ECE 590
COMPSI 590

Special Topics: Edge Computing

Wednesday January 8th, 2020

1

Duke UNIVERSITY

2

Duke UNIVERSITY

About the Instructor (1/2)

- Started at Duke in Fall 2018
- Previously: Associate Research Scholar, Princeton University, Electrical Engineering
- Ph.D. Columbia University, Electrical Engineering
- M.Sc., B.Sc. University of Ottawa, Canada

Duke UNIVERSITY

About the Instructor (2/2)

- Worked in industry before, during, and after all degrees

IBM

Disney Research SCIENCE & PLAY

D E Shaw Research

Canadian Intellectual Property Office

Telcordia Technologies

DEFENCE R&D DÉFENSE

SiriusXM SATELLITE RADIO

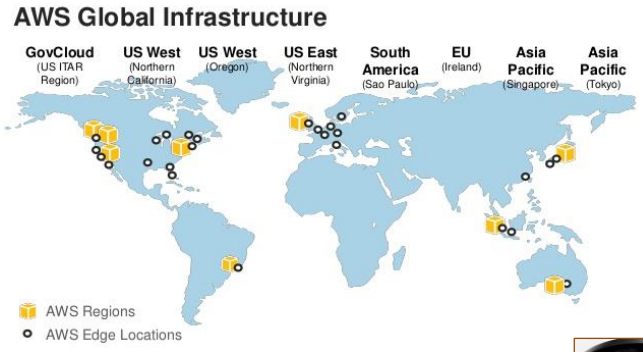
Duke UNIVERSITY

Introductions

In this Lecture

- Introduction to edge computing, part 1
- Course structure and syllabus
- Introduction to edge computing, part 2

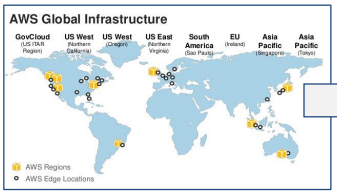
Cloud: Computing in Datacenters



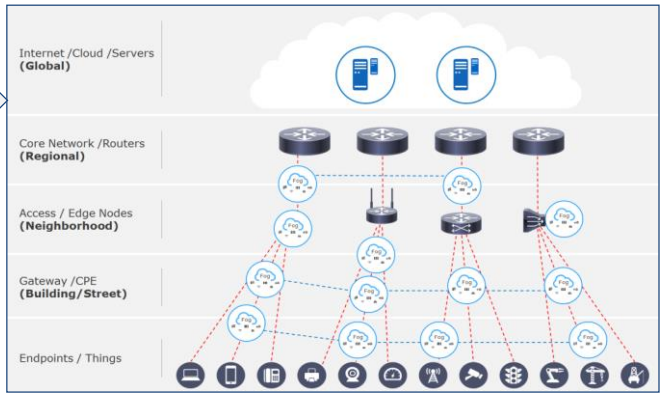
- AWS: 44 locations worldwide, MS Azure: 30
- For emerging applications: **fundamental limitations** in **latency**, **bandwidth**

7

Edge/Fog: Computing Closer to the Users



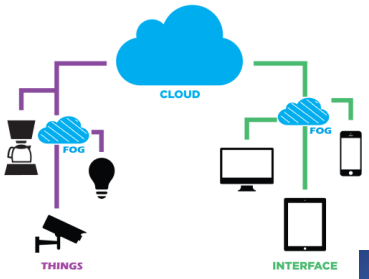
- Data processing, business logic, decision-making at multiple points in the hierarchy



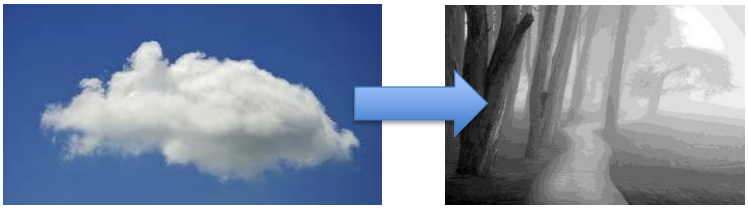
Smart city IoT deployments: computing in buildings, neighborhoods, zip codes

8

Edge Computing: Cloud to Fog



- Distribute computation, storage, communication and control services immersively closer to end-users **along the Cloud-to-Things (C2T) continuum**

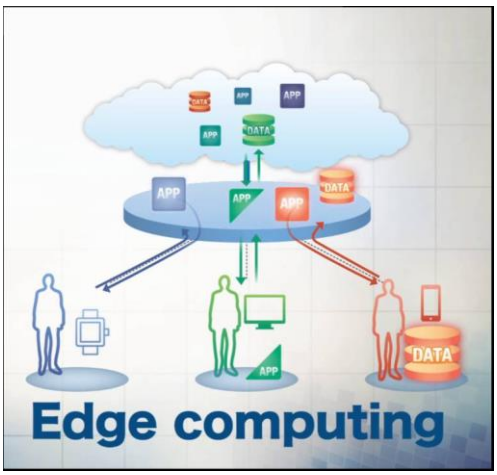


9

Duke UNIVERSITY

Edge Computing: Core to Edge

- An architecture that uses **one** or a **collaborative multitude** of **end-user clients** or **near-user edge devices** to carry out a **substantial amount** of computation, storage, communication, and control
- **Core → Edge**



10

Duke UNIVERSITY

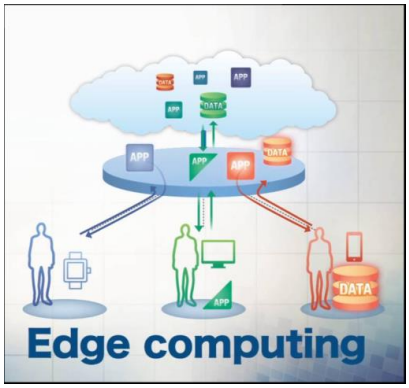
End-User Clients or Near-User Edge Devices: A Range of Options

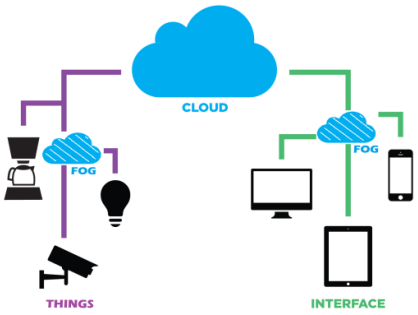
- Gateways, stationary or mobile
- Set-top boxes
- Servers, cloudlets
- Mini-datacenters

- Different properties

11

Substantial Amount of Work Near the Users: Range of (De)centralization Alternatives





12

What Does Edge Provide?

- Latency, bandwidth
- Cognition – advanced intelligence close to the users
- Privacy

- **Improve the performance of existing applications** and **enable new ones**

13

Why Edge Computing? (1/2)

- "Next billion-dollar tech market" ANDREESSEN HOROWITZ
- "Most interesting part of cloud computing"
- Fundamental enabler of the pervasive computing vision

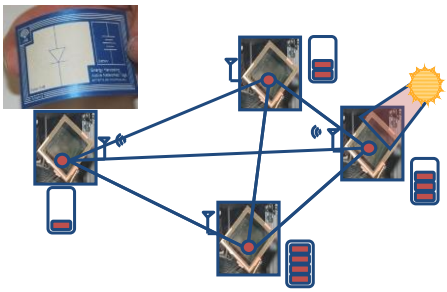
14

Why Edge Computing? (2/2)

- Advanced intelligence close to users is cool
- Emerging field
- On the boundary of domains not usually examined together
- Requires re-thinking implicit assumptions

15

My Life On the Edge: Background



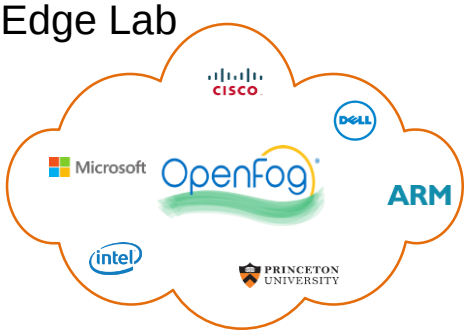
D E Shaw Research

- See edge as an enabler for the true potential of the Internet of Things
- Unusual IoT-to-cloud perspective

16

My Life on the Edge: 2016-2018

- OpenFog Consortuim technical committee member, working group co-chair, board of directors alternate
- Contributor to the IEEE 1934 Fog Computing Standard
- Associate director of Princeton Edge Lab



17

My Life on the Edge: 2019-2020

- OpenFog Consortuim technical committee member, working group co-chair, board of directors alternate
- Contributor to the IEEE 1934 Fog Computing Standard
- Associate director of Princeton Edge Lab

18

In this Lecture

- Introduction to edge computing, part 1
- Course structure and syllabus
- Introduction to edge computing, part 2

19

Duke UNIVERSITY

Course Logistics

- Lecture times: **01:25 -02:40 PM**, Mondays and Wednesdays
- Professor office hours: Mondays 03:00 – 04:00 PM, Wednesdays 11:00 – 12:00 AM, 2471 CIEMAS
- TA office hours: TBD
- Readings before every class

20

Duke UNIVERSITY

Course Structure

1. Background, core architectural principles, challenges and opportunities
   - Goal: enough “lay of the land” for the research projects
2. Domains where edge is particularly exciting
3. Notable papers in the field and techniques that need to be re-imagined for the edge

21

Upcoming Lectures

- **1/13:** The origins and the current state of edge computing
- **1/15:** Edge helping the IoT
- **1/22:** Edge helping higher-end mobile devices
- **1/27:** Edge helping the cloud

22

Grading

- Quizzes: 20%
- Research paper presentation: 20%
- Research project: 50%
- Participation in class discussions: 10%

23

Duke UNIVERSITY

Please See the Syllabus for …

- Quizzes
- Research paper presentation
  - Spread over weeks of **September 12th – October 15th**
  - Sign-up Google Sheets available
- Participation in class discussions

24

Duke UNIVERSITY

Research Project in Edge Computing

- 50% of the grade
- Teams of 1-3 people
- Research project
  - Generate and thoroughly validate a new idea
- Best-case outcome: work leading to a paper that can be published
  - But, its research – not all explorations are fruitful
  - High-risk high-rewards > incremental improvement

25

Duke UNIVERSITY

Project Grading

- **50% total**
  - 10% project proposal
  - 10% progress report
  - 15% final presentation
  - 25% final report
- For late reports, 10% penalty per day, up to 5 days.

26

Duke UNIVERSITY

Research Project: Timelines

- Teams established: Friday **January 24th**
- Proposal (10%) due: Monday **February 10th**
  - Will talk about the format next class
- Progress report (10%) due: **Friday March 20th**
- Final presentations (15%): **weeks of March 29th, April 5th, and April 12th**
- Final report (25%) due: **Friday April 17th**

27

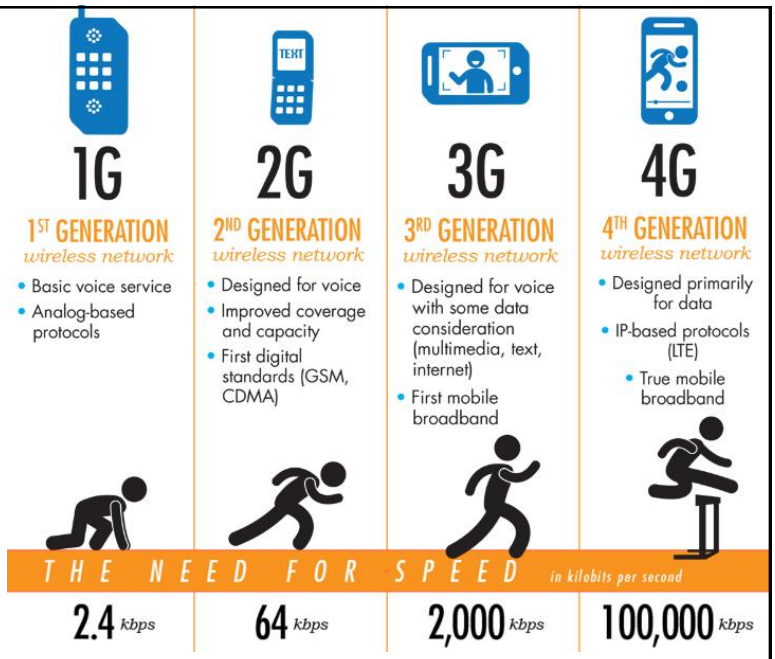
In this Lecture

- Introduction to edge computing, part 1
- Course structure and syllabus
- Introduction to edge computing, part 2
  - Edge and 5G
  - Research themes in edge computing

28

Generations of Cellular Technology

- Define both last-hop and wired backbone behavior



Duke UNIVERSITY

5G Requirements

- 1 ms end-to-end round-trip latency
  - Compared to ~ 10 ms with 4G
- 1-10 Gbps connections to mobile hosts
  - 4G Verizon: 12 Mbps downlink, 5 Mbps uplink
  - Peak downlink ~ 50 Mbps: 20 – 200x slower than 4G
- 1000x bandwidth per unit area

30

Duke UNIVERSITY

5G is Coming

- Several phone models already support 5G
  - Samsung Galaxy S10 5G
  - LG V50 THINQ
- Limited deployments ongoing world-wide
  - Limited regionally
  - Limited in capabilities
- Much more expected over the next couple of years
  - E.g., Germany expects 98% of households to have 5G access by 2022

31

Duke UNIVERSITY

Current 5G Deployments in the US

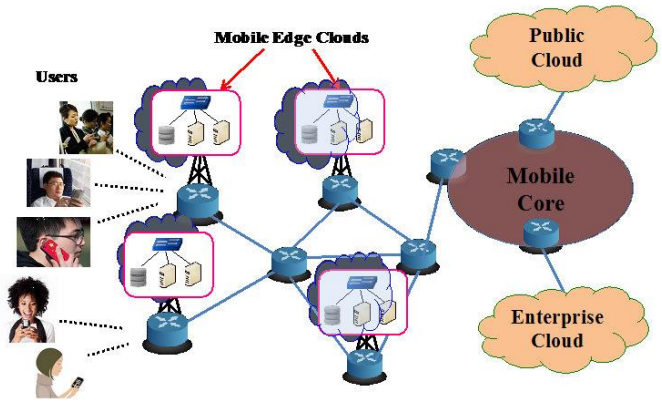
City \ Carrier	AT&T	Sprint	T-Mobile	Verizon
Atlanta	Live	Live	Live	Live
Boston	—	—	—	Planned
Charlotte	Live	—	—	Planned
Chicago	—	Live	—	Live
Cincinnati	—	—	—	Planned
Cleveland	—	—	—	Planned
Columbus	—	—	—	Planned
Dallas–Fort Worth	Live	Live	Planned	Planned
Denver	—	—	—	Live
Des Moines	—	—	—	Planned
Detroit	—	—	—	Live
Houston	Live	Live	—	Planned
Indianapolis	Live	—	—	Live
Jacksonville	Live	—	—	—
Kansas City	—	Live	—	Planned
Las Vegas	Live	—	Live	—
Little Rock	—	—	—	Planned
Los Angeles	Live	Live	Live	—
Louisville	Live	—	—	—
Memphis	—	—	—	Planned
Minneapolis–Saint Paul	—	—	—	Live
Nashville	Live	—	—	—
New Orleans	Live	—	—	—
New York	—	Live	Live	—
Oklahoma City	Live	—	—	—
Orlando	Live	—	—	—
Phoenix	—	Live	—	Live
Providence	—	—	—	Live
Raleigh	Live	—	—	—
Salt Lake City	—	—	—	Planned
San Antonio	Live	—	—	—
San Diego	Live	—	—	Planned
San Francisco	Live	—	—	—
San Jose	Live	—	—	—
Tampa	Live	—	—	—
Waco	Live	—	—	—
Washington	—	Live	—	Live

32

Duke UNIVERSITY

Edge Computing is a Part of 5G

- Edge computing
  - Computing capabilities attached to each base station
- Offers:
  - Lower latency
  - Reduced load on core network



33

Duke UNIVERSITY

Research Themes in Edge Computing (1/5)

- What should be placed where?
  - Computing, storage, decision-making
- Restructuring applications and algorithms to fit edge/fog conditions

34

Duke UNIVERSITY

Research Themes in Edge Computing (2/5)

- Edge in specific applications
  - Augmented reality, virtual reality
  - Networks of drones
  - Autonomous driving
  - Invited speaker: edge computing for autonomous trucking

35

Research Themes in Edge Computing (3/5)

- Data processing that preserves privacy
- ML training on the edge
- ML inference on the edge
- Reinforcement learning on the edge

36

Research Themes in Edge Computing (4/5)

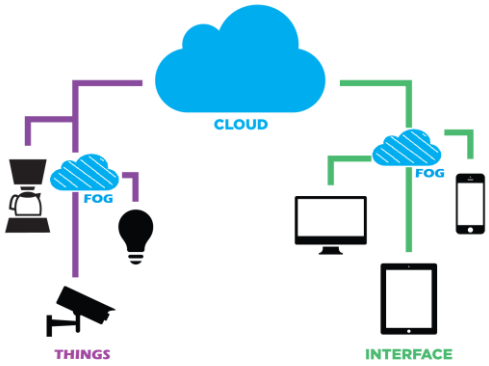
- Thinking across applications, devices, platforms
- Operating across multiple computing / storage / control / decision quality levels

Research Themes in Edge Computing (5/5)

- Multi-tenancy
- Multi-device operation
- *Uberization* of resources
- …

Lecture Recap

- Edge computing: definitions
  - ➢Different devices
  - ➢Different degrees of application centralization
- Logistics of the class
- Research themes



39

Duke UNIVERSITY

Next Class and Homework (1/2)

- *"IoT meets the cloud: the origins and the current state of edge computing"*
- Reading materials for the class

40

Duke UNIVERSITY

Next Class and Homework (2/2)

- Choose the dates for your presentations
- Start thinking about your research project
  - My office hours: 3-4 PM Mondays, 11-noon Wednesdays
  - Feel free to reach out via e-mail as well

41

Duke UNIVERSITY