



Berner Fachhochschule  
Haute école spécialisée bernoise  
Bern University of Applied Sciences

# Informatikseminar

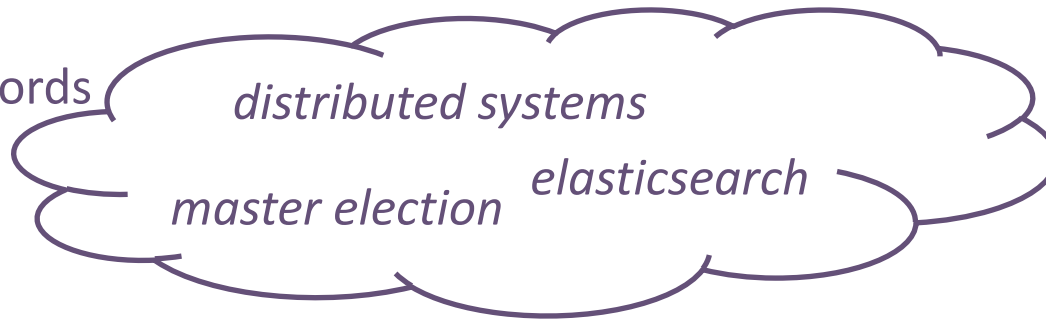
## Projects

Dr. Mascha Kurpicz-Briki

FS 2019

# Project 1 – Master Election Algorithms

Keywords



elasticsearch

Distributed systems often use a master (or leader) node for coordination. In order to determine a new master, whenever the current master is not available anymore, election algorithms are used.

One example of such a distributed system is Elasticsearch. Elasticsearch is a distributed search engine that by default uses Zen discovery and master election. Explain how this master election works and how it differentiates from other common master election algorithms (e.g. Paxos). Describe best practices for master election configurations (e.g. quorum) in Elasticsearch.

<https://www.elastic.co/guide/en/elasticsearch/reference/current/modules-discovery-zen.html>

# Project 2 – Honey Pots

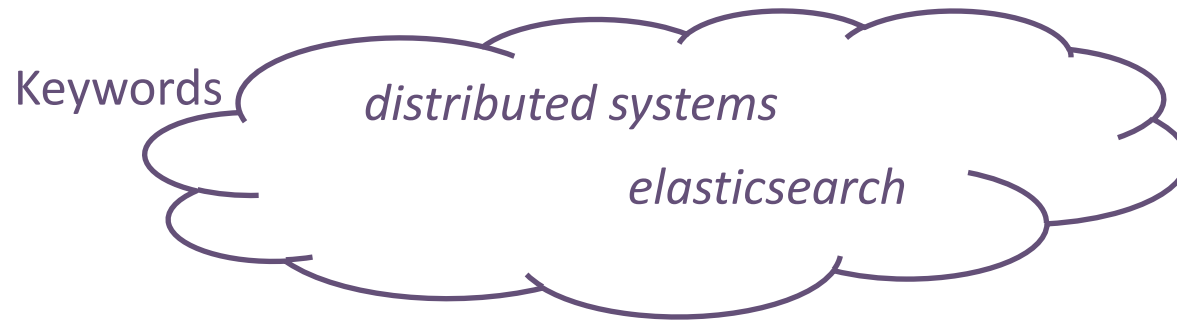


Honey pots in dedicated environments can be used to detect and study hacker attacks.

Explain how such honey pots can be implemented. What are the advantages and dangers of using such honey pots?

<https://www.sans.edu/cyber-research/security-laboratory/article/honeypots-guide>

# Project 3 – Elasticsearch Best Practices



Elasticsearch is a distributed and scalable search engine. The data stored in Elasticsearch is organized in indices and shards and distributed among the data nodes.

Explain how indices and shards are used in Elasticsearch. Describe what best practices need to be respected to build a scalable and efficient system. For example, describe how many shards should be used per index or what is an efficient size for a shard/index.

<https://www.elastic.co/guide/en/elasticsearch/reference/current/getting-started-concepts.html>

# Project 4 – Kubernetes

Keywords

*distributed systems*

*kubernetes*

*container platform*



Kubernetes is a platform for the orchestration of containers. In practice, it simplifies the administration by providing features like auto-upgrade of nodes, auto-scaling and auto-restart of failed resources.

Provide an overview of the architecture of a Kubernetes cluster. Explain the system components and how they interact. Discuss current best practices for the configuration of a Kubernetes cluster.

<https://kubernetes.io/>

<https://kubernetes.io/docs/concepts/overview/components/>

# Project 5 – NoSQL Databases



**cassandra**

Keywords

*NoSQL*

*data storage*

*databases*



To scale horizontally, typically NoSQL databases are used. The way the data is structured and stored is different from the traditional relational databases.

Explain how NoSQL databases differ from relational databases. Describe and compare different implementations of NoSQL databases used in practice.

<https://en.wikipedia.org/wiki/NoSQL>

# Project 6 – Scala vs. Java

Keywords



In this project you will compare the programming languages Java and Scala. Study the differences between the two languages and find use cases, where one or the other is typically used.

Try to find out about the popularity of the two languages, for example by using Google Trends.

<https://www.scala-lang.org/>

<https://trends.google.com/trends>

# Project 7 – Elasticsearch Mappings



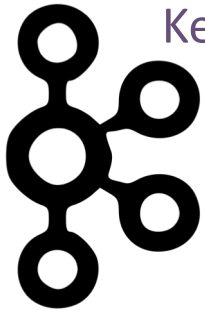
Elasticsearch is a distributed and scalable search engine. Mapping defines how the fields of a document stored in Elasticsearch are stored and indexed.

Explain what mappings are and how they are used in Elasticsearch. Describe what settings can be tuned to optimize the mapping of the data stored in Elasticsearch.

<https://www.elastic.co/guide/en/elasticsearch/reference/current/mapping.html>



# Project 8 - Apache Kafka vs. Redis



Keywords

*message queue apache kafka  
redis*



redis

Apache Kafka and Redis can both be used as a message queue. For example, they can be used to enhance the performance of an Elastic stack for log management.

Describe how Redis and Apache Kafka can be used as message queues and what are typical architectures and use cases. What are the advantages or drawbacks of using one or the other?

<https://redis.io/>

<https://aboullaite.me/using-redis-as-buffer-in-the-elk-stack/>

<https://kafka.apache.org/>

# Project 9 – Nature-Inspired Algorithms



Nature-inspired algorithms are designed based on principals and structures found in nature. This can be used for example for synchronizations in distributed systems or for optimization problems.

Read about nature-inspired algorithms and summarize your findings. Pick some of the algorithms you found and describe what they are used for.

<http://www.cleveralgorithms.com/nature-inspired/index.html>