Data lake migration to cloud
a modern infrastructure approach

Ramon Gonzalez
Cloud and Analytics Architect

REAL use case

High number of files to be stored
non-structured and semi-structured data

Automation
Solution should be reusable without "any kind"
of human effort

Collaborative
It should implement a collaborative way of work
usable by a distributed team

About me

Ramón González Castro

INDITEX — Cloud Architect (2019 -)

BI & Analytics PFE (2014 – 2019)

@ramonbgc

https://www.linkedin.com/in/**ramongonzalezcastro**

Automation

- Has been used for a while (for several tasks, processes…)
- Really(?) well defined for some of them

- Configurations: ANSIBLE

- CI/CD: **Jenkins** GitLab

- Planned executions: **bmc** Control-M

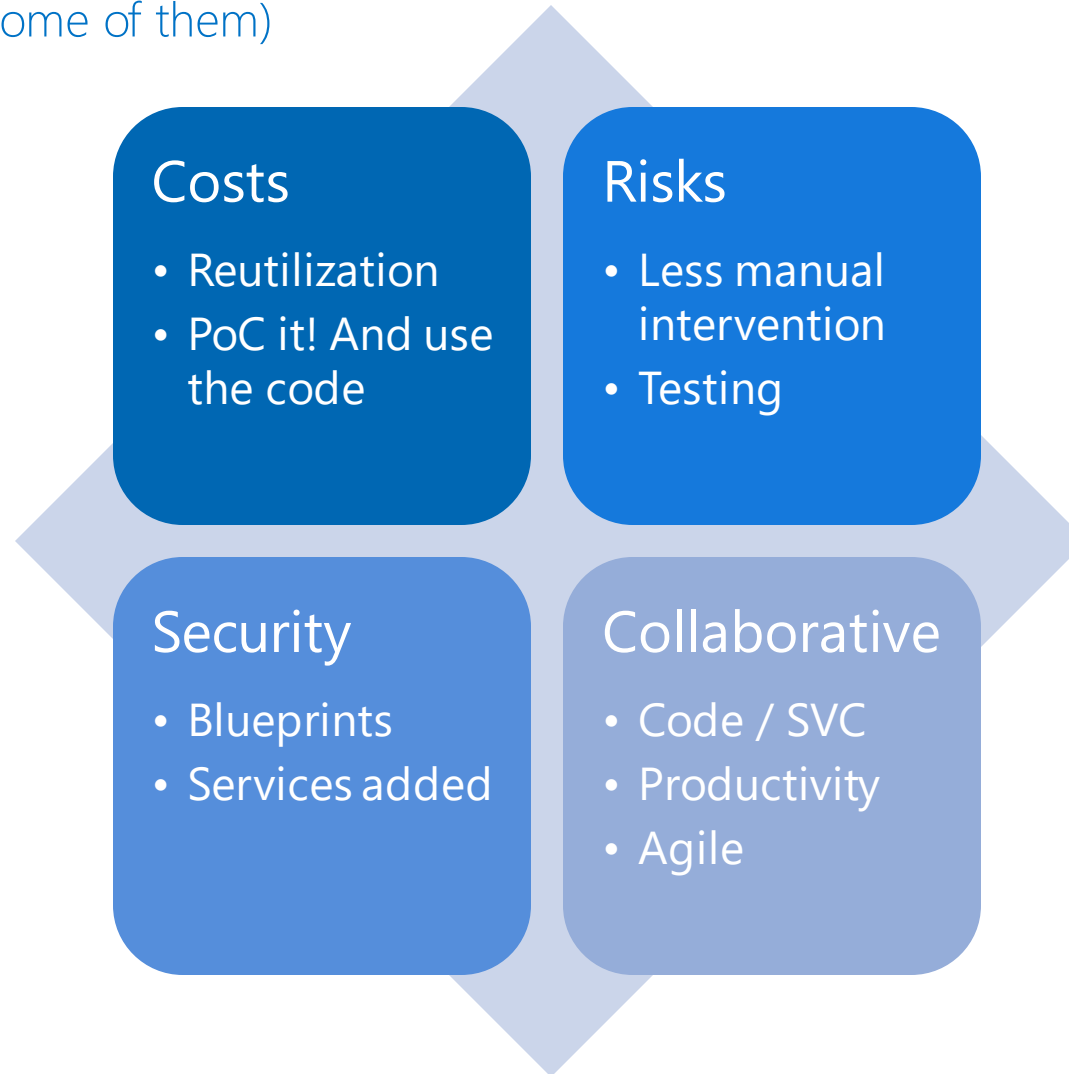# Automation

What is (was?) infrastructure automation really?

Image credits: Jobsintown.de

Introducing … IaC (**I**nfra **a**s **C**ode)

The benefits (just some of them)

Costs
- Reutilization
- PoC it! And use the code

Risks
- Less manual intervention
- Testing

Security
- Blueprints
- Services added

Collaborative
- Code / SVC
- Productivity
- Agile

Introducing … IaC (Infra as Code)

The REAL reason

Introducing … IaC (Infra as Code)

- How to start?
- What "tool" to choose?*
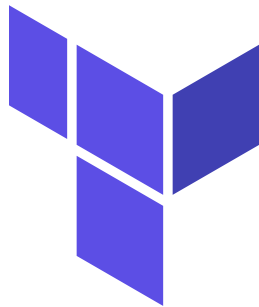
*But it's just a suggestion to start playing with IaC! I don't work for HashiCorp!!

Introducing ... IaC (Infra as Code)

- Terraform benefits (also in other tools*):
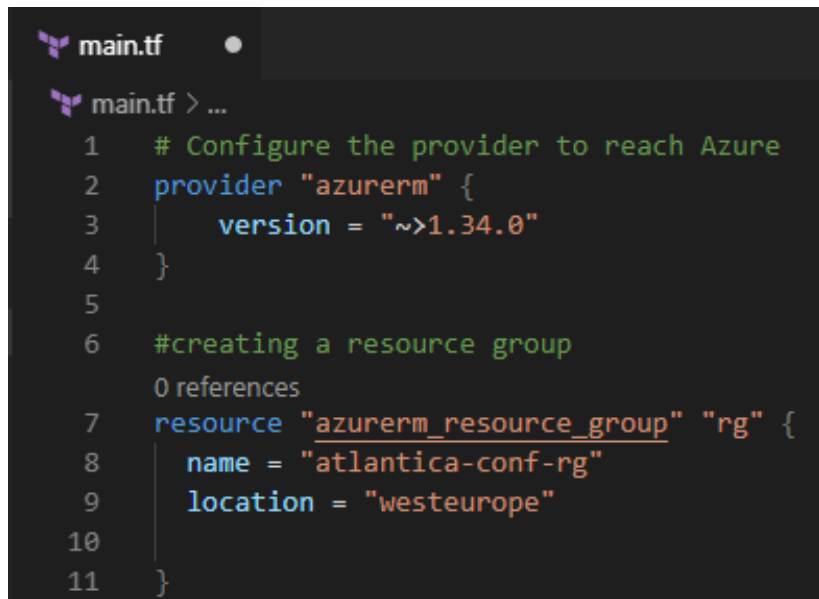
  - Works for the (multi) cloud

  - As well as for "on-prem" (virtualization)

Introducing ... IaC (Infra as Code)

- ## Terraform benefits (also in other tools):
  - Declarative

```
main.tf  ●
main.tf > ...
1    # Configure the provider to reach Azure
2    provider "azurerm" {
3        version = "~>1.34.0"
4    }
5
6    #creating a resource group
     0 references
7    resource "azurerm_resource_group" "rg" {
8      name = "atlantica-conf-rg"
9      location = "westeurope"
10
11   }
```

  - Idempotent

    Idempotence (UK: /ˌɪdɛmˈpoʊtəns/,[1] US: /ˌaɪdəm-/)[2] is the property of certain operations in mathematics and computer science whereby they can be applied multiple times without changing the result beyond the initial application. The concept of idempotence arises in a number of places in abstract algebra (in particular, in the theory of projectors and closure operators) and functional programming (in which it is connected to the property of referential transparency).

Introducing … IaC (Infra as Code)

- ## Terraform benefits (also in other tools):
  - Interpolation
  - Execution plans
  - Resource Graph

Introducing ... IaC (**I**nfra **a**s **C**ode)

Pleaaaaaaaaase stop talking and ...

show me the ~~money~~ DEMO

THANK YOU!!!