



KubeCon



CloudNativeCon

Europe 2022

WELCOME TO VALENCIA





KubeCon



CloudNativeCon

Europe 2022

Building for the (Inevitable) Next Cloud Outage

Pavel Nikolov, Section



Building for the (Inevitable) Next Cloud Outage

It is not a question of if
but rather a question of when...



Pavel Nikolov
Software Engineer
 **Section**



No cloud provider can guarantee 100% uptime

- There are external factors - like internet routes
- Human factor - we make mistakes
- Natural disasters
- Prohibitively expensive to replicate everything everywhere

It is not a question of “if” but rather “when”

- Your cloud will go down some day (when you least expect it)
- It will (likely) happen more than once
- Their upstream internet providers might go down
- Your DNS might go down
- Your team might deploy bad code by mistake or delete data by mistake

Disaster Recovery strategies anyone?

- Active / Active deployment (DNS switch)
- Active / Passive deployment (spin up + DNS switch)
- Periodic Backup (manual recovery)
- No DR strategy (figure it out when it happens)



KubeCon



CloudNativeCon

Europe 2022

Discipline and human factor

- DR requires discipline and solid processes in place
- Do you have dedicated SRE engineer(s) in your team?
- How often do you test your DR strategies?
- Does your DR strategy require manual intervention?
- The human factor during incidents

What are your RTO & RPO?

- **RTO** - Recovery Time Objective
- **RPO** - Recovery Point Objective

What if there is a different approach?

- Build a self-healing system
- Cloud native
- No single point of failure
- Expect that anything could go down at any time
... even the DNS servers



KubeCon



CloudNativeCon

Europe 2022

BGP + Anycast IPs to the rescue

What is BGP?



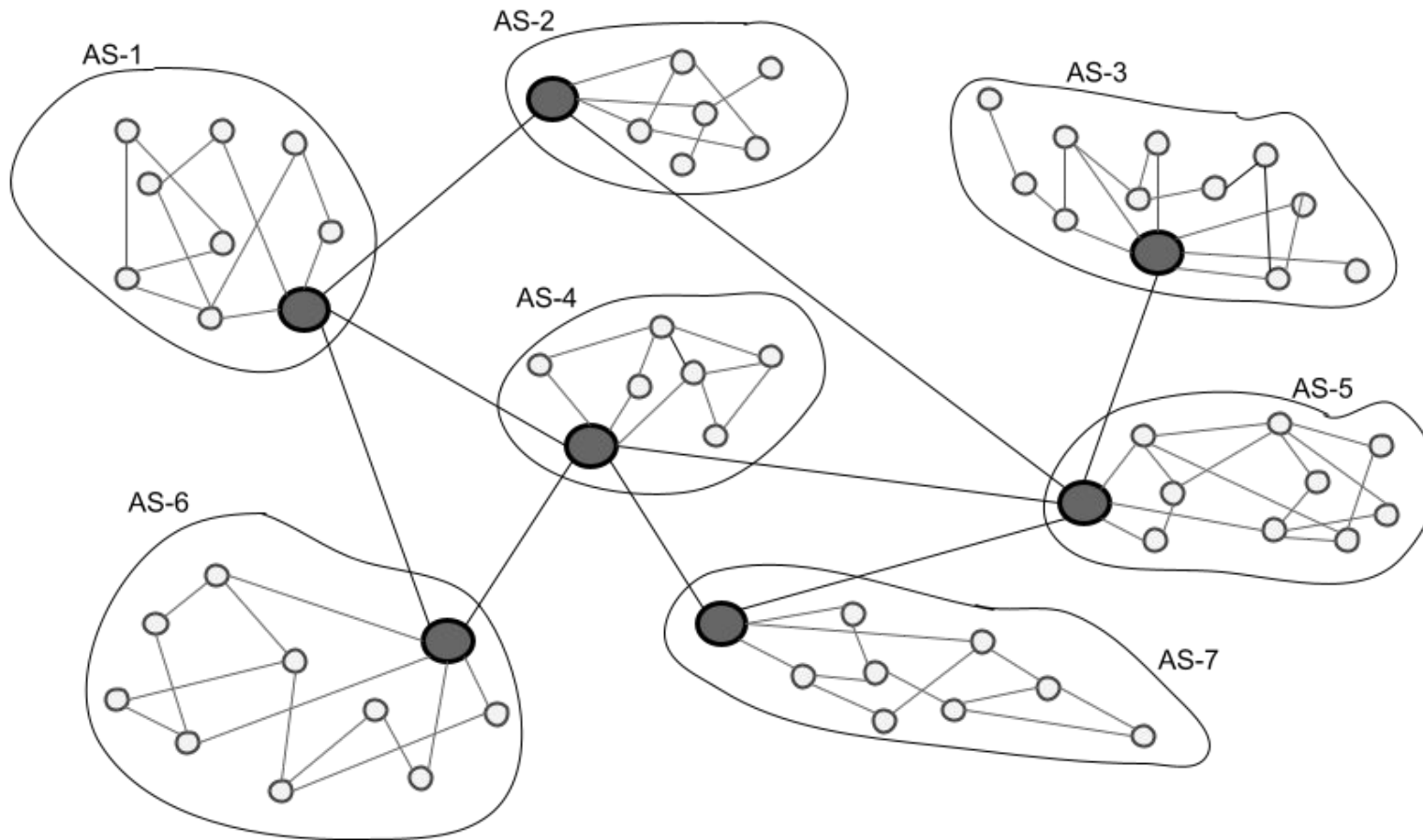
KubeCon



CloudNativeCon

Europe 2022

When you send data over the internet BGP makes sure that you reach your destination using the most efficient path.



Each AS can announce one or more IP ranges.

*If, for example, AS-1 announces **1.2.3.0/24** then packets from all other ASes to addresses like **1.2.3.1**, **1.2.3.4**, **1.2.3.105**, etc. would be sent to AS-1.*

What is Anycast IP?

There are different types of IP packets:

- Unicast - one-to-one
- Multicast - one-to-many
- Anycast - one-to-nearest

Benefits from BGP (over DNS)

- DNS has TTL (usually at least 300s)
- BGP convergence takes seconds

Downsides

- You have to own an IP address range
- Your cloud provider has to support BYO IP
- Learning curve

Anycast IP in action



KubeCon



CloudNativeCon

Europe 2022





KubeCon



CloudNativeCon

Europe 2022

Demo

(Demonstrate different servers with the same public IP)

What about data (consistency)?

The answer is (almost) always Eventual Consistency



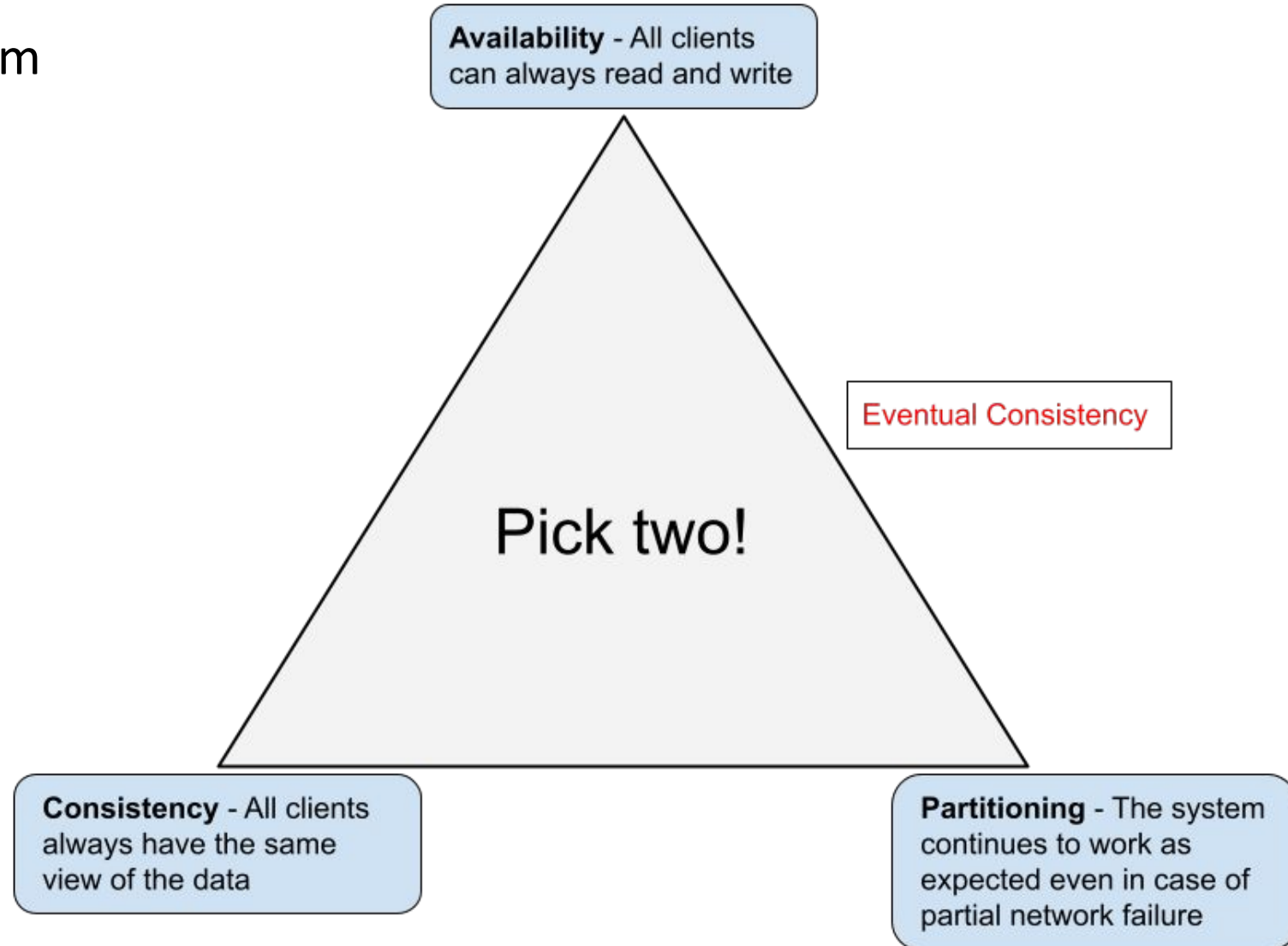
KubeCon



CloudNativeCon

Europe 2022

CAP Theorem



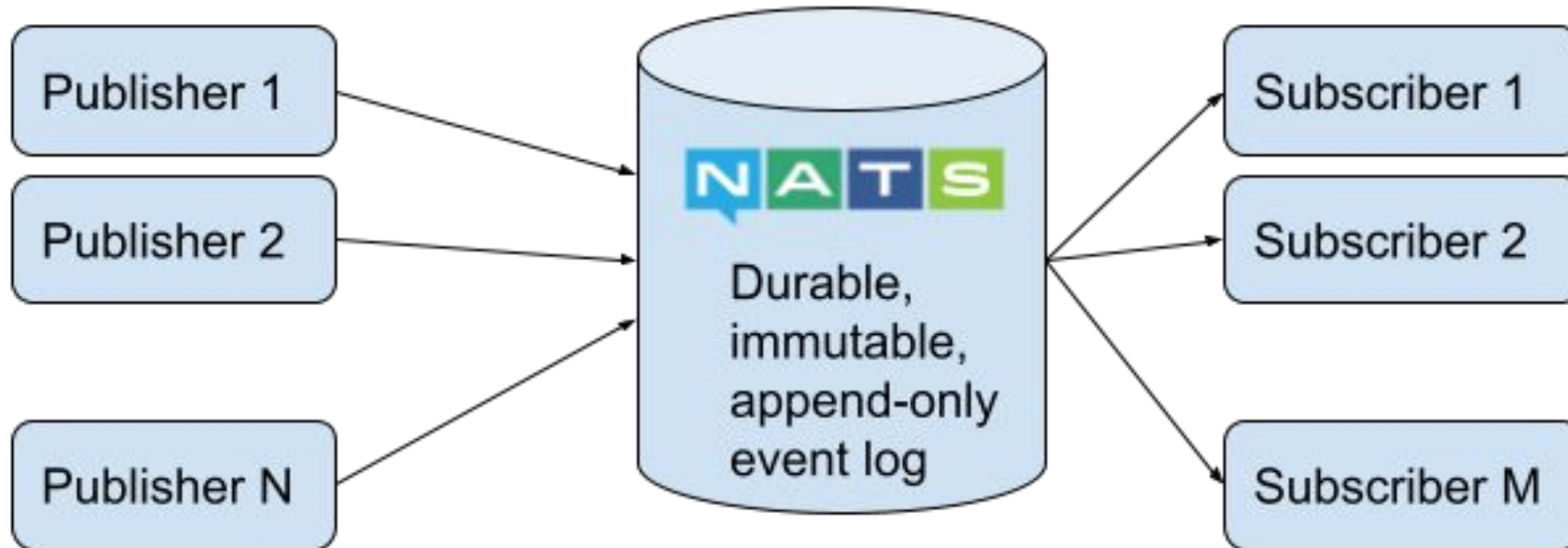
You can use a hosted DB solution

- Do they provide 100% uptime guarantee?
- Do they have 100% consistency?
- How much do they cost?
- Are they yet another SPF in your system?

Event Sourcing is a perfect fit

- Ideal for Microservice architecture
- CQRS pattern - command query responsibility segregation
- Requires a durable event store (e.g. NATS Jetstream or Kafka)
- Immutable data
- Reproducible state
- Eventual Consistency

How does it work?



The event log

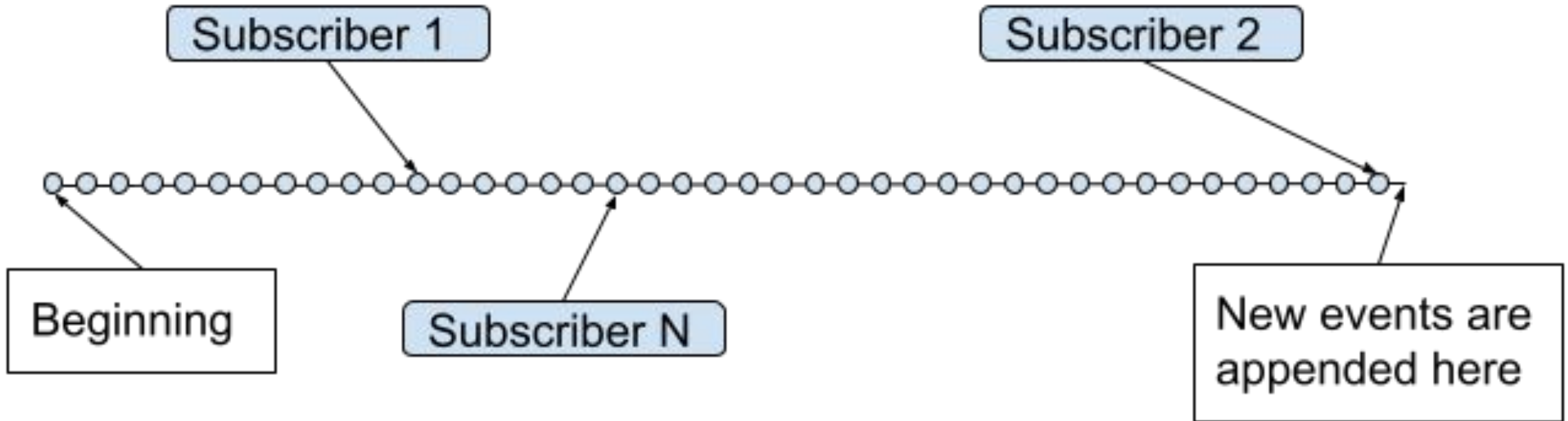


KubeCon



CloudNativeCon

Europe 2022



When does my pod become ready?

1. Start a consumer pod
2. Read last event from the event log and store its ID (let's call it X)
3. Read all events since the beginning of time
4. Turn readiness to true once you pass X
5. Continue consuming events



KubeCon



CloudNativeCon

Europe 2022

Demo

(A simple app which uses event sourcing with NATS Jetstream and is deployed to multiple Kubernetes clusters with Anycast IP)

In summary

- Outages will happen sooner or later
- It is expensive and non-trivial to have a reliable DR strategy in place
- Anycast IPs + Event Sourcing would result in a great reliability, low downtime windows (RTO) and no data loss at a reasonable cost



KubeCon



CloudNativeCon

Europe 2022

Questions?

Come talk to me at
Booth (#S28, Pavilion 2)



Section

Resources

- [Border Gateway Protocol](#)
- <https://nats.io>
- Bird - <https://bird.network.cz>
 - Alternatively - GoBGP - <https://github.com/osrg/gobgp>
- K3s - <https://k3s.io>
- <https://github.com/graph-gophers/graphql-go>