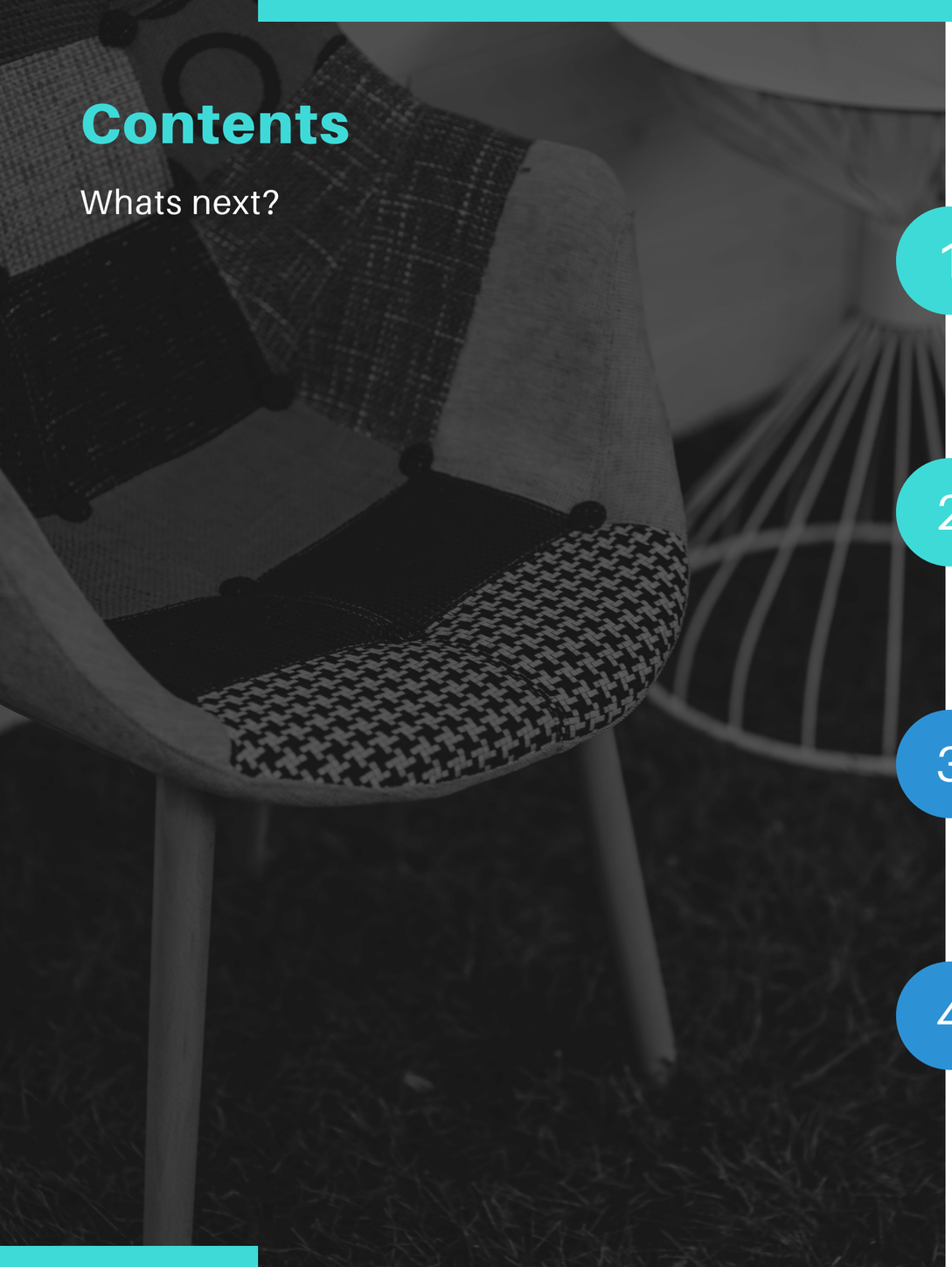
Custom Kubernetes Schduler

SINA SHARIATI

Amirkabir University Of Technology



Contents

Whats next?

1 WhoAmI

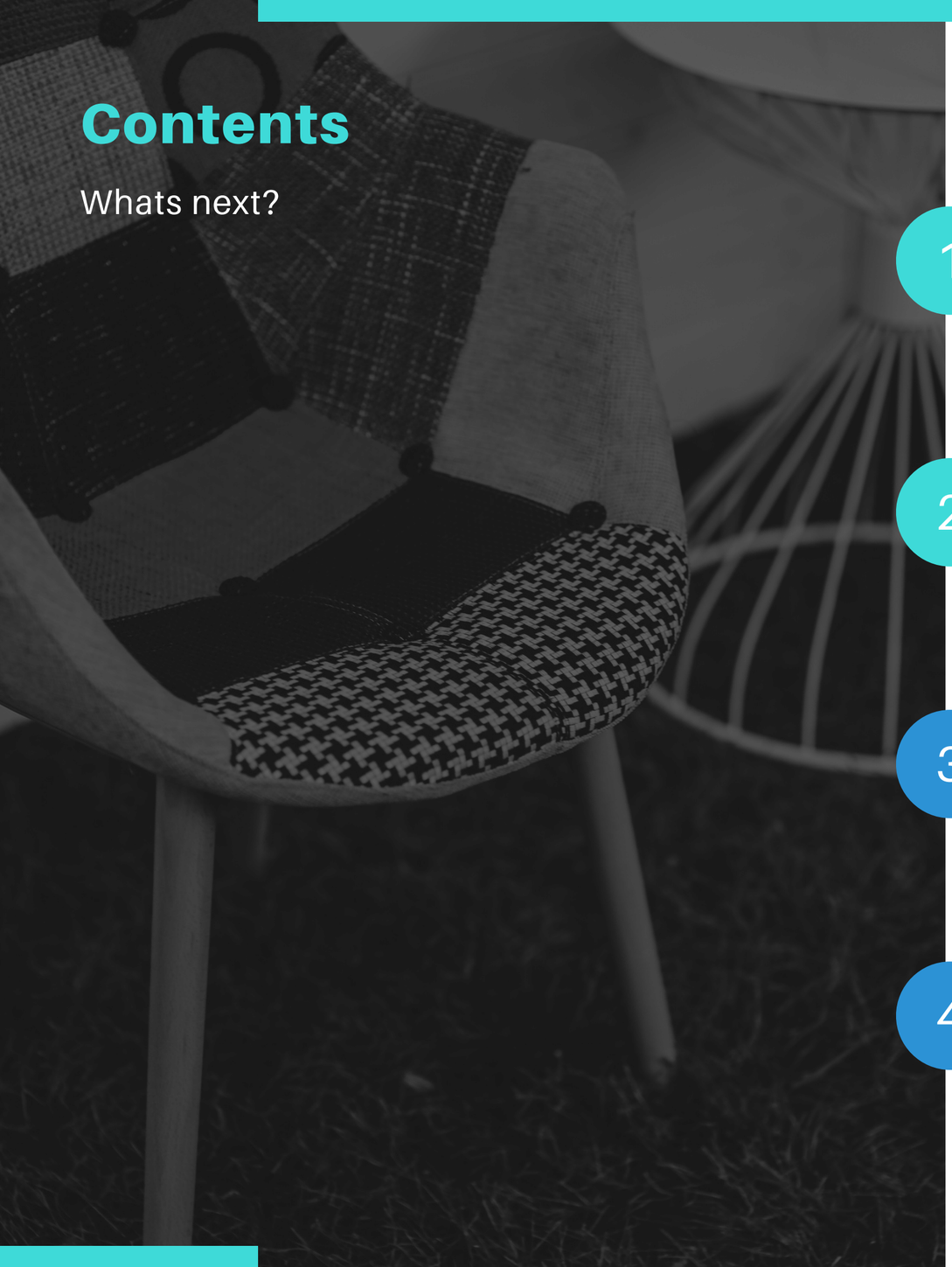
2 What?

3 Why?

4 Live Coding

5 Deployment & testing

6 Final Notes & QA



Contents

Whats next?

1 **WhoAmI**

2 What?

3 Why?

4 Live Coding

5 Deployment & testing

6 Final Notes & QA

$ whoami

SINA SHARIATI

- Currently, software engineer at snapp! - offering team

LINKS

github.com/noisyboy-9

@sina_shariati

sina.shariati@yahoo.com

sina-shariati

Noisyboy-9 - Overview

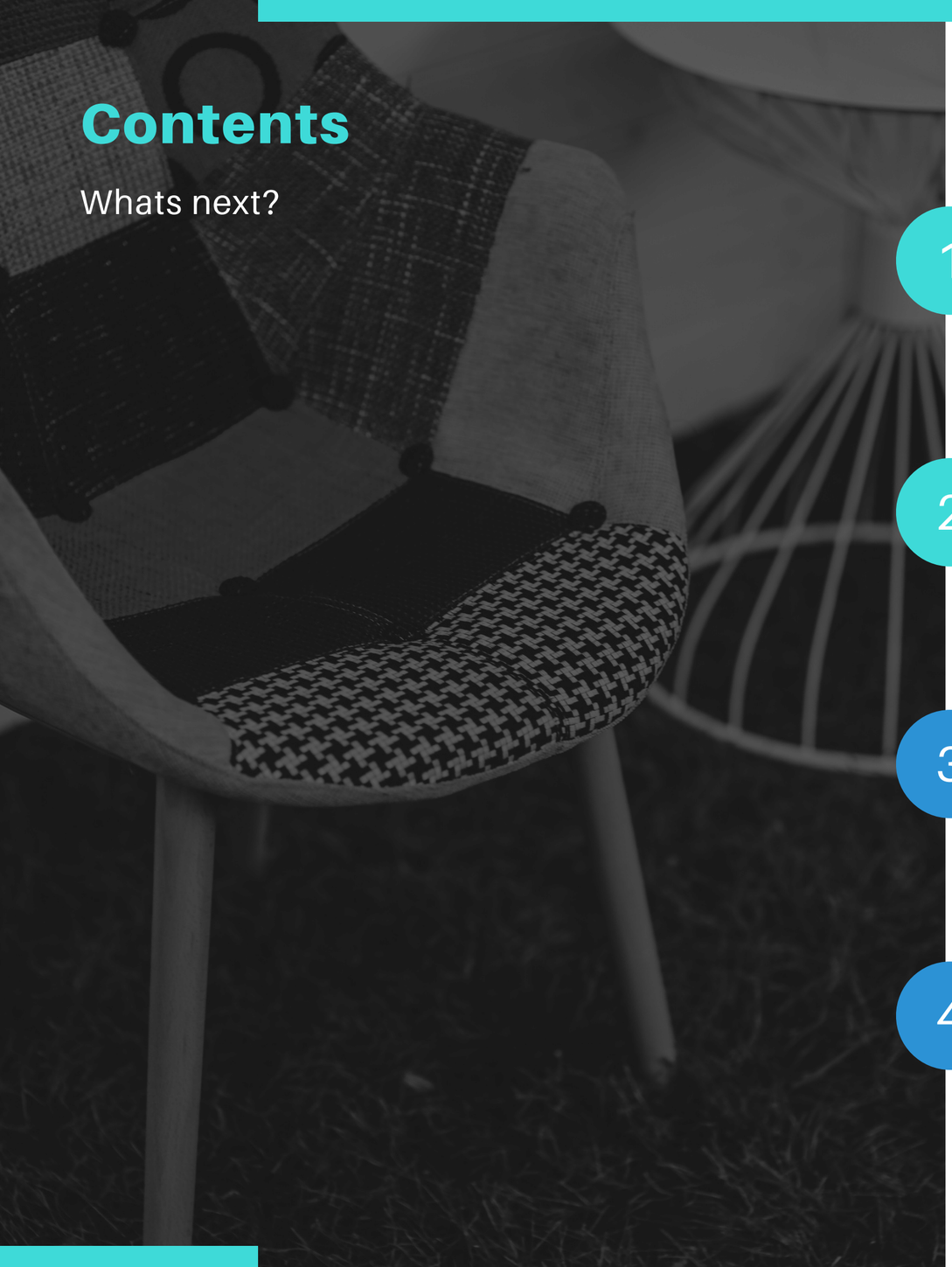
22 y.o backend web developer and CE student at AUT. - Noisyboy-9

GitHub

RELATED PROJECTS

- Colousses: Scheduling Algorithm Arena
- ECMUS: Fragmentation aware cloud-assisted edge scheduler (contributor)

AMIRKABIR UNIVERSITY OF TECHNOLOGY



Contents

Whats next?

1 WhoAmI

2 **What?**

3 Why?

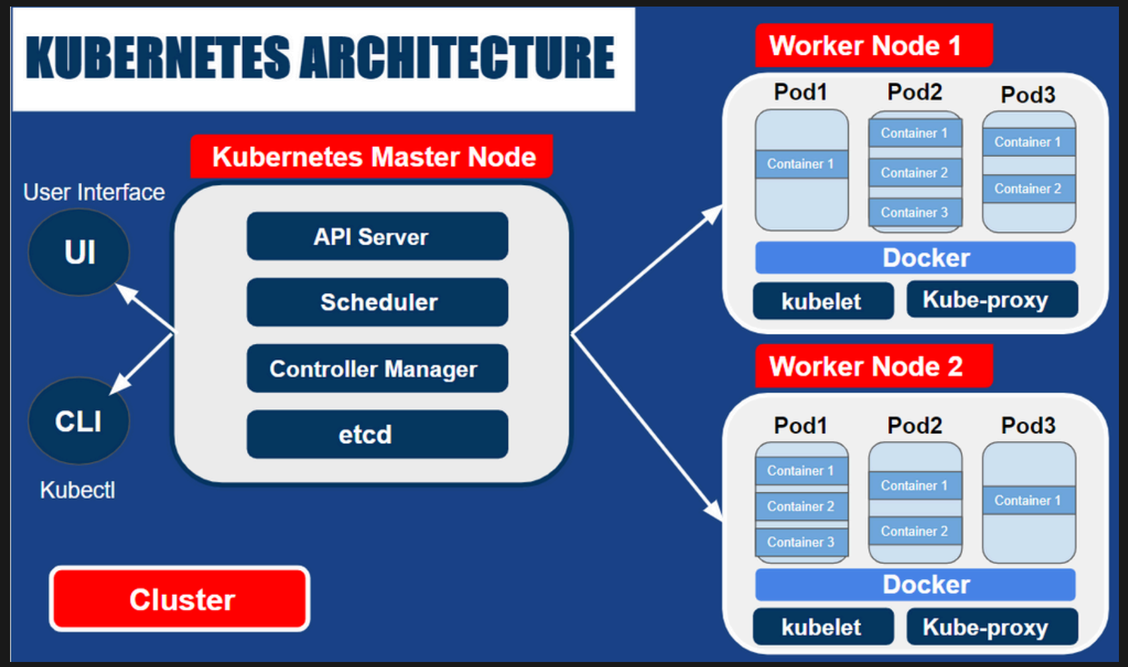
4 Live Coding

5 Deployment & testing

6 Final Notes & QA

What?

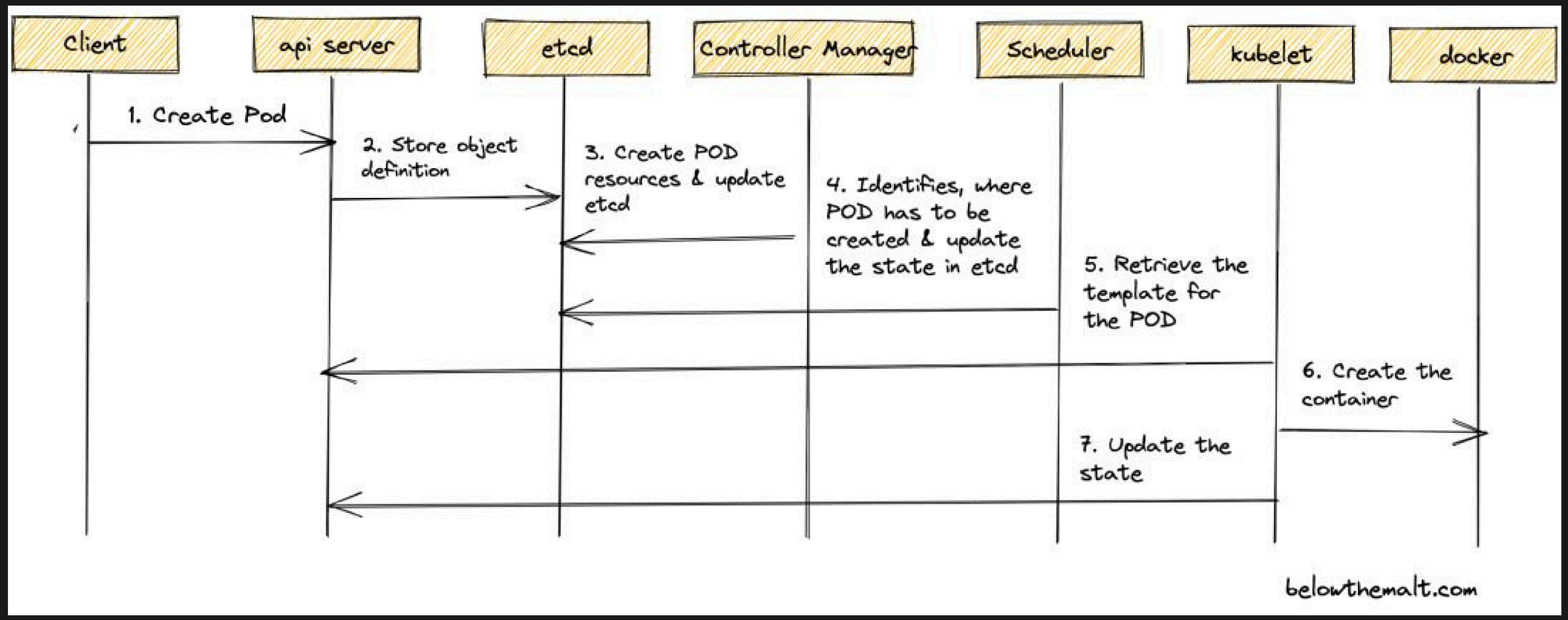
- The scheduler is a part of the control plane (master node), that schedules pods to nodes.
- In Kubernetes, scheduling refers to making sure that Pods are matched to Nodes so that Kubelet can run them.
- Kube-scheduler: K8s default scheduler **BUT REPLACEABLE!**



AMIRKABIR UNIVERSITY OF TECHNOLOGY

Scheduling Flow

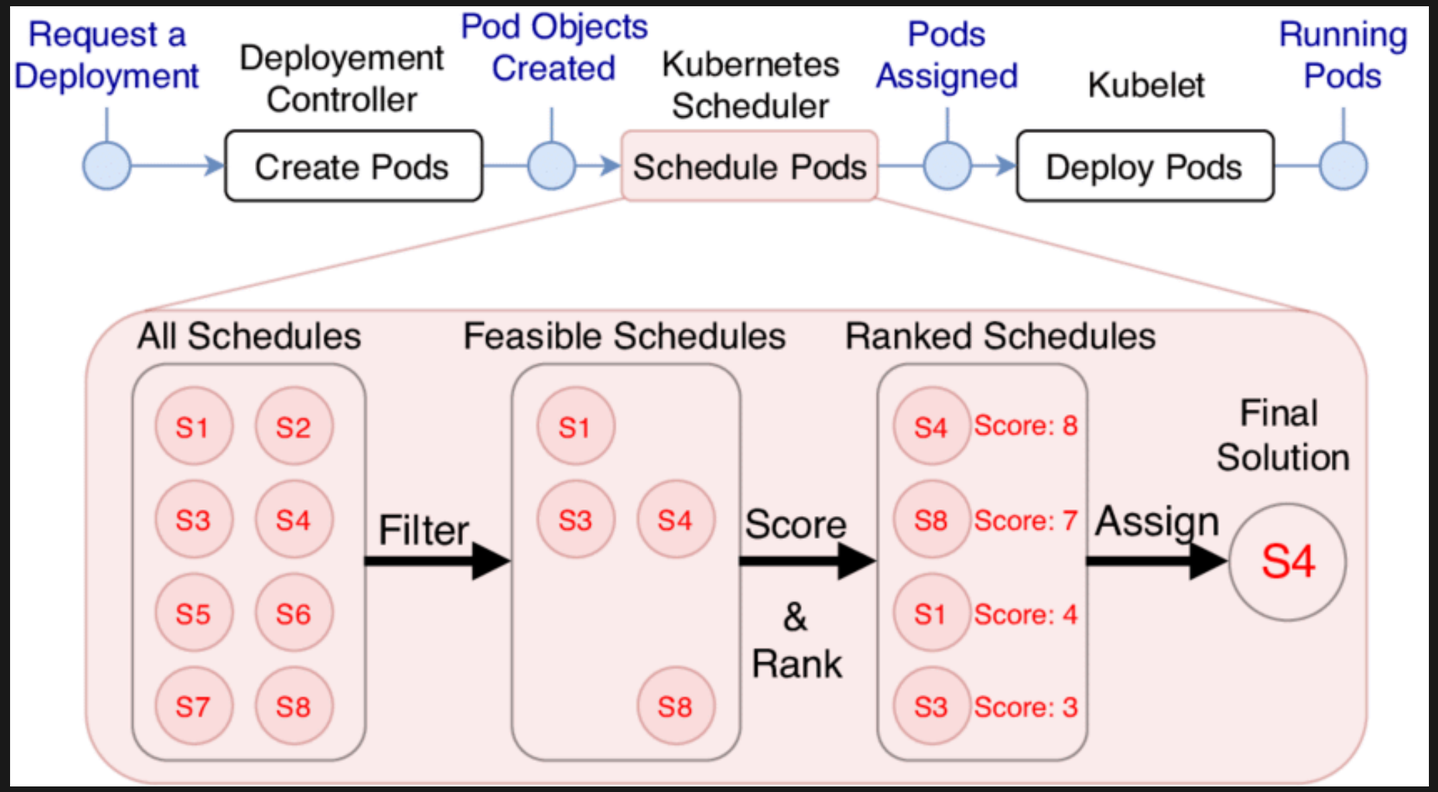
- In kubernetes, the flow that starts from deployment creation and leads to pod allocation on the node is often called **Scheduling Flow.**



AMIRKABIR UNIVERSITY OF TECHNOLOGY

Scheduling Flow (cont.)

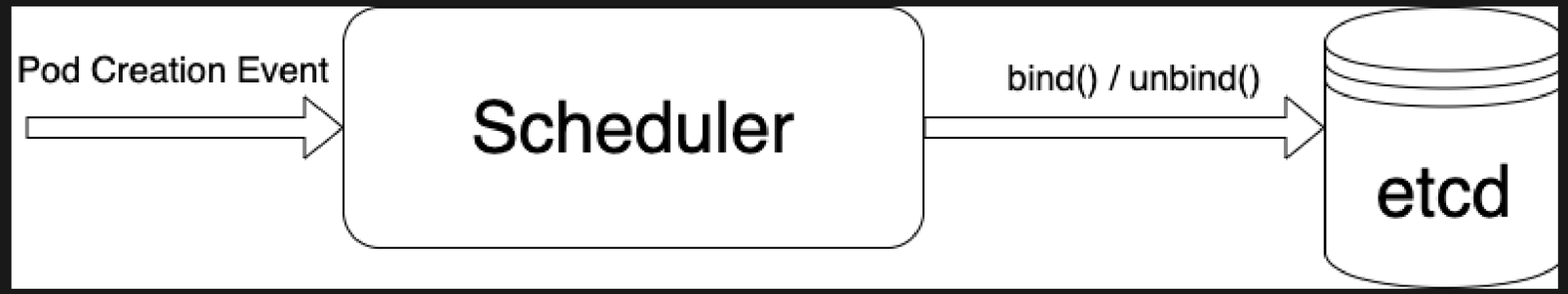
- Kube-scheduler has two phases: **Filtering and scoring**
- In the filtering phase, unfeasible nodes are filtered
- In scoring, a weighted some of scores for pod allocation is calculated
- The node with the highest score is chosen for binding.



AMIRKABIR UNIVERSITY OF TECHNOLOGY

Scheduling Flow (cont.)

- The scheduler is responsible for finding the feasible node for the pod.
- So, the scheduler isn't responsible for:
  - Notifying Kubelet of the worker node to create the new pod (etcd & kube-controllers)
  - Monitoring pod creation routines (Kubelet)
  - handle pod binding error (Kubelet, API-server, and controllers)
- Scheduler calls bind(pod, selected_node) in its source code and etcd gets updated and the rest of the process is other components' responsibilities!
- So the scheduler API is something like this:

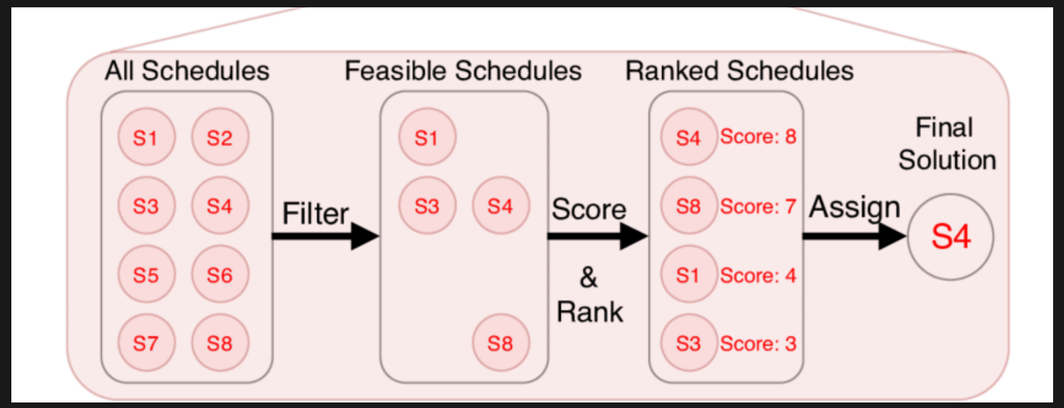


- This architecture is a huge advantage for writing a custom scheduler!

AMIRKABIR UNIVERSITY OF TECHNOLOGY

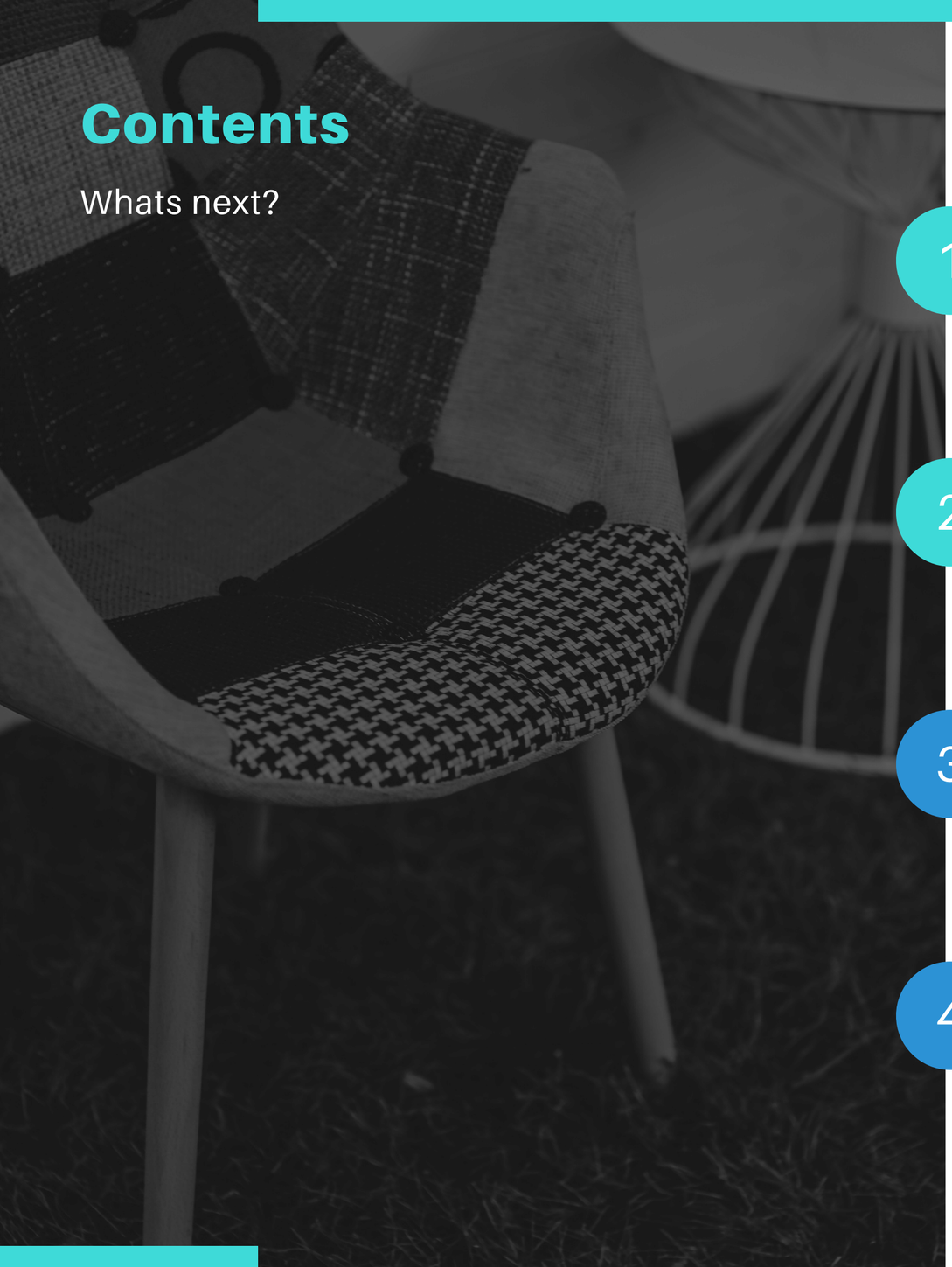
Fiddling with kube-scheduler!

- There is 2 main way to fiddle with the kubernetes scheduler:
  - Writing custom scorer plugin
  - Using k8s default options (e.g. node selector, taint and toleration, and etc)
- Default options are mostly for the **Filtering Phase**
- Custom scorer is for **Ranking Phase.**



- Kube-scheduler is mostly optimized for **pod spreading.** This its nature and can't be that much fiddled with!
- **Question:** What if we want to change this behavior?
- **Solution:** Write your own custom scheduler!

AMIRKABIR UNIVERSITY OF TECHNOLOGY



Contents

Whats next?

1 WhoAmI

2 What?

3 **Why?**

4 Live Coding

5 Deployment & testing

6 Final Notes & QA

Why?

- **Is it hard to write a custom scheduler?** Yes it is not!
- Writing a simple scheduler is easy.
- It is extremely hard when you have to take into account:
  - Scalability
  - Availability
  - Good algorithms design
  - etc

- Some use cases for writing a custom scheduler:
  - **AIaas**: Kube-scheduler isn't optimized for AI deployments
  - **Edge computing**: Using edge computing paradigm
  - **Better algorithm**: Scheduling is NP-Hard by nature. You have a better algorithm?
  - **Learning**: it is one of the best ways to learn the inner workings of k8s
  - **Fun**: it is a very cool project to have in your resume!

Contents

Whats next?

1 WhoAmI

2 What?

3 Why?

4 **Live Coding**

5 Deployment & testing

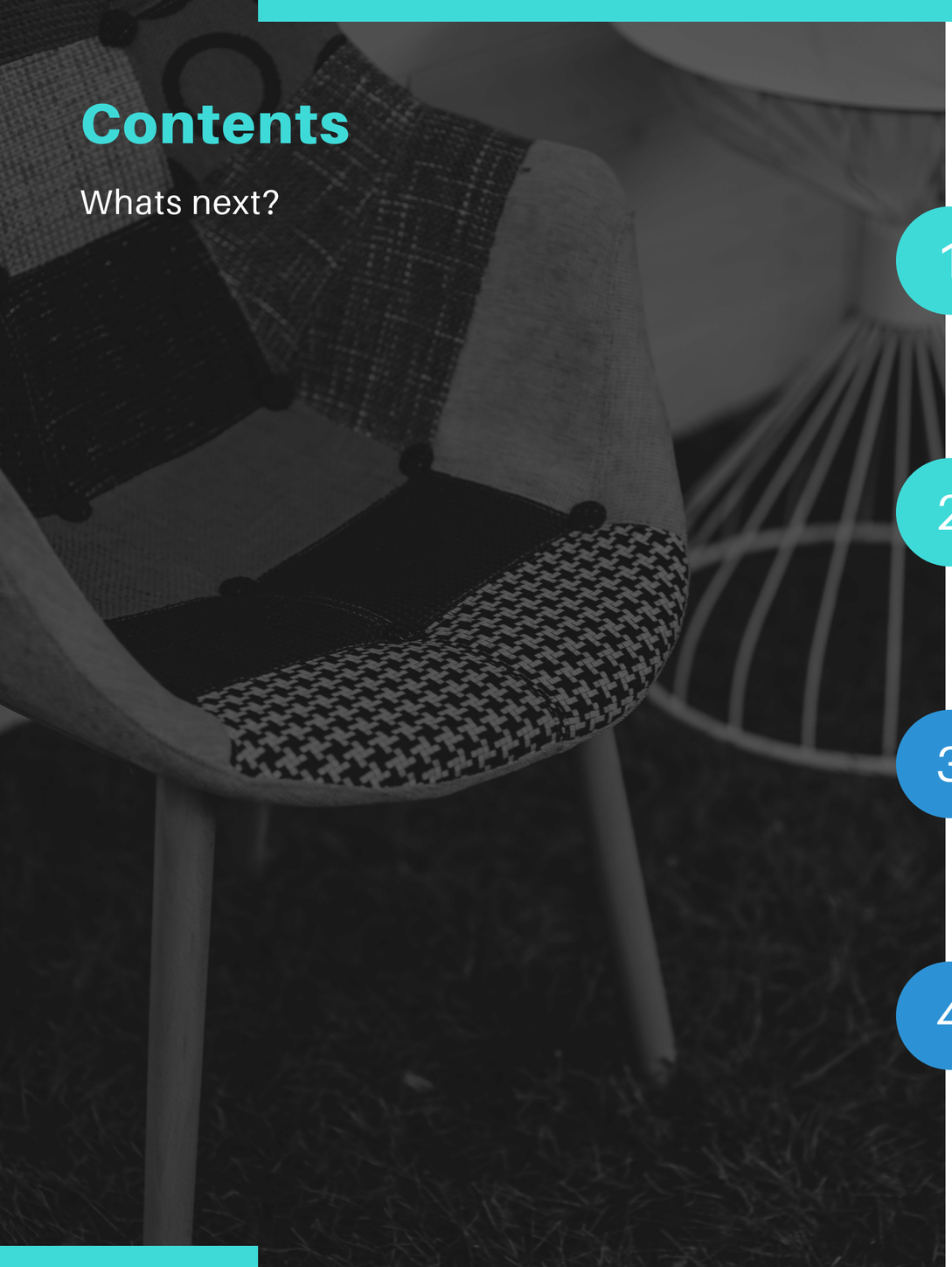
6 Final Notes & QA

But first . . .

Welcome to project Lattary!

Coding Time!

Welcome to project cutom_k8s_scheduler!



Contents

Whats next?

1 WhoAmI

2 What?

3 Why?

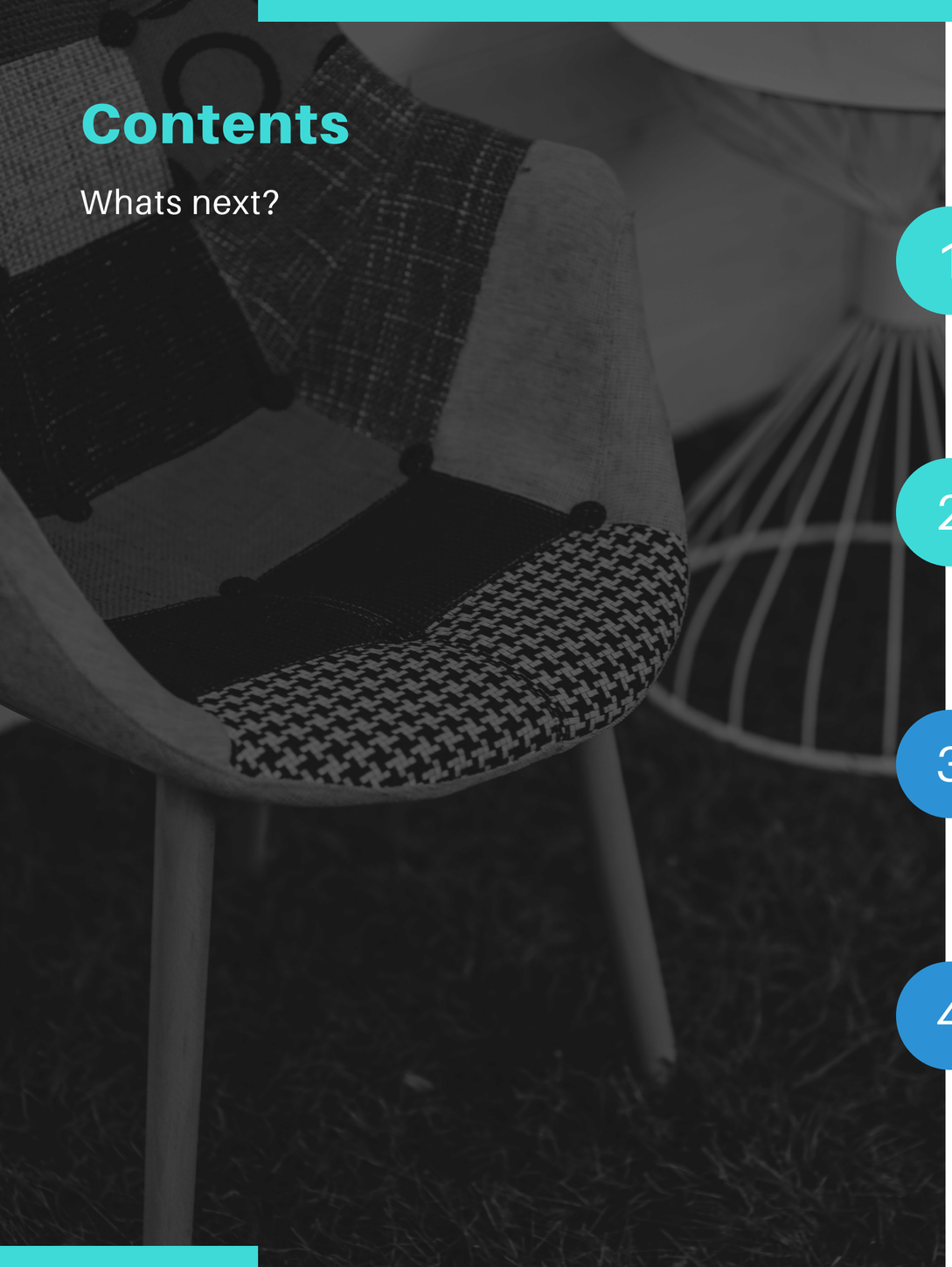
4 Live Coding

5 **Deployment & testing**

6 Final Notes & QA

Lets Deploy and test!

For the love of god... Work :)



Contents

Whats next?

1 WhoAmI

2 What?

3 Why?

4 Live Demo

5 Deployment & testing

6 **Final Notes & QA**

Final Notes

- Future of the project: there will be a V2 containing all of the project lottery algorithms and the algorithms are swappable at runtime
- The project repo is here.

Q/A

- Any questions?
- Feel free to contact me at: @sina_shariati

Thanks for your time & attention :)

AMIRKABIR UNIVERSITY OF TECHNOLOGY