cloud in the palm of your hands

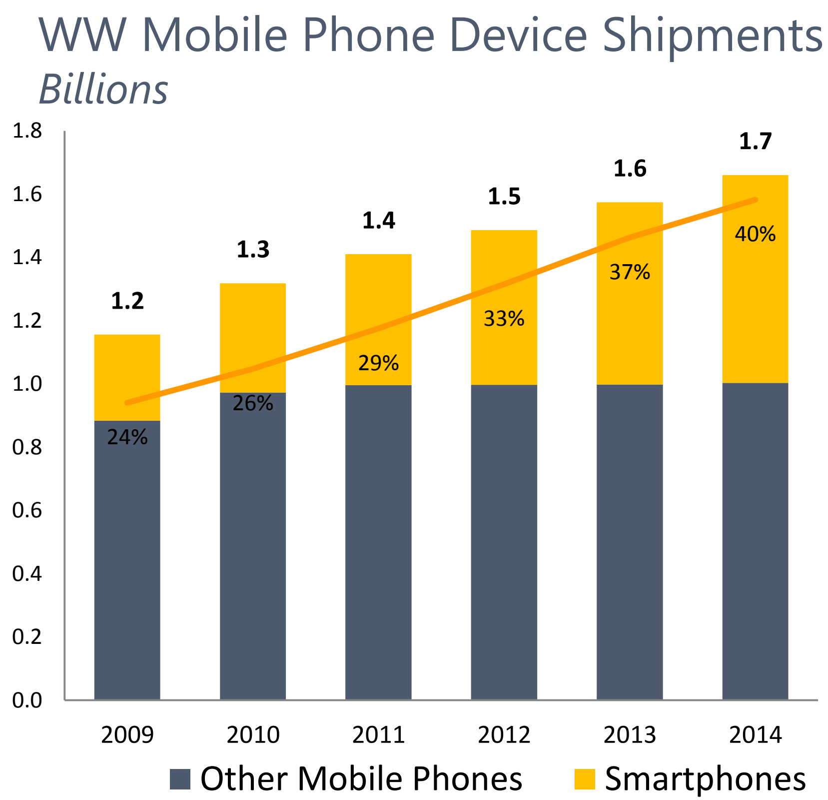
Victor Bahl

7.28.2011

mobile phone market

IDC FY12 forecast 518 million SmartPhones sold world-wide

- More smartphones shipped than PCs in FY11 Q2 (101M vs. 92M)



Source: IDC, iSuppli, Gartner, Accenture analysis.

sad reality of mobile computing

hardware limitations

- vs. static elements of same era (deskt
- weight, power, size constraints
- CPU, memory, display, ke

wireless comm ity

- band
- ity
- oney, require service agreements

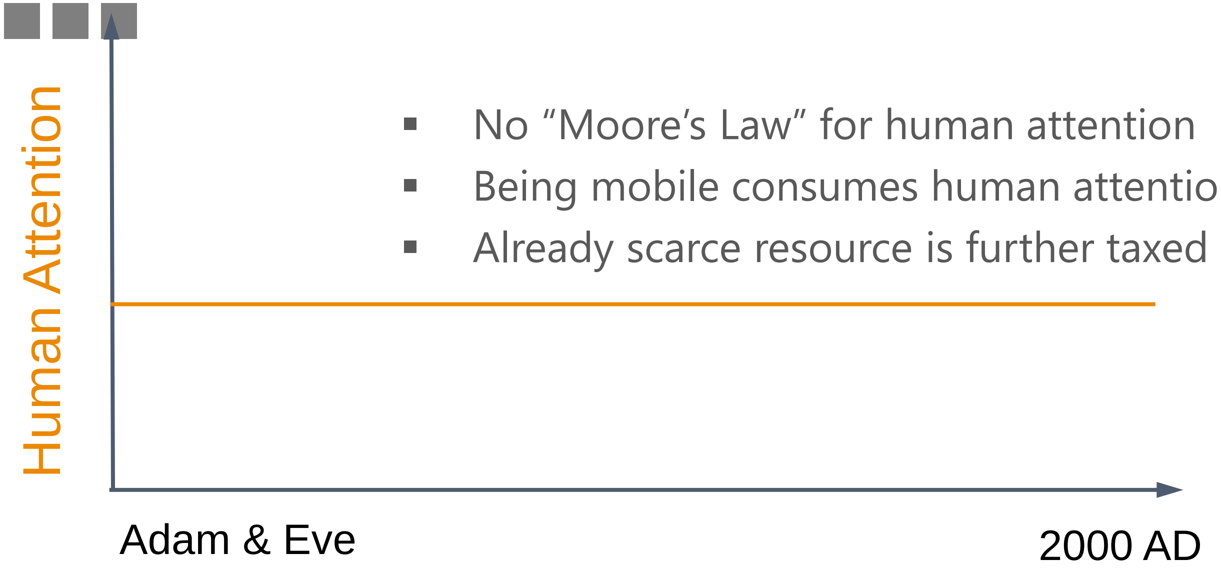
energy source

- actions may be slowed or deferred
- wireless communication costs energy

True 15+ years ago (early 1990s)

- huge hardware & wireless networking improvements since but deep essentials still the same. Will the same slide will be true in 2020?

why resource poverty hurts



- No "Moore's Law" for human attention
- Being mobile consumes human attention
- Already scarce resource is further taxed by resource poverty

Reduce demand on human attention

- Software computing demands not rigidly constrained
- Many "expensive" techniques become a lot more useable when mobile

Some examples

- machine learning, activity inferencing, context awareness
- natural language translation, speech recognition, …
- computer vision, context awareness, augmented reality
- reuse of familiar (non-mobile) software environments

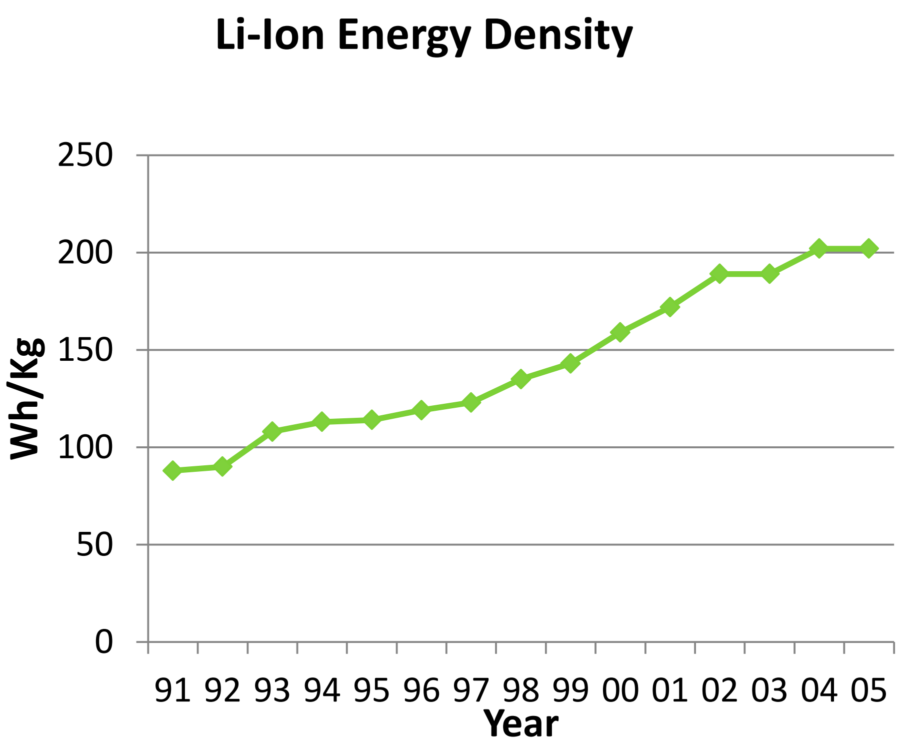
…

Clever exploitation needed to deliver these benefits

Vastly superior mobile user experience

Courtesy. M. Satya, CMU

battery trends

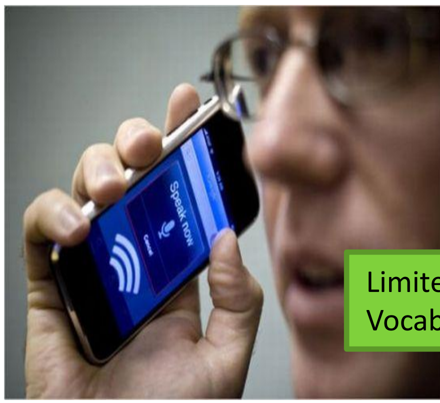


- Lagged behind
  - Higher voltage batteries (4.35 V vs. 4.2V) – 8% improvement
  - Silicon anode adoption (vs. graphite) – 30% improvement
- Trade-offs
  - Fast charging = lower capacity
  - Slow charging = higher capacity

- CPU performance improvement during same period: 246x
- A silver bullet seems unlikely

resource constraints prevent today's mobile apps from reaching their full potential

Speech recognition & synthesis



Limited Vocabulary

Augmented Reality

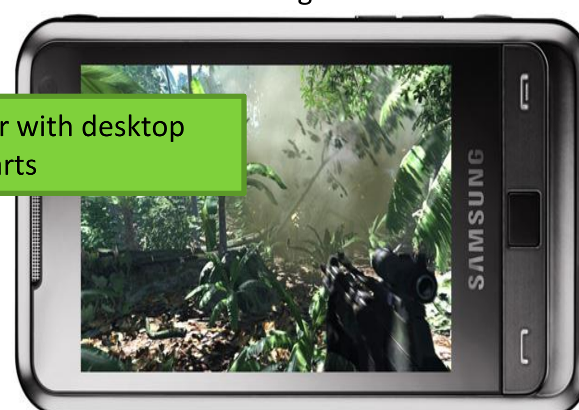
Too CPU intensive

Healthcare sensing & analysis

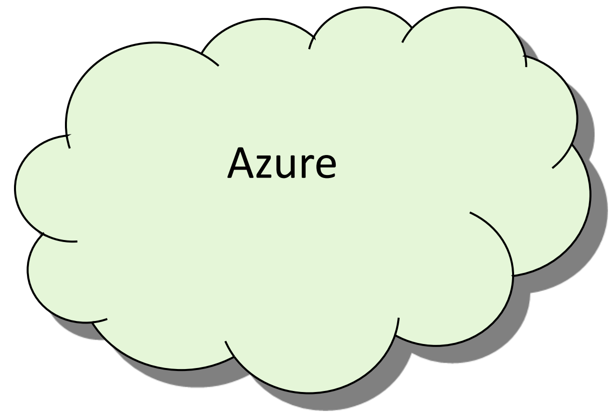
Too Energy intensive

3D Interactive Gaming

Not on par with desktop counterparts



marrying the phone to the cloud



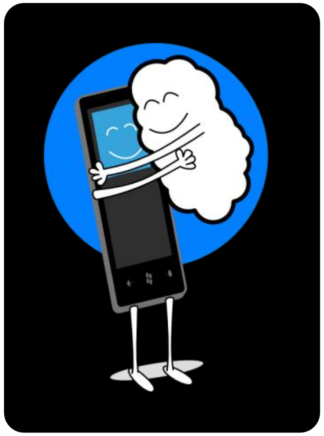
Phone offers ubiquitous connectivity and context awareness.

The cloud offers near-limitless resources

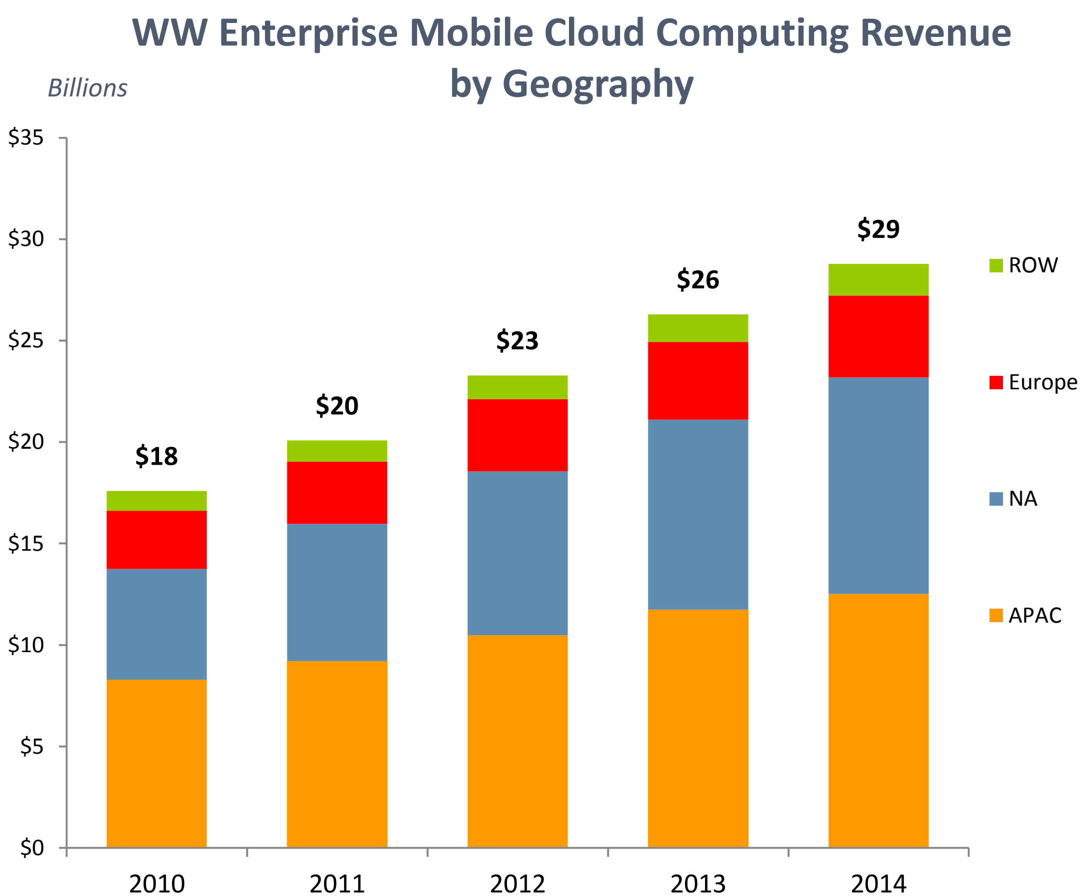
Together, they enable applications that were simply not possible before

vision: cloud in the palm of your hand

Enable mobile application developers to fully realize the potential of the cloud, and to do so quickly, reliably and easily.



mobile cloud computing market



- Today 91% of enterprise mobile cloud computing revenue derives from email and Internet accessibility applications but is expected to decline to 84% by 2014.
- Productivity enhancing applications will increase from 5% of overall spend in 2010 to 13% in 2014.

Source: ABI Research Accenture analysis

strategy: evolve the development model

- computational offloads (MAUI)
- cutting down latency & mitigating spectrum/bandwidth issues (Cloudlets)
- services & service composition framework (Hawaii)
- multi-phone programming for collaborative applications (Rainier)

Computational Offload

opportunistic use of the cloud

research challenges

- what to offload?
- how to dynamically decide when to offload?
- how to minimize programmer effort?

important for adoption: a simple programming model

- app developer community has varying expertise & skills
  - Cannot require app developers to become experts in distributed systems

strategy

- developers build standalone apps with simple annotations but no changes to program logic
- use of nearby and cloud-server resources is opportunistic
- result: applications adapt as their execution environment changes

enabling simple program partitioning

Programming Model

- Dynamic partitioning made simple for th partitioning
  – Programmer builds app as standalone phone app
  – Programmer adds .NET attributes to indi "remoteable" methods / classes

```
[Remoteable]
ArrayList GetValidMoves(Square s)
{
    if (s.IsEmpty())
    {
        return new ArrayList();
    }
    if (s.Piece.IsEnemyOf(active))
    {
        //this piece does not belong to the active side, no moves possible
        return new ArrayList();
    }
    //forward the call to the Rule-class
    return rules.getMoves(s);
}
```

Salient Point:
The model supports disconnected operations

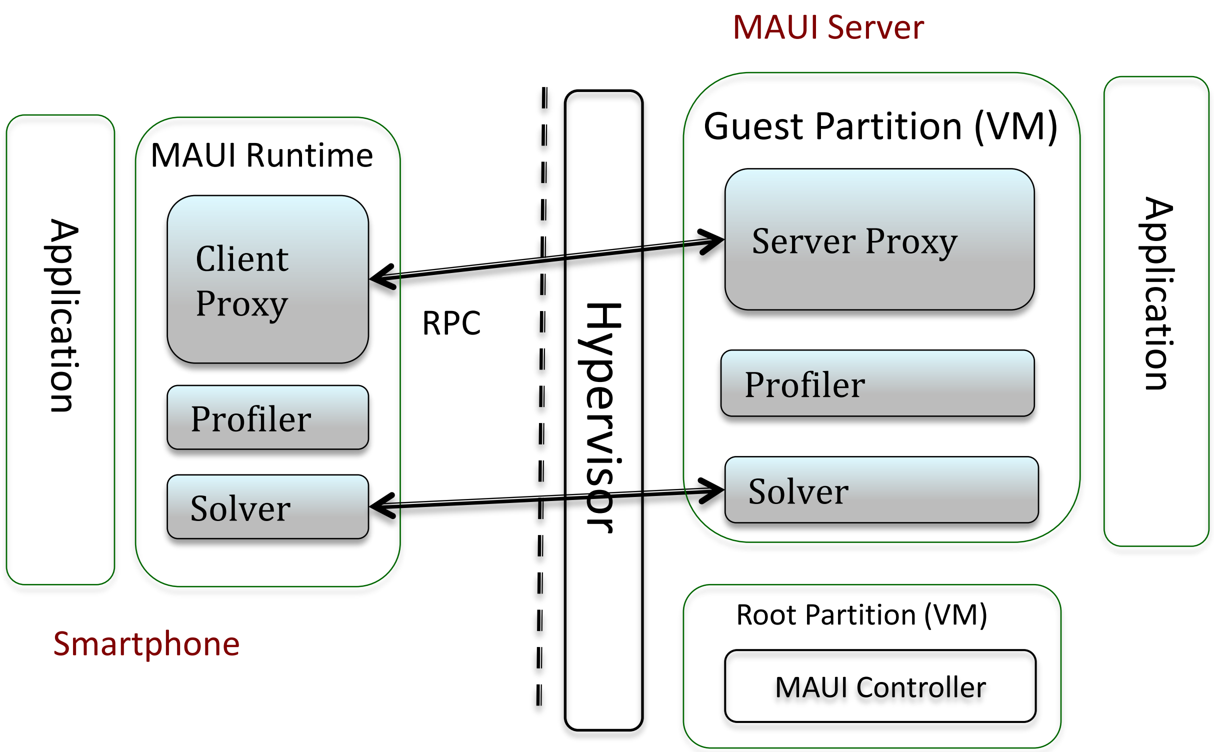
- MAUI runtime: partitions (splits) the program at run-time
  – Can optimize for energy-savings, or performance

Why not use a static client/server split?

– Developers need to revisit application structure as devices change
– Failure model: when phone is disconnected, or even intermittently connected, applications don't work
– The portion of an app that makes sense to offload changes based on the network conn. to the cloud server

dynamic offloading

Application Partitioning

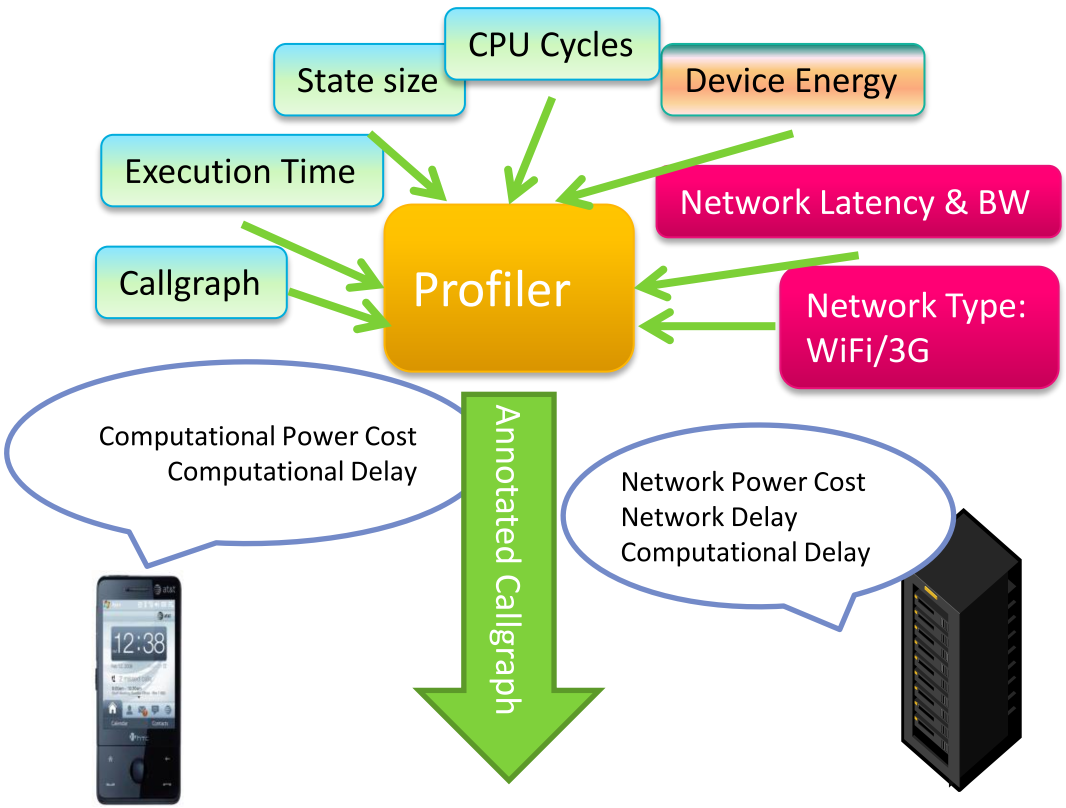


Currently supports client/server split, can be extended to multiple tiers

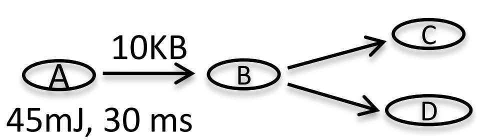
profiler and decision engine

Profiler:

Handles dynamics of devices, program behavior, and environment (Network, Server Load)



- Vertex: method annotated with computation energy and delay for execution
- Edge: method invocation annotated with total state transferred



Decision Engine:

Partition A Running App

We use an Integer Linear Program (ILP) to optimize for performance, energy, or other metrics...

Example – Maximize:

$$\sum_{v \in V} (I_v \times E_v) \quad - \quad \sum_{(u,v) \in E} (|I_u - I_v| \times C_{u,v})$$

energy saved — cost of offload

Such that:

$$\sum_{v \in V} (I_v \times T_v) \quad + \quad \sum_{(u,v) \in E} (|I_u - I_v| \times B_{u,v}) <= \text{Lat.}$$

execution time — time to offload

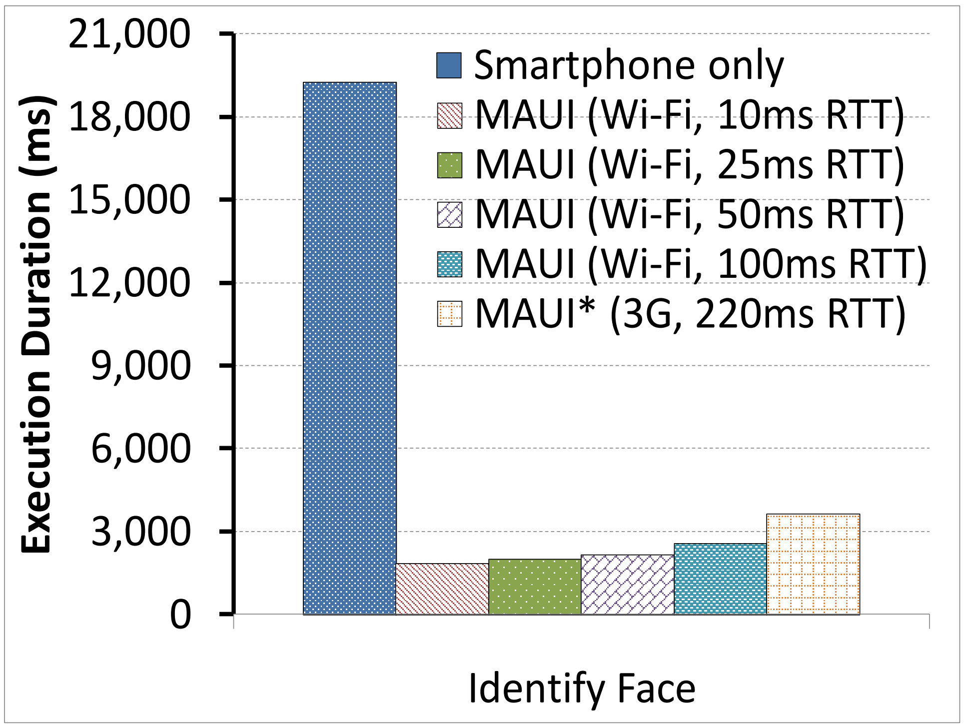
and

$I_v <= R_v$ for all $v \in V$

Offloading Performance and Energy Benefits

Performance Benefits:

Memory Assistant Face recognizer:



Face recognition becomes “interactive” w/ offload

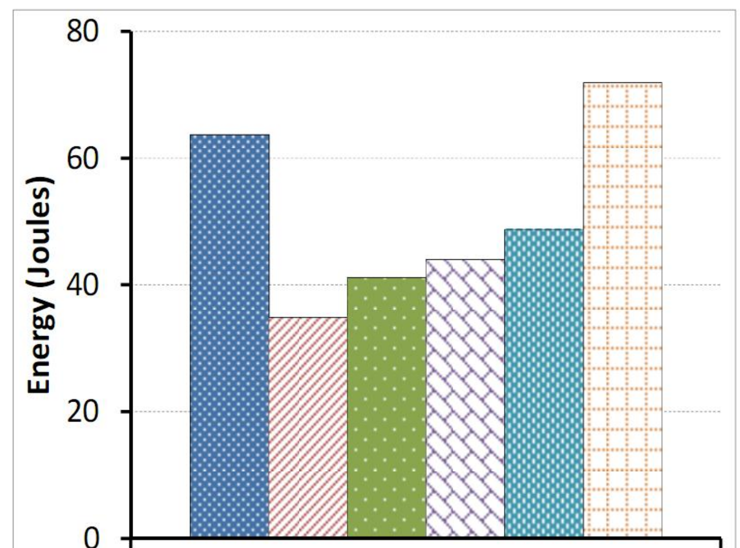
Energy Benefits:

Interactive arcade game w/physics engine:

Energy measurements from hardware power monitor

Arcade game benefits:

- Up to double the frame rate (6 -> 13 fps)
- Up to 40% energy reduction



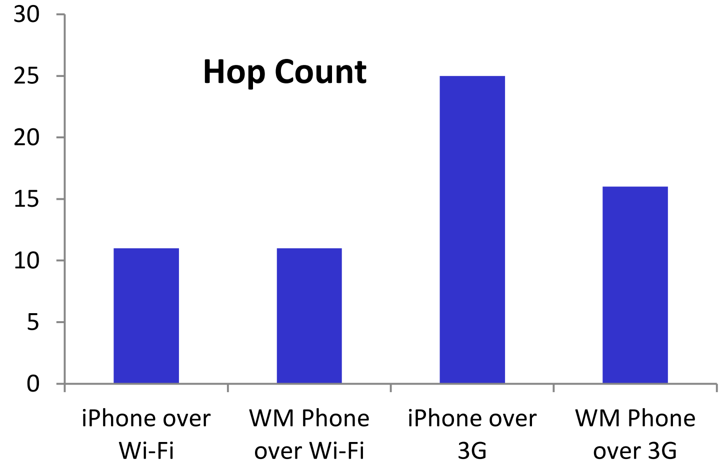
Cloudlets

latency

iPhone via Wi-Fi : 11 hop

Wi-Fi -> 209.85.225.99

1. (10.0.2.1) 8.513 ms 8.223 ms 9.365 ms
2. (141.212.111.1) 0.913 ms 0.606 ms 0.399 ms
3. (192.122.183.41) 11.381 ms 6.054 ms 5.975 ms
4. (192.12.80.69) 7.038 ms 7.353 ms 7.026 ms
5. (198.108.23.12) 12.525 ms 13.027 ms 12.619 ms
6. (198.110.131.78) 12.715 ms 9.424 ms 9.315 ms
7. (216.239.48.154) 9.974 ms (209.85.250.237) 10.295 ms (216.239.48.154) 9.405 ms
8. (72.14.232.141) 19.308 ms 22.249 ms 23.312 ms
9. (209.85.241.35) 32.987 ms 22.708 ms (209.85.241.27) 124.588 ms
10. (72.14.239.18) 22.256 ms (209.85.248.106) 29.154 ms (209.85.248.102) 21.635 ms
11. (209.85.225.99) 19.973 ms 21.930 ms 21.656 ms



traceroute to 209.85.225.99 (one of the server IPs of www.google.com)

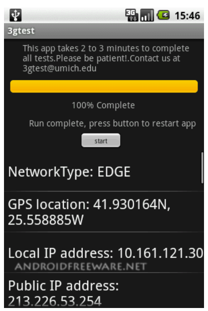
iPhone via 3G : 25 hop

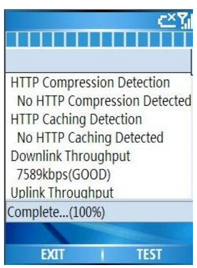
3G -> 209.85.225.99

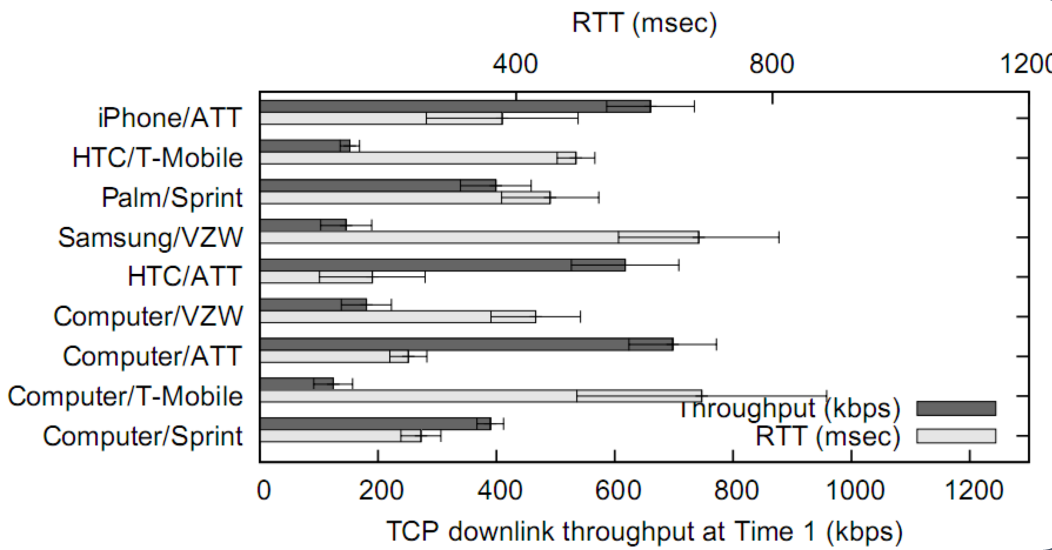
1. * * *
2. (172.26.248.2) 414.197 ms 698.485 ms 539.776 ms
3. (172.16.7.82) 1029.853 ms 719.595 ms 509.750 ms
4. (10.251.11.23) 689.837 ms 669.340 ms 689.739 ms
5. (10.251.10.2) 509.781 ms 729.746 ms 679.787 ms
6. (10.252.1.7) 719.652 ms 760.612 ms 788.914 ms
7. (209.183.48.2) 689.834 ms 599.675 ms 559.694 ms
8. (172.16.0.66) 539.712 ms 809.954 ms 689.547 ms
9. (12.88.242.189) 589.857 ms 1129.848 ms 709.784 ms
10. (12.122.138.38) 589.699 ms 1009.723 ms 769.808 ms
11. (12.122.138.21) 669.690 ms 529.758 ms 699.965 ms
12. (192.205.35.222) 699.569 ms 979.769 ms 1489.869 ms
13. (4.68.19.190) 699.435 ms (4.68.19.126) 559.875 ms (4.68.19.62) 499.598
14. (4.69.136.149) 889.946 ms (4.69.136.141) 879.443 ms (4.69.136.145) 469.601 ms
15. (4.69.132.105) 559.716 ms 539.754 ms 1219.982 ms
16. (4.69.132.38) 719.700 ms 659.613 ms 539.695 ms
17. (4.69.132.62) 549.752 ms 549.640 ms 800.128 ms
18. (4.69.132.114) 669.729 ms (4.69.140.189) 769.711 ms 959.663 ms
19. (4.69.140.193) 959.735 ms 979.674 ms 849.886 ms
20. (4.68.101.34) 649.609 ms 659.767 ms (4.68.101.98) 1119.996 ms
21. (4.79.208.18) 669.405 ms 629.574 ms (209.85.240.158) 1200.039 ms
22. (209.85.240.158) 769.538 ms (72.14.232.141) 729.505 ms (209.85.241.22) 719.715 ms
23. (209.85.241.22) 769.665 ms (209.85.241.35) 769.880 ms 859.536 ms
24. (209.85.241.29) 589.710 ms (66.249.95.138) 789.762 ms (209.85.248.106) 913.287 ms
25. (209.85.225.99) 716.000 ms (66.249.95.138) 1039.963 ms (72.14.239.18) 899.607 ms

cloud computing has its challenges

End-to-end latency hurts interaction quality (crisp interaction essential for low demand on human attention)

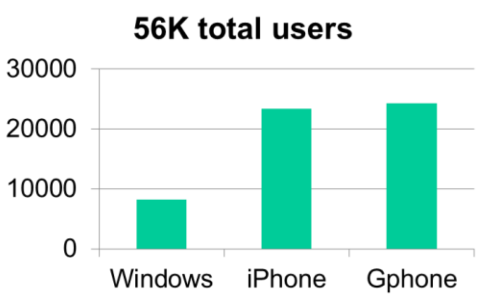






High loss rate & low throughput severely limits the scope of cloud services

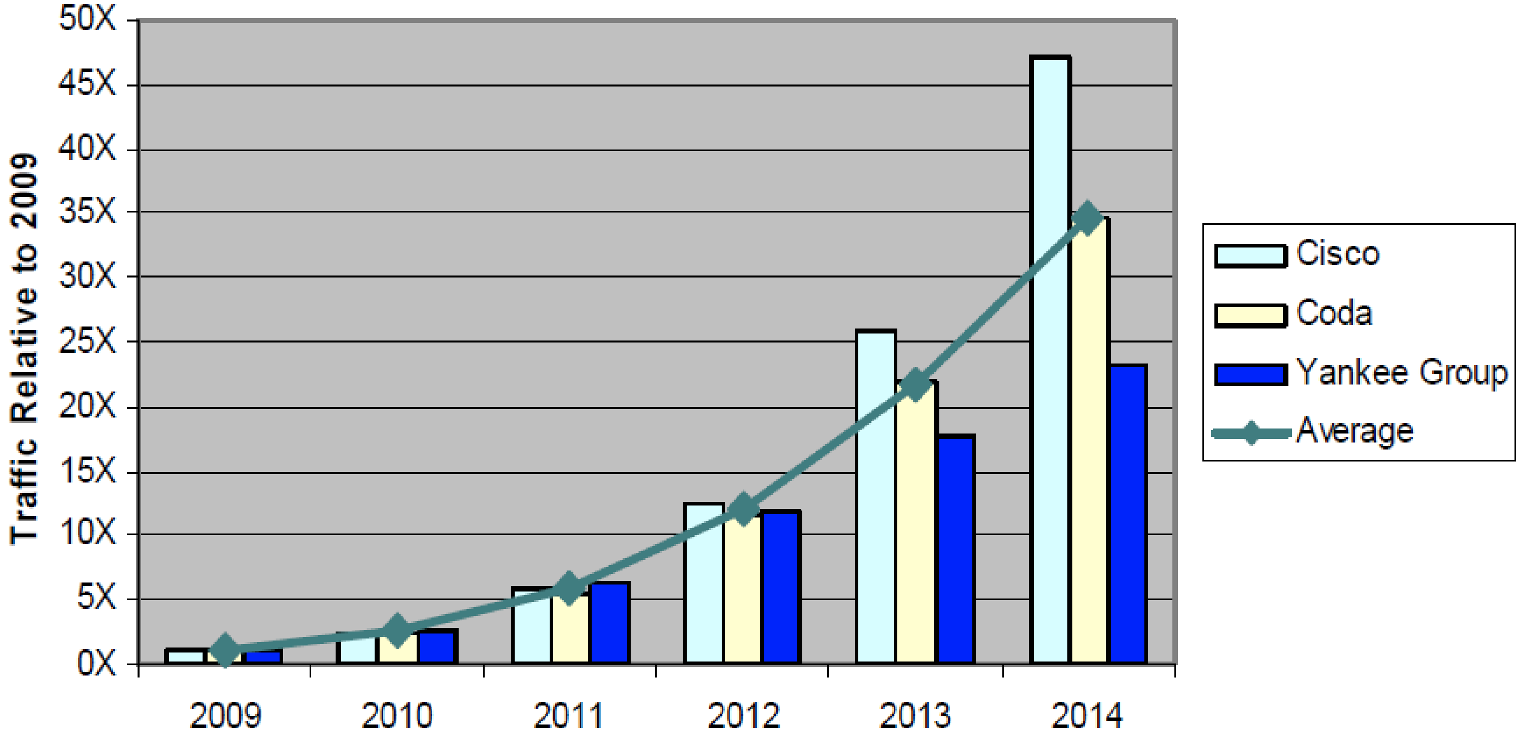
http://www.eecs.umich.edu/3gtest



demand

Industry Forecasts of Mobile Data Traffic

FCC, Staff Technical Paper, “Mobile Broadband: The Benefits of Additional Spectrum”, OBI Technical Paper No. 6 (Oct. 2010),

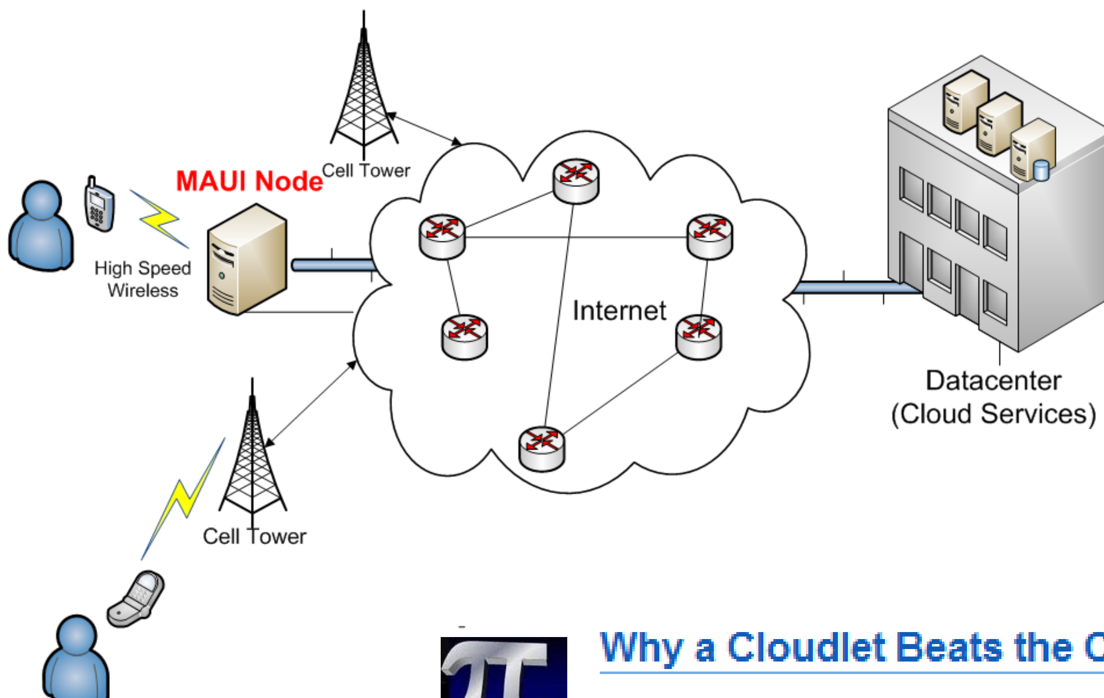


AT&T reported 106M Wi-Fi connections in 2010 Q3; versus 85M in all of 2009

AT&T’s mobile data traffic has experienced a fifty-fold increase over a three year period

cloudlets: defined

a resource rich infra-structure computing device with high-speed Internet connectivity to the cloud that a mobile device can use to augment its capabilities and enable applications that were previously not possible



Why a Cloudlet Beats the Cloud for Mobile Apps

Posted on December 13, 2009 by lewisshepherd

sample deployment scenario

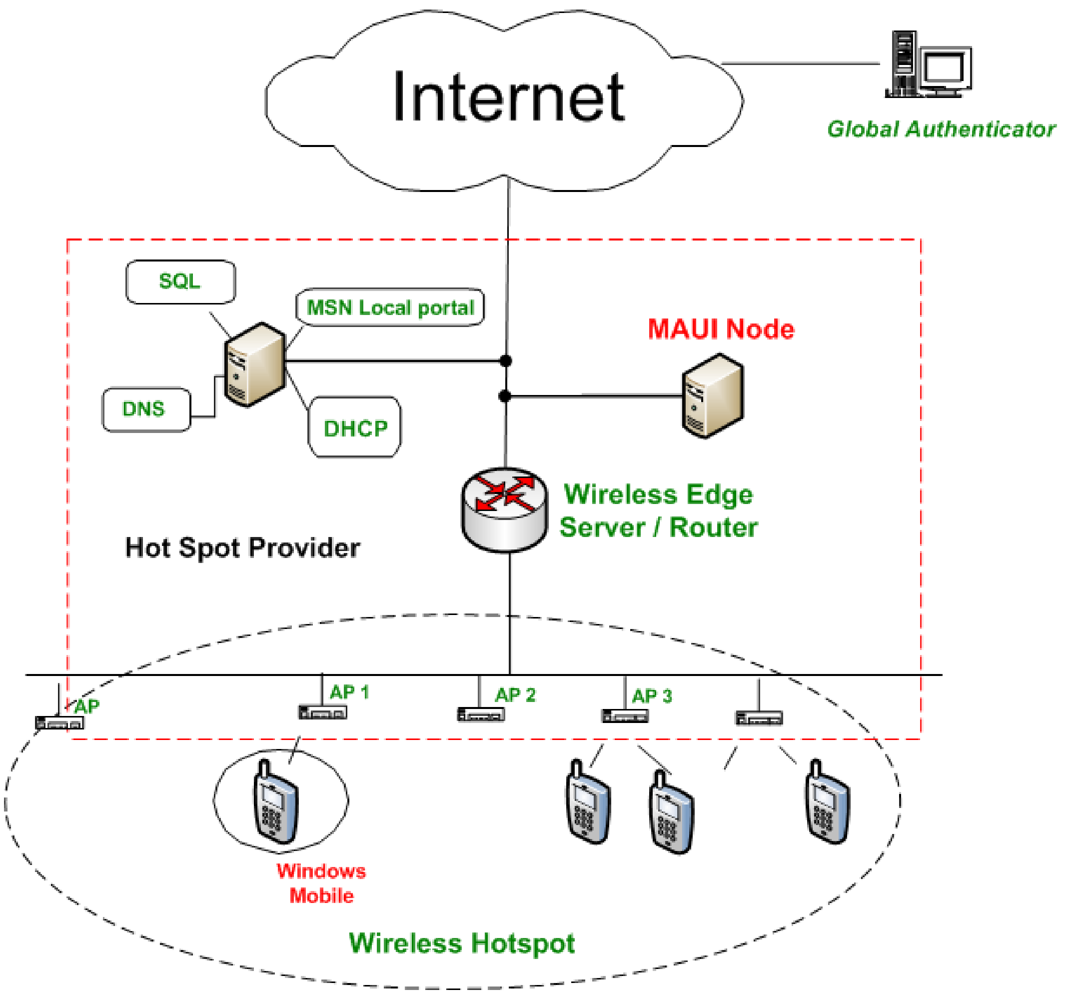
augment wi-fi hot spots with cloudlets
(in publics spaces & enterprise networks)

advantages

- does not use cellular spectrum
- short round-trip-times between mobile & cloud(let)
- optimal performance

research challenges

- security & privacy



cloudlets: properties

- better application performance
  - human attention management
- new application / behavior enablement
- reduced (manageable) latency
- extensible computing horsepower
- efficient spectrum usage - improved congestion / bandwidth management
- longer battery life

Cloud Services

cloud services and their composition

infrastructure for cloud-enhanced mobile applications

service toolbox

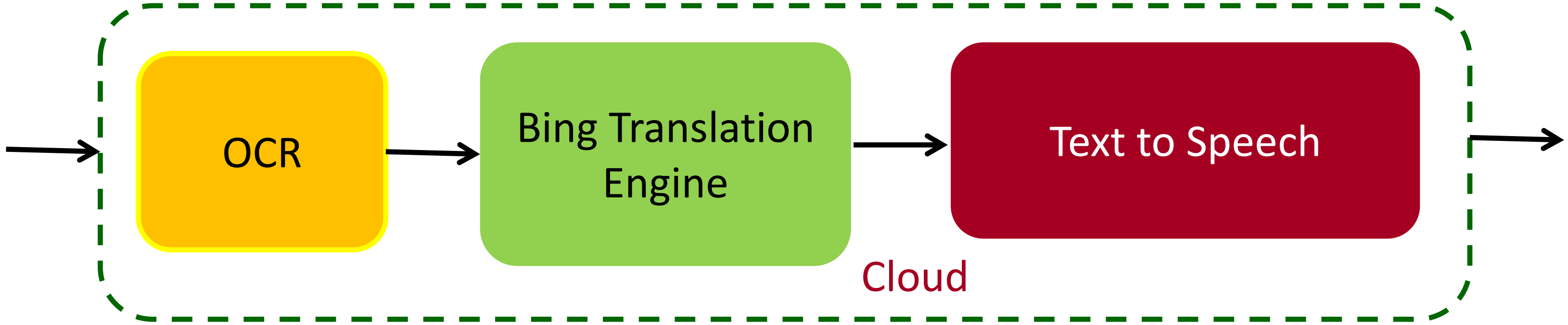
sophisticated resource intensive algorithms running in the cloud

- typically CPU, memory & storage intensive
- battery and/or bandwidth hungry

why do it?

- innovative combination of services lead to new sophisticated applications
- wp get enterprise-class applications & competitive advantage
- stickiness

language translation



cloud services

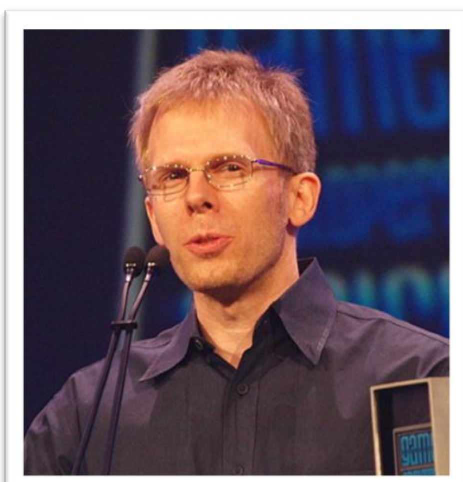
… build world-class cloud services that enable application developers to easily realize the full potential of mobile computing

Available (alpha)	In Progress	Under Consideration
Relay: Phone to phone data transfer	Adaptive App Profiling and Analytics	Face recognition
Rendezvous: Lookup for Relay endpoints	Trajectory Prediction	Location Sharing
Optical Character Recognition	Generalized Image Processing	Generic machine learning
Speech2Text	Mobile Game Matchmaking	Various wrappers: translation, text to speech etc.
Service composition framework		

Working with Windows Phone Services and Bing Mobile teams

killer application is killing time!

Windows Phone 7
Top 10+ apps are games



John Carmack (Wolfenstein 3D, Doom, Quake)…

“multiplayer in some form is where the breakthrough, platform-defining things are going to happen in the mobile space”

multiplayer mobile gaming: key challenge

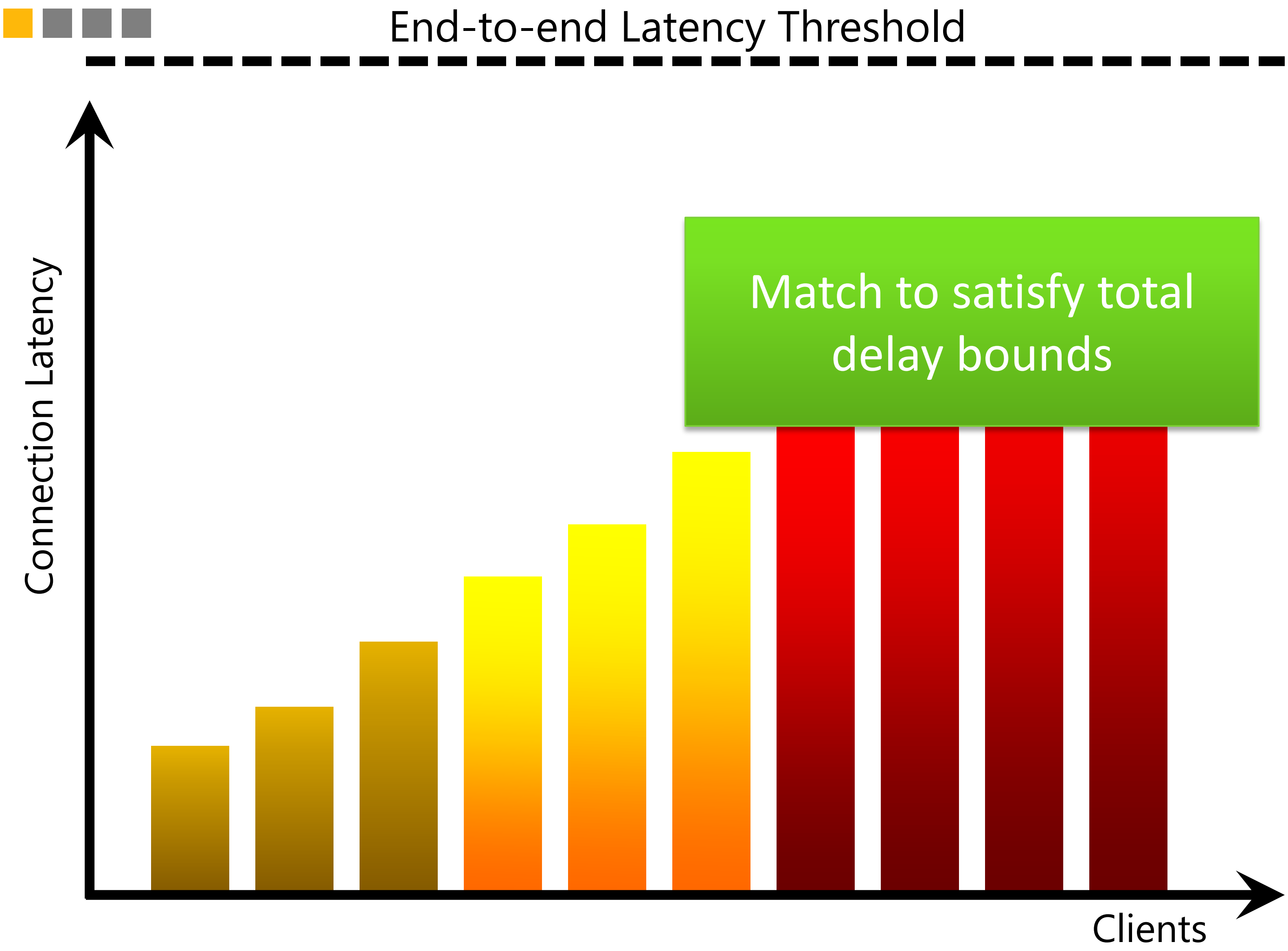
Bandwidth is fine: 250 kbps to host 16-player Halo 3 game

Delay bounds are much tighter

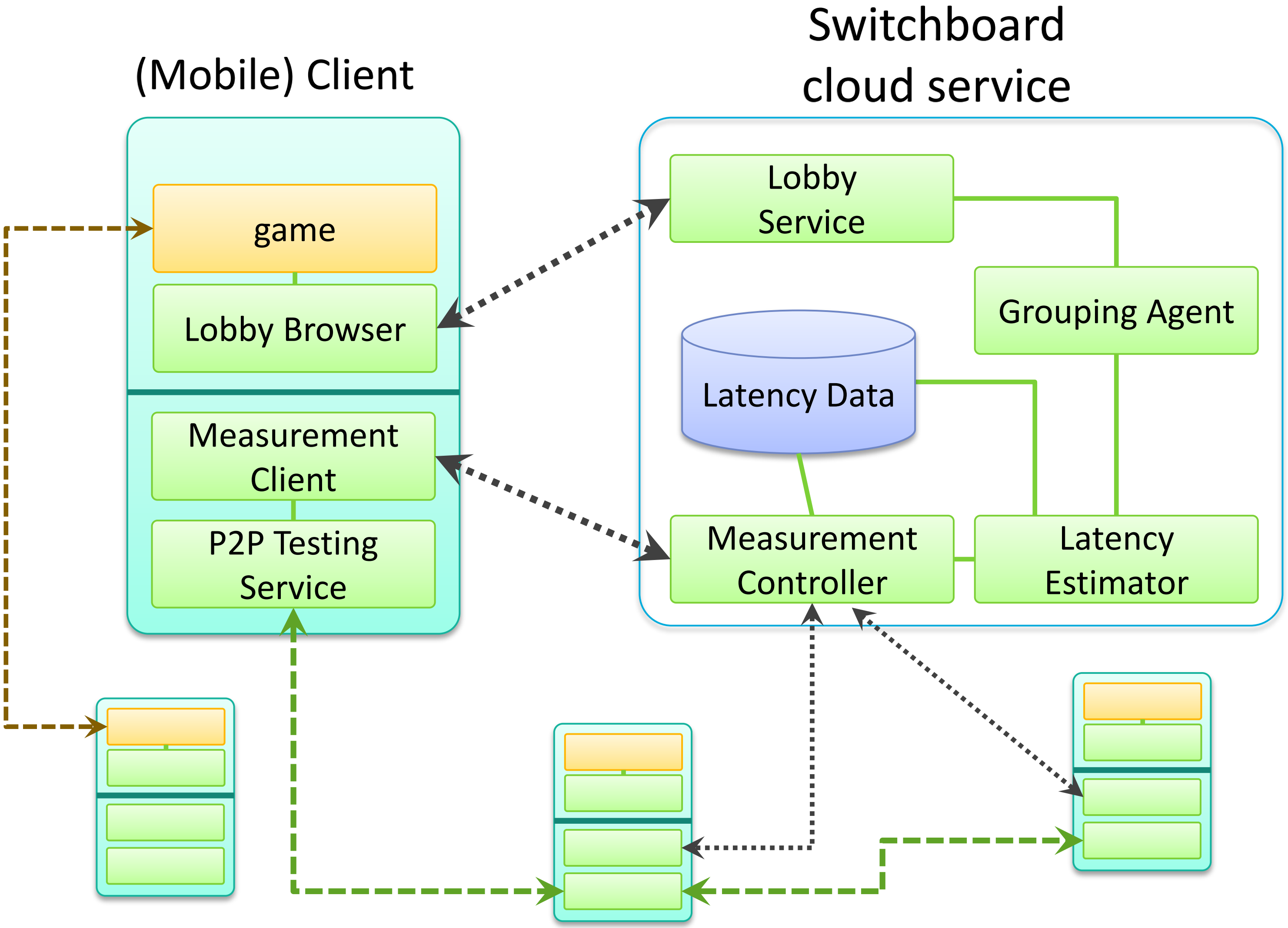
Challenge: find groups of peers than can play well together

Game Type	Latency Threshold
First-person, Racing	≈ 100 ms
Sports, Role-playing	≈ 500 ms
Real-time Strategy	≈ 1000 ms

the matchmaking problem

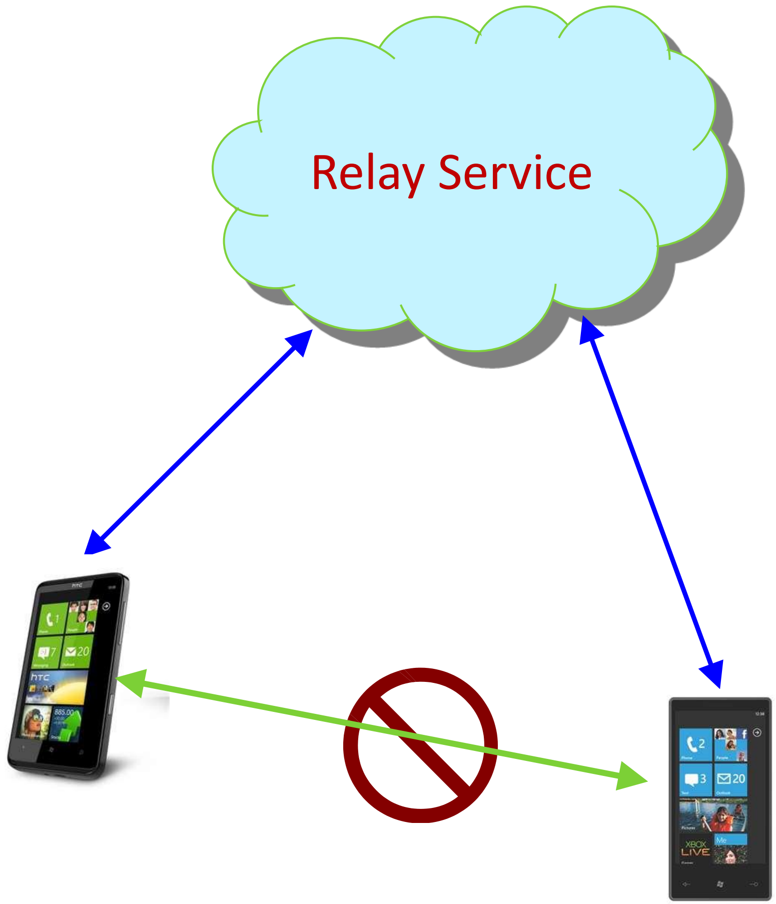


matchmaking service

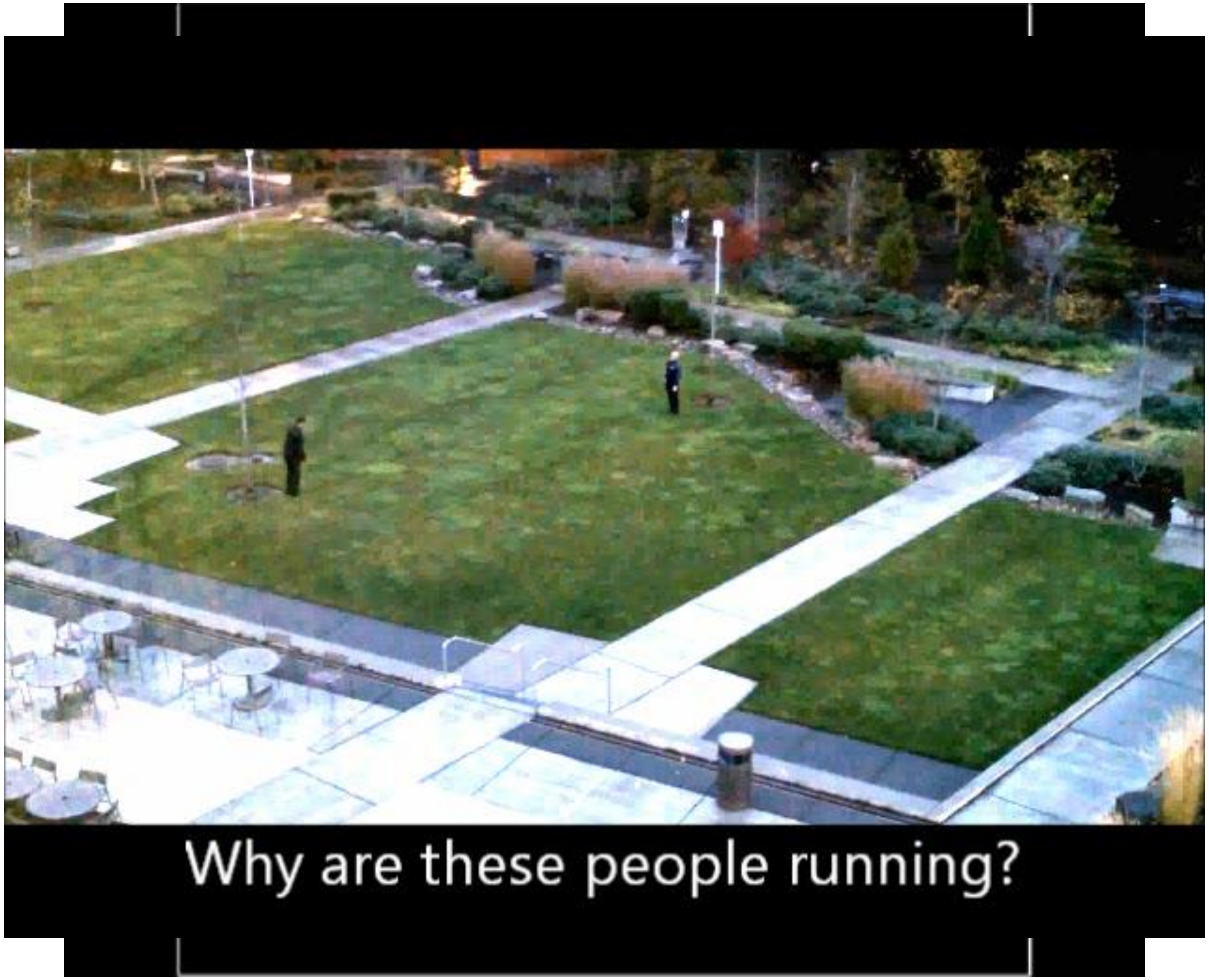


connecting phones via the relay service

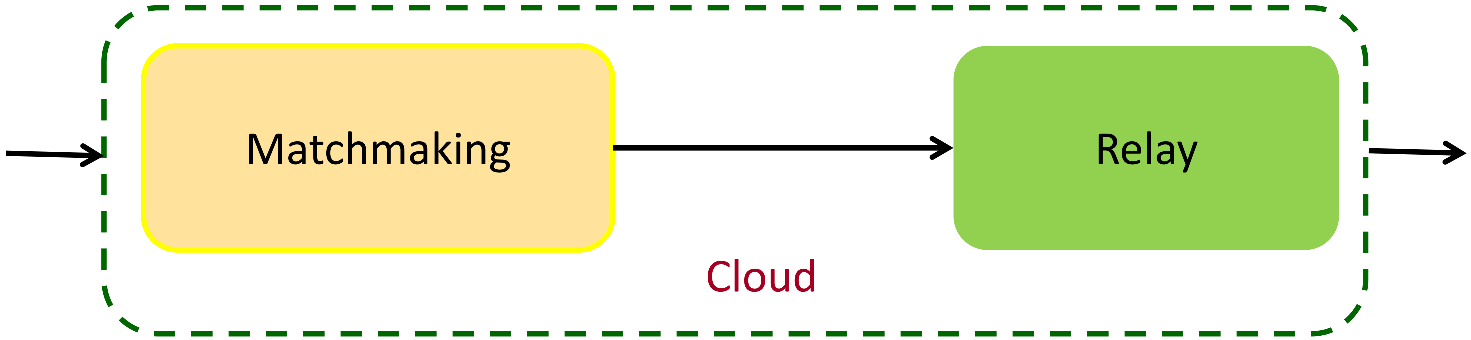
- For use when phones cannot communicate directly
- Generalized pub-sub service
  - With acks, ACLs, TTL
- Paired with Rendezvous Service discovery



light cycles: AR, multi-player, fast-action game



watching video together



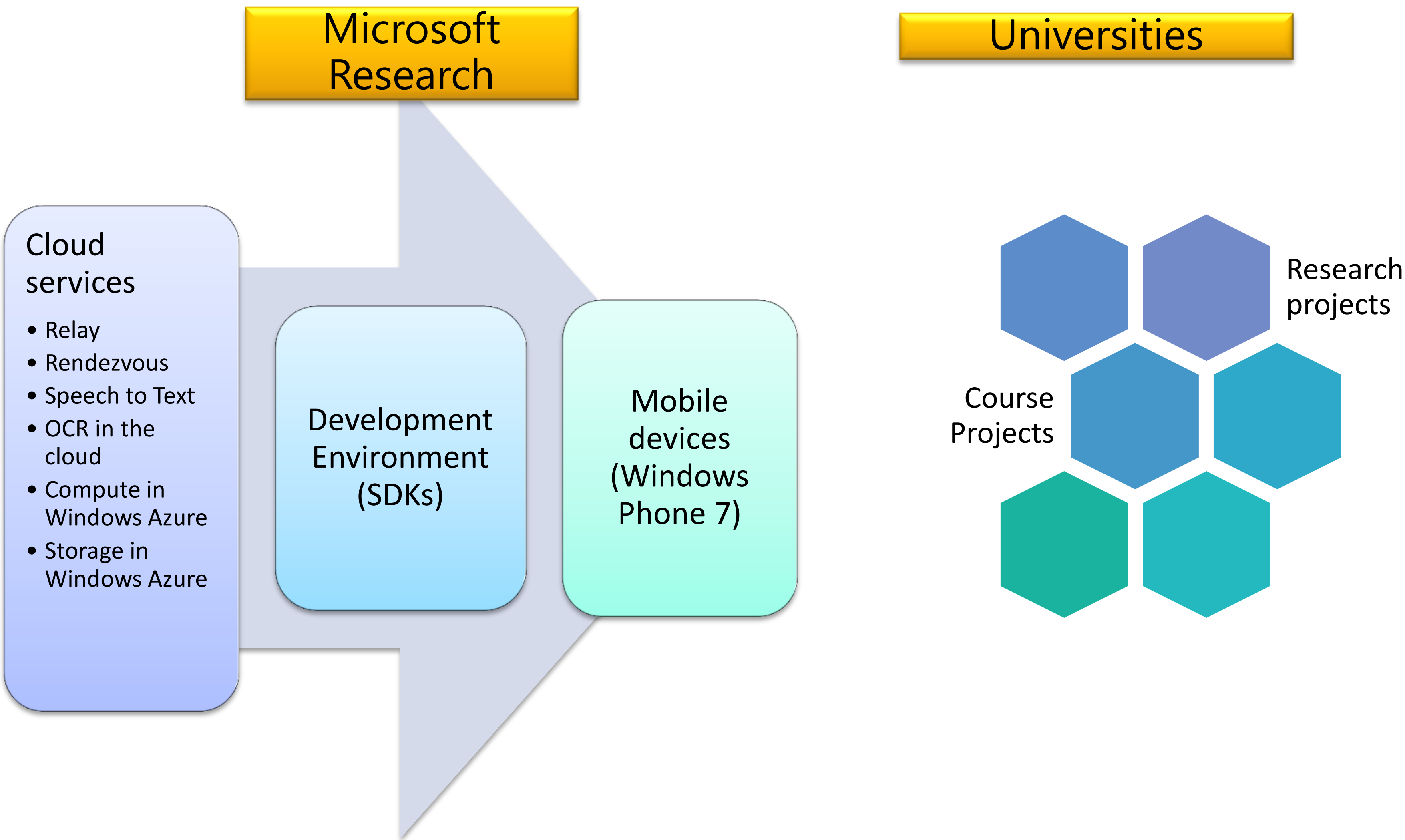
what happens to deployed apps?

testing the concept: project hawaii

Effort to investigate the ability of the **cloud** to enhance end-user experience on **mobile** devices

- To unleash the creative power of students by lowering barriers to writing mobile + cloud apps

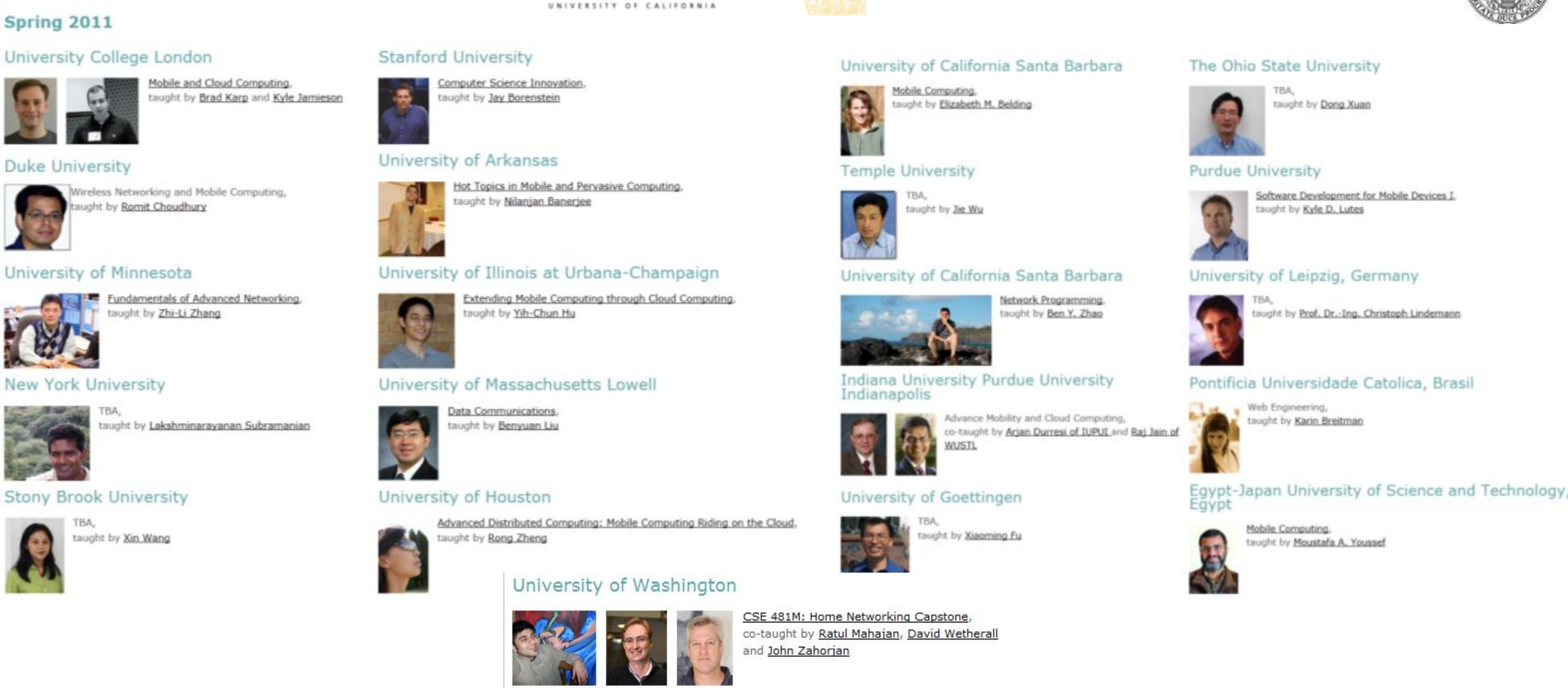
what does hawaii offer?



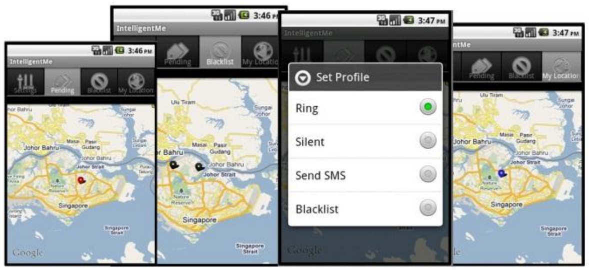
spring 2011 semester

- Launched Hawaii in **21** universities
- Close to **300** students start using Hawaii (with Windows Phone 7)

STANFORD UNIVERSITY · ILLINOIS · UNIVERSITY OF ARKANSAS · UNIVERSITY OF MINNESOTA Driven to Discover · UCL · UNIVERSITÄT LEIPZIG · PURDUE UNIVERSITY · Duke UNIVERSITY · UC SANTA BARBARA UNIVERSITY OF CALIFORNIA · UNIVERSITY OF MICHIGAN · USC Viterbi School of Engineering

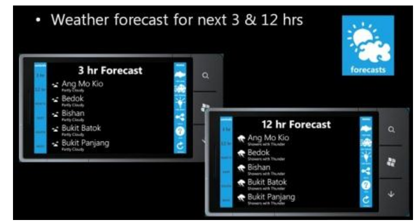


student developed applications



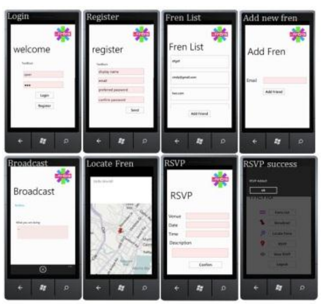
intelligentME

ReceiptManager

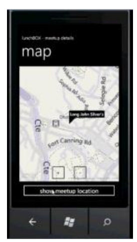


DaySaver

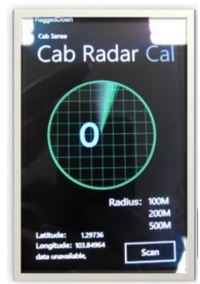
Network Forecaster



myFrens



LunchBox

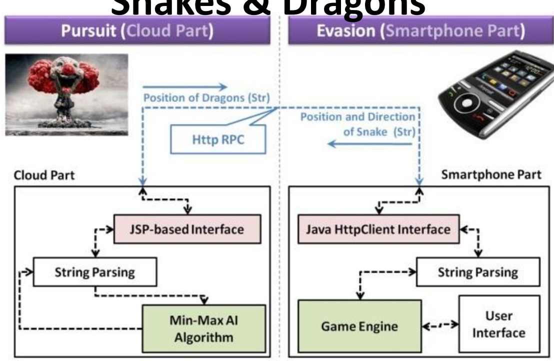


Flagged Down

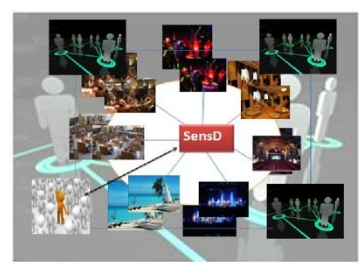
MobiProg

Parking Assistant

Snakes & Dragons



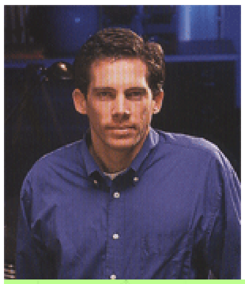
Activity Classification



SensD

Image Stitching

what are the academics saying?



"Project Hawaii, Windows Azure and Windows Phone 7 have collectively enabled my students to explore and develop new technology applications that push the boundaries of scientific data gathering beyond what was possible just a few years ago."

PROFFESSOR JAY BORENSTEIN, CS Department, Stanford University

UMASS LOWELL

Computer Science Department News

Students Demonstrate Smart Phone Applications for Data Communications II

By Martin, Fred on May 22, 2011 9:37 AM | Permalink | TrackBacks (0)

Prof. Benyuan Liu recently received an equipment grant from Microsoft Research to introduce Windows 7 phones and related mobile cloud services in 91.564 Data Communications II. The grant was received in time for the Spring 2011 semester, and Liu integrated the technologies into his course.

Prof. Liu reported that "students were very excited about the opportunity to use the advanced mobile cloud technologies from Microsoft Research, and learned a great amount about the principles and practice of wireless networking and mobile computing through the smart phone application developments."

On May 18, students demonstrated the applications they have developed for the course projects, ranging from smartphone RSS news reader to providing various services for the community. The projects were:

- *UML Parking Finder* by Peng Xia and Shan Lu, for finding parking spaces at UMass Lowell.
- *ZurianSwap* by Steve Bilozur, Swapnil Gewande and Ian White for sign translation using smartphones.
- *iBridge: Augmenting Reality with Barcode* by Ke Huang and Liuying Peng, to scan product barcodes with smartphone and obtain relevant information (e.g., stores nearby, compare price, nutrition analysis, etc).
- *UML Shuttle Tracker* by Jason Chan, I-Hsuan Lin and Xiawei Liu, a user friendly smartphone application to look up the shuttle bus's location in real time.
- *SleepSafe* by Bhanu Kaushik, a smartphone based approach for sleepwalking detection.
- *Language Translator* by Darshan Darbari and Rachit Mathur, language translator on smartphones for storing translated text.
- *UML App* by Kavya Kona, Prathiba Dyavegowda and Sunil Kumar Balaganchi Thammaiah, All-you-want to know information about UMass Lowell (shuttle services, dining services, on-campus residence, athletic centers, libraries, emergence notifications, etc)

ENGINEERING AT ILLINOIS

PROSPECTIVE STUDENTS | CURRENT STUDENTS | ALUMNI | CORPORATIONS | GIVING | FACULTY & STAFF | ABOUT US | RESEARCH

HOME \ ECE COURSE ENABLES STUDENTS TO CREATE SMARTPHONE APPS

NEWS
- Selected Publications
- Engineering Update
- Contact Us

ECE course enables students to create smartphone apps

June 6, 2011

This spring, computer engineering students capped off their undergraduate careers by developing smartphone applications that utilize cloud computing.

ECE 498HP: Extending Mobile Computing Through Cloud Computing was a new class spearheaded by Constantine Polychronopoulos, a professor of electrical and computer engineering (ECE) and assistant professor Yih-Chun Hu. In the course, students studied how to extend cloud computing resources to mobile smartphone users.

Microsoft provided access to their Windows Phone 7 devices and Windows Azure cloud services for the course. "From Microsoft's standpoint, the phone is the next big computing platform," Hu said. As battery life becomes more of an issue—with people's increasing reliance on energy-draining tasks from their smartphones—cloud computing becomes an important option.

"If your app does something sophisticated that you don't want to run on the device itself, you can push it out to the cloud," Hu said.

ECE 498 provided students with experience in developing applications using the growing remote computing technology. Of the five apps developed during the course, three received special recognition and a cash prize sponsored by Microsoft.

Yih-Chun Hu

Mapster: Cloudiest Application

next step: going broad & global

- transition from Universities to “real-world”
  - create an SLA for services to attract application developers across the world
- avoid the tragedy of the commons
  - everyone should get a fair share
- mechanism for supporting applications that become popular using our services
- challenge: avoid the dilemma of maintaining high SLA without incurring high operational or capital cost

Summarizing:

Putting the power of the cloud in the palm of your hand

services + offloading
+ rainier

+

=

plethora of enterprise class mobile computing apps

Thanks!