Pyrrha: Navigate easily into your system binaries

How to build a firmware mapper & extend it

Eloïse Brocas

ESE, Paris — 25 mai 2024

Quarkslab

Whoami

Security R&D Engineer @ Quarkslab

- Vulnerability research
- Reverse-engineering
- Tooling

@_cryptocorn_

Automated Analysis Team

Tools

Dynamic Analysis	**QBDI**	dynamic binary instrumentation framework
	Qtracer	dynamic trace generator and analysis
Symbolic Execution	**Triton**	symbolic execution framework
	TritonDSE	DSE and exploration library *(whitebox fuzzing)*
Fuzzing	**PASTIS**	collaborative/distributed fuzzing
	HF/QBDI	Honggfuzz backed by QBDI
Firmware Analysis	**Pandora**	whole firmware analysis engine
	Pyrrha	firmware cartography
	QSig	firmware 1-Day matching engine *(discontinued)*
Diffing	**python-bindiff**	python library wrapping Bindiff
	QBinDiff	Binary Differ based on machine learning algorithm
Static Analysis	**python-binexport**	python API to manipulate Binexport files
	Quokka	IDA plugin and python API to manipulate IDA disassembly
Deobfuscation	**Qsynthesis**	synthesis based deobfuscator *(targeting MBAs)*

Whoami

Security R&D Engineer @ Quarkslab

- Vulnerability research
- Reverse-engineering
- Tooling

@_cryptocorn_

Pass the Salt Organizer

- Lille, July 3-5, 2024
- 2024.pass-the-salt.org

Automated Analysis Team

Tools

Dynamic Analysis	**QBDI**	dynamic binary instrumentation framework
	Qtracer	dynamic trace generator and analysis
Symbolic Execution	**Triton**	symbolic execution framework
	TritonDSE	DSE and exploration library *(whitebox fuzzing)*
Fuzzing	**PASTIS**	collaborative/distributed fuzzing
	HF/QBDI	Honggfuzz backed by QBDI
Firmware Analysis	**Pandora**	whole firmware analysis engine
	Pyrrha	firmware cartography
	QSig	firmware 1-Day matching engine *(discontinued)*
Diffing	**python-bindiff**	python library wrapping Bindiff
	QBinDiff	Binary Differ based on machine learning algorithm
Static Analysis	**python-binexport**	python API to manipulate Binexport files
	Quokka	IDA plugin and python API to manipulate IDA disassembly
Deobfuscation	**Qsynthesis**	synthesis based deobfuscator *(targeting MBAs)*

Definition

Embedded into IoT objects

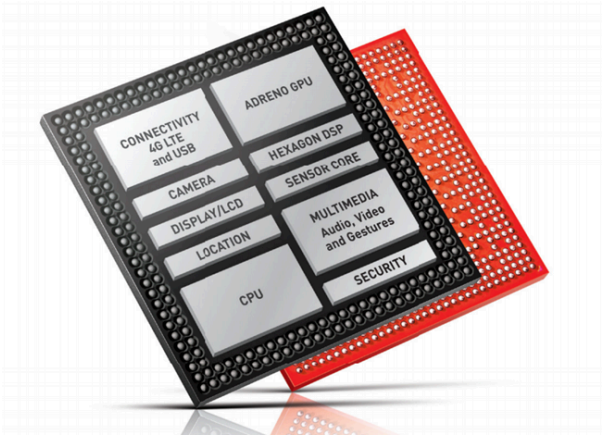
- routers
- smartphones
- automotive

Structured firmwares

- filesystem
- thousands of files

Complete OSes *(Android, Linux, OpenWRT, ...)*

- base operating system
- libraries, SDK, *etc.*
- vendor specific code



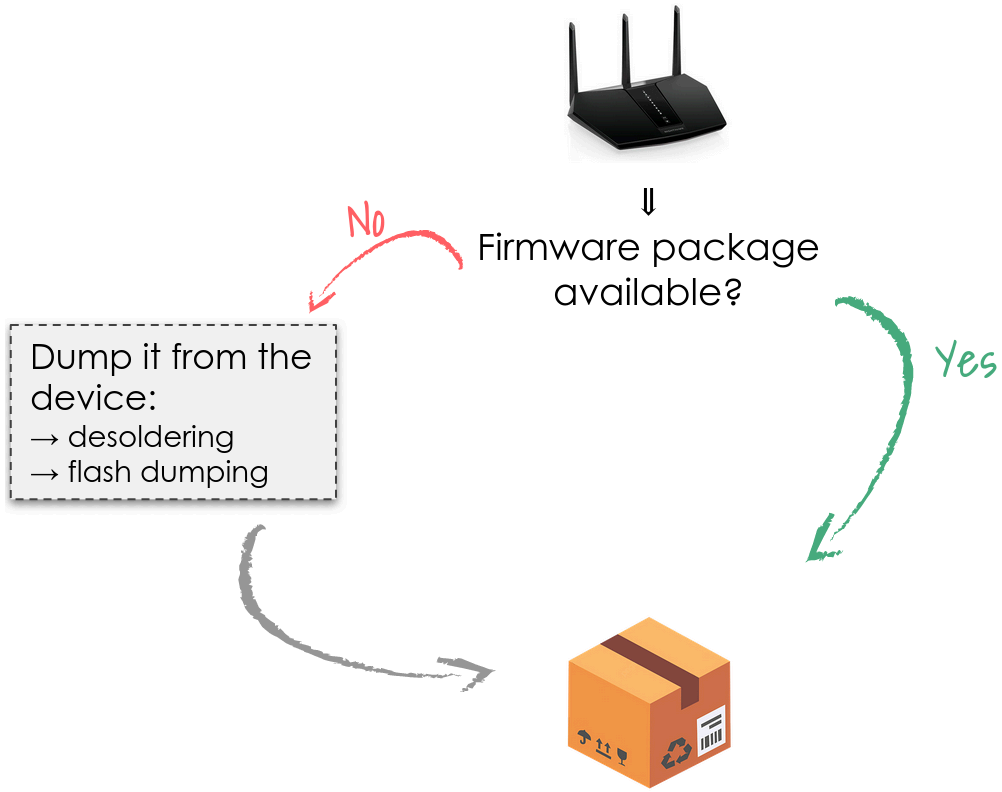
Firmware Analysis use case

- Vulnerability Research
- Artifacts left *(private keys, build files, symbols)*
- Compliance:
  - Checking patches are applied
  - Checking firmware signatures *(trust management)*
- SBOM (Software Bill-of-Materials)

Analysis workflow

1. Acquisition

- could require hardware reverse



Analysis workflow

1. Acquisition

2. Extraction

- Custom *proprietary formats*
- Encrypted firmwares

Binwalk

- developed by ReFirmLabs
- https://github.com/ReFirmLabs/binwalk

Unblob

- developed by OneKey
- https://unblob.org

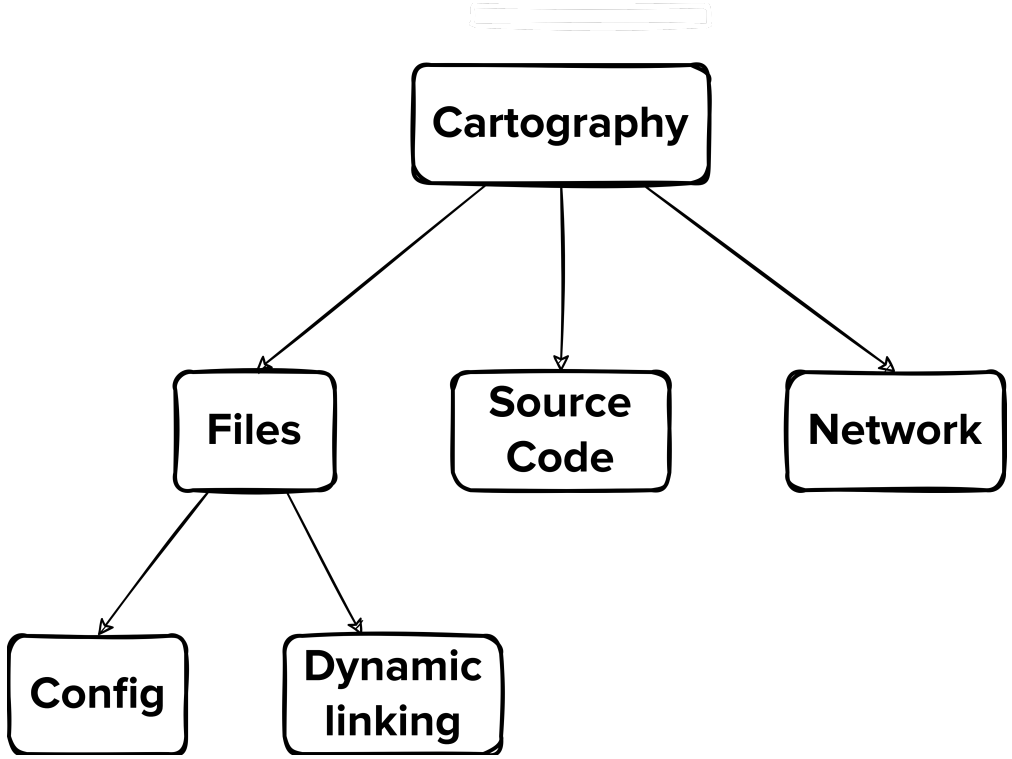
Analysis workflow

1. Acquisition

2. Extraction

3. Cartography

- Make correlations
- Understand interactions between components
- Identify points of interest



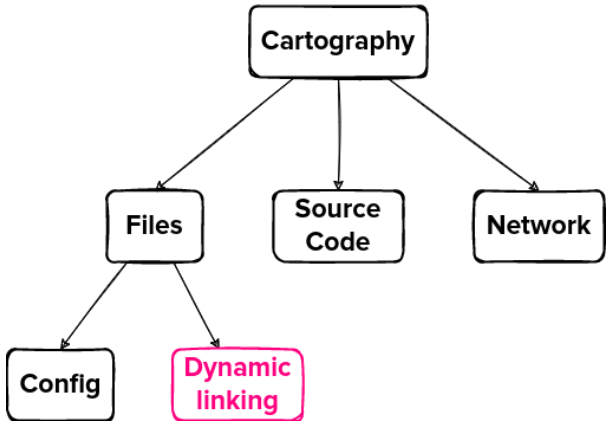
Analysis workflow

1. Acquisition
2. Extraction
3. Cartography
4. In-depth Analysis & Conclusions

Quarklsab blogpost about WAN vulnerabilities in Netgear RAX30 router.

Pyrrha specifications

Cartography tool



How to display result?

- JSON?
- console only?
- GUI?

We also need visualization

What is the family link?

Pyrrha, J.-P. A. Tassaert. Source: Musée du Louvre.

Sisyphus, Titian. Source: Wikimedia.

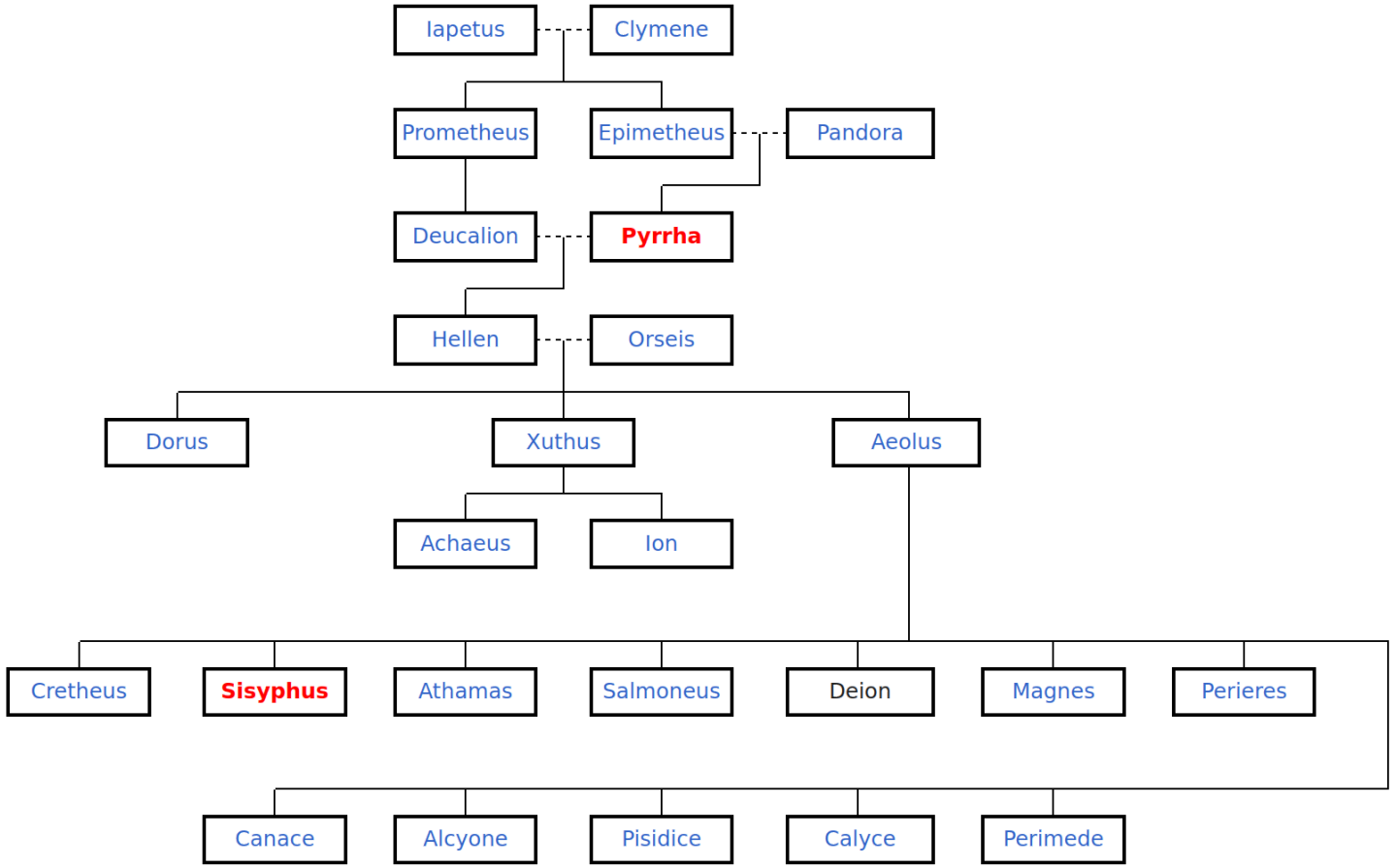
Visualization is the key

«[1.7.2] And Prometheus had a son Deucalion. He reigning in the regions about Phthia, married Pyrrha, the daughter of Epimetheus and Pandora, the first woman fashioned by the gods. And when Zeus would destroy the men of the Bronze Age, Deucalion by the advice of Prometheus constructed a chest, and having stored it with provisions he embarked in it with Pyrrha. But Zeus by pouring heavy rain from heaven flooded the greater part of Greece, so that all men were destroyed, except a few who fled to the high mountains in the neighborhood. It was then that the mountains in Thessaly parted, and that all the world outside the Isthmus and Peloponnese was overwhelmed. But Deucalion, floating in the chest over the sea for nine days and as many nights, drifted to Parnassus, and there, when the rain ceased, he landed and sacrificed to Zeus, the god of Escape. And Zeus sent Hermes to him and allowed him to choose what he would, and he chose to get men. And at the bidding of Zeus he took up stones and threw them over his head, and the stones which Deucalion threw became men, and the stones which Pyrrha threw became women. Hence people were called metaphorically people (laos) from laas, “a stone.” And Deucalion had children by Pyrrha, first Hellen, whose father some say was Zeus, and second Amphictyon, who reigned over Attica after Cranaus; and third a daughter Protogenia, who became the mother of Aethlius by Zeus

[1.7.3] Hellen had Dorus, Xuthus, and Aeolus by a nymph Orseis. Those who were called Greeks he named Hellenes after himself, and divided the country among his sons. Xuthus received Peloponnese and begat Achaeus and Ion by Creusa, daughter of Erechtheus, and from Achaeus and Ion the Achaeans and Ionians derive their names. Dorus received the country over against Peloponnese and called the settlers Dorians after himself. Aeolus reigned over the regions about Thessaly and named the inhabitants Aeolians. He married Enarete, daughter of Deimachus, and begat seven sons, Cretheus, Sisyphus, Athamas, Salmoneus, Deion, Magnes, Perieres, and five daughters, Canace, Alcyone, Pisidice, Calyce, Perimede.»

Apollodorus, 1.7.[2-3].

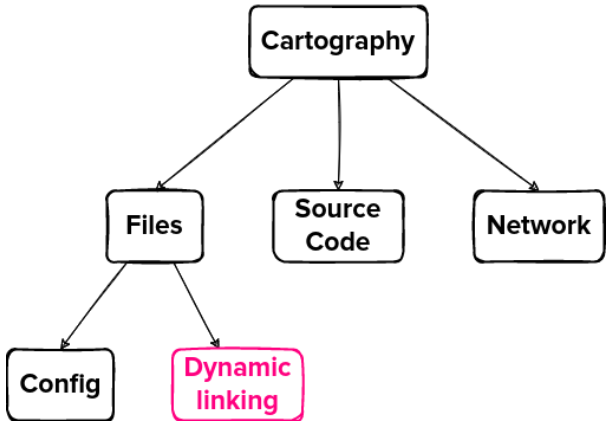
Visualization is the key



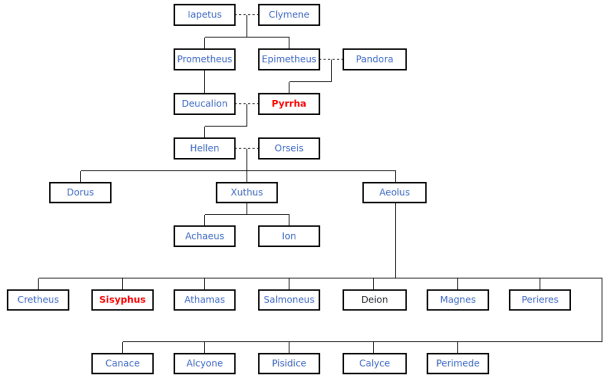
Genealogy of Hellenes. Source: Wikipedia.

Pyrrha specifications

Cartography tool

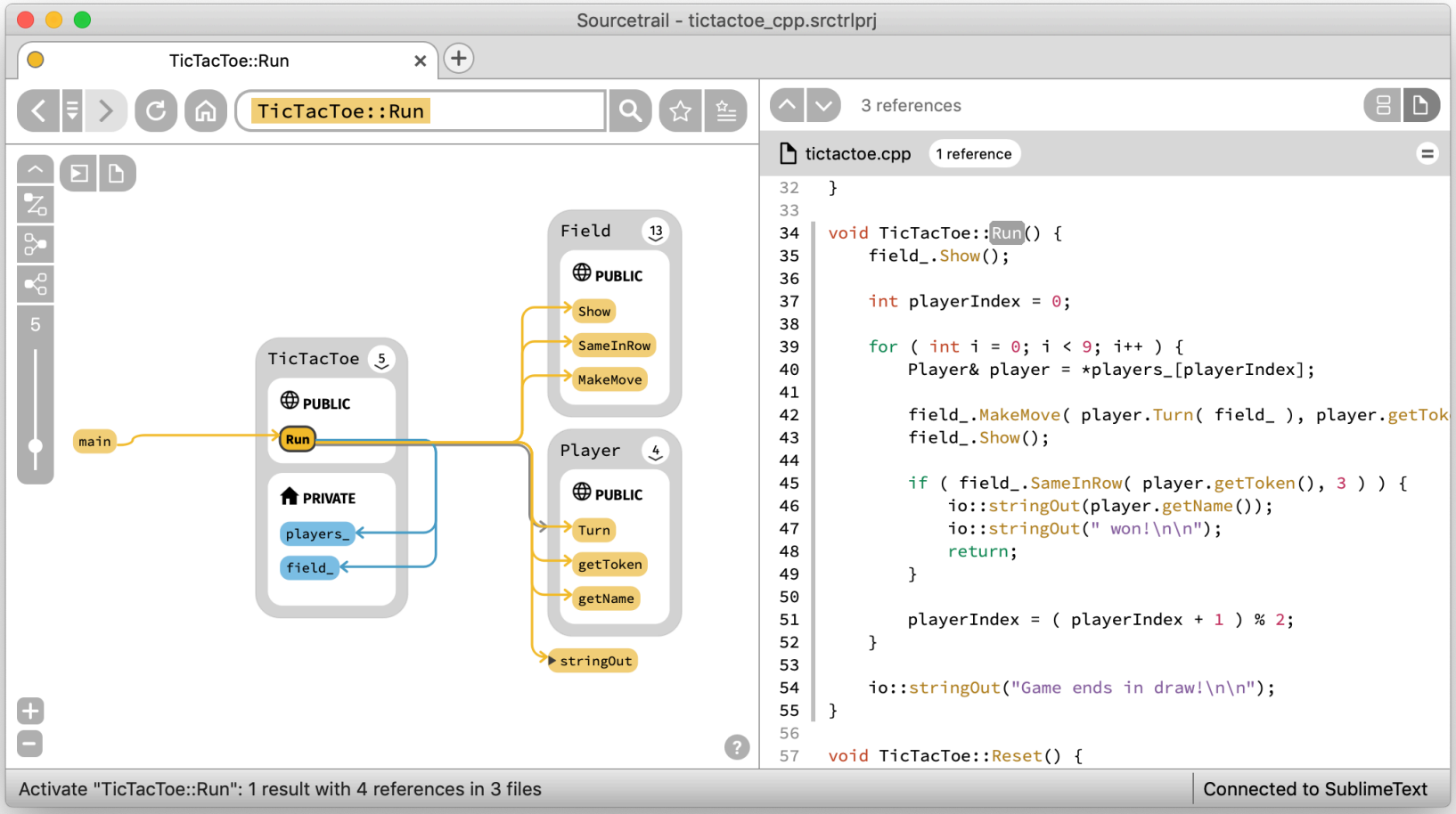


Result visualization



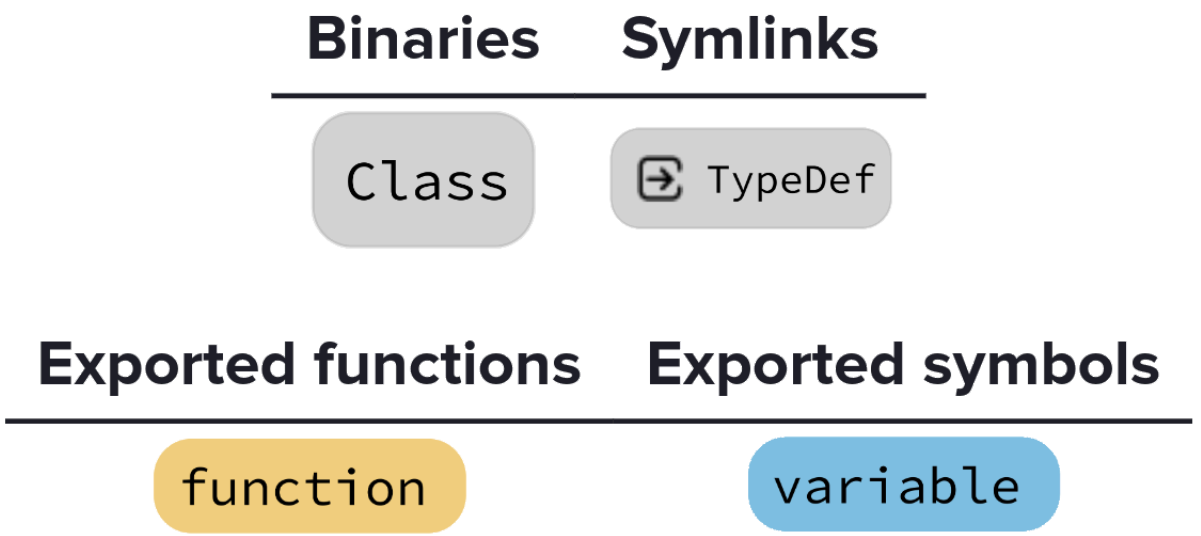
Reuse existing project?

Let's try not to reinvent the wheel!



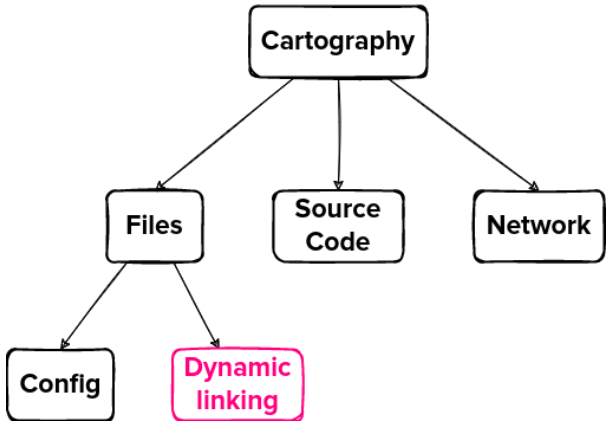
Sourcetrail interface. Source: Sourcetrail README.

Filesystem as a language

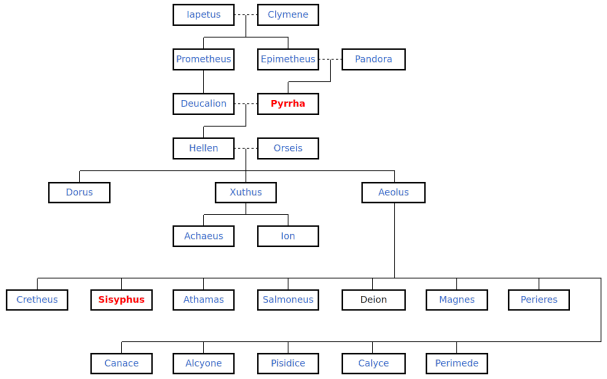


Pyrrha specifications

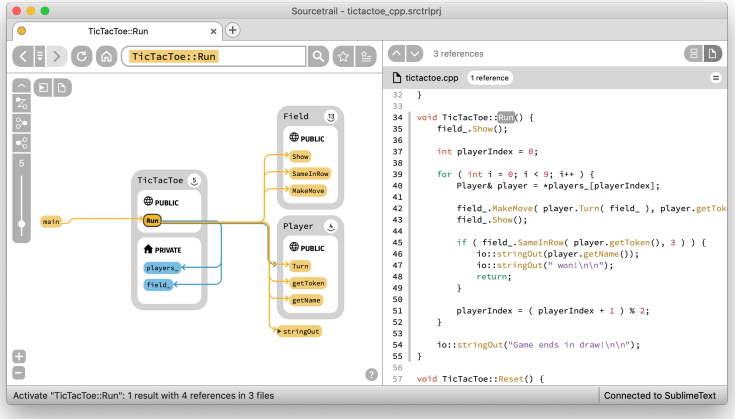
Cartography tool



Result visualization



Extend Sourcetrail

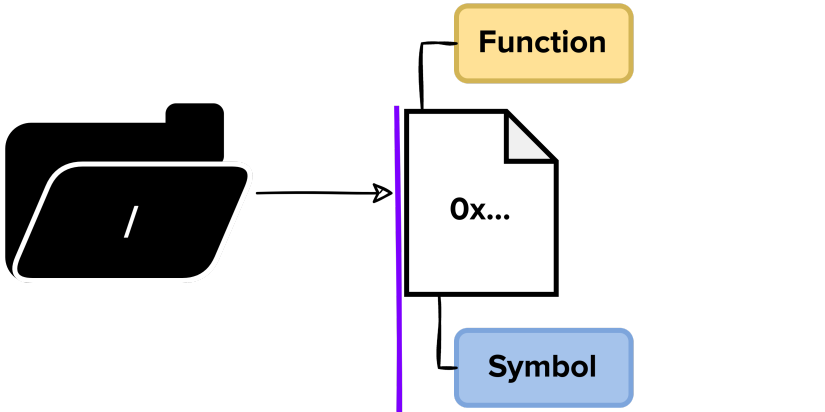


Pyrrha workflow

Filesystem to map

- already extracted
- entry: root directory

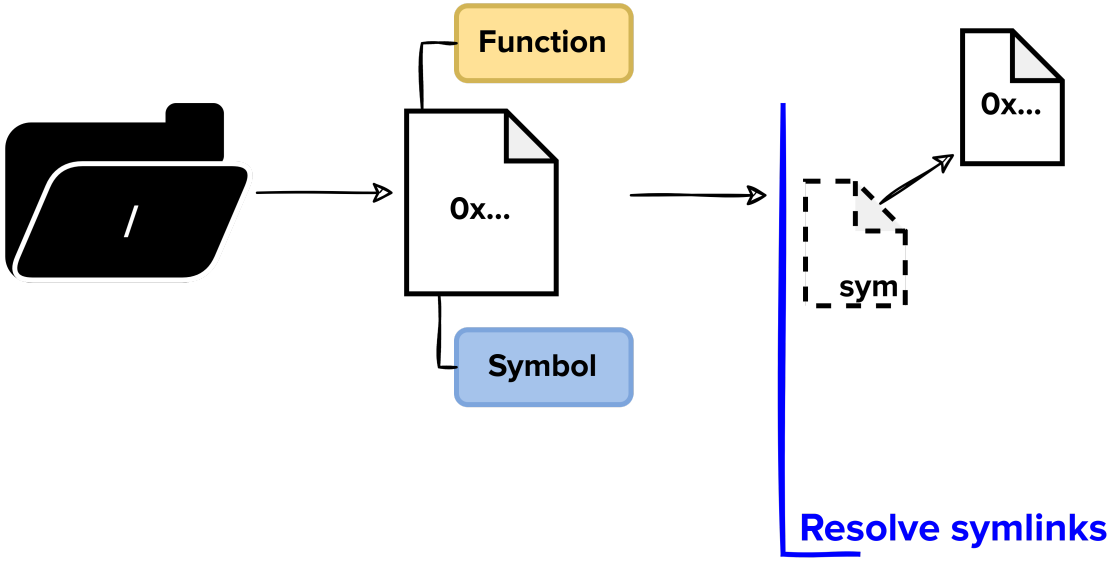
Pyrrha workflow



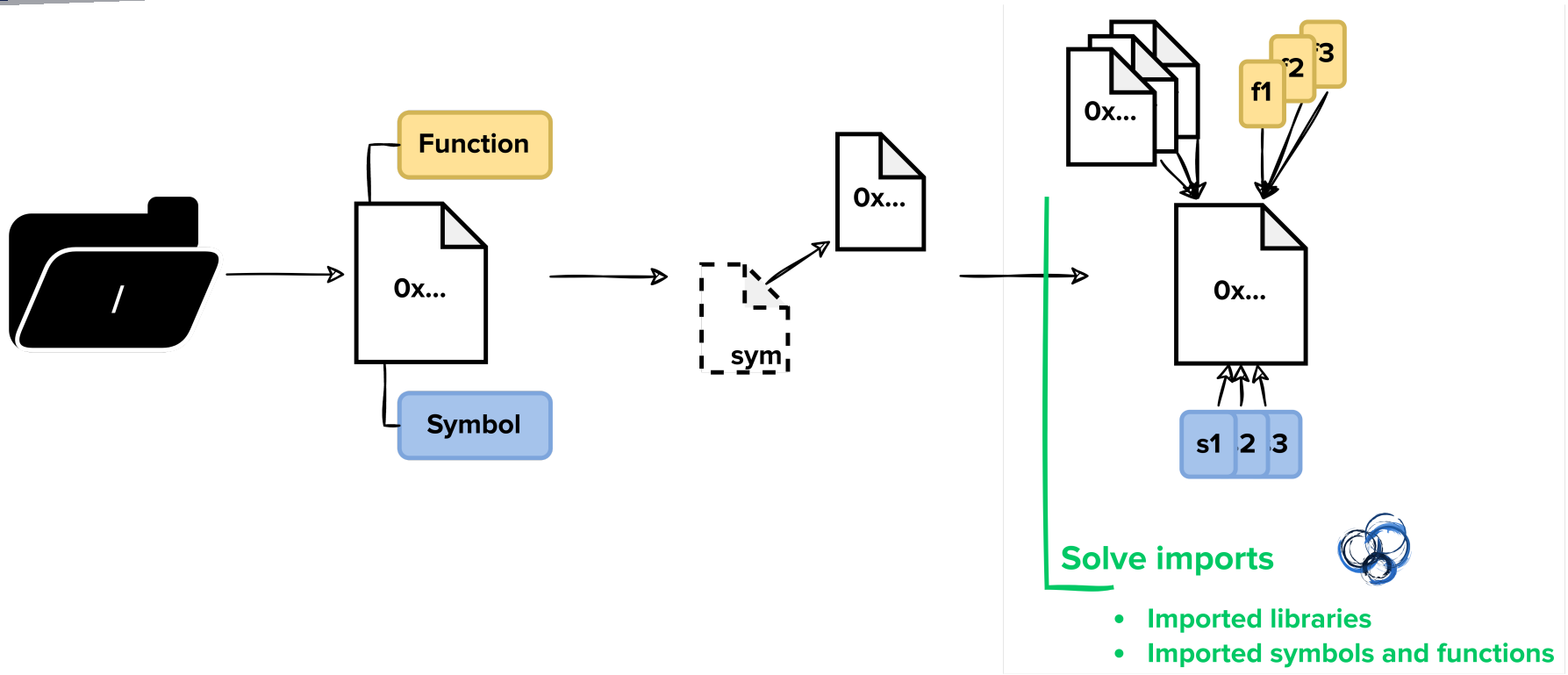
Iterate binaries

- Exported symbols
- Exported functions

Pyrrha workflow



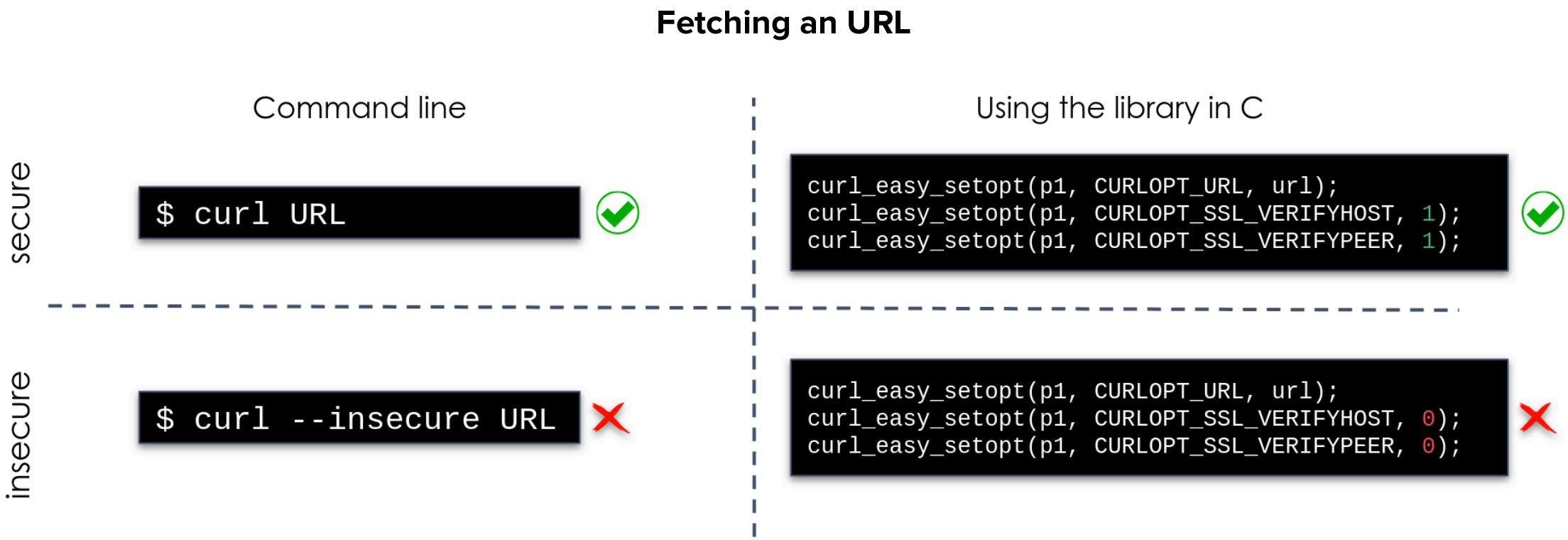
Pyrrha workflow



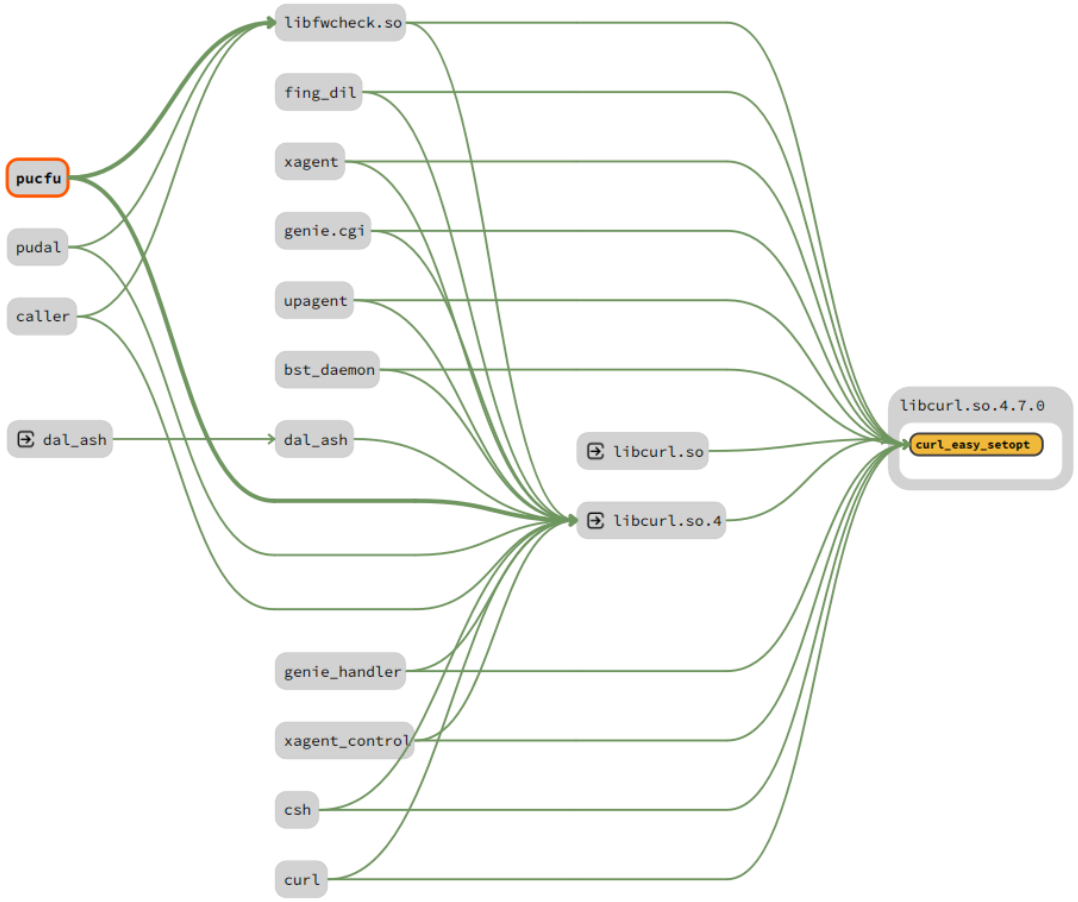
Demo Time

Netgear RAX30

Entrypoint identification: theory



Entrypoint identification: using Pyrrha



Result Export for intercompatibility

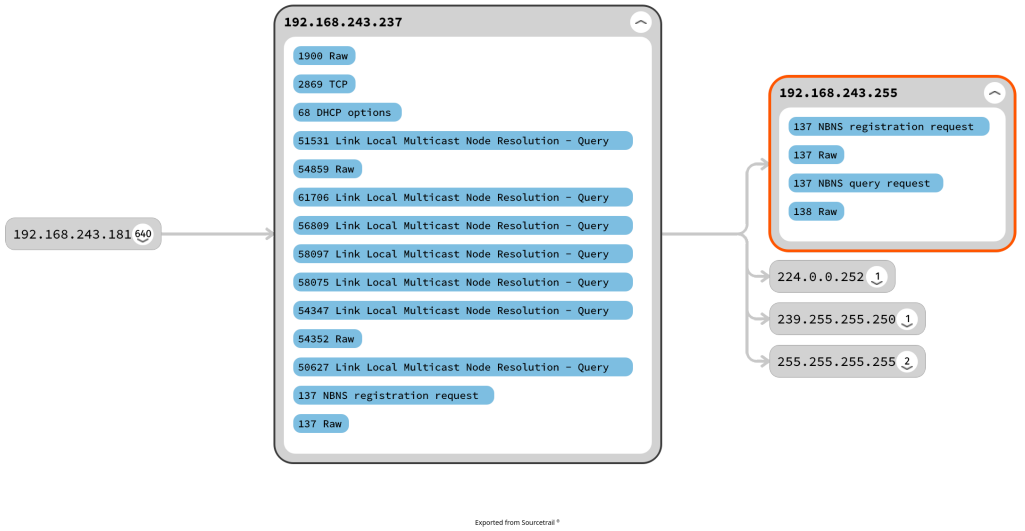
```
Binaries no longer in RAX30-V1.0.9.90_3.json:
  - /lib/libcurl.so.4.6.0
  - /bin/pppoe-relay
  - /bin/websockd

Binaries added in RAX30-V1.0.9.90_3.json:
  - /[KERNEL_VERSION]/kernel/net/netfilter/xt_connlimit.ko
  - /lib/libcurl.so.4.7.0
[...]

Common binaries that have changed:
pudil have changed:
  - symbols removed: {'sprintf'}
  - symbols added: {'sleep', 'snprintf'}
fing_dil have changed:
  - lib removed: {'libcurl.so.4.6.0'}
  - lib added: {'libcurl.so.4.7.0'}
[..]

Total having changed: 108
```

Diffing use case example.

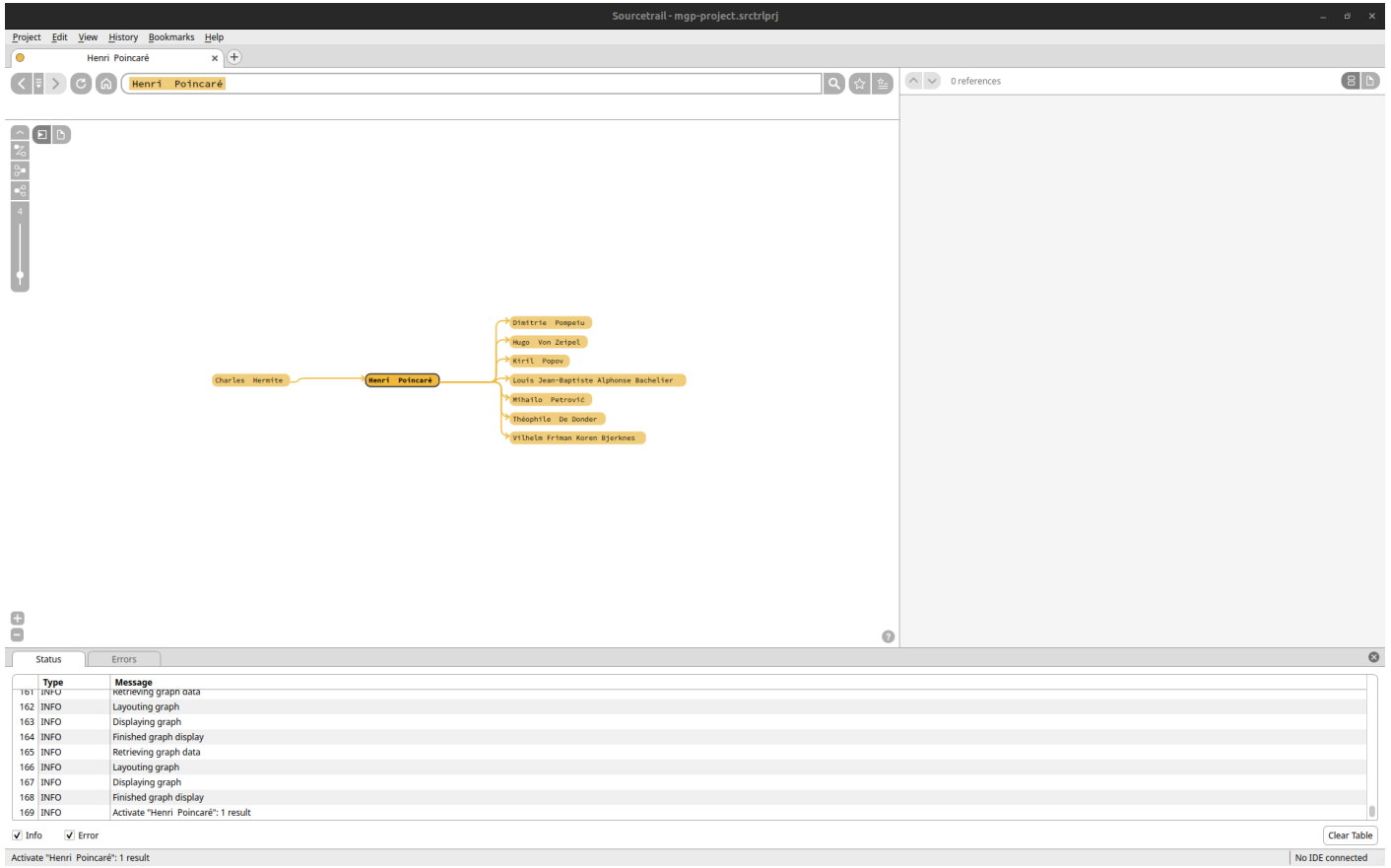
Numbat

- Quarkslab implementation of Sourcetrail API
- Fully Pythonic
  - https://github.com/quarkslab/numbat
  - `pip install numbat`

Visualize any graph-like data



Numbat Example



Data Source: Mathematics Genealogy Project.

Numbat Usage

```python
from numbat import SourcetrailDB

# Create DB
db = SourcetrailDB.open('my_db', clear=True)

# Create a first class containing the method 'main'
my_main = db.record_class(name="MyMainClass")
meth_id = db.record_method(name="main", parent_id=my_main)

# Create a second class with a public field 'first_name'
class_id = db.record_class(name="PersonalInfo")
field_id = db.record_field(name="first_name",
                           parent_id=class_id)

# The method 'main' is using the 'first_name' field
db.record_ref_usage(meth_id, field_id)

# Save modifications and close the DB
db.commit()
db.close()
```

Conclusion

Pyrrha: Firmware mapper (extensible)

- https://github.com/quarkslab/pyrrha
- pyrrha-mapper

Numbat: Map any graph into Sourcetrail

- https://github.com/quarkslab/numbat
- numbat

Automatization

- Speed up some taks to give more time for specialized analyses
- Cannot solve everything

Future Work

- Customized fork of Sourcetrail
- Add more features to Pyrrha

Thank you!

Contact information:

Email: ebrocas@quarkslab.com

Phone: +33 1 58 30 81 51

Website: https://www.quarkslab.com

Quarkslab