

# Cloud Computing – Part 2

Comprehensive Review

DSC 208R – Parallel Data Processing and the Cloud

---

## Contents

|          |   |          |
|----------|---|----------|
| <b>1</b> | <b>Motivation</b>                       | <b>2</b> |
| <b>2</b> | <b>Serverless Paradigm</b>              | <b>2</b> |
| <b>3</b> | <b>Example AWS Workloads</b>            | <b>2</b> |
| <b>4</b> | <b>Resource Disaggregation</b>          | <b>2</b> |
| <b>5</b> | <b>Is All This Complexity Worth It?</b> | <b>3</b> |
| <b>6</b> | <b>State of the Cloud Surveys</b>       | <b>3</b> |
| <b>7</b> | <b>Pros and Cons Summary</b>            | <b>3</b> |
| <b>8</b> | <b>Future Directions</b>                | <b>4</b> |

## 1 Motivation

Fixed bundles of CPU, memory, and storage can leave some resources under-used. Cloud providers therefore offer finer-grain renting models, notably serverless Function as a Service (FaaS), to boost resource efficiency and cut cost by up to 10x compared to spot instances.

## 2 Serverless Paradigm

- User supplies a function and a resource hint (CPU, RAM).
- Provider handles provisioning, autoscaling, and teardown.
- Billing is by the millisecond of execution and GB-s of memory.

### Car Analogy

Owning a car is like on-prem hardware. Ride-sharing on demand is like serverless: you pay only for miles driven.

## 3 Example AWS Workloads

1. Athena – serverless SQL over S3, schema on read, pay per TB scanned.
2. Predictive data science – SageMaker plus Data Lake for serverless ML training and inference.
3. IoT pipeline – edge devices stream data to SageMaker Neo models for real-time inference.

## 4 Resource Disaggregation

Logical next step: detach compute, memory, and storage so each is network attached and elastically added or removed in real time.



Figure 1: All resources network attached and elastic.

Ongoing research aims to hot-plug new CPUs or RAM with sub-second latency.

## 5 Is All This Complexity Worth It?

- Pros of cloud: manageability, elastic capacity, pay-as-you-go cost.
- Cons of cloud: API complexity, long-term spend crossover, accidental waste, vendor lock-in, privacy and downtime risks.:contentReference[oaicite:4]index=4

Large enterprises, health care, and academia still rely on on-prem clusters or hybrid clouds to balance these trade-offs.

## 6 State of the Cloud Surveys

Flexera surveys show

- Public cloud adoption continues to rise each year.
- Multi-cloud strategies are common.
- Top challenge is controlling cost and cloud spend.:contentReference[oaicite:5]index=5

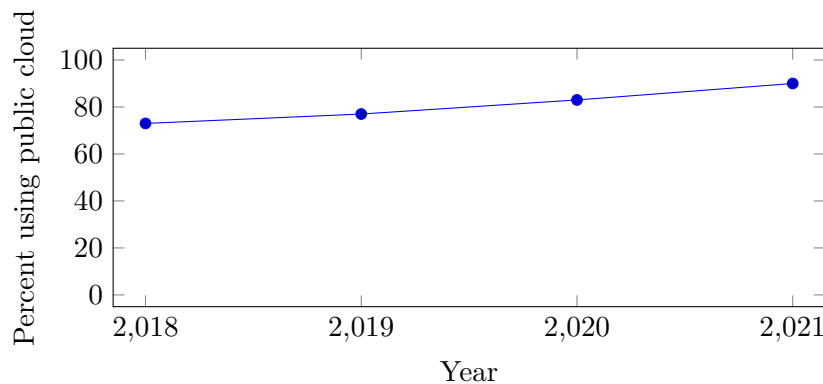


Figure 2: Trend in public cloud adoption (illustrative numbers).

## 7 Pros and Cons Summary

- **Pros**
  - No hardware maintenance.
  - Fine-grained cost aligned with use.
  - Rapid scale up and down.

- **Cons**

- Complex APIs and licenses; need CloudOps expertise.
- Long-term cost can exceed on-prem clusters.
- Vendor lock-in, outages, privacy and governance risks.

## 8 Future Directions

- Fully disaggregated clouds with sub-second elasticity.
- Automatic budget guards and cost observability tools.
- Cross-vendor orchestration to reduce lock-in risk.

## Conclusion

Cloud Computing Part 2 underscores how serverless and resource disaggregation push the pay-as-you-go vision, yet introduce new complexity and risk. Choosing between on-prem, hybrid, and cloud-native architectures requires careful cost, performance, and governance analysis.:contentReference[oaicite:6]index=6