Distributed tracing for jobs: jobsub and beyond
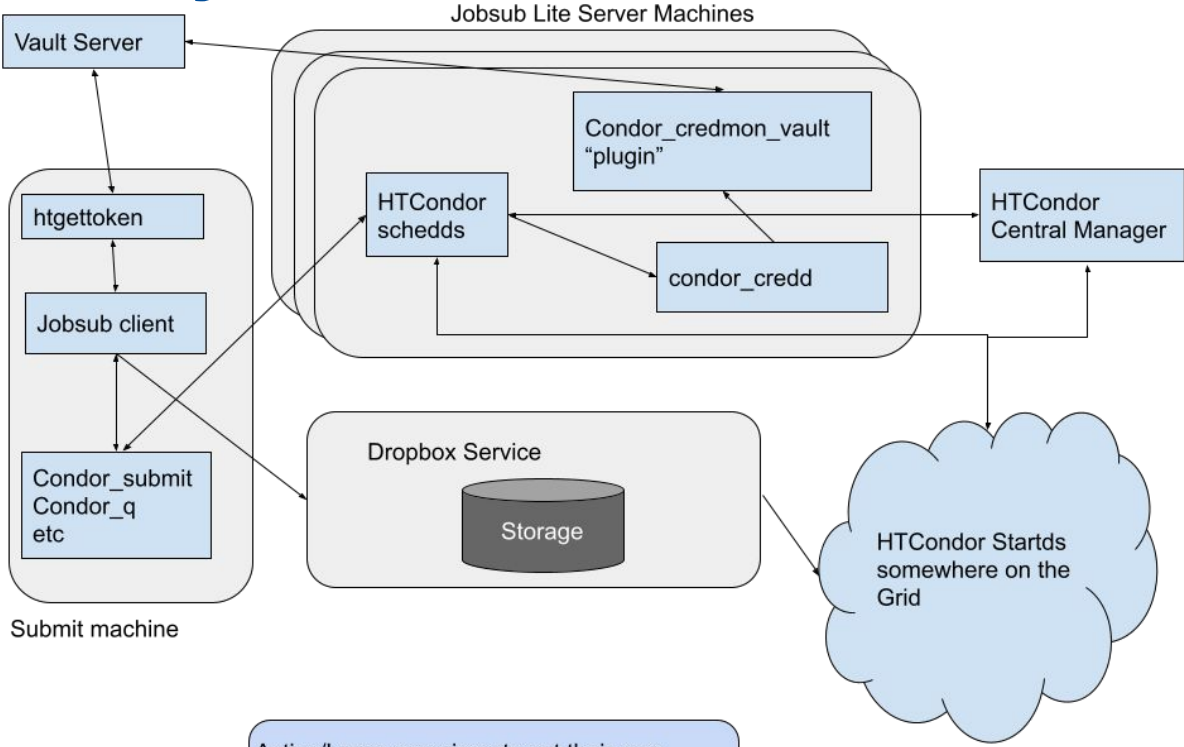
Shreyas Bhat and Marc Mengel
on behalf of the Jobsub Team
Scientific Computing Monitoring Workshop
July 24, 2024

Intro to jobsub_lite

- Software for job submission and monitoring
- Tried to keep the most-used pieces of jobsub_client, strip out unnecessary parts
- Client-only, installed on experiment interactive nodes
- For more info:

  https://fifewiki.fnal.gov/wiki/Getting_started_with_jobsub_lite

S. Bhat and M. Mengel, Distributed tracing for jobs: jobsub and beyond,Scientific Computing Monitoring Workshop, July 24, 2024

Fermilab

jobsub_lite Layout

Jobsub Lite Server Machines

Vault Server

htgettoken

Jobsub client

Condor_submit
Condor_q
etc

Submit machine

Condor_credmon_vault
"plugin"

HTCondor
schedds

condor_credd

HTCondor
Central Manager

Dropbox Service

Storage

HTCondor Startds
somewhere on the
Grid

Active/Large experiments get their own
schedd and can't be brought down by other
experiments errant users.

Original Image Credit: J. Boyd

S. Bhat and M. Mengel, Distributed tracing for jobs: jobsub and beyond,Scientific Computing Monitoring Workshop, July 24, 2024

🎄 Fermilab

What happens

- jobsub_lite takes user command, converts it to HTCondor submission file (Job Definition File), and uses HTCondor python bindings to submit the job to a remote schedd
- Following sub-operations handled by jobsub CLI:
  - Authentication
  - File upload
  - Job submission
  - Job management

S. Bhat and M. Mengel, Distributed tracing for jobs: jobsub and beyond,Scientific Computing Monitoring Workshop, July 24, 2024

🔶 **Fermilab**

Then vs. now

- In previous client/server setup, had copious server logs
- Now, no logs on client to help troubleshoot
- Debug mode can only take you so far
- If submission goes wrong - where in submission did error occur?

🎇 Fermilab

Enter tracing!

- In jobsub_lite 1.4, we instrumented the jobsub_lite code to enable tracing of jobsub_lite functions
- While clear error messages can be helpful, for troubleshooting, much easier to see a picture:
- https://landscape.fnal.gov/jaeger/trace/58702552abacbdd980d89ff34ac7c63c
  - Can see job submission processes
  - We can see that job actually submitted, but all sandbox files didn't get transferred properly (canned example where we Ctrl-C-ed)

Cool!

S. Bhat and M. Mengel, Distributed tracing for jobs: jobsub and beyond,Scientific Computing Monitoring Workshop, July 24, 2024

Fermilab

What if something fails INSIDE our job?

- Might get trace that looks like this:
  - https://landscape.fnal.gov/jaeger/trace/a026cabdc37613c3c0f7933aa5c53a4d
- But job sandbox shows:

```
Fri Jul 19 10:27:07 EDT 2024 81422934.0@jobsub02.fnal.gov BEGIN EXECUTION false
Fri Jul 19 10:27:44 EDT 2024 false COMPLETED with exit status 1
```

- So what went wrong? Can we look inside the jobs some other way than doing

```
ifdh cp -D exelogs /pnfs/myexpt/scratch/users/mylogs
```

S. Bhat and M. Mengel, Distributed tracing for jobs: jobsub and beyond,Scientific Computing Monitoring Workshop, July 24, 2024

Fermilab

Tracing INSIDE jobs?!?

- In all jobsub_lite jobs, we set an environment variable `TRACEPARENT` that holds a special Trace ID and Span ID for the jobsub_submit command that launched the job
- If we can get our job executable to use that `TRACEPARENT`, can we continue the previous traces INTO the job? Yes!!
- export `OTEL_EXPORTER_OTLP_TRACES_ENDPOINT` to "https://landscape.fnal.gov/otlp/v1/traces" - tell any underlying instrumentation code where to send traces
- "**Distributed tracing** lets you observe requests as they propagate through complex, distributed systems" - (Source: https://opentelemetry.io/docs/concepts/observability-primer/#understanding-distributed-tracing)

S. Bhat and M. Mengel, Distributed tracing for jobs: jobsub and beyond,Scientific Computing Monitoring Workshop, July 24, 2024

🪐 Fermilab

Tracing INSIDE jobs!

- To demonstrate, here is test job that works…until it doesn't. What happened??
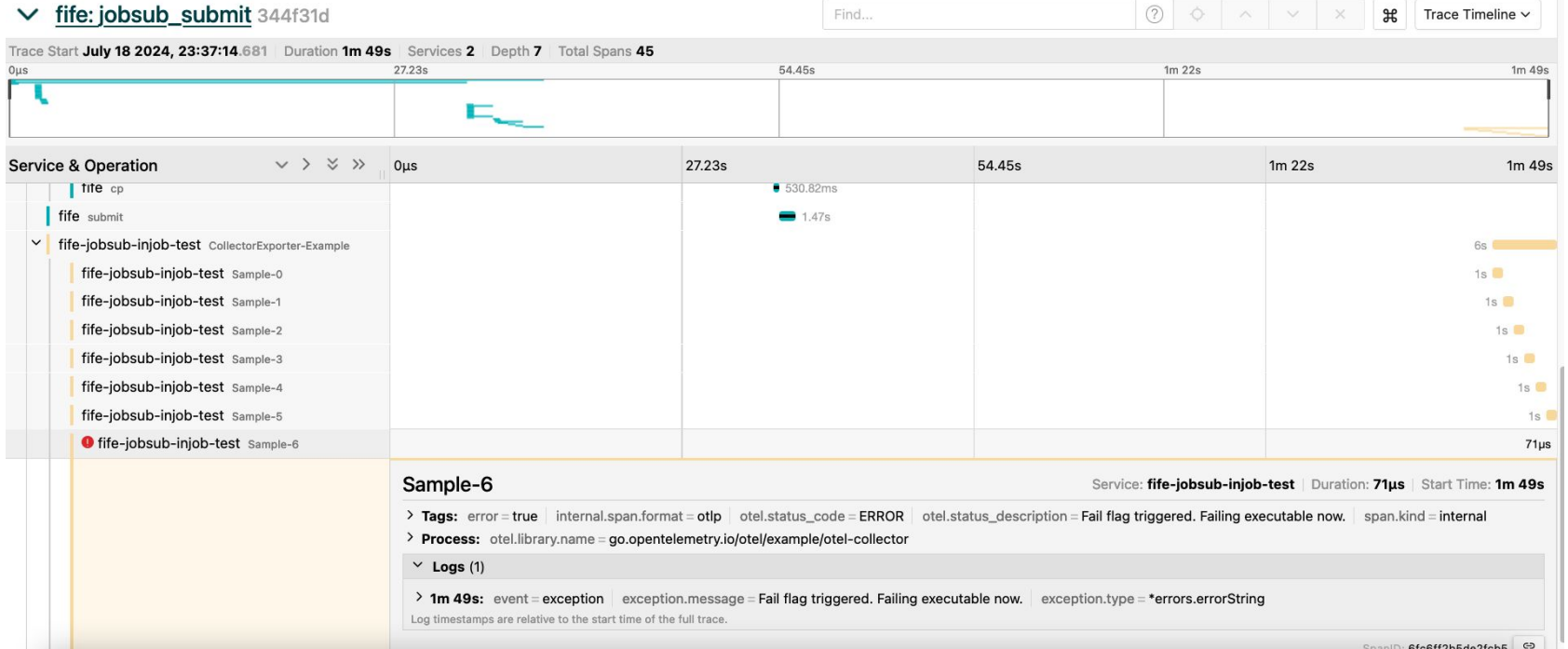
Job log stdout:

```
Looking for file /cvmfs/fifeuser3.opensciencegrid.org/sw/fermilab/e076d4781fc698bad292712e7bfde07f67d10b05ef4b09a703275e83e8a3fe33/
tracingJob_linux on RCDS.  Try 1 of 40
Found file /cvmfs/fifeuser3.opensciencegrid.org/sw/fermilab/e076d4781fc698bad292712e7bfde07f67d10b05ef4b09a703275e83e8a3fe33/
tracingJob_linux on RCDS.  Copying in.
Fri Jul 19 04:22:22 UTC 2024 81419552.0@jobsub02.fnal.gov BEGIN EXECUTION test_trace.sh -f
JobID: 81419552.0@jobsub02.fnal.gov
Running tracing executable
Error running tracingJob.exe.  See TRACEPARENT 00-6cfba14149a1af4306bf7bdcad4e0258-64285d7973a41400-01
Fri Jul 19 04:22:28 UTC 2024 test_trace.sh COMPLETED with exit status 2
```

Job log stderr:

```
Fri Jul 19 04:22:22 UTC 2024 81419552.0@jobsub02.fnal.gov BEGIN EXECUTION test_trace.sh -f
2024/07/19 04:22:22 Waiting for connection...
2024/07/19 04:22:22 TraceID given: 6cfba14149a1af4306bf7bdcad4e0258-64285d7973a41400-01
2024/07/19 04:22:22 Doing really hard work (1 / 10)
2024/07/19 04:22:23 Doing really hard work (2 / 10)
2024/07/19 04:22:24 Doing really hard work (3 / 10)
2024/07/19 04:22:25 Doing really hard work (4 / 10)
2024/07/19 04:22:26 Doing really hard work (5 / 10)
2024/07/19 04:22:27 Doing really hard work (6 / 10)
2024/07/19 04:22:28 Fail flag triggered.  Failing executable now.
cvmfs info:
sed: -e expression #1, char 31: unterminated address regex
Fri Jul 19 04:22:28 UTC 2024 test_trace.sh COMPLETED with exit status 2
```

S. Bhat and M. Mengel, Distributed tracing for jobs: jobsub and beyond,Scientific Computing Monitoring Workshop, July 24, 2024

🟦 Fermilab

Tracing INSIDE jobs! (2)

https://landscape.fnal.gov/jaeger/trace/344f31d50a7c0c07f85f7bdd97f7
2f92

S. Bhat and M. Mengel, Distributed tracing for jobs: jobsub and beyond,Scientific Computing Monitoring Workshop, July 24, 2024

Tracing INSIDE jobs! (3)

- Allows us to see:
  - Exactly where job failed
  - Any logs associated with that failure specifically
  - Duration of failed function (i.e. was there a bottleneck/timeout that caused the failure?)
- Users/software developers can decide what is important to track/log
  - Limit log noise
- Can see entire flow of job workflow from submission through to job execution - very powerful

Fermilab

Future

- This demonstration was done using a shell script that runs a Go executable (with the former setting ENVs and tracing instrumentation inside the latter)
- Jobsub team will work on recipe to do this within pure shell scripts, subject to effort availability
- Outside of that, we're happy to share as much knowledge as we've got with anyone who's interested

S. Bhat and M. Mengel, Distributed tracing for jobs: jobsub and beyond,Scientific Computing Monitoring Workshop, July 24, 2024

🎝 Fermilab

Resources

- Github repo: https://github.com/fermitools/jobsub_lite
- Roadmap: https://cdcvs.fnal.gov/redmine/projects/jobsub/wiki/Jobsub_lite_Tasks
- Getting Started: https://fifewiki.fnal.gov/wiki/Getting_started_with_jobsub_lite
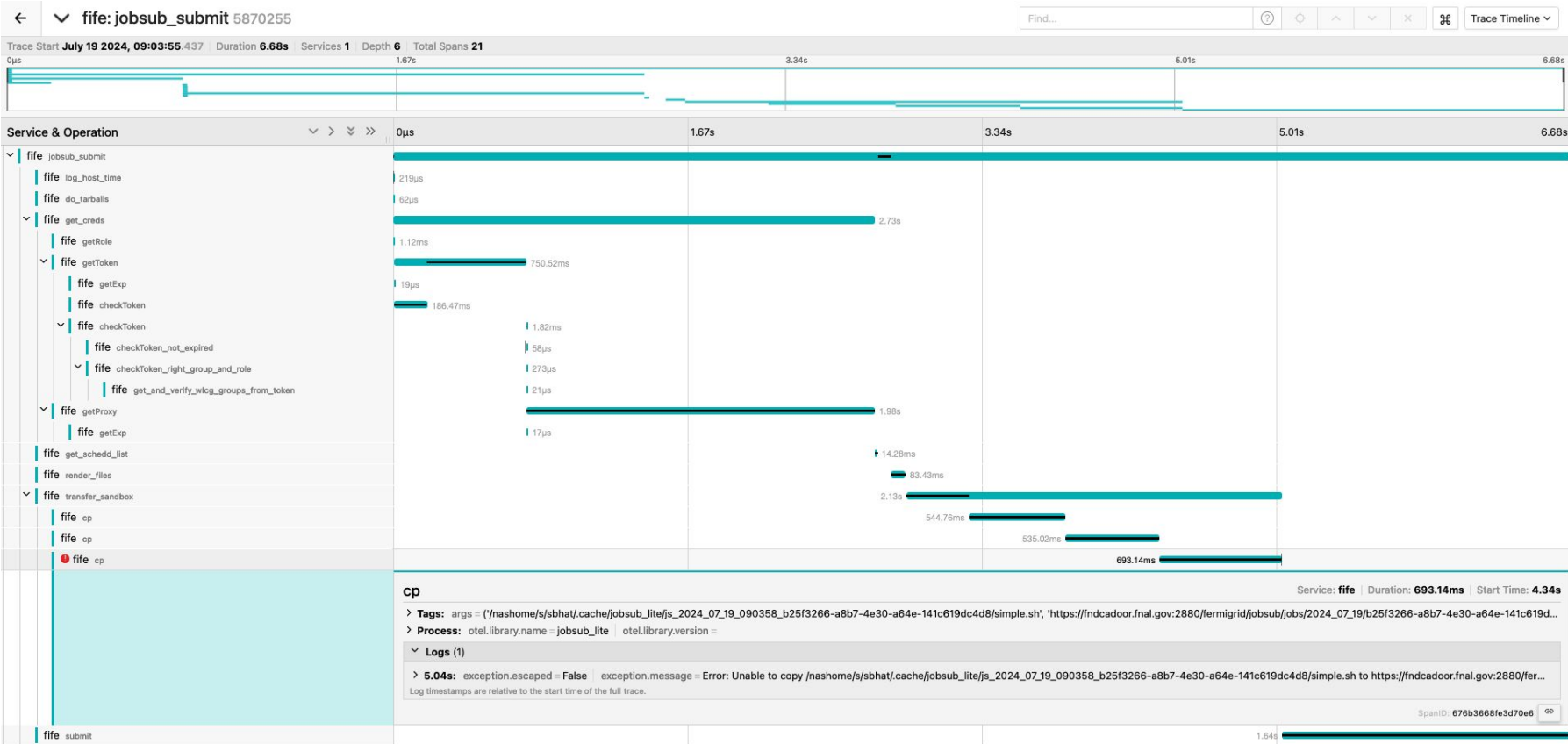- OpenTelemetry Tracing: https://opentelemetry.io/docs/

Fermilab

Thank you!

The jobsub project team:

Shreyas Bhat, Vito Di Benedetto, Lisa Goodenough, Marc Mengel, Nick Peregonow, Kevin Retzke

ᐧᐧ Fermilab

Extra Slides

S. Bhat and M. Mengel, Distributed tracing for jobs: jobsub and beyond,Scientific Computing Monitoring Workshop, July 24, 2024

jobsub_submit tracing - failed sandbox file copy

S. Bhat and M. Mengel, Distributed tracing for jobs: jobsub and beyond,Scientific Computing Monitoring Workshop, July 24, 2024

🎗️ Fermilab

jobsub_submit tracing - good submit but bad job

S. Bhat and M. Mengel, Distributed tracing for jobs: jobsub and beyond,Scientific Computing Monitoring Workshop, July 24, 2024

🟦 Fermilab