SmartFarm
Hybrid Cloud-based Sensor System for Simplifying and Automating Agriculture Analytics

Chandra Krintz (ckrintz@cs.ucsb.edu)
Rich Wolski (rich@cs.ucsb.edu)

http://www.cs.ucsb.edu/~racelab
Computer Science Department
UC Santa Barbara
Nov 2014
Motivation

• Ag Industry -- under tremendous pressure to increase food production and yields to keep up with demand
  – Population growth is outpacing production
  – Natural resources are limited (arable land, water, ...)
  – Hard to predict weather patterns, climate change

• **Goal of our work**: investigate a novel, unifying, and open source approach to agriculture analytics and precision farming called *SmartFarm*
  – Open source, secure, on-premise (private) cloud
    • Easy to use, self managing, *lights-out automation (think Tivo!)*
  – Linking disparate sensor technologies and analysis engines
  – To give farmers actionable insights, alerts, accurate prediction
The UCSB SmartFarm

- Internet data
- Climate data (NOAA)
- External data sources (ingress only)
- Aerial imagery
- UAVs
- Security boundary
- Actionable analytics, alerts
- APIs
- Private Cloud
- Public Clouds
- Many sources of farm sensor data
UCSB SmartFarm Technologies

• Private cloud is API compatible with Amazon and Google Clouds
  – Any application or service that executes in public cloud can also execute in on-farm without modification
  – Facilitated via AppScale and Eucalyptus private cloud systems
  – Hybrid configurations
    • Move computation to the data on-farm; control all data that leaves farm
      – Precision farming “app store”
    • Or perform computation in the public cloud; shared datasets and services

• Leverages emerging Internet-of-things (IOT) & big data technologies
  – Unified data ingress from disparate sensor systems
    • Images and data
  – Machine learning and data mining
  – Streaming (think Storm) and batch processing (think Hadoop/Spark)
Proposed Research Goals

A novel, unified software system that provides

- A cloud-based software architecture that links data sources with data aggregation, processing, and analytics algorithms,
- Data ingress for disparate sensor systems (local and remote), each potentially with their own data formats, time scales, etc.,
- A new data model based on document referentiality, geolocation, and eventual consistency,
- Application programming interfaces (APIs) for easy and automatic integration of modeling, analysis, alerting, auditing/reporting, and visualization technologies, and
- Privacy and security mechanisms that restrict access to data and the system to only authorized parties.