Preview of Vulture's upcoming web filtering engine

Pass the SALT 2018.

Security And Libre Talks.

2-4 juillet 2018 - Lille, France.

dDvens

SECURITY FOR THE DIGITAL AGE

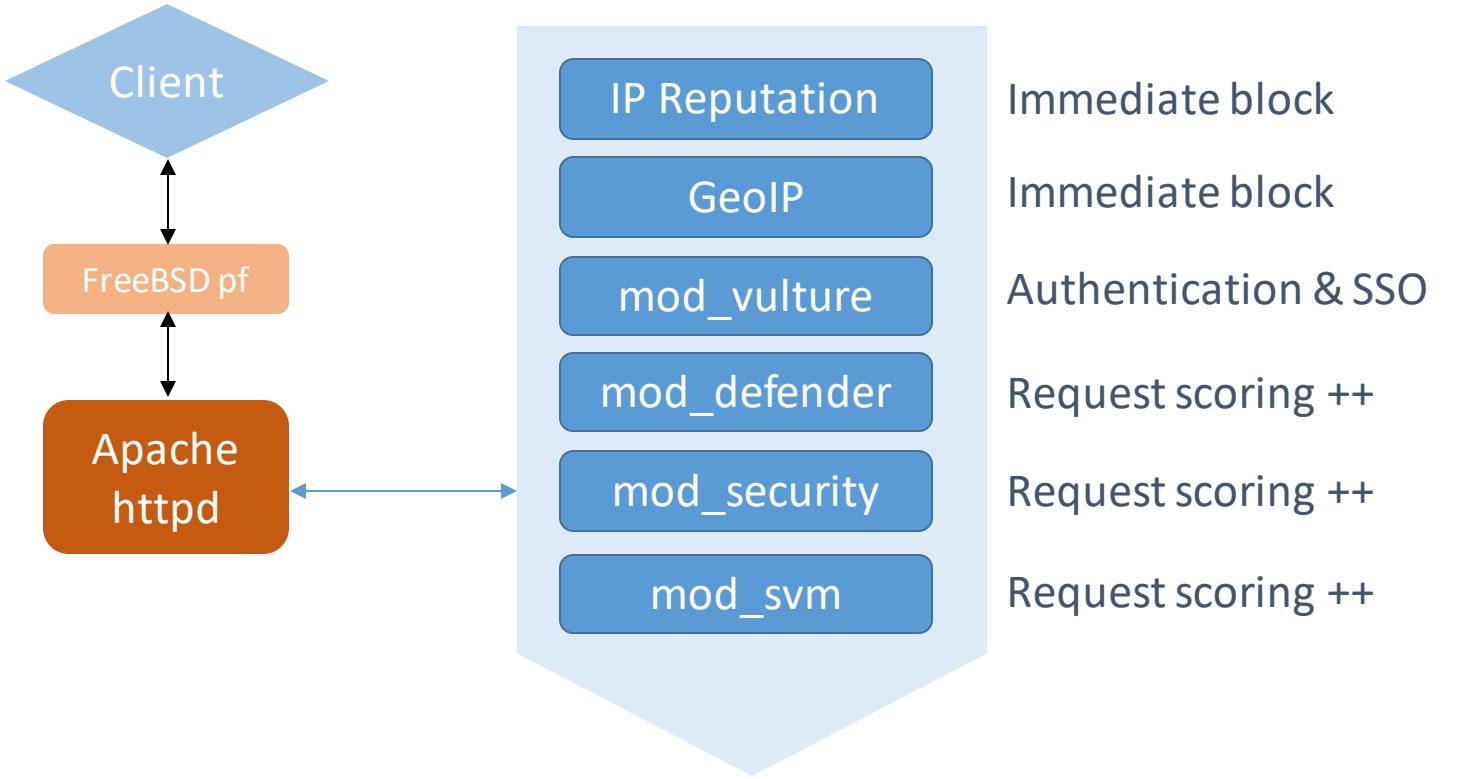
Vulture ?

- A brief history
  - 2003: Linux software (httpd / mod_perl + PHP Web UI)
  - 2016: FreeBSD Cluster (pf, haproxy, httpd + Django Web UI)

- Web SSO: mod_vulture + django portal

- Web application firewall
  - Clustered mod_security, using hiredis [blacklisting]
  - mod_defender, aka "Naxsi for Apache2" [whitelisting]
  - mod_svm [machine learning]

Vulture's current filtering engine



Current limits

- Works well
- No performance issue, but we can do much better
- 3 engines => quite complex
  - Code overlapping
  - Complex UI, httpd knowledge recommended
- Rule-based approach
- Human-based approach
  - Time consuming
  - Need tuning
  - Not mistake proof configuration

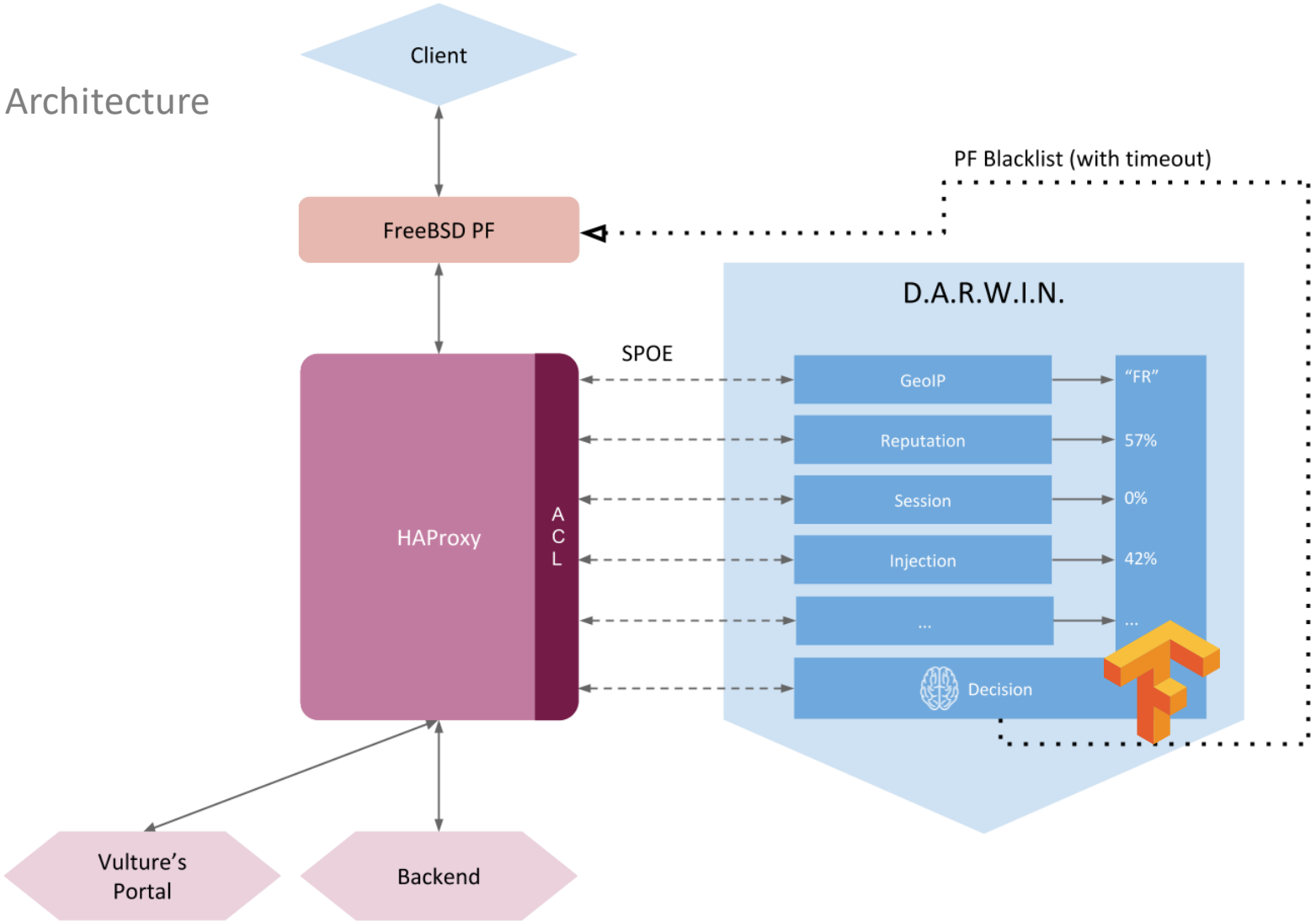
The need for a better, unified engine

- Focused on performance
- High availability required
- Precision and intelligence
- No bullshit
- Simplicity -> For users AND for filter devs

Internal name: « D.A.R.W.I.N. »

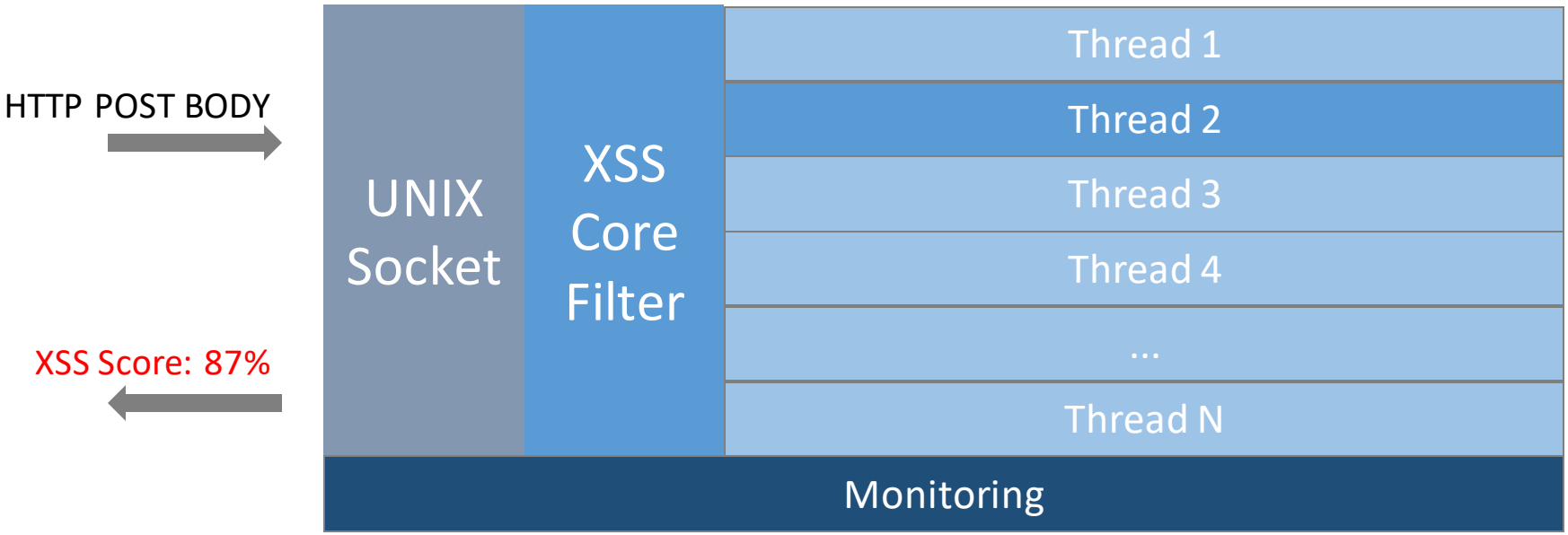
Overview

Architecture



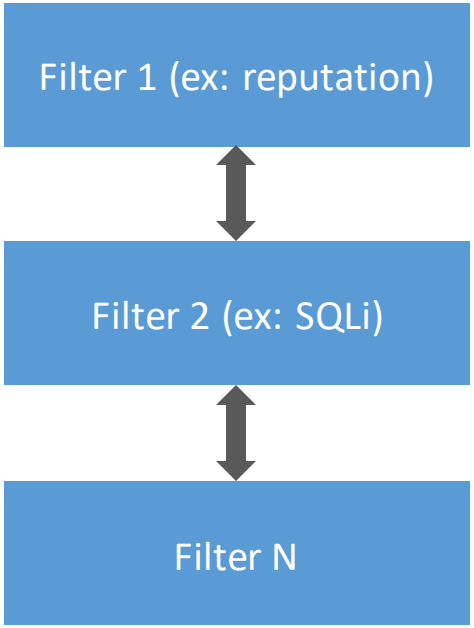
Overview: XSS Filter

Filter



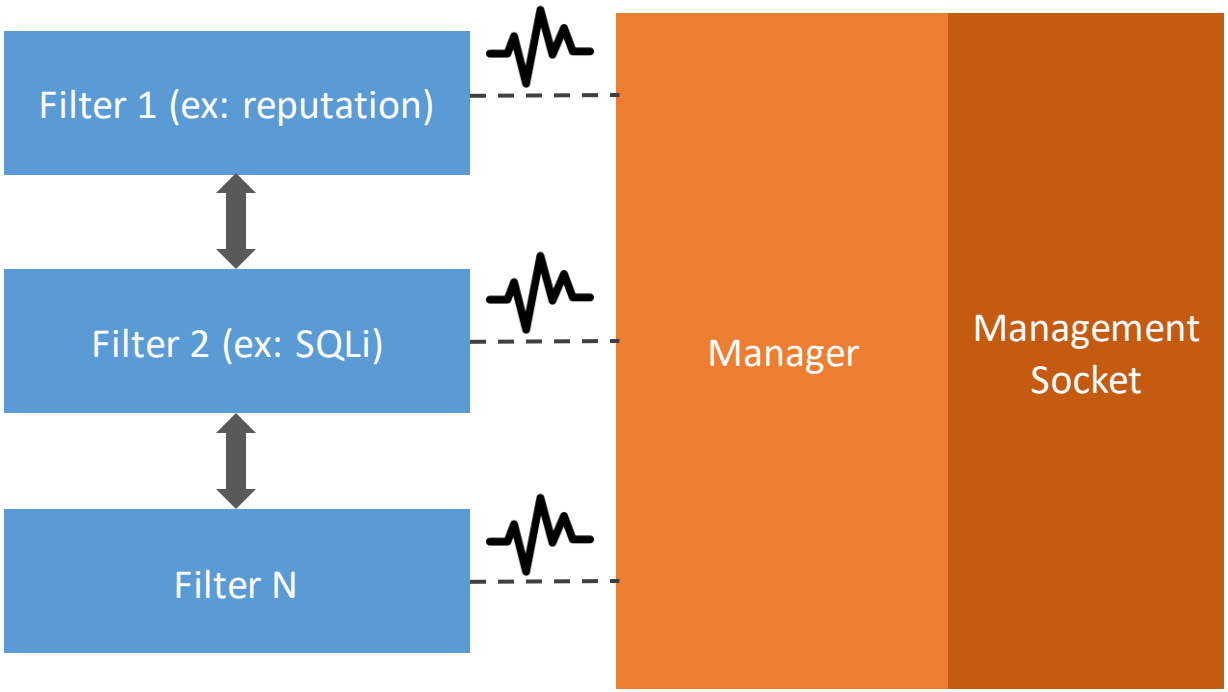
Overview: Filter Workflow

D.A.R.W.I.N.



Overview: Filter Workflow

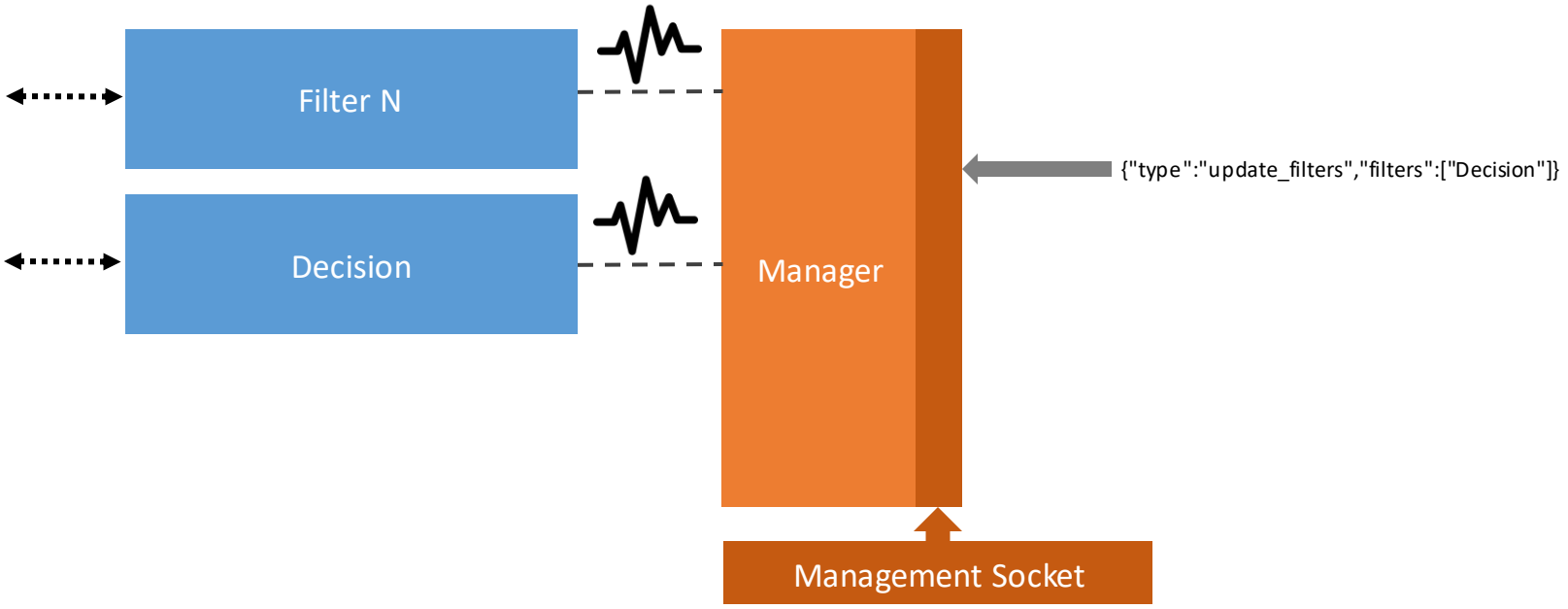
D.A.R.W.I.N.



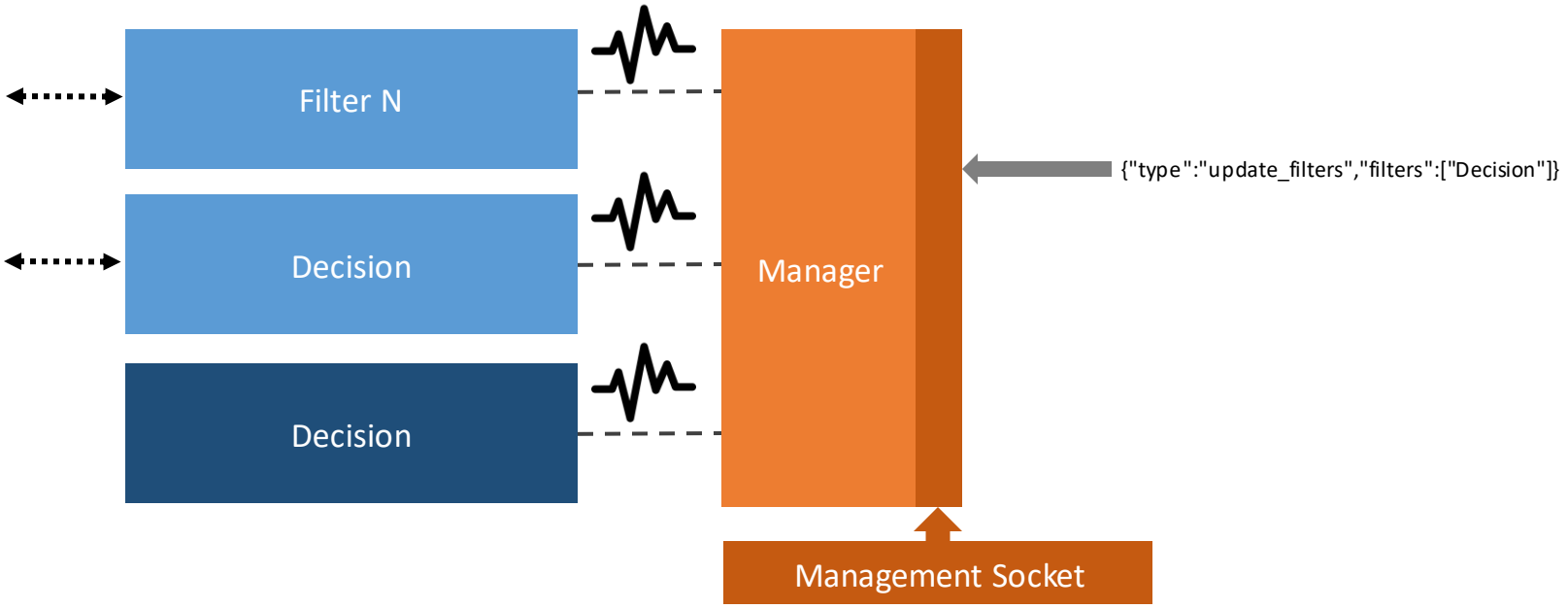
Performance ?

- HAProxy asynchronous events
- C/C++14
- UNIX socket
- Shared in-memory cache (REDIS)
  - Context-sharing between filters among the Vulture cluster
  - Used by Darwin's Neural Networks to track events in time
- Supports GPU acceleration
  - TensorFlow as AI library

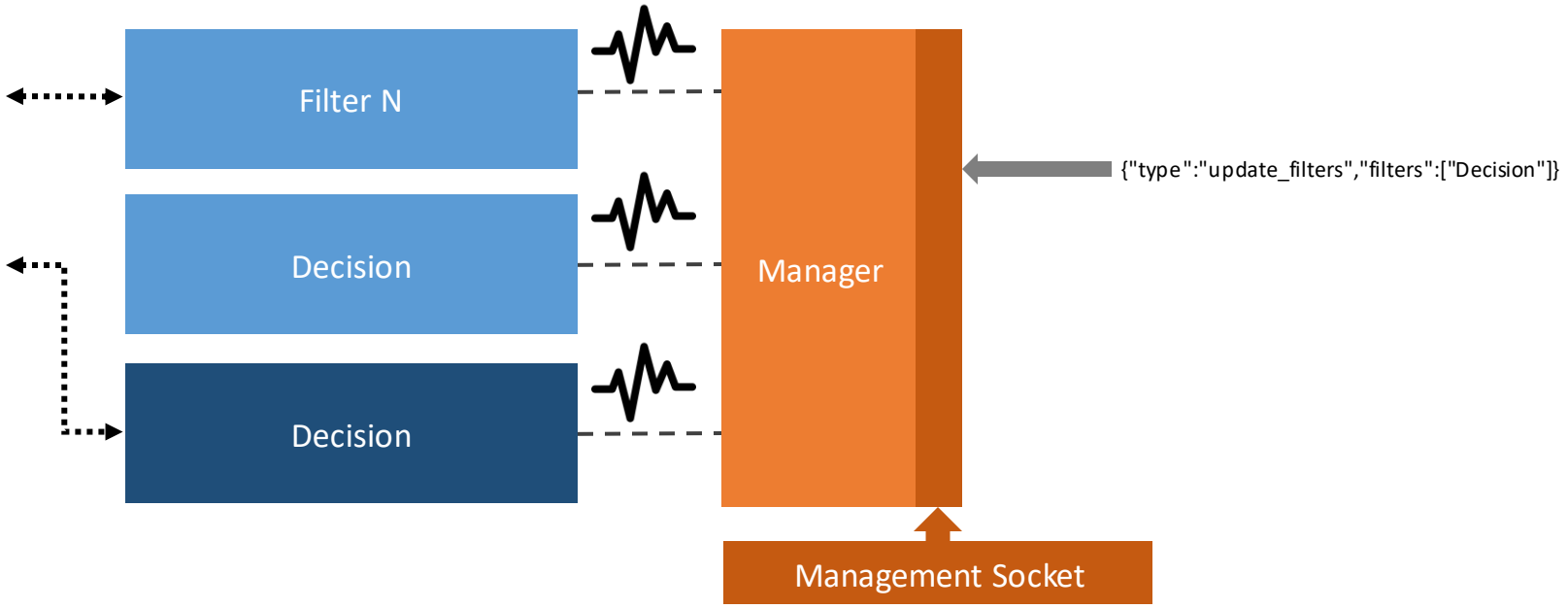
High Availability



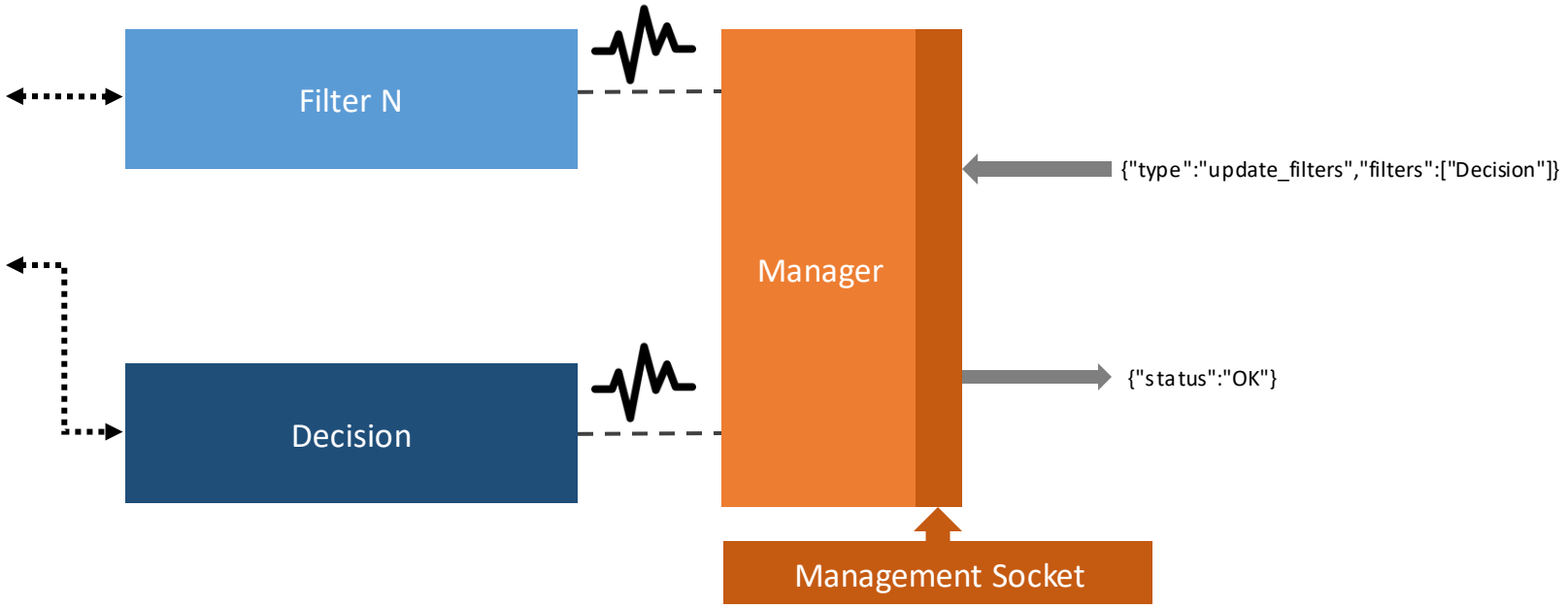
High Availability



High Availability



High Availability



Precision and Intelligence

- **Precision**
  - Multiple small filters
    - Very efficient for one unit task
    - Ability to chain filters (workflow)

- **Intelligence**
  - Decision filter based on Artificial Intelligence
    - Prediction based on filters' results
    - Active learning capabilities: Interact with human to correct itself
  - Human focuses on high-level "decisions"
    - The AI manages the technical security rules

No Bullshit

- Heuristic / basic correlation in a black box is not AI

- Those methods are promising but…
  - We use some of them in v3 (SVM, regression…)
  - Few false-positives
  - Unfortunately, few false-negatives: rules still needed

- We work hard to take it to the next level!
  - **"AI first":** by design, not an add-on component
  - Excellent results so far, beta-version coming this year

Simplicity

- Easy for users
  - Minimalist configuration
  - Autonomous system
  - Simple feedback (Normal or Malicious request)

- Easy for developers
  - Filters mostly independent
  - Simple SDK
  - On Github soon ;)

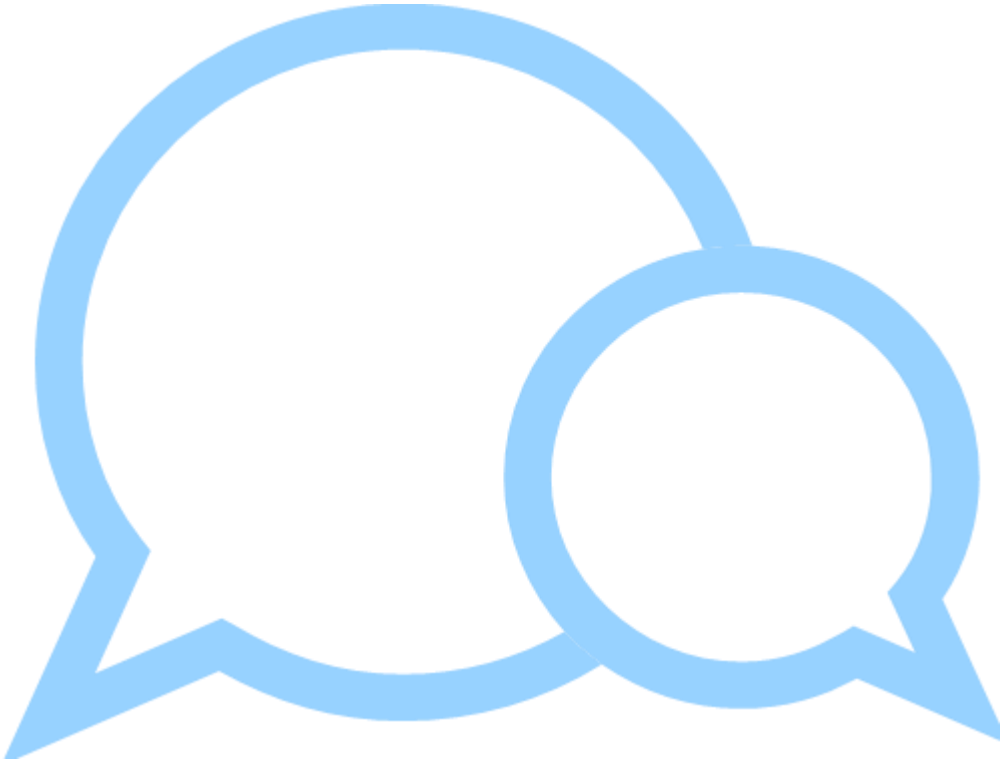
```
/// Used to configure the Task generator at startup.
bool Generator::Configure(std::string const& configFile);

/// Entry point of the execution.
/// This is what the thread will execute.
void Task::operator()();
```

Portable

- Replace HAProxy with anything you want
  - Simply develop a connector
- Not only HTTP !
- Real world example (aDvens): DARWIN + Rsyslog
  - mmdarwin plugin
  - Real time log analysis
  - Real time log enrichment
  - On Github soon… ;)

Questions ?



Thank You !

hugo.soszynski@advens.fr
jeremie.jourdin@advens.fr

https://www.vultureproject.org
https://github.com/VultureProject/mod_defender
https://github.com/VultureProject/darwin (coming soon)