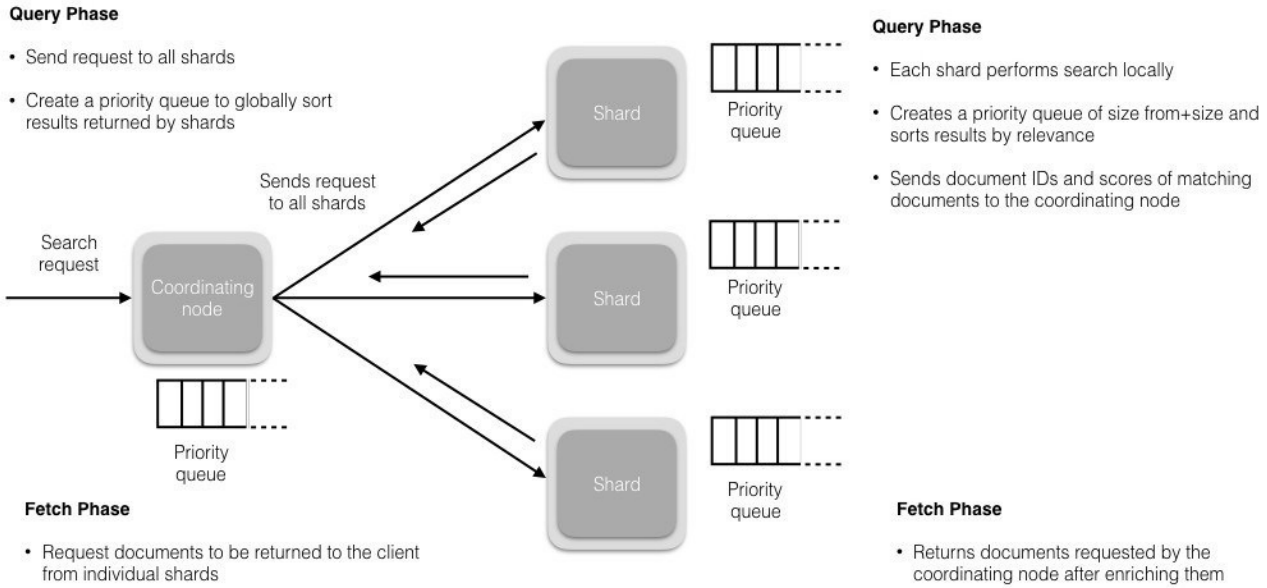
TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI
HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

Elasticsearch & Kibana

Thanh-Chung Dao
BKC Group
School of Information and Communication Technology
(https://bkc-group.github.io)

Elasticsearch

- Full-text search engine
- Based on the Lucene library
- HTTP web interface and schema-free JSON documents



Lab: Elasticsearch and Kibana

- Set up Elasticsearch to store data
- Write/Read data to Elasticsearch
- Install and run Kibana

Installation

- Install Docker and login
  - https://docs.docker.com/docker-for-windows/install/
  - https://docs.docker.com/docker-for-mac/install/
- Login to @chung-pi gitlab to pull images
  - docker login registry.gitlab.com -u bi-class -p bqp_cSsCJ2kaNjMu1U4A
- Pull images
  - docker pull registry.gitlab.com/chung-pi/bi-docker/elasticsearch
  - docker pull registry.gitlab.com/chung-pi/bi-docker/kibana:latest
- If Internet is not available
  - docker load --input elasticsearch.tar
  - docker load --input kibana.tar

Start Elasticsearch and Kibana

- Clone bi-class git project
  - https://gitlab.com/chung-pi/bi-class
- Start containers using docker-compose
  - docker-compose up -d --build elasticsearch
  - docker-compose up -d --build kibana

GUI

- Elasticsearch
  - http://localhost:9200
- Kibana
  - http://localhost:5601

Load data to Elasticsearch

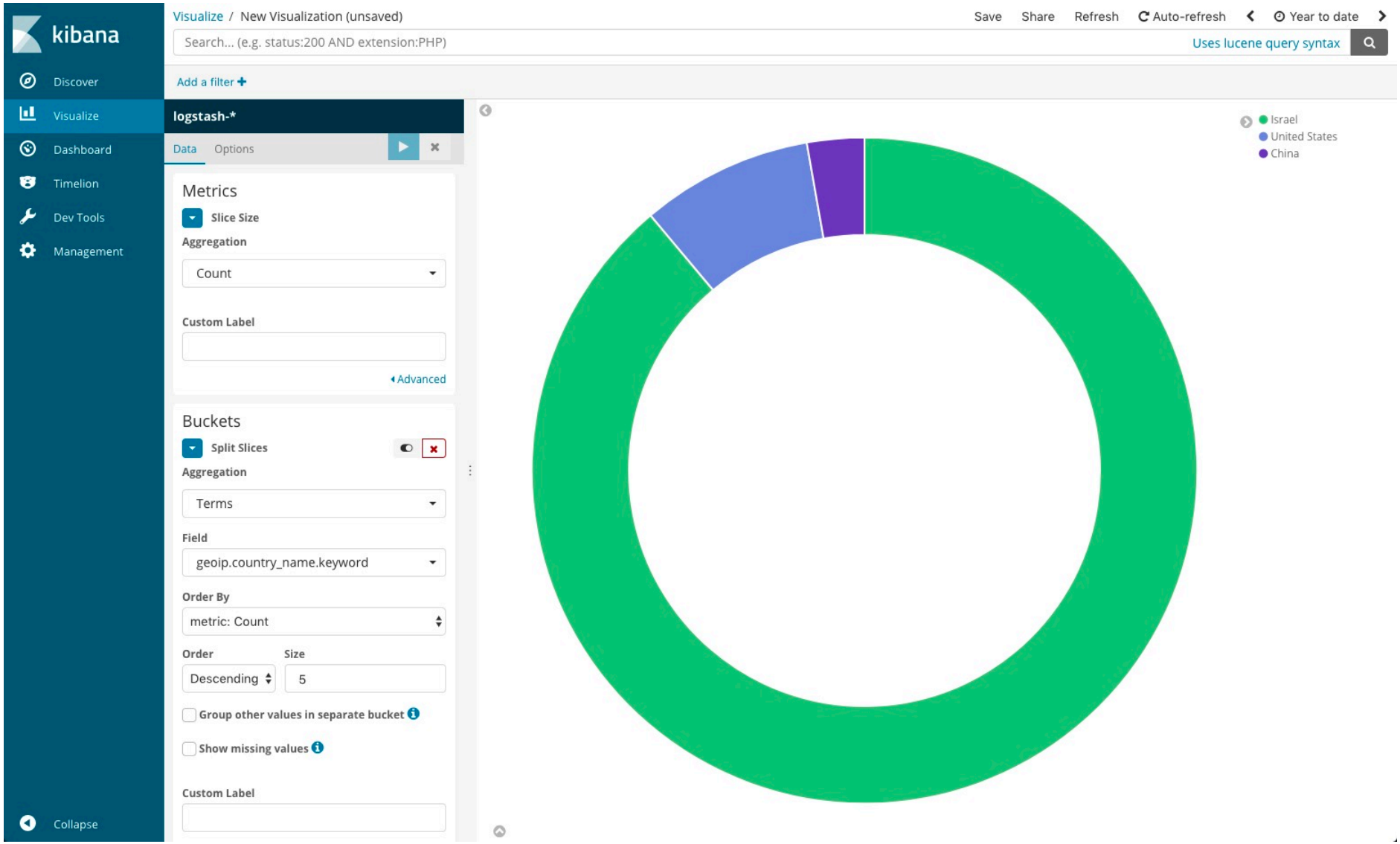
- Using CURL
  - curl -O https://download.elastic.co/demos/kibana/gettingstarted/7.x/accounts.zip
  - unzip accounts.zip
  - curl -H 'Content-Type: application/x-ndjson' -XPOST 'localhost:9200/bank/account/_bulk?pretty' --data-binary @accounts.json
- Checking data using Kibana
  - Open Kibana on browser

Understanding Kibana aggregations

- There are two types of aggregations
  - Bucket aggregations groups documents together in one bucket according to your logic and requirements
  - Metric aggregations are used to calculate a value for each bucket based on the documents inside the bucket

Metric aggregations	Bucket aggregations
• Count • Sum • Average • Media • Min • Max • Unique Count • Standard Deviation • Percentiles • Percentile Ranks	• Date Histogram • Date Range • Filters • Histogram • IPv4 Range • Range • Terms • Significant Terms • Geohash

Example


kibana
Visualize / New Visualization (unsaved)
Save
Share
Refresh
Auto-refresh
Year to date
Search... (e.g. status:200 AND extension:PHP)
Uses lucene query syntax
Add a filter
Discover
Visualize
Dashboard
Timelion
Dev Tools
Management
Collapse
logstash-*
Data
Options
Metrics
Slice Size
Aggregation
Count
Custom Label
Advanced
Buckets
Split Slices
Aggregation
Terms
Field
geoip.country_name.keyword
Order By
metric: Count
Order
Descending
Size
5
Group other values in separate bucket
Show missing values
Custom Label
Israel
United States
China

Basic Charts

	For visualizing time series data and for splitting lines on fields	Users over time
	For showing statistical outliers and are often used for latency values	Latency and outliers
	Good for showing relationships between two fields	URL and referrer
	are a simple way to show time series and are good for splitting lines to show anomalies	Average CPU over time by host
	Useful for displaying parts of a whole	Top 5 memory consuming system procs
	Great for time series data and for splitting lines across fields	URLs over time

Data

	Best way to split across multiple fields in a custom way	Top user, host, pod, container by usage
	A way to show the status of a specific metric using thresholds you define	Memory consumption limits
	Similar to a Gauge, useful for monitoring a specific metric defined as a goal	No. of errors per service
	Useful visualization for displaying a calculation as a single number	No. of Docker containers run.

Map, Time series, and Others

	Help add a geographical dimension to IP-based logs	Geographic origin of web server requests.
	Allows you to create more advanced queries based on time series data	Percentage of 500 errors over time
	**Experimental** - Allows you to create selectors or sliders for alternating between options.	Switch between
	A great way to add a customized text or image based visualization to your dashboard based on markdown syntax	Company logo or a description of a dashboard
	Helps display groups of words sized by their importance	Countries sending requests to a web server
	**Experimental** - allows you to add custom visualizations based on Vega and VegaLite	-

Other tutorials

- https://logz.io/blog/kibana-tutorial/
- https://logz.io/blog/kibana-tutorial-2/