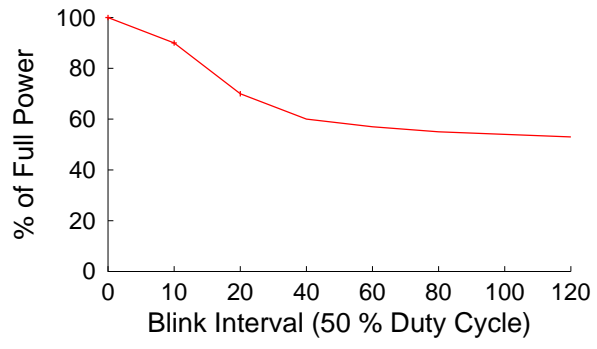


Figure 4: The popularity rank, by number of fans, for all 20 million public group pages on Facebook follows a Zipf-like distribution with $\alpha = 0.6$.

a corresponding update in persistent storage. Each memcached server uses the Least Recently Used (LRU) replacement policy to evict objects. One common example of a cached object is an HTML fragment generated from the results of multiple queries

Figure 5: To explicitly control the mapping of keys to servers, we interpose always-active request proxies between memcached clients and servers. The proxies are able to monitor per-key hit rates and migrate similarly popular objects to the same nodes.

server. One way to correct the problem is to dynamically change the built-in client mapping function to only reflect the current set



120 -

90 -

60 -

30 -

0 -

1

10

250 -

200 -

150 -

100 -

50 -

0 -

storage research, such as migrating popular objects to more active nodes [28, 41]. Additionally, power-aware caching algorithms focus on maximizing the idle time between disk accesses to reduce disk power consumption, while our work focus on controlling the power consumption of the cache itself [42].

Blinking introduces regulated churn into data center applications as nodes switch from the active to inactive state. Churn has been well-studied in decentralized, self-organizing distributed hash tables [34]. However, the type of churn experienced by DHTs is different than the churn caused by blinking, which motivates our different approach to the problem. In the former case, nodes arrive and depart unexpectedly based on autonomous user behavior and

