Clouds and the Internet of Things (Sensor Clouds)

Geoffrey Fox 24 November 2012

This project builds on existing worked funded by Wright Patterson Air Force Base (AFRL) where program manager Bill McQuay has retired and I do not expect further funding. The final report of project is at http://grids.ucs.indiana.edu/ptliupages/publications/Sensor%20Cloud%20Final%20Report.pdf while we have publications including

- Geoffrey C. Fox, Supun Kamburugamuve, Ryan Hartman "Architecture and Measured Characteristics of a Cloud Based Internet of Things API" Workshop on IoT Internet of Things, Machine to Machine and Smart Services Applications (IoT 2012) Denver, Colorado, USA, http://grids.ucs.indiana.edu/ptliupages/publications/cts_2012_paper-2.pdf

- Geoffrey Fox, Alex Ho, and Eddy Chan, "Measured Characteristics of FutureGrid Clouds for Scalable Collaborative Sensor-Centric Grid Applications" IEEE International Symposium on Collaborative Technologies and Systems CTS 2011, Waleed Smari, Editor. May 23-27 2011, Philadelphia. http://grids.ucs.indiana.edu/ptliupages/publications/cts_2011_paper_mod_6%5B1%5D.pdf

And typical presentations are:

- Geoffrey Fox "Clouds for Sensors and Data Intensive Applications" May 13 2012 Presentation at 1st International Workshop on Data-intensive Process Management in Large-Scale Sensor Systems (DPMSS 2012): From Sensor Networks to Sensor Clouds at CCGrid 2012 http://grids.ucs.indiana.edu/ptliupages/presentations/SensorClouds-CCGrid-May14-2012.pptx

- Alex Ho and Geoffrey Fox Distributed Clouds for "Scalable Collaborative Sensor-Centric Grid Applications" AMSA TO 4 Sensor Grid Final Presentation Dayton OH May 9 2012 http://grids.ucs.indiana.edu/ptliupages/presentations/AMSA%20TO4%20Final%20Presentation%20May%209%202012%20-%20v3.pptx

The project is documented by two web sites https://sites.google.com/site/sensorcloudproject/ (original) and https://sites.google.com/site/opensourceiotcloud/

The project started studying hierarchical grids of sensors showing how many sensor types could be integrated with the same framework. We demonstrated that webcams, GPS, game controllers (Wii, Kinect), Lego Robots, "hexacopters", RFID, tablets and smartphones could be tackled with same architecture. The early work was in context of DoD's "net-centric" architecture for the "Global Information Grid" with an extensive summary given at http://grids.ucs.indiana.edu/ptliupages/publications/gig/. Always sensors form a Grid which is a managed distributed system. However it has become clear that clouds are particularly well suited to perform this management which consists of controlling the sensor, updating the local sensor on demand and giving sensors access to the almost unlimited information available through the cloud. Technical advantages of clouds for this application come from sensors tending to be small and needing support on demand. This is well suited to a cloud architecture where sensors are controlled by specialized "virtual

machines" each running on a single core. There is no parallelism needed by a single sensor; rather parallel computing (which can have poor performance on clouds if tightly coupled) is used to integrate results of multiple sensors; a loosely coupled process which performs well on clouds.

Our DoD work initially focused on the net-centric aspects and for the last two years security issues (trusted sensor networks) and has been performed with two companies Ball Aerospace and Anabas but current work is solely at IU. There are of course many applications of Sensor Grids/Clouds outside DoD. Indeed one can argue that the explosion in use of tablets and smartphones today is driven by their linkage to clouds to provide back up and access to information. Note that a sensor can be defined as any source of time dependent data and so an individual interacting with a smartphone is "just a sensor" and crowd sourcing is "just the integration of this sensor input". So sensor clouds encompass much social media work and they are also relevant in many crisis management situations. For example, we were earlier funded by NASA to adapt this sensor architecture to manage the collection of GPS sensors in Southern California. This was part of an activity QuakeSim that recently was joint winner of NASA's software of the year award. Earthquakes provide an interesting test case with (GPS and seismic) scientific sensors providing data streams; first responders to the earthquake provide reports in video, text and other form; simulations are forecasting chances of damage and tsunamis and this is shared with the global community. The latter interact with their "smartphone sensors" and already the value of social media has been demonstrated by Google and others with resources like http://www.google.com/crisisresponse/japanquake2011.html that help link families (and relevant sensor data) together. Our earthquake forecasting collaborators consider this area as a promising opportunity to benefit science and society by building a sensor analytics platform linking social media and scientific instruments. It is well known that tweet feeds (a modern popular sensor) filtered correctly often signal earthquakes before they are officially announced.

Another important area is robotics and Google has also stressed the importance of clouds to manage robots. In that clouds are necessarily distant from the robot, one uses onboard software for any decisions that are needed in <~ 0.1-1 seconds but planning and other tasks that can afford waiting the round trip time to and from the cloud are in fact best done there.

At the conclusion of the AFRL funded work, we decided to generalize our framework and focus it on management of the Internet of Things where we are motivated by the estimate that there will be 24 billion devices on the internet by 2020. We are using the same successful idea that we pioneered in initial DoD work; namely that so called publish-subscribe technology is well suited to control a sensor grid. The technical work involves efficient cloud implementation, support of metadata defining nature of sensor and its data and optimizing communication with multiple transport protocols where for example webcams would use UDP protocol and GPS sensors TCP. We have defined open API's (application interfaces) that allow users to add their own sensors, transport protocols, metadata management and higher level filters. We will use it in our funded NASA work to support data gathering of SAR radar.

Future work includes integrating this with a complete cloud stack to provide "Sensors as a Service" as shown in the figure. Note pervasive importance of security. We have extensive expertise at the infrastructure level (IaaS) where we can instantiate systems on demand that can then support the sensor cloud dynamically on virtualized and non-virtualized environments. The Sensor service will also include Platform support of data storage (probably with a modern system like MongoDB) and parallel MapReduce processing. Here Indiana University has an excellent reputation as we introduced and continue to improve the first iterative MapReduce framework; an approach that is essential in several data analytics (on sensor data). The net result is a Sensor data analytics platform. An important area where one of my students is finishing a PhD is in the real time annotation of sensor data; this is a common requirement where data (say from an earthquake) is streamed to many sites and individuals are annotating selected streams and these annotations are shared with other scientists and/or first responders. There were difficult performance issues to solve here. We originally designed this for telemedicine and eSports applications. People (like cars!) are being instrumented by sensors so remote patients and athletes can be linked via the cloud to streams of data that can be used for diagnosis and monitoring. Note as all sensor data is sent to a cloud controller it is easy to support collaboration as one just broadcasts data to multiple recipients; a capability built into publish-subscribe architectures such as we use. This discussion points out that a Sensor cloud is the natural architecture to support the rapidly expanding personalized medicine field. Other features shown in the figure include cloud bursting where sensor (robot) applications automatically send all or part of their processing to the cloud. At the SaaS (application level) this platform supports X-Informatics for the different fields X using sensor data; we have expertise including robotics, radar, health and social media.

Sensors as a Service
Military Command & Control
Personalized Medicine
Disaster Informatics (Earthquakes)