Introduction

Goal: Information driven applications in developing regions.

Challenges:
- Flaky network connectivity and equipment
- Applications need to be low cost, low power

Solution: Factor common complexity of building these applications into a shared storage infrastructure.
TierStore

Distributed Storage System

- Data center is the permanent data repository.
- Proxies in the villages cache data pertinent for their users.

Internet

City

Motorcycle

Dial-up Tel.

Village

LEO satellite

Motorcycle
Hierarchical Namespace

Data objects have names similar to files and directories:
- /joe/mailbox/my_email
- /bob/mailbox/my_email

Applications place subscriptions on portions of the namespace it uses:
- Subscribed data is automatically synchronized
- Data is transmitted only on demand

Namespaces express both application semantics and the unit of replication.
Conflict Resolution

Multiple updates from different locations

- Alice modifies her inbox in two different locations. How are these two versions of the inbox reconciled?

Namespaces has application specific types

- Application specific resolution of conflicts
- Example: 2 deletes of an e-mail message
Optimizations

TierStore also can optimize based on application behavior:

- Identical data only transmitted and stored once.
- Send deltas instead of whole data.
- Merge updates at intermediate proxies.
Current Status and Plan

What has been done:
- Subscriptions and synchronization updates
- Disconnected Web Proxy
- IMAP e-mail server

Plan:
- Finish e-mail proxy this Fall
  Eat our own dog food!
- Build XML document manager
- Voicemail