Spring Boot Actuator

Spring Boot Actuator module helps you monitor and manage your Spring Boot application by providing production-ready features like

- health check-up
- auditing
- metrics gathering
- HTTP tracing etc.

All of these features can be accessed over **JMX** or **HTTP** endpoints.

Actuator also integrates with external application monitoring systems like:

- Prometheus
- Graphite
- DataDog
- Influx
- Wavefront
- New Relic and many more.

These systems provide you with excellent dashboards, graphs, analytics, and alarms to help you monitor and manage your application from one unified interface

Actuator uses **Micrometer** - an application metrics facade to integrate with these external application monitoring systems. This makes it super easy to plug-in any application monitoring system with very little configuration.

Example actuator endpoints:

- /**actuator** - list all the actuator endpoints exposed over HTTP
- /**health** - provides basic information about the application's health
- /**metrics** - shows several useful metrics information like JVM memory used, system CPU usage, open files, etc
- /**loggers** endpoint shows the application's logs and also lets you change the log level at runtime

You can see the complete list of the endpoints on the official documentation page: https://docs.spring.io/spring-boot/docs/current/reference/html/production-ready-endpoints.html

Note that every actuator endpoint can be explicitly enabled and disabled. Moreover, the endpoints also need to be exposed over HTTP or JMX to make them remotely accessible.

By default, **all endpoints** except for shutdown are **enabled**.

By default, only the **health** and **info** endpoints are **exposed** over HTTP.

For example, to enable the shutdown endpoint, add the following to your application.properties file:

```
management.endpoint.shutdown.enabled=true
```

Here is how you can expose actuator endpoints over HTTP and JMX using application properties:

```
#Exposing Actuator endpoints over HTTP
# Use "*" to expose all endpoints, or a comma-separated list to expose selected ones
management.endpoints.web.exposure.include=health,info
management.endpoints.web.exposure.exclude=

#Exposing Actuator endpoints over JMX
# Use "*" to expose all endpoints, or a comma-separated list to expose selected ones
management.endpoints.jmx.exposure.include=*
management.endpoints.jmx.exposure.exclude=
```

Health

The /health endpoint checks the health of your application by combining several health indicators. It uses these health indicators as part of the health check-up process.

Spring Boot Actuator comes with several predefined health indicators:

- DataSourceHealthIndicator
- DiskSpaceHealthIndicator
- MongoHealthIndicator
- RedisHealthIndicator
- CassandraHealthIndicator etc.

You can also disable a particular health indicator using application properties:

```
management.health.mongo.enabled=false
```

The health endpoint only shows a simple UP or DOWN status. To get the complete details including the status of every health indicator that was checked as part of the health check-up process, add the following property in the application.properties:

```
management.endpoint.health.show-details=always
```

Metrics

The /metrics endpoint lists all the metrics that are available for you to track.

To get the details of an individual metric, you need to pass the metric name in the URL like this -

```
http://localhost:8080/actuator/metrics/{MetricName}
```

For example, to get the details of system.cpu.usage metric, use the URL:

```
http://localhost:8080/actuator/metrics/system.cpu.usage
```

Loggers

The loggers endpoint, which can be accessed at http://localhost:8080/actuator/**loggers**, displays a list of all the configured loggers in your application with their corresponding log levels.

You can also view the details of an individual logger by passing the logger name in the URL like this:

```
http://localhost:8080/actuator/loggers/{name}
```

For example, to get the details of the root logger, use the URL http://localhost:8080/actuator/loggers/**root**.

The loggers endpoint also allows you to **change the log level of a given logger in your application at runtime**.

For example, To change the log level of the root logger to DEBUG at runtime, make a POST:

```
POST http://localhost:8080/actuator/loggers/root
{
  "configuredLevel": "DEBUG"
```

```
}
```

This functionality will really be useful in cases when your application is facing issues in production and you want to enable DEBUG logging for some time to get more details about the issue.

To reset the log-level to the default value, you can pass a value of null in the configuredLevel field.

Security

If Spring Security is present in your application, then the endpoints are secured by default using a form-based HTTP basic authentication.

Override existing security configuration:

```
@Override
	protected void configure(HttpSecurity http) throws Exception {
	http
		.authorizeRequests()
		  .requestMatchers(EndpointRequest.to(ShutdownEndpoint.class))
		    .hasRole("ACTUATOR_ADMIN")
		  .requestMatchers(EndpointRequest.toAnyEndpoint())
		    .permitAll()
		  .requestMatchers(PathRequest.toStaticResources().atCommonLocations())
		    .permitAll()
		  .antMatchers("/")
		    .permitAll()
		  .antMatchers("/**")
		    .authenticated()
		.and()
		  .httpBasic();
	}
```

To be able to test the above configuration with HTTP basic authentication, you can add a default spring security user in application.properties:

```
# Spring Security Default username and password
spring.security.user.name=actuator
spring.security.user.password=actuator
spring.security.user.roles=ACTUATOR_ADMIN
```

Prometheus

Prometheus is an open-source monitoring system that was originally built by SoundCloud. It consists of the following core components -

- A **data scraper** that pulls metrics data over HTTP periodically at a configured interval.
- A time-series **database** to store all the metrics data.
- A simple **user interface** where you can visualize, query, and monitor all the metrics.

Actuator endpoint: /actuator/prometheus

Spring Sleuth / Zipkin

https://howtodoinjava.com/spring-cloud/spring-cloud-zipkin-sleuth-tutorial/

Sleuth is a tool from the **Spring cloud** family. It is used to generate the trace id, span id and add this information to the service calls in the headers and MDC, so that It can be used by tools like Zipkin and ELK etc. to store, index and process log files.

As it is from the spring cloud family, once added to the CLASSPATH, it automatically integrated to the common communication channels like:

- requests made with the RestTemplate etc.
- requests that pass through a Netflix Zuul microproxy
- HTTP headers received at Spring MVC controllers
- requests over messaging technologies like Apache Kafka or RabbitMQ etc.

```
<dependency>
   <groupId>org.springframework.cloud</groupId>
   <artifactId>spring-cloud-starter-sleuth</artifactId>
</dependency>
```

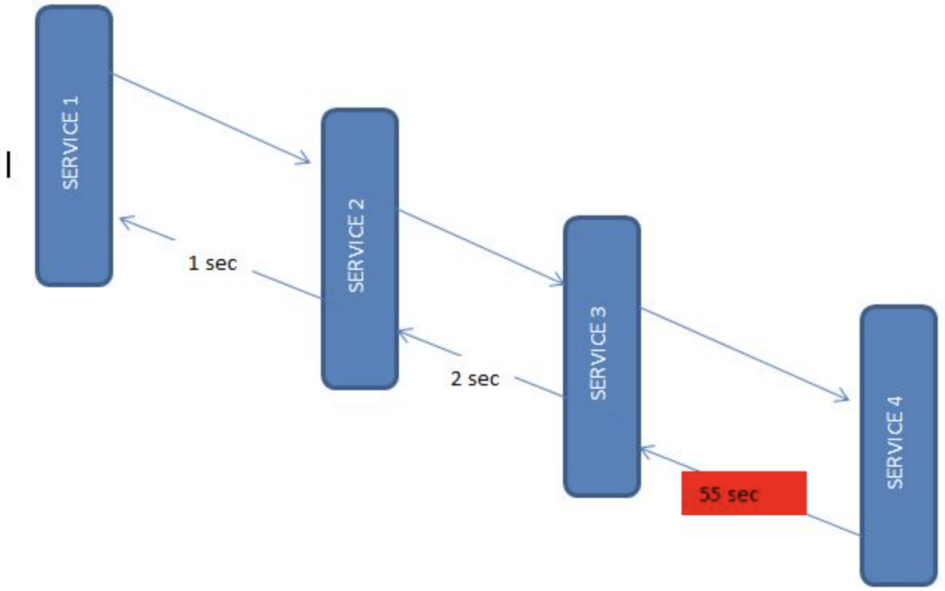
Zipkin was originally developed at Twitter, based on a concept of a Google paper that described Google's internally-built distributed app debugger – dapper. It manages both the collection and lookup of this data. To use Zipkin, applications are instrumented to report timing data to it.

If you are troubleshooting latency problems or errors in an ecosystem, you can filter or sort all traces based on the application, length of trace, annotation, or timestamp. By analyzing these traces, you can decide which components are not performing as per expectations, and you can fix them.

Internally it has 4 modules:

- Collector – Once any component sends the trace data, it arrives to Zipkin collector daemon. Here the trace data is validated, stored, and indexed for lookups by the Zipkin collector.
- Storage – This module store and index the lookup data in backend. **Cassandra, ElasticSearch and MySQL** are supported.
- Search – This module provides a simple JSON API for finding and retrieving traces stored in the backend. The primary consumer of this API is the Web UI.
- Web UI – A very nice UI interface for viewing traces.

```
<dependency>
    <groupId>org.springframework.cloud</groupId>
    <artifactId>spring-cloud-starter-zipkin</artifactId>
</dependency>
```



Reactive programming

https://www.baeldung.com/spring-webflux-concurrency
https://www.educba.com/rxjava-vs-reactor/
https://docs.spring.io/spring-framework/docs/current/reference/html/web-reactive.html#webflux-new-framework

DeferredResult

DeferredResult provides an alternative to using a **Callable** for asynchronous request processing. While a Callable is executed concurrently on behalf of the application, with a DeferredResult the application can produce the result from a thread of its choice.

Your controller is eventually a function executed by the servlet container (I will assume it is Tomcat) worker thread. Your service flow starts with Tomcat and ends with Tomcat. Tomcat gets the request from the client, holds the connection, and eventually returns a response to the client. Your code (controller or servlet) is somewhere in the middle.

Consider this flow:

- Tomcat get client request.
- Tomcat executes your controller.
- Release Tomcat thread but keep the client connection (don't return response) and run heavy processing on different thread.
- When your heavy processing complete, update Tomcat with its response and return it to the client (by Tomcat).

Because the servlet (your code) and the servlet container (Tomcat) are different entities, then to allow this flow (releasing tomcat thread but keep the client connection) we need to have this support in their contract, the package javax.servlet, which introduced in **Servlet 3.0**.

About servlet container worker threads: we can configure Tomcat's server thread pool:

```
server.tomcat.threads.min-spare: 10
server.tomcat.threads.max: 200
```

minSpare is the smallest the pool will be, including at startup. maxThreads is the largest the pool will be before the server starts queueing up requests.

Spring Boot defaults these to 10 and 200, respectively.

How and when to use DeferredResult?

With DeferredResult you can choose any thread (or threadpool) to run your async code, which has some beneficial advantages:

- Http worker threads released immediately by passing on the work to the non HTTP thread.
- The worker threads are able to handle additional HTTP request without waiting for the long-running process to finish the work.