Architectures and Systems for Mobile-Cloud Computing: A Workload-Driven Perspective

Xin Zhang, Prashant Nair, Mayur Naik and Moin Qureshi

Background

Mobile devices have become the primary computing device

New Applications

Mobile Cloud Computing

Challenge 1: Interleaved I/O and computation

Flexible Offloading Scheme I

Challenge 2: Network latency

Flexible Offloading Scheme II

Mobile Cloud Computing

“Dialing 123-456-7890”

“Call John.”

Call John

Challenges

1. Interleaved I/O and computation

2. Network latency

3. Diverse and dynamic execution environment
Architectures and Systems for Mobile-Cloud Computing: A Workload-Driven Perspective

Xin Zhang, Prashant Nair, Mayur Naik and Moin Qureshi

Analytical Models for Tradeoffs

- Solution to Challenge 3
  - Network features
  - Hardware features
  - Software features
  - Analytical model
  - Runtime decision

Offloading System

- Offloading schemes
  - 3G
  - Wi-Fi
  - 4G
- How to offload
- What to offload
- Analytical model

Roadmap

- Mobile workload tracing (Infrastructure implemented)
  - Trace mobile workloads of top 150 Google Play apps
- Workload characterization (~3 months)
  - Identify features common and unique to mobile workloads
- Analytical models of performance and (~3 months) energy usage
  - Produce tolerable error bounds compared to hardware measurement
- Mobile-cloud computing system (~6 months)
  - Show speedup and energy savings

Preliminary Findings

- Mobile devices have become the new “PC”
- Big performance gap between mobile and desktop

Our Proposal:

- Enable desktop-class performance for mobile apps
  - By:
    - Offloading computation to the cloud
    - Using robust analytical modeling
- Enable new applications and usage models

Results of Tracing

- 24 threads
- 3 million function calls
- 17 million memory reads
- 13 million memory writes

Summary

- Mobile phones are the new “PC”
- Big performance gap between mobile and desktop
- Our Proposal:
  - Enable desktop-class performance for mobile apps
    - By:
      - Offloading computation to the cloud
      - Using robust analytical modeling
  - Enable new applications and usage models