Reliability of Architectures: NODE failures
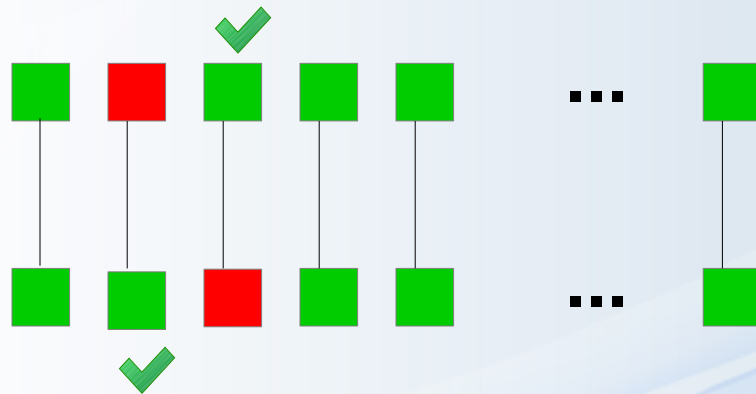
2 Node failure => ALL their disks are unreachable
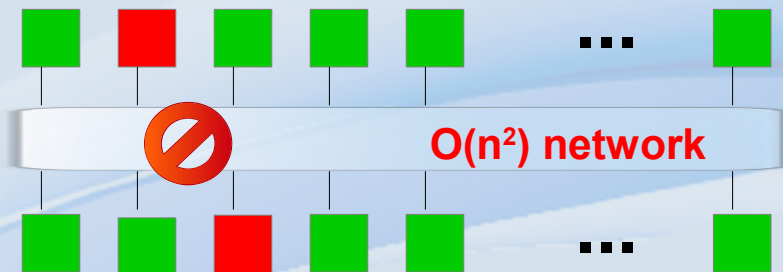
**DRBD or MARS
simple pairs**

**Big Storage Cluster
e.g. Ceph, Swift, ...**

same n

O(n²) network

=> no customer-visible incident

**k=2 replicas not enough
=> INCIDENT because objects are randomly
distributed across whole cluster**

**Low probability for hitting the *same* pair,
even then: only 1 shard affected
=> low total downtime**

**Higher probability for hitting *any* 2 nodes,
then O(n) clients affected
=> much higher total downtime**

need k >= 3 replicas here

MARS Presentation by Thomas Schöbel-Theuer