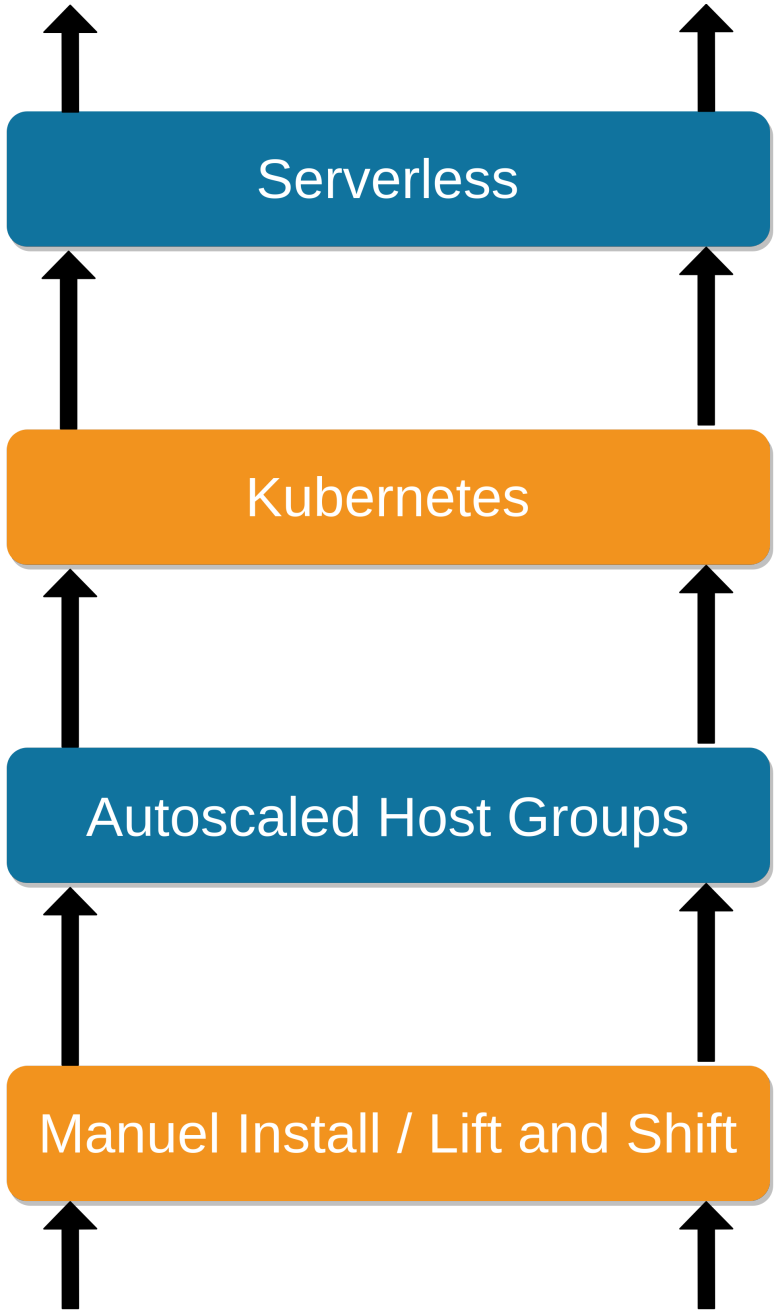
Scaling talk draft

- not just a talk, also demo at bastiandg/scaling-in-the-cloud

Scaling in the cloud

- cloud is not about cost, it is about scaling
- side effect: acceleration
- immutable infrastructure required
- Everything talks via http

Scaling Ladder Illustration



Lift and shift

- Migrate legacy VMs/Software with (almost) no modification
- Doesn't scale very well (bigger VMs, faster disks)
- Your mess for less
- Conclusion: don't do it

Autoscaled Host Groups

- Lift and Shift +
- Paradigm shift: VMs are containers
- Make infrastructure immutable
- Persist data outside VM
- vm scaling mechanisms of the cloud (gcp: managed instance group)

immutable infrastructure

- Computing infrastructure doesn't change at runtime
- Operating System images are prebuilt
- Separation of storage and computing

k8s

- Paradigm shift: There are no VMs
- Pool of resources
- services containerized
- Scaling unit: pod

Demo kubernetes

- Containerize Service
- Set up node auto scaling
- Set Pod autoscaling
- Deploy

serverless

- Paradigm shift: There is no infrastructure
- No infrastructure management
- Pay per use
- stateless
- cloudfunctions
- Knative
- cloud run

Demo serverless

- Containerize Service
- Deploy

Takeaways

- Don't do lift and shift
- separate storage from computing