

An Evaluation of Web Acceleration Techniques for Developing Countries

Amal Fahad Zhuan Chen Kai Shen Jeffrey Bigham

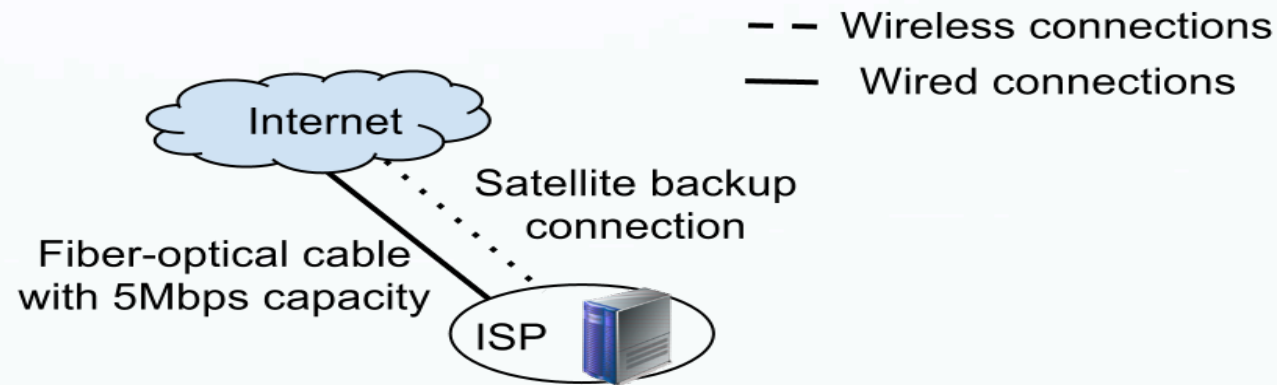
University of Rochester, U.S.A.

Assmaa Fahad

Baghdad University, Iraq

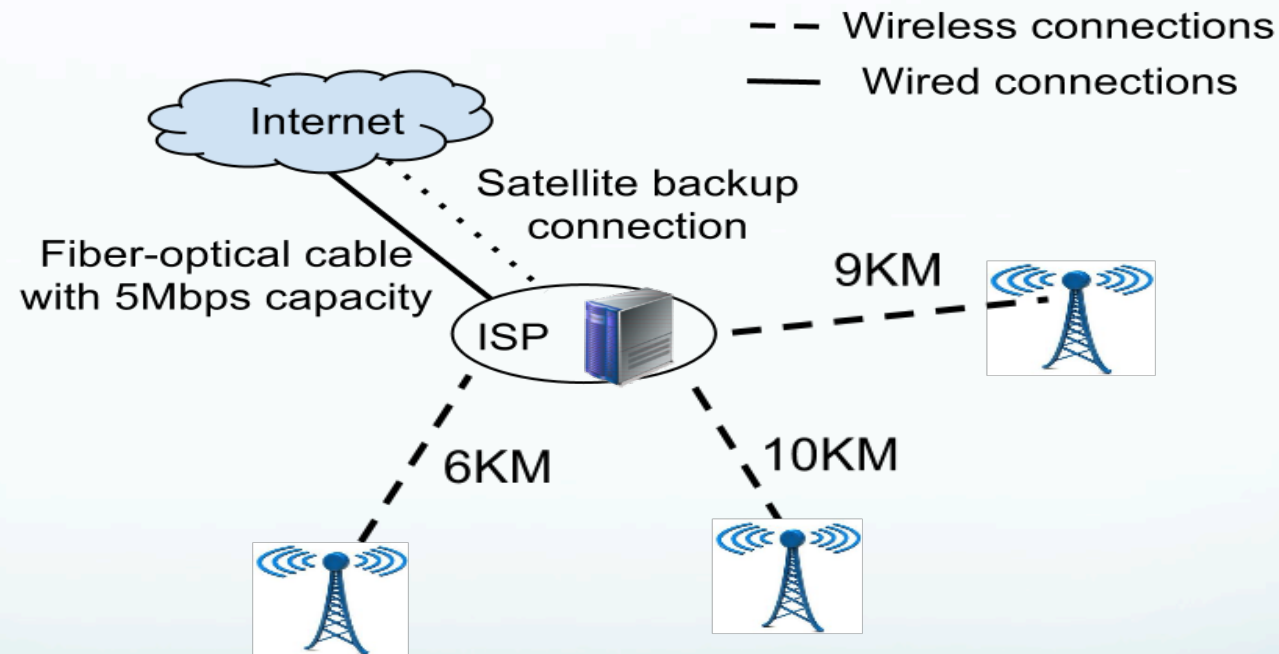
Motivation

Poor network infrastructure is the common factor across developing world.



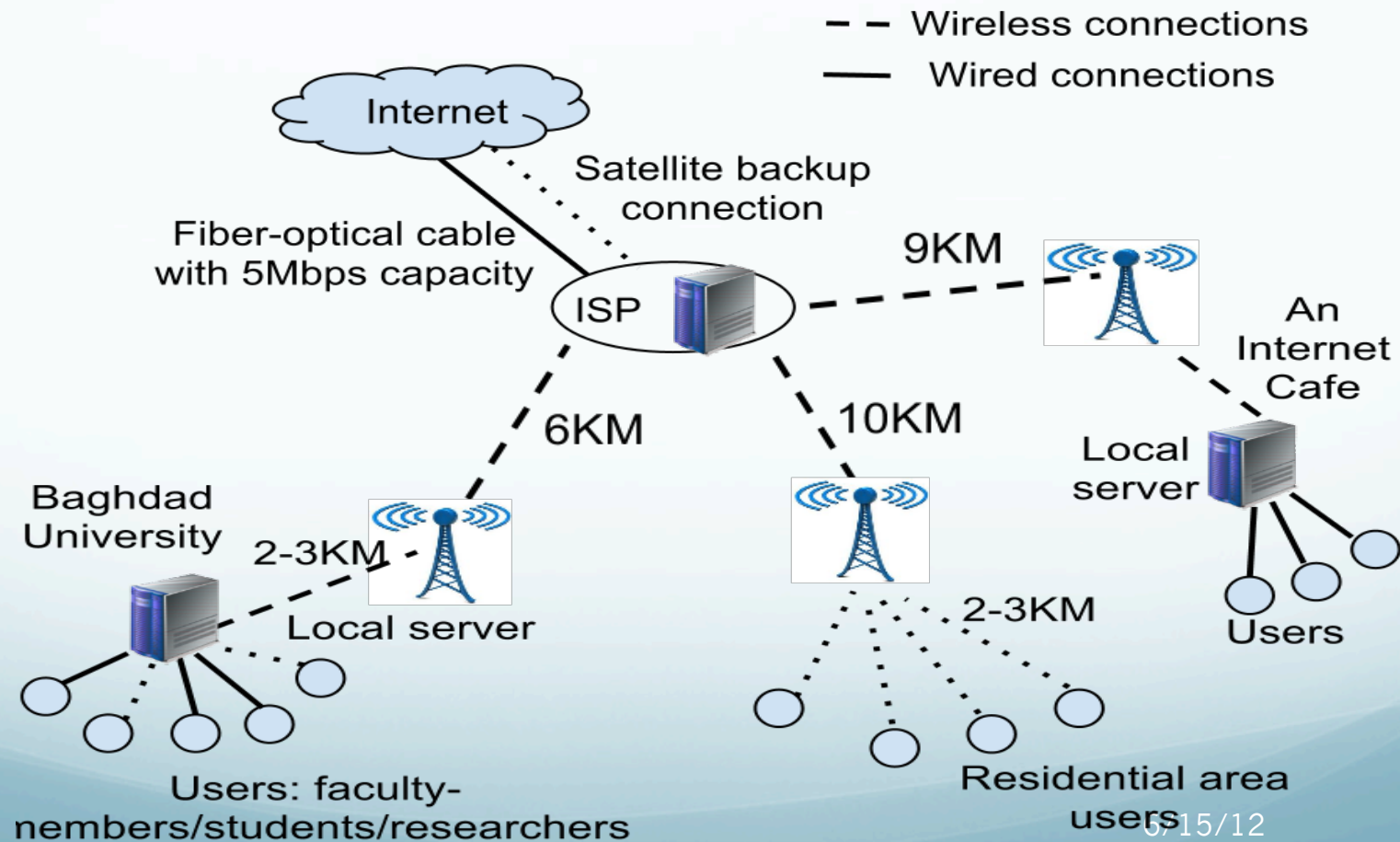
Motivation

Poor network infrastructure is the common factor across developing world.



Motivation

Poor network infrastructure is the common factor across developing world.



Motivation.. In Pictures



6/15/12

Motivation.. In Pictures



Motivation.. In Pictures





Motivation.. In Pictures



Motivation.. In Pictures



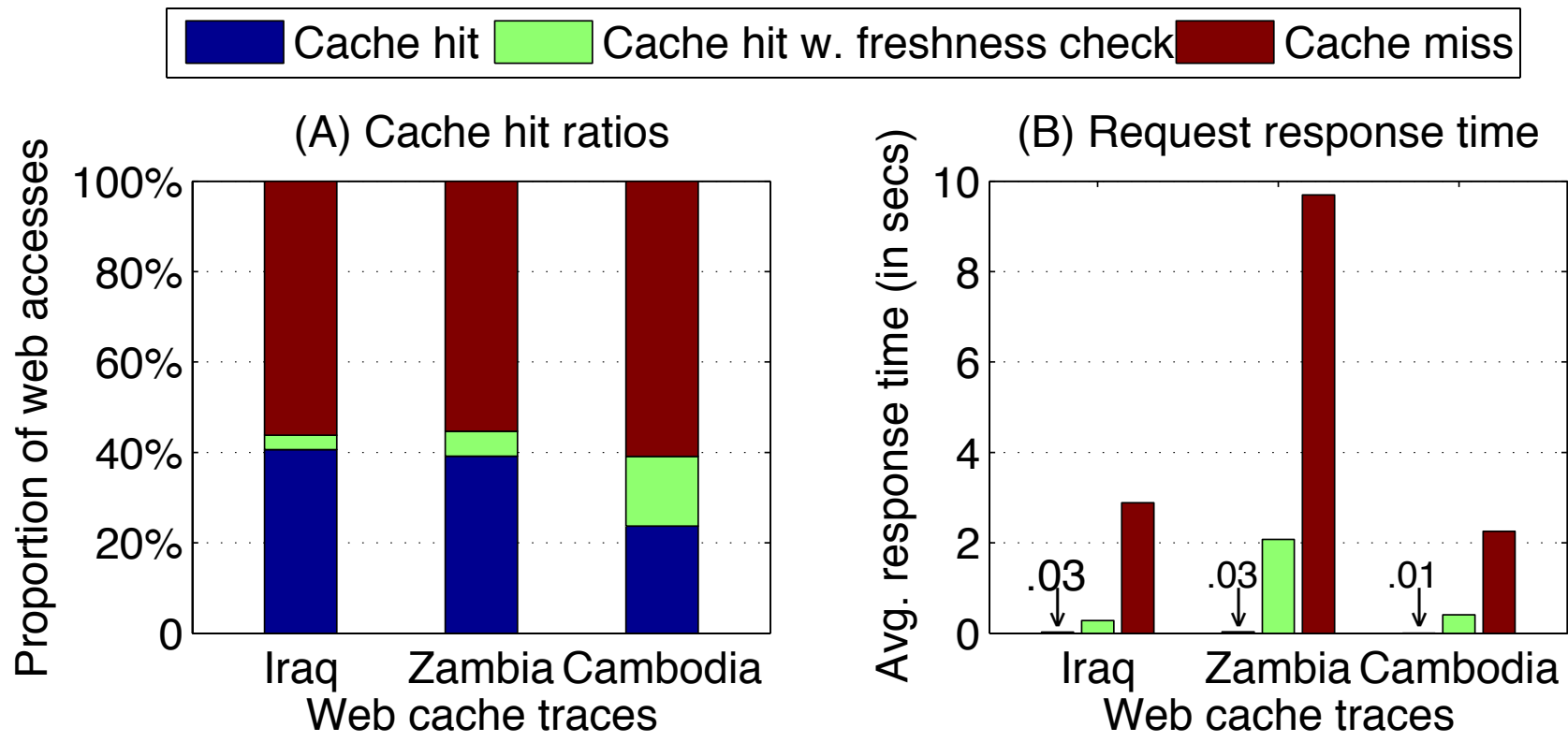
Network Acceleration Techniques

- Different acceleration techniques have been suggested to improve the Internet service for clients in developing regions, such as:
 - Web proxy caching. 
 - Web Prefetching. 
 - Offline browsing.
 - Data de-duplication.
- In this paper:
 - We evaluate the performance of some of these techniques using real web traces.
 - We also suggest a new technique based on our observations from this evaluation.

The Used Web Traces

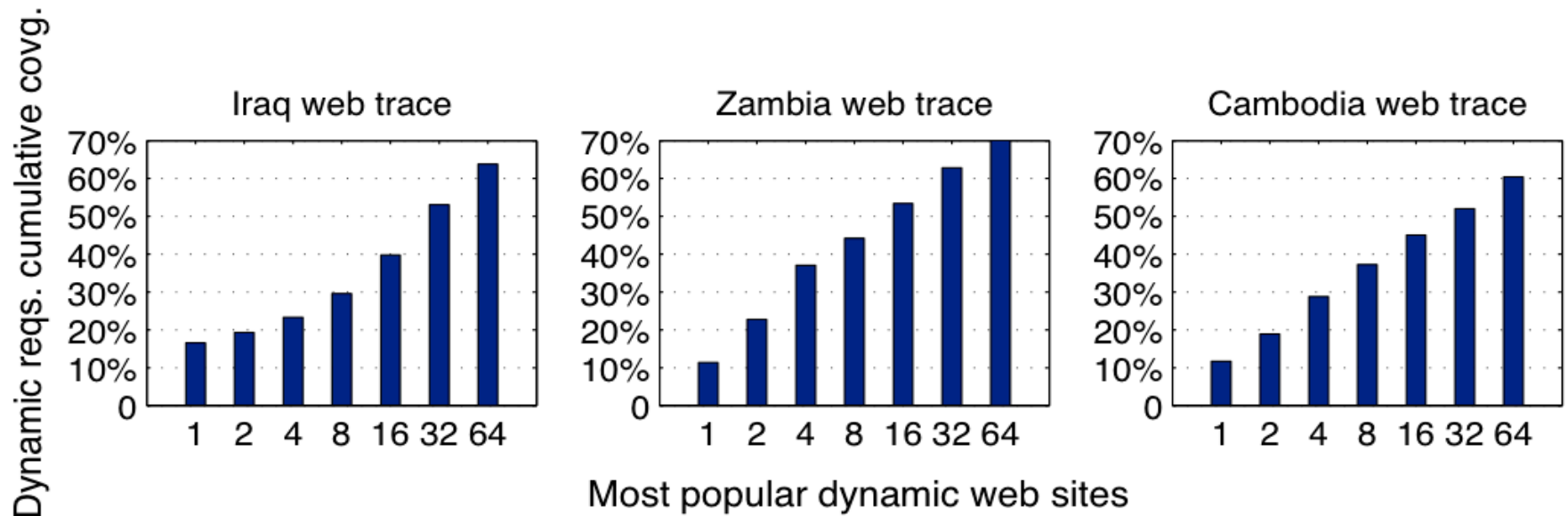
Location	Internet connection	Time period	# of web accesses	Collection site
Iraq	Cable (W/ satellite backup)	Aug.23- Aug.25 2011	2 Million	At the ISP that serves Baghdad University and the nearby residential area.
Zambia [Johnson et al. 2010]	Satellite	Jan.24- Feb.14 2010	4 Million	At a gateway that connects the satellite to the wireless network covers Macha village.
Cambodia [Du et al. 2005]	Satellite	June 1- Sep. 28 2005	11 Million	At an Internet Service Center.

Web Proxy Caching Performance



Dynamic Web Applications

- Dynamic web accesses ratio is 59%, 62%, and 41% among the cache misses respectively.



The top rank application in each of the three traces were Facebook, Gmail, and Yahoo! search engine respectively.

Prefetching

- Predicting future web accesses and prefetching them ahead of user requests.
- Prefetching techniques can be classified into:
 - Non-history based.
 - History based.
- Early work showed:
 - 2% hit rate for a non-history based technique with 22% bandwidth overhead[Chen et al. 2011].
 - 10% hit rate for a history based technique with 23% bandwidth overhead [Fan et al.1999].

Prefetching Evaluation

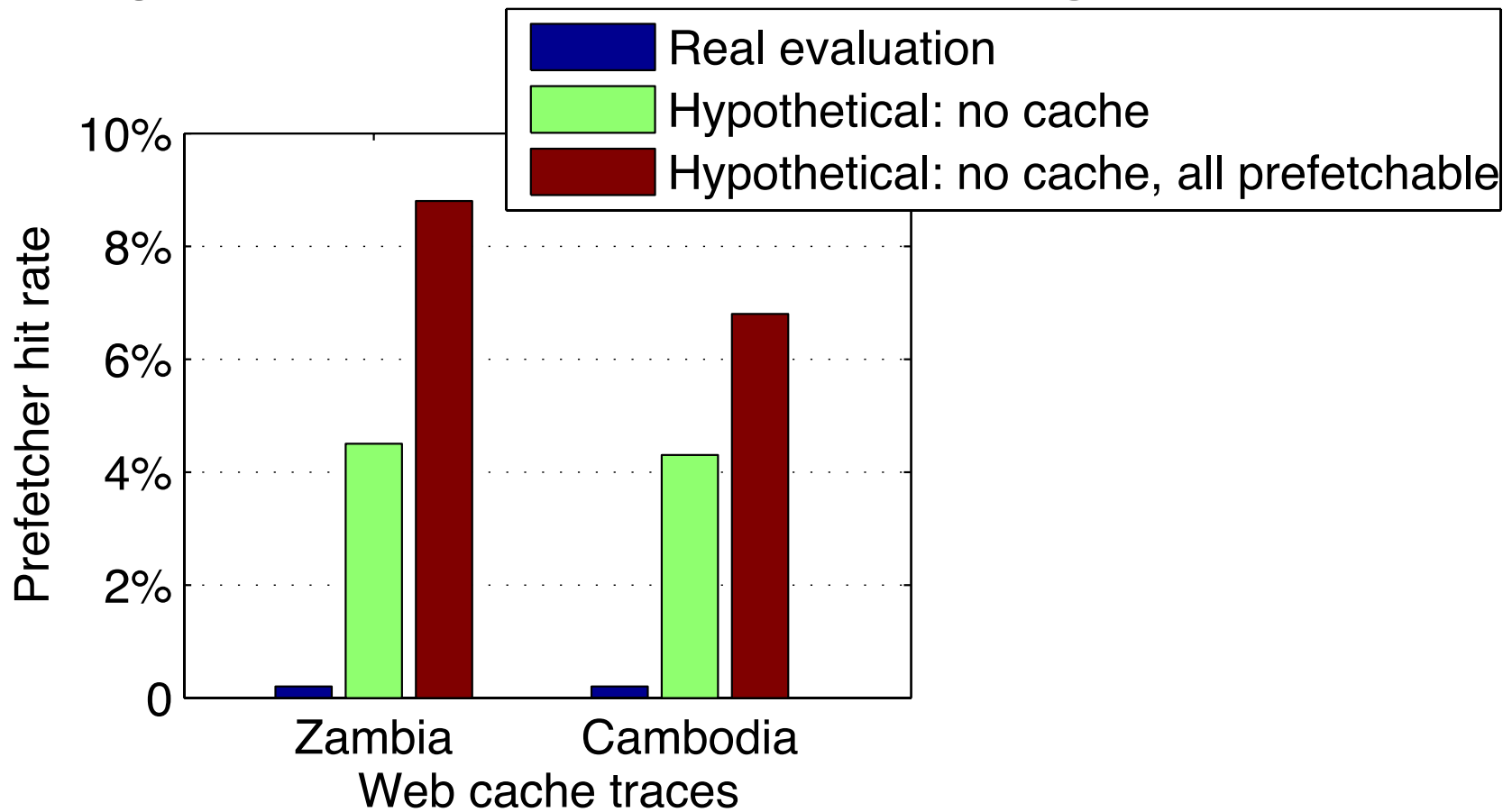
Using the Prediction by Partial Matching (PPM) [Fan et al. 1999] history based prefetching:

Traffic location	Hit rate	Overhead rate	Use rate
Zambia	0.2%	2.5%	1.7%
Cambodia	0.2%	8.7%	1.1%

Compared with 10% hit rate reported earlier, our evaluation gave poor results... why?

Results Analysis

- Web proxy caching effectiveness.
- Dynamic requests ratio is increasing.



Network Acceleration Performance

- So far we have seen that:
 - Web proxy caching is an effective solution to speed up static pages accesses.
 - The ratio of dynamic requests is increasing in today's network traffic.
 - Requests from small number of applications dominate most of the network traffic (a set of 8 applications is responsible of ~40% of the dynamic requests).

Local Application Mirroring

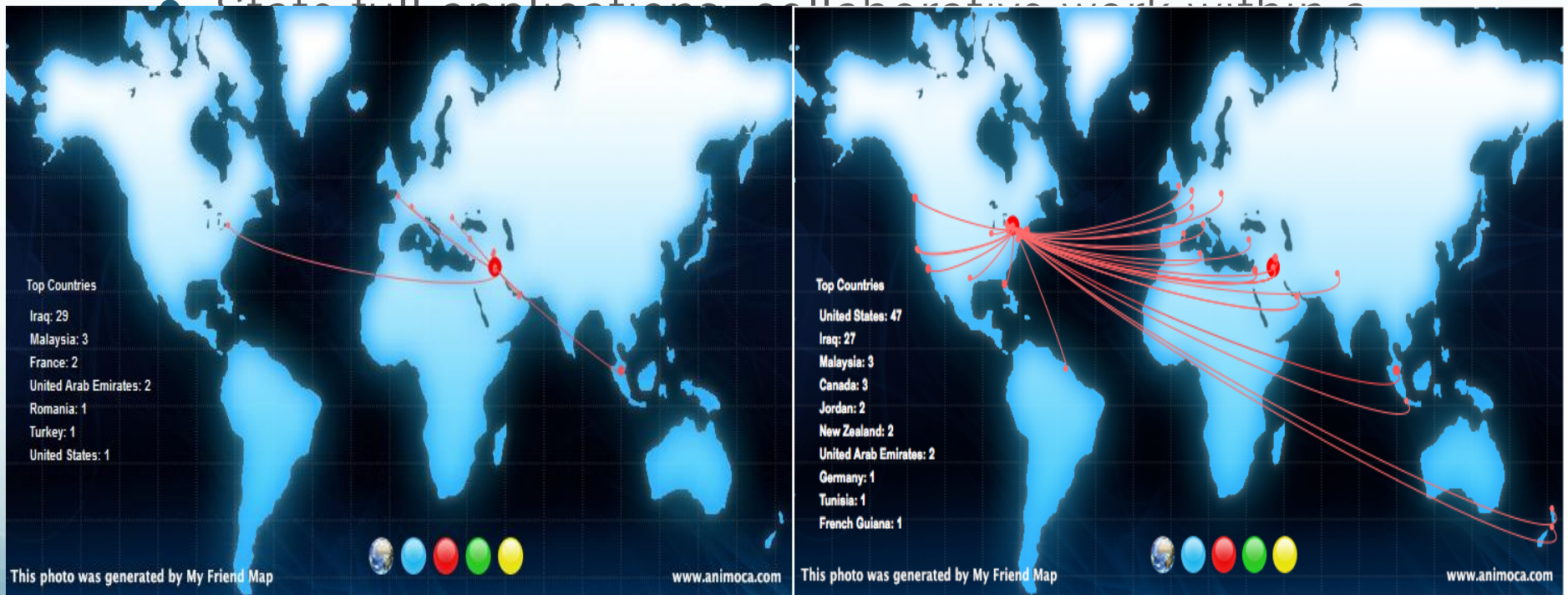
- Local Application Mirroring: Caching dynamic web applications locally in order to reduce the latency at the client's side as well as the network traffic.
- Similar in concept to Akamai and Amazon CloudFront but closer to their users than both.
- Challenges.....?

Data Management

- Web based applications can be classified into:
 - Read-only state applications, content transformation:
 - Perfect for application mirroring.
 - State-full applications, collaborative work within a group.
 - Locality.

Data Management

- Web based applications can be classified into:
 - Read-only state applications, content transformation:
 - Perfect for application mirroring.



Data Management

- Web based applications can be classified into:
 - Read-only state applications, content transformation:
 - Perfect for application mirroring.
 - State-full applications, collaborative work within a group.
 - Locality.
- Hierarchical updating: local hosts report their update to the higher level nodes in the hierarchy.

Deployment and Security

- Applications have different requirements:
 - Runtime environment.
 - Required libraries.
 - Configuration.
- Diversity between local and global platforms.
- Security requirements:
 - Among locally hosted applications.
 - Between an application and its hosting environment.

Virtual Machine

- Native machine-level virtualization, which is used in cloud computing, can be leveraged to implement our technique.
- Migration over unreliable network, unlike the VM migration process in the cloud.
- Limited bandwidth:
 - Transferring only the difference between the application's VM image and a local stock image. Using rsync, this difference can be as small as 43MB for a VM image that is 725MB in size.

Evaluation of a Collaborative Application Performance

- We used Feng office, an open source web-based application for collaborative software in our measurement.



The screenshot displays the Feng Office web interface. On the left, a sidebar contains a 'Workspaces' section with a search filter and a list of workspaces including 'All', 'zchen Personal', 'Archived Objects', and 'Trash'. Below this is a 'Tags' section with a search filter and a list of tags including 'All'. The main content area features a top navigation bar with tabs for 'Overview', 'Notes', 'Email', 'Contacts', 'Calendar', 'Documents' (selected), 'Tasks', 'Web Links', 'Time', and 'Reporting'. Below the navigation bar is a toolbar with buttons for 'Back', 'Save', and 'Rename', followed by a rich text editor toolbar with various formatting options. The main content area displays the 'feng OFFICE' logo. Below the logo, a paragraph of text describes the application: 'Feng Office Community Edition is an open source collaboration platform developed and supported by Feng Office and the OpenGoo community. It is a fully featured online office suite with a similar set of features as other online office suites, like Google Apps, Microsoft Office Live, Zimbra, and Zoho. The application can be downloaded and installed on a server. Feng Office could also be categorized as collaborative software and as personal information manager software.'

Feng Office Community Edition is an open source collaboration platform developed and supported by Feng Office and the OpenGoo community. It is a fully featured online office suite with a similar set of features as other online office suites, like Google Apps, Microsoft Office Live, Zimbra, and Zoho. The application can be downloaded and installed on a server. Feng Office could also be categorized as collaborative software and as personal information manager software.

Evaluation Setup

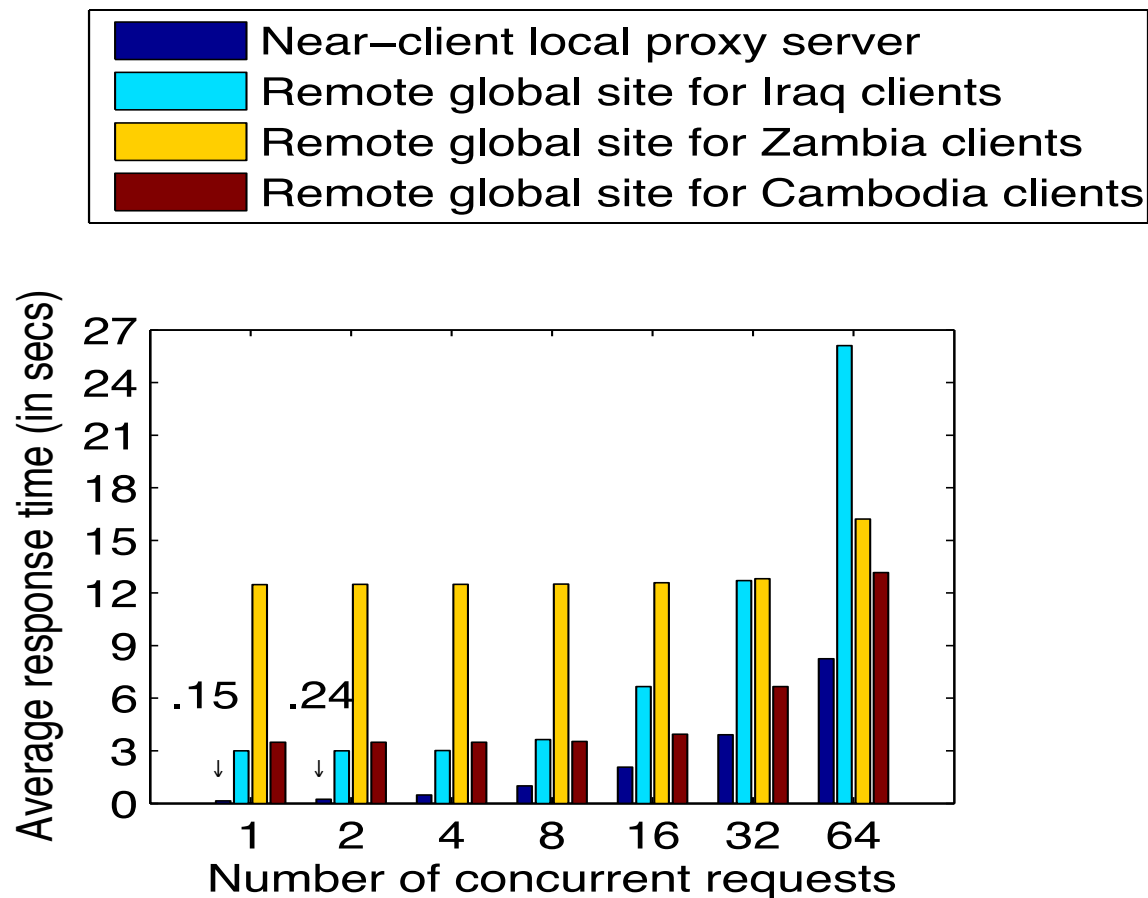
- Workload: Wikipedia document revision history for collaborative editing with a 9:1 read to write ratio.
- We use Linux Traffic Control tool (tc) to simulate the network bandwidth and latency for users in the three different sites.

Traffic location	Latency	Bandwidth
Iraq	1.48 seconds	106Kbps
Zambia	6.03 seconds	196Kbps
Cambodia	1.77 seconds	208Kbps

Evaluation Setup

- We use a low end machine as a local server, a single-core 2.66GHz Pentium 4 processor and 512MB memory, and a high end machine as a remote server, two dual-core (four cores total) 3.00 GHz Xeon CPUs and 4 GB DRAM memory.
- Simulate up to 64 concurrent local clients.

Evaluation results



Conclusions

- Web proxy is efficient for caching static pages.
- Web prefetching has limited benefit since its effectiveness is diminished by the web proxy and the high ratio of the dynamic web pages.
- Local Application Mirroring is a promising technique to accelerate the response time of dynamic web requests.