to become the most used Web server in the world (measured by the total number of servers running Apache), a distinction that it has held ever since (according to Apache's Web site). Incidentally, during the same period that Apache's use was spreading, NCSA's popularity was plummeting, and by 1999, NCSA was officially discontinued by its maintainers.

> *For more information on the history of Apache and its developers, see* `http://httpd.apache.org/ABOUT_APACHE.html`.

Today, the Apache Web server is available on just about any major operating system (in addition to the source code download, Apache binaries are available for over a dozen operating systems). Apache can be found running on some of the largest server farms in the world, as well as on some of the smallest devices (including several hand-held devices). In UNIX data centers, Apache is as ubiquitous as air conditioning and UPS systems.

While Apache was originally a somewhat mangy collection of miscellaneous patches, today's versions are rock-solid production quality servers. The only real competitor to Apache in terms of market share and feature set is Microsoft's Internet Information Server (IIS), which is bundled free with certain versions of the Windows operating system. As of this writing, Apache's market share is estimated at around 60 percent, with IIS at 30 percent (statistics courtesy of `http://news.netcraft.com/archives/web_server_survey.html`).

It is also worth noting that Apache has a reputation for being much more secure than Microsoft IIS. When new vulnerabilities are discovered in either server, the Apache developers fix Apache far faster than Microsoft fixes IIS.

# The Apache Software Foundation

In 1999, the same folks who wrote the Apache server formed the Apache Software Foundation (ASF). The ASF is a nonprofit organization that was created to facilitate the development of open source software projects. Tomcat is developed under the auspices of the ASF. According to their Web site, the ASF accomplishes this goal by doing the following:

- Providing a *foundation* for open, collaborative software development projects by supplying hardware, communication, and business infrastructure
- Creating an independent legal entity to which companies and individuals can *donate resources* and be assured that those resources will be used for the public benefit
- Providing a means for individual volunteers to be sheltered from *legal suits* directed at ASF projects
- Protecting the Apache *brand* (as applied to its software products) from being abused by other organizations

In practice, the ASF does indeed sponsor a great many open source projects. While the best-known of these projects is likely the aforementioned Apache Web server, the ASF hosts many other well-respected and widely used projects, including such respected industry standards as the following:

- **Xerces:** A Java/C++ XML parser with JAXP bindings
- **Ant:** A Java-based build system (and much more)
- **Axis:** A Java-based Web services implementation

The number of ASF-sponsored projects is growing fast. Visit `www.apache.org` to see the latest list.

# Tomcat

The Tomcat project has its origins in the earliest days of Java's servlet technology. *Servlets* are a certain type of Java application that plug into special Web servers, called *Servlet containers* (originally called Servlet engines). Sun created the first Servlet container, called the Java Web Server, which demonstrated the technology but wasn't terribly robust. Meanwhile, the ASF folks created the JServ product, which was a Servlet engine that integrated with the Apache Web server.

In 1999, Sun donated its Servlet container code to the ASF, and the two projects were merged to create the Tomcat server. Today, Tomcat is used by Sun in its *reference implementation* (RI), which means that Tomcat's first priority is to be fully compliant with the Servlet and *JavaServer Pages* (JSP) specifications published by Sun. This is discussed in more detail in Chapter 2.

The first version of Tomcat was the 3.*x* series, and it implemented the Servlet 2.2 and JSP 1.1 specifications. The Tomcat 3.*x* series was descended from the original code that Sun provided to the ASF in 1999.

In 2001, Tomcat 4.0 (code-named Catalina) was released. Catalina was a complete redesign of the Tomcat architecture, and built on a new code base. The Tomcat 4.*x* series was used in the RI of the Servlet 2.3 and JSP 1.2 specifications.

The latest version of Tomcat, Tomcat 6, implements the Servlet 2.5 and JSP 2.1 specifications. In addition, it boasts of an improved clustering implementation over the previous iteration (Tomcat 5.5).

> **Tomcat used to be a subproject under the Apache Jakarta project. The Jakarta project is an umbrella under which the ASF sponsors the development of many Java sub-projects, such as JMeter, Log4j, and Struts. However, Tomcat has now been promoted to a top-level project.**

## *Distributing Tomcat: The Apache License*

Tomcat is open source software, and, as such, is free and freely distributable. However, if you have much experience in dealing with open source software, you're probably aware that the terms of distribution can vary from project to project.

Most open source software is released with an accompanying license that states what may and may not be done to the software. At least 40 different open source licenses are in use, each of which has slightly different terms.

Providing a primer on all of the various open source licenses is beyond the scope of this chapter, but the license governing Tomcat is discussed here and compared with a few of the more popular open source licenses.

Tomcat is distributed under the Apache License, which is listed at `apache.org/licenses`. The key points of this license state the following:

- The Apache License must be included with any redistribution of Tomcat's source code or binaries.
- Any documentation included with redistribution must give a nod to the ASF.
- Products derived from the Tomcat source code can't use the terms "Tomcat," "The Jakarta Project," "Apache," or "Apache Software Foundation" to endorse or promote their software without prior written permission from the ASF.
- Tomcat has no warranty of any kind.

However, through omission, the license contains the following additional implicit permissions:

- Tomcat can be used by any entity (commercial or noncommercial) for free without limitation.
- Those that make modifications to Tomcat and distribute their modified version do not have to include the source code of their modifications.
- Those who make modifications to Tomcat do not have to donate their modifications to the ASF.

Thus, you're free to deploy Tomcat in your company in any way you see fit. It can be your production Web server or your test Servlet container used by your developers. You can also redistribute Tomcat with any commercial application that you may be selling, provided that you include the license and give credit to the ASF. You can even use the Tomcat source code as the foundation for your own commercial product.

## Comparison with Other Licenses

Among the previously mentioned and rather large group of other open source licenses, two licenses are particularly popular at the present time: the GNU General Public License (GPL) and the GNU Lesser General Public License (LGPL). Let's take a look at how each of these licenses compares to the Apache License.

### GPL

The GNU Project created and actively evangelizes the GPL. The GNU Project is somewhat similar to the ASF, with the exception that the GNU Project would like all of the non-free (that is, closed source or proprietary) software in the world to become free. The ASF has no such (stated) desire and simply wants to provide free software.

Free software can mean one of two entirely different things: software that doesn't cost anything and software that can be freely copied, distributed, and modified by anyone (thus, the source code is included or is easily accessible). Such software can be distributed either free or for a fee. A simpler way to explain the difference between these two types of free is to compare "free," as in "free beer," and "free," as in "free speech." The GNU Project's goal is to create free software of the latter category. All uses of the phrase "free software" in the remainder of this section use this definition.

The differences between the Apache License and the GPL thus mirror the distinct philosophies of the two organizations. Specifically, the GPL has the following key differences from the Apache License:

- No "non-free" software may contain GPL-licensed products or use GPL-licensed source code. If non-free software is found to contain GPL-licensed binaries or code, it must remove such elements or become free software itself.
- All modifications made to GPL-licensed products must be released as free software if the modifications are also publicly released.

These two differences have huge implications for commercial enterprises. If Tomcat were licensed under the GPL, any product that contained Tomcat would also have to be free software.

Furthermore, while the Apache License permits an organization to make modifications to Tomcat and sell it under a different name as a closed source product, the GPL would not allow any such act to occur; the new derived product would also have to be released as free software.

### LGPL

The GNU Lesser General Public License (LGPL) is similar to the GPL, with one major difference: Non-free software may contain LGPL-licensed products. The LGPL license is commonly referred to as the "library" GPL because it is intended primarily for software libraries that are themselves free software, but whose authors want them to be available for use by companies who produce non-free software.

If Tomcat were licensed under the LGPL, it could be embedded in non-free software, but Tomcat could not itself be modified and released as a non-free software product.

For more information on the GPL and LGPL licenses, see `www.gnu.org`.

### Other Licenses

Understanding and comparing open source licenses can be a rather complex task. The preceding explanations are an attempt to simplify the issues. For more detailed information on these and other licenses, the following two resources can help you:

- The Open Source Initiative (OSI) maintains a database of open source licenses. Visit them at `www.opensource.org`.
- The GNU Project has an extensive comparison of open source licenses with the GPL license. See it at `www.gnu.org/licenses/license-list.html`.

# The Big Picture: Java EE

As a Servlet container, Tomcat is a key component of a larger set of standards collectively referred to as the Java Enterprise Edition (*Java EE*) platform. The Java EE standard defines a group of Java-based *APIs* that are suited to creating Web applications for *enterprises* (that is, large companies). To be sure, companies of any size can take advantage of Java EE, but many Java EE technologies are especially designed to solve the problems associated with the creation of large software systems.

Java EE is built on the Java Standard Edition (*Java SE*), which includes the Java binaries (such as the JVM and bytecode compiler), as well as the core Java code libraries. Java EE depends on Java SE to function. Both the Java SE and Java EE can be obtained from `http://java.sun.com`. Both Java SE and Java EE are referred to as *platforms*, because they provide core functionality that acts as a sort of platform or foundation upon which applications can be built.

**Since the middle of 2005, Sun has been re-branding some of the Java platform names. Java Enterprise Edition, previously called J2EE, is now called Java EE. Java Standard Edition, previously called J2SE, is now Java SE. Similarly, the mobile edition (previously J2ME) has been renamed to Java ME.**

## Java APIs

As mentioned, Java EE is a standardized collection of Java APIs. The term *API* (or *application programming interface*) is used by software developers in general to describe services made available to applications by an underlying service provider (such as an operating system). In the Java world, this term is used to describe many of the services that the Java Virtual Machine (JVM) and its code libraries make available to Java programs.

An important characteristic of APIs is that they are separated from the services that provide them. In other words, an API is a kind of technical contract defining the functionality that two parties must provide: a service provider (often called an *implementation*) and an application. If both parties adhere to the contract, an API is *pluggable* (that is, a new service provider can be plugged into the relationship). Of course, if a service provider fails to conform to the contract, the applications that use the API will fail to function properly.

### The Java Community Process

APIs in the Java world are created and modified by a standards body known as the Java Community Process (*JCP*). The JCP is composed of hundreds of *Java Specification Requests (JSRs)*. Each JSR is a request to either change an existing aspect of Java (including its APIs) or introduce a new API or feature to Java. New JSRs can be submitted by a *member* of the JCP. Anyone can become a member of the JCP and, notably, individuals may do so at no cost (organizations pay a nominal fee). Once submitted, the JCP *Executive Committee* must approve the JSR. The Executive Committee consists of JCP members who have been elected to three-year terms in an annual election.

When a JSR is approved, the submitter becomes the *Spec Lead*. The Spec Lead forms an *Expert Group* composed of JCP members who assist the Spec Lead in creating a specification detailing the change or addition to the Java language. The Expert Group shepherds the specification along through various review processes (to other JCP members and to the public) until, finally, the JSR is judged completed and is approved by the Executive Committee. If a JSR results in an API, the Expert Group must also provide a reference implementation of the API (discussed earlier in this chapter in the context of Tomcat) and a *technology compatibility kit* (TCK) that other implementers can use to verify compatibility with the API.

Thus, via the JCP, any Java developer can influence the Java platforms, by submitting a JSR, becoming a member of an existing JSR's Expert Group, or by simply giving feedback to JSR Expert Groups. While not the first attempt to create a technology standards body, the JCP is probably the world's best combination of accessibility and influence. As a contrast, the influential World Wide Web Consortium (W3C) standards body charges almost $6,000 for individuals to join. Visit the JCP at `www.jcp.org`.

## The Java EE APIs

As mentioned, the Java EE 5 platform consists of many individual APIs. The Servlet and JSP APIs are two of these. The following table describes some of the other Java EE APIs, and a complete list can be found at `http://java.sun.com/javaee/technologies/`.

| Java EE API | Description |
|---|---|
| Enterprise JavaBeans (EJB) | Provides a mechanism that is intended to make it easy for Java developers to use advanced features in their components, such as remote method invocation (RMI), object/relational mapping (that is, saving Java objects to a relational database), distributed transactions across multiple data sources, statefulness, and so on. |
| Java Message Service (JMS) | Provides high-performance asynchronous messaging. Among other things, it enables Java EE applications to communicate with non-Java systems on top of various transports. |
| Web service APIs | A set of APIs for Web services and XML processing. These include JAX-WS, JAX-RPC, JAXB, SAAJ, and StAX. |
| Java Management Extensions (JMX) | Standardizes a mechanism for interactively monitoring and managing applications at runtime. |

*Table continued on following page*

| Java EE API | Description |
|---|---|
| Java Transaction API (JTA) | JTA enables applications to gracefully handle failures in one or more of their components by establishing transactions. During a transaction, multiple events can occur, and if any one of them fails, the state of the application can be rolled back to the way it was before the transaction began. JTA provides the functionality of database-transactions technology across an entire distributed application. |
| JavaMail | Provides the capability to send and receive e-mail via the industry-standard POP/SMTP/IMAP protocols. |

In addition to the Java EE–specific APIs, Java EE applications also rely heavily on Java SE APIs. In fact, over the years, several of the Java EE APIs have been migrated to the Java SE platform. Two such APIs are the Java Naming and Directory Interface (JNDI), used for interfacing with LDAP-compliant directories (and much more), and the Java API for XML Processing (JAXP), which is used for parsing and transforming XML (using XSLT). The vast collection of Java EE and Java SE APIs form a platform for enterprise software development unparalleled in the industry.

## Java EE Application Servers

As mentioned, an API simply defines services that a service provider (i.e., the implementation) makes available to applications. Thus, an API without an implementation is useless. While the JCP does provide RIs of all the APIs, using them piecemeal is not the most efficient way to build applications. Enter the *Java EE application server*.

Various third parties provide commercial-grade implementations of the Java EE APIs. These implementations are typically packaged as a Java EE application server. Whereas Tomcat provides an implementation of the Servlet and JSP APIs (and is thus called a Servlet container), application servers provide a superset of Tomcat's functionality: the Servlet and JSP APIs plus all the other Java EE APIs, and some Java SE APIs (such as JNDI).

Dozens of vendors have created *Java EE–compatible* application servers. Being called "Java EE–compliant" means that a vendor of an application server has paid Sun a considerable sum, and has passed various compatibility tests. Such vendors are said to be *Java EE licensees*.

The two most widely used commercial Java EE application servers are Websphere from IBM and Weblogic from BEA. Other than these, there are a number of open source implementations too, such as the following:

- JBoss (`www.jboss.org`)
- JOnAS (`jonas.objectweb.org`)
- Geronimo (`geronimo.apache.org`)
- Glassfish (`glassfish.dev.java.net`)

## "Agree on Standards, Compete on Implementation"

Developers who use the Java EE APIs can use a Java EE–compatible application server from any vendor, and it is guaranteed to work with their applications. This flexibility is intended to help customers avoid

vendor lock-in problems, enabling users to enjoy the benefits of a competitive marketplace. The Java slogan along these lines is "Agree on standards, compete on implementation," meaning that the vendors all cooperate in establishing universal Java EE standards (through participation in the JCP) and then work hard to create the best application server implementation of those standards.

That's the theory, at least. In reality, this happy vision of vendor neutrality and open standards is slightly marred by at least two factors. First, each application server is likely to have its own eccentricities and bugs. This leads to a popular variation on the famous "Write Once, Run Anywhere" Java slogan: "Write Once, Test Everywhere." Second, vendors are rarely altruistic. Each application server typically includes a series of powerful features that are outside the scope of the Java EE APIs. Once developers take advantage of these features, their application is no longer portable, resulting in vendor lock-in. Developers must, therefore, be vigilant to maintain their application's portability, if such a capability is desirable.

## *Tomcat and Application Servers*

Up to this point, Tomcat has been referred to as an implementation of the Servlet/JSP APIs (i.e., a Servlet container). However, Tomcat is more than this. It also provides an implementation of the JNDI and JMX APIs. However, Tomcat is not a complete Java EE application server; it doesn't provide support for even a majority of the Java EE APIs.

Interestingly, many application servers actually use Tomcat as their implementation of the Servlet and JSP APIs. Because Tomcat permits developers to embed Tomcat in their applications with only a one-line acknowledgment, many commercial application servers quietly rely on Tomcat without emphasizing that fact. The JBoss and JOnAS application servers mentioned previously make explicit use of Tomcat.

Developers seeking to create Java Web applications that utilize the Servlet, JSP, JNDI, and JMX APIs will find Tomcat an excellent solution. However, those seeking support for additional APIs will probably be better served to either find an application server, or use Tomcat in addition to an application server. A third option is to find an implementation of the individual Java EE APIs required and use them in conjunction with Tomcat. This piecemeal approach is perfectly valid, although integration problems may manifest themselves.

Do you always need a full-fledged Java EE application server to develop enterprise applications? The short answer is "It depends on your requirements." An increasing number of Web sites eschew the traditional Java EE technologies — especially EJB — and develop fairly complex applications with "lightweight" and often open source components. These typically use an application framework such as Struts or Spring, or an object-relational mapping framework such as Hibernate — all running in a state-of-the-art Servlet container, i.e., Tomcat!

# Tomcat and Web Servers

Tomcat's purpose is to provide standards-compliant support for Servlets and JSPs. The purpose of Servlets and JSPs is to generate Web content such as HTML files or GIF files on demand, using changing data. Web content that is generated on demand is said to be *dynamic*. Conversely, Web content that never changes and is served up as is, is called *static*. Web applications commonly include a great deal of static content, such as images or Cascading Style Sheets (CSS).

While Tomcat is capable of serving dynamic and static content, many production deployments use a native Web server, such as Apache HTTP Server or IIS, to handle the static content. There are many reasons for choosing to do this, some of which relate to performance and others relate to support of legacy code. Chapters 11 and 12 address these issues in greater detail.

Recognizing that Tomcat could enjoy a synergistic relationship with conventional Web servers, the earliest versions of Tomcat included a "Connector" that enabled a Tomcat and Apache Web server to work together. In such a relationship, Apache receives all of the HTTP requests made to the Web application. Apache then recognizes which requests are intended for Servlets/JSPs, and passes these requests to Tomcat. Tomcat fulfills the request and passes the response back to Apache, which then returns the response to the requestor.

The Apache Connector was initially crucial to the Tomcat 3.*x* series, because Tomcat's support for both static content and its implementation of the HTTP protocol were somewhat limited.

Starting with the 4.*x* series, Tomcat featured a much more complete implementation of HTTP and better support for serving up static content, and should by itself be sufficient for most deployments.

If you're not using either Apache or IIS or any other Web server officially supported by Tomcat, then don't give up hope entirely. It is still very possible to integrate Tomcat with other Web servers, even one that resides on the same machine. If you wish to run them on the same machine for instance, all you have to do is to set up Tomcat and the Web server to run on different network ports. You then can then design your Web application to request its static resources from the Web server, and have Tomcat handle the requests for dynamic content.

> **In many situations it might be simpler to just use Tomcat's own Web server implementation. Tomcat has an "HTTP Connector" — i.e., a component that implements an HTTP server. More on this, including when it makes sense to use this, and when a native Web server is a better choice, is explained in Chapter 10.**

# Summary

To conclude this chapter overview of Tomcat, let's review some of the key points we discussed:

- The Apache Software Foundation (ASF) is a nonprofit organization created to provide the world with quality open source software.
- The ASF maintains an extensive collection of open source projects. Many of the ASF's Java projects are collected under the umbrella of a parent project called Jakarta.
- Tomcat started as a subproject of the Jakarta project, but now is independent of it.
- Tomcat can be freely used in any organization. It can be freely redistributed in any commercial project so long as its license is also included with the redistribution and proper recognition is given.

- Java EE is a series of Java APIs designed to facilitate the creation of complex enterprise applications. Java EE–compatible application servers provide implementations of the Java EE APIs.
- Tomcat is a Java EE–compliant Servlet container and is the official reference implementation for the Java Servlet and JavaServer Pages APIs. Tomcat also includes implementations of the JNDI and JMX APIs, but not the rest of the Java EE APIs, and is not, thus, a complete Java EE application server.
- While Tomcat can function as a Web server, it can also be integrated with other Web servers.
- Tomcat has special support for integrating with the Apache, IIS, and Netscape Enterprise Server (NES) servers, among others.

This chapter has provided a basic introduction to Tomcat. Chapter 2 describes what Tomcat-served Web applications look like and what files they comprise. It also provides a quick background to Web applications, which should be useful for administrators who do not have a background in Java technologies.