

Cloud infrastructure for high concurrency

Jordi Molina, Teradisk

Antoni Bertran, 3ipunt

Agenda

- Who we are?
- Introduction
- The unicorn
- The reality
- Splitting the architecture
- Elasticity
- CI/CD
- Costs skyrocketing
- Plugins and Themes
- Stress it!!!
- Check environment version

Who we are?



CTO Teradisk (2014)

 @warper
j.molina@teradisk.com



CTO 3ipunt (2004)
Moodle Partner

 @tunyafix
antoni@tresipunt.com



The unicorn and the reality



Unlimited Scalability

Deploy your current code without changes

Forget about DBAs!! Save Money!

Everything is covered by an API!

Provider sets limits and technologies have their own as well

PaaS put their own restrictions in place, changes are needed

You only forget about Backups and Updates, kinda

Every service has an unrelated API



So, you got a visit from a sales representative of <name_of_partner_or_public_cloud_provider_here>.

Did he tell you something like:

- Her cloud offers almost unlimited scalability
- You'll be able to deploy your code without worrying about underlying infrastructure
- You'll be able to forget about DBA and other managed services
- You'll save money
- It'll be extremely easy, because everything has been designed from the API

Scalability is there but:

- There are limits everywhere and for every service
- Loadbalancers are not as magic as they say

Easy deploying is easy to achieve, relatively:

- PaaS offerings can ease the task, but only if your application is flexible enough to be accommodated there
- There are still computing limits beneath a PaaS or a CaaS service

Managed services get part of the burden but:

- You still need to optimize queries and DBMS parametrization, and PaaS DB services have their own specifics

You'll save money: eventually

API's make everything super easy: after you learn them

Splitting the architecture

>99,5 % SLA and tons of users	< 99,5% and not a lot of users
Load Balancers Pool of application servers Networked Moodle Data Separate Cache service for each cache layer Separate search accelerator Separate database	a) Single VM a) Single VM for app + caches VM or managed DBMS



 moodlemoot

Azure, GCE and AWS offer between 99% and 99,5% SLA in most resources

If you are looking for high availability (over 99%) you'll need to split (ideally between AZs in the same region):

- Load balancers: think of balancers that favor routing, HTTP/2 and certificate management
- Application server(s): use latest generation whenever possible
- Moodle Data: Cloud vendors have NFS offerings with a lot of caveats, and sometimes NFS clustering solutions offer better performance and pricing
- Cache systems
- Search accelerators
- Database

Don't kill me for this tip:

If you don't require high availability, best shot is to go for a single VM with all services inside:

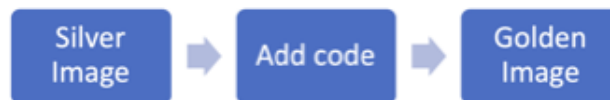
- No network latency
- Disk per service if I/O bound
- Moodle Data is a local FS, no NFS issues
- If required at a certain stage, you can decouple one of the components

CI/CD

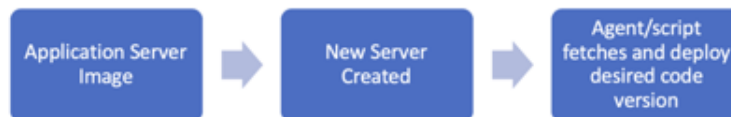
Single VM: Nothing changes

Autoscaling group:

- Golden Image approach:



- Deploy code on boot approach:





Autoscaling group:

- Golden Image approach: fastest boot time, requires you to have a CI/CD pipeline that merges a Silver Image with all requirements and the latest version of your code (easy to do with things like Packer and GitlabCI). Container Based autoscaling looks like this.

This approach allows you to easily perform blue/green and canary deployments.

- Deploy code on boot approach: requires you to write a process that downloads the code every time a new machine boots
- Explain the flow of creating a silver image, a golden image and how long it takes to

boot and warm a new instance

- Explain blue/Green , canary deployments
- Explain deploy code on boot strategies: using an agent (like AWS's code deploy, chef, or puppet) to ensure rollback capacity and control of which version of your code is there

Elasticity

So, your Moodle instance requires >99% availability and has literally tons of users... use this checklist!

- Base minimum capacity
- Maximum expected capacity
- Healthchecks
- Increase fast, decrease slow
- Warmup newly spun instances
- Bigger instances > Small instances (when dealing with PHP monoliths)
- Separate pool of instances for users and admin/scheduled processes
- Know your cloud: do balancers need something before being hit by users?
- Scheduled capacity



- Base minimum capacity: capacity needed to survive the non-peak hours without spinning new machines
- Maximum expected capacity: how much your boss allows you to spend.
- Both base and maximum can change over time
- Healthchecks are critical to ensure that an application server is "healthy"

- Increase in 2 or 3 VM's, decrease by 1 and take your time
- ALB's in AWS allows you to use a warmup policy on newly registered servers, Azure and GCE have similar functions or allows some form of time allowance.
- It makes no sense to spin 6 or 8 4 cpu VM's within an hour when a peak of users come, better spin 2 16cpu vm's when the surge is detected
- LoadBalancers allow you to redirect traffic for the admin site to a specific instance or set of instances
- Some cloud providers require you to prewarm your loadbalancers if you are expecting a non linear traffic surge, and sometimes you need a support paying account (as little as 50USD) to be able to open a ticket for this
- If you know a lot of users come at 9AM, increase the pool by 8:45-8:50 AM

Costs do skyrocket

Costs will eventually and unexpectedly rise on public cloud, until you learn how costs work in your public cloud provider at a 100%. Watch out for:



\$++



\$++



\$++



\$--



\$++



\$++

Things to avoid:
The finance guy
feeling like this -->



moodle moot

Bidding image by yieldbird.com; capacity image by tashatuvango; cleanup image by openstreetmaps;

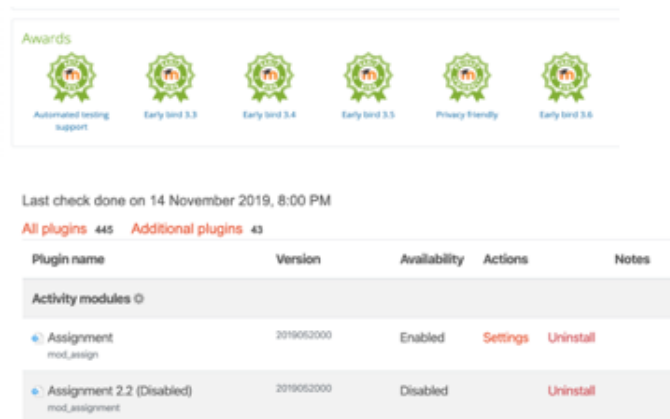
- All outgoing traffic costs money (even between availability zones in the same region)
- Most modern cloud load balancers bill you by processed data (both in throughput and traffic)
 - You can do nothing here but pay, in some strange cases an ASG with self managed load balancers in a VM has become cheaper than a managed load balancer
- Buy or reserve your base capacity once you get it sorted (AWS has simplified this recently)
 - AWS has simplified a lot the way you reserve capacity since last week, the other players have similar tools or options to grant you savings in long term resources
- Take advantage of unused capacity of your provider (Think about Fleet Management in AWS: mixing Reserved Capacity, On demand instances and Spot offerings)
 - Again, AWS has the better offering for this kind of machines, but GCE's preemptible VM's can be leveraged somehow and Azure Low-

Prio VM's too

- **Cleanup after use**
 - Golden Images can rack up easily with multi level environments (think staging, test, prod). They use space. Space costs money.
- **Maintain an eye on your bill**
 - Depending on the activity in your cloud account, you'll need to check daily, weekly or monthly, but it is always a good idea to set up a cost alarm when things start to look ugly and another one when they are really ugly. Make sure people in finance get the alert as well, they use to scare easier than tech management.

Plugins and Themes

- Analyze
- Plugins core vs additional
- Themes core vs additional
- Additional
 - Who develops it?
 - Has support?
 - Usage?
 - Versions updated?
 - Awards?
- Do we need all these plugins?



Awards

Automated testing support Early bind 3.3 Early bind 3.4 Early bind 3.5 Privacy friendly Early bind 3.6

Last check done on 14 November 2019, 8:00 PM

All plugins 445 Additional plugins 43

Plugin name	Version	Availability	Actions	Notes
Activity modules 0				
Assignment mod_assign	2019052000	Enabled	Settings Uninstall	
Assignment 2.2 (Disabled) mod_assignment	2019052000	Disabled	Uninstall	

moodle moot

1. Analyze => go to yoursite/admin/plugins.php and check!
2. We need all these Additional plugins? Review the instances
3. Themes we need 4 additional themes (we tested in the past)
4. Additional

Plugin Benchmark

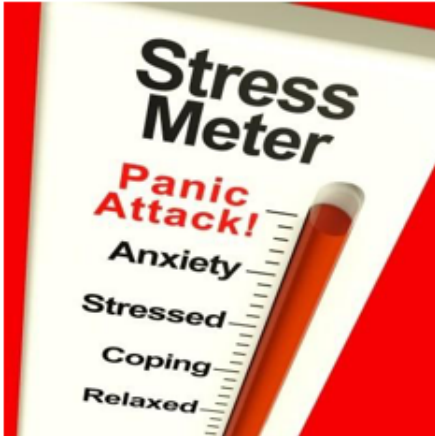
Performs various tests to determine the performances of your Moodle
https://moodle.org/plugins/report_benchmark

Very good approach but we need more analysis



This plugin helps giving a result of your server performance

Stress it!!



Source: <https://jtway.co/>

3 | with a blue highlight

Site administration

[Site administration](#) [Users](#) [Courses](#) [Grades](#) [Plugins](#) [Appearance](#) [Server](#) [Reports](#) [Development](#)

Development

- [Debugging](#)
- [Web service test client](#)
- [Purge caches](#)
- [Third party libraries](#)
- [Acceptance testing](#)
- [Make test course](#)
- [Make JMeter test plan](#)
- [PHPUnit tests](#)
- [Template library](#)
- [XMLDB editor](#)



How to stress it?



1. Create new environment (*never in production*)
 - Same HW like production
 - Same Architecture (Load Balancer, Caching, DB, etc) like production
 - Same Moodle version and plugins like production
2. Server monitoring tool (recommended)
3. Enable development
4. Make test course
5. Make test plan
6. Launch de client JMeter
7. Analyze it!



Server monitoring tool Zabbix for instance

Make test course

php

admin/tool/generator/maketestcourse.php

Make test plan

php admin/tool/generator/maketestplan.php

Questions?



 moodlemoot

Server monitoring tool Zabbix for instance

Make test course

php

admin/tool/generator/maketestcourse.php

Make test plan

php admin/tool/generator/maketestplan.php



THANKS!

