Cloud-based Infrastructures for Large-scale Smart Transportation Systems

Aniruddha Gokhale, William R. Otte and Abhishek Dubey
{gokhale,wotte,dabhishe}@isis.vanderbilt.edu

Institute for Software Integrated Systems, Dept of EECS,
Vanderbilt University, Nashville, TN, USA

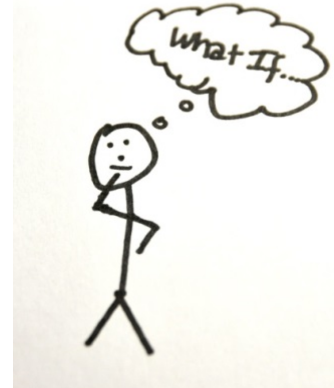
NSF CPS Transportation Workshop, Jan 23-24, 2014, Arlington, VA

Session 3: Infrastructure, Cloud-based Systems and Green Approaches

VANDERBILT UNIVERSITY

INSTITUTE FOR SOFTWARE INTEGRATED SYSTEMS

AF 447: A crash that should never have occurred



- An Airbus A330 plane crashed in the Atlantic killing all 228 on board
- Pilots lost situational awareness and made wrong decisions until it was too late to rectify the mistake
  - Pitot tubes froze => no indication of speed, altitude readings
  - Auto pilot got disengaged
  - Pilots decided on a nose-up, which led to a stall and a continuous loss of altitude
  - Plane crashed belly first in the ocean

- We know that the plane was constantly sending various parameters via the Aircraft Communications and Addressing System (ACARS) to France
  - What if real-time data analytics on this info raised observed anomalies?
  - Could pilots have been alerted through other out-of-band mechanisms, e.g., Reverse path? Satellite radio?
- Could other planes in the vicinity have monitored each other and run local analytics to detect anomalies?

Pacific Highway 101 Vehicle Crash: Jan 2012

- Jan 2012, Buellton, California
- Northbound gravel truck struck a BMW from behind, trampling it and dragging it underneath along a bridge
- Truck fell in the deep ravine and burst into flames; the sedan remained hanging on the bridge
- Truck driver had high levels of certain drug
- The 3 passengers in the sedan miraculously survived after being extricated

- We are living in an increasingly smart world
- Smart appliances are already being used for monitoring patient's vital signs
- Could these smart appliances observe human behavior
- Run analytics to deduce anomalies:
  - Lack of sleep, excessive drinking
- Disseminate potential hazards to interested subscribers (e.g., traffic enforcement)
- Such drivers can then proactively be caught before they cause harm

Oklahoma City Tornado: May 31, 2013

- May 31, 2013 EF3 tornado in Oklahoma City killed 13 people including 3 veteran storm chasers
- Commuters on I-40 during rush hour were caught unaware due to sudden change in direction of tornado
- No timely warnings were available; escaping was impossible due to traffic jams
- A few days earlier, 51 people were killed in an EF5 tornado in the same region.

- Collected sensor data and resulting real-time analytics could predict the path of tornado at shorter time scales
- Leverage the fact that the world is more interconnected now.
  - Smartphone-based connectivity
  - Onboard capabilities with V2V and V2I capabilities
- Need for interconnectivity among multiple agencies
  - Weather, traffic, police, hospitals

Service Disruptions: Maintenance/Sabotage/Wear & Tear



- Service disruptions are caused by many events (normal wear/tear, weather, sabotage)
- Often the solutions are reactive and are put in place after some tragedy occurs (e.g., accidents and loss of life)

- Sensors of different kinds continuously monitor health
  - E.g., train wheels can measure rail width and wear/tear, road sensors can detect emerging potholes,
- Analytics
  - Localized (e.g., on the train), global (in the cloud)
- Real-time dissemination and proactive corrective actions

Many More Examples





- Calls for improved situational awareness, proactive alerting capabilities, autonomous control and corrective actions to improve safety
- Calls for minimal disruptions and downtime => high availability
- Calls for sensing-collecting-analytics-dissemination-actuation cycle
- Reliable and resilient architectures with effective control over cyber resources (compute nodes, storage, networks)

Observed Traits

- Massive amounts of sensory data from variety of sources needs to be collected
- The data must be acted upon in real-time
  - Data processing can take place locally, remotely, part local-part remote
- The kinds of analytics needed will depend on the situation
- Information must be disseminated in real-time
- Proactive and semi/fully autonomous corrective actions are necessary
- Reliability and resilience of the system is critical
- Privacy preservation
- Security and trust

Challenges Galore (1/2)

- Heterogeneity of involved systems, devices and technologies
- Mobile and stationary sources of information
  - Number and locations of sources may fluctuate over time
- Massive amounts of information from variety of sources
  - Collection, aggregation, dissemination, processing challenges
  - Real-time stream processing, queries, etc. needed
  - Data models, structured vs unstructured data
- Type of analytics is not pre-defined
  - Situation may demand changing what information is sought
  - May result in different resource demands
- Where to execute these analytics?
  - All locally, distribute it across ad hoc resources, in traditional cloud, part local, part distributed?
  - Distributing and placement of the “intelligence” or the “smarts” is important
  - Needs physics-aware dynamic resource management
  - How to aggregate the results from distributed sources and make sense out of it?

Challenges Galore (2/2)

- Existing cloud platforms are not yet suitable for real-time and highly reliable performance
  - RT Xen, our efforts at real-time VM failover, etc are promising directions
- Information dissemination challenges
  - Pub/sub is attractive (e.g., DDS) but many challenges to make it work in such set ups, across heterogeneous n/ws, discovery, etc
  - No control over networking resources; Software-defined Networking shows some promise
- Deployment and configuration hurdles
  - Need predictable and resilient capabilities
  - Orchestrated vs choreographed approaches
- Massive integration challenges
  - Differences in APIs, hosting models, resource models across cloud providers
  - Lack of standardization; Efforts such as FACE are promising
- Security and privacy concerns
  - How much to monitor? What is allowed? How much can be shared and used?

Concluding Remarks

- Many unresolved challenges in CPS Smart Transportation
- Existing solutions continue to be independent, vertical silos
- Needs integration across heterogeneous CPS
- **CPS Middleware frameworks** that provide reusable and composable capabilities are an important consideration
- Need to leverage technological advances and adopt emerging paradigms, such as IoT
- Collaborative R&D is a must
- Need good testbeds for experimentation – not just transportation testbeds, but sensor systems, cloud platforms, networking testbeds, mobile devices, etc all integrated

Thank You.

Questions?