Container Native FS Interposer

Julia Hua, Jiawei Xiang, Hilario Gonzalez, Juncheng Cao
Mentors: Vasily Tarasov, Alex Merenstein

November 20, 2024

Recap

Last sprint

- Package and manage deployment with Helm/Helmfile
- Implement metric collection
- Integrate Faulty fs with Jaeger
- Fix TokenBucket implementation

This sprint

- Add postgresql and fio as test programs
- Implement metric collection FUSE layer
- Make FUSE layers individually configurable
- Implement tracing across multiple related operations
- Adding robustness to FaultyFs

PostgreSQL

PostgreSQL

The World's Most Advanced Open Source Relational Database

- One of the most widely used database
- Good representative of real world workloads
- Builtin benchmarking tool: pgbench

FIO

FIO

Flexible I/O tester

- Generic testing tool for generating various IO patterns
- Capable of recording and replaying IO traces
- A large collection of existing testing scripts

Limitations

IO tracing is done by `blktrace` at block level, not nessarily suitable for virtual filesystems

Configuring utilities


```
{
    "metrics": {
        "enabled": true
    },
    "traces": {
        "enabled": true,
        "nestFileSpans": true,
        "otelLibName": "fstracing",
        "otelServiceName": "fs-workload-tracing",
        "otelHostName": "local-host",
        "otelEndpt": "localhost:4317"
    },
    "faultyIO": {
        "enabled": false
    },
    "throttleIO": {
        "enabled": false
    },
    "fakeIO": {
        "enabled": false
    }
}
```

Workload Tracing

Feedback from mentors

- Nesting spans

Adding robustness to FaultyFs

More Testing

- Previously simulated faults in read, write, fsync, and get and set attributes
- Adding faults in certain functions can be unpredictable or breaks FS
- Some bugs with truncation indexing out of buffer
- Alternatively, more functions are well suited for faults that didn't have them before

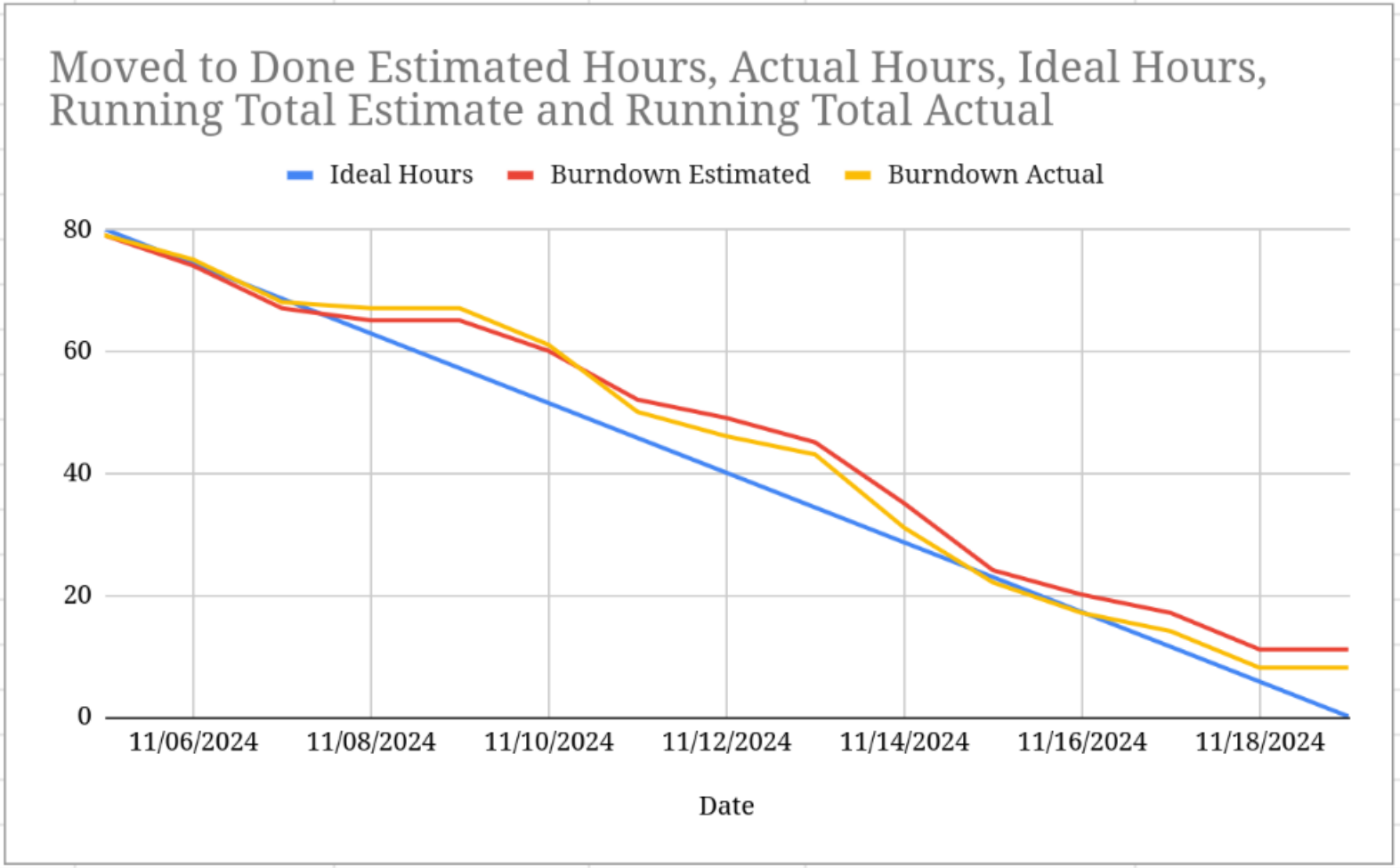
More features

- Error log also gets file descriptor so more info available
- Implemented flags that can set the likelihood of faults and optionally increase when seeding happens

Review

1	Add prometheus-cpp and nlohmann_json dependency to container image #107	Nov 5, 2024
2	Add postgresql/pgbench as a test program #108	Nov 6, 2024
3	Add fio as a test program #109	Nov 6, 2024
4	Generate FUSE config from volume attributes #110	Nov 7, 2024
5	Address dependabot vulnerability report #113	Nov 7, 2024
6	Post processing script for resolving file path from inode number #81	Nov 8, 2024
7	Workload Tracing: Nest spans #120	Nov 10, 2024
8	Make filesystem attributes configurable #83	Nov 11, 2024
9	Create parent spans for multiple operations on the same file #80	Nov 11, 2024
10	Implement metric collection layer for FUSE fileystem #117	Nov 12, 2024
11	Make the FUSE layers individually configurable #118	Nov 14, 2024
12	Prepare for presentation #119	Nov 18, 2024
13	Combine all fuse filesystems into a single binary #82	
14	Grafana dashboard for the metrics #84	
15	Test the installation in alternative kubernetes environments (like EKS) #87	
16	Push mapping from inode number to file name to external database #116	
17	OTLP in docker container for Faulty #121	

Burndown


Moved to Done Estimated Hours, Actual Hours, Ideal Hours, Running Total Estimate and Running Total Actual
Ideal Hours
Burndown Estimated
Burndown Actual
80
60
40
20
0
11/06/2024
11/08/2024
11/10/2024
11/12/2024
11/14/2024
11/16/2024
11/18/2024
Date