

Final Report on: Pretix Redundant Operation

Filipe Pires [85122], João Alegria [85048]

Computational Infrastructures Management
Department of Electronics, Telecommunications and Informatics
University of Aveiro

June 12, 2020

Introduction

This report aims to describe the work developed for the final phase of the practical assignment of the discipline of Computational Infrastructures Management (*I*) from the Msc. degree in Informatics Engineering of the University of Aveiro at the Department of Electronics, Telecommunications and Informatics. It is assumed that the reader has knowledge about the previous report, and it is here included: the characterization of the service level agreements for each service; the description of the load balancing mechanisms among software components; the redundancy strategy; the listing of metrics and computational resources monitored, along with the respective defined alarms; and other relevant aspects such as horizontal scalability and component fault tolerance.

The service provided is Pretix, an online shop, box office and ticket outlet already successfully used by other service providers for conferences, festivals, exhibitions, workshops and more. The previously delivered work focused on the product presentation, distributed installation, and resource and performance analysis. All code developed is publicly accessible in our GitHub repository:

<https://github.com/FilipePires98/GIC>.

1 System Architecture

1.1 Infrastructure

To provide the Pretix Ticketing Software to the public, it was required to set up an infrastructure containing their Django-based web application, an instance of a Web Server Gateway Interface (WSGI) hosting the application and of a reverse-proxy for the web application deployment in production mode, as well as a database management system (DBMS) (where PostgreSQL was used) for handling disk storage and a caching server (where Redis was chosen) for in-memory storage and asynchronous queuing of tasks.

Our strategy is thoroughly described in the previous report, which already considered minimum redundancy to ensure availability. It is achieved through the usage of Docker to isolate each component for greater control and easier configuration, and Docker Swarm for the orchestration of the entire infrastructure. Nevertheless, as it was a still primitive solution, naturally several upgrades were applied, mostly internally to each container. The latest architecture is visually presented in Figure 1. This solution considers quality attributes whose assurance is described in the sections below.

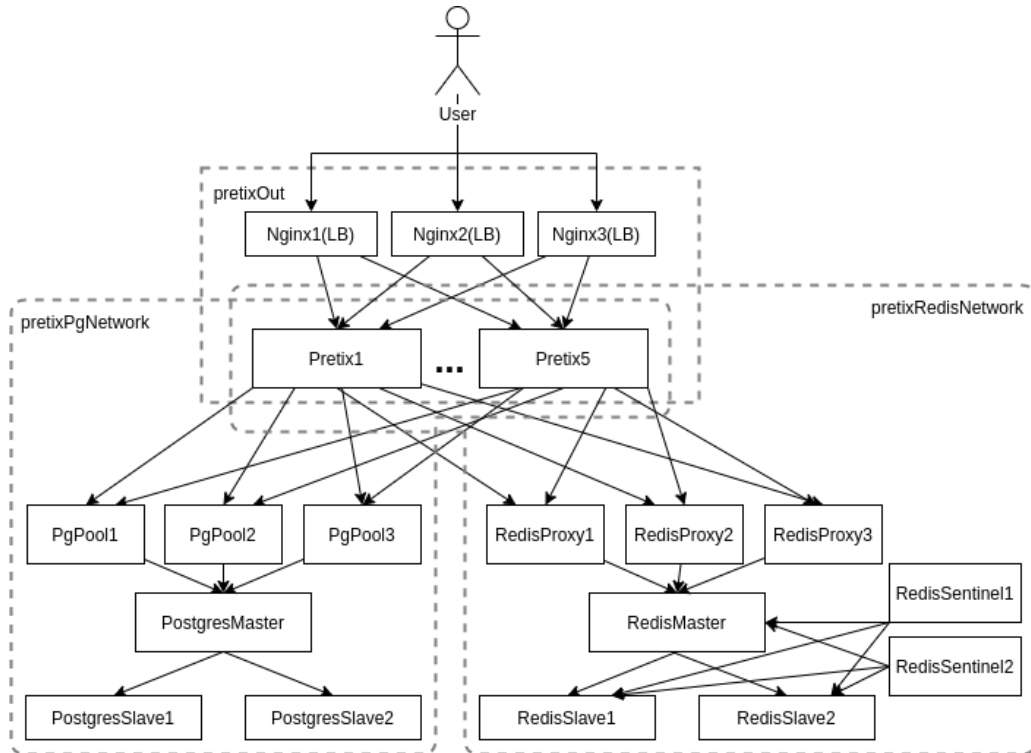


Figure 1: Infrastructure architecture diagram deployed in Docker Swarm.

1.2 Load Balancing

The knowledge gathered so far regarding the operation of Pretix affirms that user activity peaks are to be expected when in production. This has many implications and requires the implementation of mechanisms that maximize the infrastructure's efficiency.

With this in mind, we have come up with mechanisms for load balancing based on proxy servers. These servers, when strategically placed, act as intermediaries for requests seeking resources from the system itself. This not only reduces the complexity of requests made to our services, but also provides additional benefits with regards to security and load balancing, as they allow a controlled and intelligent distribution of requests among components.

An NGinX proxy was installed between the end users and the web application. With the help of Docker Swarm, replicas of both the proxy and the web application were deployed. This is transparent for the user and it works as if he/she is sending requests to one single access point.

The storage clusters also resort to proxies. For the PostgreSQL server cluster, we adopted Pgpool-II. As our database is replicated in master/slave mode, Pgpool is able to take advantage of the replication feature in order to reduce the load on each PostgreSQL server by evenly distributing queries among available servers and consequently improving the overall throughput. Pgpool provides other useful features such as connection pooling, where it maintains established connections to the servers and reuses them whenever a new connection with the same properties comes in, and automated fail over, mentioned in greater detail in section 2.3.

For the Redis server cluster, the adopted solution was the well-known HAProxy. The instances installed spread requests evenly across the Redis cluster. Here, a master/slave mode was also used, meaning that the proxies interact with master instances that then delegate work among slaves.

Load tests executed during development proved that such additions had a significant impact on the performance of the system. The usage of networks, seen in Figure 1, where the proxies mentioned above are inserted help orchestrating and logically grouping components.

1.3 Redundancy

In order for us to provide Pretix, it is only required to maintain one instance of each component running. However, this is highly inadvisable as it is susceptible to failure from any point, i.e. if any of the system components fails, the entire system goes down. This is referred to as a redundancy measure of N .

Ensuring actual redundancy of a system requires at least a measure of $N+1$, where each component can be replaced by another if required. In most cases this is achievable through the replication feature of Docker Swarm. Pretix's stack has K replicas of NGinX, of the Pretix web

application, of Pgpool proxy and of HAProxy, where K can be defined on deployment with a value greater than 2.

Redundancy was also the reason why a master/slave mode was adopted on both clusters. By instantiating 1 master and K slaves, it is possible to fulfill the requirements implicit to N+1 since if any slave fails the others can replace it and if the master fails a slave is elected the new master. Fortunately, for the PostgreSQL cluster this reelection process is done automatically, but for the Redis cluster the use of sentinels is required: sentinels observe the behavior of masters and slaves, detect failures and reassign roles when needed; here we insist on providing N+1 redundancy by having multiple instances of sentinels as well.

1.4 Scalability

Any project intended to be successful must consider the scalability of the product in the future. Ours is no exception since, if the ticketing platform becomes popular, the infrastructure must be prepared to respond to greater loads of income traffic.

There were 2 means to ensure scalability: vertical or horizontal. As we had access to limited computational resources and adding more physical RAM or CPUs was not an option, the only alternative left was to support horizontal scaling.

The idea here is to make possible the deployment of new replicas of any component and integrate them to the existing infrastructure without significant negative impact on insertion and in a completely transparent way to the users. This is exactly what we achieved. Additionally, each component identifies itself to the monitoring server on initialization so that the data related to what is being managed is constantly up-to-date.

Scaling a portion of the stack can be done both manually or in an automated way. By default we support manual scaling, but we found it was logical to implement automated mechanisms to scale up or down according to the online users and the resources used. This was achieved through the monitoring of meaningful metrics collected from each container instance, as we present in section 2.3.

2 Infrastructure Management

2.1 Service Level Agreement

As service providers, we are supposed to fulfill a commitment with our clients, whether it is formally defined or not. Obviously it is a good practice to elaborate an agreement where particular aspects of the service are agreed between the service provider and the service user - this is called a Service Level Agreement (SLA). We describe ours in this section.

Considering the scope of our project, we've defined the following commitment key points:

- Pretix is to be provided through web browsers, by accessing a particular address, and it will only be accessible to those within the university's virtual private network (VPN).
- Its usage is fully dependent on the usability of the Pretix service itself and no complexity should be added by the infrastructure we provide with regards to this.
- We are responsible for providing free access to the ticketing platform through our servers and keeping data integrity even when the service is unavailable
- The installations hosting our deployment belong to a third party (the university) and thus we are not accountable for possible availability issues related to networking and machine up time. Nevertheless, if this is not considered, we ensure an availability of 100% under normal circumstances and of 70% for high activity peaks ¹, with a maximum response time of 50 seconds for ticket purchasing requests.
- No support is guaranteed for issues about features related to the internal implementation of the ticketing software.

It is worth mentioning that our analysis to the (free version of the) Pretix product showed that the software had strong limitations on maximum load capacity and response time that, according to the authors, could not be solved through horizontal scaling.

Focusing on the infrastructure itself and the computational resources, we have defined a list of Service Level Objects (SLO) - the metrics to be observed - representative of the stack's state, and a list of Service Level Indications (SLI) - the thresholds or functions applied to the metric values - that will help us react to situations where the SLA compliance is threatened. These SLOs revolve around: memory capacity, operation latency, data transfer throughput, availability percentage, recovery time and reliability.

...

2.2 Monitoring

Understanding the state of the infrastructure and components is essential for ensuring the correct functioning and stability of the service. Information about the health and performance of

¹Over 50 requests per minute.

Category	Metric	Lower Threshold	Upper Threshold
Hardware (SNMP)	System Up Time	-	-
	% User CPU Time	-	90%
	Total RAM	-	-
	Total RAM Used	-	-
	% RAM Usage	-	90%
	Total Disk	-	-
	Total Disk Used	-	-
	% Disk Usage	-	90%
Pretix	Total # of Events	-	-
	Total # of Orders	-	-
	Total # of Purchases	-	-
	Response Time	-	50 seconds
Redis	Redis Up Time	-	-
	Memory Used	-	-
	# of Connections	-	-
	# of Connected clients	-	-

Table 1: More rows.

deployments not only helps us react to issues, it also gives security to make changes with confidence. One of the best ways to gain this insight is with a robust monitoring system that gathers logs and metrics, visualizes data, and alerts us operators when undesired events occur or when the maximum capacity of some resource is almost reached.

Our monitoring system is hosted on It consists of a virtual machine (VM), running

2.2.1 Log Collection

...

2.2.2 Metric Collection

2.3 Alerts and Automation

...

3 Additional Remarks

3.1 Swarm Stability

...

3.2 Assignment Contributions

|||||| HEAD Due to external issues that prevented physical proximity among students and professors, a confined approach had to be adopted during this second development phase. Nevertheless, the close-contact strategy was kept as this project required a broad perspective on the infrastructure and the tasks to be implemented by both team members. Regarding the work distribution among developers, tasks were assigned to each according to a predefined plan and a common working schedule was applied. ===== |||||| HEAD Regarding the work distribution amongst developers, a close-contact strategy was defined where each worked on a cluster component or piece of software according to a predefined plan. ===== Regarding the work distribution among developers, a close-contact strategy was defined where each worked on a cluster component or piece of software according to a predefined plan. |||||| 733311d3d5d304f0545f0cf5008b0bde8fced17e The cluster strategy and respective details were decided in conjunction, as well as the key objectives and tasks to be achieved before the final delivery deadline. |||||| 7a7ddb2fed714a84a50822704febe51539a1dcee

The technology-related decisions were first made in conjunction. Then, Filipe set up the monitoring VM to gather the data from the stack while João began to work on the mechanism to collect the logs from each container. Filipe moved on to installing metric exporters on each container, and João found means to collect lower level metrics using the SNMP. The container identification mechanism of each instance to the monitoring system was then implemented. At this stage, both operators worked on defining the SLA and João designed the monitoring dashboards. Finally, the alarms and automated response mechanisms were implemented. It is needless to say that bug and error solving was made along the development phase by both developers any time it was required. |||||| HEAD Once a stable deployment fulfilling all predefined requirements was completed, this report became our primary concern, with both contributing equally. ===== Once performance benchmarking and bottleneck identification were completed, this report and the code documentation became our primary concern, with both contributing equally. |||||| 7a7ddb2fed714a84a50822704febe51539a1dcee

Conclusions

...

References

1. J. P. Barraca, *GIC - Report no.2: Simple Product Operation*, University of Aveiro, 2019/20.
2. *About Pretix*, <https://pretix.eu/about/en/>. Pretix.eu, retrieved in April 2020.
3. *Welcome to Rami.io*, <https://rami.io/>. Rami.io, retrieved in April 2020.
4. *Pretix Code Repository*, <https://github.com/pretix/pretix>. GitHub, Inc., retrieved in April 2020.
5. *Welcome to pretix' documentation!*, <https://docs.pretix.eu/en/latest/>. Pretix.eu, retrieved in April 2020.
6. *Docker Homepage*, <https://www.docker.com/>. Docker Inc., retrieved in April 2020.
7. *pretix/standalone*, <https://hub.docker.com/r/pretix/standalone>. pretix, retrieved in April 2020.
8. *Meet Django*, <https://www.djangoproject.com/>. Django Software Foundation and individual contributors, retrieved in April 2020.
9. *Gunicorn Homepage*, <https://gunicorn.org/>. Gunicorn.org, retrieved in April 2020.
10. *NGinX News*, <https://nginx.org/>. NGinX.org, retrieved in April 2020.
11. *PostgreSQL: The World's Most Advanced Open Source Relational Database*, <https://www.postgresql.org/>. The PostgreSQL Global Development Group, retrieved in April 2020.
12. *About Redis*, <https://redis.io/>. Redis Labs, retrieved in April 2020.
13. *Welcome to Pgpool-II*, https://www.pgpool.net/mediawiki/index.php/Main_Page. PgPool Global Development Group, retrieved in April 2020.
14. *HAProxy: The Reliable, High Performance TCP/HTTP Load Balancer*, <https://www.haproxy.org/>. HAProxy.org, retrieved in April 2020.
15. *Locust, an open source load testing tool*, <https://locust.io/>. Locust.io, retrieved in April 2020.
16. *About Selenium*, <https://www.selenium.dev/about/>. Software Freedom Conservancy, retrieved in April 2020.
17. *Ansible*, <https://www.ansible.com/>. Red Hat, retrieved in April 2020.