

[Open in app](#)[Get started](#)

You have **1** free member-only story left this month. [Sign up for Medium and get an extra one](#)



Chouaieb Nemri

[Follow](#)

Jan 31 · 4 min read

[Listen](#) [Save](#)

Your 2022 Data Engineering Roadmap



What's up dear readers ? I've been absent for almost two months. It was for a good reason, I was working on one of my Dreams (I'll talk about it in a Future blog post)

Introduction

A new year means, new resolutions. A good resolution to have in 2022 is doing a career



[Open in app](#)[Get started](#)

the tech industry so feel free to skip the steps and concepts you're already comfortable with.

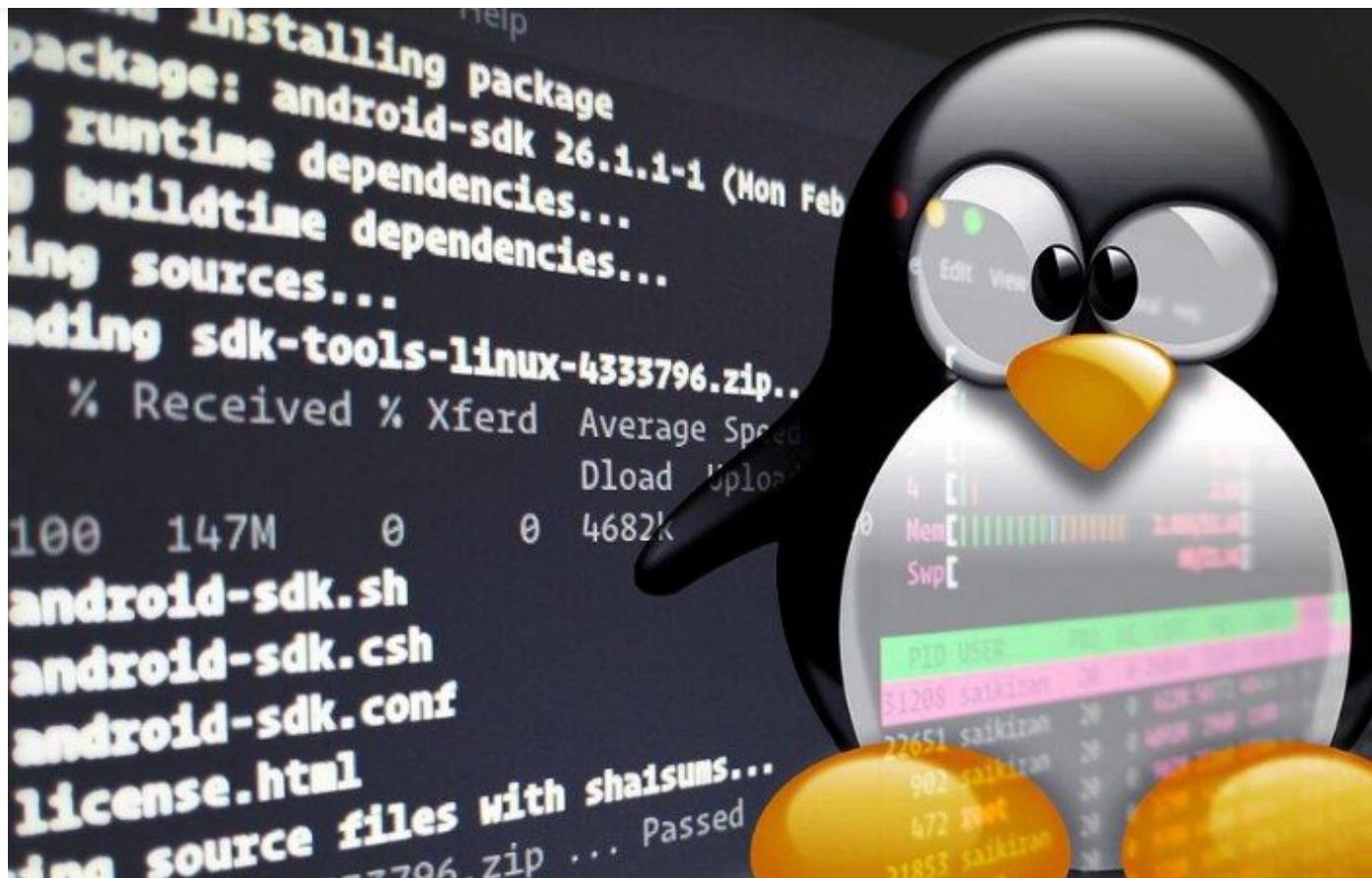
1. Programming Languages



Learn one / or two of the following Programming Languages (Python — Java — Scala)

- **Python:** <https://www.udemy.com/course/complete-python-developer-zero-to-mastery/>
- **Scala Beginners:** <https://www.udemy.com/course/rock-the-jvm-scala-for-beginners/>



[Open in app](#)[Get started](#)

- System Administration: <https://www.udemy.com/course/linux-administration-bootcamp/>
- Shell Scripting: <https://www.udemy.com/course/linux-shell-scripting-projects/>

3. Learn about Data Structures and Algorithms





Open in app

Get started

any shortest-length paths are get from your house to the cut shop?

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}$$

$$e^{i\pi} + 1 = 0$$

P	Q	R	$P \vee Q$	$P \vee R$	$(P \vee Q) \wedge (P \vee R)$
T	T	T	T	T	T
T	T	F	T	T	T
T	F	T	T	T	T
F	T	F	T	T	T
F	T	T	F	F	F
F	F	T	T	F	F
F	F	F	F	F	F

$\sum_{n=1}^{\infty} n^2 = 1^2 + 2^2 + 3^2 + \dots = \infty$

$S_n = 1^2 + 2^2 + 3^2 + \dots + n^2$

$\sum_{n=1}^{\infty} n^2 = 1^2 + 2^2 + 3^2 + \dots + n^2$

$2S_n = 1^2 + 2^2 + 3^2 + \dots + n^2 + (n+1)^2 + \dots + \infty$

$2S_n = 1^2 + 2^2 + 3^2 + \dots + n^2 + (n+1)^2 + \dots + \infty$

$S_n = \frac{1^2 + 2^2 + 3^2 + \dots + n^2}{2}$

$S_n = \frac{n(n+1)(2n+1)}{6}$

Original: $\exists x \forall y (x \geq 2y \rightarrow x)$

Converse: $\exists x \forall y (x > y + 1 \rightarrow x)$

Negation: $\neg [\exists x \forall y (\neg(x \geq 2y) \vee x)]$

$\forall x \exists y (x \geq 2y \wedge x \leq y + 1)$

Contrapositive: $\exists x \forall y (x \leq y + 1 \rightarrow x \geq 2y)$

$v - e + f = 2$

- <https://www.udemy.com/course/master-the-coding-interview-data-structures-algorithms/>

4. Learn about Core DBMS (Data Base Management Systems)



[Open in app](#)[Get started](#)

- <https://www.udacity.com/course/database-systems-concepts-design--ud150>

5. Learn SQL



- <https://www.udacity.com/course/sql-for-data-analysis--ud198>

6. Data Exploration Libraries (Pandas — NumPy — Spark)



[Open in app](#)[Get started](#)

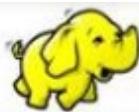
7. Data Warehousing Concepts



- General Concepts: <https://www.coursera.org/specializations/data-warehousing>
- Snowflake: <https://www.udemy.com/course/snowflake-masterclass/>

8. Learn about Hadoop ecosystem

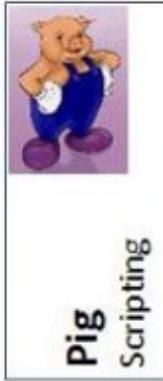
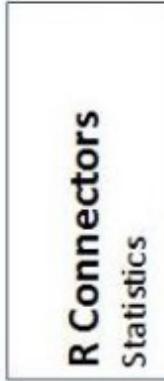
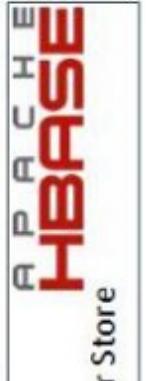


[Open in app](#)[Get started](#)

Apache Hadoop Ecosystem

**Ambari**

Provisioning, Managing and Monitoring Hadoop Clusters

**Sqoop**
Data Exchange**Zookeeper**
Coordination**Oozie**
Workflow**Pig**
Scripting**Mahout**
Machine Learning**R Connectors**
Statistics**Hive**
SQL Query**Hbase**
Columnar Store

YARN Map Reduce v2

Distributed Processing Framework

**HDFS**
Hadoop Distributed File System

- Hadoop Ecosystem: <https://www.udemy.com/course/the-ultimate-hands-on-hadoop-tame-your-big-data/>
- Spark (Scala): <https://www.udemy.com/course/spark-essentials/>
- Spark Streaming: <https://www.udemy.com/course/spark-streaming/>

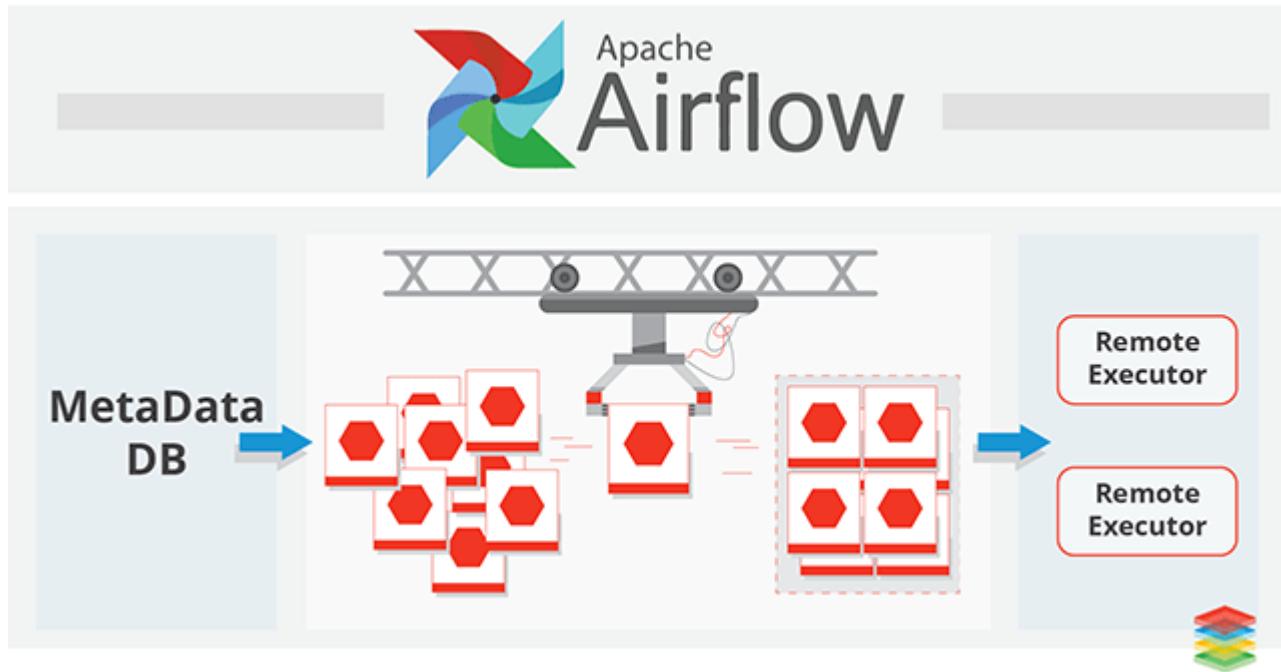
9. Workflow schedulers





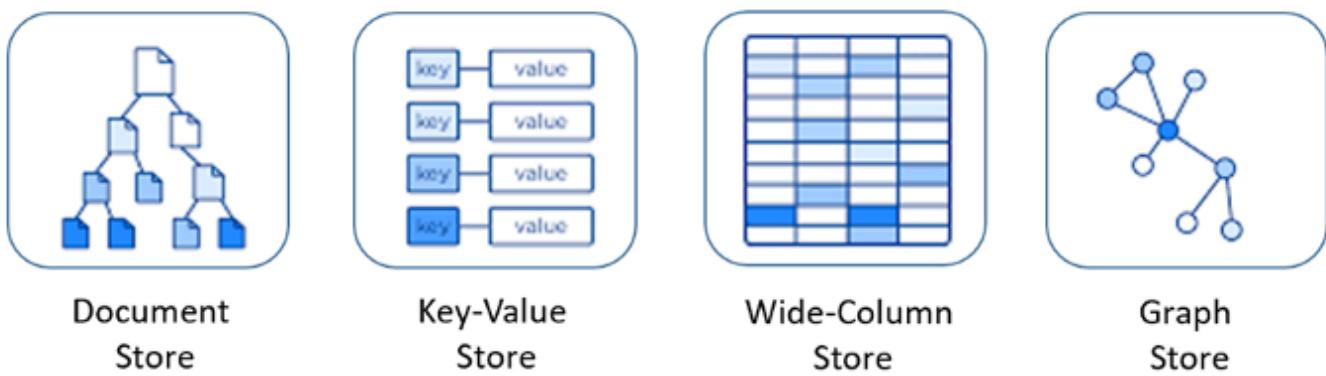
Open in app

Get started



- Apache Airflow: <https://www.udemy.com/course/the-ultimate-hands-on-course-to-master-apache-airflow/>

10. NoSQL Databases



- Read this article: <https://www.journaldunet.com/solutions/dsi/1194284-base-nosql-laquelle-choisir-pour-quels-besoins/>
- Cassandra: <https://www.udemy.com/course/learn-cassandra-from-scratch/>

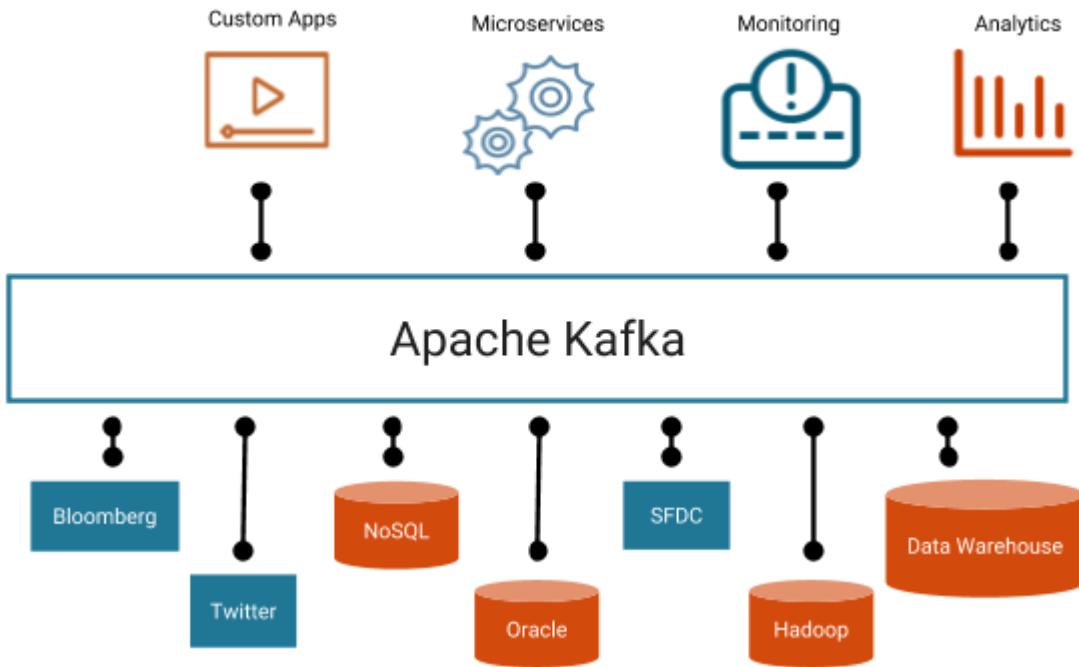




Open in app

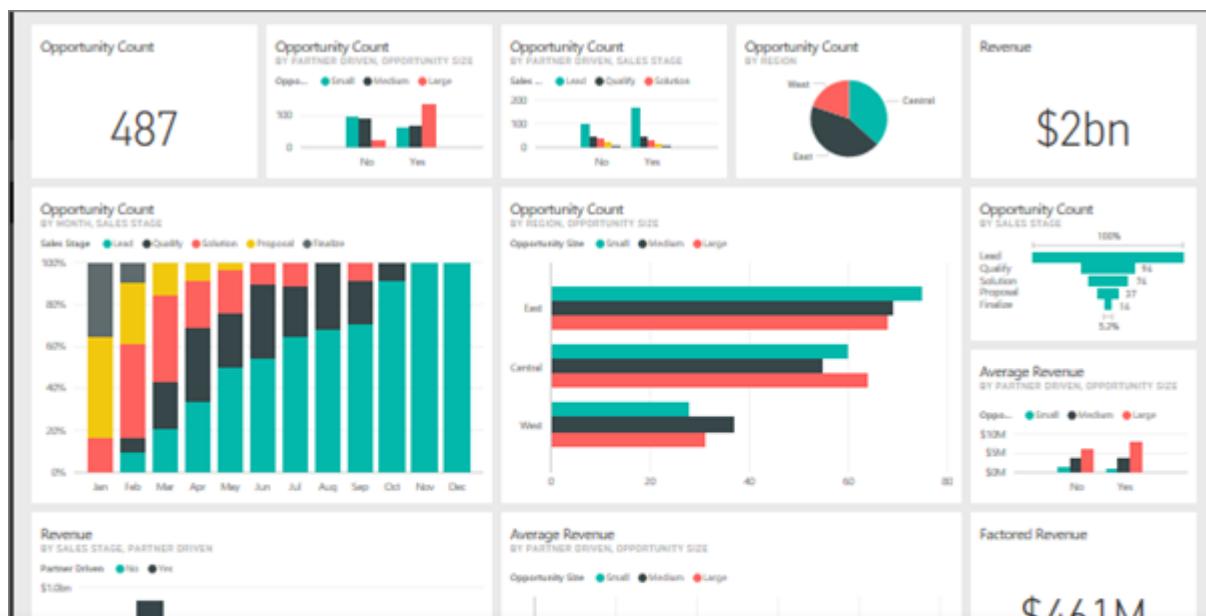
Get started

11. MQ (Streaming frameworks)



- Apache KAFKA: <https://www.udemy.com/course/apache-kafka/>

12. Dashboarding tools



[Open in app](#)[Get started](#)

- Kibana: Look at ElasticSearch part
- Tableau: <https://www.udemy.com/course/tableau-10-u/>
- Python: <https://www.udemy.com/course/learning-python-for-data-analysis-and-visualization/>
- Matplotlib
- Seaborn

13. Cloud Computing:



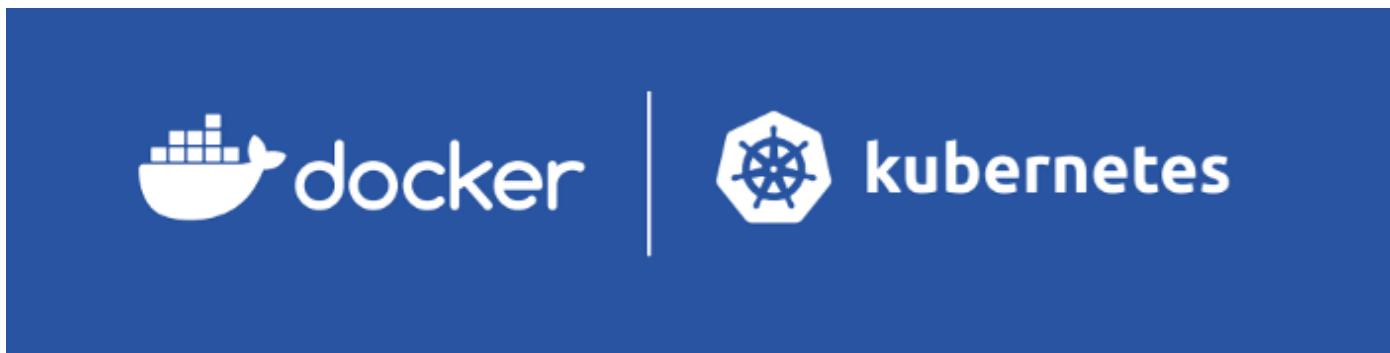
- AWS: <https://www.coursera.org/learn/aws-cloud-technical-essentials>
- Google: <https://www.coursera.org/learn/gcp-foundamentals>



[Open in app](#)[Get started](#)

- AWS: <https://www.udemy.com/course/aws-data-analytics/>
- Google: <https://www.coursera.org/professional-certificates/gcp-data-engineering>
- Azure: <https://www.coursera.org/specializations/microsoft-azure-dp-203-data-engineering?>

15. DevOps (Docker — Kubernetes)



- Docker: <https://www.udemy.com/course/learn-docker/>
- Kubernetes: <https://www.udemy.com/course/learn-kubernetes/>

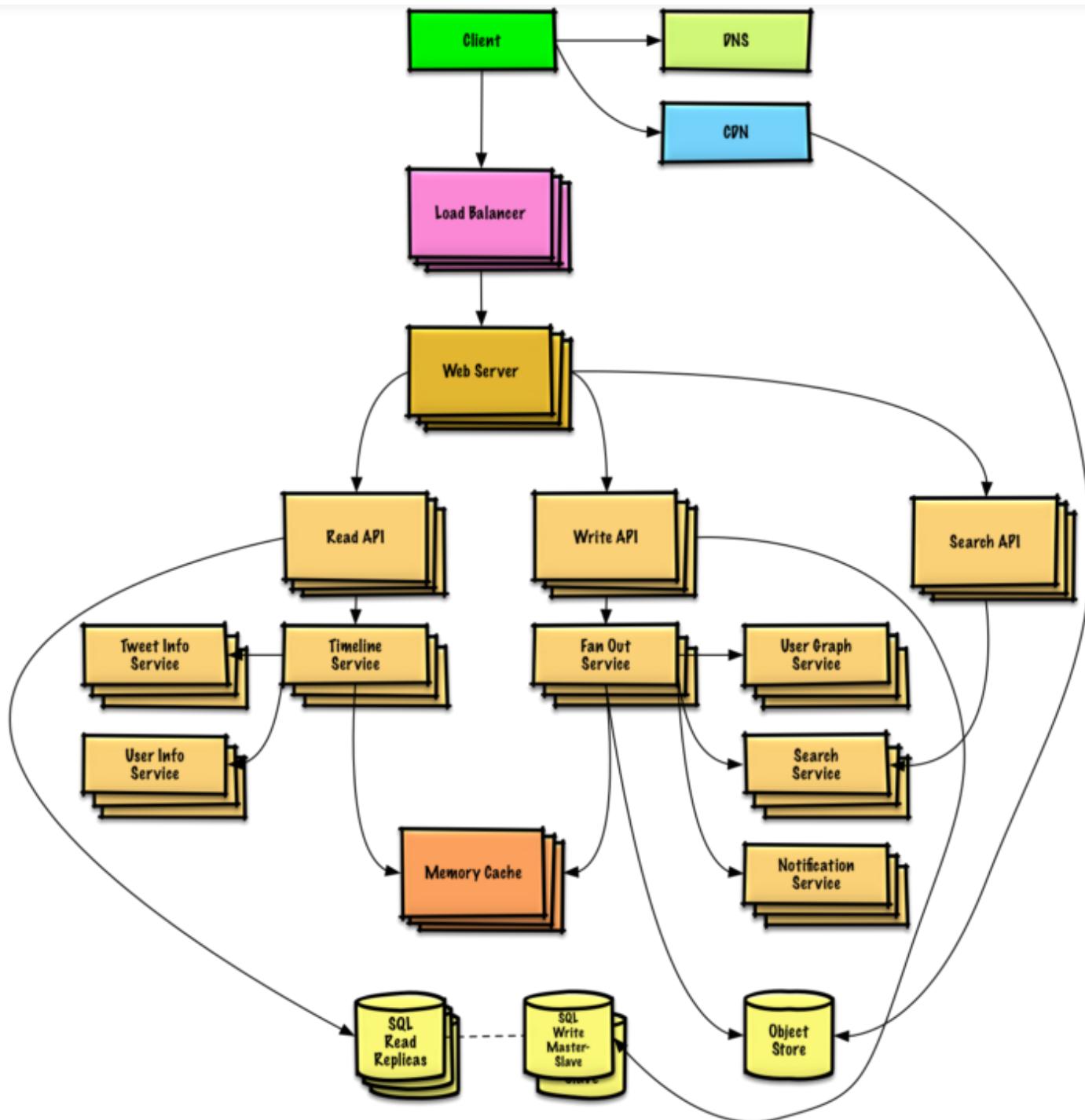
16. System Design





Open in app

Get started



- <https://www.udemy.com/course/system-design-interview-prep/>

About me



[Open in app](#)[Get started](#)

[Let's connect on LinkedIn](#). I work as a Lead Cloud Data Engineer for [DataScientest.com](#), a leading Data-driven French EdTech Startup specialized in Data and Cloud Training. I am a Certified Data Engineer on the Top 3 leading Cloud Platforms: AWS, Google, Microsoft Azure. In addition, I hold 2x Kubernetes and 6x additional AWS certifications. I happen to suffer from a severe physical disability tying me to the use of an Electric Wheelchair. I am a Disability Rights advocate and a happily married father of two boys.

Get an email whenever Chouaieb Nemri publishes.



[Open in app](#)[Get started](#)

By signing up, you will create a Medium account if you don't already have one. Review our [Privacy Policy](#) for more information about our privacy practices.

[About](#) [Help](#) [Terms](#) [Privacy](#)

Get the Medium app

