Advanced Topics in Computer Systems, CS262A
Prof. Eric Brewer

Lessons from Giant-Scale Services

Experience paper on how to build and operate very large Internet sites...

Background:

- CAP Theorem: can pick any two of Consistency, Availability, Partition-tolerance
- ... but at most two.
- Clusters pick C and A; disconnected operation and leases picks AP, locks pick CP
- Availability defined by the view at the data center -- if you can't connect that is outside the scope!

Key ideas:

- Load management
- Partitioning vs. Replication, load redirection
- Availability metrics: yield and harvest, MTTR emphasis
- Online evolution
- Graceful degradation
- the DQ principle

Basics:

symmetry

data center

backplane

Load Management

smart clients

disaster recovery

Availability Metrics: uptime, MTBF, MTTR, yield, harvest

DQ Principle

Giant Scale

Replication vs. Parititioning

- replication maintains D but not Q
- partitioning maintains Q but not D
- which is better?

Load Redirection problem: replicating the data is not enough -- must replicate the DQ access to get to the data

Graceful Degradation

- major drop in DQ
- typically try to maintain Q by reducing D significantly
- but many more sophisticated options: skip hard queries, turn on non-critical services, cache more (with stale data)

Disaster Tolerance

- loss of many replicas plus graceful degradation

Online evolution

- need a process
- staging: maintain two full versions, fast swap among them
- three ways to upgrade

Moving sites!

- possible but not easy...