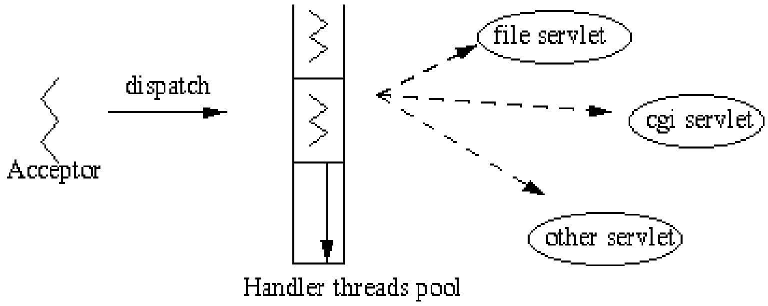
Overview of the Java HTTP Server Architecture

This document describes briefly the architecture of the generic server framework and how the http server is built using this framework. It also describes the servlet api currently used by the http server.

Base server framework

The base server framework provides a generic set of classes for implementing connection-oriented servers which manage multiple threads for handling client connections. There is one server acceptor thread that continually sits in a tight loop accepting the connections, keeping the backlog to minimum. Once it gets a connection, it dispatches it to one of the server handler threads from the worker pool. The size of the worker pool is dynamically configurable, as are a variety of additional server parameters.

The server begins by creating worker threads for handling new connections. The server then repeatedly accepts new connections and places them on a connection queue for subsequent retrieval by the handler threads. Before a new connection is dispatched to the pool of handler threads, the server checks to make sure that at least one handler thread is waiting for a connection. This ensures that new connections will not be accepted unless there is an available handler. If no handlers are currently available, then a new handler thread will be forked, as long as the limit on the number of such handlers is not exceeded. If that limit would be exceeded, the server will block until a new connection is available. Additionally, if a handler thread has been waiting a "long" time


dispatch
Acceptor
Handler threads pool
file servlet
cgi servlet
other servlet