The Concurrency Utilities

Java’s original support for multithreading is, it is not

ideal for all applications—especially those that make intensive use of multiple threads. For

example, the original multithreading support does not provide several high-level features,

such as semaphores, thread pools, and execution managers, that facilitate the creation of

intensively concurrent programs.

The term concurrent program refers

to a program that makes extensive, integral use of concurrently executing threads. An example

of such a program is one that uses separate threads to simultaneously compute the partial

results of a larger computation. Another example is a program that coordinates the activities

of several threads, each of which seeks access to information in a database.

To begin to handle the needs of a concurrent program, JDK 5 added the concurrency

utilities, also commonly referred to as the concurrent API. The original set of concurrency

utilities supplied many features that had long been wanted by programmers who develop

concurrent applications. For example, it offered synchronizers (such as the semaphore),

thread pools, execution managers, locks, several concurrent collections, and a streamlined

way to use threads to obtain computational results.

Although the original concurrent API was impressive in its own right, it was significantly expanded by JDK 7. The most important addition was the Fork/Join Framework. The Fork/Join Framework facilitates the creation of programs that make use of multiple processors (such as those found in multicore systems). Thus, it streamlines the development of programs in which two or more pieces execute with true simultaneity (that is, true parallel execution),not just time-slicing.

**The Concurrent API Packages**

The concurrency utilities are contained in the **java.util.concurrent** package and in its two

subpackages: **java.util.concurrent.atomic** and **java.util.concurrent.locks**

java.util.concurrent

java.util.concurrent defines the core features that support alternatives to the built-in

approaches to synchronization and interthread communication. It defines the following

key features:

• Synchronizers

• Executors

• Concurrent collections

• The Fork/Join Framework

Synchronizers offer high-level ways of synchronizing the interactions between multiple

threads. The synchronizer classes defined by java.util.concurrent are

|  |  |
| --- | --- |
| Semaphore | Implements the classic semaphore. |
| CountDownLatch | Waits until a specified number of events have occurred. |
| CyclicBarrier | Exchanges data between two threads. |
| Phaser | Synchronizes threads that advance through multiple phases of an  operation. |

*Executors* manage thread execution. At the top of the executor hierarchy is the Executor

interface, which is used to initiate a thread. **ExecutorService** extends **Executor** and provides

methods that manage execution. There are three implementations of **ExecutorService**:

**ThreadPoolExecutor**, **ScheduledThreadPoolExecutor**, and **ForkJoinPool**. java.util.concurrent

also defines the Executors utility class, which includes a number of static methods that

simplify the creation of various executors.

Related to executors are the **Future** and **Callable** interfaces. A **Future** contains a value

that is returned by a thread after it executes. Thus, its value becomes defined “in the

future,” when the thread terminates. **Callable** defines a thread that returns a value

**java.util.concurrent** defines several concurrent collection classes, including

**ConcurrentHashMap**, **ConcurrentLinkedQueue**, and **CopyOnWriteArrayList**. These offer

concurrent alternatives to their related classes defined by the Collections Framework.

The *Fork/Join Framework* supports parallel programming. Its main classes are **ForkJoinTask**,

**ForkJoinPool**, **RecursiveTask**, and **RecursiveAction**.

**java.util.concurrent.atomic**

**java.util.concurrent.atomic** facilitates the use of variables in a concurrent environment.

It provides a means of efficiently updating the value of a variable without the use of locks.

This is accomplished through the use of classes, such as **AtomicInteger** and **AtomicLong**,

and methods, such as **compareAndSet( )**, **decrementAndGet( )**, and **getAndSet( )**. These

methods execute as a single, non-interruptible operation.

**java.util.concurrent.locks**

**java.util.concurrent.locks** provides an alternative to the use of synchronized methods. At

the core of this alternative is the **Lock** interface, which defines the basic mechanism used

to acquire and relinquish access to an object. The key methods are **lock( )**, **tryLock( )**, and

**unlock( )**. The advantage to using these methods is greater control over synchronization.

The remainder of this chapter takes a closer look at the constituents of the concurrent API.

**Semaphore**

A semaphore controls access to a shared resource

through the use of a counter. If the counter is greater than zero, then access is allowed. If

it is zero, then access is denied. What the counter is counting are *permits* that allow access to

the shared resource. Thus, to access the resource, a thread must be granted a permit from

the semaphore.

In general, to use a semaphore, the thread that wants access to the shared resource tries

to acquire a permit. If the semaphore’s count is greater than zero, then the thread acquires

a permit, which causes the semaphore’s count to be decremented. Otherwise, the thread

will be blocked until a permit can be acquired.

When the thread no longer needs access

to the shared resource, it releases the permit, which causes the semaphore’s count to be

incremented. If there is another thread waiting for a permit, then that thread will acquire

a permit at that time. Java’s **Semaphore** class implements this mechanism.

**Semaphore** has the two constructors shown here:

Semaphore(int *num*)

Semaphore(int *num*, boolean *how*)

Here, *num* specifies the initial permit count. Thus, *num* specifies the number of threads

that can access a shared resource at any one time.

By default, waiting threads are granted a permit in an

undefined order. By setting *how* to **true**, you can ensure that waiting threads are granted a

permit in the order in which they requested access.

To acquire a permit, call the **acquire( )** method, which has these two forms:

void acquire( ) throws InterruptedException

void acquire(int *num*) throws InterruptedException

The first form acquires one permit. The second form acquires *num* permits. Most often, the

first form is used. If the permit cannot be granted at the time of the call, then the invoking

thread suspends until the permit is available.

To release a permit, call **release( )**, which has these two forms:

void release( )

void release(int *num*)

The first form releases one permit. The second form releases the number of permits

specified by *num.*