Large-Scale Analytics and Visualization of Military Medical Information
Chen Li, Jianfeng Jia, Michael Carey, Aditi Majumder (UC Irvine)
Simon Su (Army Research Laboratory)

http://cloudberry.ics.uci.edu/

Overview

- **Goal**: Study how to integrate, store, index, query, and analyze medical information using a big data management system called AsterixDB and develop tools to support powerful visualization of the information.
- Based on Cloudberry, a platform for interactive analytics and visualization of Spatial, Temporal, and Text data, using the Apache AsterixDB system.
- Allows users to explore multi-granularity data from the highest level to the finest level
- Implemented using a full-stack three-tier architecture including
  - **Front-end**: a modularized JS library to generate the user interface;
  - **Middleware**: an intelligent server to efficiently transform a front-end request to an AsterixDB query;
  - **Database**: using an AsterixDB cluster to continuously ingest, store, index, and query large data sets.
- Interactive visualization on a large display with multiple projectors.

### Implementation

- Twitter Dataset
  - Number of records: > 60,000,000
  - Collection period: From 2016-06-30
  - Total data size: > 60G
- Technology Stack
  - Open source: https://github.com/ISG-ICS/cloudberry

Sample Tweets

- The Heller School @TheHellerSchool
  - Don Shepard lends his expertise in the economics of dengue fever to projections about Zika in the U.S.
  - moneynews.com/2016/07/19/zika

- Zika Virus Spreads to NYC, Ab... New York City is now host to the...