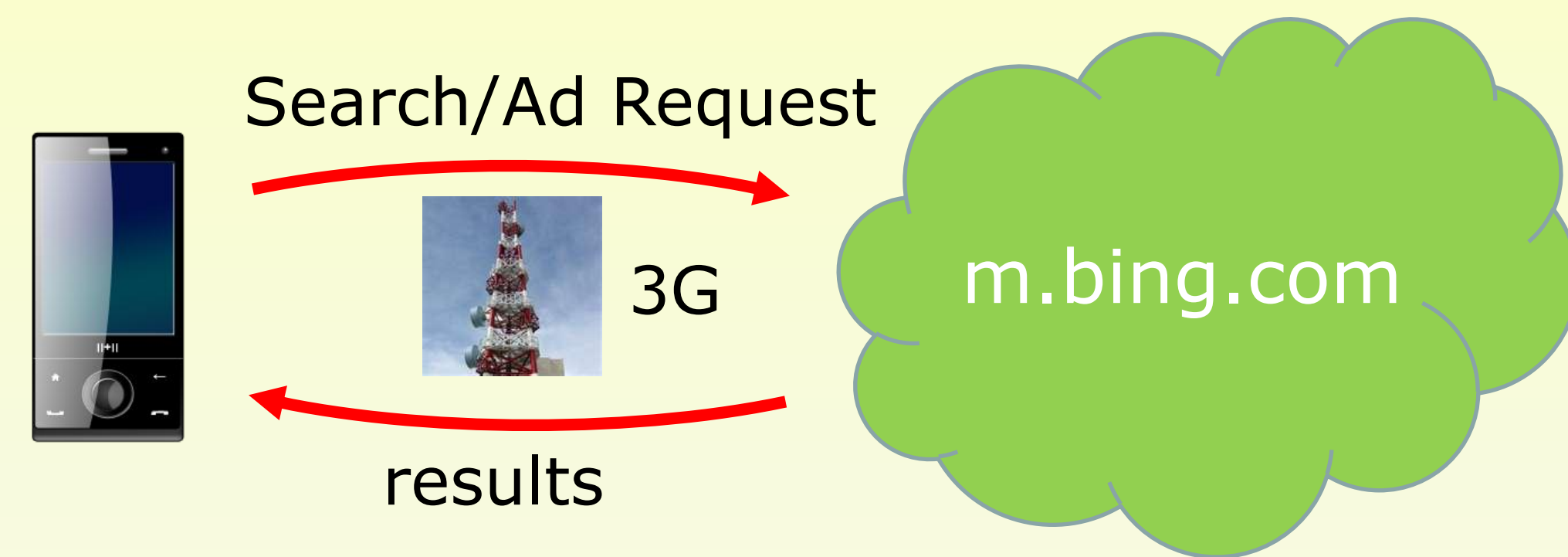


Mobile Search and Advertisement Cache Architecture

Dimitrios Lymberopoulos, Emmanouil Koukoumidis, Jie Liu, Doug Burger (MSR Redmond)
Varun Kansal, Chen Xia, Kenny Chien, Bin Wu, Fang Wang, Melissa Dunn (MAXPLAT)

Mobile Search Experience

Performance Bottleneck

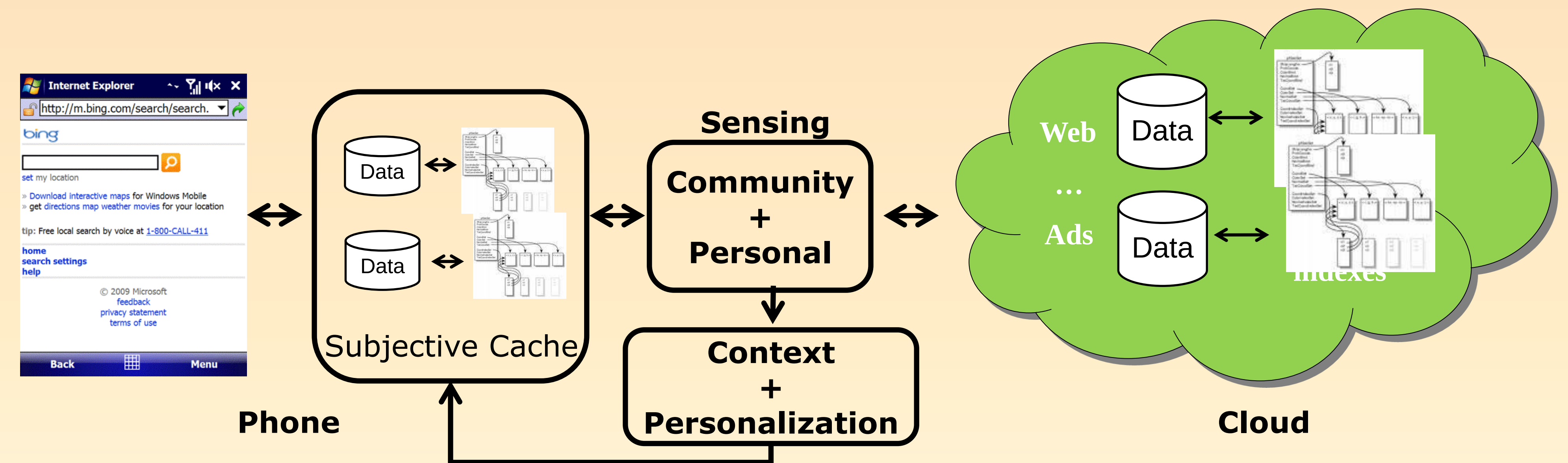


3G link is slow

- Average end-to-end delay: 4-8 seconds
- 3G connection time: 10 - 30 seconds

Goal: Faster Mobile Search Experience

Push Search/Ads to the Phone



Leverage mobile trends

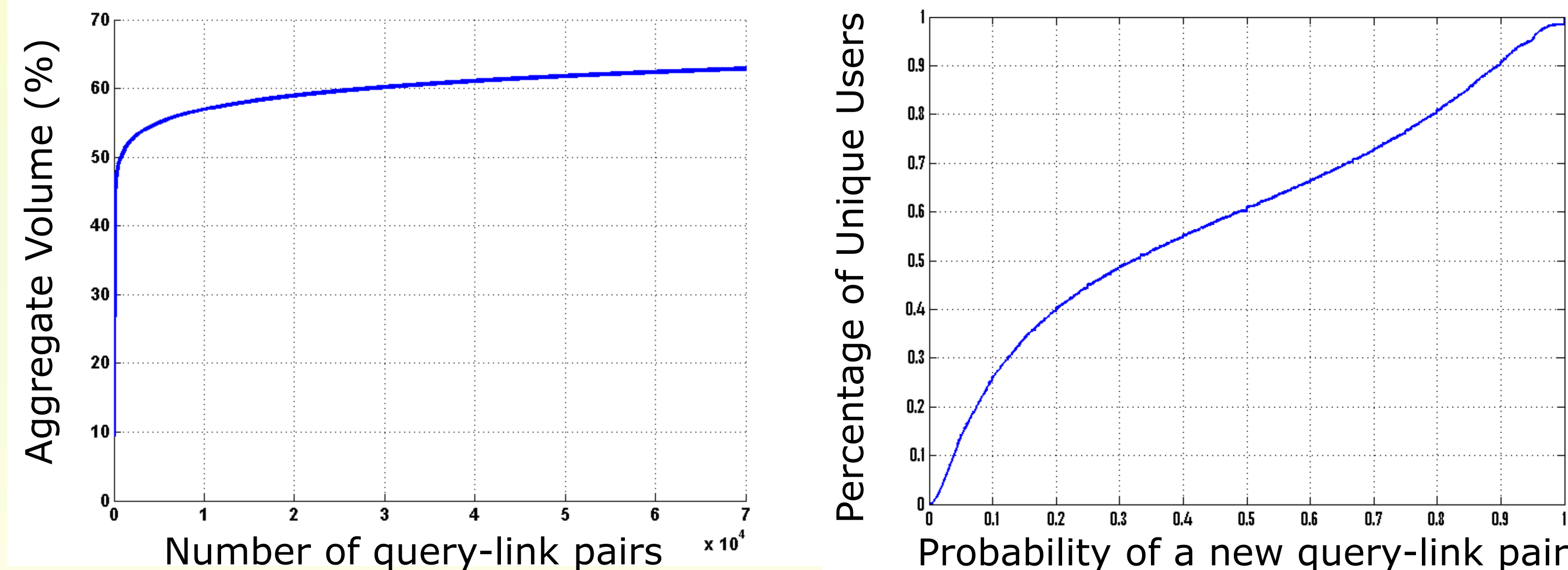
- Increasing flash density at lower cost
- Increasing processing power

Challenges

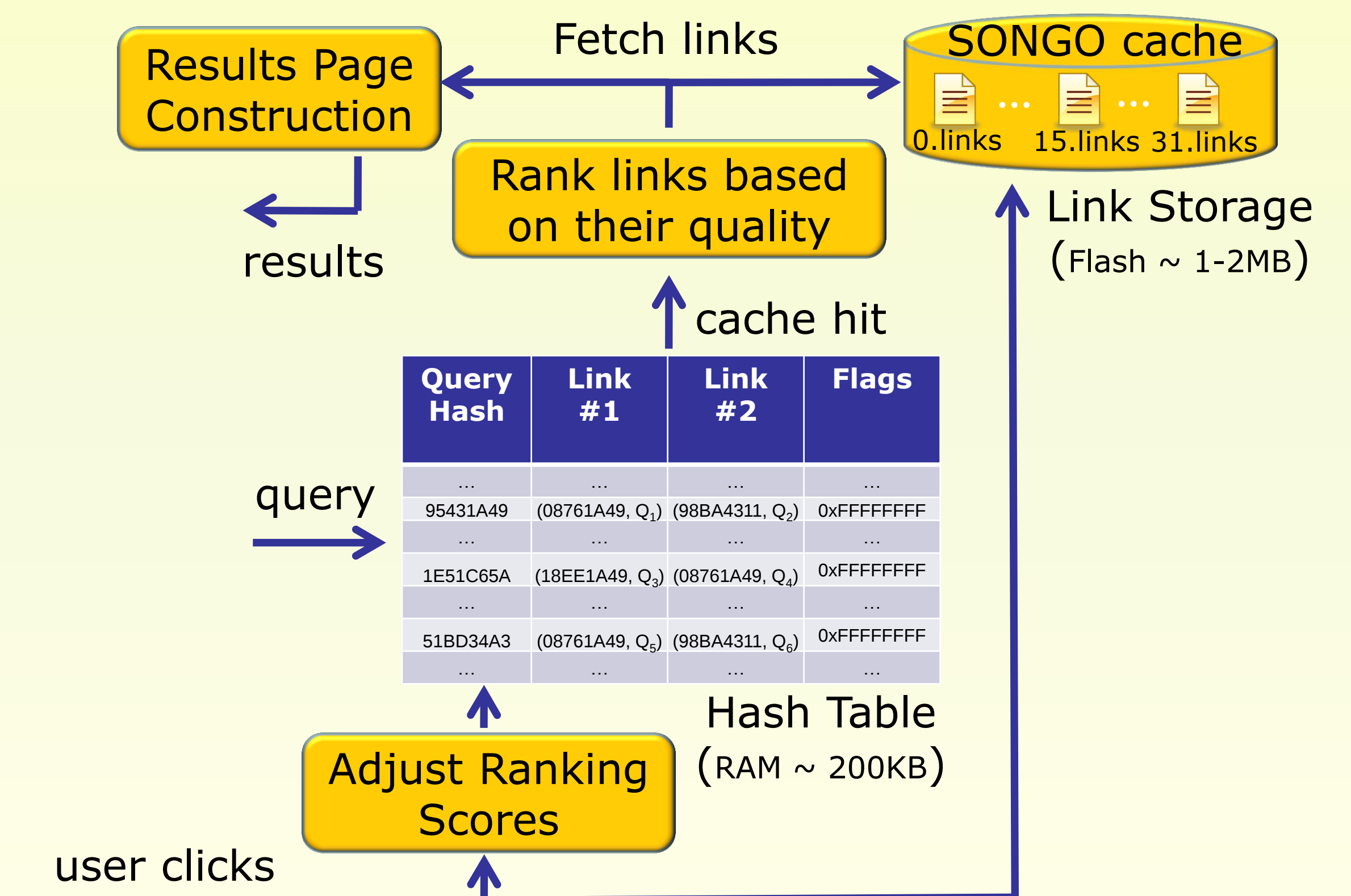
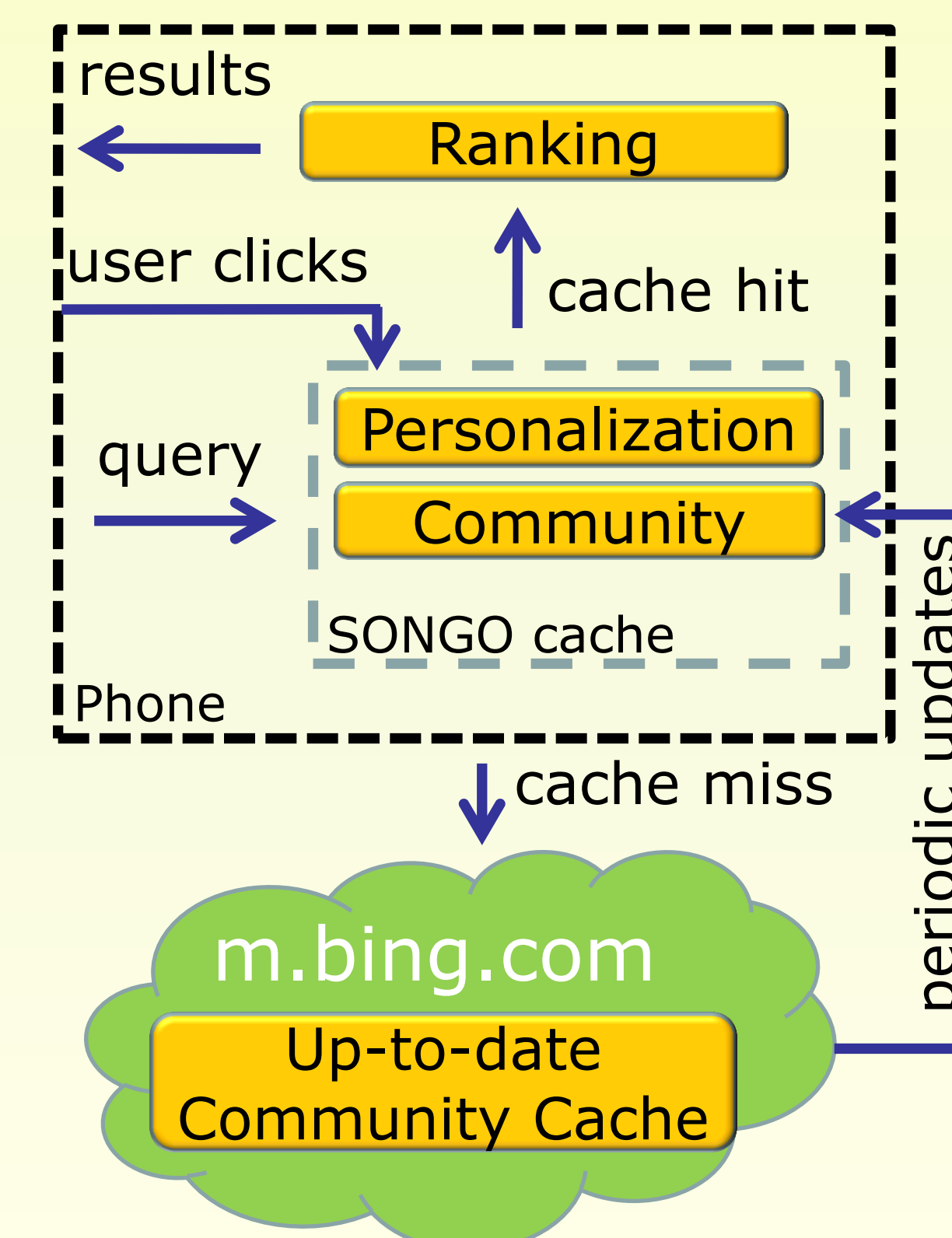
- What information do we cache?
- How do we cache it?
- How do we manage it over time?

SONGO (Search ON the GO): A Data-driven Architecture

Mobile Search Log Analysis - 100M Queries

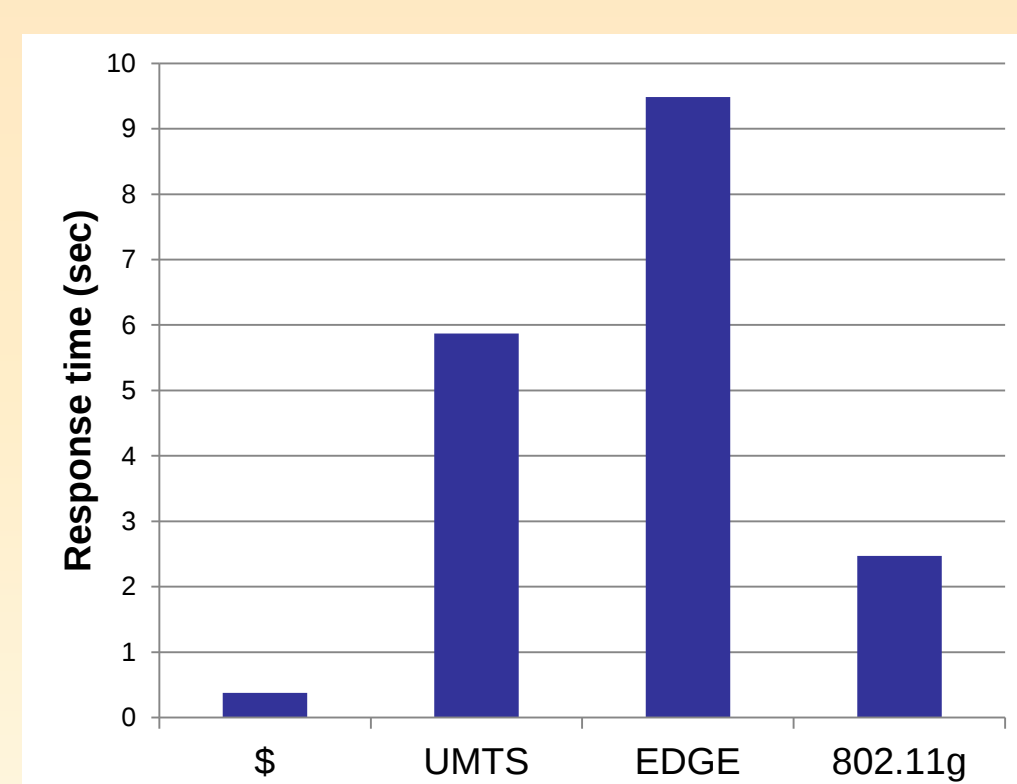


60% of mobile query volume hits 6K queries and 4K links!
50% of the users repeat a query at least 70% of the time!

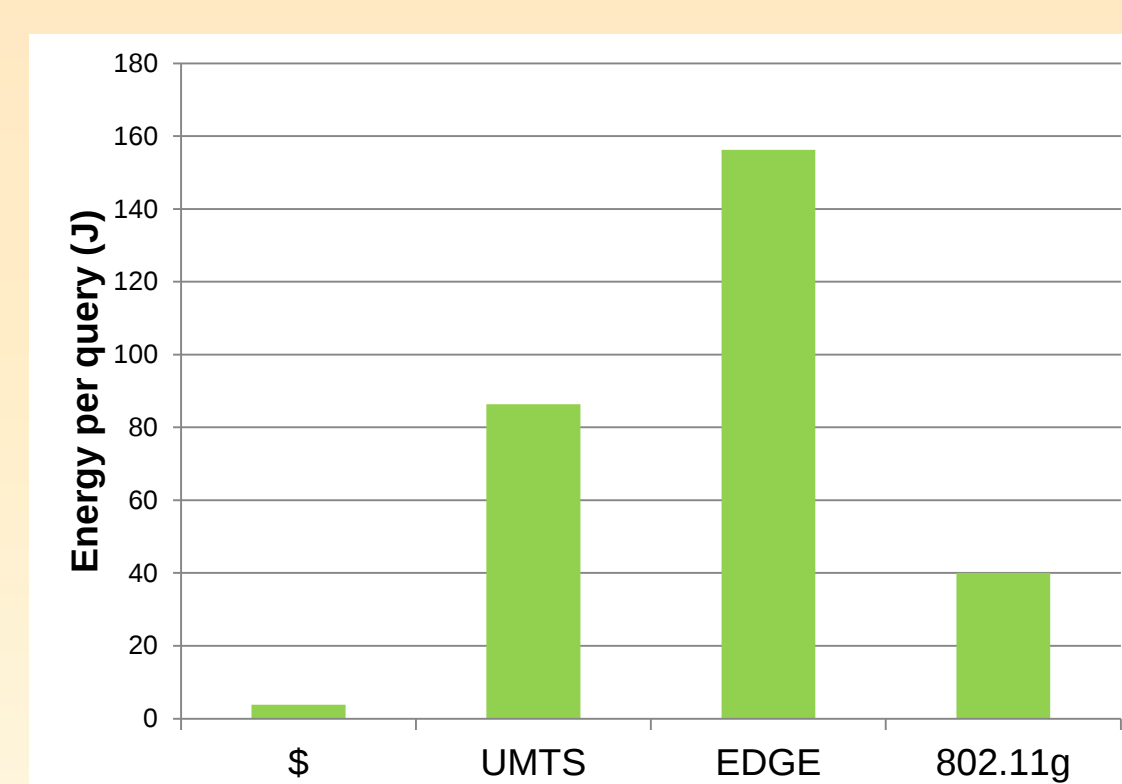


Experimental Prototypes

On average 66% of the queries a user submits hit the cache

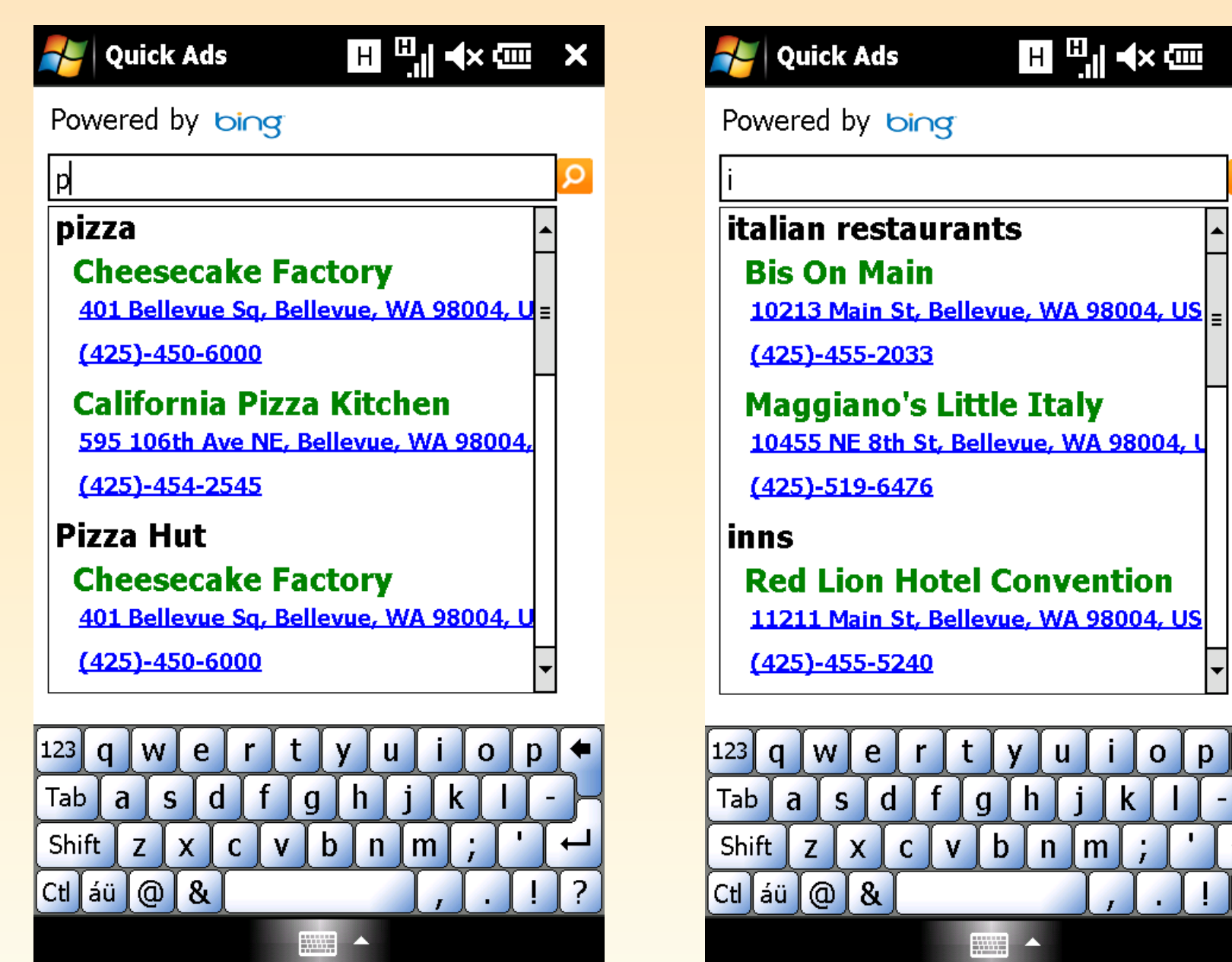


16x faster!



23x more energy efficient!

Beyond Web Search



Real-time:

- search results
- business lookups
- ad delivery

Opportunities:

- Faster user experience
- Monetization of autosuggest
- Fastest mobile ad delivery engine
- Personalized ranking for search/ads
- Privacy: profile on the phone