Let’s roll by 09:05 !!!

System Design Fundamentals

Lets learn how Big Billion Days works smoothly!

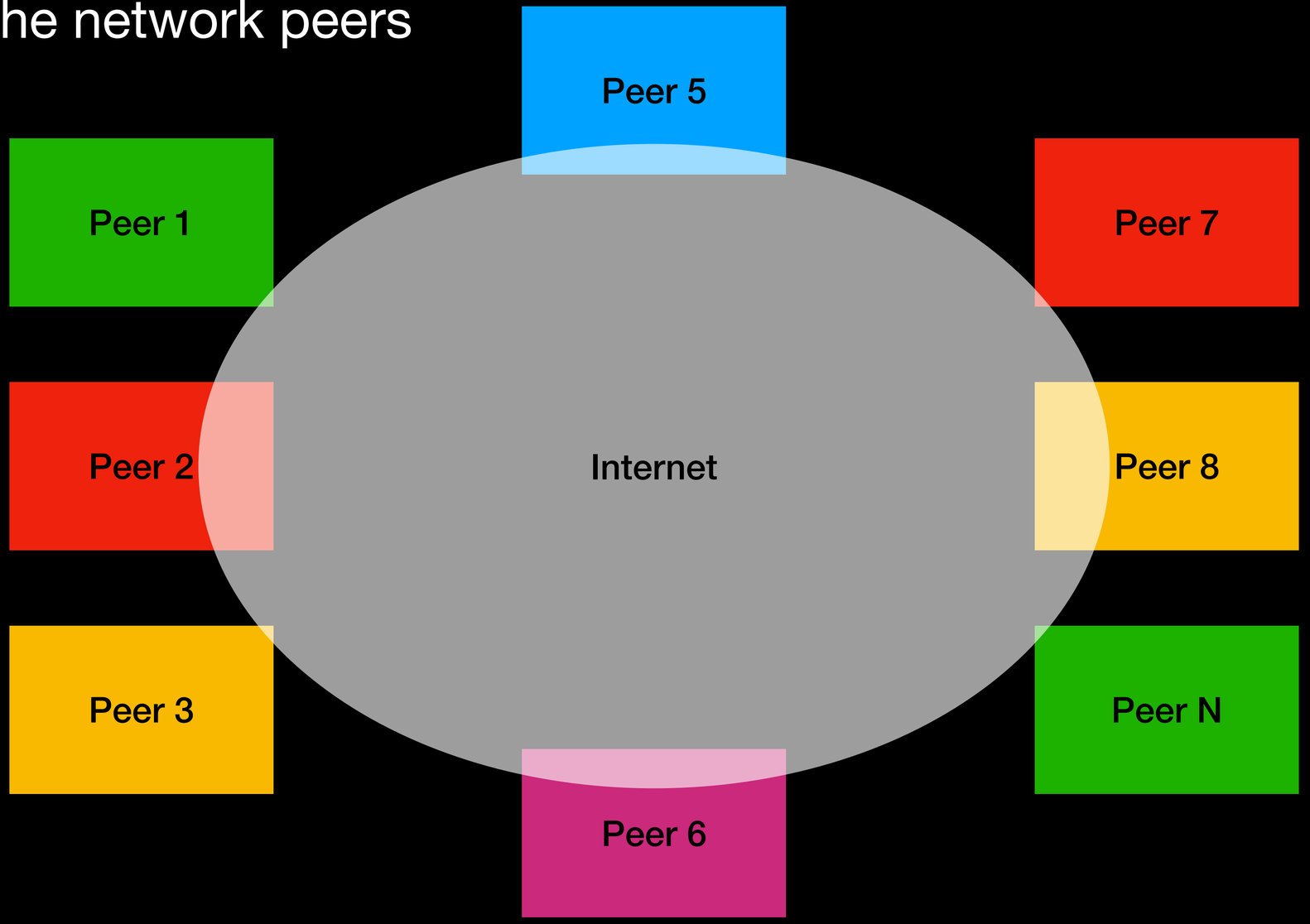
Gopi Krishnan R

Topics to be Discussed

- Internet
- What is SYSTEM?
- Why System designing?
- Distributed System
    - Scalability
    - Reliability
    - Availability
    - Efficiency
    - Maintainability
- Load Balancing
- CAP Theorem
- Stateless vz Stateful System

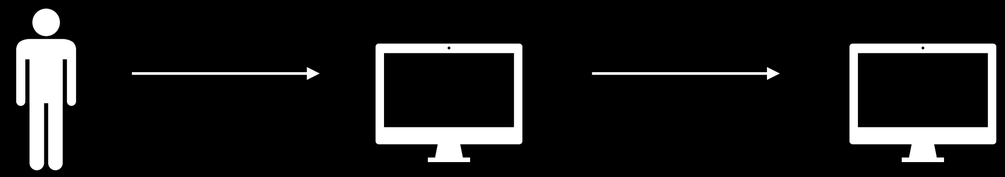
Internet

- Communicate with the network peers

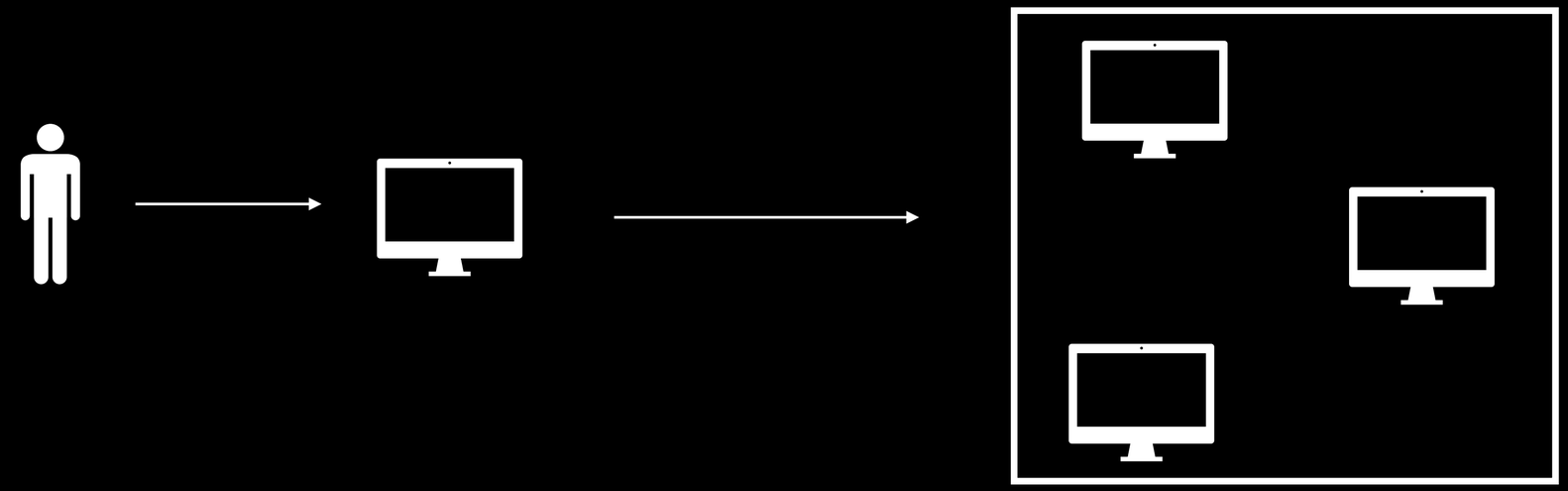


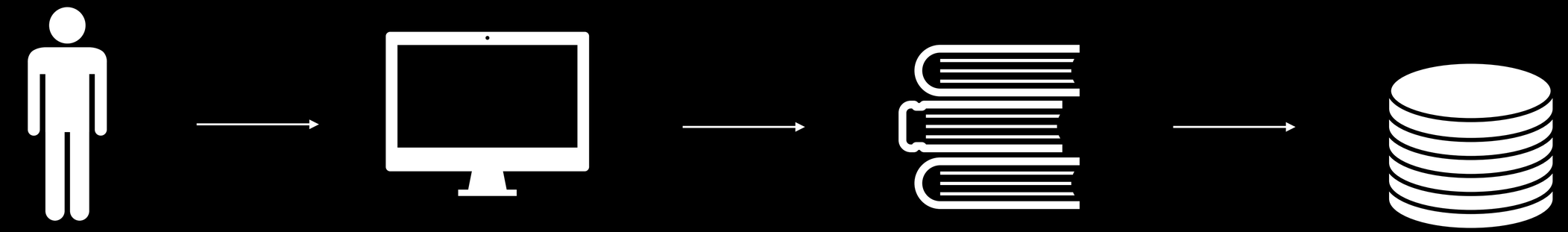
What is a System?

Centralised System



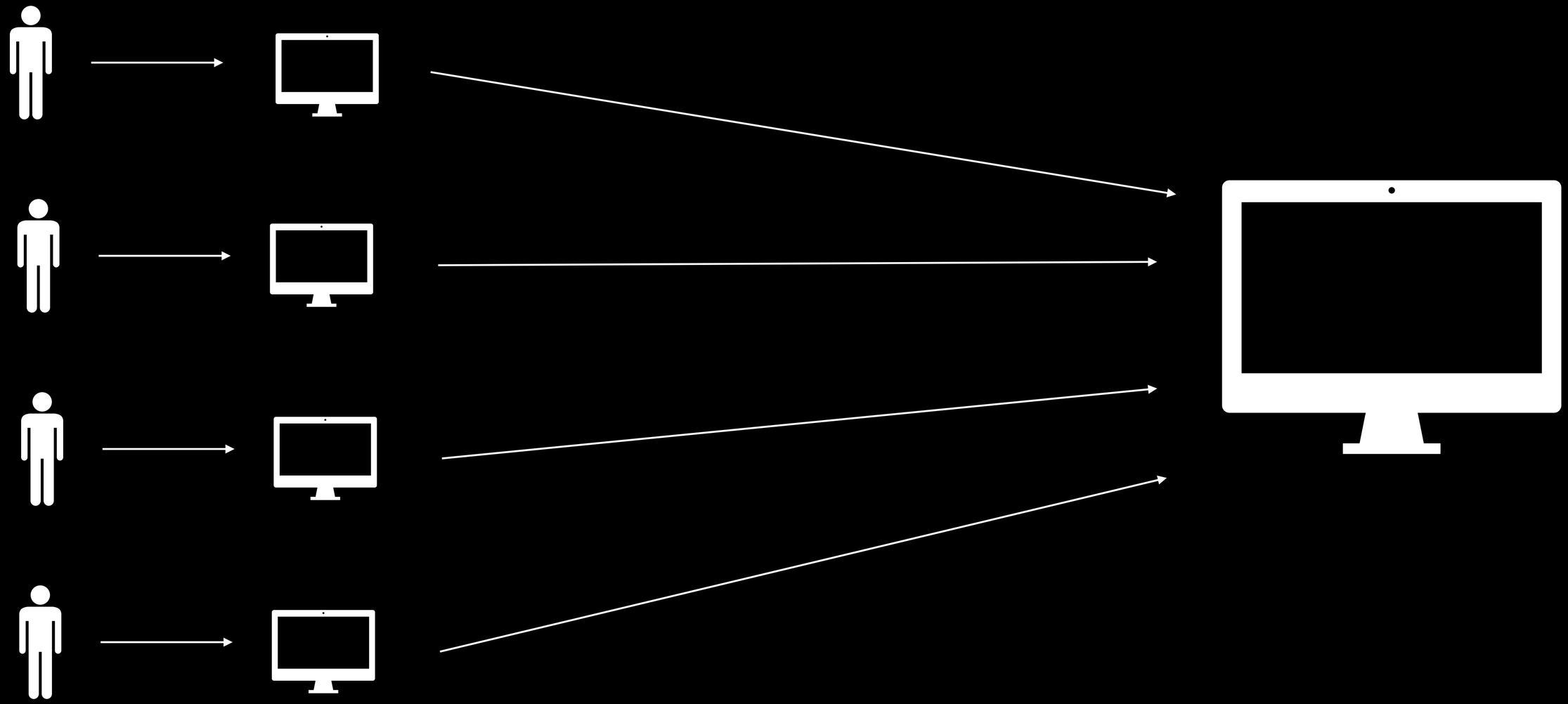
Distributed System





Client Front End Back End Database

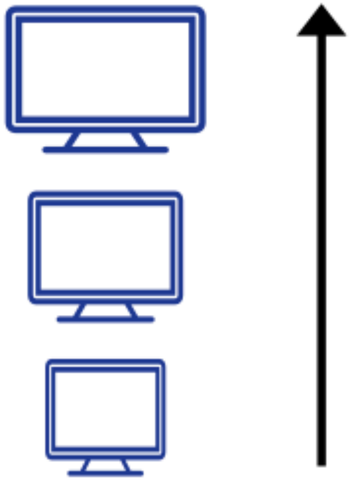
How Centralised System work during BBS?



Single point of failure

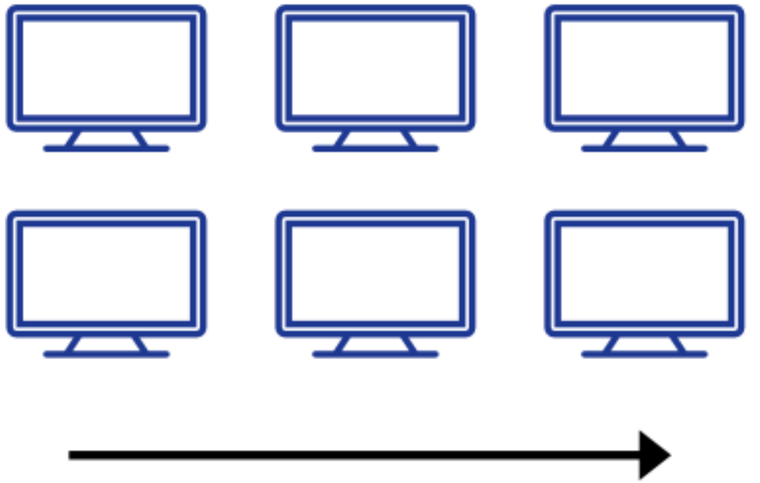
VERTICAL SCALING

Increase size of instance
(RAM, CPU etc.)

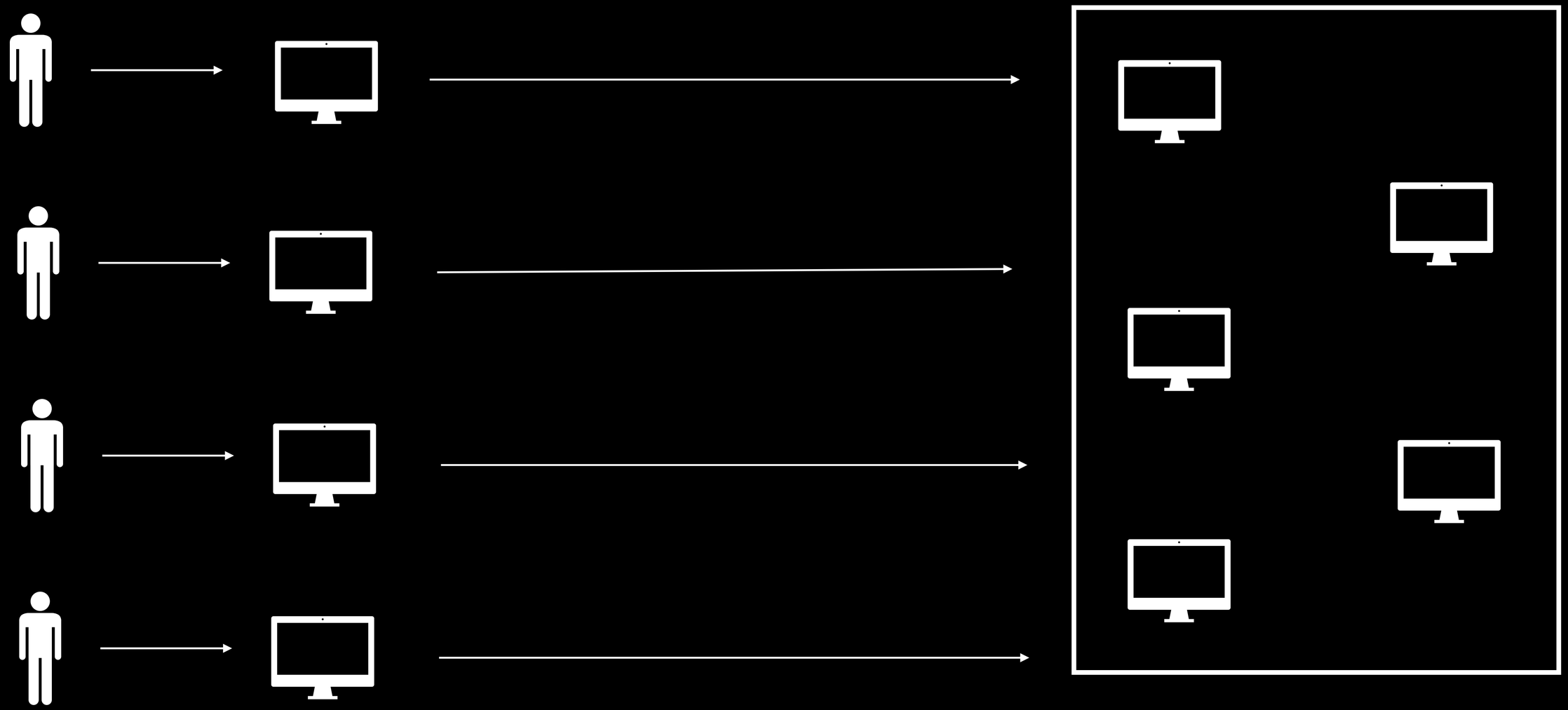


HORIZONTAL SCALING

(Add more instances)



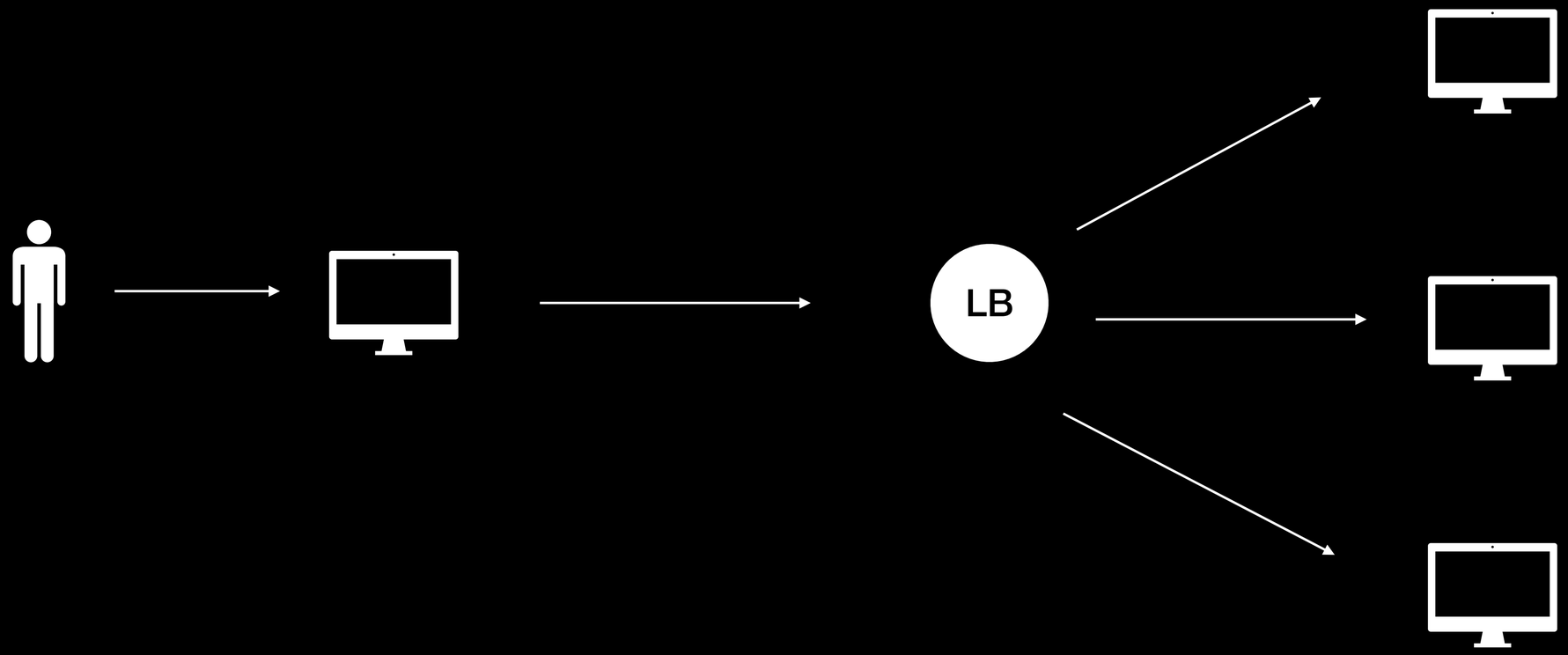
How Distributed System work during BBS?



Which one to take up?

Load Balancer

- Component that acts and distributes network or application traffic across a number of servers to improve responsiveness and availability of applications, websites or databases.



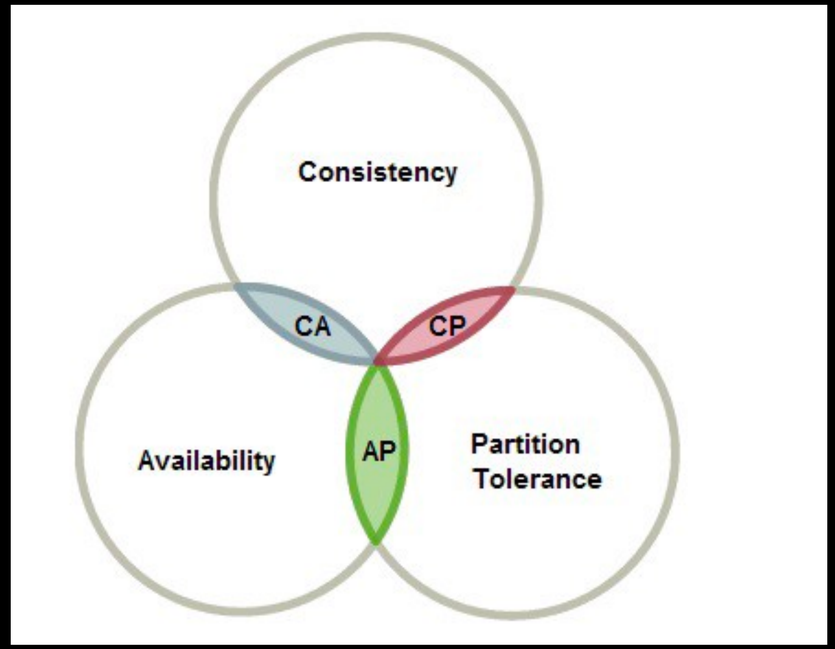
Characteristics of DS

- **Stability** or **Serviceability** or **Manageability** = simplicity and speed with which a system can be repaired or maintained.
- **Reliability** = *probability a system will fail in a given period*. A distributed system is considered reliable if it keeps delivering its services even when one or several of its software or hardware components fail.
- **Scalability** = capability of a system, process, or a network to grow and manage increased demand.
- **Efficiency/Performance** = *response time* (or latency = delay to obtain the first item) + *throughput* (or bandwidth = the number of items delivered in a given time unit)
- **Availability** = simple measure of the percentage of time that a system, service, or a machine remains operational under normal conditions.
- Fault tolerance and Catastrophe-preparedness

CAP Theorem

Consistency - Availability - Partition Tolerance

CAP Theorem is a concept that a distributed database system can only have 2 of the 3: Consistency, Availability and Partition Tolerance.



Consistency

Consistency means that all clients see the same data at the same time, no matter which node they connect to. For this to happen, whenever data is written to one node, it must be instantly forwarded or replicated to all the other nodes in the system before the write is deemed ‘successful.’

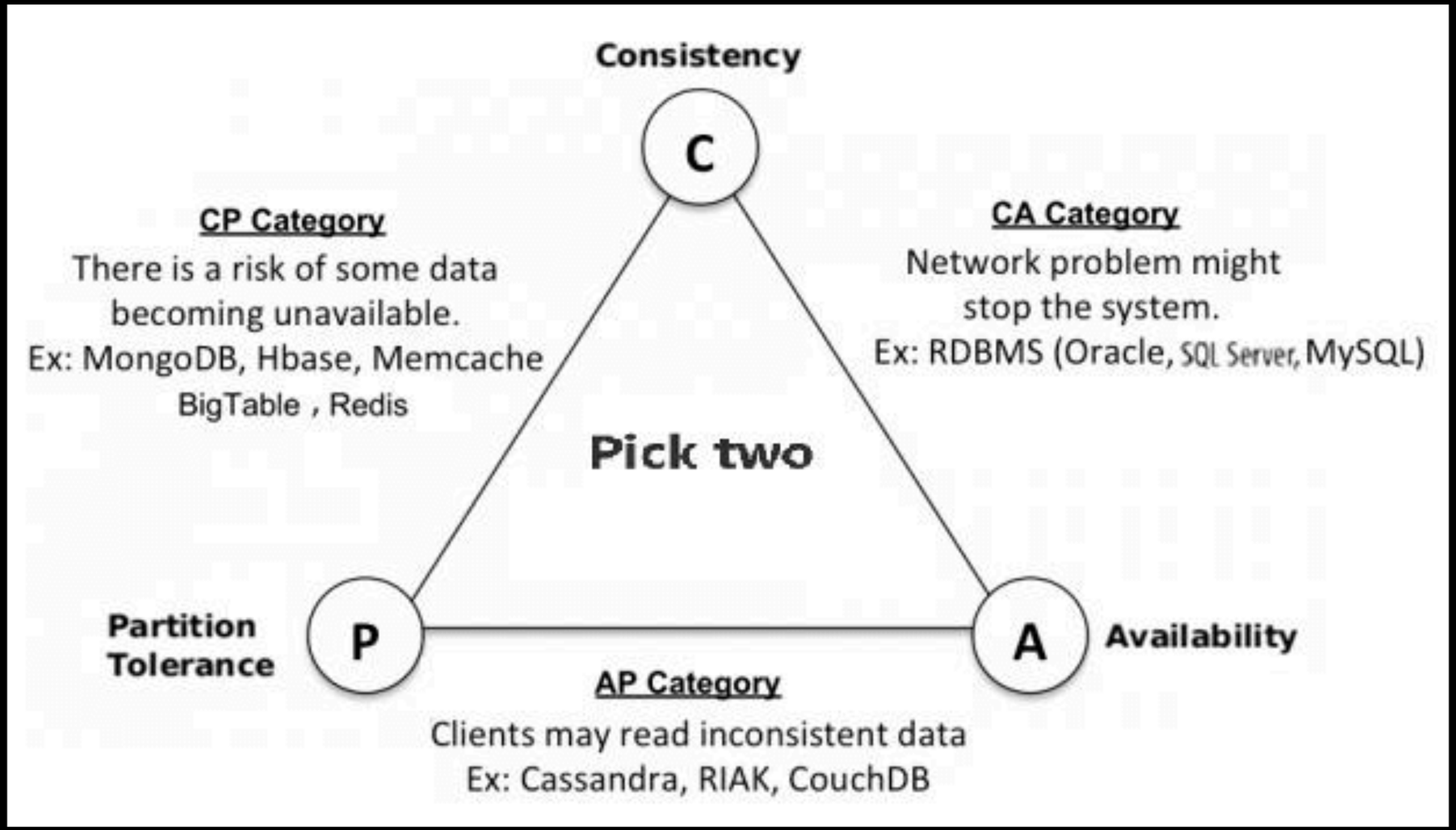
Availability

Availability means that that any client making a request for data gets a response, even if one or more nodes are down. Another way to state this—all working nodes in the distributed system return a valid response for any request, without exception.

Partition Tolerance

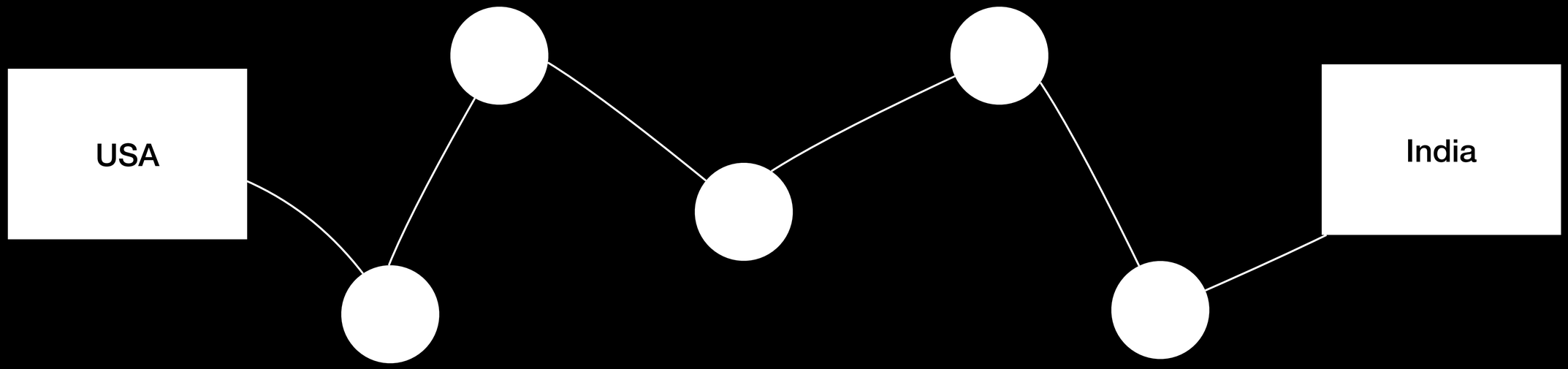
A *partition* is a communications break within a distributed system—a lost or temporarily delayed connection between two nodes. Partition tolerance means that the cluster must continue to work despite any number of communication breakdowns between nodes in the system.

CAP Theorem - In Depth



Latency

In computer networking, latency is an expression of how much time it takes for a data packet to travel from one designated point to another. Ideally, latency will be as close to zero as possible. Network latency can be measured by determining the round-trip time (RTT) for a packet of data to travel to a destination and back again



Stateful vz Stateless System

- Stateful System —> records or stores user data

- Stateless System —> no user data is recored or stored