Grigorios Dimopoulos
M1548

Dimitris Sartzetakis
M1561

LogDb-2

20th January 2020

OVERVIEW

LogDb-2 is a web application that given some log files of specific form will parse them, add them in mongodb and give you an endpoint interface to query/add/edit/delete the logs, and execute more advanced queries as well.

Technologies used for the project

We implemented the project using, MongoDb, Gradle, Java 8 with SpringBoot, Spring Data. Our IDEA of preference was IntelliJ.

Database schema

logs collection:

```
[
  {
    "_id" : ObjectId,
    "userId" : String,
    "httpMethod" : HttpMethodEnum,
    "resource" : String,
    "userAgent" : String,
    "status" : Integer,
    "timestamp" : ISODate,
    "type" : TypeEnum,
    "size" : Long,
    "sourceIp" : String,
    "upvoters" : List<Upvoter(username: String, email: String)>,
    "_class" : "com.logdb2.document.Access"
  },
  {
    "_id" : ObjectId,
    "blockIds" : Long,
    "destinationIps" : String,
    "timestamp" : ISODate,
    "type" : TypeEnum,
    "sourceIp" : String,
    "upvoters" : List<Upvoter(username: String, email: String)>,
    "_class" : "com.logdb2.document.Dataxceiver"
  },
  {
    "_id" : ObjectId,
    "blockIds" :List<Long>,
    "destinationIps" : List<String>,
    "timestamp" : ISODate,
```

```
        "type" : TypeEnum,

        "sourceIp" : String,

        "upvoters" : List<Upvoter(username: String, email: String)>,

        "_class" : "com.logdb2.document.Namesystem"

    }

]
```

admin collection:

```
[

    {

        "_id" : ObjectId,

        "username" : String,

        "email" : String,

        "phoneNumber" : String,

        "timestamp" : ISODate,

        "upvotes" : List<Upvote(log: ObjectId, sourceIp: String)>,

        "_class" : "com.logdb2.document.Admin"

    }

]
```

Indices in logs collection:

- referer: {"referer" : 1}

- httpMethod: {"httpMethod" : 1}

- timestamp: {"timestamp" : 1}

- type: {"type" : 1}

- upvoters: {"upvoters.username" : 1, "upvoters.email" : 1}

- blockIds: {"blockIds" : 1}

Indices in admin collection:

- username: {"username" : 1}

- email: {"email" : 1}

- phoneNumber: {"phoneNumber" : 1}

- timestamp: {"timestamp" : 1}

- upvotes.log: {"upvotes.log" : 1}

- upvotes.sourceIp: {"upvotes.sourceIp" : 1}

Notes

We changed the schema according to our needs for the queries we needed to implement.
Also the indexes that have been added helped to make the execution faster.

LogDb 2 sample outputs

- End-point 1:

GET
http://localhost:9090/logs/total/per-type/created?start=2015-01-13T00:00:00&stop=2020-01-01T0
0:00:00

```json
[
    {
        "type": "access",
        "total": 3353488
    },
    {
        "type": "receiving",
        "total": 1723232
    },
    {
        "type": "served",
        "total": 428726
    },
    {
        "type": "delete",
        "total": 21784
    },
    {
        "type": "received",
        "total": 7097
    },
    {
        "type": "replicate",
        "total": 7002
    }
]
```

- End-point 2:

GET
http://localhost:9090/requests/total/per-day?start=2016-01-13T00:00:00&stop=2017-01-01T00:00:00&log-type=access

```
1    [
2        {
3            "timestamp": "2016-01-25T22:26:37",
4            "total": 103
5        },
6        {
7            "timestamp": "2016-01-28T22:28:00",
8            "total": 82
9        },
10       {
11           "timestamp": "2016-08-16T11:38:35",
12           "total": 67
13       },
14       {
15           "timestamp": "2016-01-28T22:28:01",
16           "total": 59
17       },
18       {
19           "timestamp": "2016-07-08T08:55:38",
20           "total": 59
21       },
22       {
23           "timestamp": "2016-05-02T16:01:46",
24           "total": 59
25       },
26       {
27           "timestamp": "2016-08-19T08:31:41",
28           "total": 58
```

● End-point 3:

GET http://localhost:9090/logs/three-most-common/per-source-ip?date=2015-12-12

```
1    [
2         {
3             "sourceIp": "191.182.199.16",
4             "total": 37
5         },
6         {
7             "sourceIp": "188.45.108.168",
8             "total": 31
9         },
10        {
11            "sourceIp": "83.167.113.100",
12            "total": 8
13        }
14   ]
```

● End-point 4:

GET
http://localhost:9090/http-methods/two-least-common?start=2015-01-13T00:00:00&stop=2020-01-01T00:00:00

```
1    [
2         "DELETE",
3         "OPTIONS"
4    ]
```

- End-point 5:

GET http://localhost:9090/referers/with-resources

```
 1  [
 2      {
 3          "referer": "http://almhuette-raith.at/",
 4          "total": 131474
 5      },
 6      {
 7          "referer": "almhuette-raith.at",
 8          "total": 1239
 9      },
10      {
11          "referer": "http://site.ru",
12          "total": 1214
13      },
14      {
15          "referer": "set(['https://www.google.com/', 'https://www.youtube.com', 'https://duckduckgo.com/'])",
16          "total": 362
17      },
18      {
19          "referer": "sfish'\\\"'\\\"'\\\"'\\\"",
20          "total": 260
21      },
22      {
23          "referer": "sfish'''\\\"\\\"\\\"\\\"",
24          "total": 260
25      },
26      {
27          "referer": "sfish\\\\\\\\'\\\\\\\\\\\"",
28          "total": 258
29      },
30      {
31          "referer": "sfish\\\\'\\\\\\\\\"",
32          "total": 258
33      },
```

- End-point 6:

GET http://localhost:9090/blocks/replicated-served-same-day

```
1   [
2       {
3           "blockIds": 2806786215629483376,
4           "year": 2017,
5           "month": 11,
6           "day": 8,
7           "types": [
8               "served",
9               "replicate"
10          ]
11      },
12      {
13          "blockIds": -7722587184407799771,
14          "year": 2017,
15          "month": 11,
16          "day": 8,
17          "types": [
18              "replicate",
19              "served"
20          ]
21      },
22      {
23          "blockIds": 307873766972496113,
24          "year": 2017,
25          "month": 11,
26          "day": 8,
27          "types": [
28              "served",
29              "replicate"
30          ]
31      },
```

- End-point 7:

GET http://localhost:9090/logs/fifty-most-upvoted?date=2020-01-01

```
 1  [
 2      {
 3          "upvotes": "5e2581fe0291de204934a123",
 4          "total": 1
 5      },
 6      {
 7          "upvotes": "5e25816e0291de2049f31e44",
 8          "total": 1
 9      },
10      {
11          "upvotes": "5e2581d00291de20491df944",
12          "total": 1
13      },
14      {
15          "upvotes": "5e2581e50291de20492752a6",
16          "total": 1
17      },
18      {
19          "upvotes": "5e2581ee0291de20492beae7",
20          "total": 1
21      },
```

- End-point 8:

GET http://localhost:9090/admins/fifty-most-active

```
3029        {
3030            "_id": "5e25834f0291de2049363ab8",
3031            "username": "bert.heathcote",
3032            "email": "erin.labadie@hotmail.com",
3033            "phoneNumber": "(197) 022-6687 x2038",
3034  >         "upvotes": [ ...
4035            ],
4036            "size": 1000
4037        },
4038        {
4039            "_id": "5e25834f0291de204936403e",
4040            "username": "bobbie.zulauf",
4041            "email": "sterling.grant@yahoo.com",
4042            "phoneNumber": "354-849-3991 x557",
4043  >         "upvotes": [ ...
5043            ],
5044            "size": 999
5045        },
5046        {
5047            "_id": "5e25834f0291de2049363b78",
5048            "username": "dustin.gutmann",
5049            "email": "marc.schiller@yahoo.com",
5050            "phoneNumber": "743-187-6607 x1955",
5051  >         "upvotes": [ ...
6051            ],
6052            "size": 999
6053        },
```

- End-point 9:

GET http://localhost:9090/admins/most-upvotes-ips/top-fifty

```
 1   [
 2       {
 3           "_id": "5e25834f0291de2049363b03",
 4           "set_size": 416
 5       },
 6       {
 7           "_id": "5e25834f0291de20493638f0",
 8           "set_size": 415
 9       },
10       {
11           "_id": "5e25834f0291de2049363bca",
12           "set_size": 415
13       },
14       {
15           "_id": "5e25834f0291de204936387c",
16           "set_size": 414
17       },
18       {
19           "_id": "5e25834f0291de20493635dc",
20           "set_size": 414
21       },
22       {
23           "_id": "5e25834f0291de2049363a17",
24           "set_size": 413
25       },
```

- End-point 10:

GET http://localhost:9090/logs/same-email-different-usernames-logs

```
 1   [
 2       {
 3           "_id": "5e25814d0291de2049e1a9d9",
 4           "email": "nona.schowalter@yahoo.com",
 5           "usernames": [
 6               "tamela.hamill",
 7               "chelsey.jones"
 8           ]
 9       },
10       {
11           "_id": "5e25814d0291de2049e20790",
12           "email": "nona.schowalter@yahoo.com",
13           "usernames": [
14               "tamela.hamill",
15               "chelsey.jones"
16           ]
17       },
```

- End-point 11:

GET http://localhost:9090/blocks/in-upvoted-logs?username=reginia.bayer

```
 1    [
 2          {
 3              "_id": "5e2581c20291de2049161a6c",
 4              "blockId": -449562607211085345
 5          },
 6          {
 7              "_id": "5e2581c20291de204917680c",
 8              "blockId": -300098183527097334
 9          },
10          {
11              "_id": "5e2581c80291de20491a77cb",
12              "blockId": -3417298775104673827
13          },
14          {
15              "_id": "5e2581c80291de20491a94e0",
16              "blockId": 5832786780927065674
17          },
18          {
19              "_id": "5e2581c80291de20491ad31d",
20              "blockId": 2947156443023254610
21          },
22          {
23              "_id": "5e2581d00291de20491c6027",
24              "blockId": -5946869697429969008
25          },
26          {
27              "_id": "5e2581d00291de20491d4a08",
28              "blockId": -9123023868494514948
29          },
```

● Create:

POST http://localhost:9090/logs/access

REQUEST BODY:

```
{
    "httpMethod" : 1,
    "resource" : "dimitris",
    "userAgent" : "Mozilla/5.0 (Windows NT 6.0; rv:34.0) Gecko/20100101 Firefox/34.0"
}
```

```
{
    "_id" : ObjectId("5e25b1d40291de3dee8c01f9"),
    "httpMethod" : 1,
    "resource" : "dimitris",
    "userAgent" : "Mozilla/5.0 (Windows NT 6.0; rv:34.0) Gecko/20100101 Firefox/34.0",
    "_class" : "com.logdb2.document.Access"
}
```

● Update:

POST http://localhost:9090/logs/dataxceiver

REQUEST BODY:

```
{
    "_id": "5e2581c20291de204914d236",
    "size" : 1,
    "destinationIps" : "dimitris"
}
```

Before:

```
{
    "_id" : ObjectId("5e2581c20291de204914d236"),
    "blockIds" : NumberLong(-3544583377289625738),
    "destinationIps" : "10.250.19.102",
    "timestamp" : ISODate("2016-11-08T20:35:23.000Z"),
    "type" : 1,
    "sourceIp" : "10.250.11.100",
    "upvoters" : [],
    "_class" : "com.logdb2.document.Dataxceiver"
}
```

After:

```
{
    "_id" : ObjectId("5e2581c20291de204914d236"),
    "destinationIps" : "dimitris",
    "size" : NumberLong(1),
    "_class" : "com.logdb2.document.Dataxceiver"
}
```

● DirtyUpdate:

Notes: Updates only the fields that we pass (only the dirty fields).

POST http://localhost:9090/dirty/logs/namesystem

REQUEST BODY:

{

   "_id": "5e2582030291de204935c3f8",

      "sourceIp" : "IP",

      "destinationIps" : ["dimitris", "grigoris"]

}

Before:

```
{
    "_id" : ObjectId("5e2582030291de204935c3f8"),
    "blockIds" : [
        NumberLong(-1608999687919862906)
    ],
    "destinationIps" : [
        "10.251.74.79",
        "10.251.107.19"
    ],
    "timestamp" : ISODate("2016-11-08T20:35:24.000Z"),
    "type" : 4,
    "sourceIp" : "10.251.215.16",
    "upvoters" : [],
    "_class" : "com.logdb2.document.Namesystem"
}
```

After:

```
{
    "_id" : ObjectId("5e2582030291de204935c3f8"),
    "blockIds" : [
        NumberLong(-1608999687919862906)
    ],
    "destinationIps" : [
        "10.251.74.79",
        "10.251.107.19",
        "dimitris",
        "grigoris"
    ],
    "timestamp" : ISODate("2016-11-08T20:35:24.000Z"),
    "type" : 4,
    "sourceIp" : "IP",
    "upvoters" : [],
    "_class" : "com.logdb2.document.Namesystem"
}
```

- Upvote:

Notes: /admins/{adminId}/upvote/{logId}

POST

http://localhost:9090/admins/5e25834f0291de2049363469/upvote/5e25814d0291de2049e1a69
9

Admin before upvote:

```
{
    "_id" : ObjectId("5e25834f0291de2049363469"),
    "username" : "tamela.hamill",
    "email" : "nona.schowalter@yahoo.com",
    "phoneNumber" : "729.595.1281",
    "timestamp" : ISODate("2020-01-17T00:00:00.000Z"),
    "upvotes" : [
        {
```

Notes: On admin after the upvote, the specified log will be appended in the upvotes list..

Log upvote before:

```
{
    "_id" : ObjectId("5e25814d0291de2049e1a699"),
    "userId" : "-",
    "referer" : "http://almhuette-raith.at/administrator/",
    "httpMethod" : 1,
    "resource" : "/administrator/index.php",
    "userAgent" : "Mozilla/5.0 (Windows NT 6.0; rv:34.0) Gecko/20100101 Firefox/34.0",
    "status" : 200,
    "timestamp" : ISODate("2015-12-12T17:25:11.000Z"),
    "type" : 0,
    "size" : NumberLong(4494),
    "sourceIp" : "109.169.248.247",
    "upvoters" : [
        {
            "username" : "carlee.schamberger",
            "email" : "tamatha.gulgowski@gmail.com"
        }
    ],
    "_class" : "com.logdb2.document.Access"
}
```

Log upvote after:

```
{
    "_id" : ObjectId("5e25814d0291de2049e1a699"),
    "userId" : "-",
    "referer" : "http://almhuette-raith.at/administrator/",
    "httpMethod" : 1,
    "resource" : "/administrator/index.php",
    "userAgent" : "Mozilla/5.0 (Windows NT 6.0; rv:34.0) Gecko/20100101 Firefox/34.0",
    "status" : 200,
    "timestamp" : ISODate("2015-12-12T17:25:11.000Z"),
    "type" : 0,
    "size" : NumberLong(4494),
    "sourceIp" : "109.169.248.247",
    "upvoters" : [
        {
            "username" : "carlee.schamberger",
            "email" : "tamatha.gulgowski@gmail.com"
        },
        {
            "username" : "tamela.hamill",
            "email" : "nona.schowalter@yahoo.com"
        }
    ],
    "_class" : "com.logdb2.document.Access"
}
```

Build and Run:

To run the application you can simply run:

Edit the application.properties file to match your mongoDb configuration.
Then you can start the application with gradle using:

./gradlew bootRun