Internet of Things - Part Deux

Smart Devices
Smart Sensors
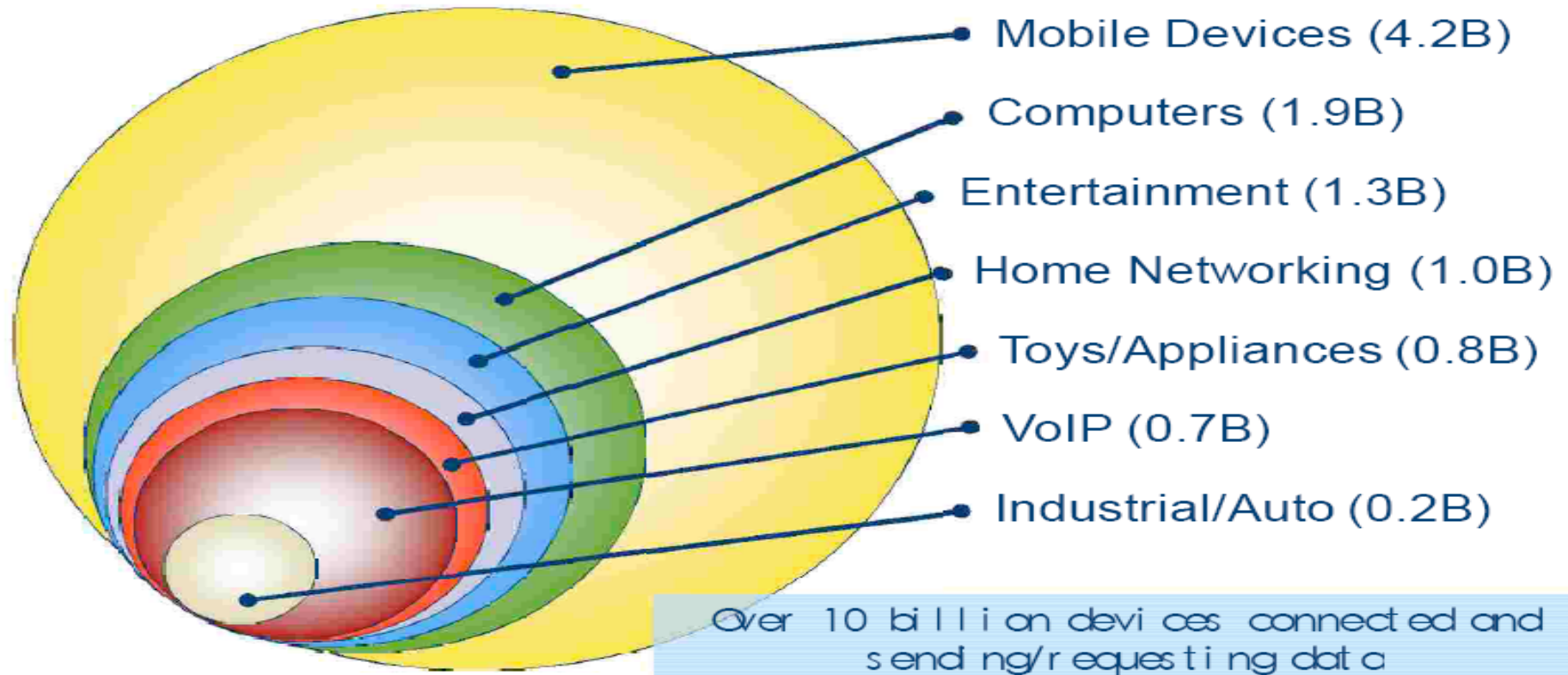# Big Gobs of Data

The rise of Embedded Systems

Wayne Duquaine
Grandview Systems
grandvu@sonic.net

Internet of Things: We're Now There !

10 Billion machines vs 6.9 Billion Humans

Devices Communicating at any given time (2012) – excluding enterprise datacenters

IDC — Analyze the Future

- Mobile Devices (4.2B)
- Computers (1.9B)
- Entertainment (1.3B)
- Home Networking (1.0B)
- Toys/Appliances (0.8B)
- VoIP (0.7B)
- Industrial/Auto (0.2B)

Over 10 billion devices connected and sending/requesting data

© IDC   Feb-11

And over 7 Billion Sensors are connected to those machines

Many New Apps are Sensor Based Apps - we just don't recognize them as such

- Geo-Location Apps
  - Google Maps and Turn-by-Turn Voice Navigation
  - Lookup Local Hotels/Restaurants via Smartphone

- New Financial Apps
  - Google Wallet   (uses NFC sensors on Smartphone)

- Distributed Data Collection Services
  - Flume   (Hadoop)
  - openPDC   (National Grid Monitoring)
  - Lumberjack   (Time-Series Data)
  - Pachube   (Hookup Sensors over the Internet)

Major Drivers for the Internet of Things

- ## Smart Phone / Tablets + Sensors
  - Social Networks   (Geo-location data)
  - Tethered Sensors
  - Electronic Wallets

- ## Green Technology

- ## eHealth / Medical

- ## Industrial Applications and System "Health Monitoring"

Embedded Systems and Smart Phone/Tablet

- The bulk of the Internet of Things are <u>Embedded Systems</u>
  - Embedded System = Computer system designed to do a few dedicated functions, in real-time, using sensors for data I/O

- Technically, all Smart Phones and Tablets are Embedded Systems, not PCs

- Smart Phone / Tablet Characteristics
  - Runs some variant of Unix
  - Have an <u>embedded relational database</u> (MySQL, …)
  - Have <u>inboard comm capability</u> (3G/4G and/or WiFi to Internet)
  - Have 3 or more <u>built in sensors</u> (Accelerometer, GPS, Photodiode)
  - All of them have on-board DSPs (Digital Signal Processors)

Tethering Sensors

- ADK (Android) and RedCable (iPhone/iPad) make it easy to tether sensors to your Smart Phone/Tablet

- In many cases, sensors are tied from Smart Phone/Table to Arduino Platform
  - Open Source Hardware/Software
  - Cheap ($29) and Easy to program
  - Is Google's reference platform for ADK
  - Used by <u>Tens of Thousands</u> of DIYers

- Explosion of Applications using Arduino
  - Programming and hooking up <u>sensors</u> is <u>no longer a black art</u>
  - Deployed in <u>thousands of experimental applications</u> from Art projects to Biology Experiments to Environmental Monitoring
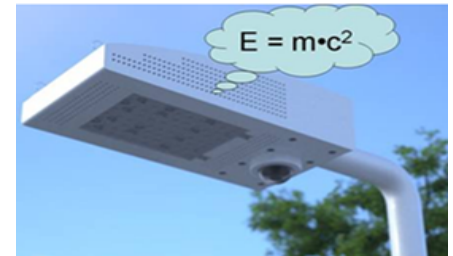
Green Technology

- As per HPTS 2009, the drive to improve efficiency and eliminate <u>wasting 50-90% of our electricity</u> continues:
    - Smart Motors
    - Smart Power Supplies
    - Solar Inverters/Micro-Inverters
    - Intelligent Lights (MCU driven high-brightness LEDs)

- All of these technologies critically depend upon DSPs and Calculus based algorithms

- DSPs are now down to $2 per pop, and every smart phone/tablet, smart motor, smart power supply, solar inverter, etc, … has one or more DSPs

eHealth / Medical

Home Health Tablet

- Baby boomers are going to bust Medicare
  - Costs continue spiraling upward
  - Conventional Doctor office visits are becoming WOMBATs
    - First 10 minutes of office visit are redundant - measure BP, HRT, weight
    - Only gives <u>a few data data points 2-3 times per year</u>
      => <u>not very useful trend analysis</u>
- Reducing costs will be crucial
  - Home Health Hubs
    - Connect to inexpensive ($ 25-75) health sensors (HRT, BP, …)
    - Record couple times a day, or continuously via wearable sensors
    - Data is saved in Smart Device's DBMS, and uploaded to Dr periodically
  - Telemedicine via real-time video will become standard practice

- MCU controlled Iontophoresis Bio-Patches will begin to replace replace Shots (reduces need for RN and going to Dr)

WOMBAT = Waste Of Money, Brains, And Time

System "Health Monitoring" - Better Sensor Data Can Prevent Failures

- Most Major Power Outages can be avoided
  - Major Outages usually take many minutes to hours to build up
  - But existing SCADA based data feeds provide:
    - Too little data: 2-3 data points
    - Too infrequently: Once every 10 seconds
  - Smart grid based Sensors can provide dozen+ data points at 10 samples per second
    - Power Gen operators can get a true real-time pulse of the network
    - Have Time to Throttle back/re-route the choke points
  - Key Tools needed:
    - Lots of real time data + Time-Series Trend Analysis
    - Calculus based computations for Instantaneous and Aggregate Power at Nodes   +   Data Visualization tools to show the hot spots

- Ditto goes for other HM apps: bridges, airplanes, …
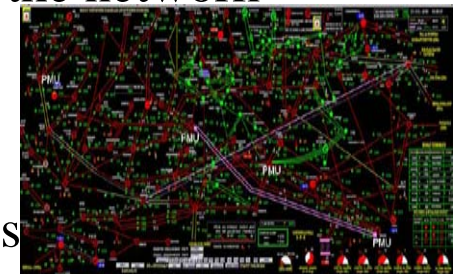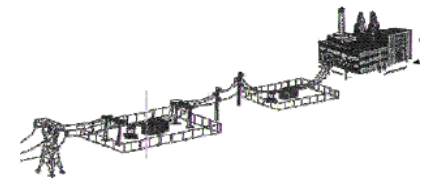
Sensor Data - We're going to Zetabytes

Digital Information Created, Captured, Replicated Worldwide
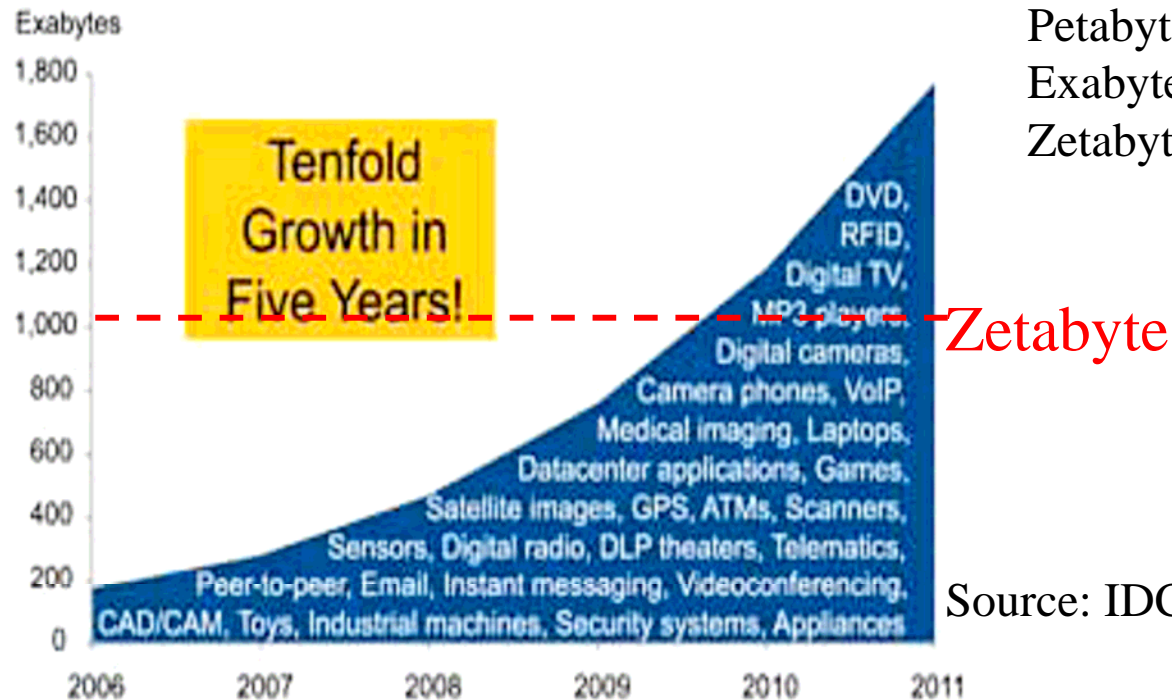
Terabyte = 1 Trillion Bytes
Petabyte = 1 Thousand Terabytes
Exabyte = 1 Million Terabytes
Zetabyte = 1 Billion Terabytes
            = stack of DVDs halfway to Mars

Exabytes

1,800
1,600
1,400
1,200
1,000 — — — — — — — — — — — Zetabyte
800
600
400
200
0

Tenfold Growth in Five Years!

DVD,
RFID,
Digital TV,
MP3 players,
Digital cameras,
Camera phones, VoIP,
Medical imaging, Laptops,
Datacenter applications, Games,
Satellite images, GPS, ATMs, Scanners,
Sensors, Digital radio, DLP theaters, Telematics,
Peer-to-peer, Email, Instant messaging, Videoconferencing,
CAD/CAM, Toys, Industrial machines, Security systems, Appliances

Source: IDC

2006   2007   2008   2009   2010   2011

- Much of the new data coming on stream has either a Location, a Timestamp, or Both

- Most of that data is coming from Sensors or Mobile Apps
       ==> Useful Queries on such data involve spatial or time dimensions

"More data is being generated than a researcher can wrap his head around"

HPTS 2011

New Computational Skill sets and Tool boxes

- ## Business Analytics
  - <u>Heavy Statistics</u>
  - Correlating data based on Location and/or User Preferences
  - Tools such as:
    - R, CRAN, Python, SAS

- ## Sensor Analytics
  - <u>Heavy Calculus</u>
  - Heavy Signal Analysis
  - Tools such as:
    - MatLab, SciLab, Octave, Maxima, SciPy, GSL
    - Flume + Time-series (Lumberyard)
    - Data Visualization Tools - Humans are lousy at grok'ing large reams of numbers

Summary

- Sensor based apps are exploding, right under our noses

- Data storage demands will continue to explode - "the size of the data becomes part of the problem"

- Mathematicians are going to clean up
  - All of the advanced Data Analytics rely on Statistics and Calculus
  - All of the underlying Green Technolgies and DSP technologies rely on Calculus
  - Calculus rules !

$$E = mc^2$$

$$e^{-ti\theta}$$

$$x = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n$$

$$y = \iint \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$x + y = \sum_{n}^{5} z + 1$$