- ✓ Database Partitioning
- ✓ Database Sharding

SQL vs NoSQL

# SQL

- relational
- tables
- structured

# No-SQL

→ unstructured data can be stored

e.g. MongoDB

No-SQL

↳ document based,
key-value

↳ flexibility

SQL → vertically scalable

No-SQL → horizontally scalable

# No-SQL

## Strengths

1.) **flexible** → no data relationships (no tables)

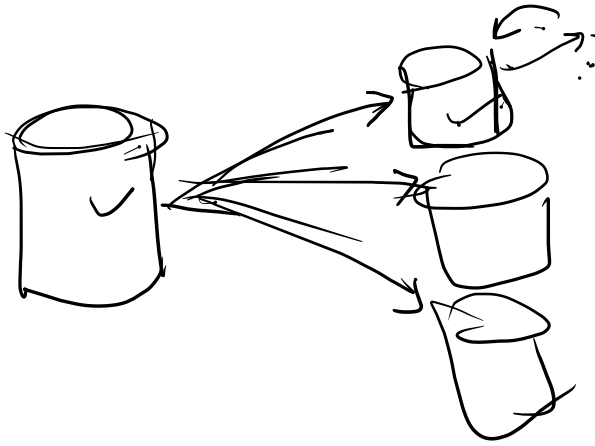
↳ better choice for storing unstructured data.

2.) **Data sharding**

↳ unstructured data can be divided across multiple servers.

→ Horizontal scaling is easier

# Weaknesses



1.) loss of consistency

Data sharding

↳ there can be delay for update propagation in replicas

* Eventual consistency

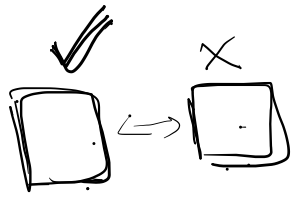
# Examples

1. MongoDB
2. Cassandra
3. Redis
4. neo4j
5. Hbase

# CAP Theorem

→ trade-offs while designing shared-data system

# CAP




**Consistency**
(among replicated copies)

**Availability**
(for read/write operations)

**Partition tolerance**
(in case of node fault)

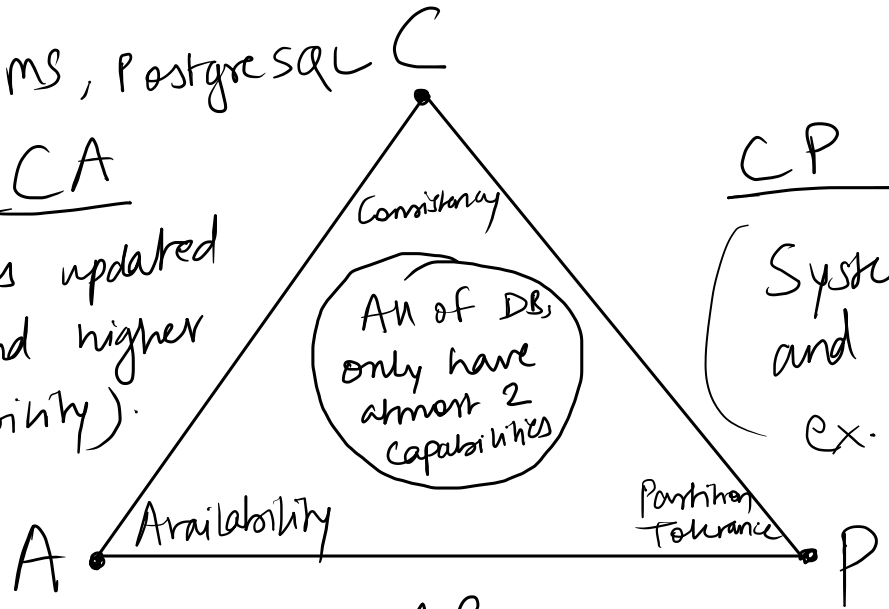
(ANY 2)

* CAP theorem states

It is not possible to guarantee all three of the desirable properties at the same time in a Distr. system with data replication.

→ shared data system can only strongly support 2 out of 3 properties

# CAP Theorem




ex RDBMS, Postgresql

## CA

(Promises updated data and higher availability).

## CP

(System can be distributed and promise updated data)

ex. mongoDB, HBase

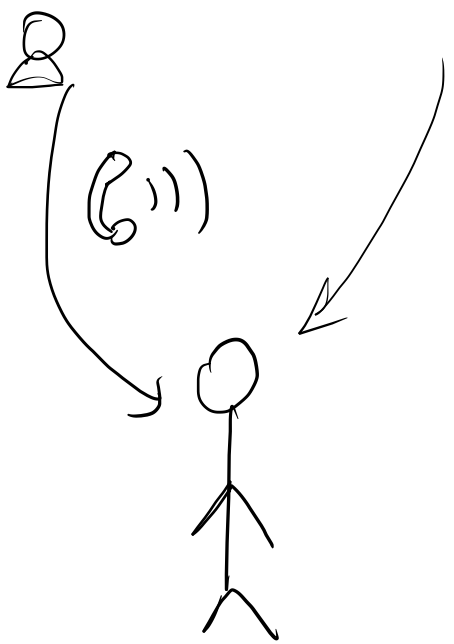
## AP

(System can be distributed and promise higher availability)

ex. DynamoDB, cassandra

Next : Realistic example – CAP theorem

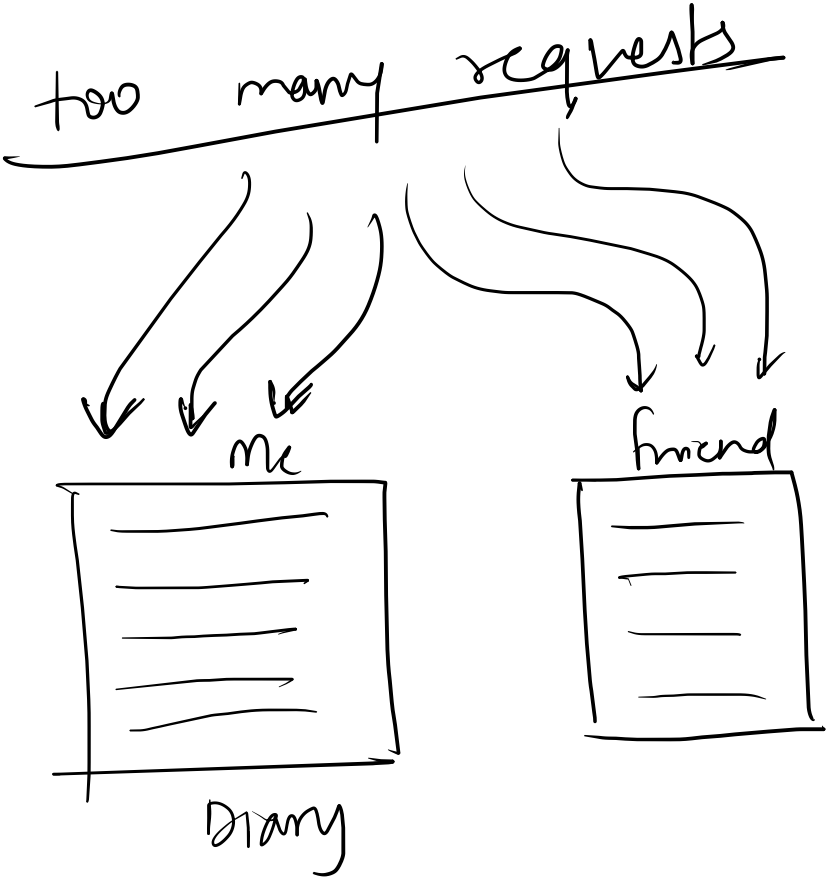
# Reminder Service

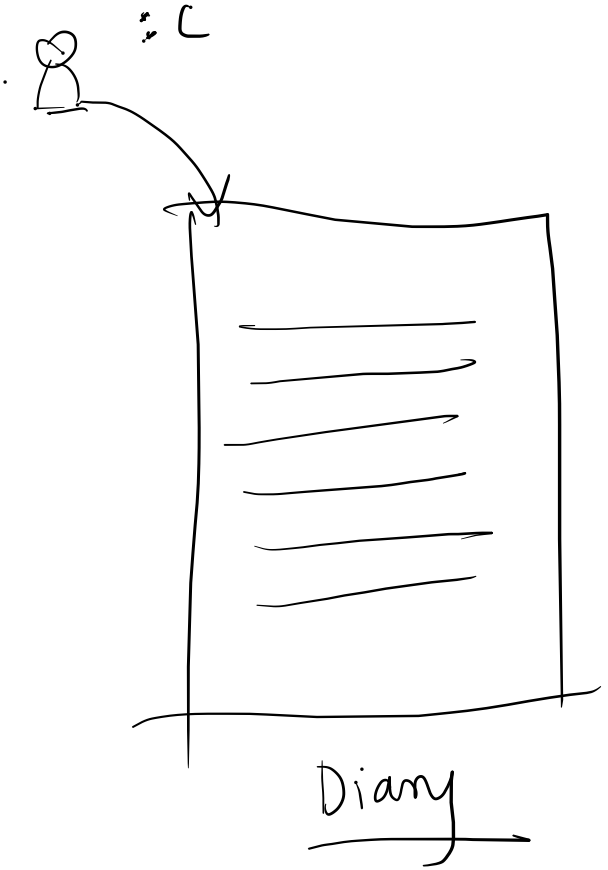


1.) meeting at 10 am

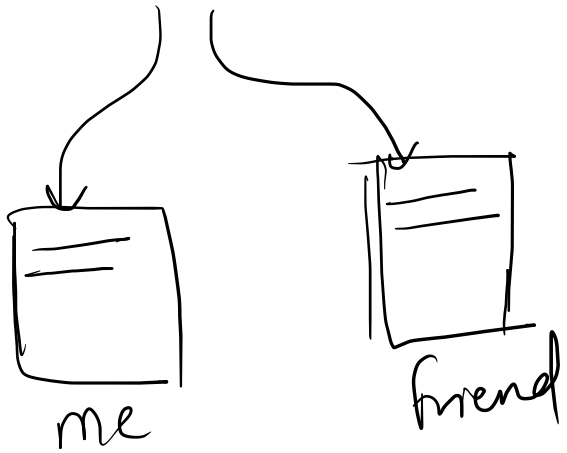
2.) Flight at 7pm

# Scale it



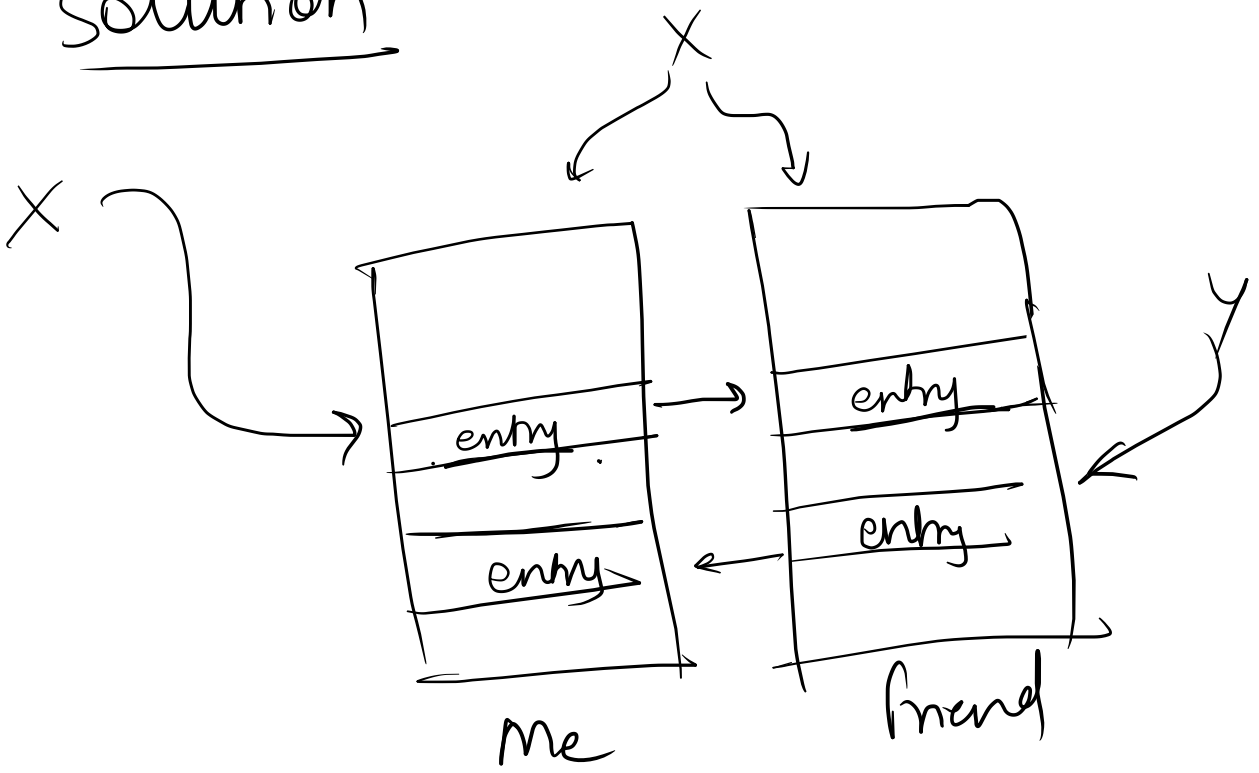


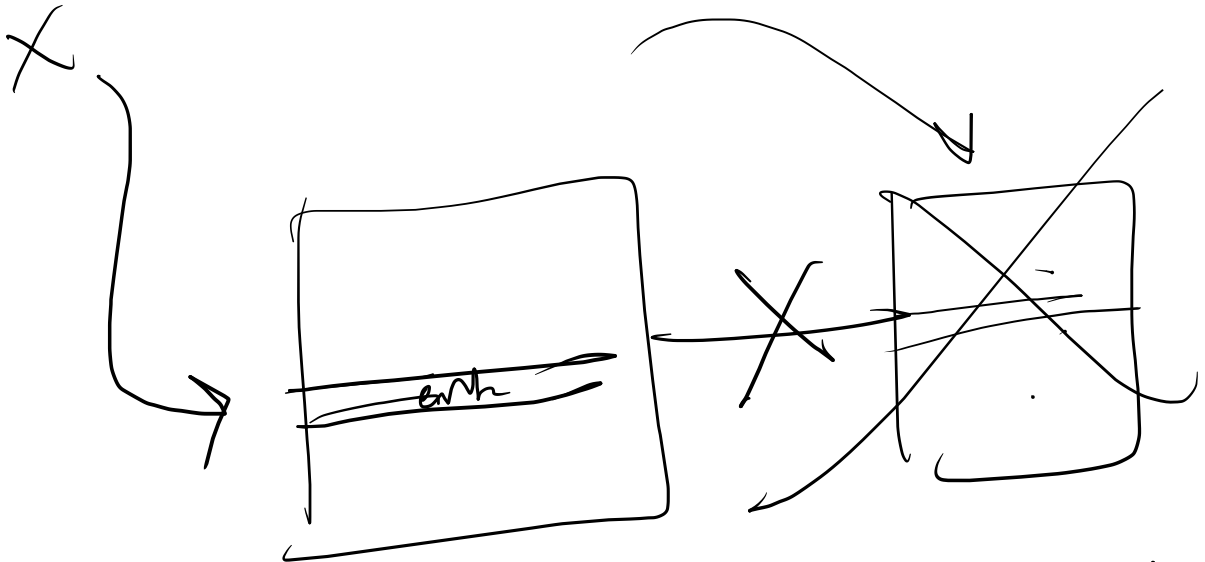

fault ?



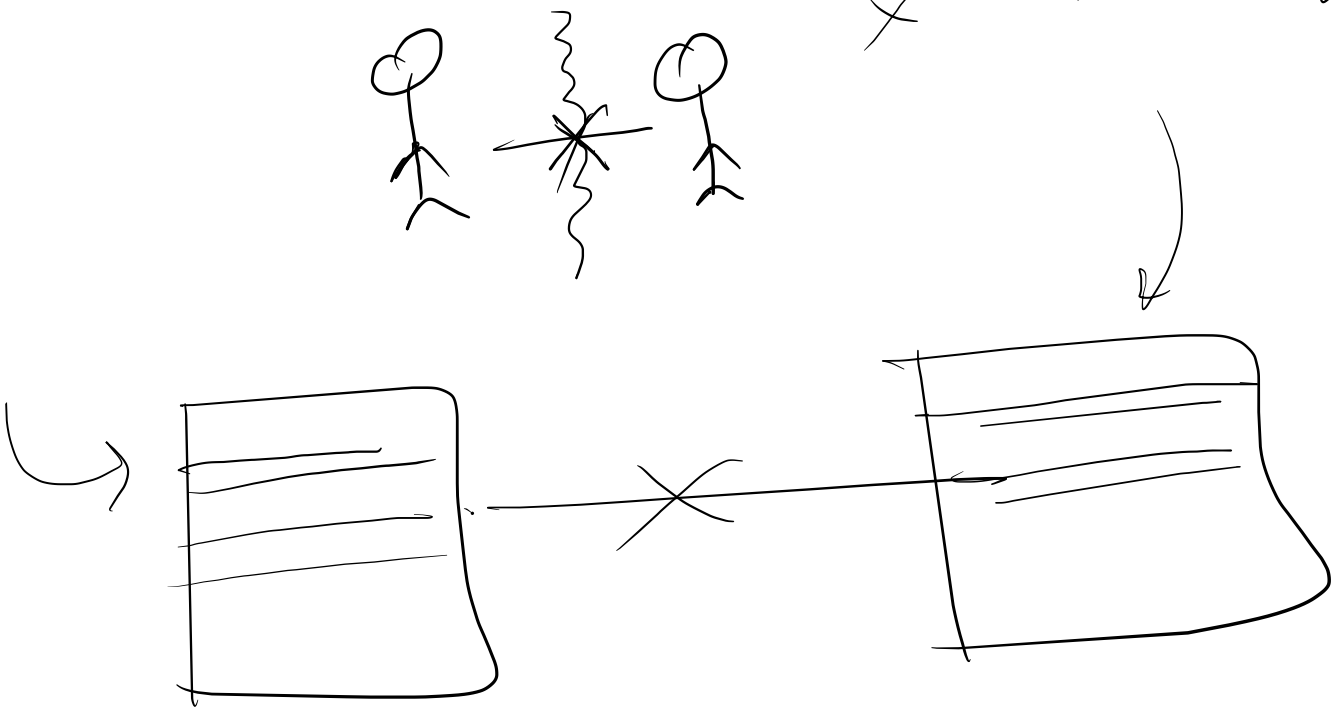

Data consistency problem

# Solution



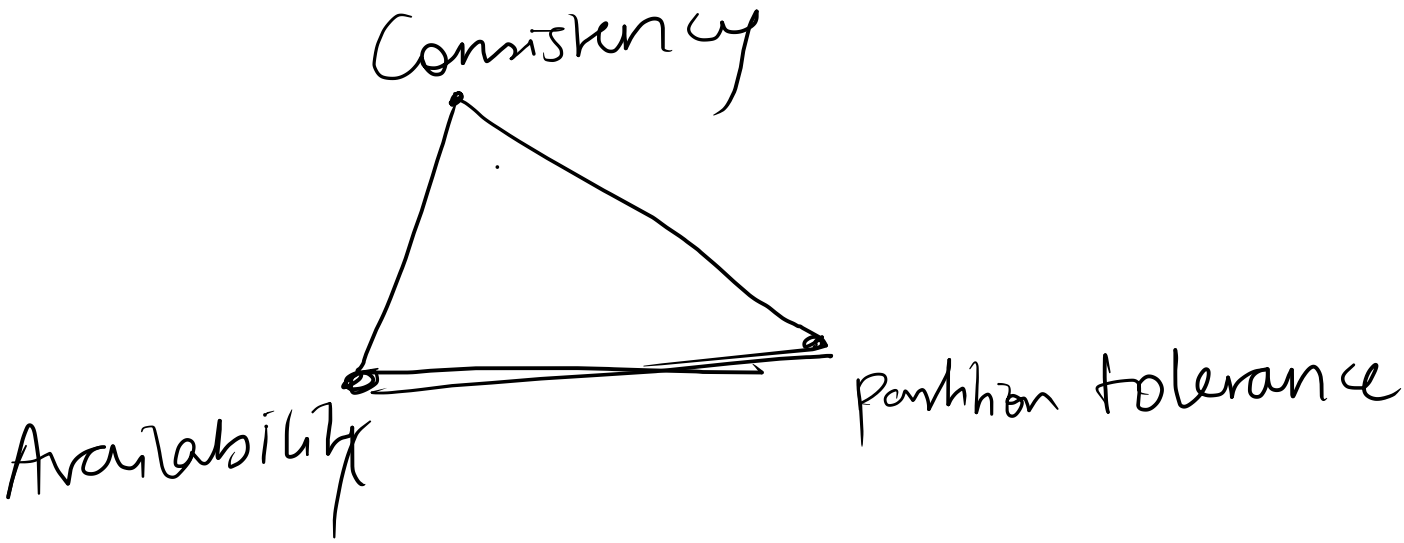

earth
friend
Availability problem

Choose between consistency / availability

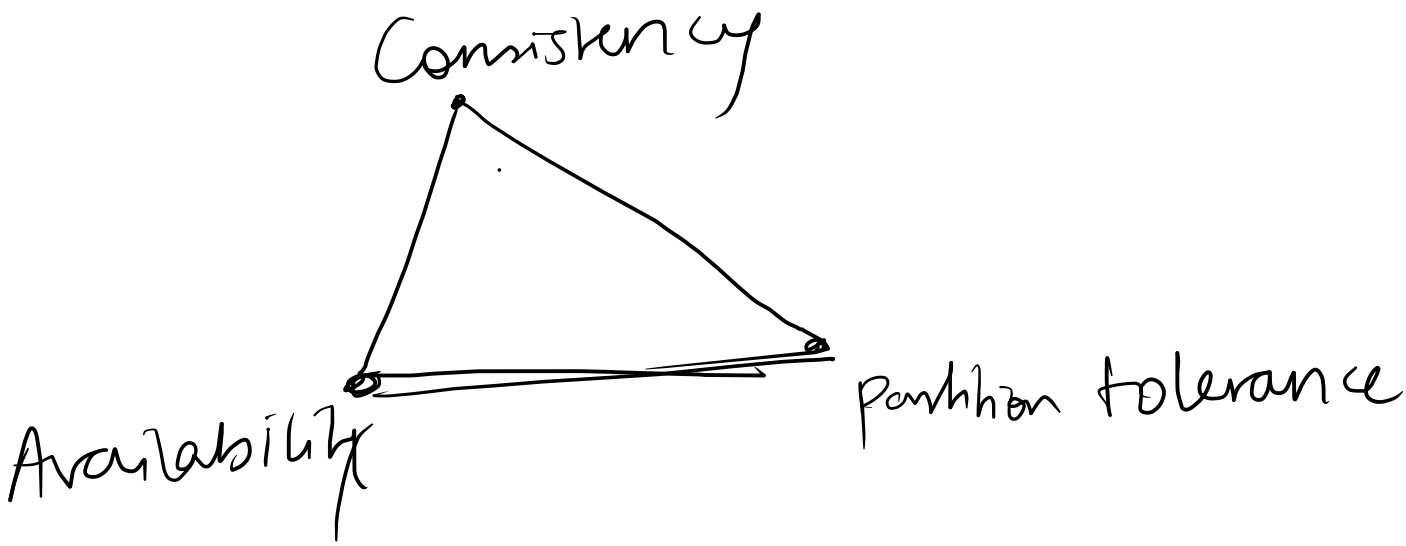


network partition

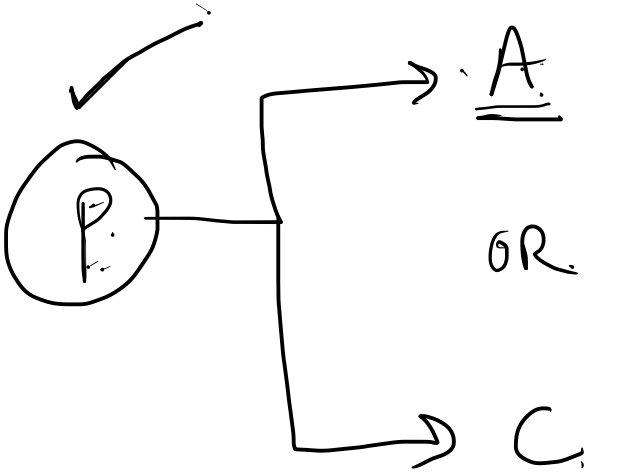
# Choose any 2



# Choose any 2



CAP → extension → PACELC



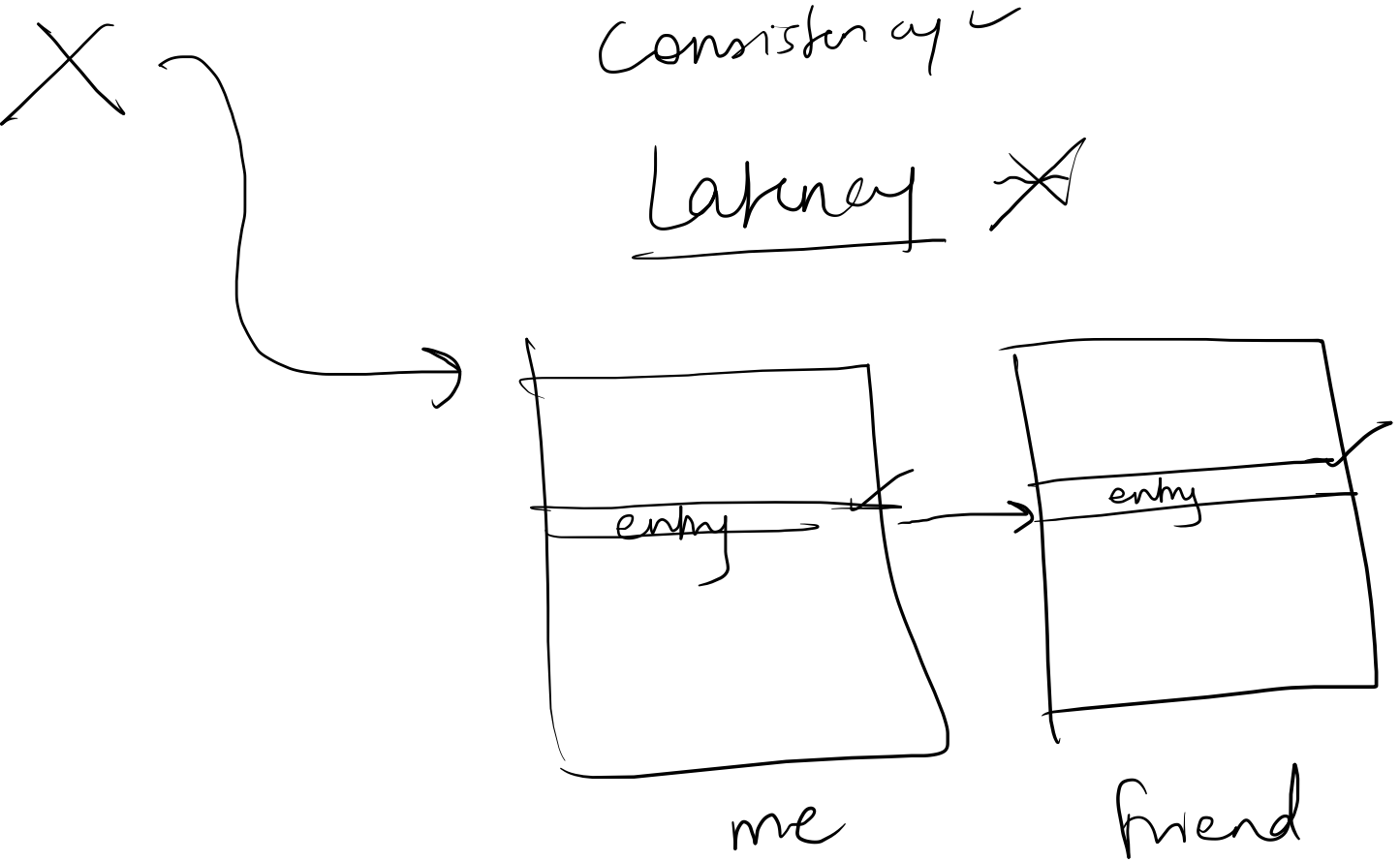

Else

×

tradeoff

Latency

consistency


Consistency
Latency
entry
entry
me
friend