

Intermediate Report on: Pretix Cluster Deployment & Management

Filipe Pires [85122], João Alegria [85048]

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

April 29, 2020

Introduction

This report aims to describe the work developed for the intermediate 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 here included: a product description and the respective expectations regarding client capacity, derived from a thorough analysis; the adopted clustering strategy, considering matters such as virtualization, storage, network and load distribution mechanisms; an explanation of the performance analysis executed based on the department's support platform for representative use cases; and the results drawn from such analysis, to be considered during the second phase of development.

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. All code developed is publicly accessible in our GitHub repository:

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

1 Pretix Ticketing Software

Pretix (2) is an all-in-one product, serving as an online shop, box office and ticket outlet, meant for companies that aim to host events of numerous kinds. This ticketing software created by Rami.io (3) has a wide variety of features that range from ticket shop customization (domain definition, multiple language support, product structure with multiple categories, packages and possibility for donations, seating plans, website embeddings, etc.) and marketing (voucher system, email, campaign tracking, etc.), to payment definition (invoicing and multiple payment methods) and on-site management (check-in lists, ticket scans, etc.), and even administrative control (with statistics and event-spanning reports, team permissions and data export and API).

As a complete solution, Pretix has already been successfully used for conferences, festivals, concerts, shows, exhibitions, workshops, barcamps, and more according to the creators. The software's source code is completely open source (4), and the authors make available extensive documentation (5), an installation guide and a support blog.

1.1 Our Product

What we offer, as a developers team, is the Pretix' infrastructure deployment and management as a service that could potentially be provided to businesses interested in hosting events by taking advantage of Pretix' many advantages, without having resources or knowledge of how to deploy it themselves. Our contract is simply to ensure the correct, reliable, secure and constant functioning of Pretix, capable of supporting usage scenarios of the same magnitude as of the one described next.

1.2 Use Case Scenario

In order to ensure the fulfillment of our contract, a representative scenario of usage was defined to guide us throughout development. A simulated organization intends to host a WebSummit-like event by the end of this year. It has analysed the feature list of Pretix and wishes to adopt it in the future. However, as it is its first time using the software, the company decided to use it simply as a ticket purchase platform and reduced the payment methods to on-site manual payment only. The organization expects to sell 10 000 tickets in a week and warned us that in previous events they witnessed a peak close to the end of the selling period.

In practical terms this means that we only need to concern ourselves with the online ticket purchasing features and do not even need to consider the actual payment process, although it was our intent from the beginning to support most of Pretix's features even if most would not be used. It also means that we had to build the system considering our hardware limitations and the activity peak expectation.

1.3 Support Requirements

As stated in their documentation, to use Pretix it is required:

- Pretix and the python packages it depends on
- An WSGI application server
- A periodic task runner
- A database
- A reverse proxy
- A server for caching, session storage and task queuing

As we intended to deploy the product in containers, Docker became an additional requirement.

Regarding the hardware remotely available at our department, it was allocated for us an infrastructure of a total of 58 nodes and 115 cores, with 5.6TB of storage and 235.8GB of RAM available. However, this is destined for several developer teams, meaning that we must only use a small portion of such resources, by our estimates of less than 10% of the total capacity.

Also according to Pretix' documentation, the software has some scalability limitations. In some cases their software needs to "fall back to event-global locking for some actions which are likely to run with high concurrency and cause harm". For every event, only one of such locking actions can be run at the same time. With this in mind, they estimate a maximum capacity of approximately 500 orders per minute per event, "even if you add more hardware". However, they also mention that if the event to be hosted has an unlimited number of tickets (which is our case), fewer locking is applied and thus it is possible to reach approximately 1500 orders per minute per event.

Considering all of these aspects, we made an estimation of how many ticket purchase requests we would be able to support per minute. At this preliminary stage, we believed it could be achieved a capacity of 800 ticket purchases per minute without a relevant amount of failed attempts. In chapter 3 we explore our own system in an attempt to determine whether our estimations were correct and feasible or not.

2 Clustering Strategy

2.1 Architecture

Docker (?) allows us to deliver the infrastructure in isolated containers with well-defined communication channels. This dramatically reduces infrastructure resources required and facilitates its management, while ensuring consistency within the whole environment. Also, its simplicity in configurations and rapid deployment makes it perfect for its use in any system with a reasonably large size and complexity.

By taking advantage of Pretix' support for Docker, its dependencies were covered after analyzing which technologies would be most appropriate for our scenario. With an already fully functional Docker Image publicly available (?), the authors of the product made the deployment of the web application easier. In it is the Django-based (?) web application that serves all the services and resources needed for Pretix, a instance of Gunicorn (?), a Web Server Gateway Interface (WSGI) hosting the mentioned Django application and lastly an instance of NGinX (?), a reverse-proxy needed to allow the for the Django application in production mode, since that framework when in production divides the static resources(that should be served by a reverse-proxy) and the dynamic resources that are assured by Django itself. Pretix' container configuration file also simplified the connection between the product and the storage containers, namely for the required database and caching server.

To handle all storage we chose PostgreSQL (?). This relational Database Management System (DBMS) had powerful tools to deal with scalability and redundancy, so it was quickly understood as a good option, although others like MariaDB were strongly considered as well. It is the PostgreSQL that stores the data related to organizations, events, tickets, clients, invoices, etc. In our case a PostgreSQL cluster is instantiated for redundancy purposes in a Master-Slave mode. Currently, a single PostgreSQL master and 2 slave nodes are instantiated, and if the master fails the system is capable of assigning one of the slaves as the new master.

A similar approach was followed for the server for caching. This server has the responsibility of storing in-memory relevant data such as open user sessions, live events, etc. and of asynchronously queuing tasks. In this case, the chosen tool was Redis (?), instantiated in a Master-Slave cluster mode with 1 master and 2 slave nodes as well. The failover process is also supported with the help of sentinels. Sentinels are nodes natively supported by redis which have the responsibility of over-watching the cluster and in case any change occurs, such as the addition or removal of a node(slave or master), it registers that modification and notifies the necessary nodes so that they can do the appropriate action. In our case we opted to instantiate two sentinels.

A problem encountered during the infrastructure development and configuration was the fact

that if a master node really crashed in one of the storage clusters, the connection to the storage infrastructure were lost since the connection to the cluster is through it's master(s). Being in a distributed environment and the master machine being down, the Pretix service was incapable of locating the new master. This fact in addition to the efforts made in balancing the load to the service, cause in we making the additional effort of finding a proxy for both clusters.

Pgpool (?) was introduced for that reason. This is a middleware that works between PostgreSQL servers and PostgreSQL clients, and enables connection pooling, replication, load balancing and more, but the mentioned features were the ones interesting for us. This component served as a proxy that by connecting to it always redirected to the master of the PostgreSQL cluster and if there were more than one master in the cluster the requests will be distributed in a balanced fashion over the available masters.

For Redis, an HAProxy (?) was found to be the best option for our scenario, since it provided us a reliable and high performing load balancer. This component, similarly to the one mentioned above, works as a proxy that forwards the request to the redis cluster always to one of it's main, as well as balanced the load in case more than one master is present.

The last component of our architecture is a reverse-proxy of our own, implemented using NginX, with the objective of balancing the requests made to the Pretix service itself. This component, as suggested by the name reverse-proxy, also enabled us to access multiple instances of the Pretix server through a unique endpoint, facilitating the interaction between clients and service.

To provide this service we focused on some aspects that we considered paramount: availability, consistency, maintainability, scalability, redundancy, fault tolerance and load balancing. Those were the reasons we implemented the storage infrastructures as clusters, we implemented the mentioned proxies, both for the storages systems and the service itself, and we made every service redundant by having replicas, or through clusters.

The result is visible in Figure 1.

2.2 Deployment

For this project there are 2 methods available to deploy the service stack: in a normal docker through the usage of Docker Compose or in Docker Swarm mode. The intended method was the second one, a deployment in Docker Swarm since a shared powerful infrastructure hosted in the DETI's department was made available for all groups. We then implemented a stack deployment for Docker Swarm but also developed a deployment for the normal Docker Compose, since the two are very similar. Both alternatives respect the architecture mentioned above but are naturally deployed with some differences in configurations, since each mode supports different features.

To create this stack much research was made to find the best already available solutions

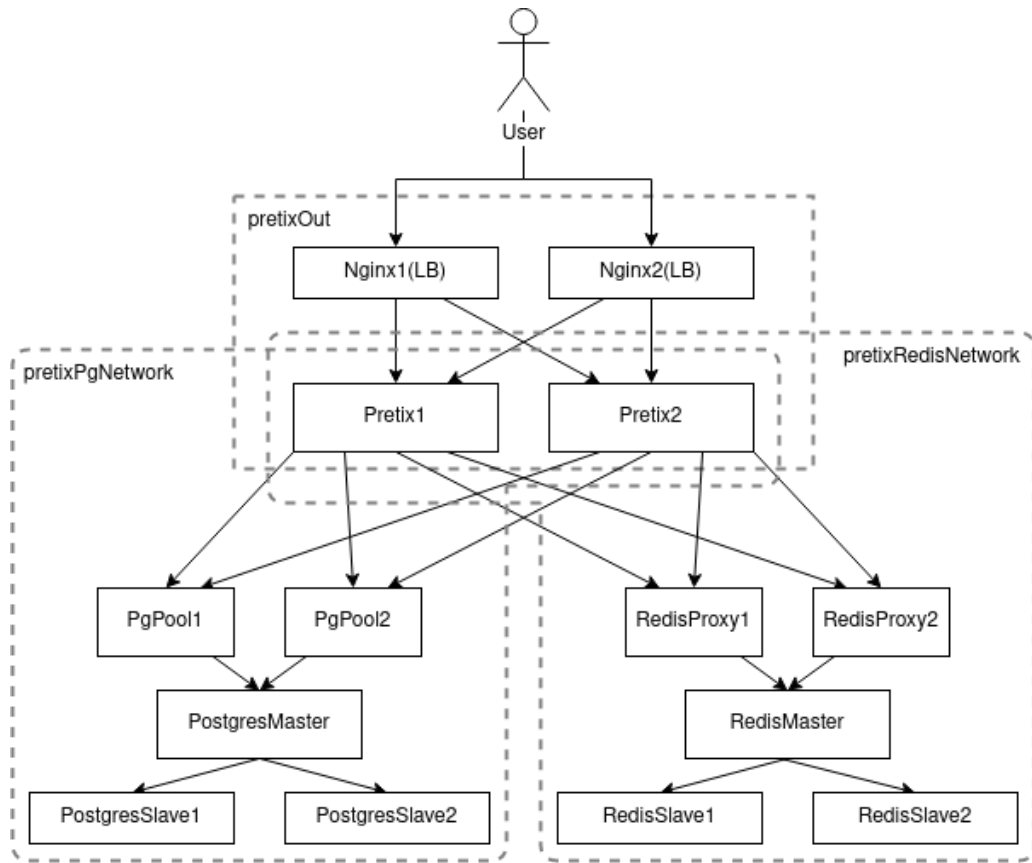


Figure 1: Infrastructure architecture diagram.

and containers that suited our needs. This as well as the integration of every component to the remaining ones turned out to be quite more challenging and complex than previously assumed. In the ended we settled in the following images for our components:

- PostgreSQL Master/Slave: <https://hub.docker.com/r/bitnami/postgresql-repmgr>
- PostgreSQL Pool(PgPool): <https://hub.docker.com/r/bitnami/pgpool>
- Redis Master/Slave/Sentinel: https://hub.docker.com/_/redis
- Redis Proxy: https://hub.docker.com/_/haproxy
- Pretix: <https://hub.docker.com/r/pretix/standalone>
- NginX: https://hub.docker.com/_/nginx

Obviously, some of this images serves as base for self created images. This was necessary and desirable because in certain situation we needed a higher degree of control, e.g., for a container to wait for other containers to start first. This necessity for self-created images happen in the Redis Slave to configure the Redis server to be a slave of a given machine and wait for

the cluster master, in the Redis Sentinel to configure the Redis server to behave as a sentinel of the cluster and wait for the cluster master, the Redis Proxy to configure the HAProxy so that the proxy would have in consideration all redis nodes and access always the master and in Pretix for the container to wait for the PostgreSQL and the Redis Proxies.

With each component defined we needed a way to connect them all, and a unique network didn't seem the best option. We decided to divide the stack in a logical way, each cluster had it's own network, NginX had its own network and the Pretix service was the linking element between the three networks.

Another aspect that we took in consideration was the data produced by the application itself. We needed a consistent and fault-tolerant system, so the data produced by the system could not be lost in any situation. For that reason we needed create and assign volumes to the containers so that no information was lost even if all services crashed. All volumes were defined and mount at a previously defined Network File System(NFS) provided by the discipline's professor.

Lastly, we had some configuration files that needed to be injected to some containers, namely the Pretix server and the NginX reverse-proxy. We decided that the best way to inject those files to those containers is by secrets. In our stack definition file we defined those files as secrets and to which container each secret should be assigned. Those secrets are generated automatically when the stack is deployed.

Both Docker Compose and Docker Swarm accepts the same stack configuration file. This is quite useful since the same configuration file we developed for the Swarm mode allows the deployment in both of them. But as already mentioned, there are significant differences between them, that in practical terms translates to certain fields being proceed in a mode and not in the other, namely the `deploy`, `configs` and `secrets` fields are not processed in the Compose mode. This meant that the automatic replication provided by Swarm is not available as well as the inject of configuration files and secret files. For that reason, since we mostly debugged and tested locally in normal docker with Docker Compose, we felt the need to create a auxiliary stack configuration file that would deploy the system with exact copy of the one in Swarm(missing the NginX replication). In this file the only changes were the mounting point for the volumes, since this configuration was not intended to be deploy in the remote infrastructure, the replication had to be made manually by creating different containers and the otherwise configuration files injected by secrets API needed to be injected as volumes.

2.3 Distribution

Docker Swarm not only adds some quite important features in the quality assuring realm, such as replication, load balancing and more, but it does that in a distributed way. This mode allows the Swarm mode to spread across as much machines as the user wants. The provided remote

infrastructure is spread across three machine, where in each one several virtual machines were then created so that the swarm had even more agents to use, simulating a truly heterogeneous and distributed environment.

Because we deploy our service stack in the Docker Swarm established in the remote infrastructure, our stack was distributed by design, which was one of the feature that appealed us in the beginning. This basically means that independently of the available machines, the number of machines, the machines' operational system, the location of the machine, and so on, installing docker on all of them and establishing the Swarm mode across all of them, we can safely guarantee that our stack can be deploy in that scenario and it will work as well as in the current one.

Through the usage of the provided platform, we concluded that the distribution is made automatically by assigning a task to an agent. A task is normally a replica of a service and the agent is a machine running a docker daemon somewhere in the Swarm network. Since each service consists in one or more replicas, they are spread across the Swarm network as the platform finds fit.

This distribution is quite important since it enables the user to gain access to the power of not only one machine, but all assigned to tasks of his stack, making possible smaller response times. This does not happen when a local deployment with Docker Compose is made, since this approach uses only one machine for all services needed, overloading all computational resources and possibly increasing the response times since the machine is worker more when compared to the ones in Swarm mode.

3 Current Cluster Performance

In order to determine the true capacity of our deployed solutions, a set of measures were performed both locally and on the remote, shared and more powerful infrastructure. Although replication and redundancy were already concerns of ours, we were not yet focused on determining availability statistics. Rather, our aim was to test our services in terms of load capacity.

We wished to determine how many users were supported simultaneously accessing Pretix and attempting to purchase tickets to our fictitious event. To do this, we performed benchmarking tests and tracked the number of failures both in the act of requesting a purchase and in the act of storing purchases in our databases. In this chapter we describe the efforts made for benchmarking and the bottlenecks discovered from our performance evaluation.

3.1 Benchmarking

After doing research on available tools and frameworks, we reached the conclusion that the most adequate would be Locust (6). Locust is an open source, intuitive, distributed, user load testing framework intended to "figure out how many concurrent users a system can handle" by programmatically simulating thousands of users called locusts. During tests, systems are "attacked" by a swarm of locusts whose behavior is defined by the developer in Python and the swarming process is monitored from a web UI in real-time.

When defining the behavior of our locusts, some considerations had to be dealt with. The job of a locust wasn't simply to send an HTTP request to purchase a ticket; real users need to go through a purchasing process on Pretix' web platform, and in that process the user's browser requests far more data from the server than the mere purchase action. So this process needed to be somehow simulated as well. Several possibilities were considered to achieve this.

Selenium (7), a suite of tools for automating web browsers, is a widely used solution for testing purposes; however, we intended to launch thousands of virtual users and so, even with options such as launching browsers in headless mode (without rendering page elements), this created a bottleneck of processing elements that were not relevant to us and did not represent the true capacity of the system, thus rendering the whole benchmark invalid. Other alternatives like wrappers to drive browsers in Python were tried but proved to be too limited for our purposes.

The solution was to manually go through the purchase process once on the Pretix website and collect all network requests that were being made, including CSS resources, image files, Javascript code and requests sent to Pretix' REST API. Amongst these requests was the actual purchase action in a message of type POST. Then, we replicated all requests on the locust behavior and, by reading Pretix' API documentation, were able to manipulate the POST request to seem to come from different users. We also included optional data on such requests, like invoices, to increase the load the server needed to process and so establishing a worst-case-scenario usage that would give us more comfort regarding the faithfulness of the results.

3.2 Execution

Having established the behavior of the fictitious users, we proceeded to actually executing our benchmarking. This was divided in 2 groups of tests of 4 phases each. The first group was entirely performed locally in one of our personal computers, the second was performed remotely using our Docker Swarm. This was done in order to clearly comprehend the differences of performance between the 2 deployment methods and detect which bottlenecks were common to both, which were surpassed with greater hardware capacity and which appeared only in the second group (if any).

The 4 phases consisted in load testing the solution with 4 different activity peaks: one with half the expected load, one with the expected load, one with twice as many expected users and one with 10 times as many users. These were defined a priori and the expected load was set according to the analysis described in section 1.3.

The results of the first group of tests are presented next. ...

With the system fully deployed using Docker Swarm, the second group of tests was executing, resulting in the following values. ...

3.3 Bottlenecks

...

4 Additional Remarks

4.1 Pretix Quickstart

During development, a need for a faster basic setup of Pretix emerged for debugging purposes. As many of our tests assumed the existence of organizers and events already inserted on the databases, each time these had to be cleared the objects had to be manually created before proceeding with some implementation process.

In order to respond to this need, we developed a Python script capable of completely dispose of the manual labour. This script, which we called `quickstart.py` and placed under the scripts folder, uses Selenium (7) to create an organization inside Pretix, generate a secret API key for sending REST requests (storing it in a text file), create a default event and prepare it for load testing with Locust. It also contains configurable variables to personalize its output.

Such quickstart tool was not only found useful for the obvious advantage in development speed it brought to us, but was also considered a good help tool for any person that wishes to quickly experiment our infrastructure without having to know much about Pretix itself.

4.2 Planning Phase 2

Although with a service stack already replicated, redundant and tolerant to failures, now that we have a better understanding of the service itself, it's bottlenecks and it's behavior when handling with higher volumes of data and incoming requests we can personalize and optimize the stack for Pretix itself. For that reason our concerns shifted towards making the solution more fault tolerant, scalable, reliable and in general more robust, taking advantage of the already obtained conclusions on what components limited our service and needed more attention.

A big aspect that we already considered, but was from the beginning destined to be tackled in the second and last delivery of this project is the handling of errors, crashes and traffic fluctuations. In practical terms this is translated in the development of automatic mechanisms that allows the infrastructure manager to be certain that a service will never fail or to quickly scale the service's stack according to incoming traffic. Some efforts related to this theme were already made on this first delivery, namely the implementation of clusters for both of the storage systems necessary for Pretix; our implementation of both clusters already guarantees replication and redundancy of information and tolerant to failures since any node can fail, redefining the cluster if necessary.

The proposed further work in this area is to develop Ansible scripts that will enable us to scale automatically our service stack and possibly handle automatically with service crashes and their respective redeployment. Ansible (8) is a IT automation engine that automates cloud provisioning, configuration management, application deployment, intra-service orchestration,

and many other IT needs, exactly the type of engine we needed to further improve our service infrastructure.

Some more questions will be taken in consideration, namely security. By injecting any password needed to services through the secrets Docker API we aim to achieve a more secure and penetration resistant system. This was already done with configuration files needed to be injected in some of our services, but we intend to extend this process to any critical information, such as passwords.

4.3 Documentation

A great portion of the code we deal with isn't of our authorship, so the documentation it contains is the documentation we get. Nevertheless, the code we developed has specific purposes and, although usually very much self explanatory due to its nature, should be easy to interpret by any person. Our goal is to ensure that what we deliver can be easily reused or even continued by other developers.

With this in mind, we made significant efforts on ensuring that all code developed by us followed a common structure with the same coding style. Also, we made sure all source files have a description of their purpose and relevant information on the first lines, and contain comments on key points of the code explaining snippets considered of greater complexity/importance.

4.4 Assignment Contributions

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. 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.

Pretix' features were explored by both developers in order to better comprehend the web platform and the REST API. Then, João focused on the deployment of our solution on the department's infrastructure while Filipe focused on creating and executing the load tests. The analysis of the results were carefully conducted in conjunction as well, in order to maintain a common knowledge for better workflows. It is needless to say that bug and error solving was made along the development phase by both developers any time it was required.

Once performance benchmarking and bottleneck identification were completed, this report and the code documentation became our primary concern, with both contributing equally.

Conclusions

...

References

1. J. P. Barraca, *GIC - Report no.1: 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>. GuitHub, Inc., retrieved in April 2020.
5. *Welcome to pretix' documentation!*, <https://docs.pretix.eu/en/latest/>. Pretix.eu, retrieved in April 2020.
6. *Locust, an open source load testing tool*, <https://locust.io/>. Locust.io, retrieved in April 2020.
7. *About Selenium*, <https://www.selenium.dev/about/>. Software Freedom Conservancy, retrieved in April 2020.
8. *Ansible*, <https://www.ansible.com/>. Red Hat, retrieved in April 2020.