

Fast Data Analytics for the Internet of Things

Yogesh Simmhan



BiDa 2014

23-Aug-14

©DREAM:Lab, 2014

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)





Big Data in the Age of IoT

*Large
Hadron
Collider*

PS1 telescope



Illumina NGS @ IISc



*Bluetooth Mote
@ IISc*



*Smart Meter
@ LADWP*



Few Instruments,
Large Data **Volume**

10^2 Sources
TB's Data
Days to Proc.

Many Devices,
Volume & Velocity

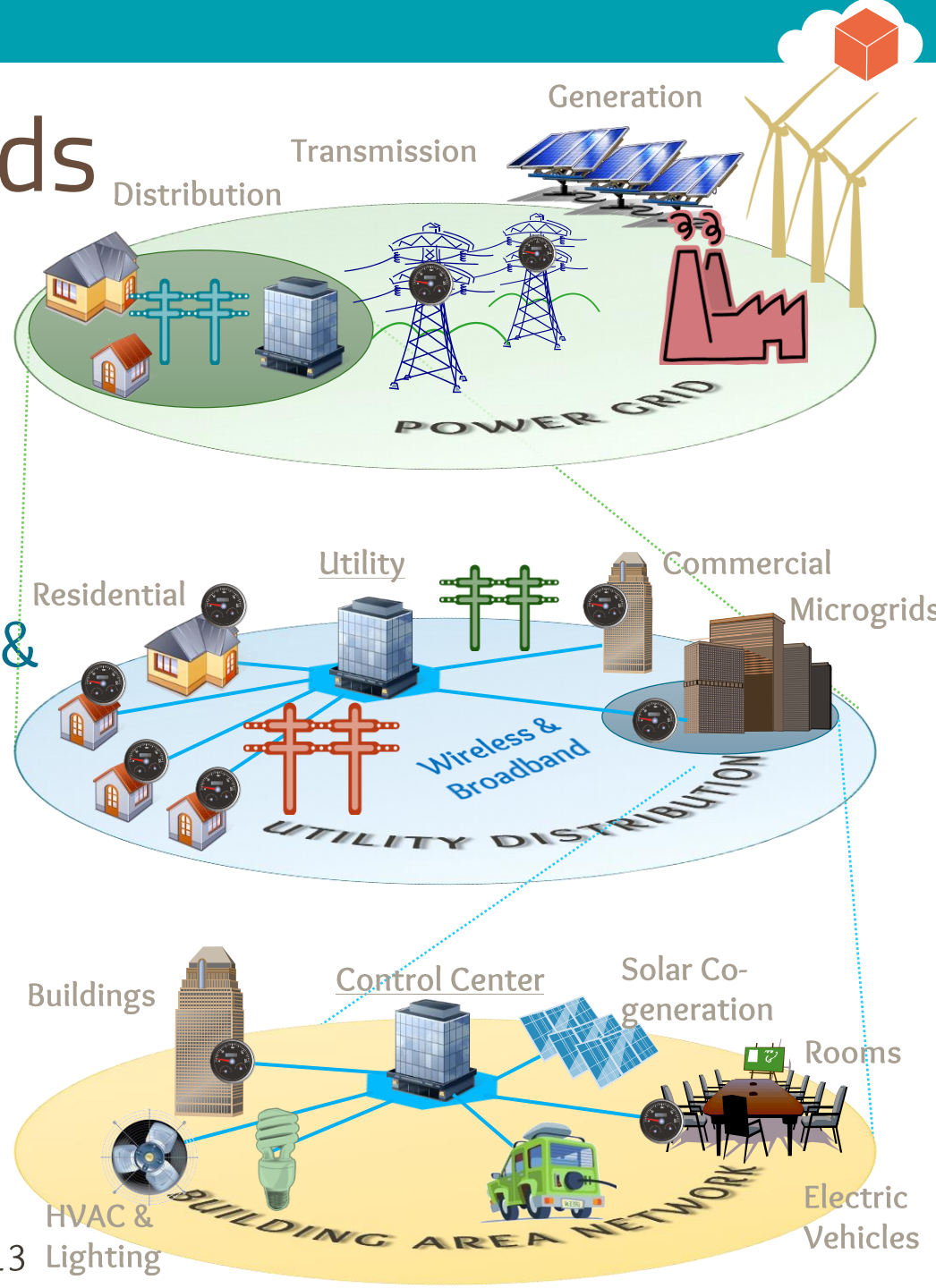
10^5 Sources
GB's Data
Hours to Proc.

Numerous Sensors,
High data **Velocity**

10^8 Sources
MB's Data
<Mins to Proc.

Smart Power Grids

- Integration of Renewables
- Advanced Instrumentation
- Bi-directional communication
- Real-time data acquisition & control
- Self-contained 'Micro Grids'...like USC
- LADWP: largest US public utility

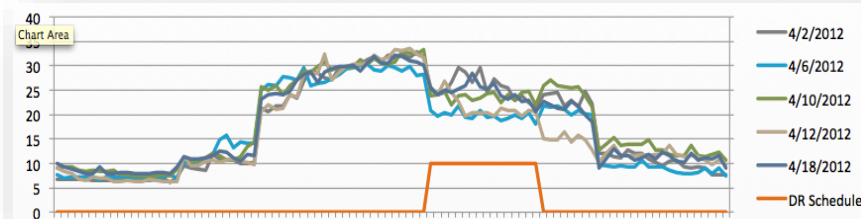
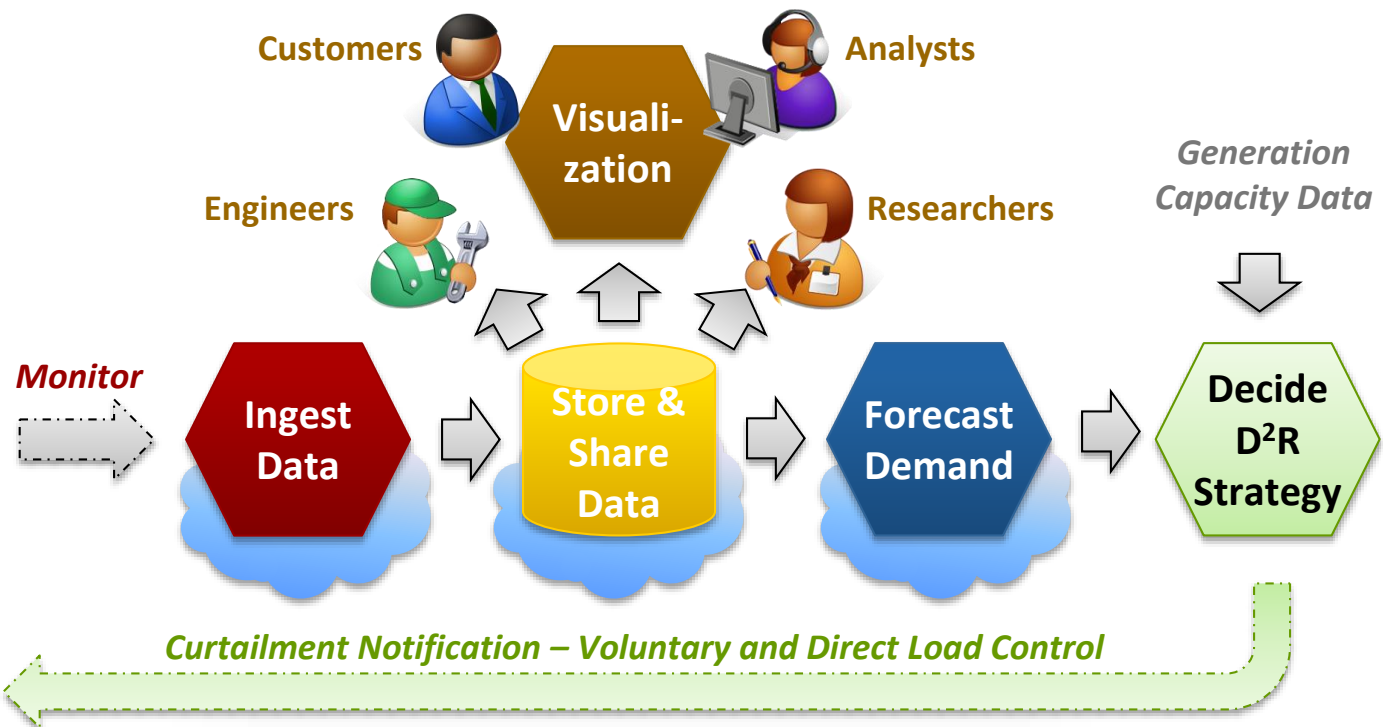
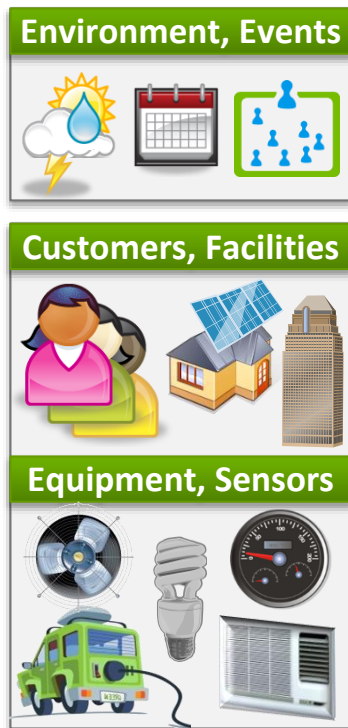


Cloud-based software platform for data-driven smart grid management, Simmhan, et al, CiSE, 2013



***Demand Response Optimization:** Reduce consumer demand for electricity during periods of peak usage to relieve stress on power grid*

When → By How Much → How/Whom ... Predict, Adapt, Evolve





USC Smart Grid Project

smartgrid.usc.edu

Web Portal Admin Panel

USC Smart Grid Web Portal

USC Power Map

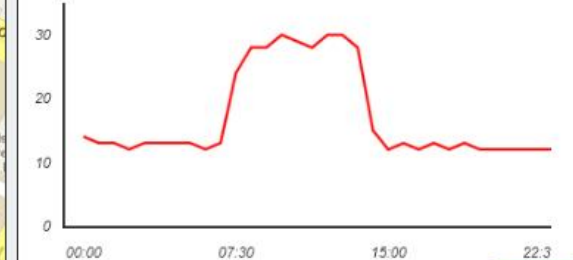
Map Satellite

PERTUSATI UNIVERSITY BOOKSTORE (BKS)

Square Foot (Net): 44,218

1989

Current kWh: 29.23



Export Data

LEGEND:

Power Use

Square Feet (Net)

23-Aug-14

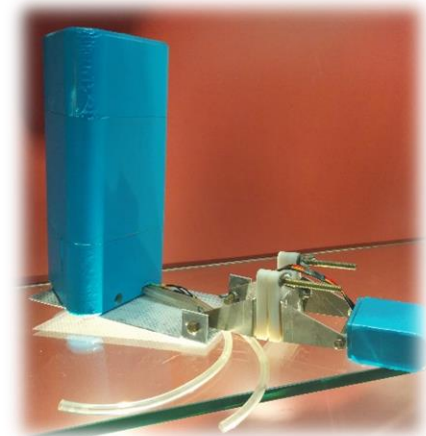
BiDa 2014

5



IISc Smart Water Project

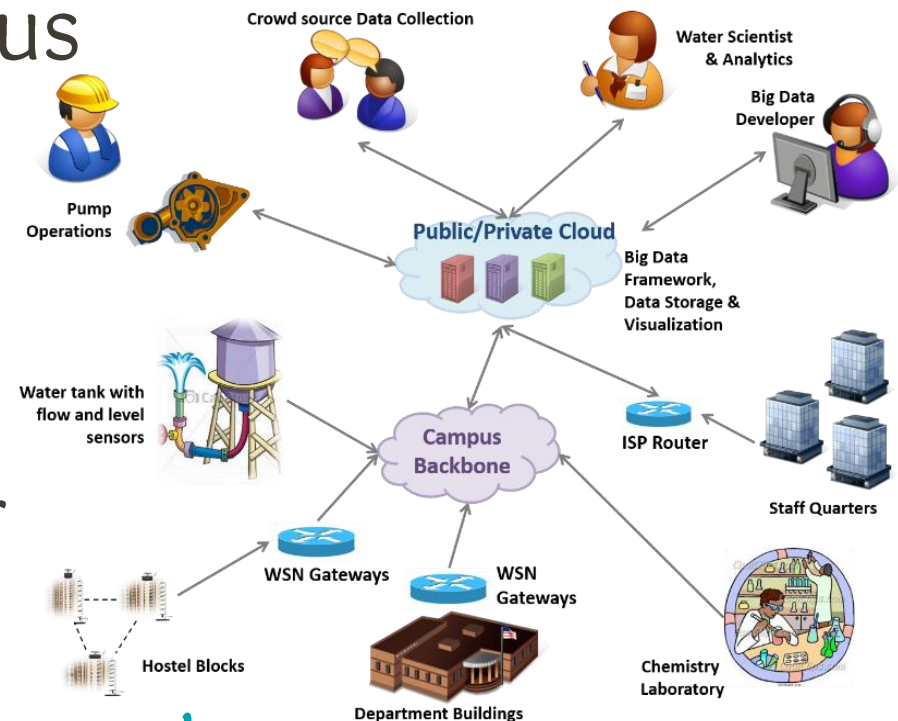
- Sustainable use of water on campus
- 100+ **Water level & quality sensors** on tanks & reservoirs, **Flow sensors** on inlets
 - Height of water
 - Total Dissolved Solids, Temperature
- **Crowd-sourced** data collection
 - 100's Water samples, Usage reports
 - Data Mules thru' Smart Phone Bluetooth





IISc Smart Water Project

- Covers 40% of campus
- Hostels, Depts, Staff Quarters
- Examine Big Data & Cloud Computing for Practical IoT
- *Reduce usage, Improve quality*
- *Extensible IoT Infrastructure!*



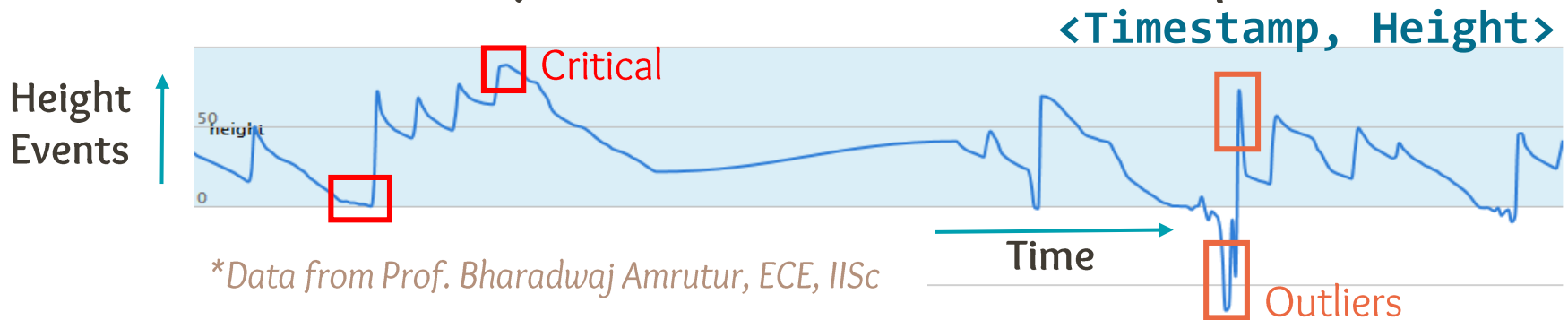


Distributed Analytics



Complex Event Processing

- Extract info from realtime, event sources to help *decision-making*
 - Specify *queries* on situations, patterns, causal relationships
 - Online analysis of 1000's of Events per Sec



```
SELECT e FROM STREAM dese_oht WHERE e.height > 90%
SELECT e_hi, e_lo FROM STREAM rbccps_oht
WHERE (e_hi.height - e_lo.height) > 5%
WITHIN WINDOW(5mins)
```



Analytics from Edge to Cloud

- Traditional CEP Processing has been centralized
 - But IoT Event sources are *distributed*

CEP only on Cloud?

Latency, Privacy of moving Data from the Edge

Longer time to Respond

CEP only on Edge?

Limited Expressivity & Compute Capability

Need to integrate realtime with offline Big Data



- CEP in a distributed IoT environment
 - Capable edge devices, Smart Phones
 - Heterogeneous computing: Cloud + Edge
 - Distributed realtime analytics for IoT

Can we process event streams across Cloud & edge through efficient query partitioning to meet QoS Goals?



Optimization Problem

- *Match a query Q within time T of input events*
- **Constraints**
 - Network latency
 - Data privacy
 - Compute capability
 - Expressivity
- **Objective to Minimize**
 - Execution costs
 - Energy consumption





Solution Approach

- Solve optimization problem using *dynamic programming*
 - » Model query as a DAG.
 - » Decide edge cut that meets objectives.
- **Distributed CEP on Android & Cloud***
 - » CEPLite engine on Android
 - » Full featured Siddhi CEP on Cloud
- **Deployment on IISc for Smart Campus**
 - » Sustainable water management
 - » Tank Overflow, Refill, Leakages, Quality

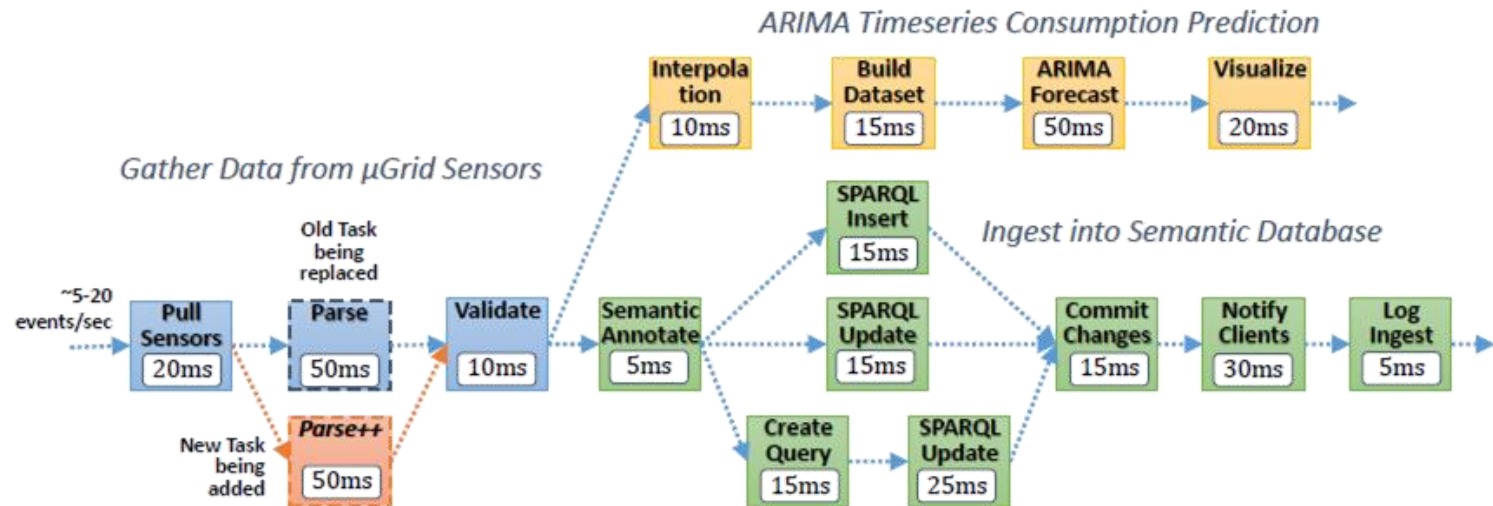


Reliable Analytics



Mission-critical Dataflows

- **Continuous Dataflows:** Directed Acyclic Graphs that perform stream processing
 - Tweet Analysis (Twitter's Storm), Smart Grids

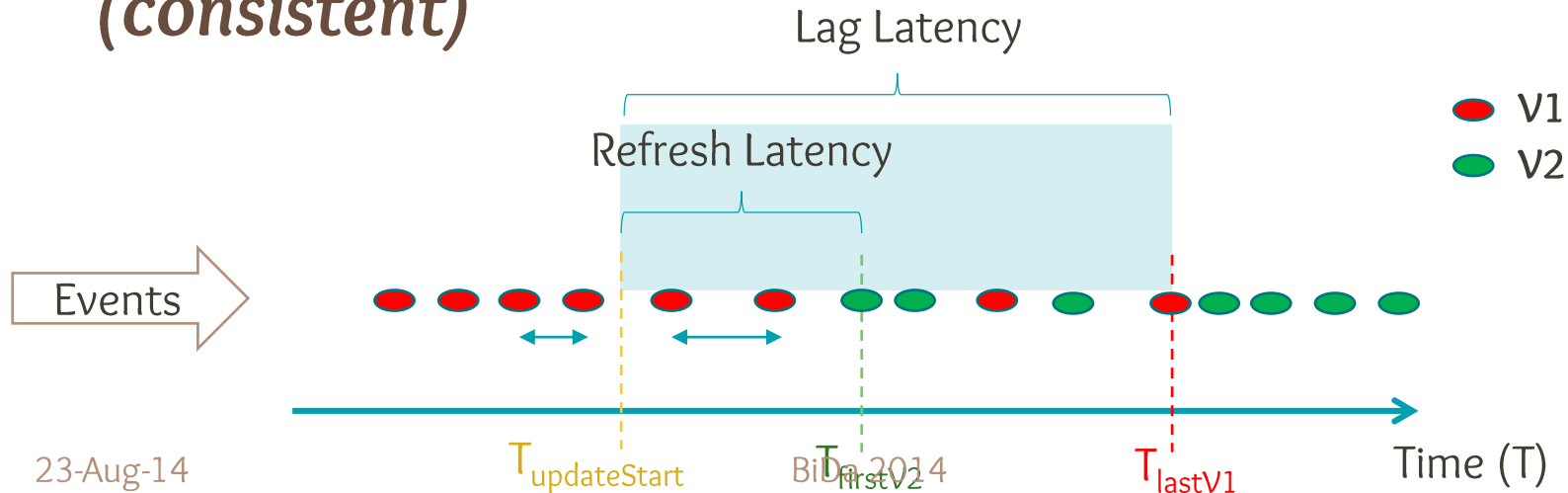


- **Mission Critical Apps cannot have downtime**
 - How to update apps in realtime
 - Algo upgrade, bug fixes, new analytics












Online Updates with Guarantees

- **Challenges** → Complex coordination in distributed environment
→ Need *formal bounds & metrics* for updates
- *Existing frameworks like Storm lack these*
- **Goals** → Reduce time for new task to come online (**refresh**) → Reduce time old task remains (**lag**) → Ensure events are processed by either, not both (**consistent**)





Solution Approach & Evaluation

	Consistency	Throughput	Refresh Latency
Naïve Consistent			
Naïve Inconsistent			
Versioned Consistent			

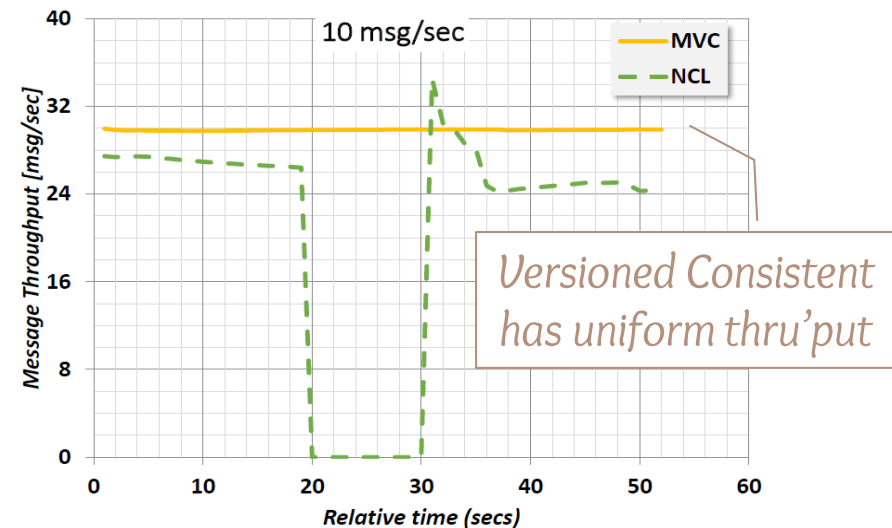
Apache Storm

Flush DAG. Update. Restart.

In-place update of DAG.

In-place update. Tag event.
Route to old or new task.

- Incorporating strategies into Storm
- Benchmarking on Clouds

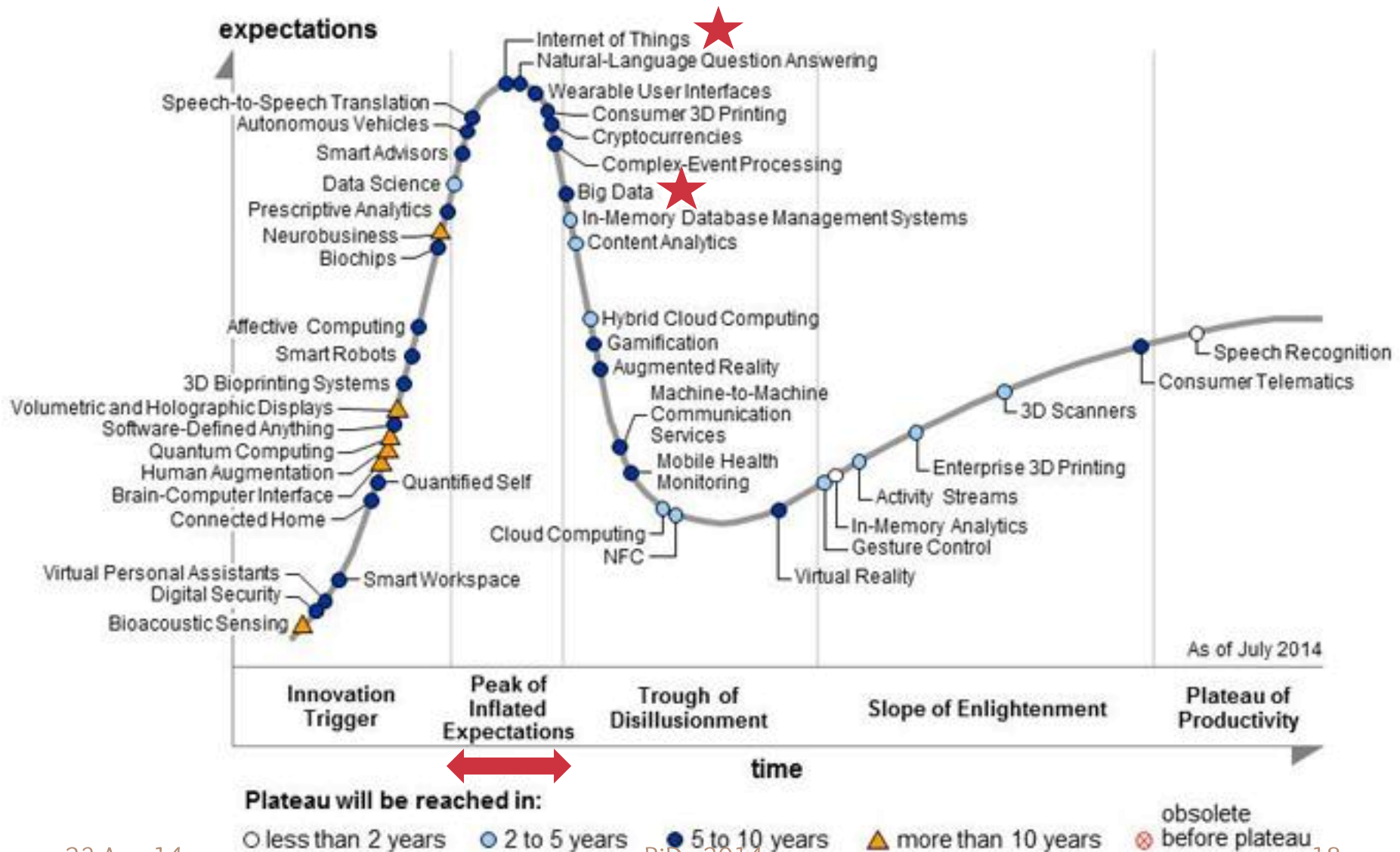




IoT for India



The Hype Cycle





india budget 2014 smart cities - Google Search - Mozilla Firefox



india budget 2014 smart cities

Smart Cities

Web

News

Images

Videos

Maps

More ▾

Search tools

About 22,500 results (0.32 seconds)



Union Budget 2014-15 talks of Digital India, Skill India an...

India Today - 10-Jul-2014

A strong and vibrant **India**, the catch word of Finance Minister Arun ... key policies - Digital **India**, Skill **India** and **Smart Cities** - to reach the goal.

Budget 2014: Government allocates Rs 7060 crore for 100 smart cities

Times of India - 10-Jul-2014

Union **Budget 2014**: Govt allocates Rs 7060 crore for 100 **smart cities**

Zee News - 10-Jul-2014

Budget 2014: Rs. 7060 crore set aside for PM's '**smart cities**' dream

Firstpost - 10-Jul-2014

Budget 2014: Government Allocates Rs 7060 Crore for 100 **Smart ...**

NDTV - 10-Jul-2014

[Explore in depth](#) (42 more articles)



The Hindu

Union Budget 2014: Go city! 100 smart urban hubs on dr...

Times of India - 10-Jul-2014

Keeping its poll promise, the Narendra Modi government allocated Rs 7,600 crore for developing 100 **smart cities**. The government also ...



Union Budget 2014: Makings of a true 'Digital India'

Livemint - 10-Jul-2014

Smart cities, too, are a need of the hour since **India's** urban population is estimated to reach 590 million by 2030, and they will live in at least 60 ...



IoT for India[†], with RBCCPS

- *Human-centric rather than Thing-centric*
 - Span **Virtual** and Physical Worlds
- *Analytics from the **Edge to the Cloud***
 - **Big-Little Data**
 - **Context** determines the Action
 - **Privacy**: Whose data is it anyway?
- *How “Low” can you go ?*
 - Bring the Network to the Sensor
 - When “**good enough**” is enough ?

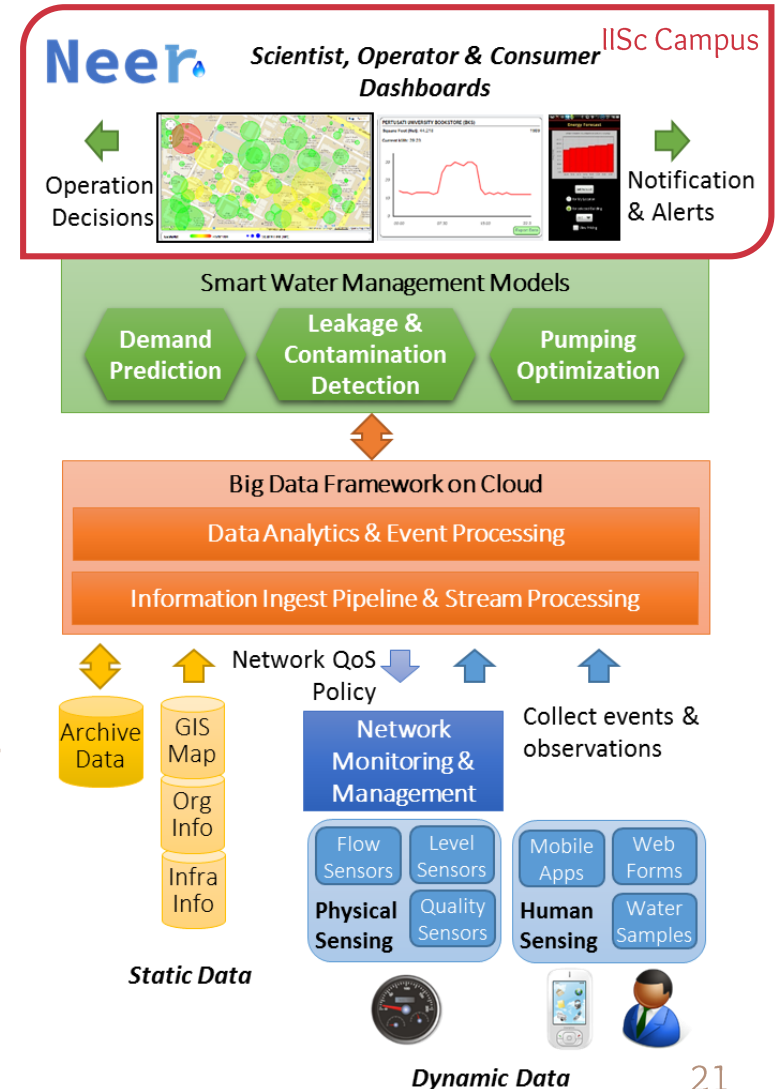




Open, integrated & extensible IoT Technology Stack for Smart Campus Management

1. Hybrid sensing
2. Adaptive networking
3. Realtime Big Data Analytics
4. Science-driven decision making

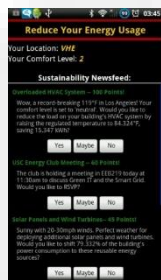
- » Experts to *Close the Loop from Network to Knowledge*
- » Validate using *real-world deployment @ IISc*





Physical & Crowd-sourced **Hybrid Sensing**

- Physical sensing infra's cost & maintenance
 - *Leverage existing infra (smart phones) & humans*
- Low-cost Physical Water Sensor & Actuator
 - Water-level & quality, flow control
 - *Expands on prior funded project from DeitY*
- Smart Phone Platform for Crowdsourcing
 - Smart phones as data mules using *Beacons*
 - *Subjective & objective* tests + Physical samples
 - Respond to signals to control pumps, usage
 - *Actively engage consumer. Soft incentives.*





Summary



Take Aways

- Big Data has immense possibilities
 - But tease out the signal from the noise
 - Find the “next big thing” in Big Data
- Big Data is not just technology
 - Fundamental research into algorithms & systems necessary
- IoT offers novel challenges to Big Data
 - Pervasive, Distributed, Velocity
 - And has significant social impact



We need you!

- *Motivated students for Ph.D., M.Sc. @ IISc*
 - **DREAM:Lab** – Research on Distributed Systems, Clouds, Big Data & IoT
 - **SERC** – Computer Systems Research, Leadership class supercomputing labs
 - **IISc** – The Top grad school in India, Research focus, Faculty with global profile
- *GATE Apps due end of **Sep**, 2014*
- *IISc Apps due **Mar**, 2015*



DREAM:Lab

DISTRIBUTED RESEARCH ON EMERGING APPLICATIONS & MACHINES

dream-lab.in

Indian Institute of Science

Thank you!



©DREAM:Lab, 2014

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

