Cloud Computing Final Project

Piero Zappi

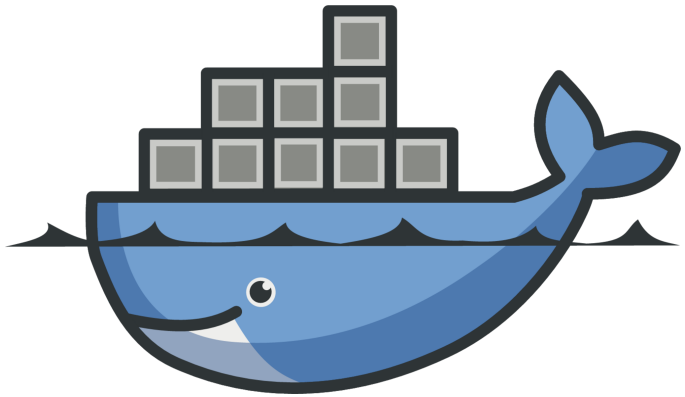
September 2024

Nextcloud Platform:

- Open source
- Extensive documentation
- User-friendly web interface
- Many built-in features
- Availability of a `Docker` image

`Docker` Deployment: `docker-compose.yaml` file

- Nextcloud
- MariaDB database
- Redis caching
- Locust for load testing



Each service is configured with its respective `docker` image, environment variables, volumes for persistent storage and network settings; all containers share the same network.

Cloud Basic - Cloud-Based File Storage System

User Authentication and Authorization:

- Sign up, log in, log out
- Role-based access control (admins and regular users)
- Admins can manage users

Storage and File Operations

- Private storage space
- Upload, download, and delete files

Cloud Basic - Cloud-Based File Storage System

Security:

- Server-side encryption
- Two-factor authentication
- Password policies
- Strong password practices

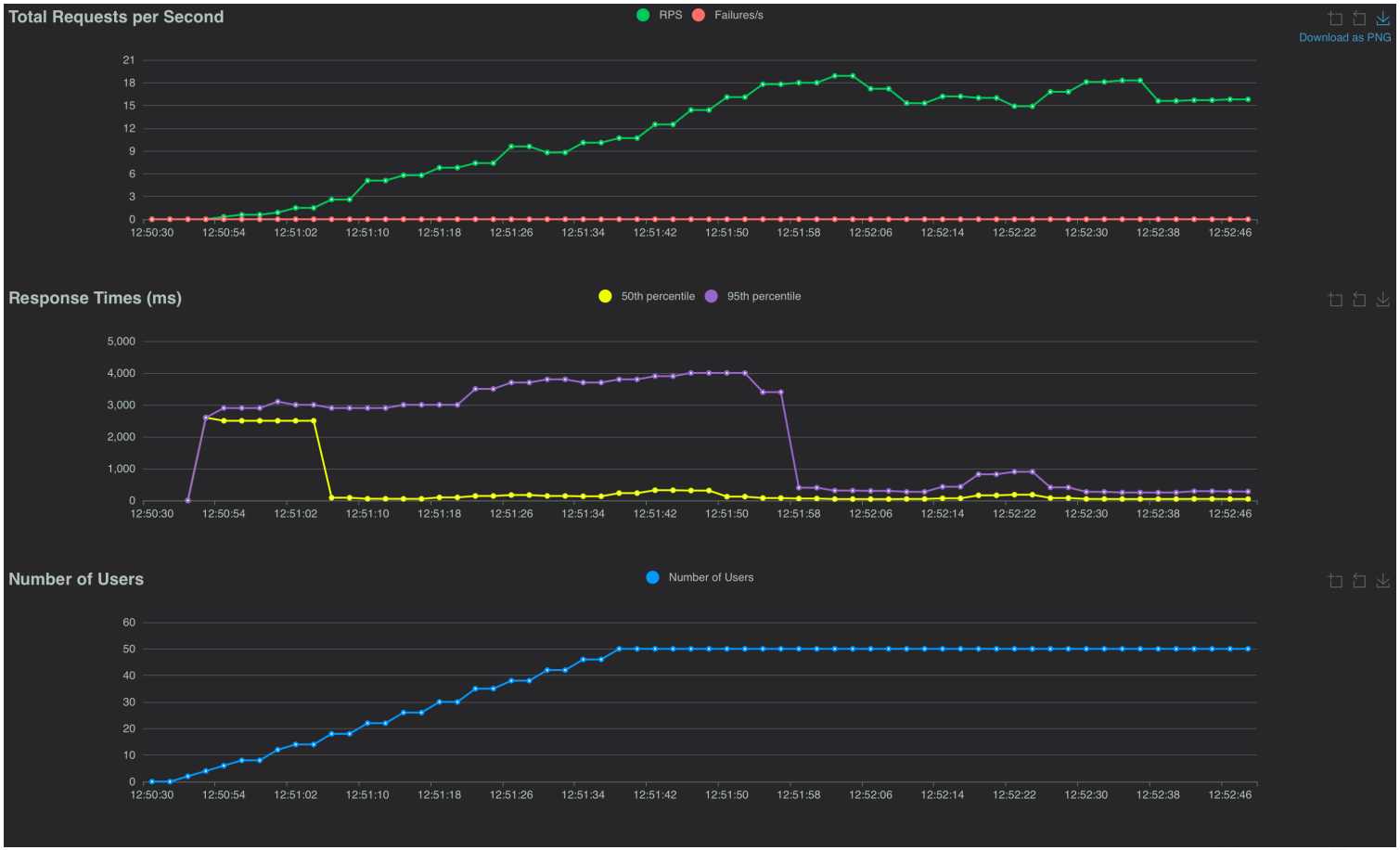
Monitoring:

- System metrics
- Logging capabilities

Locust Load Test:

Task	Probability
List files	16.1%
Read a file	16.1%
Upload a 5 KB file	32.3%
Upload a 5 MB file	16.1%
Upload a 1 GB file	3.2%
Download a 5 MB file	16.1%

Locust Load Test: executed on a MacBook Pro M2, with Locust configured to spawn one user per second, up to a maximum of 50 concurrent users. Each second, all spawned users attempt to perform one task. The test was scheduled to run for 120 seconds.



Scalability and Cost-Efficiency: Cluster Deployment

- *Advantages*:
  - hardware control
  - data residency
  - no third-party providers
  - custom policies and security measures
  - predictable costs, lower in the long run
- *Disadvantages*:
  - big initial investment
  - maintenance and upgrades costs
  - technical expertise

Scalability and Cost-Efficiency: Cloud Provider Deployment - AWS

- Virtually unlimited scalability
- Robust autoscaling features
- Composability
- Advanced security features
- Integrated services (S3, RDS)
- Pay-as-you-go model
- Monitoring tools

Scalability and Cost-Efficiency: Cloud Provider Deployment - Iaas, Paas, SaaS

- *Advantages*:
  - no big initial investment
  - virtually unlimited scalability
  - pay-as-you-go model
  - no maintenance and management of physical infrastructure
- *Disadvantages*:
  - third-party providers
  - no data residency
  - less predictable costs
  - fee costs, depending on the model
  - less control over the infrastructure, depending on the model

Cloud Advanced - Cloud-Based File Storage System in Kubernetes

Kubernetes **Deployment**: official Nextcloud Helm chart, `values.yaml file`

- Single-node Kubernetes cluster
- Nextcloud pods with 3 containers, linked to a PVC:
  1. Nextcloud application
  2. sidecar container for cron jobs
  3. nginx web server container
- External PostgreSQL database pod linked to a PVC
- Redis pod for caching
- MetalLB to expose the service
- `Secrets`

Cloud Advanced - Cloud-Based File Storage System in Kubernetes

Back-End Storage Limitations:

- Single point of failure
- Lack of dynamic provisioning
- Node affinity constraints
- Limited scalability

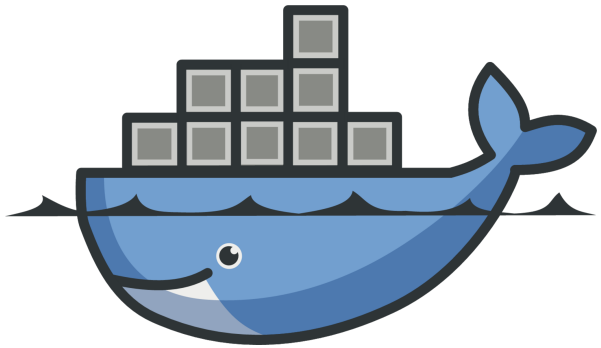
Cloud Advanced - Cloud-Based File Storage System in Kubernetes

High Availability Considerations:

- Adopt a distributed storage solution
- Enable dynamic provisioning
- Expand to a multi-node cluster
- Deploy multiple replicas
- Improve load balancing
- Horizontal pod autoscaler

Comparison with the Docker Solution:

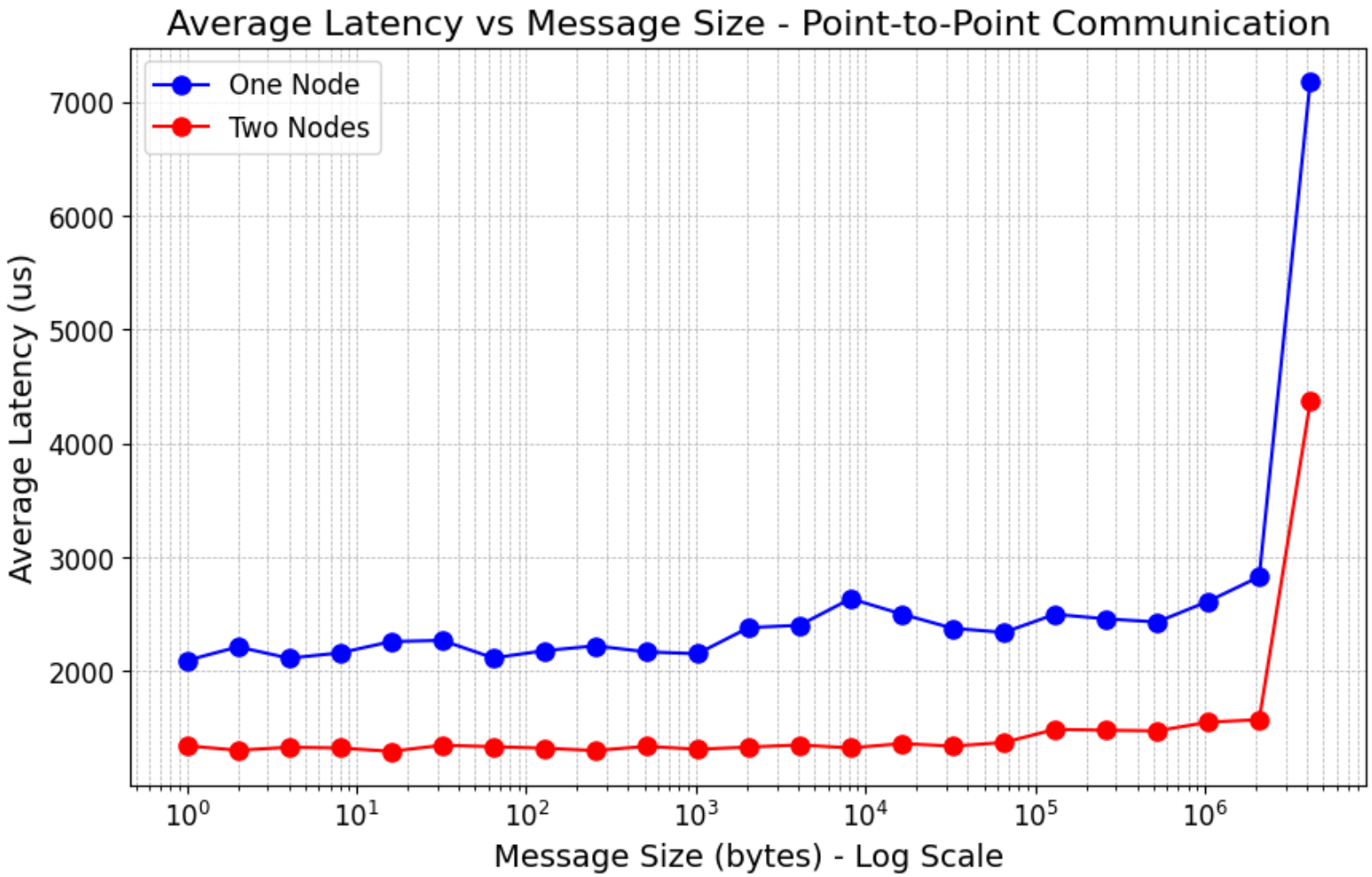
- *Advantages*:
  - scalability
  - high availability
  - self-healing
  - flexibility
  - load balancing
- *Disadvantages*:
  - more complex to set up
  - requires more resources



Solution Structure:

- Two-nodes Kubernetes cluster
- The nodes talk via Flannel
- MPI operator
- OSU benchmark container

OSU Benchmark: Point to Point Latency



OSU Benchmark: Scatter Collective Operation Latency