



AmazingStore: Available, Low-cost Online Storage Service Using Cloudlets

Zhi Yang,
Yuanjian Xing,
Song Ding,
Feng Xiao
Yafei Dai
Peking University

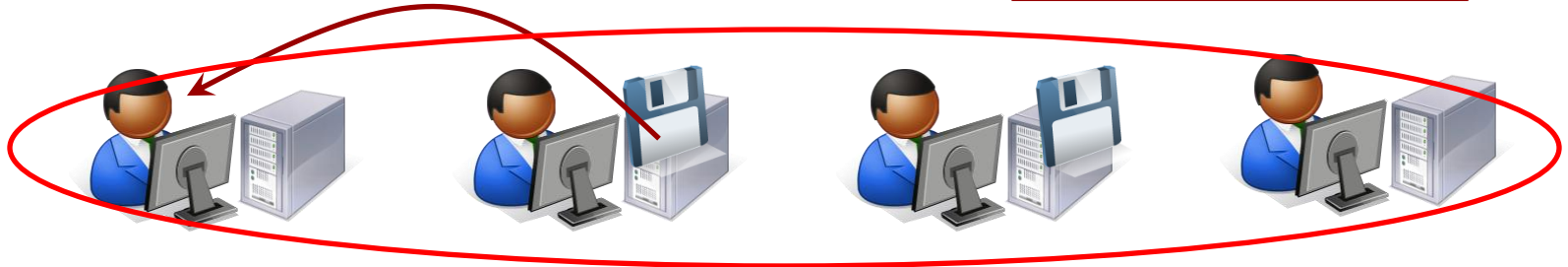
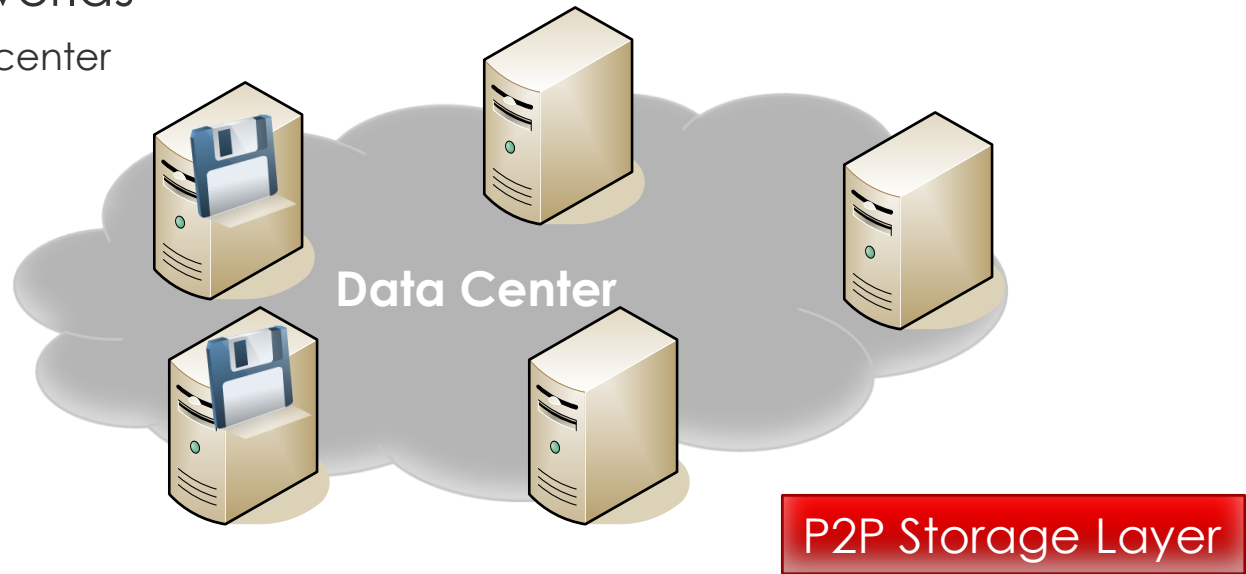
Ben Y. Zhao
U. C. Santa Barbara

Motivation

- Online storage services are getting increasingly popular
 - Amazon's S3, EMC's Mozy ...
 - Rely on data centers.
- Challenges
 - Threatened by the single point of failure.
 - Amazon suffers outages (3 times); Gmail is down (4 times) ...
 - Social networks make downtime harder to hide.
 - Incur high hardware, network and cooling costs.
- P2P storage
 - Use idle resource of users to avoid costs
 - Provide low availability because of churn.

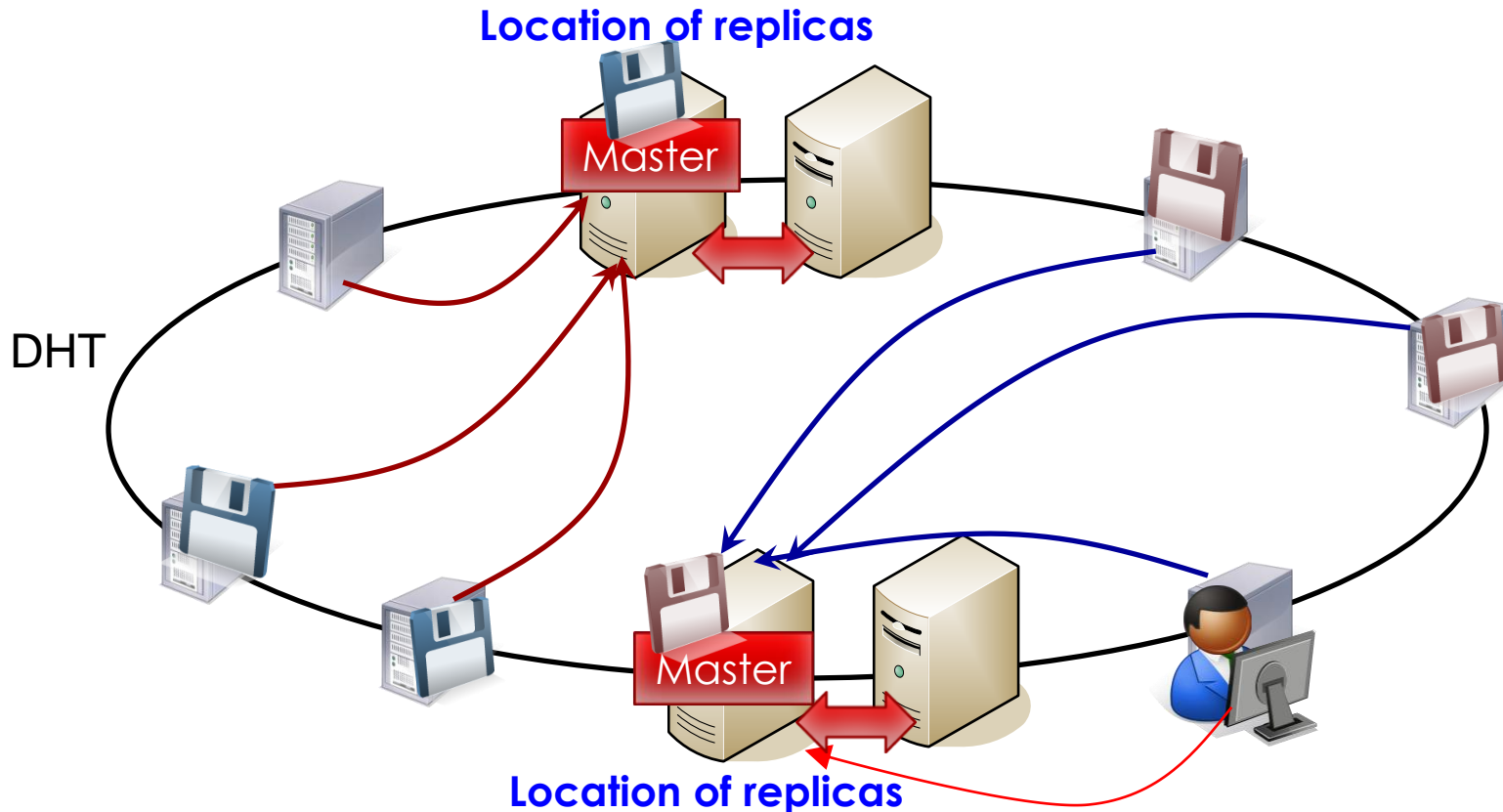
Motivation

- Best of both worlds
 - Stability of data center
 - Low cost of P2P



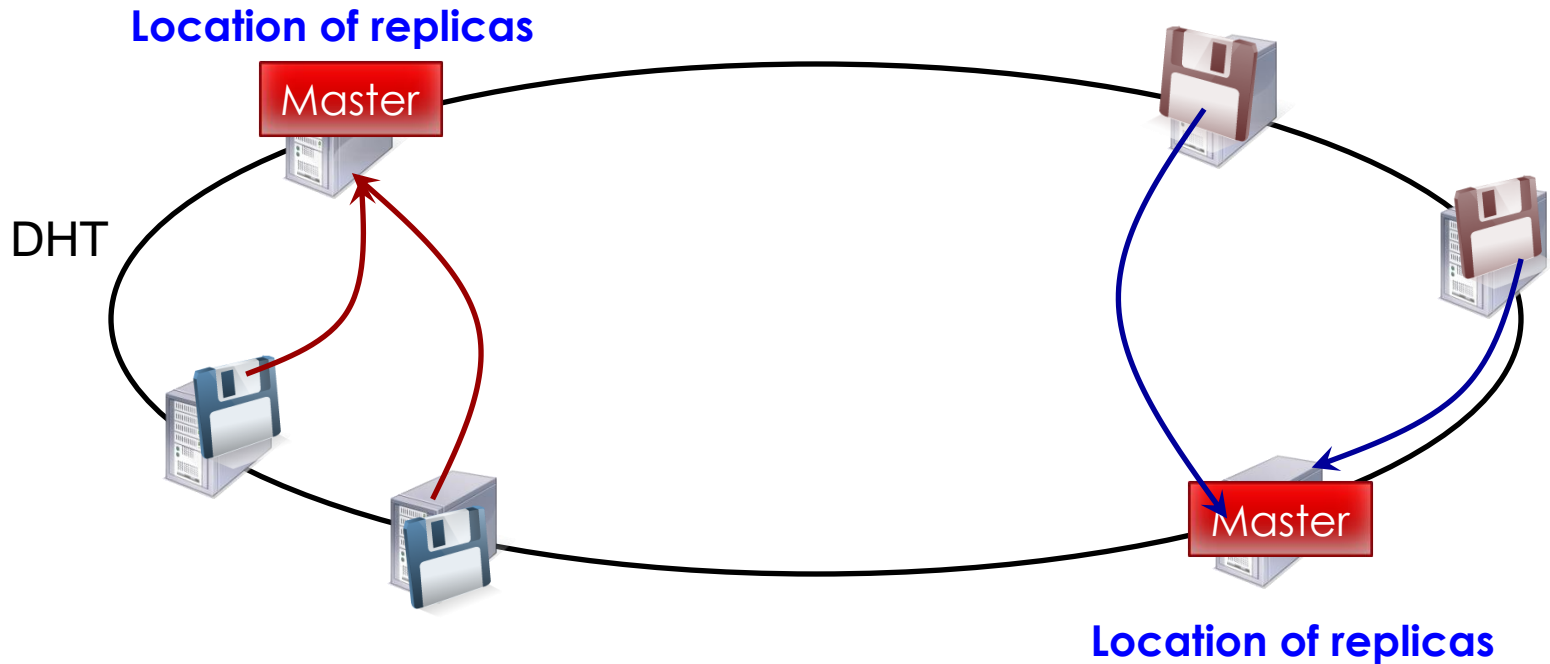
AmazingStore Design

- Combine data center and P2P storage system



AmazingStore Design

- During data center outage
 - Degrade to pure P2P storage
 - Peers closest to servers are assigned as new master nodes.



AmazingStore Design

- Determine replication degree at the peer layer
 - Threshold c derived from hybrid availability model

$$c = \frac{\log(1 - A) - \log(1 - \mu_S)}{\log(1 - \mu_P)}$$

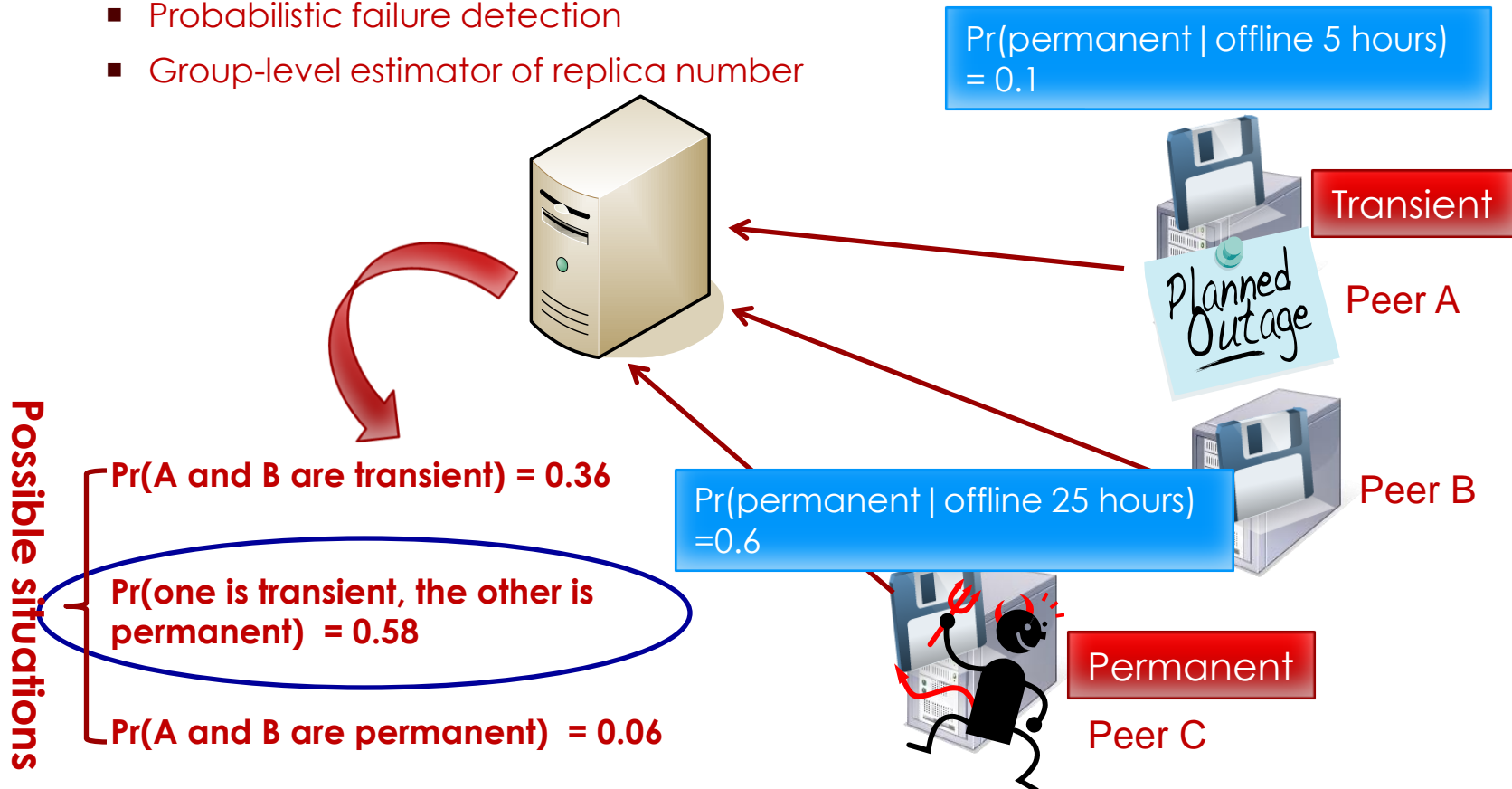
Target Availability

Data center availability

Peer availability

AmazingStore Design

- Maintain replication level at the peer side
 - Probabilistic failure detection
 - Group-level estimator of replica number



AmazingStore: Sharing & Storage

- Upload files important to you

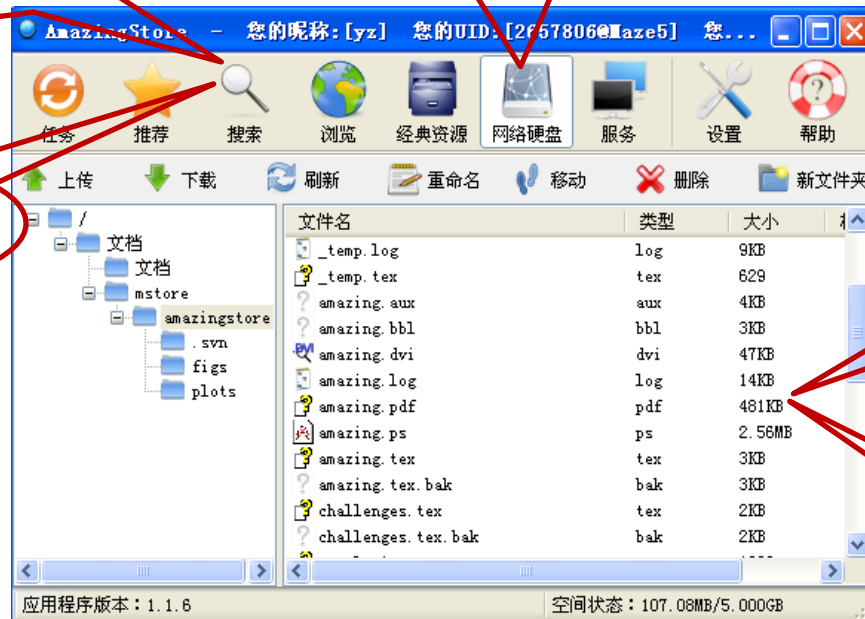
Sharing Files locally with DHT

File Storage with guarantee

No guarantee

SLA: Availability guarantee

Upload files

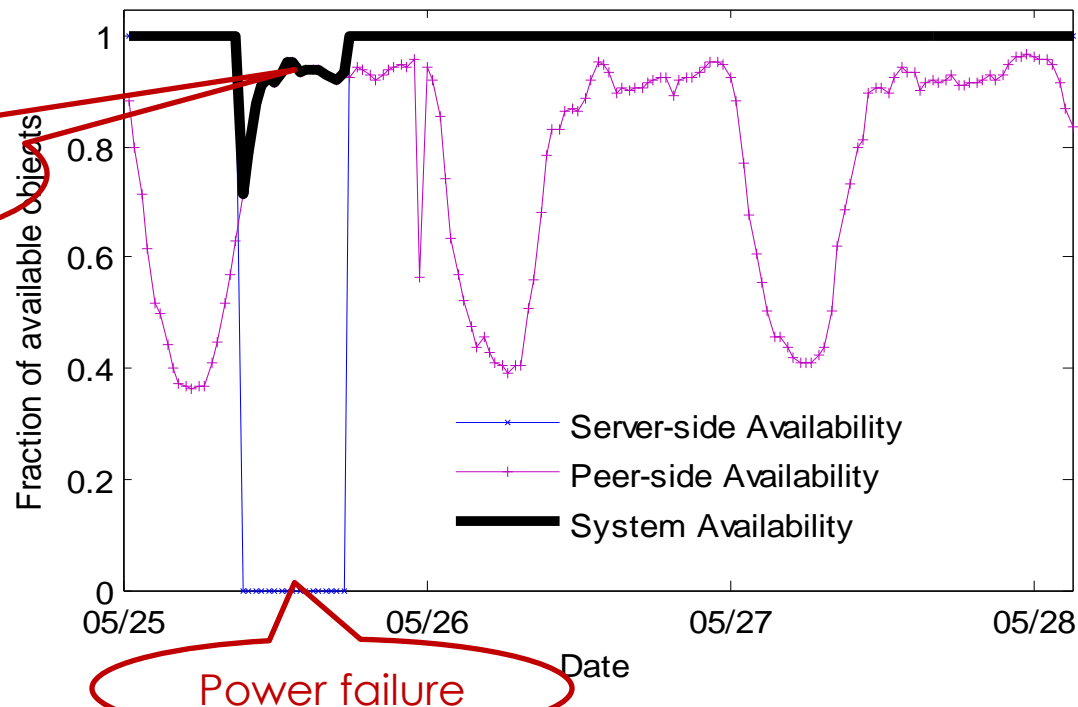


Preliminary Measurements

- composed of users and data center containing PKU servers.
- As of early April, 2010
 - Registered users > 11,820
 - Daily peak of online users > 1000
 - Data objects > 52,055.
- Provides a target of two nines availability
 - The data center availability is only 0.932
 - maintain at least 6 replicas at peer side.

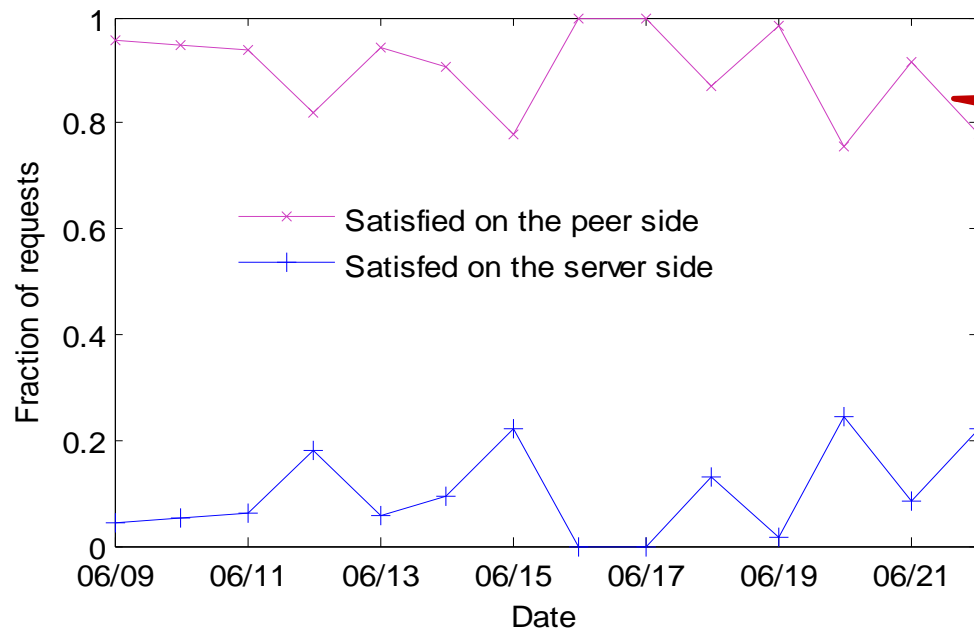
Availability Improvement

- Overall availability jumps from 93.22% to 99.13%
 - Availability gained at peer side is 83.8%



Bandwidth offloading

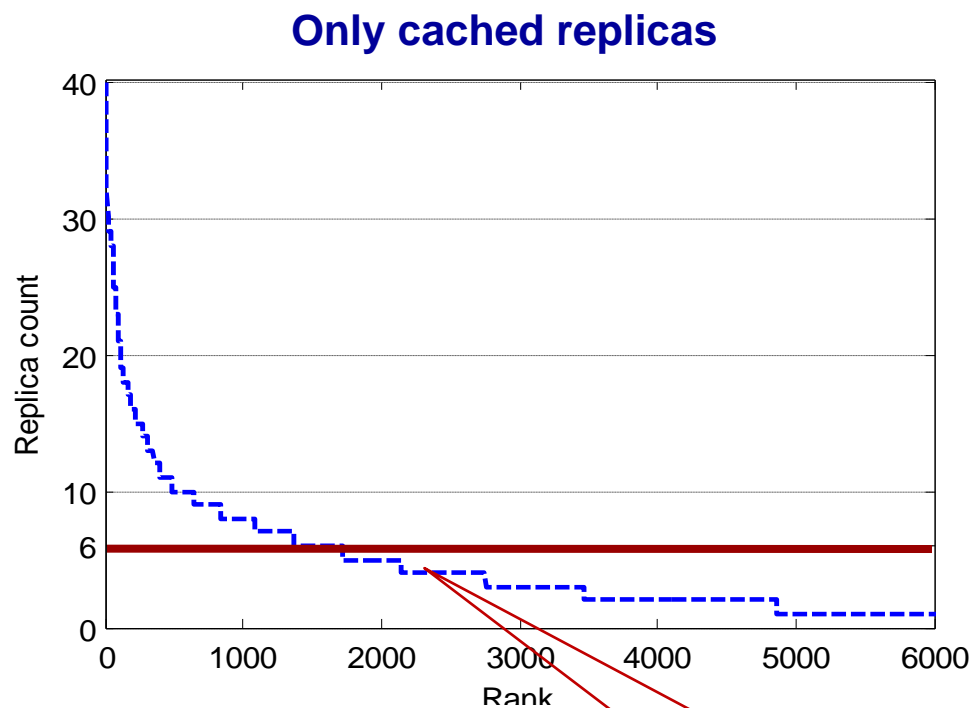
- 90.38% requests were handled by the peer layer.
- Average download bandwidth is 2.1MB/sec



Peer-side offloading

Discussion

- Edge-cached system cannot work alone.

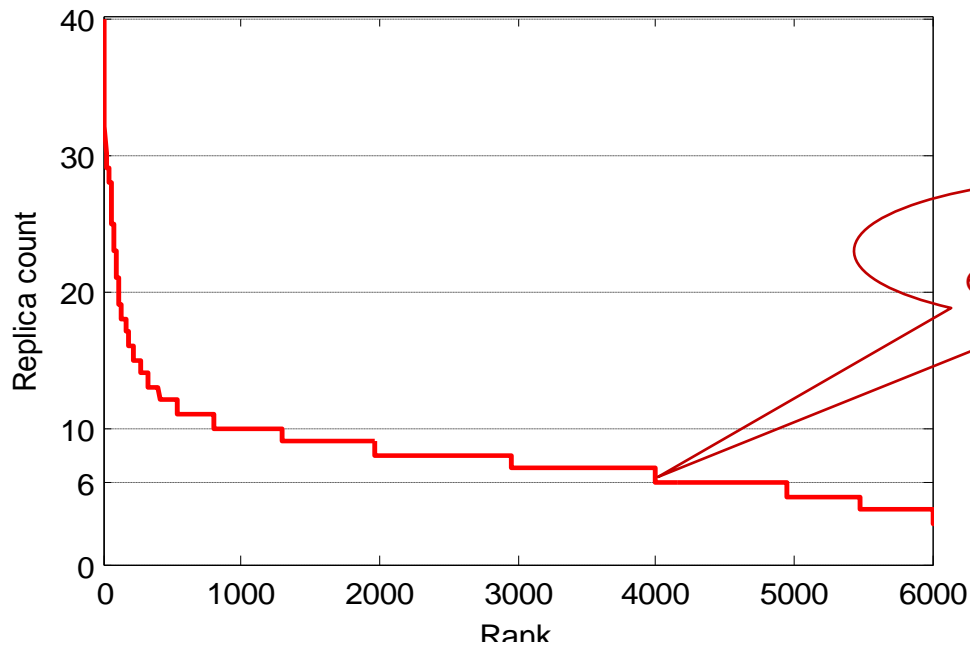


Fall below
the target

Discussion

- AmazingStore works well with data recovery

Cached replicas + Repaired replicas



Conclusion

- We advocate that data center and peers can complement well.
- We describe a deployed prototype called AmazingStore.



Questions?

Thank you!