

THE CLOUD BASICS

15 cloud concepts every AI builder should know.

What each one is, when your AI app actually needs it, and one free way to try it. In plain words, with no jargon.

1

Run your app

Virtual machines, containers, block storage

2

Handle traffic

Load balancing, autoscaling, DNS

3

Connect it safely

Subnets, routing, NAT gateway

4

Lock it down

Least privilege, workload identity, encryption

5

Keep it running

Observability, infrastructure as code, disaster recovery

CHAPTER 1 OF 5 · WHERE YOUR APP LIVES

1

Run your app

Rent a computer, pack your app so it runs anywhere, and keep its data safe.



CONCEPT 1 OF 15

Virtual machines

WHAT IT IS

A computer you rent by the hour. You pick the size, install what you need, and it runs in a data centre instead of on your desk.

YOUR AI APP NEEDS IT WHEN

Your agent or API has to run all day, not only while your laptop is open.

TRY IT FREE

Launch a small virtual machine on a cloud free tier or trial, then connect to it from your terminal. Check the provider's current free-tier terms first.



CONCEPT 2 OF 15

Containers

WHAT IT IS

Your app packed together with everything it needs, so it runs the same on any machine.

YOUR AI APP NEEDS IT WHEN

"It works on my laptop" has to become "it works on the server".

TRY IT FREE

[Docker's Get Started tutorial.](#)

3

CONCEPT 3 OF 15

Block storage

WHAT IT IS

A hard drive you plug into a server. If the server is replaced, the drive and its data stay.

YOUR AI APP NEEDS IT WHEN

Your app keeps files or a database that must survive a restart.

TRY IT FREE

Attach a small volume to your practice virtual machine and save a file on it. Restart the machine and check the file is still there.

CHAPTER 2 OF 5 · WHEN MORE PEOPLE SHOW UP

2

Handle traffic

Share the load, grow and shrink on demand, and let people find you by name.



CONCEPT 4 OF 15

Load balancing

WHAT IT IS

Spreads incoming requests across several servers, and skips any server that is down.

YOUR AI APP NEEDS IT WHEN

More than one copy of your API or model server is running.

TRY IT FREE

Run two copies of a small app with Docker and put Nginx in front of them.



CONCEPT 5 OF 15

Autoscaling

WHAT IT IS

Adds servers when it is busy and removes them when it is quiet, so you pay for what you use.

YOUR AI APP NEEDS IT WHEN

Traffic is hard to predict, like a chatbot after one post goes viral.

TRY IT FREE

The Kubernetes Basics tutorial. Scale an app up and down and watch what happens.



CONCEPT 6 OF 15

DNS

WHAT IT IS

The phone book of the internet. It turns a name like buildwithmrj.com into the address of a server.

YOUR AI APP NEEDS IT WHEN

Real people need to reach your app by a name, not a number.

TRY IT FREE

Cloudflare's free DNS. Point a domain you own at a test server.

CHAPTER 3 OF 5 · WHO CAN TALK TO WHAT

3

Connect it safely

Split your network into rooms, decide the routes, and let private servers reach out without letting anyone in.

7

CONCEPT 7 OF 15

Subnets

WHAT IT IS

Your private network cut into sections. Public sections face the internet, private ones do not.

YOUR AI APP NEEDS IT WHEN

Your database and agents should never be reachable from outside.

TRY IT FREE

In a cloud free tier, create a network with one public and one private subnet.

8

CONCEPT 8 OF 15

Routing

WHAT IT IS

The rules that decide where each request goes next.

YOUR AI APP NEEDS IT WHEN

Parts of your app live in different networks and need to find each other.

TRY IT FREE

Open the route table of the network you just made. Add a route and watch what changes.

9

CONCEPT 9 OF 15

NAT gateway

WHAT IT IS

Lets private servers reach the internet for updates and AI services, while the internet cannot reach them.

YOUR AI APP NEEDS IT WHEN

Private agent servers need to call outside services such as a model API.

TRY IT FREE

Add one to your private subnet and test an outgoing request. These usually cost a little even on free tiers, so delete it when you are done.

CHAPTER 4 OF 5 · KEEP THE WRONG PEOPLE OUT

4

Lock it down

Give each part only the access it needs, drop stored passwords, and scramble the data.

10

CONCEPT 10 OF 15

Least privilege

WHAT IT IS

Every person and every program gets only the access it needs, and nothing more.

YOUR AI APP NEEDS IT WHEN

An agent can touch files, databases or payments. If it makes a mistake, the damage stays small.

TRY IT FREE

Create a role that can only read one storage bucket, then try to use it to change something. It should fail.



CONCEPT 11 OF 15

Workload identity

WHAT IT IS

Apps prove who they are to the cloud with short-lived credentials, instead of passwords stored in files.

YOUR AI APP NEEDS IT WHEN

Your app calls cloud services and you do not want keys sitting in your code.

TRY IT FREE

Your cloud provider's workload identity guide. Replace one stored key in a test app.



CONCEPT 12 OF 15

Encryption

WHAT IT IS

Data scrambled so it is useless without the key, on the disk and while it travels over the web.

YOUR AI APP NEEDS IT WHEN

You store customer data, or send prompts that contain private details.

TRY IT FREE

[Let's Encrypt gives free HTTPS certificates. Set one up on a test site.](#)

CHAPTER 5 OF 5 · WHEN SOMETHING GOES WRONG

5

Keep it running

See problems early, rebuild your setup from a file, and keep a copy for the bad day.



CONCEPT 13 OF 15

Observability

WHAT IT IS

Logs, numbers and traces that show what your system is doing and why.

YOUR AI APP NEEDS IT WHEN

Your AI app answers slowly or wrongly and you need to know where it went wrong.

TRY IT FREE

[Grafana Cloud's free tier. Send logs from a small app and build one chart.](#)



CONCEPT 14 OF 15

Infrastructure as code

WHAT IT IS

Your servers and networks described in files. One command builds them, and you can rebuild them the same way.

YOUR AI APP NEEDS IT WHEN

You want a setup you can rebuild after a mistake, or copy for a client.

TRY IT FREE

[HashiCorp's Terraform tutorials. Start with the Docker one, it runs on your own laptop.](#)

15

CONCEPT 15 OF 15

Disaster recovery

WHAT IT IS

A plan and a copy kept somewhere else, so you can recover when something breaks or gets deleted.

YOUR AI APP NEEDS IT WHEN

Losing the data, or being down for a day, would really hurt.

TRY IT FREE

Back up a small database to a second storage location, then restore it onto a fresh machine. Practise the restore, not just the backup.

WHERE TO START

Learn them in **this order.**

Fifteen at once is too much. This is the order I would take them in, one month, about an hour a day.

WEEK 1**Containers and virtual machines**

Learn to pack an app and run it on a rented computer. Everything else builds on this.

WEEK 2**DNS, encryption, least privilege**

Put your app on a real name with HTTPS, and decide who and what can touch it.

WEEK 3**Observability and infrastructure as code**

Watch what your app does, then write your setup down so you can rebuild it.

WEEK 4**Networking, storage and backups**

Subnets, routing and the NAT gateway, then block storage and a backup you have restored once.

**AFTER
THAT****Load balancing, autoscaling, workload identity**

These matter once you have real traffic and a real team.

You do not need to master any of these. Understand what each one does, try it once, and move on. It will make sense when you need it.

CLOUD IS ONE PART

Now put it to **work.**

These concepts are the ground your AI apps stand on. The next step is knowing what to build on top of it, and in what order.

THE NEXT STEP

The AI Engineering Roadmap

A step-by-step path from beginner to building real AI systems, with the projects and resources in a clear order, so you always know what to learn next.

[See the roadmap](#)

I'm Mr J

I'm an AI engineer and software developer in Germany. I build AI systems and agents, and I share how they work in plain words.

MORE FREE GUIDES

buildwithmrj.com

FOLLOW FOR MORE

[@igbokwe_inr](#)

Cloud tools and free tiers change often. Links and tutorials here were checked in October 2026. Always check a provider's current free-tier terms and pricing before you create anything.