MLBench Dashboard Documentation

MLBench development team

Jun 09, 2021

MLBENCH

1 MLBench Dashboard **1**
- 1.1 Dashboard Functionality 1
- 1.2 REST API 2

2 Indices and tables **11**

HTTP Routing Table **13**

CHAPTER

ONE

MLBENCH DASHBOARD

MLBench comes with a dashboard to manage and monitor the cluster and jobs.

1.1 Dashboard Functionality

1.1.1 Main Page

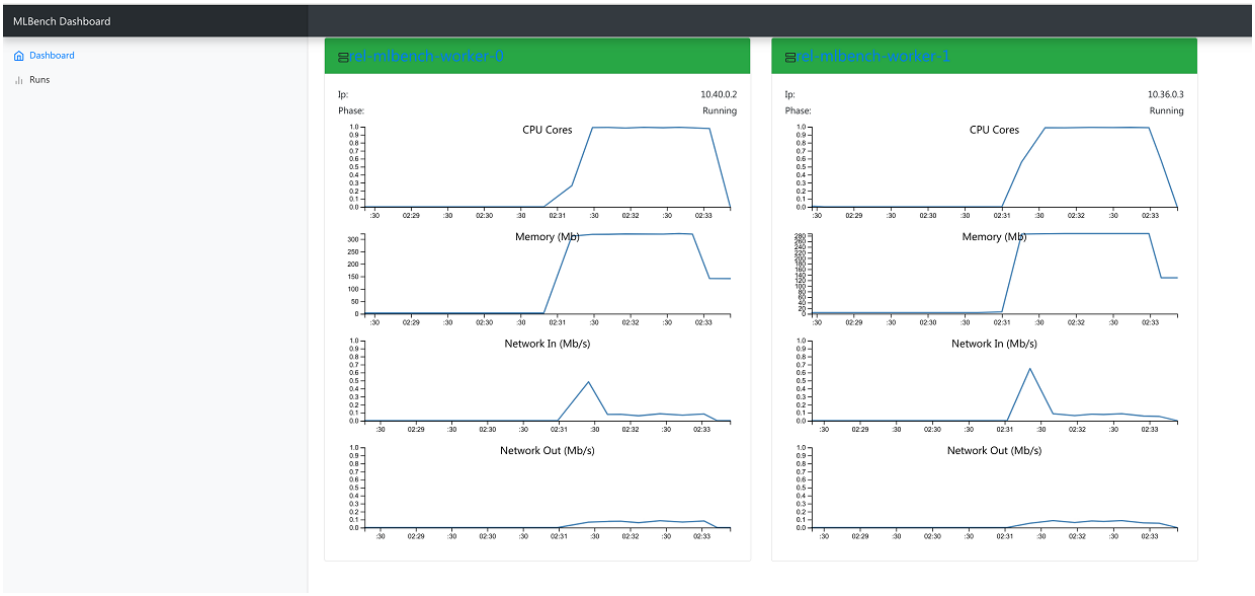

Fig. 1: Dashboard Main Page

The main view shows all MLBench worker nodes and their current status

1.1.2 Runs Page

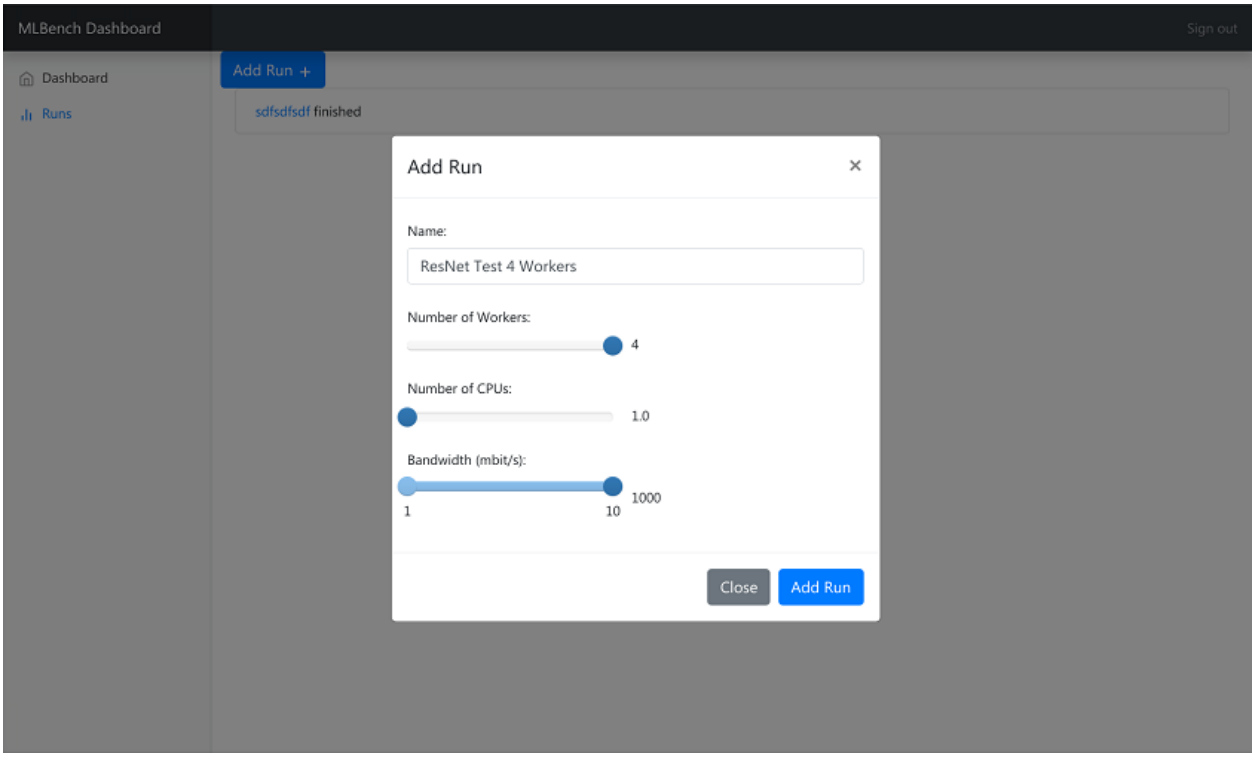

Fig. 2: Dashboard Runs Page

The Runs page allows you to start a new experiment on the worker nodes. You can select how many workers to use and how many CPU Cores each worker can utilize.

1.1.3 Run Details Page

The Run Details page shows the progress and result of an experiment. You can track metrics like `train loss` and `validation accuracy` as well as see the `stdout` and `stderr` logs of all workers.

It also allows you to download all the metrics of a run as well as resource usage of all workers participating in the run as json files.

1.2 REST API

MLBench provides a basic REST Api though which most functionality can also be used. It's accessible through the /api/ endpoints on the dashboard URL.

Fig. 3: Dashboard Run Details Page

1.2.1 Pods

GET /api/pods/

All Worker-Pods available in the cluster, including status information

Example request:

```
GET /api/pods HTTP/1.1
Host: example.com
Accept: application/json, text/javascript
```

Example response:

```
HTTP/1.1 200 OK
Vary: Accept
Content-Type: text/javascript

[
  {
    "name":"worn-mouse-mlbench-worker-55bbdd4d8c-4mxh5",
    "labels":"{'app': 'mlbench', 'component': 'worker', 'pod-template-hash':
↪'1166880847', 'release': 'worn-mouse'}",
    "phase":"Running",
    "ip":"10.244.2.58"
  },
  {
    "name":"worn-mouse-mlbench-worker-55bbdd4d8c-bwwsp",
    "labels":"{'app': 'mlbench', 'component': 'worker', 'pod-template-hash':
↪'1166880847', 'release': 'worn-mouse'}",
```

(continues on next page)

(continued from previous page)

```
    "phase":"Running",
    "ip":"10.244.3.57"
  }
]
```

Request Headers

- Accept – the response content type depends on *Accept* header

Response Headers

- Content-Type – this depends on *Accept* header of request

Status Codes

- 200 OK – no error

1.2.2 Metrics

GET /api/metrics/

Get metrics (Cpu, Memory etc.) for all Worker Pods

Example request:

```
GET /api/metrics HTTP/1.1
Host: example.com
Accept: application/json, text/javascript
```

Example response:

```
HTTP/1.1 200 OK
Vary: Accept
Content-Type: text/javascript

{
  "quiet-mink-mlbench-worker-0": {
      "container_cpu_usage_seconds_total": [
          {
              "date": "2018-08-03T09:21:38.594282Z",
              "value": "0.188236813"
          },
          {
              "date": "2018-08-03T09:21:50.244277Z",
              "value": "0.215950298"
          }
      ]
  },
  "quiet-mink-mlbench-worker-1": {
      "container_cpu_usage_seconds_total": [
          {
              "date": "2018-08-03T09:21:29.347960Z",
              "value": "0.149286015"
          },
          {
```

(continues on next page)

(continued from previous page)

```
                "date": "2018-08-03T09:21:44.266181Z",
                "value": "0.15325329"
            }
        ],
        "container_cpu_user_seconds_total": [
            {
                "date": "2018-08-03T09:21:29.406238Z",
                "value": "0.1"
            },
            {
                "date": "2018-08-03T09:21:44.331823Z",
                "value": "0.1"
            }
        ]
    }
}
```

Request Headers

- Accept – the response content type depends on *Accept* header

Response Headers

- Content-Type – this depends on *Accept* header of request

Status Codes

- 200 OK – no error

GET /api/metrics/(str: *pod_name_or_run_id*)/

Get metrics (Cpu, Memory etc.) for all Worker Pods

Example request:

```
GET /api/metrics HTTP/1.1
Host: example.com
Accept: application/json, text/javascript
```

Example response:

```
HTTP/1.1 200 OK
Vary: Accept
Content-Type: text/javascript

{
    "container_cpu_usage_seconds_total": [
        {
            "date": "2018-08-03T09:21:29.347960Z",
            "value": "0.149286015"
        },
        {
            "date": "2018-08-03T09:21:44.266181Z",
            "value": "0.15325329"
        }
    ],
```

(continues on next page)

(continued from previous page)

```
  "container_cpu_user_seconds_total": [
      {
          "date": "2018-08-03T09:21:29.406238Z",
          "value": "0.1"
      },
      {
          "date": "2018-08-03T09:21:44.331823Z",
          "value": "0.1"
      }
  ]
}
```

Query Parameters

- **since** – only get metrics newer than this date, (Default *1970-01-01T00:00:00.000000Z*)
- **metric_type** – one of *pod* or *run* to determine what kind of metric to get (Default: *pod*)

Request Headers

- Accept – the response content type depends on *Accept* header

Response Headers

- Content-Type – this depends on *Accept* header of request

Status Codes

- 200 OK – no error

`POST /api/metrics`

Save metrics. "pod_name" and "run_id" are mutually exclusive. The fields of metrics and their types are defined in *mlbench/api/models/kubemetrics.py*.

Example request:

```
POST /api/metrics HTTP/1.1
Host: example.com
Accept: application/json, text/javascript

{
  "pod_name": "quiet-mink-mlbench-worker-1",
  "run_id": 2,
  "name": "accuracy",
  "date": "2018-08-03T09:21:44.331823Z",
  "value": "0.7845",
  "cumulative": False,
  "metadata": "some additional data"
}
```

Example response:

```
HTTP/1.1 201 CREATED
Vary: Accept
Content-Type: text/javascript
```

(continues on next page)

(continued from previous page)

```
{
  "pod_name": "quiet-mink-mlbench-worker-1",
  "name": "accuracy",
  "date": "2018-08-03T09:21:44.331823Z",
  "value": "0.7845",
  "cumulative": False,
  "metadata": "some additional data"
}
```

Request Headers

- Accept – the response content type depends on *Accept* header

Response Headers

- Content-Type – this depends on *Accept* header of request

Status Codes

- 201 Created – no error

1.2.3 Runs

GET /api/runs/

Gets all active/failed/finished runs

Example request:

```
GET /api/runs/ HTTP/1.1
Host: example.com
Accept: application/json, text/javascript
```

Example response:

```
HTTP/1.1 200 OK
Vary: Accept
Content-Type: text/javascript

[
  {
    "id": 1,
    "name": "Name of the run",
    "created_at": "2018-08-03T09:21:29.347960Z",
    "state": "STARTED",
    "job_id": "5ec9f286-e12d-41bc-886e-0174ef2bddae",
    "job_metadata": {...}
  },
  {
    "id": 2,
    "name": "Another run",
    "created_at": "2018-08-02T08:11:22.123456Z",
    "state": "FINISHED",
    "job_id": "add4de0f-9705-4618-93a1-00bbc8d9498e",
    "job_metadata": {...}
```

(continues on next page)

(continued from previous page)

```
  },
]
```

Request Headers

- Accept – the response content type depends on *Accept* header

Response Headers

- Content-Type – this depends on *Accept* header of request

Status Codes

- 200 OK – no error

GET /api/runs/(int: *run_id*)/

Gets a run by id

Example request:

```
GET /api/runs/1/ HTTP/1.1
Host: example.com
Accept: application/json, text/javascript
```

Example response:

```
HTTP/1.1 200 OK
Vary: Accept
Content-Type: text/javascript

{
  "id": 1,
  "name": "Name of the run",
  "created_at": "2018-08-03T09:21:29.347960Z",
  "state": "STARTED",
  "job_id": "5ec9f286-e12d-41bc-886e-0174ef2bddae",
  "job_metadata": {...}
}
```

:run_id The id of the run

Request Headers

- Accept – the response content type depends on *Accept* header

Response Headers

- Content-Type – this depends on *Accept* header of request

Status Codes

- 200 OK – no error

POST /api/runs/

Starts a new Run

Example request:

```
POST /api/runs/ HTTP/1.1
Host: example.com
Accept: application/json, text/javascript
```

Request JSON Object

- **name** (*string*) – Name of the run
- **num_workers** (*int*) – Number of worker nodes for the run
- **num_cpus** (*json*) – Number of Cores utilized by each worker

Example response:

```
HTTP/1.1 200 OK
Vary: Accept
Content-Type: text/javascript

{
  "id": 1,
  "name": "Name of the run",
  "created_at": "2018-08-03T09:21:29.347960Z",
  "state": "STARTED",
  "job_id": "5ec9f286-e12d-41bc-886e-0174ef2bddae",
  "job_metadata": {...}
}
```

Request Headers

- Accept – the response content type depends on *Accept* header

Response Headers

- Content-Type – this depends on *Accept* header of request

Status Codes

- 200 OK – no error
- 409 Conflict – a run is already active

CHAPTER

TWO

INDICES AND TABLES

- genindex
- modindex
- search

HTTP ROUTING TABLE

/api

GET /api/metrics/, 4
GET /api/metrics/(str:pod_name_or_run_id)/, 5
GET /api/pods/, 3
GET /api/runs/, 7
GET /api/runs/(int:run_id)/, 8
POST /api/metrics, 6
POST /api/runs/, 8