Internet Data Streams

Lecture 1: Introduction

Gregory S. DeLozier, Ph.D
June 13, 2017

Introductions...

Gregory S. DeLozier, Ph.D.

gdelozie@kent.edu

GitHub: gregdelozier

Twitter: gregdelozier

I am not this guy…

- http://amarillo.com/stories/1999/01/29/usn_LA0602.002.shtml#.VXTpHVnBzGc

I am...

- A part-time faculty member and researcher
- A full-time research computer scientist
  - software engineering metrics and methodologies
  - testing and reliability methodology
  - software tools and development systems
- A *working* software developer
  - Medical and scientific software, software tools
- Husband to one, father of five, sailor, cyclist...
  - ...computer geek, photographer, etc.

I am…

Change photo

Gregory S DeLozier

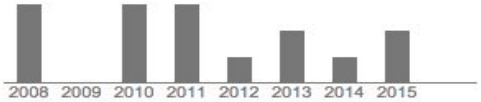
Edit | Follow

Adjunct Faculty, Computer Science, Kent State University
Computer Science, Software Engineering, Data Science
Verified email at kent.edu - Homepage
My profile is public

Title	Add	More	1–15	Cited by	Year
A method for measurement of integrated foot kinematics G DeLozier, I Alexander, R Narayanaswamy International Symposium on Three-Dimensional Analysis of Human Movement ...				17	1991
Biochemical and X-ray diffraction analysis of concanavalin B crystals from jack bean R Morrison, G Delozier, L Robinson, A McPherson Plant physiology 76 (1), 175-183				14	1984
Using personal digital assistants to collect survey data SM Nusser, DM Thompson, GS DeLozier 1996 Proceedings of the Section on Survey Research Methods, American ...				11	1996
A computer graphics program for the three-dimensional reconstruction of plant organs from serial sections G Delozier, K Eckard, M Greene, EM Lord American journal of botany, 136-140				8	1987

Google Scholar

Citation indices	All	Since 2010
Citations	62	12
h-index	4	2
i10-index	3	0



Add co-authors

Mark D. Grabiner + ×
Herbert L.Axelrod + ×

Co-authors Edit…

No co-authors

I am not…

Suspect says tea sediment led to stabbing

The Associated Press (1999)

TRENTON, N.J. (AP) - "The iced tea made me do it."

That's the defense offered by lobbyist Gregory DeLozier, 35, who was charged with the attempted murder of his wife, Michelle. She was stabbed by a masked intruder Tuesday at a salon where she was having her hair done.

During a videotaped statement he gave police after his arrest, DeLozier said sediment in a bottle of iced tea he drank had weird side effects. That, coupled with the stress of an upcoming mortgage deal the couple was pursuing, made him snap, he said.

Course Information

Course Description

This course examines IOT data streams.

- The devices that make up the data collection and control presence of the IoT.
- The data stream technology required to collect, analyze, and respond to the data generated by millions of data collecting devices in the environment.

Class organization

- Classes are on Tuesday and Thursday, 7:00 - 9:30
- All classes will have some lecture and some lab content
  - Bring a laptop at a minimum
- Attendance is mandatory without prior arrangements.
- Office hours before or after class by appointment.
- I will often stay after class for a while to work on things

Class Outline (Approximately)

- Week 1: IoT concepts, devices, and tools.
- Week 2: Data collection methods.
- Week 3: Networking – devices.
- Week 4: Networking – streams.
- Week 5: Dashboard management and presentation.
- Week 6: Data analysis and decision support.
- Week 7: Economic, legal, and social issues.
- Week 8: Project presentations and final exam.

  (Final will be in class during the last week of class.)

Class grading

- Homework project and experiments (30%)
- Create a significant and practical data streams project, and present a paper and presentation to the class on their work and its value. (40%)
- Students will demonstrate knowledge of the area in a final examination (30%)
- Extra credit (10%)

Class grading

- Not doing your own work -100%

Plagiarism and similar things

Read this:

http://www.kent.edu/plagiarism

Read this from the honor pledge:

I am aware that sanctions may be imposed at the discretion of my instructors and/or through the university's Judicial Affairs system for a violation of principles outlined in this statement, including:

Refusal to accept the work.

Failure of a specific assignment.

Failure in the course.

Explusion from the university.

...and similar things...

Using content from the web, article, or any work other than your own, and

1. Presenting it as your own effort
2. Presenting it as the answer to a test problem or assignment, even if cited
3. Using it as the basis for a "thin copy"
4. Generally not doing your own work

...and this includes...

- Turning in wikipedia articles as test answers.
- Reading IEEE journal articles and turning in paragraphs from them in your paper.
- Turning in a final exam consisting almost exclusively of cited web page excerpts.
- Taking a tutorial project off the web and turning in as your class project.

...so please...

- Do your own work.
- Don't make me into the test police.
- Trust that my expectations are reasonable.
- Ask for help.
- And...

- Again, *do your own work*.

Ok, enough about that.

Official Class Content

- *Official* Content is on *Blackboard*.
  - Class Roster
  - Grades
  - Class emails

Social Class Content

- *Social* Content is on *Twitter*.
  - Question-and-Answer
    - Instructors
    - Other students
  - Discussions between students
  - Use *#kentdata*
  - I'm *gregdelozier*
  - I rarely post anything interesting.

Instructional Class Content

- *Instructional* Content is on *GitHub*.
  - Repository:
    - *https://github.com/gregdelozier/*data_streams_class
    - Project files
    - Class demo files
  - Wiki:
    - *https://github.com/gregdelozier/data_streams_class/wiki*
    - How-to and tutorial pages
    - Instructions for labs or projects
    - Links to course notes
  - Demo time...

Accounts you will need

- Accounts
  - GitHub - free or inexpensive (demo)
  - PythonAnywhere - free or inexpensive (demo)
  - Google (i.e. gmail)
  - Twitter
  - Cloud providers as needed by labs

Things you will need

- Skills
  - Python, Git, Bash competency
  - Basic understanding of web technology
- Stuff
  - Laptop Computer (Mac, Windows, or maybe Linux)
    - Need to run Python (Anaconda Python 3.6)
    - Ethernet and USB ports
    - Administrative privileges
  - Raspberry Pi B 3 & Arduino gear as needed

Course Objectives

This is a fairly new course.

We may modify or adapt as we go.

The course is about *you doing things.*

These following objectives are approximate...

Objective 1/4 - Data Collection

The course will cover the creation of devices to participate in the IoT, and the tools and technologies needed to do so using industry-standard building blocks.

- Microcontrollers (like the Arduino)
- Single board SOC computers (Like the RPi)
- Sensors and interfacing

Objective 2/4 - IoT Networking

Next, the course will cover the necessary technology to connect IoT devices to the internet, making it possible to communicate with data aggregation systems.

- Connecting to the network
- Data protocols
- Data streaming and buffering
- JSON & REST

Objective 3/4 - Data Streams

Additionally, the course will cover the technologies necessary to aggregate and process the millions of data streams created by such devices, including:

- Web server and services
- Cloud services
- Data analysis and presentation
- Statistical methods

Objective 4/4 - Ethics and Economics

Finally, we will look at the economic and social significance of these capabilities and their effect on information value, privacy, scientific progress, and our ability to use these data to drive decisions.

- Economic valuation models
- Cost and ROI for IoT scenarious
- Privacy and ethics considerations

Questions?

Break time!

The Internet of Things

The Internet of Things (IoT) is the network of physical objects or "things" embedded[1] with electronics, software, sensors[2] and connectivity to enable it to achieve greater value and service by exchanging data with the manufacturer, operator and/or other connected devices. Each thing is uniquely identifiable through its embedded computing system but is able to interoperate within the existing Internet infrastructure.

-- Wikipedia

The Internet of Things

Kevin Ashton

- MIT
- Helped create RFID as an Open Standard
  - ...where everything has an ID
- Started "Internet of Things" concept
- (He's also interesting in other ways...)

If everything has an address...

- *Things* can have and report status
- Status is
  - immediate
  - useful
  - associated with a thing
  - associated with the thing's attributes
- *Things* can accept individual input
  - Control themselves
  - Control other things in the environment

What kind of information?

- Internal state
  - Temperature,Pressure,Speed
  - Location
- Environmental state
  - Weather
  - Traffic
  - Counts
- Financial, economic, network data
- Anything...

Even more things to measure...

- Traffic patterns
- Shopping habits
- Medical readings
- Wearables - health, mental state, locations
- Inventory
- Energy use
- Water quality
- Wildlife habits
- Airline locations

Feedback and control

- Devices can receive commands
  - Control themselves
  - Control directly connected things
  - Send control messages to other things
- Direct feedback and control
  - Microcontrollers, no networked control
- Networked feedback (or not) and control
  - Networking capability is needed (e.g. SOC)
  - Control decisions are made elsewhere

Things to control…

- Thermostats
- Speeds
- Alarms
- Positions
- Dosages
- Feeding
- Trades and economic activity
- Inventory orders
- ...and so forth...

Measurement devices

- Analog devices measure the real world
  - Temperature, light, speed, etc.
  - Usually a chemical, electrical, or magnetic sensor
- Analog to digital converters (ADCs)
  - Convert analog voltage to digital data
  - Which can be handled by a computer
  - Which can be sent via a digital network

Control Devices

- Mechanical analog devices control the real world
  - Driven by voltage levels
    - Presence of voltage
    - Intensity and frequency of voltage
  - Frequently electomagnetic or peizoelectric
- Digital to analog converter (DACs)
  - DACs create analog voltages from digital signals
  - Digital signals can be computed and sent
  - Analog voltages drive analog devices.

Computational Devices

- Requirements
  - small and inexpensive
  - simple to program
  - durable, physically and evironmentally
  - low power draw
  - easy to incorporate into products
- Devices
  - These are basically small computers
  - History and details later

How much of this is there?

- More things now than people on earth
- 20-30 billion things by 2020
- But these are "interesting" things.
- What about _everything_?

The Internet of Objects

- RFID is extremely inexpensive
  - printable circuit costs pennies.
  - doesn't require inherent power.
  - not very smart. Just transmits a number.
  - what about a 'smart' RFID?
- Internet communication for _everything_
  - *Trillions* of things -- soup cans, toothpaste tubes
  - Can you imagine the data?

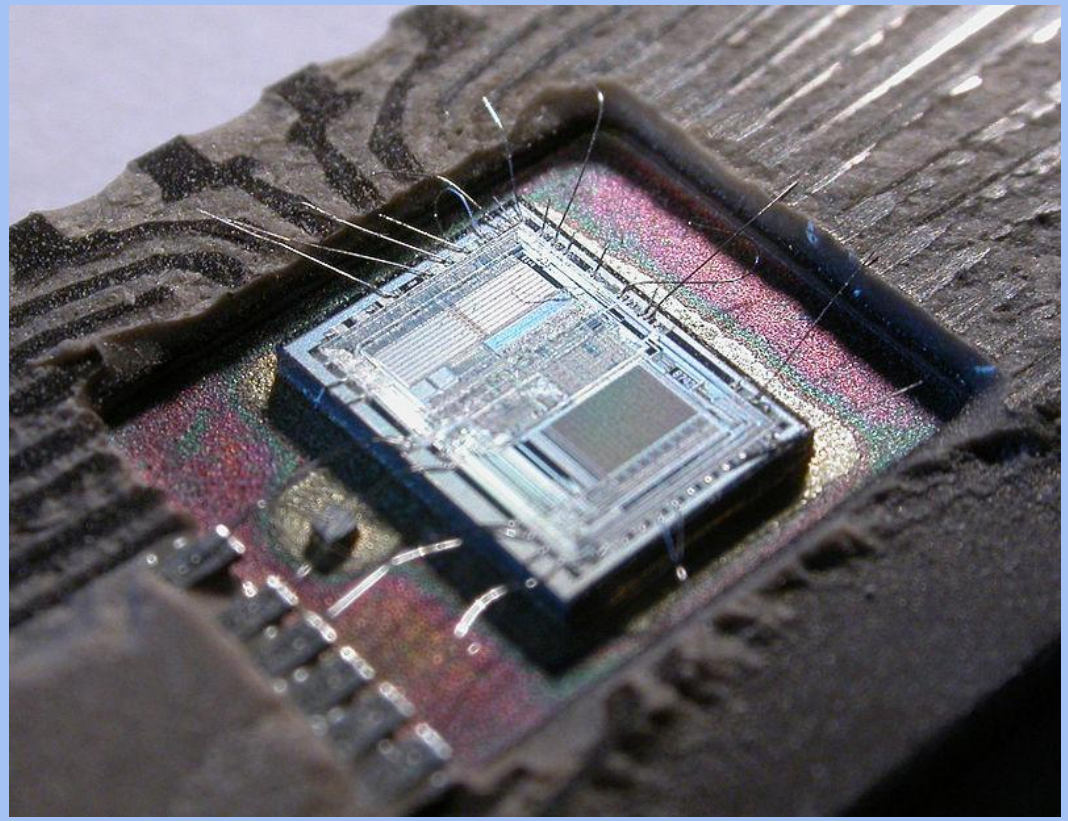
Back to Interesting Things

- Is the data or control worth $20?
- If so, you can afford a small computer
  - Processor
  - Memory
  - Storage
  - Input/output
- Early *microcomputers*
  - used external chips for memory, I/O, storage
  - expensive

Microcontrollers - system on a chip

- Intel 8742
- 12 MHz
- 128 Bytes (yes, bytes) RAM
  - Variables go here
- 2048 Bytes of EPROM
  - Put programs there
- Digital I/O ports

Put assembly program into EEPROM
(electrically erasable ROM)
Power on - starts running
Power off - stops running



Microcontroller programs

If thermometer says too cold

    Turn on heater

If thermometer says too hot

    Turn on air conditioner

Otherwise

    Turn them both off

Microcontroller characteristics

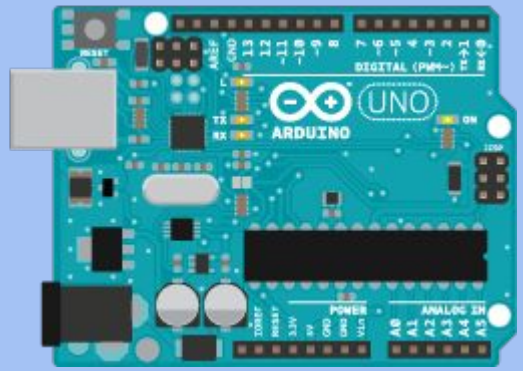
- Local data collection
- Local decisions
- Some LAN communication
  - E.g. cars, fly-by-wire, environmental control
- Little WAN communication
  - Networking is *hard*.
- Hard to program complexity
- Cheap and durable, cheaper than wires

Microcontroller usage

- Tiny computers in everything
- Cheaper than conventional circuitry
  - E.g. cheaper in TVs than analog tuning
  - Sensors and servos cheaper than long wires
- Smaller than other options
  - Ever seen an original iPad nano? A Tamagochi?
- Great for operating devices
- Not so great for Internet networking
  - Often networking is harder than data collection

Current Generation Microcontrollers

- Standardized circuits
- Easier to program
- Development kits available
- We will look at the Arduino
  - Very common for hobby use
  - Industrial versions for real world
  - http://www.arduino.cc/
  - Subject of this week's lab (on a simulator)



Microcontroller advantages

- Versatile
- Common
- Cheap
- Durable
- Low power consumption
- Easy to learn

Microcontroller limitations

- Harder to network
  - (Networking modules are available)
- Require external IDEs
- Nonstandard tools and languages (maybe)
- Relatively slow and not very smart
- Not much local processing power

We want an *Internet* of things

- Microcontrollers are toy computers
- Cell phones need *real* computers
- Cell phones need to communicate
- So what powers cell phones?

System on a Chip (SOC)

A system on a chip or system on chip (SoC or SOC) is an integrated circuit (IC) that integrates all components of a computer or other electronic system into a single chip. It may contain digital, analog, mixed-signal, and often radio-frequency functions—all on a single chip substrate.

-- Wikipedia

System on a Chip (SOC)

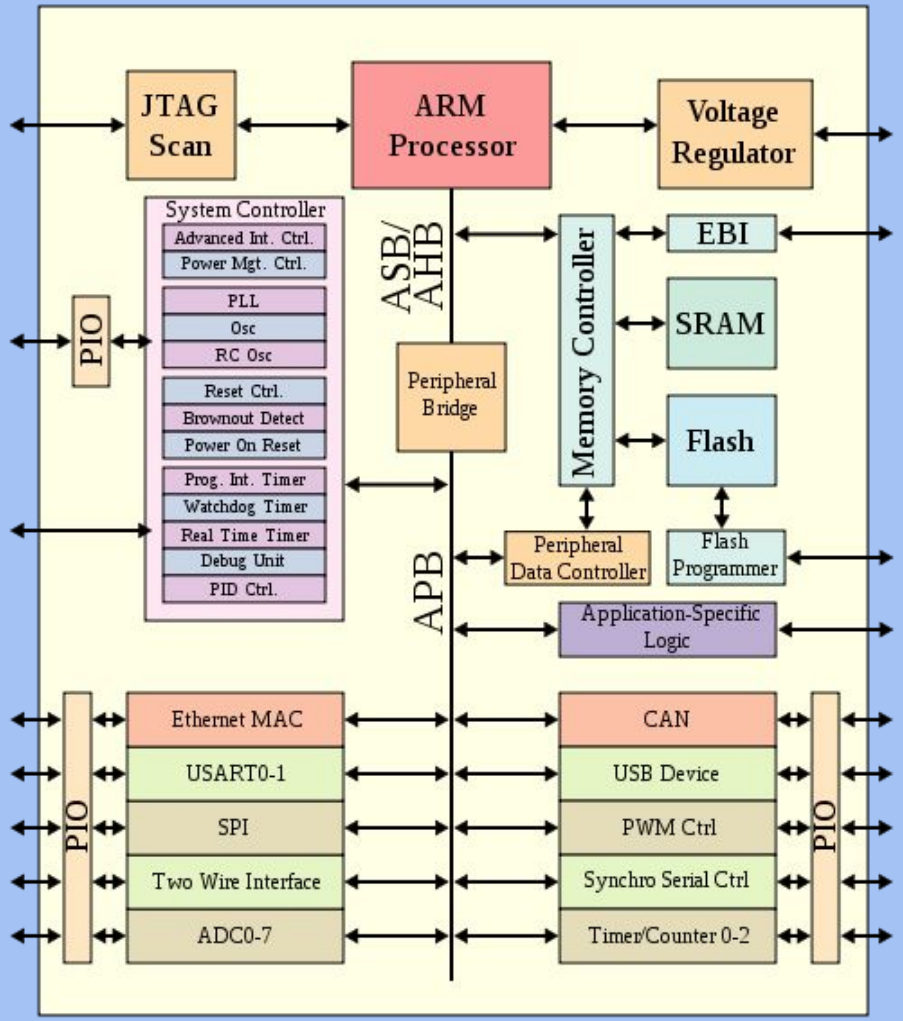
- System on a Chip (SOC)
  - Significant amounts of RAM
  - Local storage as flash-ram
  - Ability to run conventional software
    - Linux
    - Windows
    - Well-known software development stacks
  - Reasonably low power consumption
  - Easy to incorporate into designs

ARM SOC

ARM devices

- RISC
- Low power
- Open design
- Fairly fast

Seems useful...



SOCs become popular

Check this out:

http://en.wikipedia.org/wiki/List_of_system-on-a-chip_suppliers

- lots of options
- sold in manufacturing quantities
- a few hobbyist SOCs designs emerge
-

A Teaching Problem...

- Students in England getting worse at CS
- Why?
  - Home computers used to be toys
  - Modern home computers are _*not*_ toys.
  - Kids can't tear apart Mom and Dad's MacBook Pro
- Solution
  - Raspberry Pi Foundation formed
  - Design and build a cheap, educational SOC board
  - Make it very flexible and easy to use

The Raspberry Pi

1 GHz

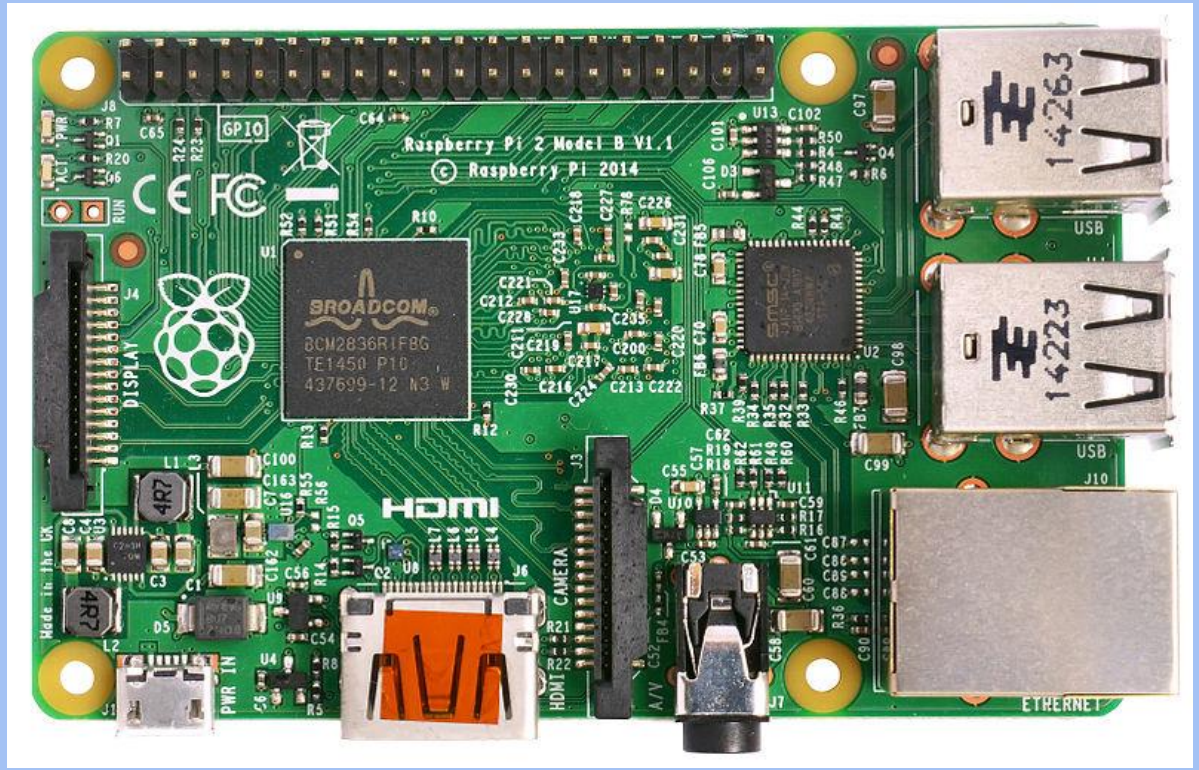
1 Gb

Linux

Ethernet

USB

GPIO



Lab Preparation

- Bring your laptop running Chrome & Python
  - Make sure WiFi works for you
- Read a little about the Arduino
  - http://www.arduino.cc/
  - http://en.wikipedia.org/wiki/Arduino
- We will be programming a simulated Arduino

Questions?

See you next time!