ForeWork: Implementation and Validation (M2/M3)

—

Andrea Barberio

What is ForeWork?

A distributed framework for the analysis of forensic artifacts.
Aimed at getting what matters as soon as possible.

Main goals

- Be parallel

- Be parallel
- Be distributed

- Be parallel
- Be distributed
- Prioritize on what matters

- Be parallel
- Be distributed
- Prioritize on what matters
- Be interactive

Methodology

Be parallel

Stream processing is the key

- IPyParallel, load balanced
- Asynchronous scheduling
- Granular, independent, isolated tasks

Be distributed

Ten is better than one

- IPyParallel, multiple hosts
- Efficient serialization with ZeroMQ
- One (or more) scheduler, many workers

Prioritize on what matters

Sometimes a cigar is just a cigar

- Every investigation has a configuration
- The scheduler knows what matters
- The tasks know how deep to dive

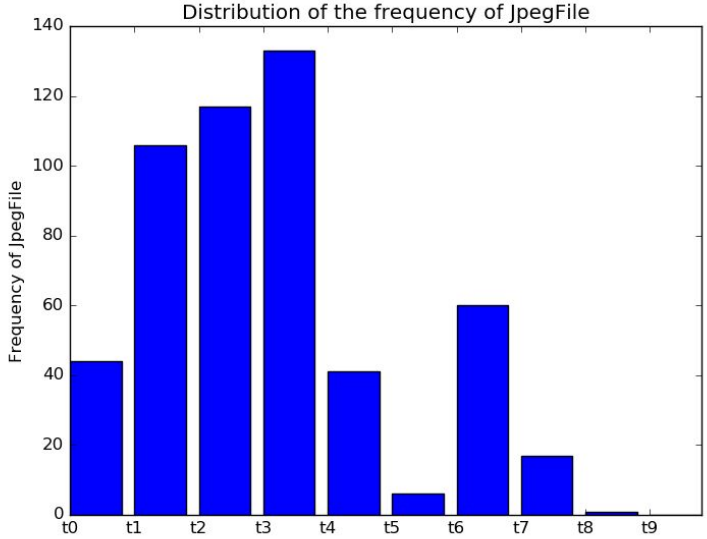
Be interactive

Because robots can fail

- IPython interface
- Every task can be executed individually
- Can do deep dive into each object

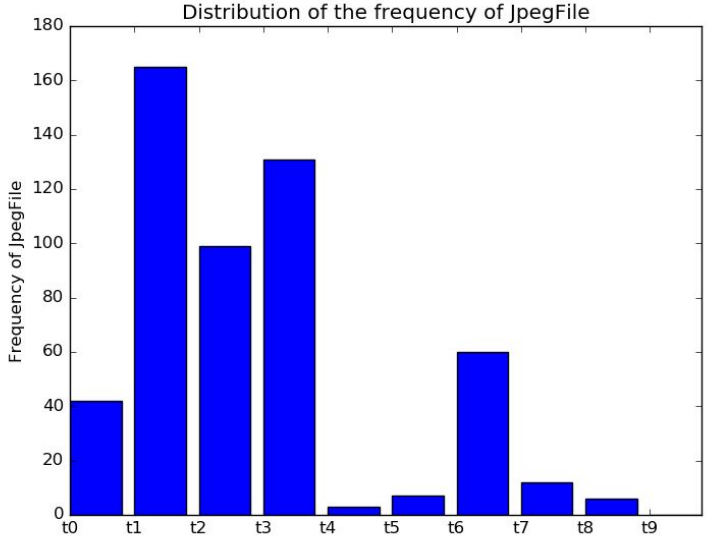
Some numbers

Frequency



- X axis: time
- Y axis: frequency of JPEGs per task aggregation
- With no prioritization

Frequency



- X axis: time
- Y axis: frequency of JPEGs per task aggregation
- With prioritization enabled

The good

- Prioritization shows higher frequency in the early phases
- Streaming lets us access results early

The bad

- Network I/O is an issue
(but MattockFS may change the things)

But above all

IT IS A PROTOTYPE

Future work

Going from prototype to production

- Make it robust
- Support more file types
- Support carving
- Checking for known checksums
- Stegoanalysis
- Entropy analysis
- import/export in DFXML

Demo time

Playing with the ForeWork shell,
showing the execution charts

Conclusion

Forework is a working prototype.

Stream processing lets us triage and prioritize.

It is easy to extend and to use for who has a Python programming background.

It is open source, publicly available, and written in a modern and widely adopted language, Python 3.

Questions?

Andrea Barberio

3 August 2016