Centralised logfile analysis using Elastic, Logstash and Kibana (ELK)

Tomas Krajca

⟨tomas@repositpower.com⟩

PUG Meetup
Canberra

June 1, 2016

Today

1. Reporting state of (python) programs.
2. Storing these reports.
3. Processing/indexing these reports.
4. Searching/analyzing these reports.
5. Elasticsearch.
6. Kibana.
7. ELK setups.
8. Alerting from these reports.
9. Lessons learned.
10. QA.

Why and how?

1. example1.py
2. example2.py
3. example3.py
4. example4.py
5. logging.ini
6. https://docs.python.org/2/library/logging.html
7. https://docs.python.org/2/library/logging.config.html

Logstash inputs

1. File
2. AWS S3
3. Console
4. Rsyslog
5. Heroku
6. ...

```
input {
  file {
    type => "python"
    path =>
      ["/var/log/apps/app.stderr.log"]
    add_field => [ "program", "app" ]
    sincedb_path =>
      "/var/lib/logstash/sincedb"
  }
}
```

Logstash filters

1. logstash.conf
2. https://github.com/elastic/logstash/blob/v1.4.2/patterns/grok-patterns

Logstash filters

```
{
           "message" => "2016-06-01 12:52:41,129 DEBUG base zmq.auth
                         PID: 1230  ZAP reply code=200 text=OK",
          "@version" => "1",
        "@timestamp" => "2016-06-01T12:52:41.182Z",
              "host" => "dauvmfes001.repositpower.net",
              "path" => "/var/log/apps/SecureTSDBProxy.stderr.log",
              "type" => "cc_python",
           "program" => "SecureTSDBProxy",
         "timestamp" => "2016-06-01 12:52:41,129",
          "loglevel" => "DEBUG",
            "module" => "base",
            "logger" => "zmq.auth",
       "received_at" => "2016-06-01T12:52:41.182Z",
     "received_from" => "dauvmfes001",
      "@source_host" => "dauvmfes001",
          "@message" => "ZAP reply code=200 text=OK",
              "tags" => [
        [0] "_parsed"
    ]
}
```

Logstash outputs

1. Zabbix
2. Redis
3. Elasticsearch

```
output {
  if "zabbix" in [tags] {
    zabbix {
      host => "zabbix.repositpower.net"
      port => "10051"
      zabbix_sender =>
        "/usr/bin/zabbix_sender_retry"
    }
  }
  elasticsearch {
    host => "search.repositpower.net"
    protocol => "http"
    flush_size => 1000
    max_retries => 7
  }
}
```

NoSQL database

1. NoSQL database
2. full-text search
3. index, insert and query examples in python
4. index, insert and query examples in curl
5. elasticsearch-curator

Demo

1. Search.
2. Filter.
3. Analyze.

Distributed

Centralized

Bosun

1. response time alert
2. too many warnings/errors alert
3. anomaly alert (number of errors)

Tips & tricks

1. python logging with multiprocessing
2. advantages of async logging
3. multiline logstash codec being stuck/old alerts
4. HW resources for logstash and elasticsearch
5. queuing in logstash when elaslticsearch is down
6. consensus on meaning of log levels for alerting
7. educate on how to use kibana
8. bosun very young/rough
9. python stacktraces in the logs
10. SIGUSR1/SIGUSR2
11. logstash testing/debugging (rubydebug)

Questions/discussion

Thank you

tomas@repositpower.com
https://github.com/tkrajca/PUG-ELK