|  |  |
| --- | --- |
|  |  |

**Content**

[1. Virago project 3](#_Toc5201087)

[1. CloudWatch – General 3](#_Toc5201088)

[2. Push snapshot metrics to CloudWatch? ? 3](#_Toc5201089)

[3. CW – Event Bus 3](#_Toc5201090)

[4. CloudWatch Rules ! 3](#_Toc5201091)

[5. CW Agents 3](#_Toc5201092)

[6. Difference between DC and DCGW 3](#_Toc5201093)

[7. Active Directory 4](#_Toc5201094)

[8. CloudTrail Events 4](#_Toc5201095)

[9. Secrets Manager 4](#_Toc5201096)

[10. Prometheus 4](#_Toc5201097)

[2. Terraform 5](#_Toc5201098)

[1. Docker 5](#_Toc5201099)

[2. Terraform Deep-Dive 7](#_Toc5201100)

[2.2.1. Interpolation Syntax 7](#_Toc5201101)

[3. Wine 7](#_Toc5201102)

[3. 8](#_Toc5201103)

[1. PacBot ! 8](#_Toc5201104)

[2. Pupped vs Chef vs Ansible vs Saltstack – Configuration management tools 8](#_Toc5201105)

[4. Elasticsearch 9](#_Toc5201106)

[1. ElasticSearch concepts 9](#_Toc5201107)

[4.1.1. Aggregation – Zoskupenie – Osszesites, halmaz, egyesules 11](#_Toc5201108)

[4.1.2. Round-robin 12](#_Toc5201109)

[4.1.3. Shard and partition 12](#_Toc5201110)

[2. Full-Text Search file (FTS file) 12](#_Toc5201111)

[3. Elasticsearch clusters 12](#_Toc5201112)

[4. X-Pack 12](#_Toc5201113)

[5. Bonus: Angular + Node + Elasticsearch + Sequelizer ORM + PostgreSQL 12](#_Toc5201114)

[5. Others 15](#_Toc5201115)

[1. Baseline – GH 15](#_Toc5201116)

[2. Nico 15](#_Toc5201117)

[3. Pupetteer 15](#_Toc5201118)

[5.3.1. Expression – animation tool 15](#_Toc5201119)

[5.3.2. Linux -.- 15](#_Toc5201120)

[6. Service Now 16](#_Toc5201121)

1. Virago project
   1. CloudWatch – General
   2. Push snapshot metrics to CloudWatch? ?
   3. CW – Event Bus

You can set up your AWS account to send events to other AWS accounts, or to receive events from other accounts. This can be useful if the accounts belong to the same organization, or belong to organizations that are partners or have a similar relationship.

* 1. CloudWatch Rules !
* **Events**—An event indicates a change in your AWS environment. AWS resources can generate events when their state changes. For example, Amazon EC2 generates an event when the state of an EC2 instance changes from pending to running, and Amazon EC2 Auto Scaling generates events when it launches or terminates instances. AWS CloudTrail publishes events when you make API calls. You can generate custom application-level events and publish them to CloudWatch Events. You can also set up scheduled events that are generated on a periodic basis. For a list of services that generate events, and sample events from each service, see CloudWatch Events Event Examples From Supported Services.
* **Targets**—A target processes events. Targets can include Amazon EC2 instances, AWS Lambda functions, Kinesis streams, Amazon ECS tasks, Step Functions state machines, Amazon SNS topics, Amazon SQS queues, and built-in targets. A target receives events in JSON format.
* **Rules !** — *Rules route events from your AWS resources for processing by selected targets*

A rule matches incoming events and routes them to targets for processing. A single rule can route to multiple targets, all of which are processed in parallel. Rules are not processed in a particular order. This enables different parts of an organization to look for and process the events that are of interest to them. A rule can customize the JSON sent to the target, by passing only certain parts or by overwriting it with a constant.

* 1. CW Agents
  2. Difference between DC and DCGW

**AWS Direct Connect** is a service aimed at allowing enterprise customers easy access to their AWS environment. Enterprises can leverage the AWS Direct Connect to establish private connectivity to the AWS global network from their data centers, office locations or co-location environments. AWS Direct Connect supports two bandwidth levels: 1 G and 10 G. higher bandwidth levels can be provisioned by having multiple 10 G interfaces connected in tandem. Lower bandwidth levels of 50 M, 100 M, 200 M, 300 M, 400 M and 500 M can only be provisioned through an AWS partner supporting AWS Direct Connect. There are two aspects of Direct Connect pricing: the per hour port fee and the data transfer pricing. Port fees depend on the port speed selected.

**AWS Direct Connect gateway** is aimed at making it easier to connect from a single Direct Connect location to multiple AWS regions or VPCs.

* 1. Active Directory
  2. CloudTrail Events
  3. Secrets Manager
  4. Prometheus

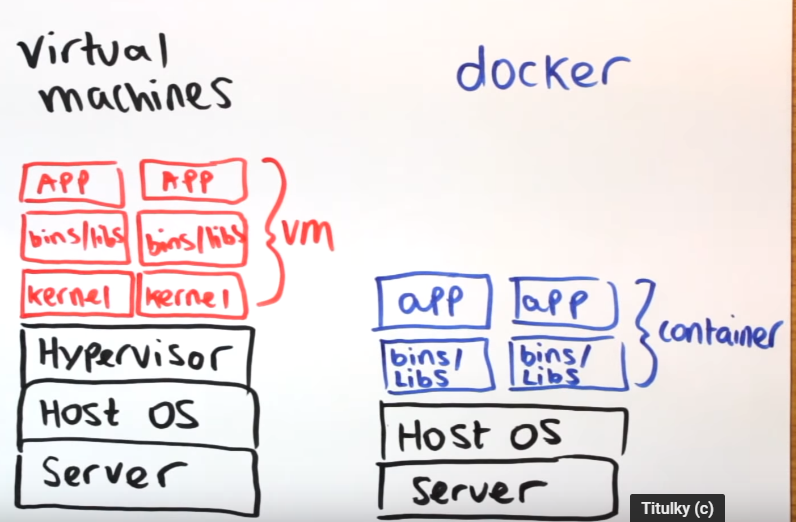
1. Terraform

* Tool for manging infrastructure
* Version contron, multi platform(AWS, AZURE, GCD, OpenStack)
* There is an Enterprise Version
* **Idempontent** – will not deploy the same resource once it is deployed
* **Cloud-agnostic -** Agnostic, in an information technology (IT) context, refers to something that is generalized so that it is interoperable among various systems.
  1. Docker

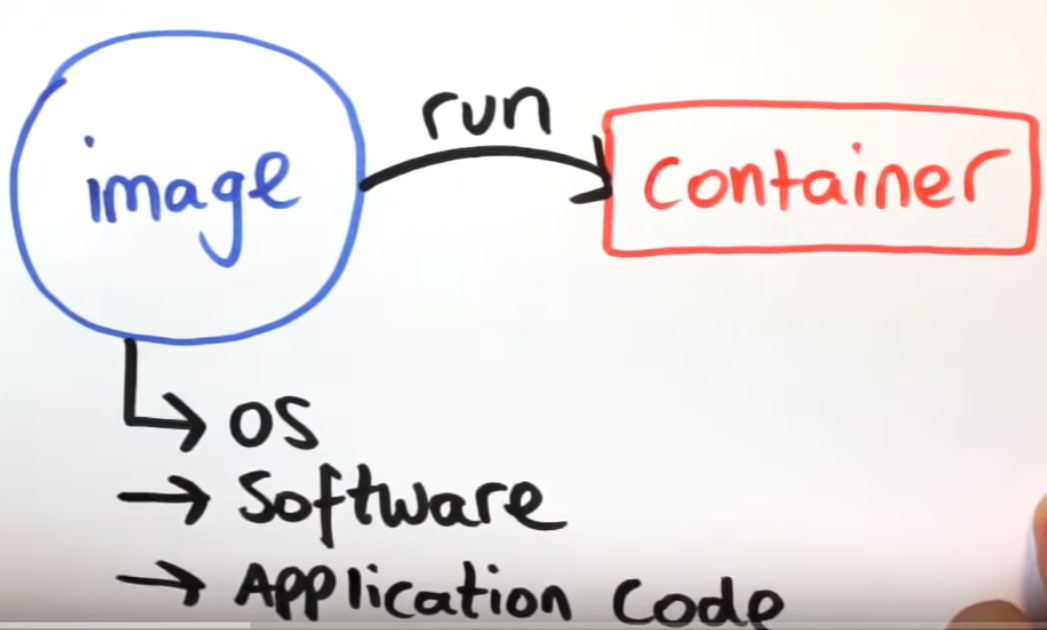
Advantages:

* Your app always runs in exactly the same invirnoment so you dont get inconsistencies in how it behaves, if it works on your computer it owrks on every computer, it works on a live server, it always acts the same.
* Sandbox Projects – eliminates the conflicts between different projects
* It is easy to get going with someboy else’s project, you dont have to install all of tools and dependencies that the project needs, you just spin up the VM, you call inside and it works.

Container – not a full VM, but htere is the code wrapped up.

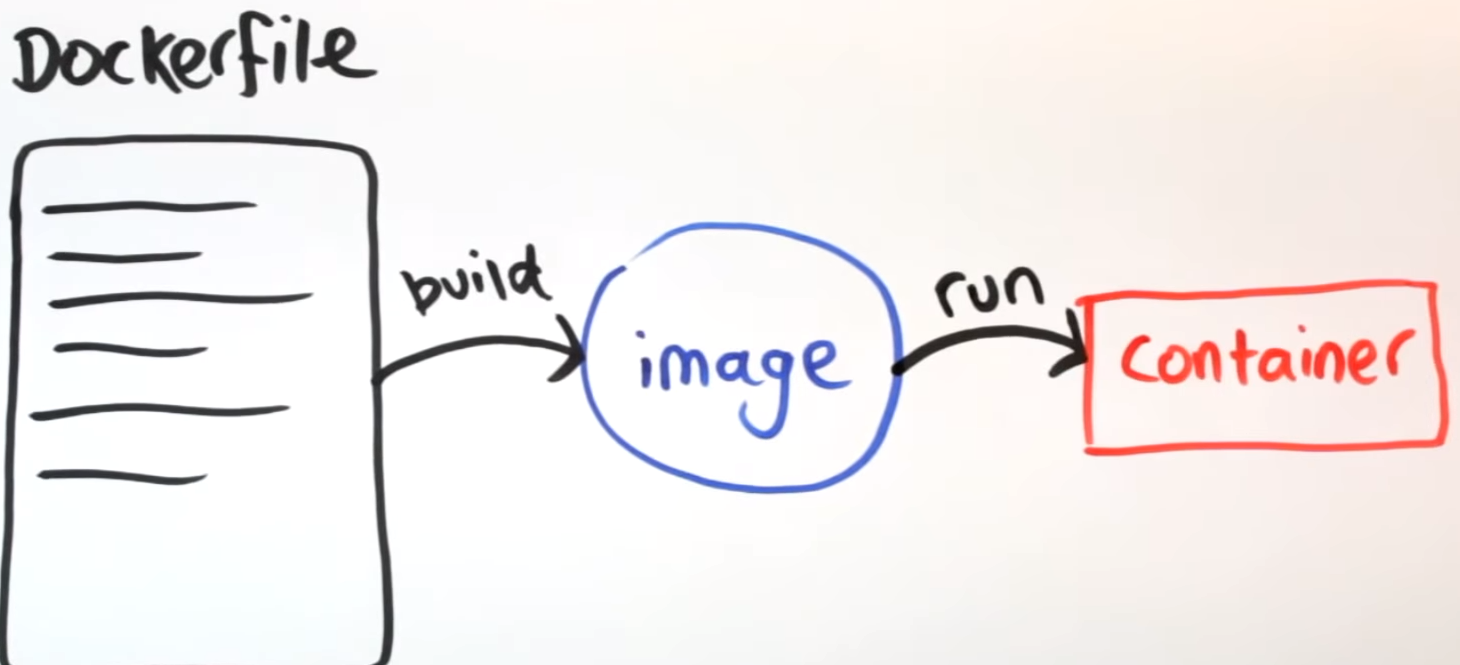


Container is a running instance of an image, and the image is a template for creating the environment you wanted to snapshot of the system at a particular time So its got the OS, the software the application code all bundled up in a file.



Images are defined with a Dockerfile, with a list of steps to perform to create that image (e.g. condigure the OP install the software you need copy the project files into the right places etc..).

So you wirte a dockerfile than you build that and you get an image which you can run to get containers .



Docker: if you build the image and run it and later change the src and mbe you refresh the localhost it wil not work because when we built the image it made a copy of that file to see the change you need to rebuild the image and spin up a new container from the updated image. During development this is obviously a massive pain and tihs is where **volumes** come in. So there are 2 types of volumes:

1. Want to persist and share data between containers
2. Lets you share folder between the host and the container, lets you mount a local directory on your computer as a volume inside of the container. Then the container when its running well be able to see the files were working on.

*sudo docker run -p 80:80 hello-world -v /home/noro/Desktop/docker/src/:/var/www/html*

* 1. Terraform Deep-Dive

Terraform is a tool for building, changing, and versioning infrastructure safely and efficiently. Terraform can manage existing and popular service providers as well as custom in-house solutions.

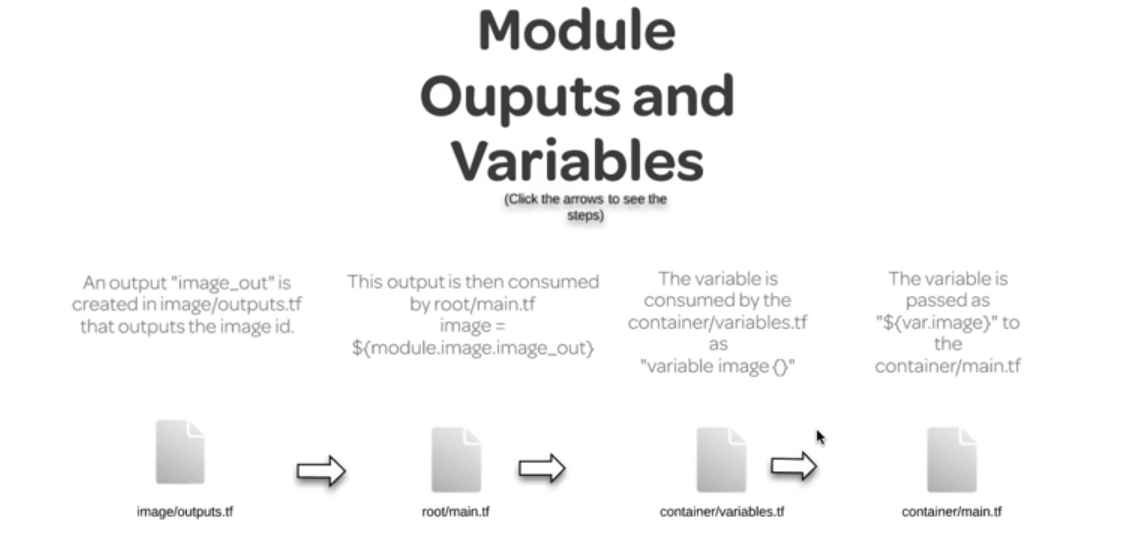
* Infrastructure as Code
* Execution Plans
* Resource Graph
* Change Automation

1. *terraform init*
2. *terraform plan*
3. *terraform apply*
4. *terraform show*

**Interpolation (interpolation syntax)** – how to reference a variable from within you infrastructure

**Ghost** is a free and open source blogging platform written in JavaScript and distributed under the MIT License, designed to simplify the process of online publishing for individual bloggers as well as online publications.

1. *terraform taint* (mark the resource to redeploy)
2. *terraform untaint ‚resource:name‘*
3. *terraform validate*
4. *terraform graph*
5. *terraform console*

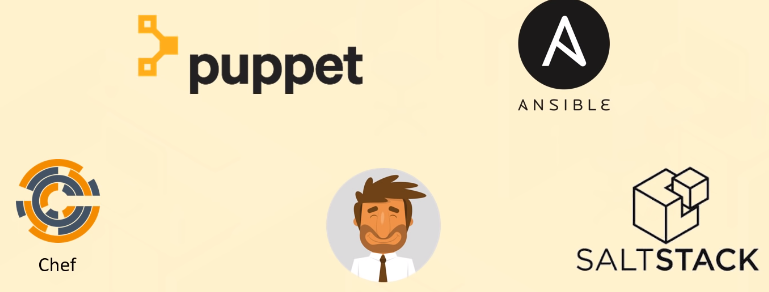


Maps and Lookups (map is in general a key value pair of a series of it)

* 1. Wine

Wine (recursive backronym for Wine Is Not an Emulator) is a free and open-source compatibility layer that aims to allow computer programs (application software and computer games) developed for Microsoft Windows to run on Unix-like operating systems. Wine also provides a software library, known as Winelib, against which developers can compile Windows applications to help port them to Unix-like systems.

1. Configuration management
   1. PacBot
   2. Pupped vs Chef vs Ansible vs Saltstack – Configuration management tools



**Puppet**: Heterogenous env., Ruby, DSL,recipes, cookbooks

**Chef**:Heterogenous env., Ruby, DSL, master/slave, SSL

**Ansible:** push, YAML, get infrastructure up really fast, playbook, inventory, SSL,no agent needed

Update from GH: focus on ansible (linux) however on windows it can be painful.

1. Elasticsearch

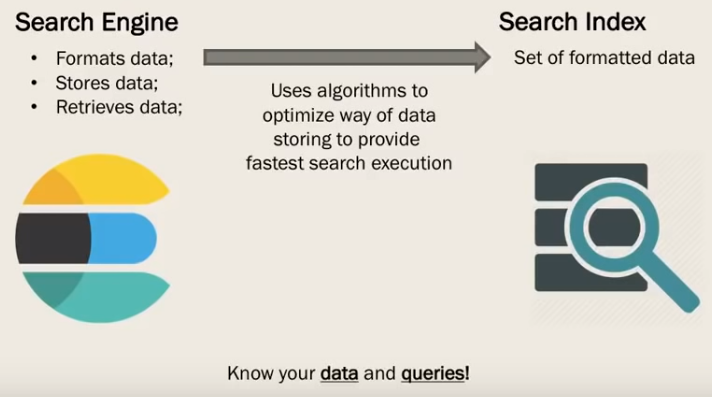
Elasticsearch is a search engine based on the **Lucene** library. It provides a **distributed**, **multitenant-capable** **full-text search engine** with an **HTTP web interface** and schema-free **JSON** documents.

Java, .NET (C#), PHP, Python, Apache Groovy, Ruby and many other languages. According to the DB-Engines ranking, Elasticsearch is the most popular enterprise search engine followed by **Apache Solr**, also based on Lucene.

Elasticsearch is developed alongside a data collection and **log-parsing engine called** **Logstash**, and an **analytics and visualisation platform called Kibana**. The three products are designed for use as an integrated solution, referred to as the "Elastic Stack" (formerly the "ELK stack").

Possible problems w SQL:

* Not Optimized Operators
* JOINs
* No way to query native FTS



* 1. ElasticSearch concepts

**• MySQL => Databases => Tables => Columns/Rows**

**• Elasticsearch => Indices => Types => Documents with Properties**

* **Document :** If you use’d to thinking of things in terms of databases, the coument is a lot like a row and database it represensts a given entity, something that you are searching for. In ES its not just about text, any structured data can work. ES works on top of JSON formatted data (JSON is bassically just a way of encoding structured data that may contain string or numbers of dates or what have you in way you can actually transmit it accross the web cleanly)
* **Types (RDS table):** defines the schema and mapping shared by a bunch of documents that represent the same sort of thing.(e.g.log entry, encyclopedia article, etc..) I might define a mapping that says an Apache access log contains things like a request URL, status code, request time .**.** [ES is movin toward eliminating the concept of types. In ES6, only one type is allowed per index.
* **Indices : Index** is like a **database** in RDS.An index powers search into all documents within a collection of types. They contain inverted indices that let you search accross everything within them at once. An index is a logical namespace which maps to one or more primary shards and can have zero or more replica shards.

**An index is split into shards.** **Documents** are **hashed** to a particular **shard**. Each shard may be on different **node** in a **cluster**. Every shard is a self-contained **Lucene** **index** of its own. We distribute our index to many differenct shards, which can live on different servers in the cluster.

**Primary and replica shards.** Your application should round-robin requests among nodes. Nodei s basically an installation of elasticsearch , usually 1 node / physical server in the cluster. If any given node of the cluster goes down, you wont even see it as an end user.

**Primary shard:** the primary copies of my index data and thats where write requests are gonna be routed to initally, that data will then be replicated to the replicas shards, which can alsho handle read requests whenever we want

The number of primary shards cannot be changed later, but you can add more read replicat shards for more read throughput.

* **TF-IDF** – **Term Frequency** \* **Inverse Document Frequency**
* **TF** is how often a term appears in a given document
* **Document Frequency** is how often a term appears on all douments
* **Term Frequency** / **Document Frequency** measures the **relevance** of a term in a document

How to use an index in ES?

1. RESTful API – it’s just like how you request the web page from a web server from a web browser in your desktop. So when you are requesting a webpage on your browser whats happening is that your browser is sending a REST request to aweb server somewhere and every rest request has a verbl ike get or put or post and some sort of body that specifies what it is that you want ot get back. So if you are looking for a webpage you would send a get verb and then that get would request a specific URL that you want to retreive from that web server.
2. Client API’s – if you are accessing ES from a web application you are writing often there will be clien API that provbide a level of abstraction on top of those rest queries. So instead i think on how do i cosntruct the right json format for the type of search that I want or In general w Elasticsearch can talk any language which is capable to send rest queries.
3. Analytic tools – higher level tools, like Kibana. Web-based graphical UI’s such as Kibana let you interact with your indices and explore them without writing code.

What purpose do inverted indices serve? – they quickly map search terms to documents.

5 primary and 3 replica shards. How many in total? -> 20

ES – not only for FTS of documents. Can index any kind of scrutured data nay kind of mapping you can deam up. Its not jsut for searching encyclopedias and websites and blogs, it can also be used for searching and even aggregating and visualizing numerical data or time-based data. Aggregating web logs from web servers or something like Google Analytics.

* + 1. Aggregation – Zoskupenie – Osszesites, halmaz, egyesules
* the formation of a number of things into a cluster.

*"a single dose of aspirin irreversibly inhibits the normal aggregation of platelets"*

* a cluster of things that have come or been brought together.

*"the pelican is the other bird likely to be found in large aggregations in East Africa's wetlands"*

* (on the Internet) the collection of related items of content so that they can be displayed or linked to.

*"a desktop aggregation app that brings together Facebook, Twitter, and LinkedIn"*

* + 1. Round-robin

*A tournament in which each competitor plays in turn againts every other.*

The name of the algorithm comes from the round-robin principle known from other fields, where each person takes an equal share of something in turn.

* + 1. Shard and partition

A **partition** is a division of a logical database or its constituent elements into distinct independent parts. Database partitioning is normally done for manageability, performance or availability reasons, or for load balancing.

A database shard is a horizontal partition of data in a database or search engine. Each individual partition is referred to as a shard or database shard. Each shard is held on a separate database server instance, to spread load.

Some data within a database remains present in all shards, but some appears only in a single shard. Each shard (or server) acts as the single source for this subset of data.

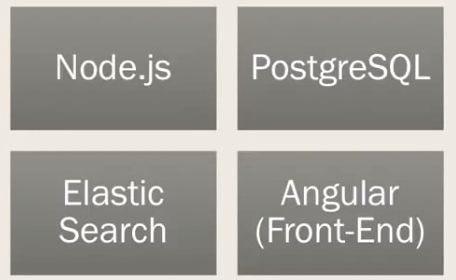
* 1. Full-Text Search file (FTS file)

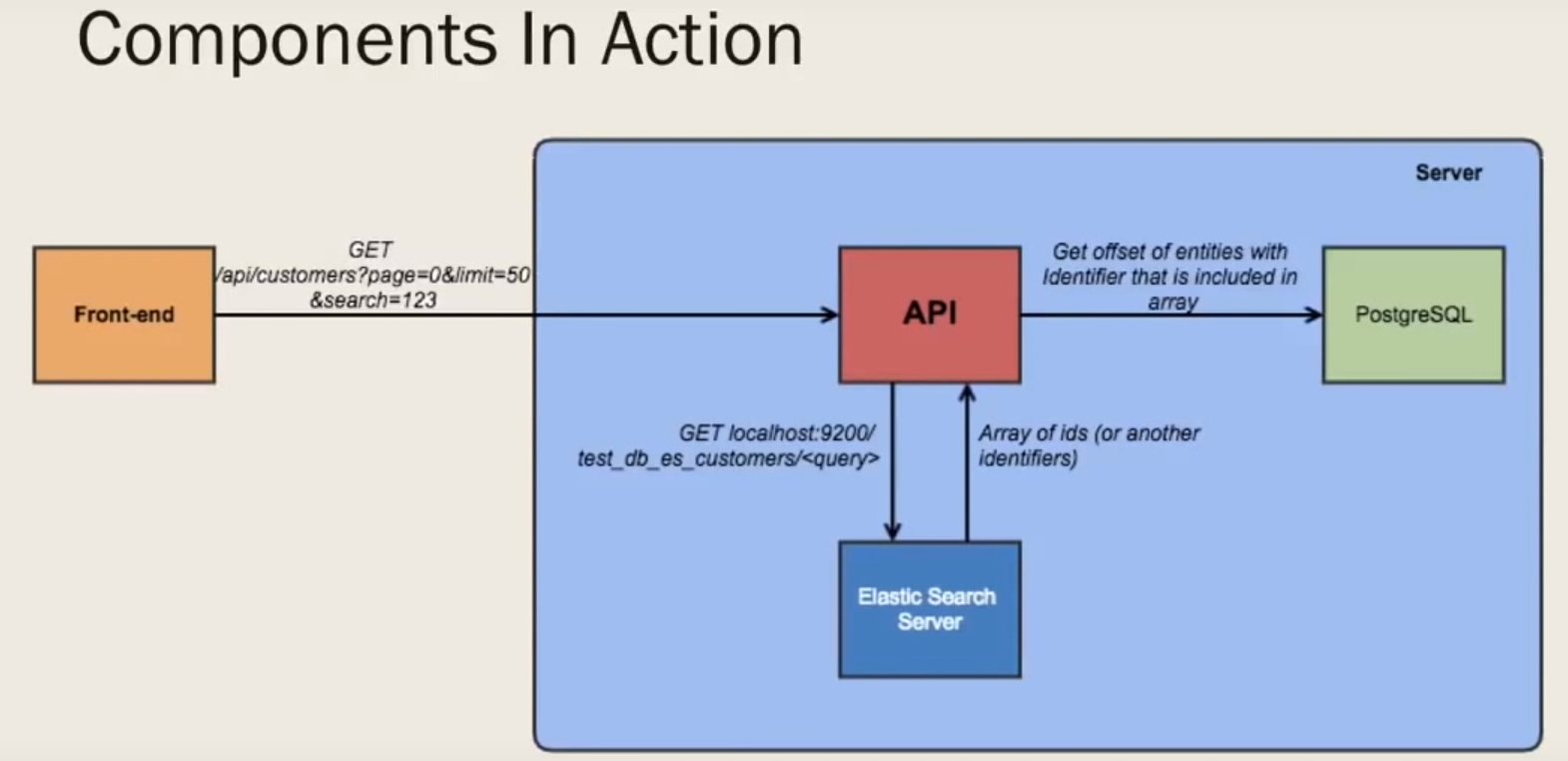
In text retrieval, full-text search refers to techniques for searching a single computer-stored document or a collection in a full-text database. Full-text search is distinguished from searches based on metadata or on parts of the original texts represented in databases (such as titles, abstracts, selected sections, or bibliographical references).

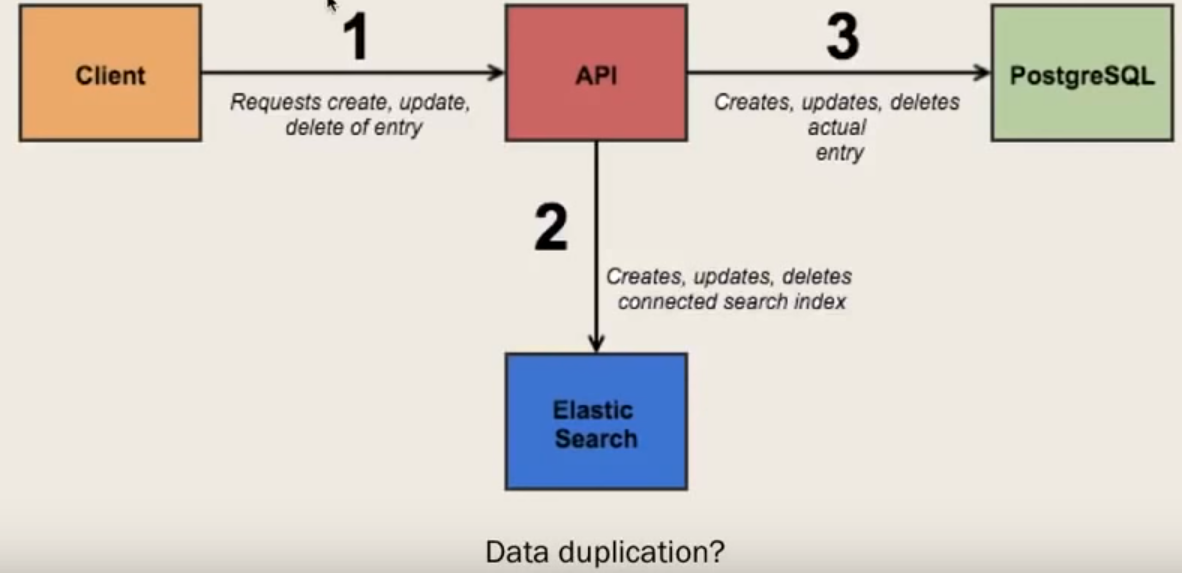
* 1. Elasticsearch clusters
  2. X-Pack

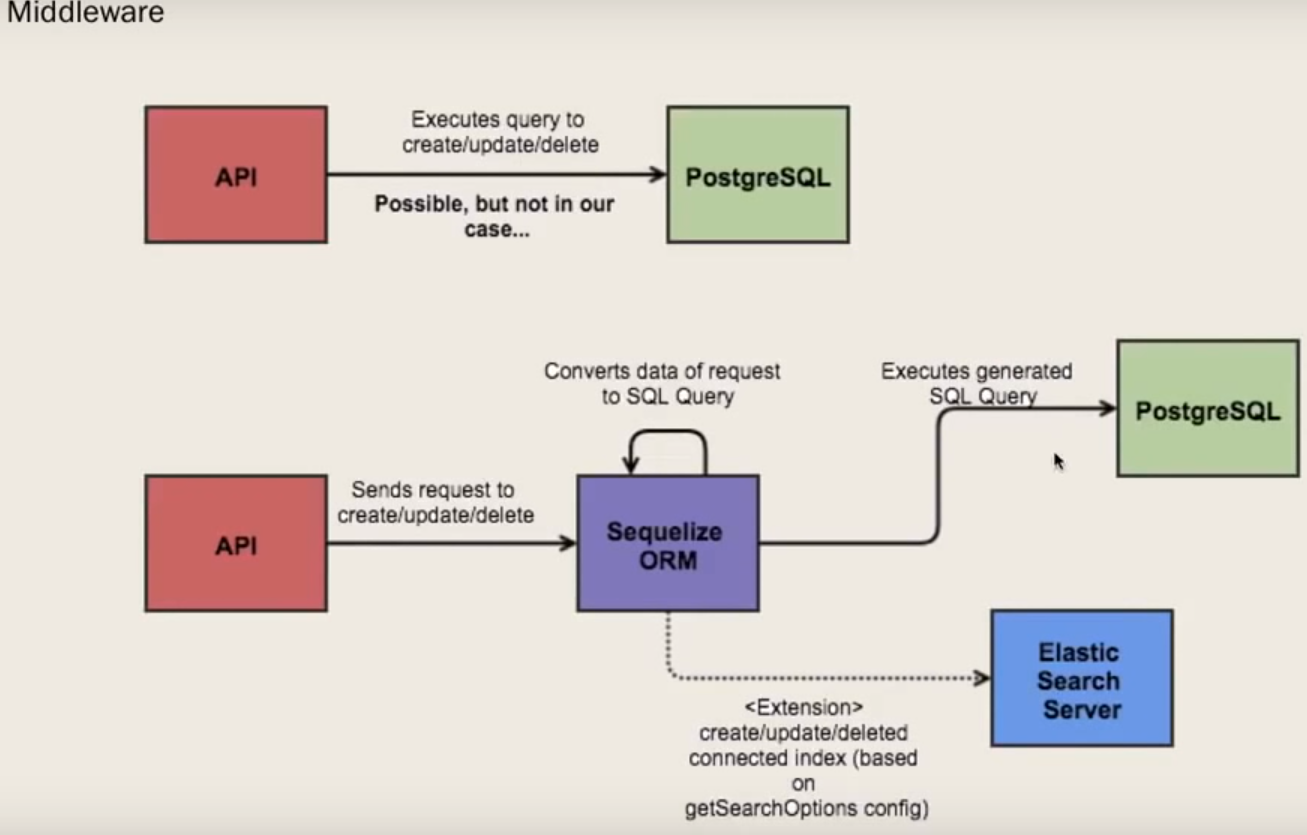
Security, Alerting, Monitoring, Reporting, Machine Learning, Graph Exploration

* 1. Bonus: Angular + Node + Elasticsearch + Sequelizer ORM + PostgreSQL









1. Others
   1. Baseline – GH

Master accc, step fucntions , prod provision baseline prod, change the input parameer then run it . thats all . to new execution copy the json structure.

* 1. Nico
* Active directory
* Azure act dir integration w AWS
* AWS Landing Zone
  1. Pupetteer
* **pb** stands for **protobuf**. In TensorFlow, the protbuf file contains the graph definition as well as the weights of the model. Thus, a pb file is all you need to be able to run a given trained model.
* [oliver.knape@telekom.de](mailto:oliver.knape@telekom.de)
  + 1. Expression – animation tool
    2. Linux -.-

Sudo systemctl daemon-reload

1. Service Now