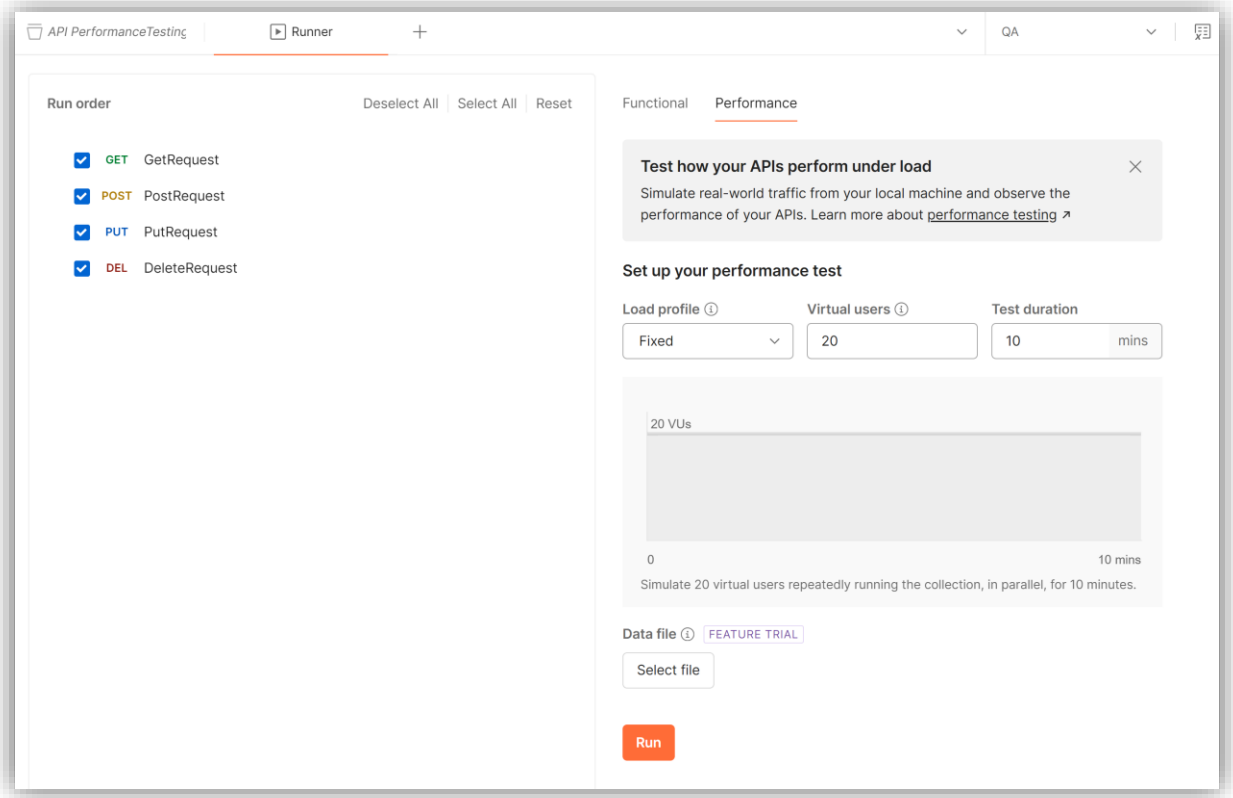
API Performance Testing

- Use the **Collection Runner** to test the **performance of your API** with the same requests and collections you use for functional API tests.
- When you run a performance test, Postman uses the requests in the selected collection to simulate the activity of your API users.
- In the Collection Runner, you can set the **duration of the test** and the **number of virtual users**.
- Each virtual user runs the requests in the specified order in a repeating loop.
- All of the virtual users operate in parallel to simulate real-world load on your API.
- You can choose whether the number of virtual users is **fixed** for the duration of the test or **ramps up** and down during the test.



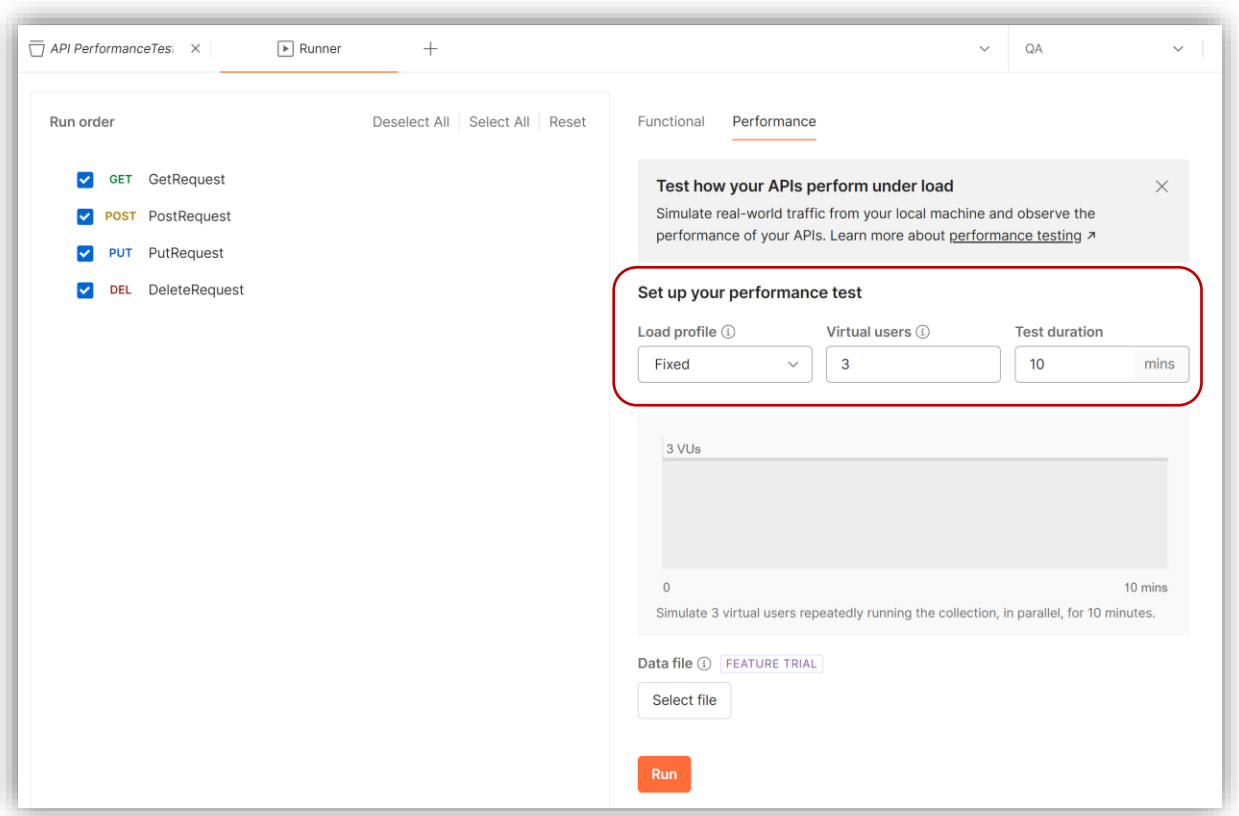
Load profile:

Fixed - The maximum number of virtual users is used throughout the test.

Ramp up - Enter an **Initial load** and drag the handles to adjust the length of the ramp up period. During the ramp up period, the number of virtual users increases from the initial load to the maximum.

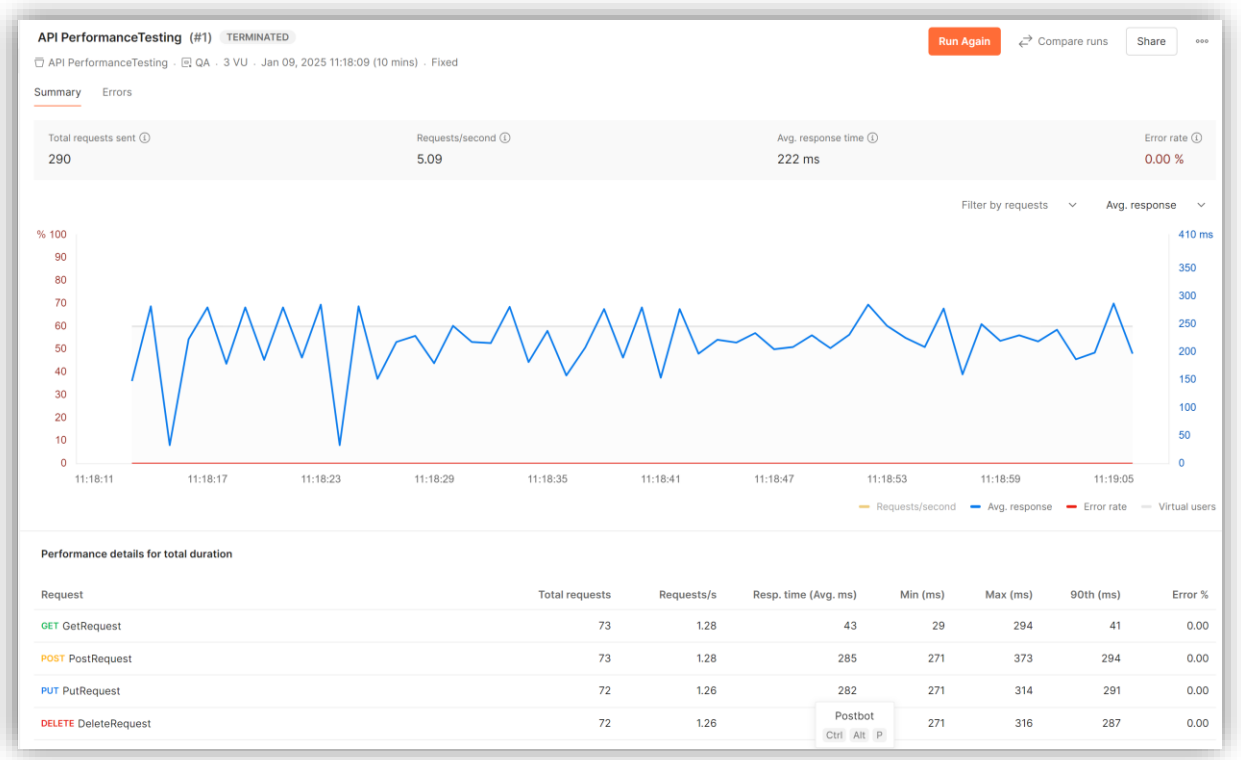
Spike - Enter a **Base load** and drag the handles to adjust the duration of the spike. During the spike, the number of virtual users increases from the base load to the maximum, then decreases back to the base load.

Peak - Enter a **Base load** and drag the handles to adjust the duration of the peak. During the peak, the number of virtual users increases from the base load to the maximum, holds steady, then decreases back to the base load.



View performance metrics

The **Summary** tab displays performance metrics in real time so you can observe how your API performs under load.



You can view the following information about the performance test:

- Name of the collection being run and the active environment (if any). Select the name to open the collection or environment.
- The number of virtual users (VU). Each virtual user simulates the behavior of a real-world user by running the collection in a repeating loop, in parallel with other virtual users.
- Start time, duration, load profile (fixed, ramp up, spike, or peak), and data file (if used).
- **Total requests sent** - The total number of requests sent across all virtual users.
- **Requests/second** - The number of requests sent each second during the performance test, a measure of throughput.
- **Avg. response time** - The response time in milliseconds of all requests averaged together.
- **Error rate** - The percentage of requests that result in an error. Responses other than 2xx responses are considered errors.

From the Summary tab, we can take the following actions:

- Select **Stop** to stop the performance test.
- Select **Run Again** to configure and start the performance test again.
- To get a link to the performance report, select **Share**. Select **Copy Link** to copy the link to your clipboard. You can share the link with your teammates so they can view the results of the performance test.
- To download a copy of the performance report in PDF or HTML format, select the more actions icon ⁰⁰⁰ and select **Export PDF report** or **Export HTML report**. The report shows response time and throughput, slowest requests, requests with the most errors, metrics for each request, and a list of errors.
- To view a list of all past performance runs for the collection, select the more actions icon ⁰⁰⁰ and select **View all runs**.
- To open a request, select the name of a request.