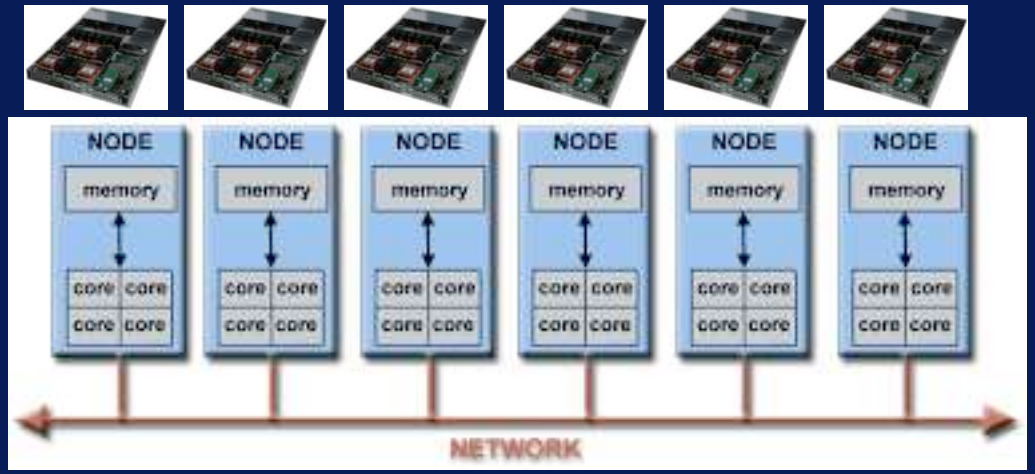
# Scalable Computing over the Internet

# After this video you will be able to..

- Use the terms “commodity cluster” in your conversations with your IT team
- Draw a simple commodity cluster architecture
- Explain data parallelism and why it is a useful concept to know
- Describe common failures of commodity clusters and name two solutions to them

# Single Compute Node

# Parallel Computer



SDSC Gordon

**Expensive**

# Commodity Cluster

# Commodity Cluster

**Affordable**

**Less-specialized**

# Commodity Cluster

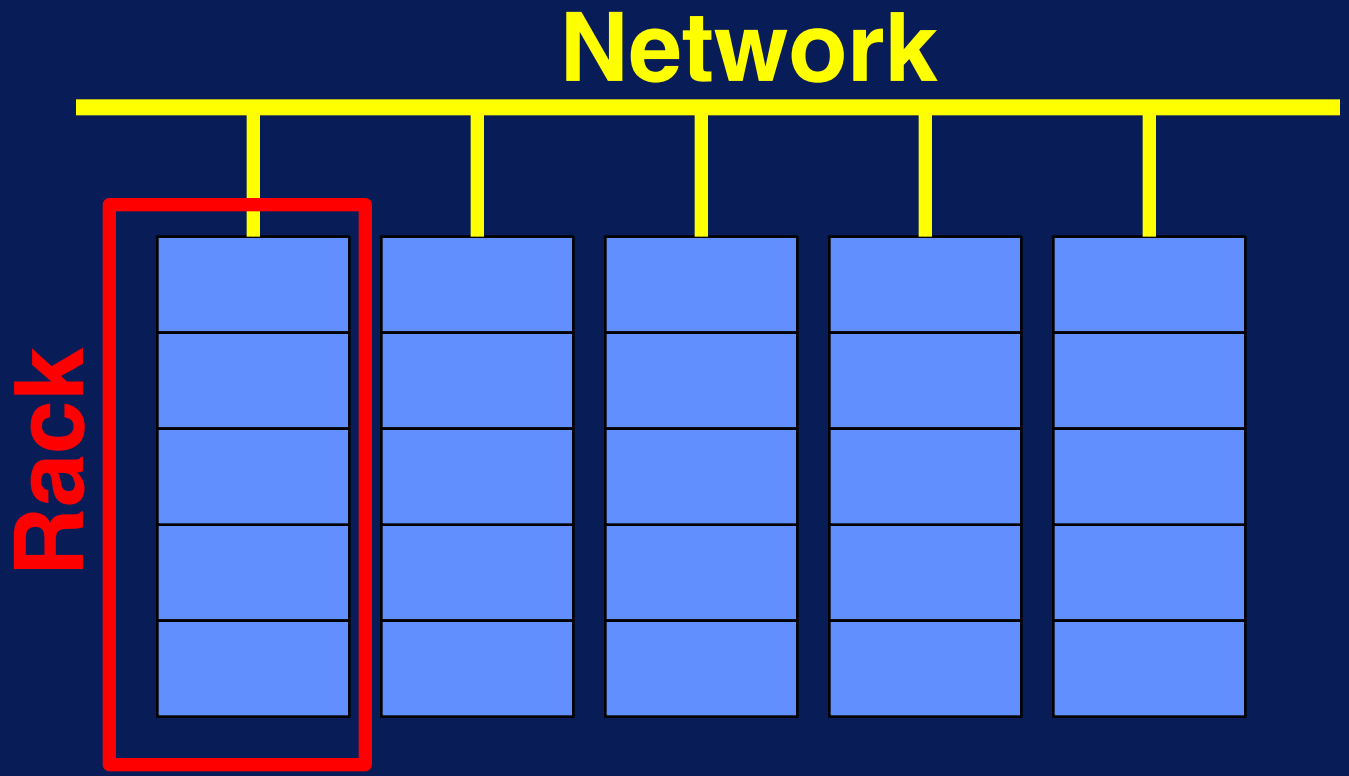
“Distributed computing“
over the Internet

# Commodity Cluster

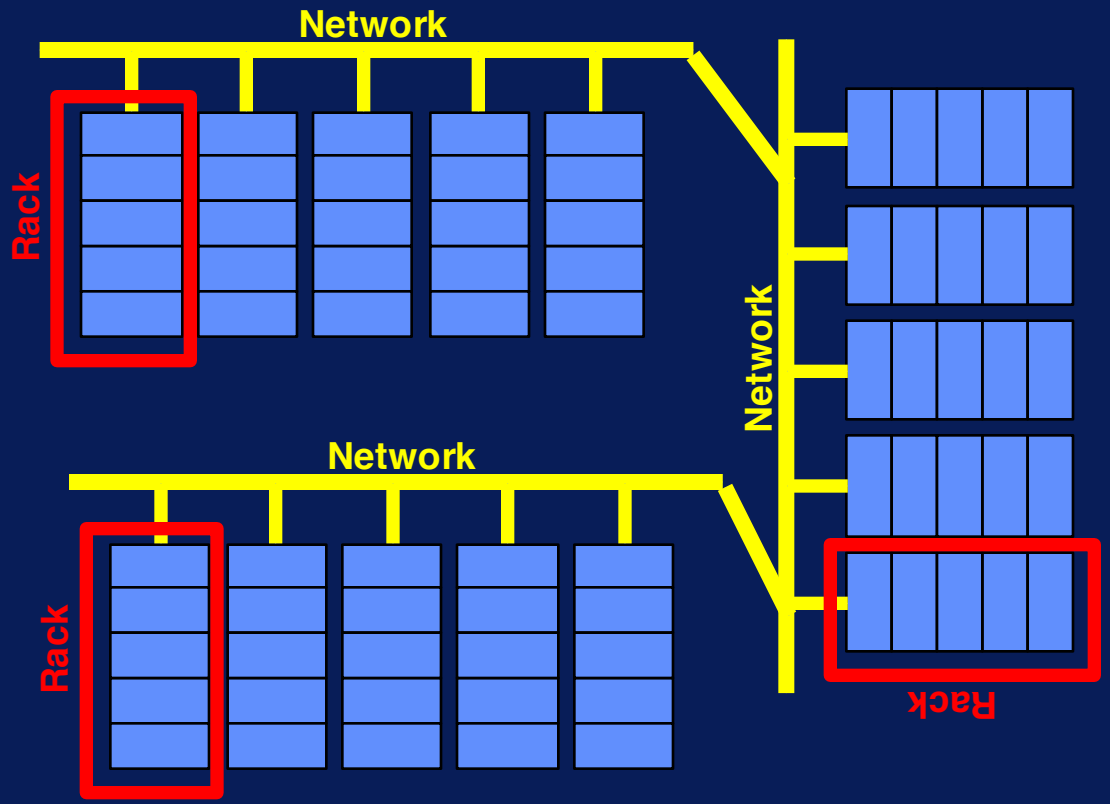
Reduced computing cost

“Distributed computing“ over the Internet

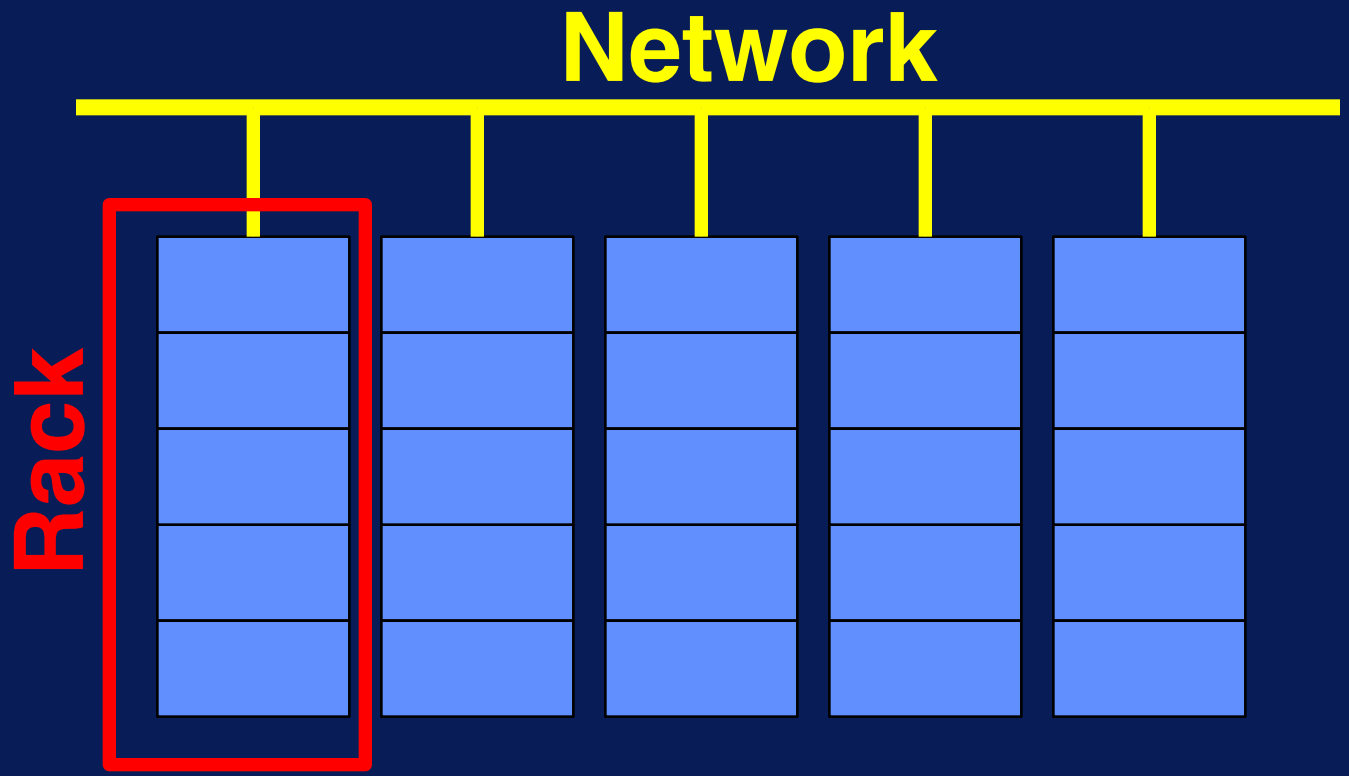
# Architecture of a Commodity Cluster



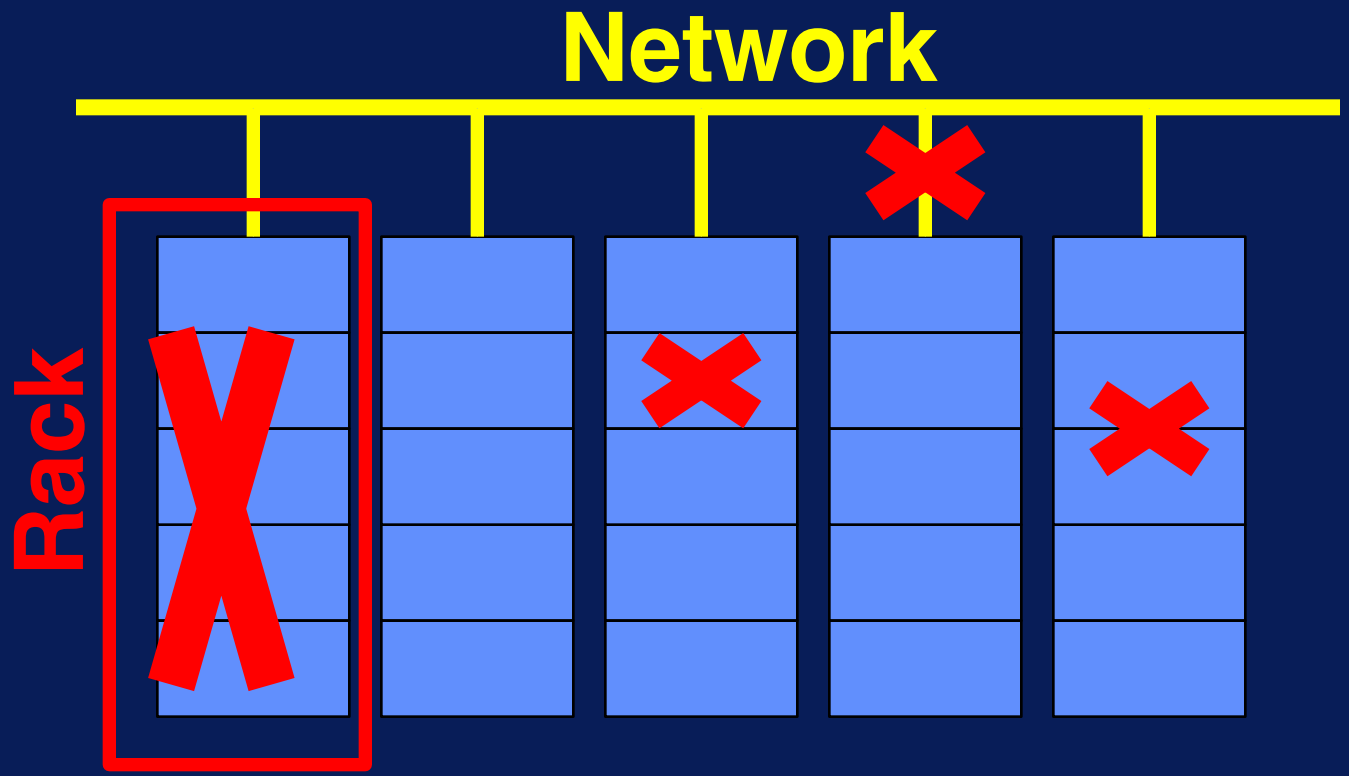
# Distributed Computing



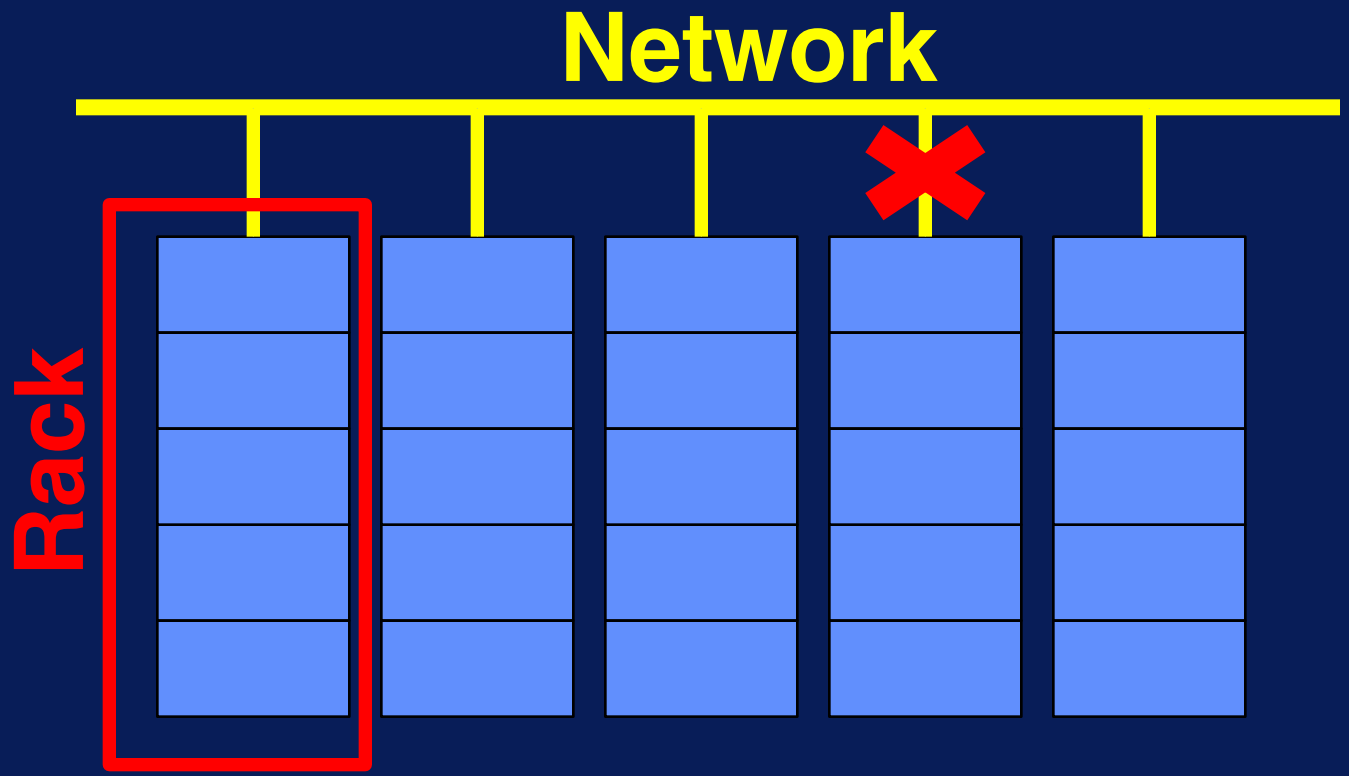
# Enables data-parallelism



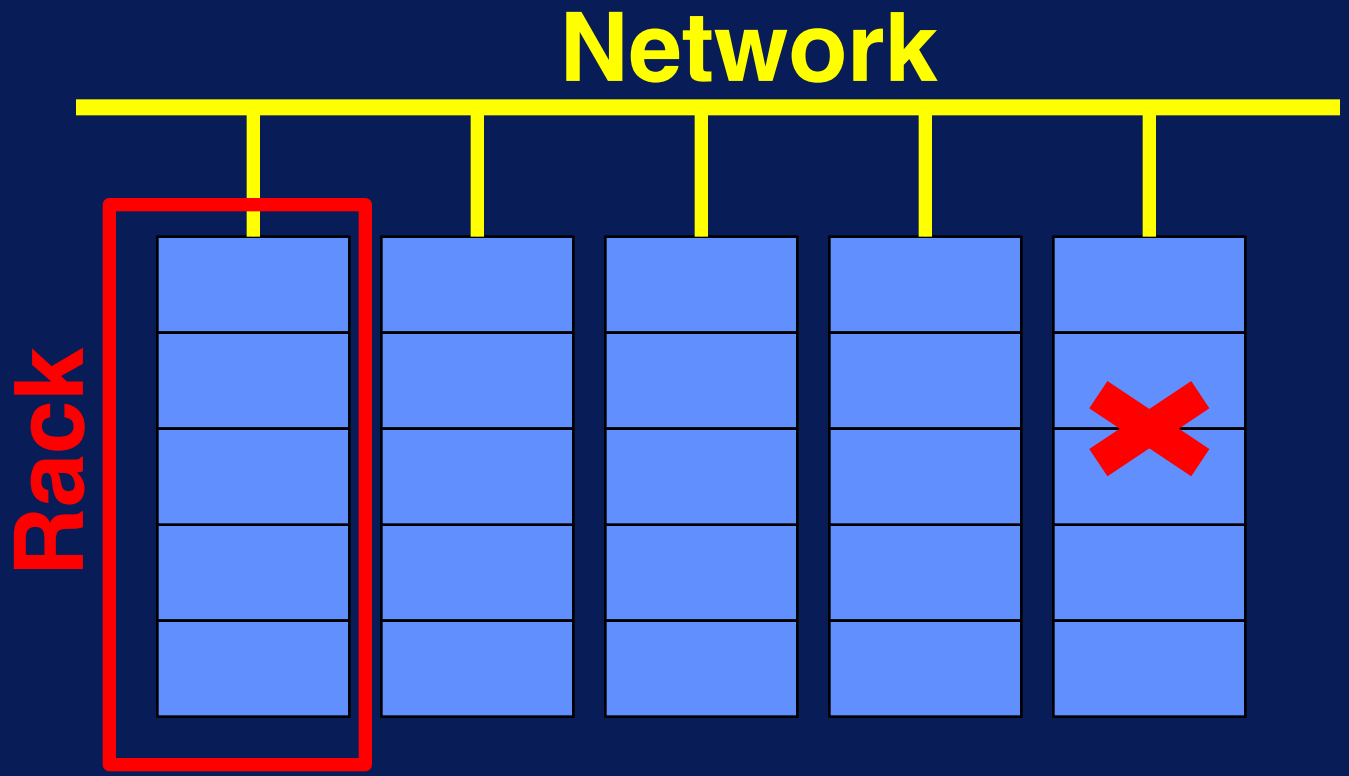
# Common failures in commodity clusters



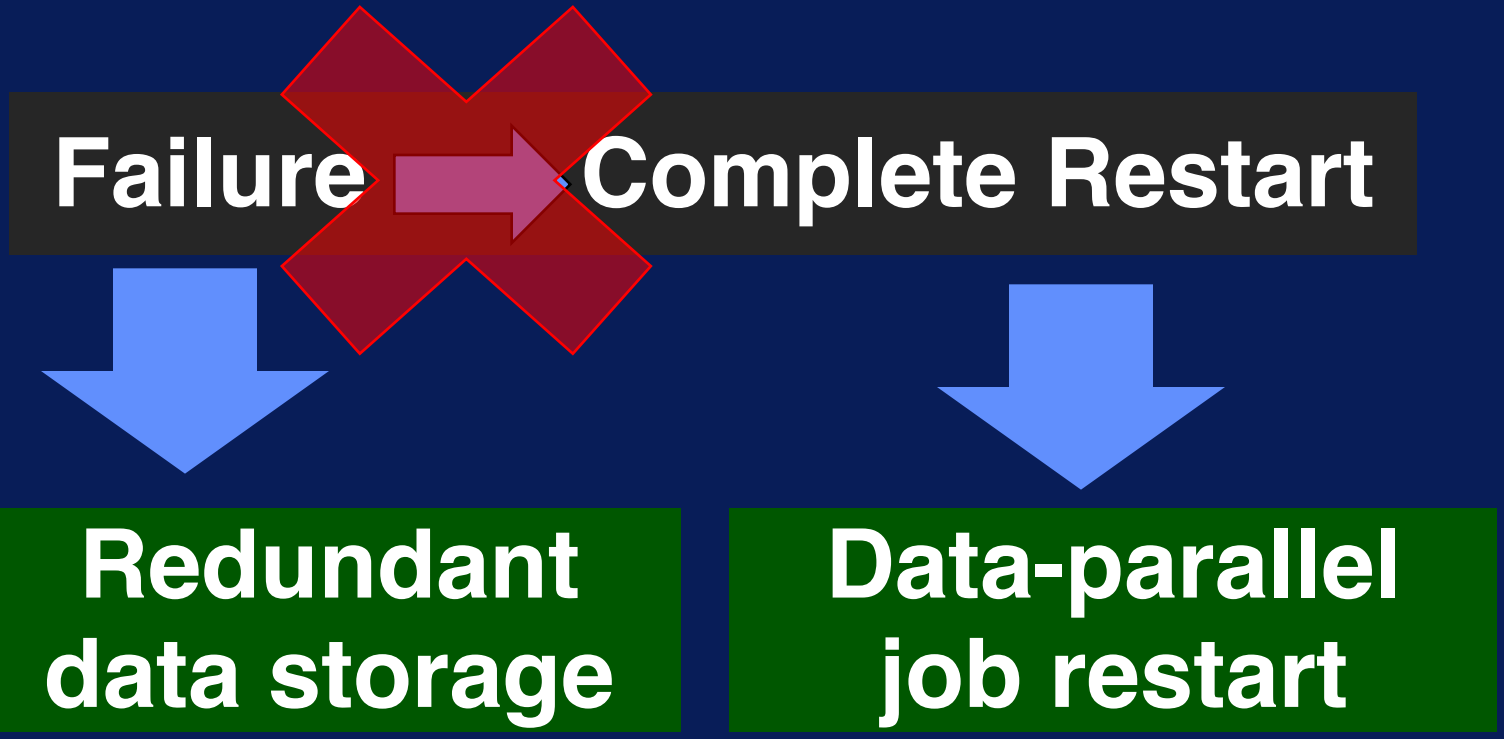
# Common failures in commodity clusters

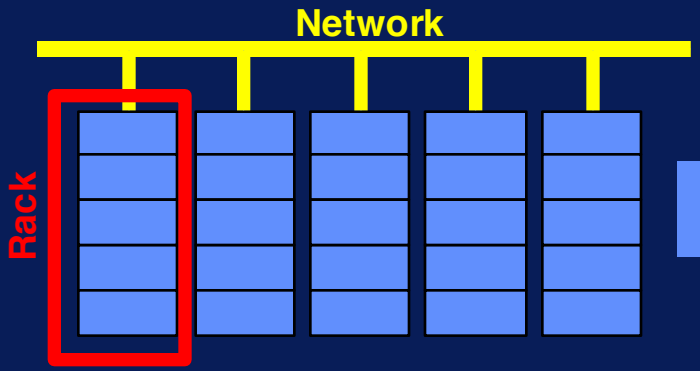


# Common failures in commodity clusters



Failure
Complete Restart


Failure
Complete Restart
Redundant data storage
Data-parallel job restart




Data-parallel scalability

Potential for node-level system failures

Redundant data storage

Data-parallel job restart