# Edge Computing

## What is it Good For?

**Mahadev Satyanarayanan**

**School of Computer Science**

**Carnegie Mellon University**

Joint work with: Brandon Amos, Yoshihisa Abe, Zhuo Chen, Benjamin Gilbert, Kiryong Ha, Jan Harkes, Martial Hebert, Wenlu Hu, Miike Kozuch (Intel), Roberta Klatzky, Padmanabhan Pillai (Intel), Wolfgang Richter, Dan Siewiorek, Junjue Wang

http://www.istc-cc.cmu.edu/

Intel Science & Technology Center for Cloud Computing

# Precious Real Estate

New York

London

Shanghai

Tokyo

San Francisco

# Another Kind of Precious Real Estate


# Edge of the Internet



NOKIA
Flex base station and multicontroller

European Telecommunications Standards Institute
**Mobile Edge Computing Initiative**
Industry Specification Group (ISG)

**Bringing Compute and Storage to Base Stations**

# 3 Rules for Success in Real Estate

1. **Location**

2. **Location**

3. **Location**

# What is a Cloudlet?

*aka "micro data center"*
*aka "mobile edge cloud"*

***Small data center at the edge of the Internet***

- **located one wireless hop away from mobile devices**
- **multi-tenant, just as in cloud computing**
- **heterogeneous environments (VM-based guests)**

**Non-constraints (relative to mobile devices)**

- **typically wired connection to Internet**
- **energy**
- **weight/size/heat**

***Catalyst for new mobile applications***

# Unique Value of Cloudlets

| | |
|---|---|
| 1. Highly responsive cloud services | Latency |
| 2. Edge analytics in IoT | Bandwidth |
| 3. Exposure firewall in the IoT | Privacy |
| 4. Mask disruption of cloud services | Availability |

# Growing Mindshare

**First Mobile Edge Computing Congress (Sept 29-30, 2015 in London)**

**OpenStack Summit (October 29 2015, Tokyo)**

**First IEEE Symp. on Edge Computing (Oct 27-28, 2016 in Washington DC)**

## Original Motivation for Cloudlets:

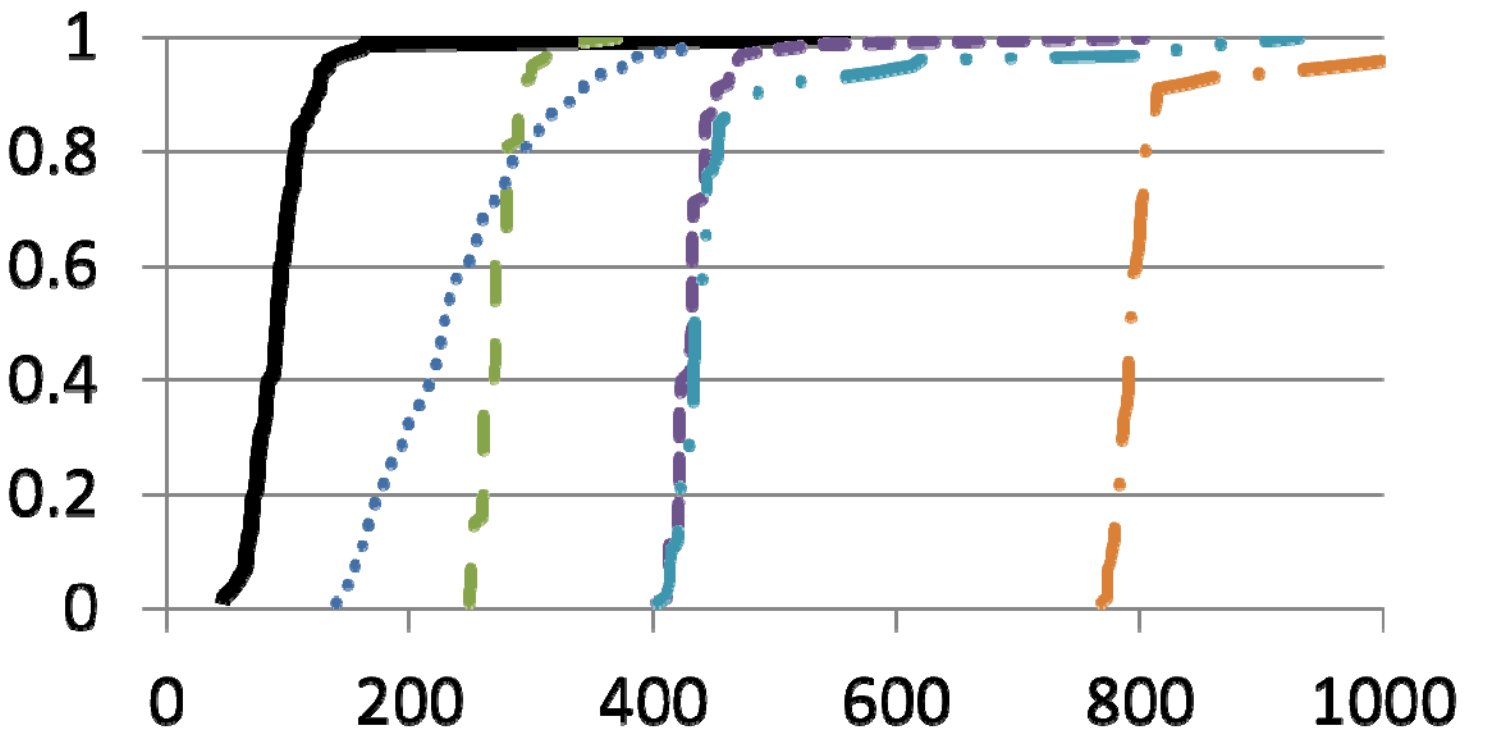
# Latency, Latency, Latency

# Does Latency Really Matter?

# Augmented Reality

CDF of Response Times (milliseconds)

Amazon EC2 offload

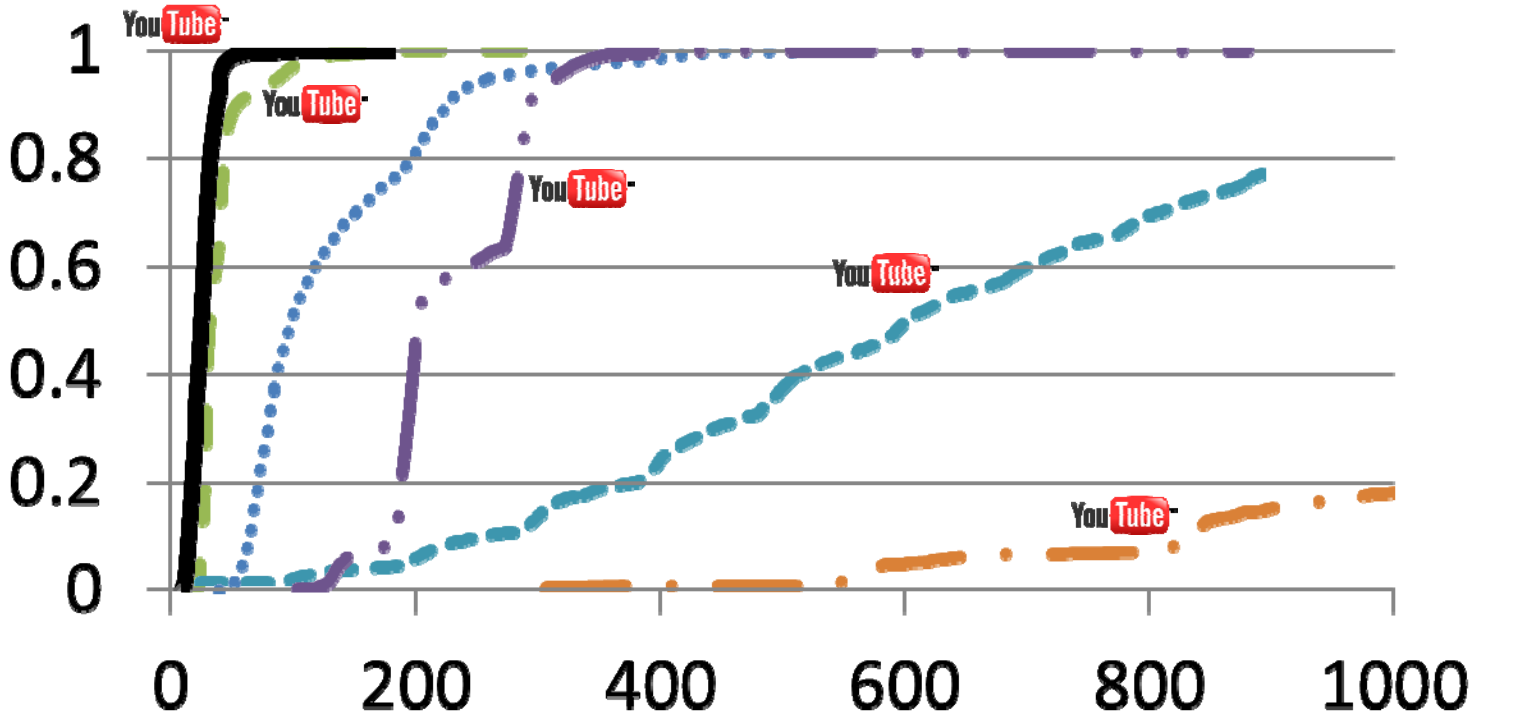


# Simulation-based Graphics

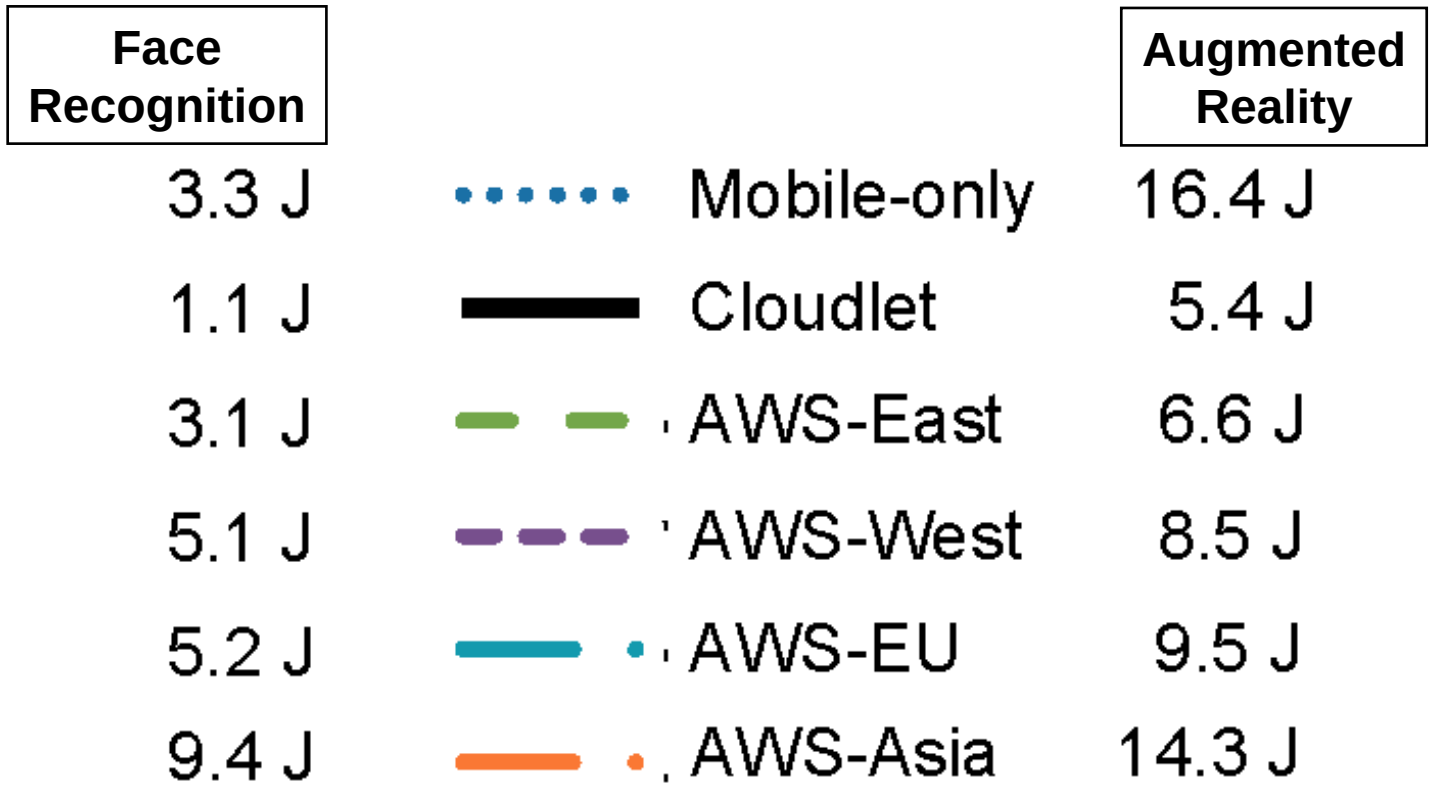
CDF of Response Times (milliseconds)

Amazon EC2 offload

See YouTube links at http://elijah.cs.cmu.edu/demo.html



# Per-Operation Energy Use by Device



# What is the Killer Use Case?

# Wearable Cognitive Assistance

***A new modality of computing***

**Entirely new genre of applications**

**Combine mobile and cloud with real-time cognitive engines**

**scene analysis, object/person recognition, speech recognition, language translation, planning, navigation, question-answering technology, voice synthesis, …**

***Seamlessly integrated into inner loop of human cognition***

**“Early Implementation Experience with Wearable Cognitive Assistance Applications”**

**Zhuo Chen, Lu Jiang, Wenlu Hu, Kiryong Ha, Brandon Amos, Padmanabhan Pillai, Alex Hauptmann, and Mahadev Satyanarayanan**

**Proceedings of WearSys 2015, Florence, Italy, May 2015**

# Baby Steps: 2D Lego Assembly

**Very first proof-of-concept (September 2014)**

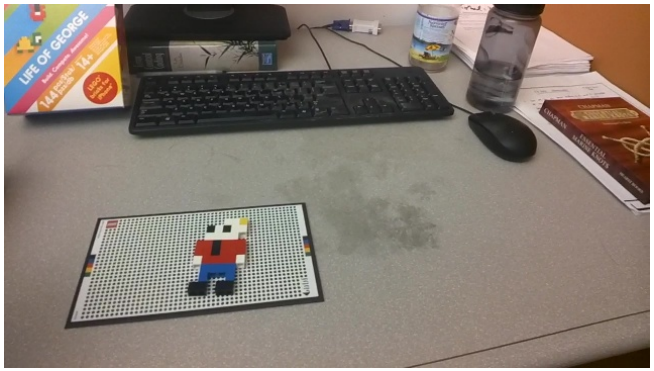
**Deliberately simplified task to keep computer vision tractable**

***2D Lego Assembly*** **(YouTube video at http://youtu.be/uy17Hz5xvmY)**

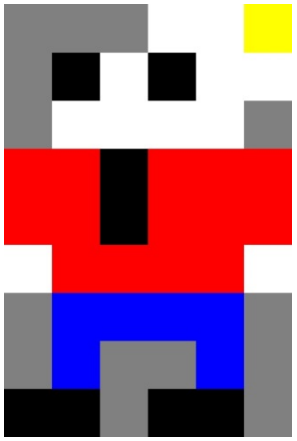
# How Does it Work?

*Two-phase processing: applies to all our applications*

**Video frame**



**Task-specific Symbolic Representation**



[[0, 0, 0, 1, 1, 3],
[0, 6, 1, 6, 1, 1],
[0, 1, 1, 1, 1, 0],
[4, 4, 6, 4, 4, 4],
[4, 4, 6, 4, 4, 4],
[1, 4, 4, 4, 4, 1],
[0, 5, 5, 5, 5, 0],
[0, 5, 0, 0, 5, 0],
[6, 6, 0, 6, 6, 0]]

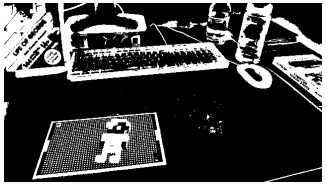
**Visual + Verbal Guidance**

- **"Analog to Digital Conversion"**
- **Tolerant of different types of lighting, background, occlusion, clutter, etc.**

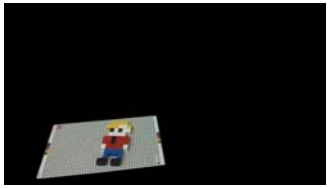
**Match current state to known state space from task description to derive guidance**

# Extracting Symbolic Representation

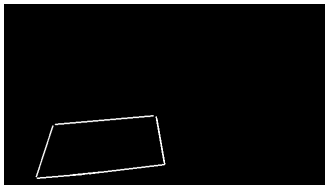
(a) Input image



(b) Detected dark parts

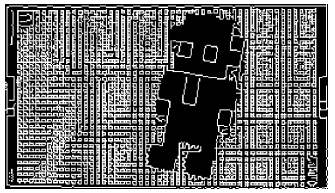


(c) Detected board



(d) Board border

(e) Perspective corrected



(f) Edges detected

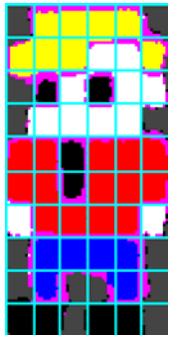
(g) Background subtracted

(h) Side parts added

(h) Lego detected

(i) Unrotated

(i) Color quantized



(j) Partitioned

```
[[0, 3, 3, 3, 3, 0],
 [3, 3, 3, 1, 1, 3],
 [0, 6, 1, 6, 1, 1],
 [0, 1, 1, 1, 1, 0],
 [4, 4, 6, 4, 4, 4],
 [4, 4, 6, 4, 4, 4],
 [1, 4, 4, 4, 4, 1],
 [0, 5, 5, 5, 5, 0],
 [0, 5, 0, 0, 5, 0],
 [6, 6, 0, 6, 6, 0]]
```

(j) Matrix

(k) Synthesized

# Example 2: Drawing Assistant

**"Drawing by observation"**

- **corrective feedback for construction lines**
- **original version uses pen tablet and screen**

**Software developed at INRIA**

**"The Drawing Assistant: automated drawing guidance and feedback from photographs"**

**Iarussi, E., Bousseau, A. and Tsandilas, T.**

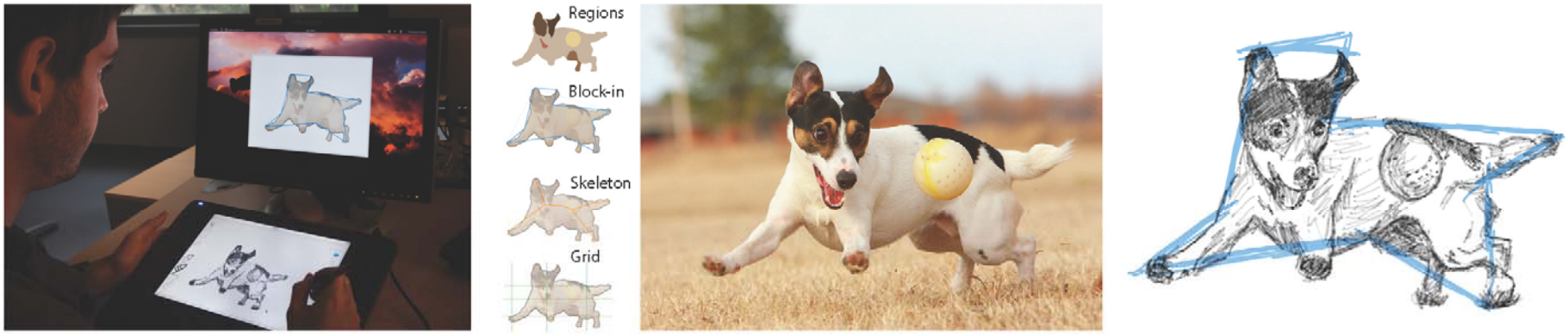
**In ACM Symposium on User Interface Software and Technology (UIST), 2013.**

# The Drawing Assistant: Automated Drawing Guidance and Feedback from Photographs

**Emmanuel Iarussi**
REVES / Inria Sophia Antipolis

**Adrien Bousseau**
REVES / Inria Sophia Antipolis

**Theophanis Tsandilas**
Inria - Univ Paris-Sud & CNRS (LRI)




(a) Interaction setup (b) Model and extracted guides (c) User construction lines and drawing

**Figure 1. Our drawing assistant provides guidance and feedback over a model photograph that the user reproduces on a virtual canvas (a). We use computer vision algorithms to extract visual guides that enhance the geometric structures in the image (b). In this example, the user first sketched the block-in construction lines (c, blue) before drawing the regions and adding details. This guidance helps users produce more accurate drawings.**

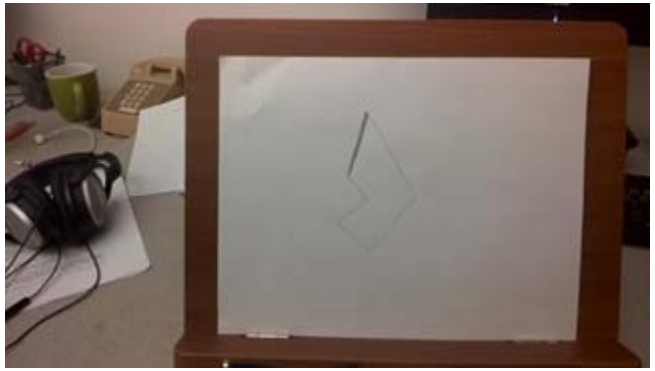
## Our goal

- **use Google Glass to untether this application**
- **allow drawing using any medium in the real world**
  **(paper, whiteboard, oil paint and brush on canvas, etc.)**

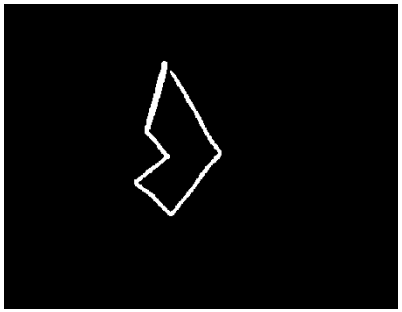
**Visual Feedback**

# Drawing Assistant Workflow

Video frame



Symbolic Representation



Visual Feedback

- find paper
- locate sketches
- remove noise

Feed to almost unmodified logic in original software

## Drawing assistant

(https://www.youtube.com/watch?v=nuQpPtVJC6o)

# Example 3: Ping-pong Assistant

***highly latency-sensitive application***

**Modest goal**

- **guide novice on where to hit the ball (left or right)**
- **not for professionals**
- **not for visually impaired people**
- **assumes decent hand-eye coordination**

# Ping-pong Assistant Workflow

**Video frame**
**(pairs of frames analyzed)**

**Symbolic Representation (3-tuple)**

**<is_playing, ball_pos, opponent_pos>**

**Verbal Feedback**

**“Left”/ “Right”**

- **Table detection**
- **Opponent detection**
- **Ball detection**

**Guidance based on recent state history**

# Ping-pong assistant

**(https://www.youtube.com/watch?v=_lp32sowyUA)**

# Many Use Cases …

Assembly instructions

Industrial troubleshooting

Medical training

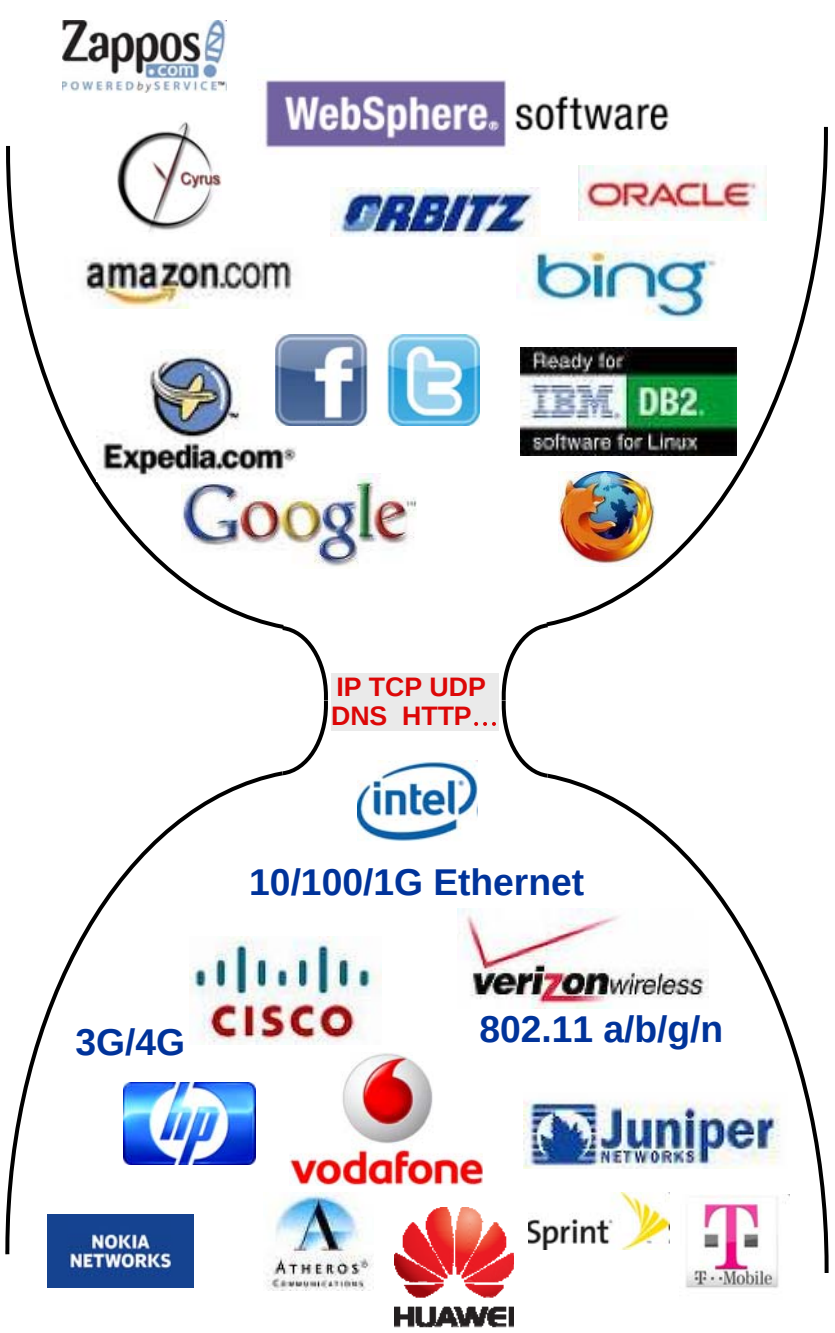
Correct Self-Instrumentation

Strengthening willpower

# How Can Edge Infrastructure Deliver Highest Value?

**~$4T in 2013**

**G-20 Internet Economy**




**~$1T in 2013**

**Total Market Cap**

**Open Internet Ecosystem**

# The Secret of the Internet

*How does something so simple create such high value?*

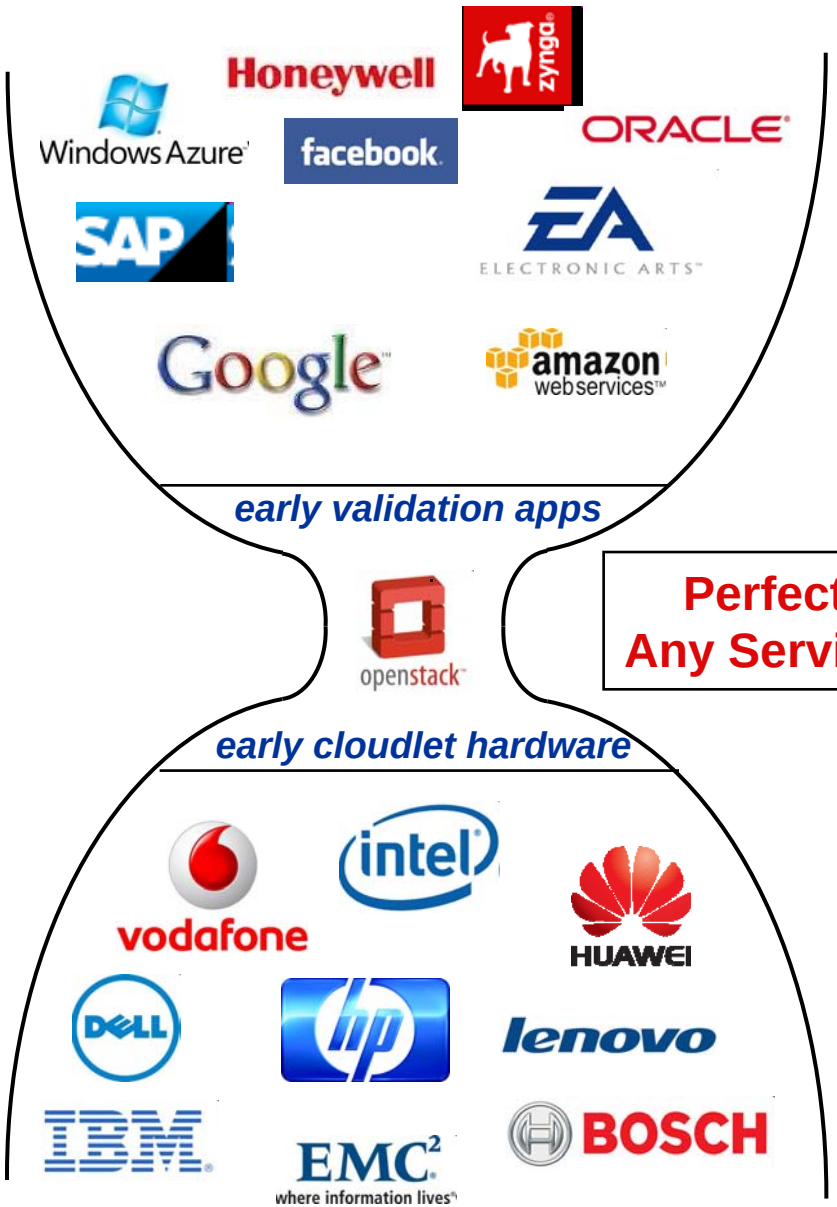
*As long as authentication, authorization and billing criteria are met, any E2E service that is running on any mobile device can leverage any available IP network.*

*The use of a particular IP network is based solely on*

- *pricing of the network*
- *performance criteria such as latency and bandwidth*

*rather than software compatibility.*

**cloudlet-based apps & services**




**Perfect Interoperability: Any Service on Any Cloudlet**

**cloudlet infrastructure**

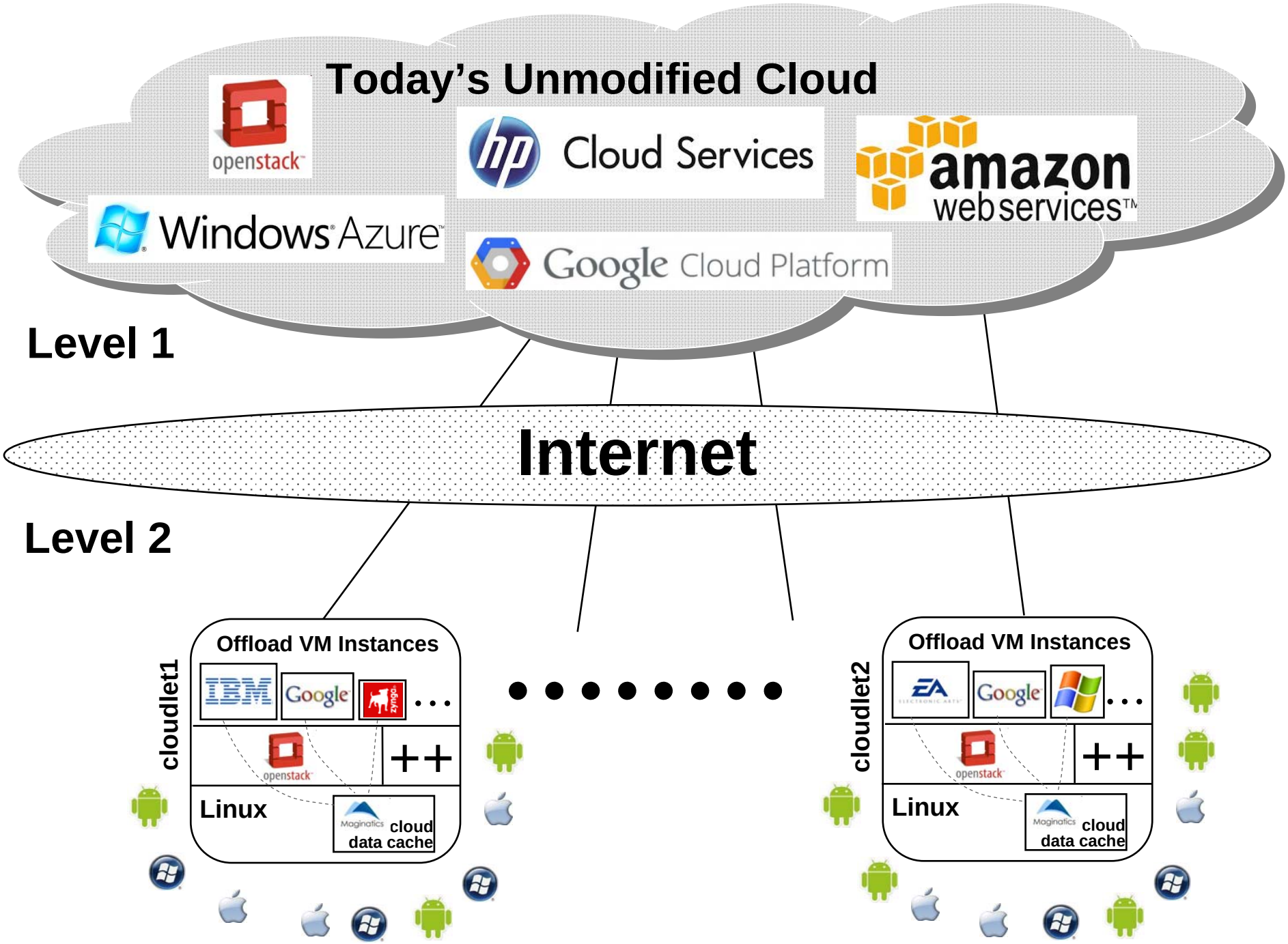
# Cloudlet Manifesto

*As long as authentication, authorization and billing criteria are met, any E2E service that is running on any mobile device should be able to leverage any cloudlet in the world.*

*The decision to use a particular cloudlet at a specific point in time should be based solely on*

- *current pricing of cloudlet resources and*
- *service-relevant performance criteria (such as E2E latency, multi-tenancy load, storage cache state, and availability of hardware accelerators)*

*rather than software compatibility.*



# OpenStack


**Open Stack ≈ EC2-like cloud services and REST API**

- **Apache v2 open source license**
- **widely used in industry**
  **(HP, Dell, IBM, Intel, Oracle, NetApp, CloudBase, CloudByte, CloudScaling, Piston Cloud, …)**
- **APIs for commonly used cloud services and management**
  **(identity, compute, image, object storage, networking, block storage, …)**

**We have created extended version for cloudlets**

***Goal: make OpenStack the platform of choice for edge computing***

# OpenStack Extensions

*(Kiryong Ha, CMU)*

**Bare minimum additional functionality**

1. **Cloudlet discovery**
2. **Rapid cloudlet provisioning (dynamic VM synthesis)**
3. **Adaptive VM handoff across cloudlets**

**No changes to OpenStack, only additions**

**Minimalistic approach: all other functionality inherited from Openstack**

# Status

**Extensions integrated into “Kilo” release of OpenStack (April 30, 2015)**

**Invitation to present at Tokyo OpenStack Summit (Oct 2015)**

**Ideal: acceptance into May 2016 release of OpenStack**