Cloud-Based Perception and Control of Sensor Nets and Robot Swarms

AFOSR FA9550-13-1-0225

DDDAS Program Review

IBM T.J. Watson Research Center

December 3 2014

Geoffrey Fox, David Crandall, Kris Hauser

Students: Supun Kamburugamuve, Hengjing He, Jangwon Lee, Jingya Wang

gcf@indiana.edu

http://www.infomall.org

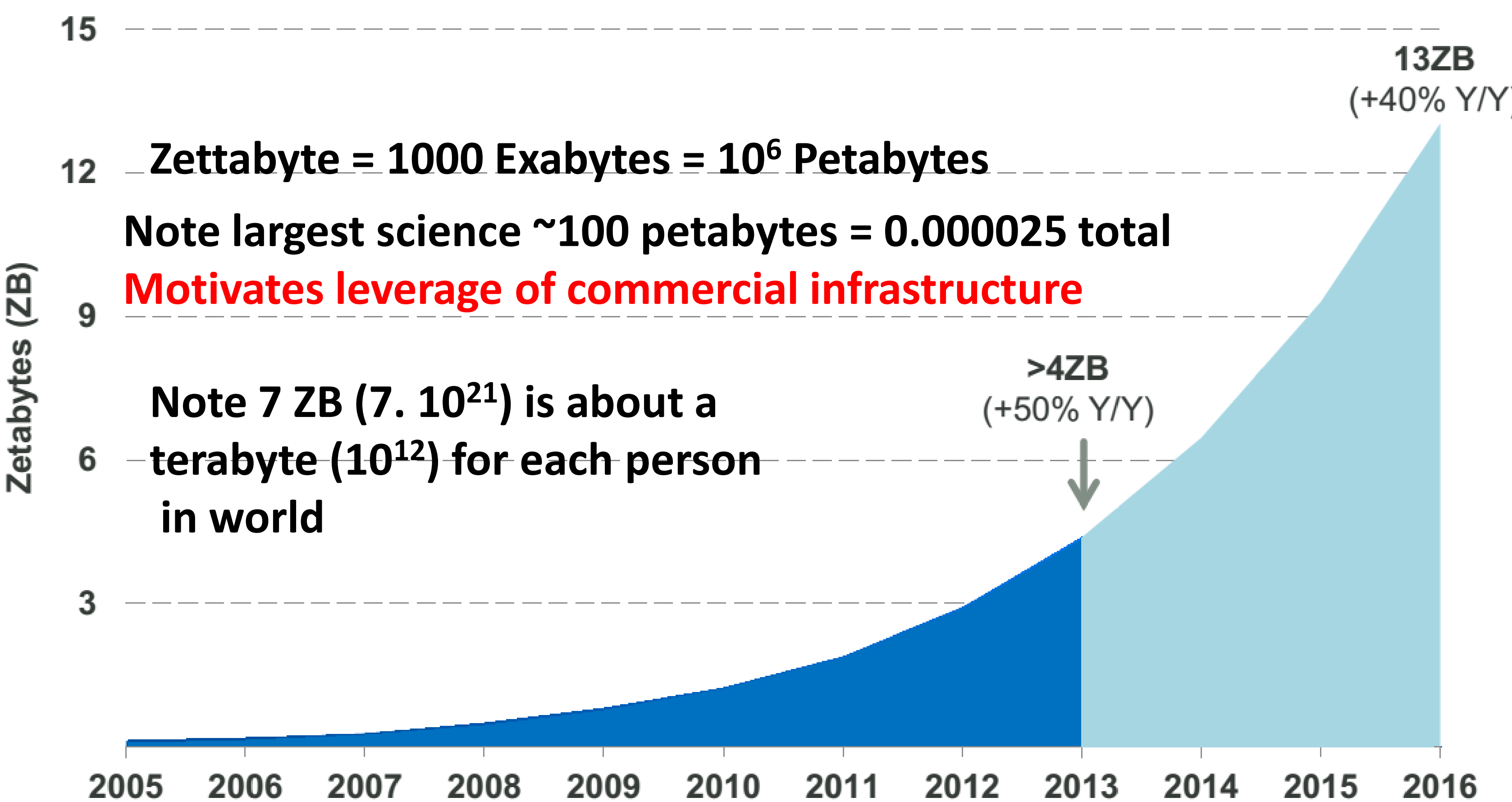
School of Informatics and Computing

Digital Science Center

Indiana University Bloomington

‘Digital Universe’ Information Growth = Robust... +50%, 2013

2/3rd's of Digital Universe Content = Consumed / Created by Consumers ...Video Watching, Social Media Usage, Image Sharing...



Zettabyte = 1000 Exabytes = 10^6 Petabytes

Note largest science ~100 petabytes = 0.000025 total

Motivates leverage of commercial infrastructure

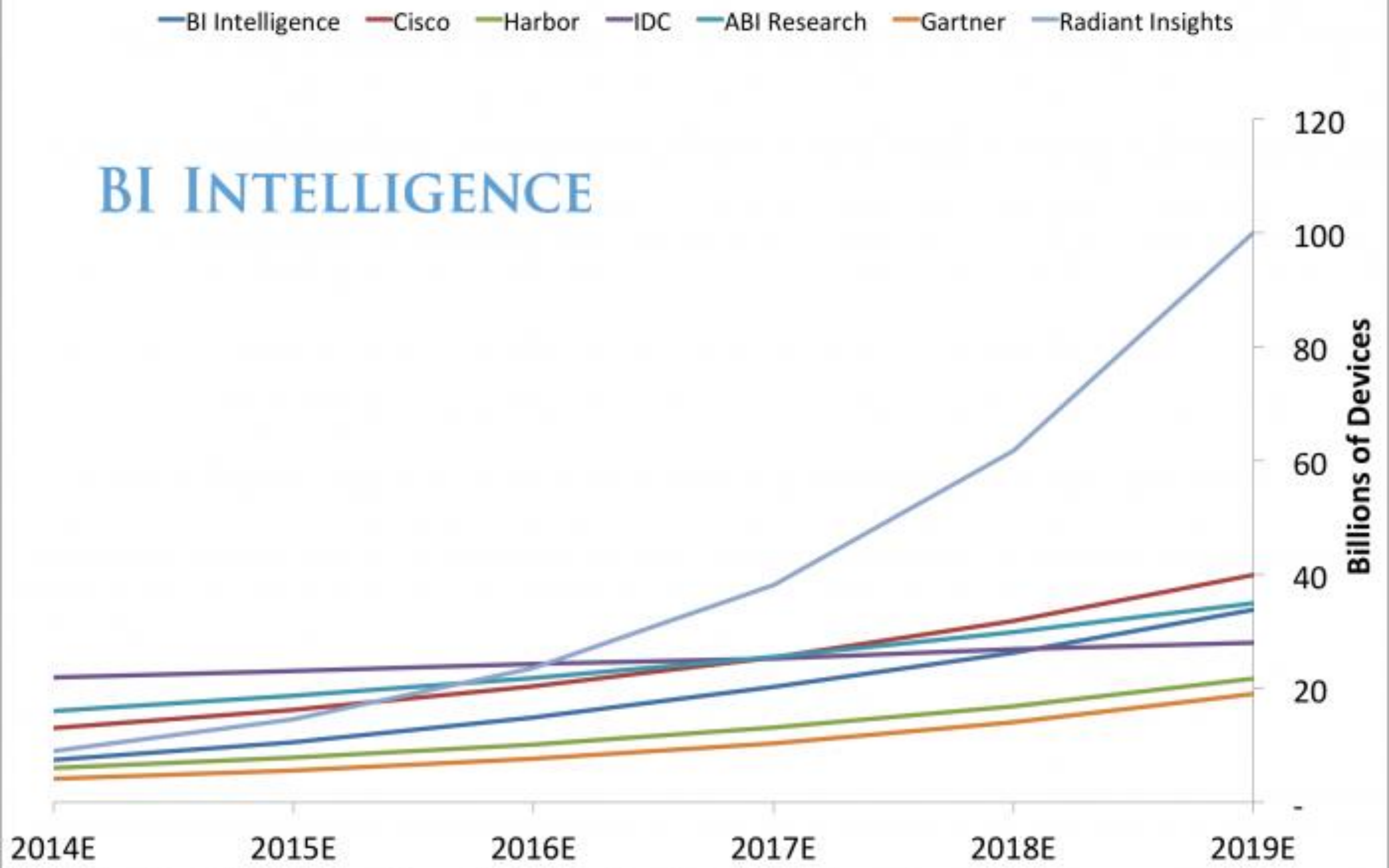
Note 7 ZB (7.10^{21}) is about a terabyte (10^{12}) for each person in world

@KPCB

Note: 1 petabyte = 1MM gigabytes, 1 zetabyte = 1MM petabytes.
Source: IDC Digital Universe, data as of 5/14.

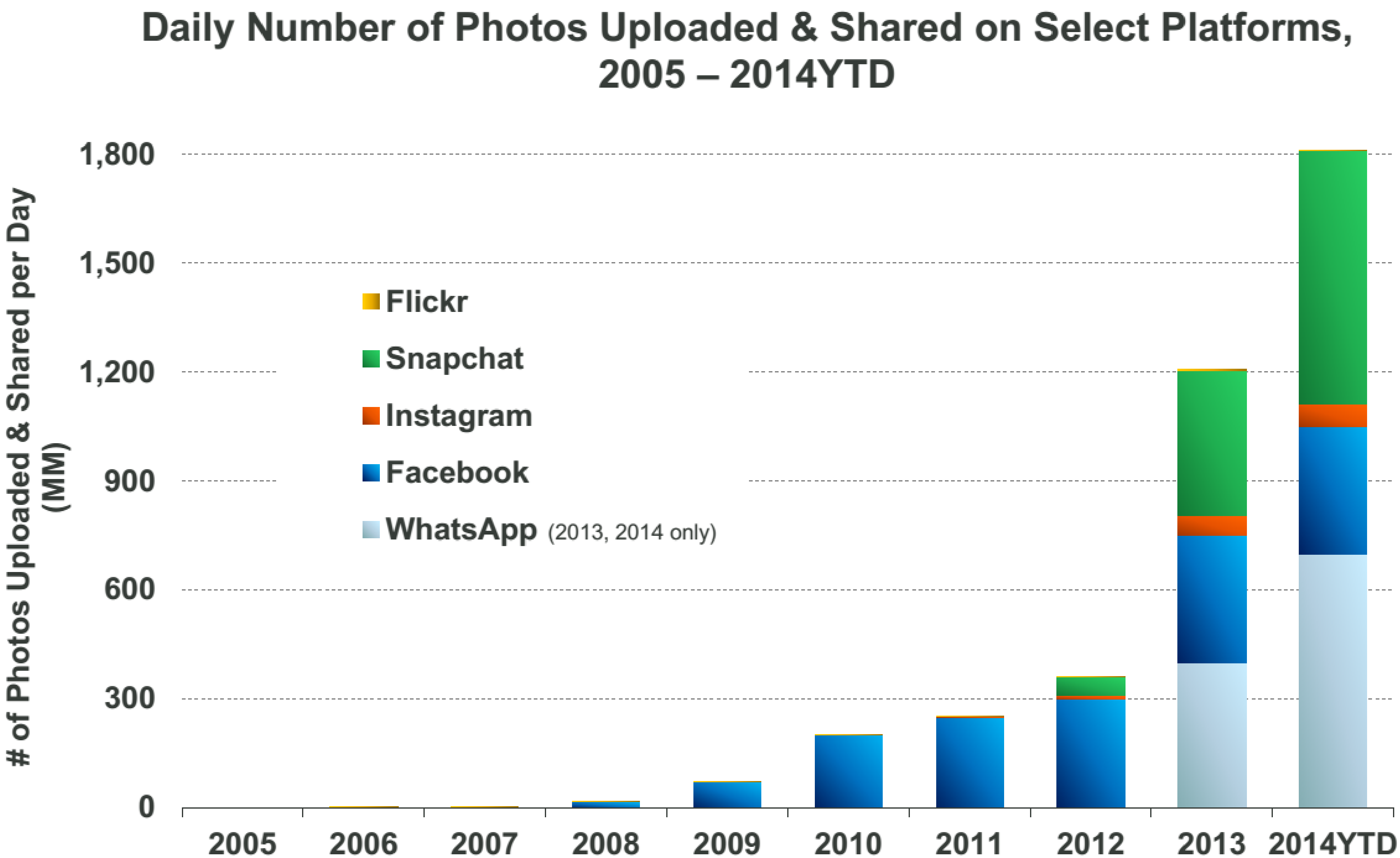
http://www.kpcb.com/internet-trends

Industry Estimates of Internet of Things Devices



**Note: Definitions of what is included in the Internet of Things varies by firm. For example, some firms include remotes (i.e., smartphones, tablets, etc.). BI Intelligence does not include remotes in our estimates.*
Source: Individual Research Firm's Repo

Photos Alone = 1.8B+ Uploaded & Shared Per Day...
Growth Remains Robust as New Real-Time Platforms Emerge



@KPCB

Source: KPCB estimates based on publicly disclosed company data, 2014 YTD data per latest as of 5/14.

http://www.kpcb.com/internet-trends

51 Detailed Use Cases: Contributed July-September 2013

http://bigdatawg.nist.gov/usecases.php, 26 Features for each use case

- **Government Operation(4):** National Archives and Records Administration, Census Bureau
- **Commercial(8):** Finance in Cloud, Cloud Backup, Mendeley (Citations), Netflix, Web Search, Digital Materials, Cargo shipping (as in UPS)
- **Defense(3):** Sensors, Image surveillance, Situation Assessment
- **Healthcare and Life Sciences(10):** Medical records, Graph and Probabilistic analysis, Pathology, Bioimaging, Genomics, Epidemiology, People Activity models, Biodiversity
- **Deep Learning and Social Media(6):** Driving Car, Geolocate images/cameras, Twitter, Crowd Sourcing, Network Science, NIST benchmark datasets
- **The Ecosystem for Research(4):** Metadata, Collaboration, Translation, Light source data
  Astronomy and Physics(5): Sky Surveys including comparison to simulation, Large Hadron Collider at CERN, Belle II Accelerator in Japan
- **Earth, Environmental and Polar Science(10):** Radar Scattering in Atmosphere, Earthquake, Ocean, Earth Observation, Ice sheet Radar scattering, Earth radar mapping, Climate simulation datasets, Atmospheric turbulence identification, Subsurface Biogeochemistry (microbes to watersheds), AmeriFlux and FLUXNET gas sensors
- **Energy(1):** Smart grid

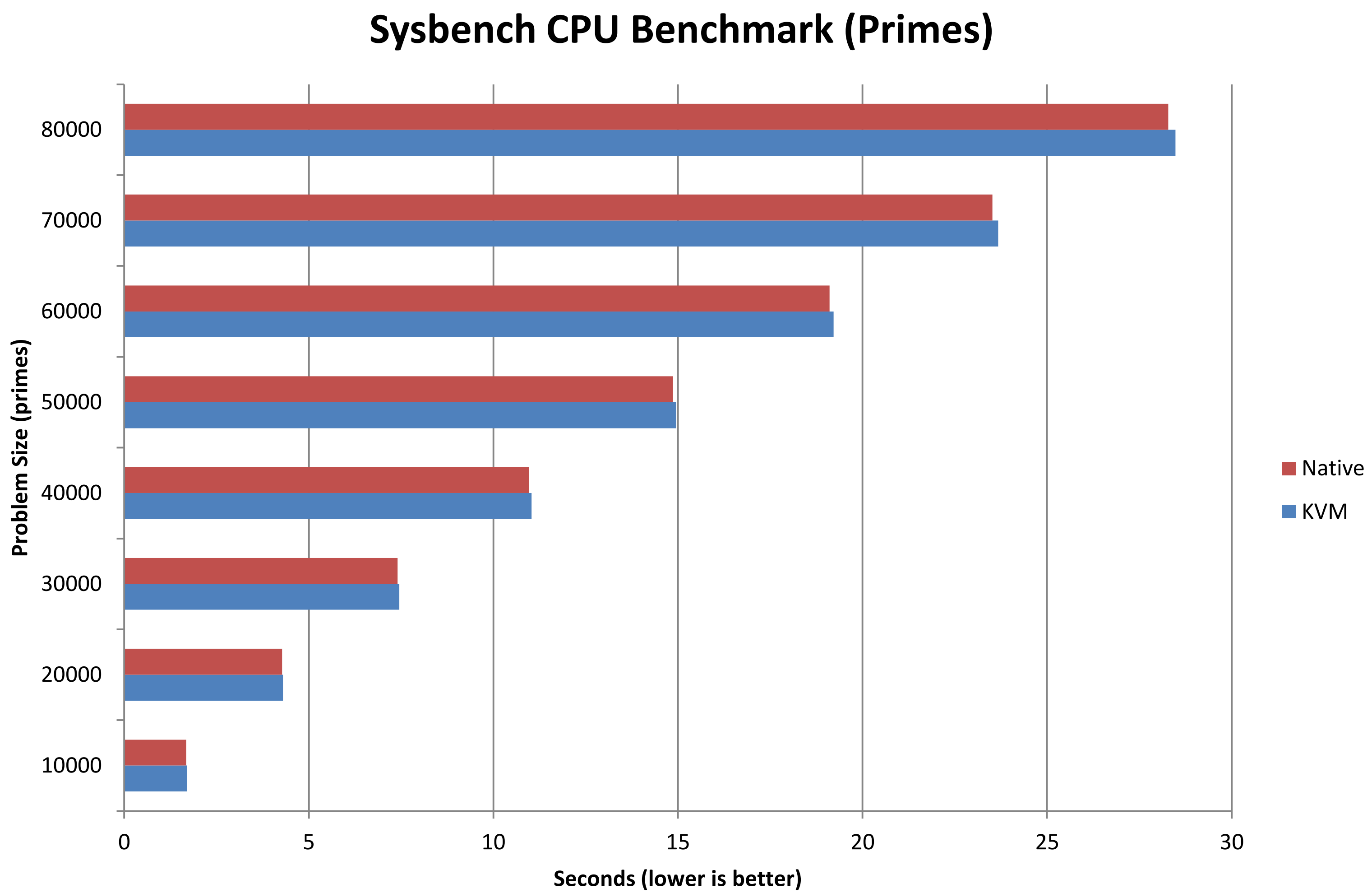
Largest open collection of Big Data Requirements?
Analyzed for common characteristics

Computer Cloud Assumptions I

- Clouds will continue to grow in importance
- Clouds consists of an “infinite” number of compute/storage/network nodes available on demand
- Clouds can be public and/or private with similar architectures (but different security issues)
- Clouds have some overheads but these are decreasing using SR-IOV and better hypervisors
- Clouds are getting more powerful with better networks but
  - Exascale Supercomputer will not be a cloud although most other systems will be!
- Performance of clouds can be (easily) understood using standard (parallel computing) methods
- Streaming and Internet of Things applications (80% NIST use cases) particularly well suited to clouds
- Can deploy “arbitrary parallel resources” to address DDDAS/IoT

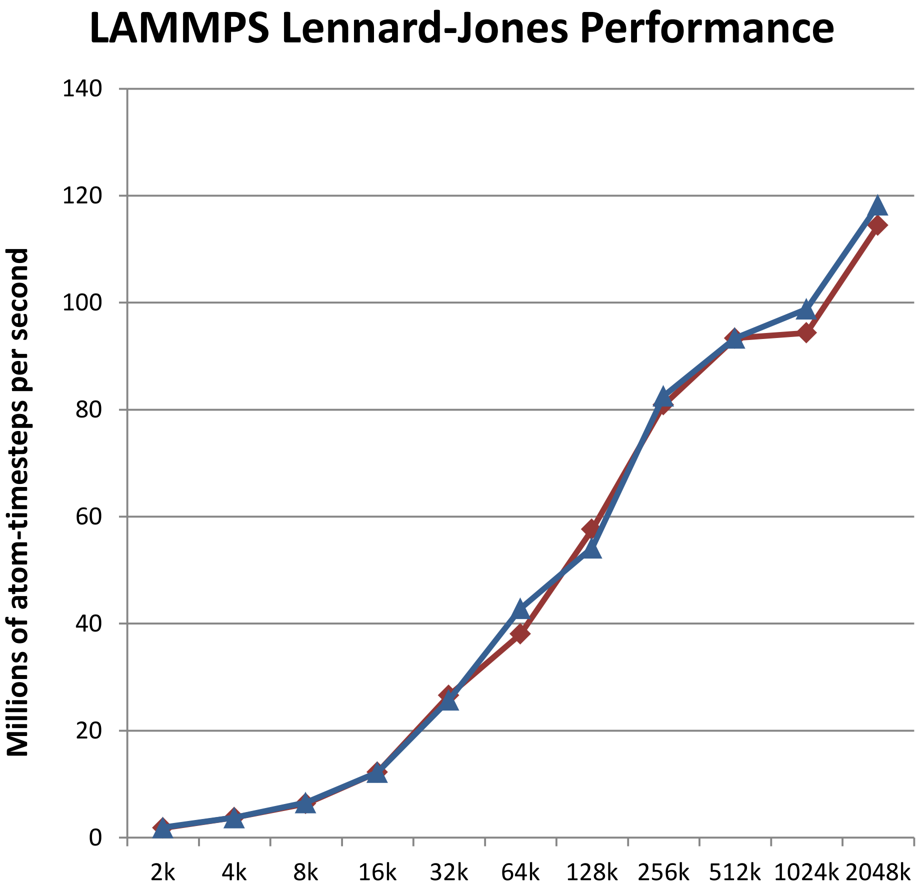
Computer Cloud Assumptions II

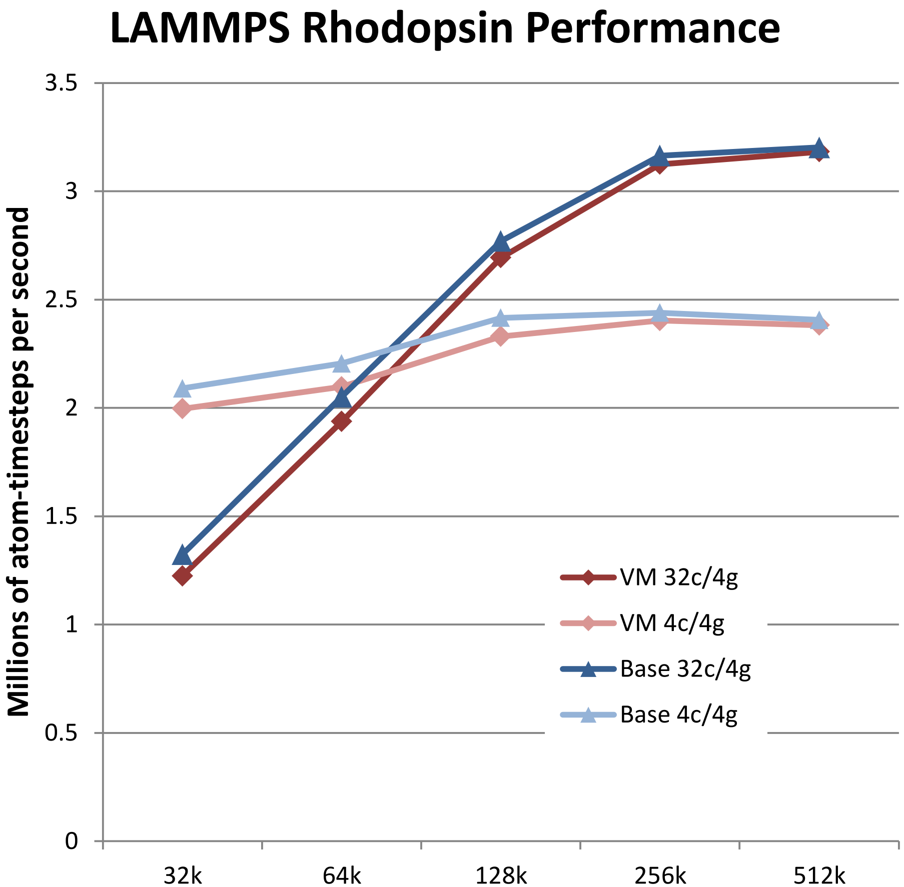
- Big data revolution built around cloud processing
- Incredibly powerful software ecosystem (the “Apache Big Data Stack” or ABDS) emerged to support Big Data
- Much of this software is open-source and at all points in stack at least one good open source choice
- DevOps (Chef, Cobbler ..) deploys dynamic virtual clusters
- Research (science and engineering) similar big data needs to commercial but less search and recommender engines
  - Both have large pleasingly parallel component (50%)
  - Less classic MapReduce and more iterative algorithms
- Streaming (DDDAS) dominant (80%) and similar needs in research and commercial
- HPC-ABDS links classic parallel computing and ABDS and runs on clouds or HPC systems



- Overhead of KVM between 0.8% and 0.5% compared to native Bare-metal

SR-IOV Enhanced Chemistry on Clouds





- SR-IOV is single root I/O virtualization and cuts through virtualization overhead
- VMs running LAMMPs achieve near-native performance at 32 cores & 4GPUs
  - 96.7% efficiency for LJ
  - 99.3% efficiency for Rhodo

There are a lot of Big Data and HPC Software systems
Challenge! Manage environment offering these different components

Kaleidoscope of (Apache) Big Data Stack (ABDS) and HPC Technologies December 2 2014

Cross-Cutting Functions

1) Message and Data Protocols: Avro, Thrift, Protobuf

2) Distributed Coordination: Zookeeper, Giraffe, JGroups

3) Security & Privacy: InCommon, OpenStack Keystone, LDAP, Sentry, Sqrrl

4) Monitoring: Ambari, Ganglia, Nagios, Inca

17 layers >266 Software Packages

17) Workflow-Orchestration: Oozie, ODE, ActiveBPEL, Airavata, OODT (Tools), Pegasus, Kepler, Swift, Taverna, Triana, Trident, BioKepler, Galaxy, IPython, Dryad, Naiad, Tez, Google FlumeJava, Crunch, Cascading, Scalding, e-Science Central, Azure Data Factory

16) Application and Analytics: Mahout , MLlib , MLbase, DataFu, mlpy, scikit-learn, CompLearn, Caffe, R, Bioconductor, ImageJ, pbdR, Scalapack, PetSc, Azure Machine Learning, Google Prediction API, Google Translation API, Torch, Theano, H_2O, Google Fusion Tables, Oracle PGX, GraphLab, GraphX, CINET, Elasticsearch, IBM System G, IBM Watson

15A) High level Programming: Kite, Hive, HCatalog, Databee, Tajo, Pig, Phoenix, Shark, MRQL, Impala, Presto, Sawzall, Drill, Google BigQuery (Dremel), Google Cloud DataFlow, Summingbird, SAP HANA, IBM META

15B) Frameworks: Google App Engine, AppScale, Red Hat OpenShift, Heroku, AWS Elastic Beanstalk, IBM BlueMix, Ninefold, Aerobatic, Azure, Jelastic, Cloud Foundry, CloudBees, Engine Yard, CloudControl, appfog, dotCloud, Pivotal

14A) Basic Programming model and runtime, SPMD, MapReduce: Hadoop, Spark, Twister, Stratosphere (Apache Flink), Reef, Hama, Giraph, Pregel, Pegasus

14B) Streams: Storm, S4, Samza, Google MillWheel, Amazon Kinesis, LinkedIn Databus, Facebook Scribe/ODS, Azure Stream Analytics

13) Inter process communication Collectives, point-to-point, publish-subscribe: Harp, MPI, Netty, ZeroMQ, ActiveMQ, RabbitMQ, QPid, Kafka, Kestrel, JMS, AMQP, Stomp, MQTT, Azure Event Hubs

Public Cloud: Amazon SNS, Google Pub Sub, Azure Queues

12) In-memory databases/caches: Gora (general object from NoSQL), Memcached, Redis (key value), Hazelcast, Ehcache, Infinispan

12) Object-relational mapping: Hibernate, OpenJPA, EclipseLink, DataNucleus, ODBC/JDBC

12) Extraction Tools: UIMA, Tika

11C) SQL: Oracle, DB2, SQL Server, SQLite, MySQL, PostgreSQL, SciDB, Apache Derby, Google Cloud SQL, Azure SQL, Amazon RDS, rasdaman

11B) NoSQL: HBase, Accumulo, Cassandra, Solandra, MongoDB, CouchDB, Lucene, Solr, Berkeley DB, Riak, Voldemort, Neo4J, Yarcdata, Jena, Sesame, AllegroGraph, RYA, Espresso, Sqrrl, Facebook Tao

Public Cloud: Azure Table, Amazon Dynamo, Google DataStore

11A) File management: iRODS, NetCDF, CDF, HDF, OPeNDAP, FITS, RCFile, ORC, Parquet

10) Data Transport: BitTorrent, HTTP, FTP, SSH, Globus Online (GridFTP), Flume, Sqoop

9) Cluster Resource Management: Mesos, Yarn, Helix, Llama, Celery, HTCondor, SGE, OpenPBS, Moab, Slurm, Torque, Google Omega, Facebook Corona

8) File systems: HDFS, Swift, Cinder, Ceph, FUSE, Gluster, Lustre, GPFS, GFFS, Haystack, f4

Public Cloud: Amazon S3, Azure Blob, Google Cloud Storage

7) Interoperability: Whirr, JClouds, OCCI, CDMI, Libcloud, TOSCA, Libvirt

6) DevOps: Docker, Puppet, Chef, Ansible, Boto, Cobbler, Xcat, Razor, CloudMesh, Heat, Juju, Foreman, Rocks, Cisco Intelligent Automation for Cloud, Ubuntu MaaS, Facebook Tupperware, AWS OpsWorks, OpenStack Ironic

5) IaaS Management from HPC to hypervisors: Xen, KVM, Hyper-V, VirtualBox, OpenVZ, LXC, Linux-Vserver, VMware ESXi, vSphere, OpenStack, OpenNebula, Eucalyptus, Nimbus, CloudStack, VMware vCloud, Amazon, Azure, Google and other public Clouds,

Networking: Google Cloud DNS, Amazon Route 53

Maybe a Big Data Initiative would include

- We don’t need 266 software packages so can choose e.g.
- **Workflow:** IPython, Pegasus or Kepler (replaced by tools like Crunch, Tez?)
- **Data Analytics:** Mahout, R, ImageJ, Scalapack
- **High level Programming:** Hive, Pig
- **Parallel Programming model:** Hadoop, Spark, Giraph (Twister4Azure, Harp), MPI;
- **Streaming:** Storm, Kapfka or RabbitMQ (Sensors)
- **In-memory:** Memcached
- **Data Management:** Hbase, MongoDB, MySQL or Derby
- **Distributed Coordination:** Zookeeper
- **Cluster Management:** Yarn, Slurm
- **File Systems:** HDFS, Lustre
- **DevOps:** Cloudmesh, Chef, Puppet, Docker, Cobbler
- **IaaS:** Amazon, Azure, OpenStack, Libcloud
- **Monitoring:** Inca, Ganglia, Nagios

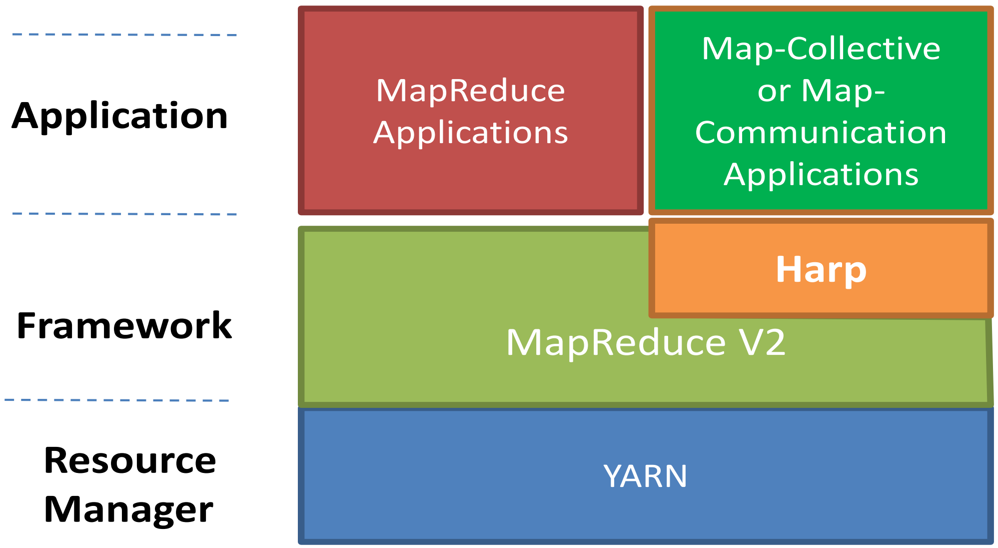
	ABDS Big Data	HPC
Orchestration	Crunch, Tez, Cloud Dataflow	Kepler, Pegasus
Libraries	Mllib/Mahout, R, Python	Matlab, Eclipse, Apps
High Level Programming	Pig, Hive, Drill	Domain-specific Languages
Platform as a Service	App Engine, BlueMix, Elastic Beanstalk	XSEDE Software Stack
Languages	Java, Erlang, SQL, SparQL	Fortran, C/C++
Streaming	Storm, Kafka, Kinesis	
Parallel Runtime	MapReduce	MPI/OpenMP/OpenCL
Coordination	Zookeeper	
Caching	Memcached	
Data Management	Hbase, Neo4J, MySQL	iRODS
Data Transfer	Sqoop	GridFTP
Scheduling	Yarn	Slurm
File Systems	HDFS, Object Stores	Lustre
Formats	Thrift, Protobuf	FITS, HDF
Infrastructure	CLOUDS	SUPERCOMPUTERS

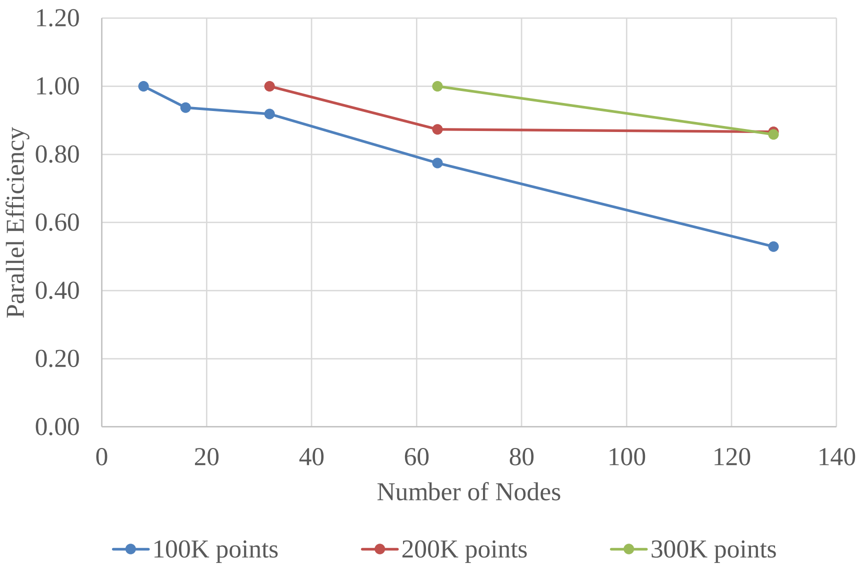
Integrated Software Ecosystem

Using ABDS gives sustainable software
Deploy with Python+Chef **Cloudmesh** DevOps on public/private cloud, container or bare-metal as Software defined system (virtual cluster)

Harp Plug-in to Hadoop

Make ABDS high performance – do not replace it!



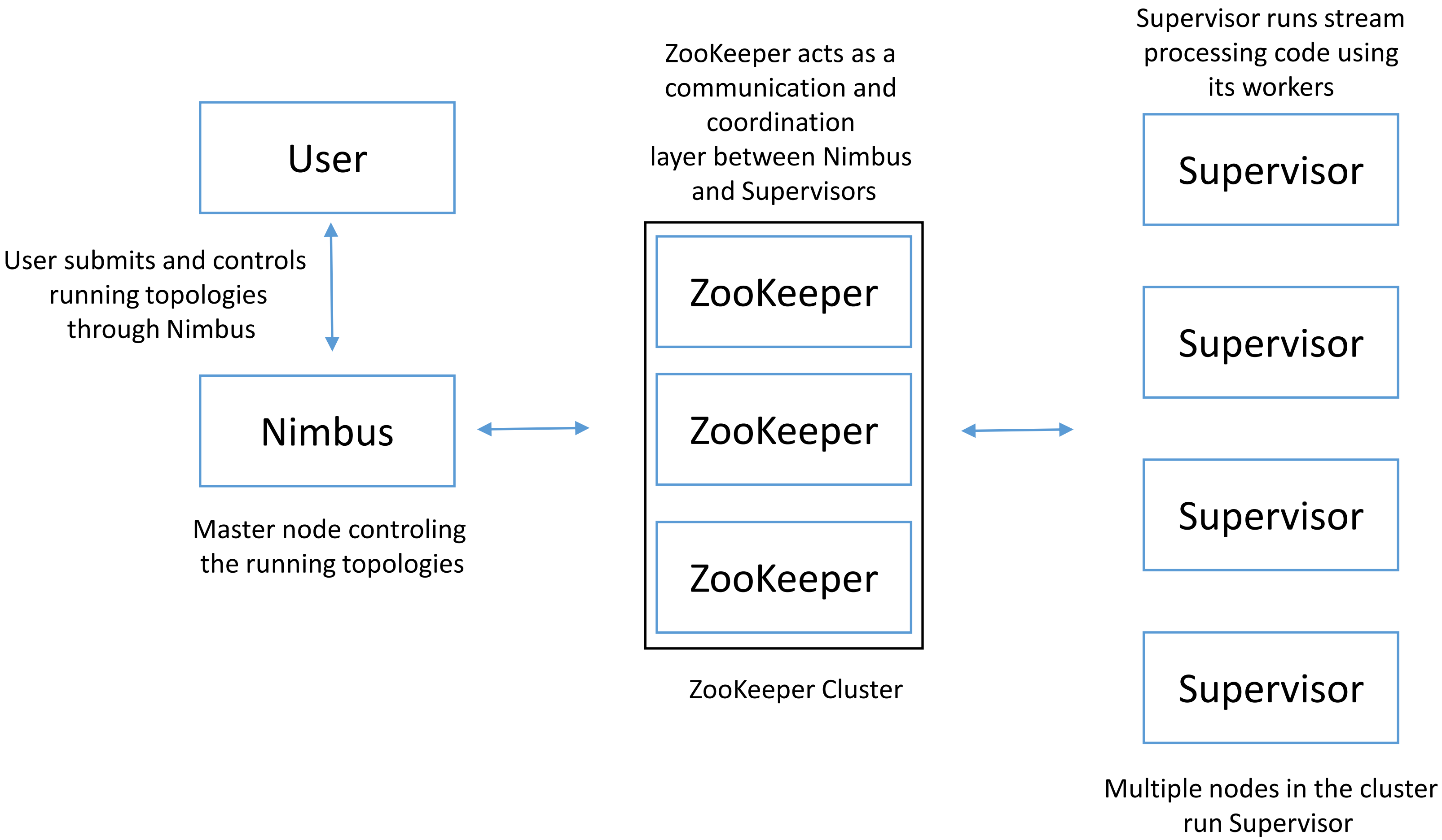


- Work of Judy Qiu and Bingjing Zhang.
- Left diagram shows architecture of Harp Hadoop Plug-in that adds high performance communication, Iteration (caching) and support for rich data abstractions including key-value
- Alternative to Spark, Giraph, Flink, Reef, Hama etc.
- Right side shows efficiency for 16 to 128 nodes (each 32 cores) on WDA-SMACOF dimension reduction dominated by conjugate gradient
- WDA-SMACOF is general purpose dimension reduction

Open Source Apache Storm

- http://storm.apache.org/ Written in Clojure (Lisp to Java) & Java
- Apache Storm is a distributed real time (DDDAS/streaming) computation framework for processing streaming data exposing a **coarse grain dataflow model**
- Storm is being used to do real time analytics, online machine learning (Spark, Hadoop …..), distributed RPC etc.
- Provides scalable, fault tolerant and guaranteed message processing.
- Trident is a high level API on top of Storm which provides functions like stream joins, groupings, filters etc. Also Trident has exactly-once processing guarantees.
- The project was originally developed at Twitter for processing Tweets from users and was donated to ASF in 2013.
  - S4 (Yahoo) and Samza (LinkedIn) are also Apache Streaming systems
  - Google MillWheel, Amazon Kinesis, Azure Stream Analytics
- Storm has being used in major deployments in Twitter, Yahoo, Alibaba, Groupon etc.

Apache Storm - Architecture



Storm uses Tuples & Streams

Tuple

“user”, 1000, Point Object

Tuple is a ordered list of elements and
storm should know how to serialize each element

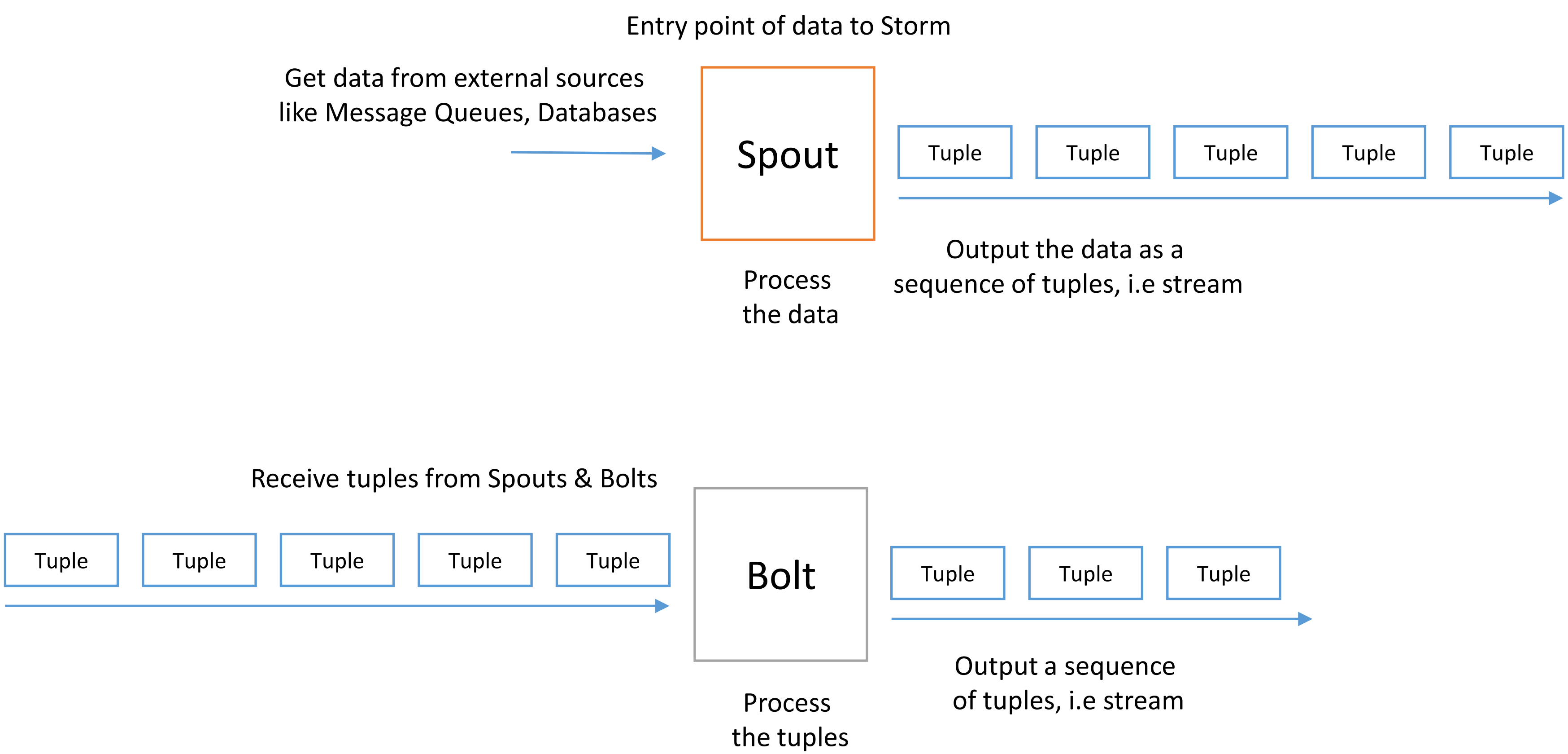
Stream

Tuple	Tuple	Tuple	Tuple	Tuple

Stream is an unbounded sequence of tuples

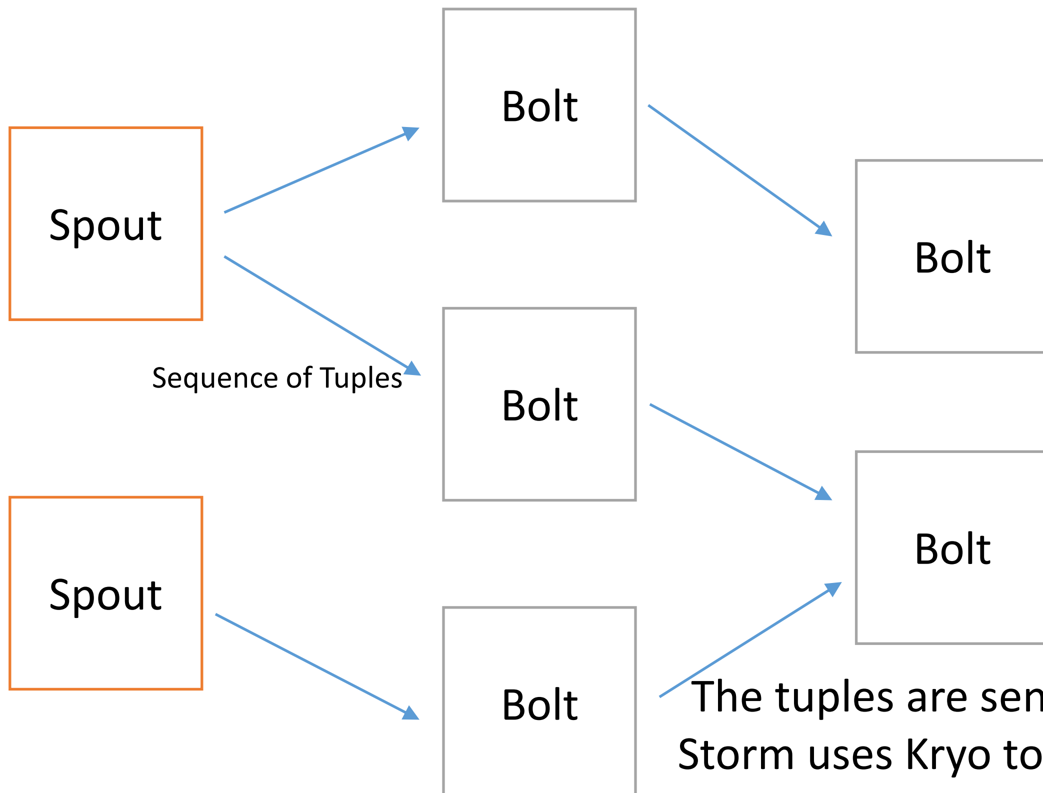
Storm – Spouts & Bolts

Spouts & Bolts are written by user



Storm Dataflow Topology

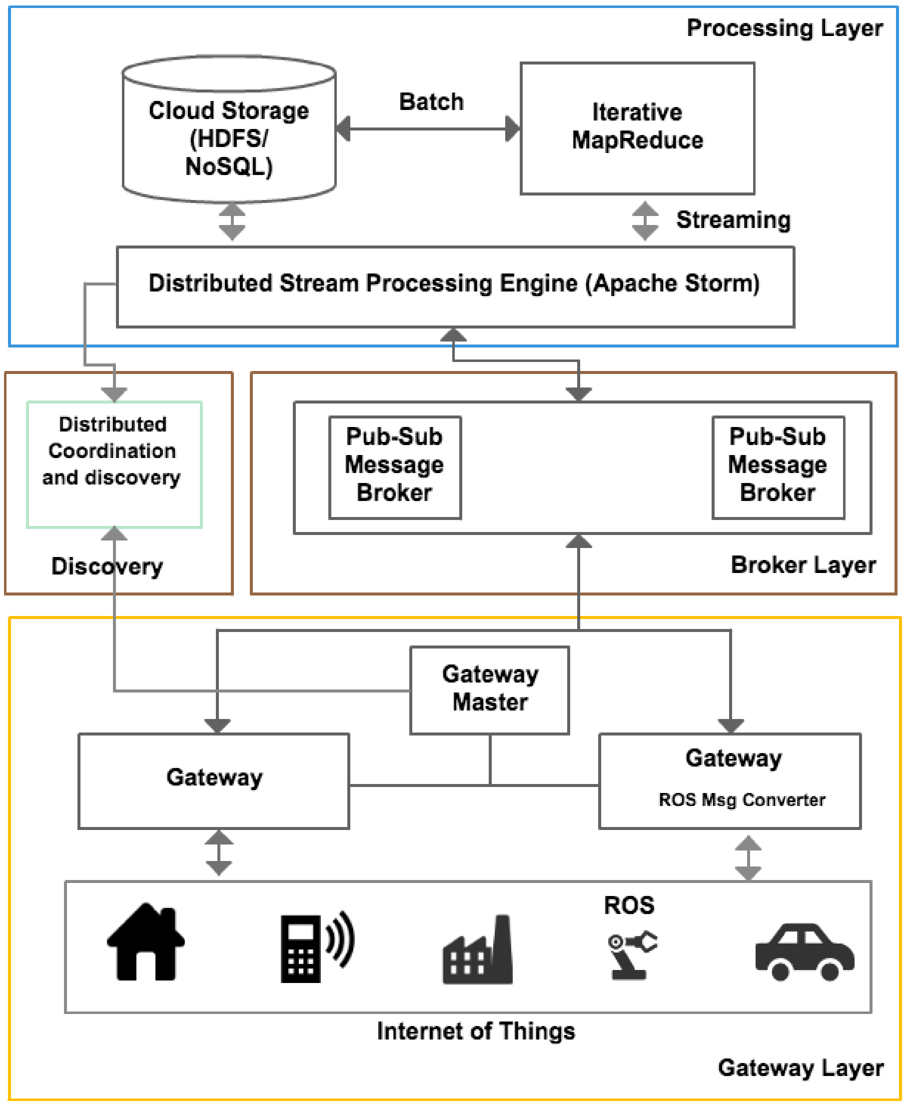
A user defined arrangement of Spouts and Bolts



The tuples are sent using messaging, Storm uses Kryo to serialize the tuples and Netty to transfer the messages

The topology defines how the bolts receive its messages using Stream Grouping

Scalable DDAS IoTCloud Architecture

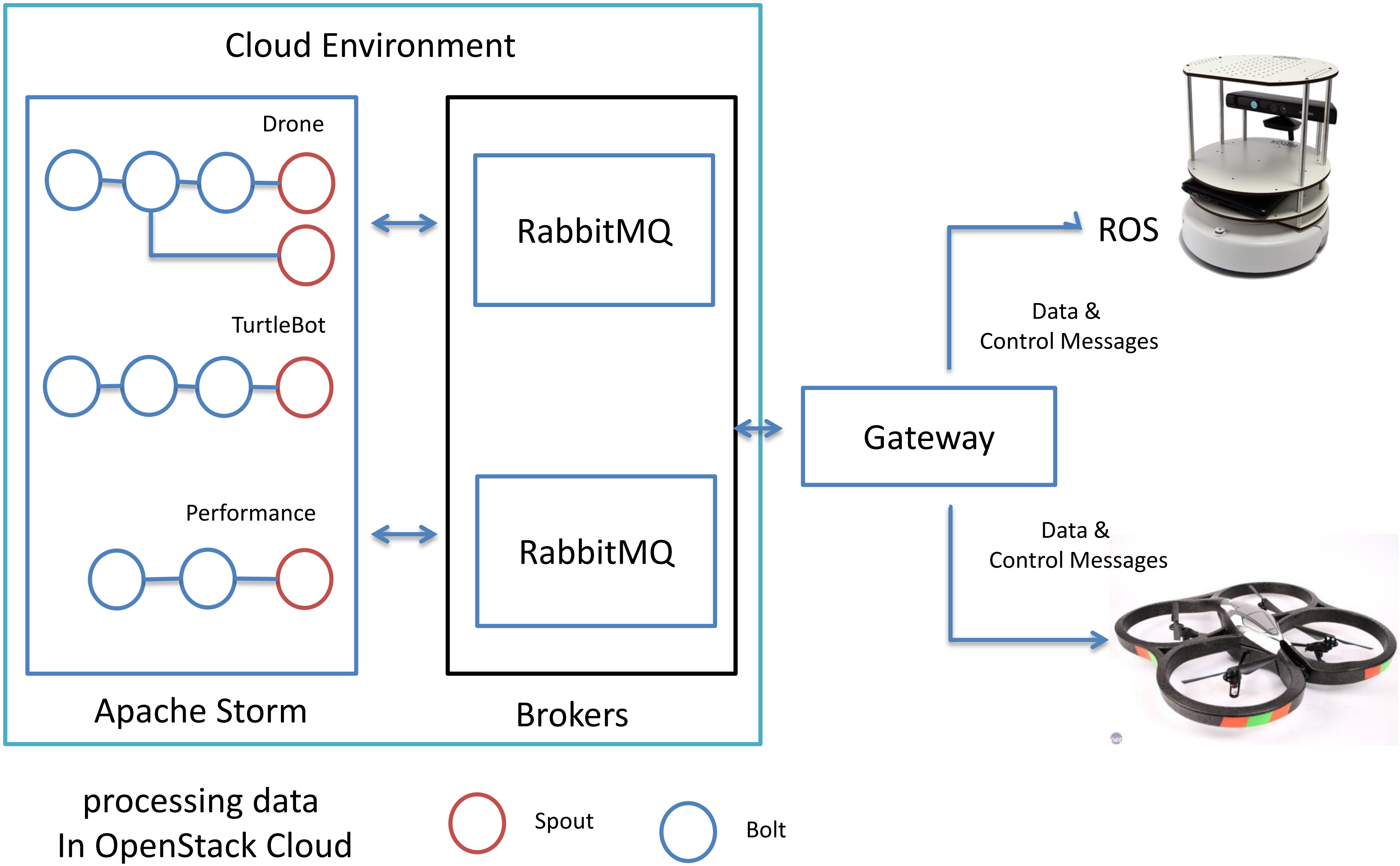


Basic Pub-Sub IoT system developed under AFRL WPAFB funding

This year extended to use Apache Storm and test

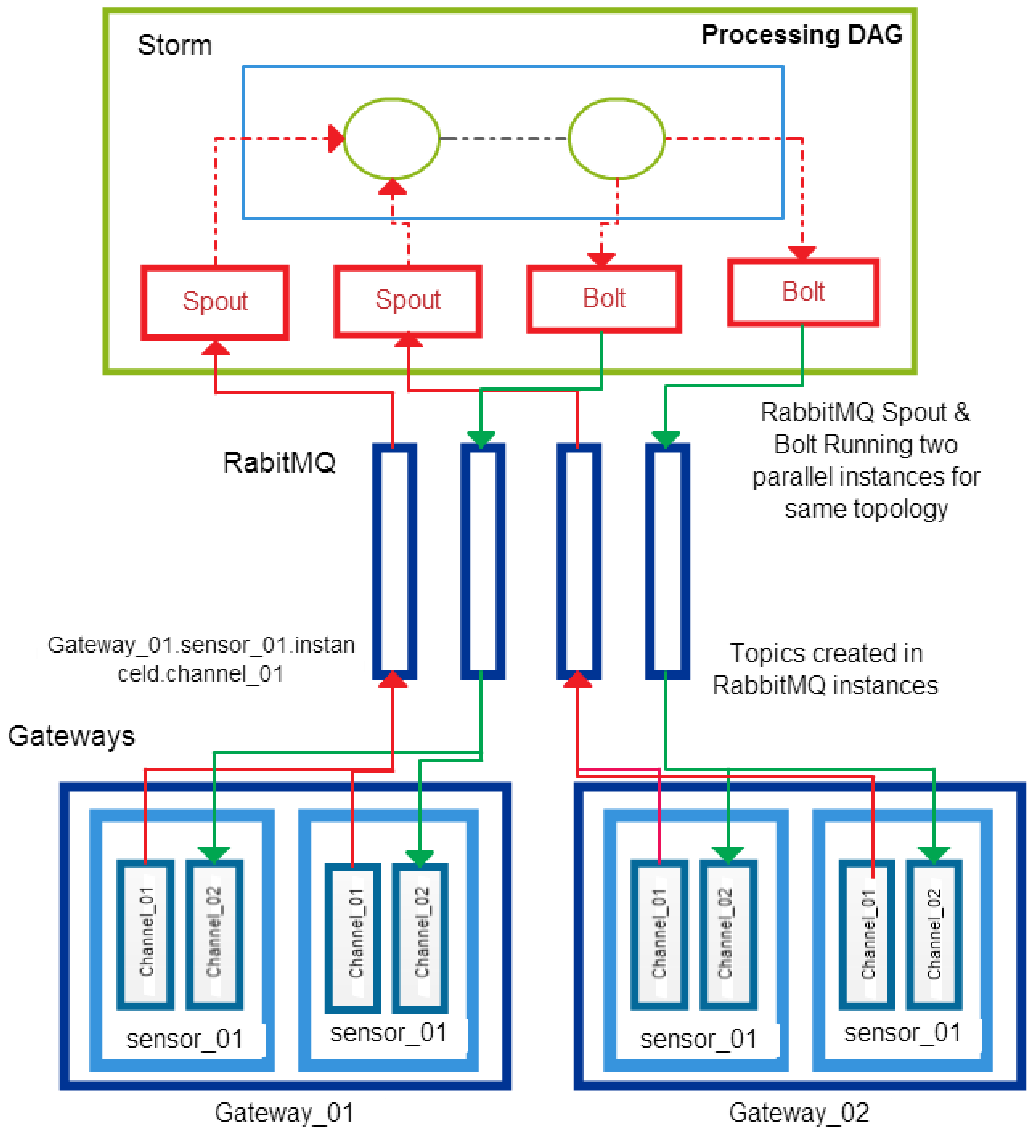
https://github.com/iotcloud
http://iotcloud.github.io/source.html

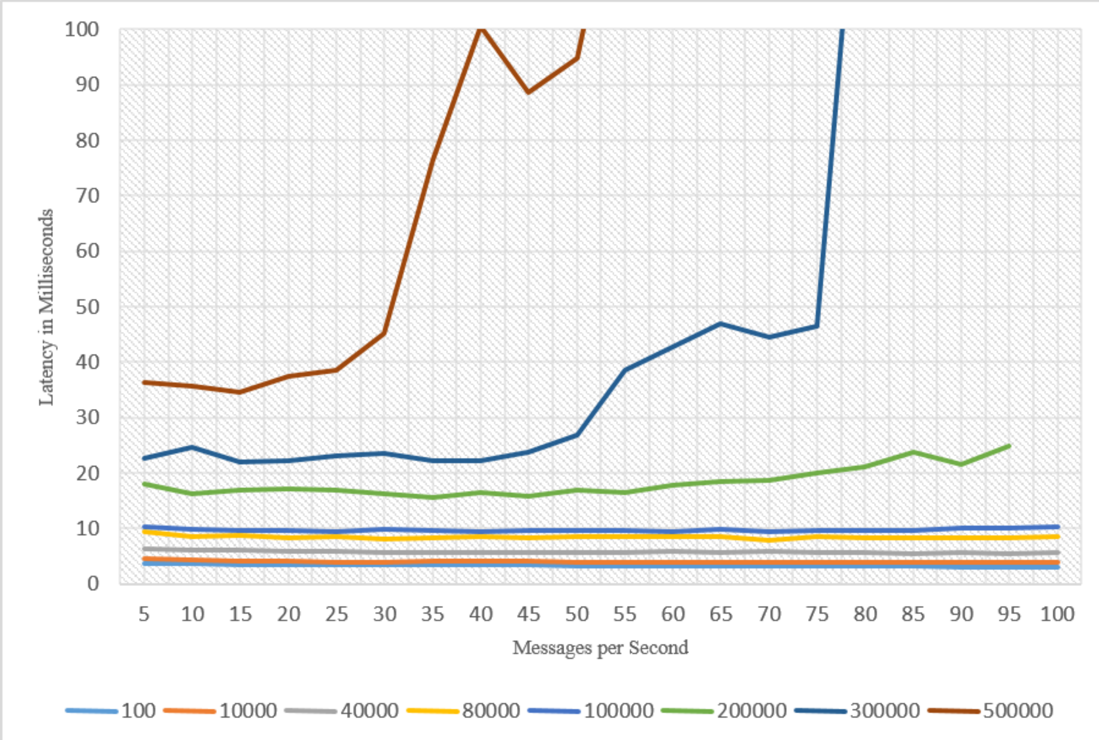
Deployment



processing data
In OpenStack Cloud

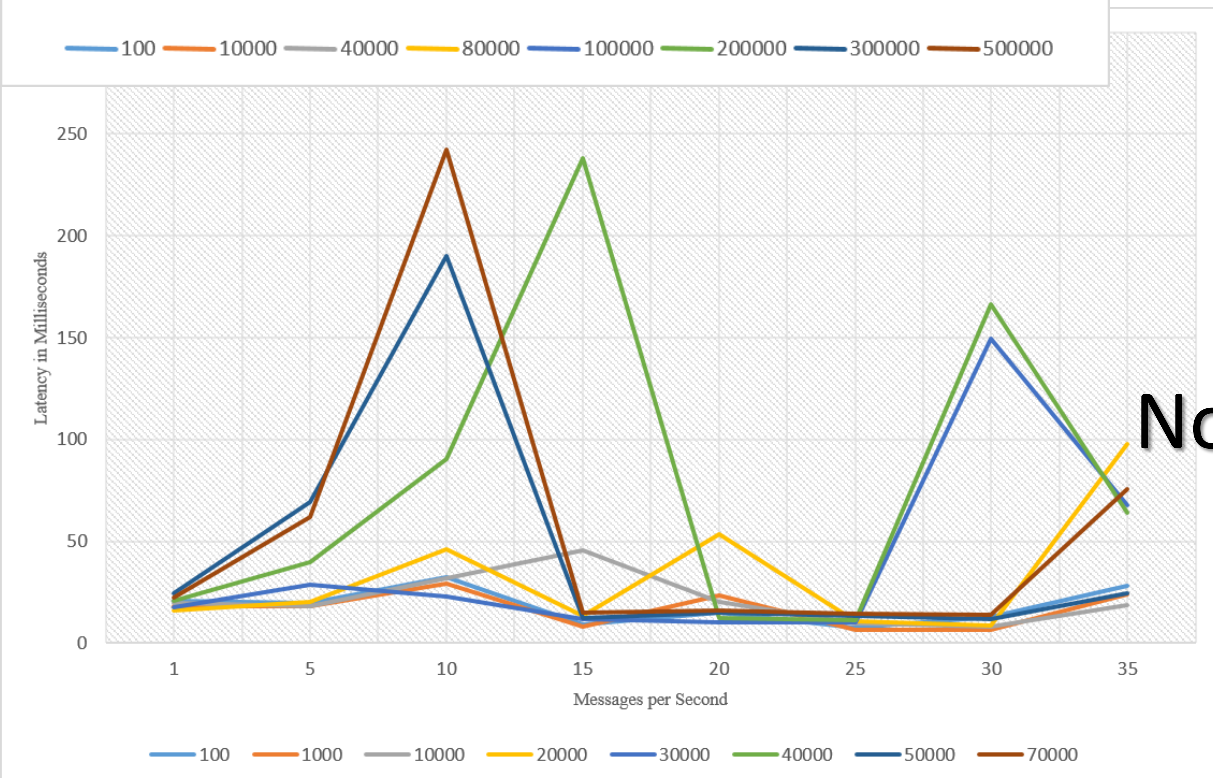
Set up with RabbitMQ – Shared Channels





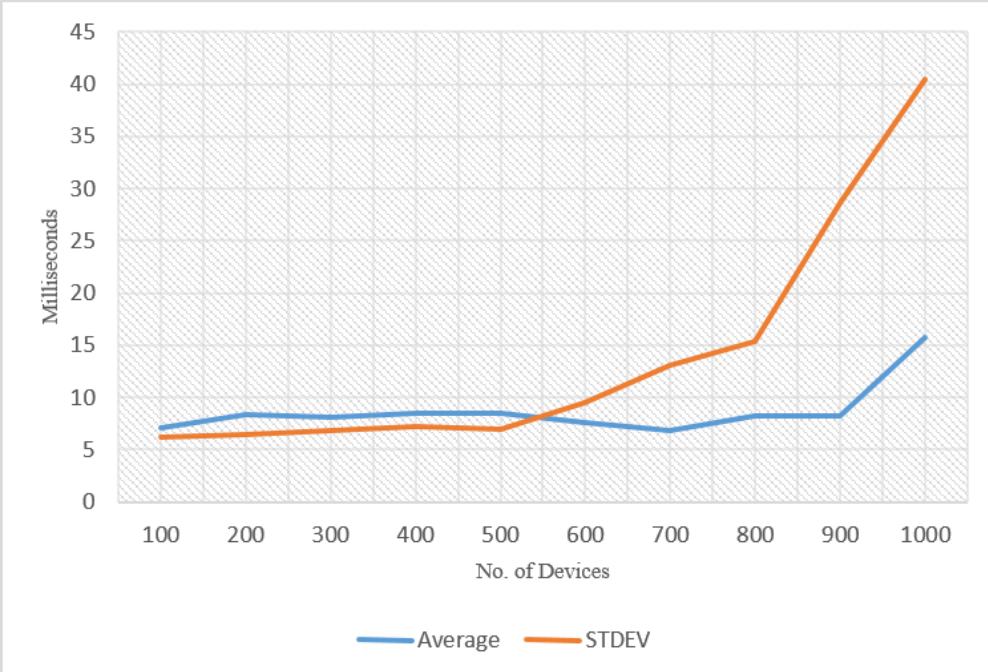
Latency with RabbitMQ

Message sizes in bytes

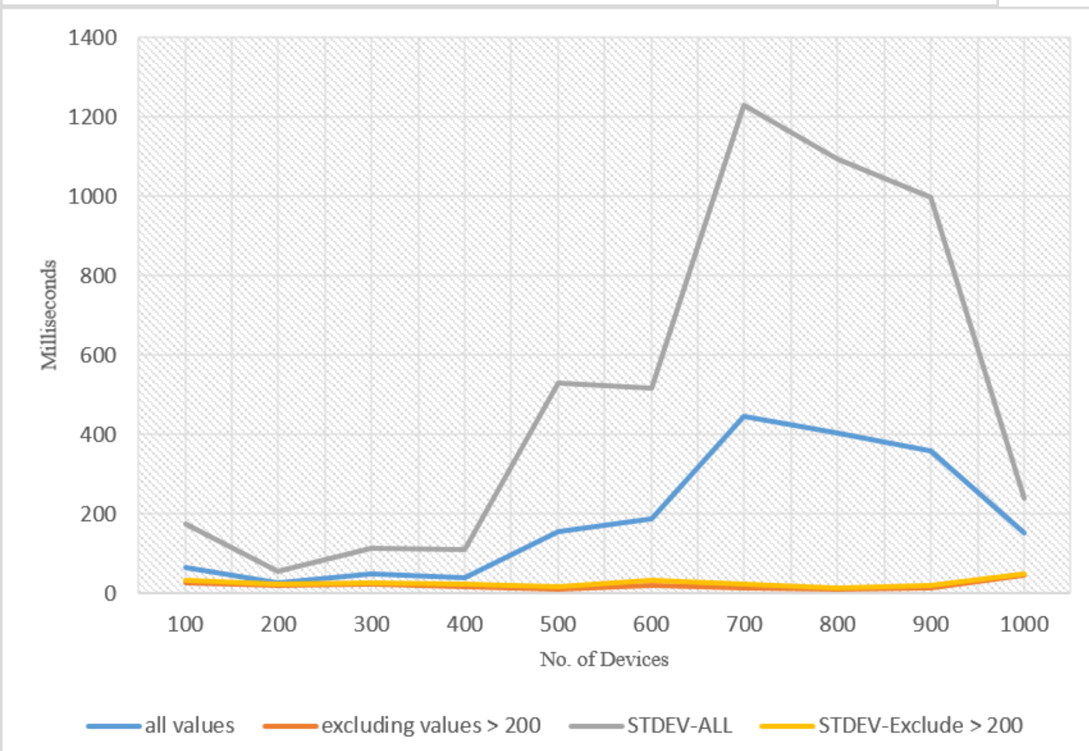


Latency with Kafka

Note change in scales



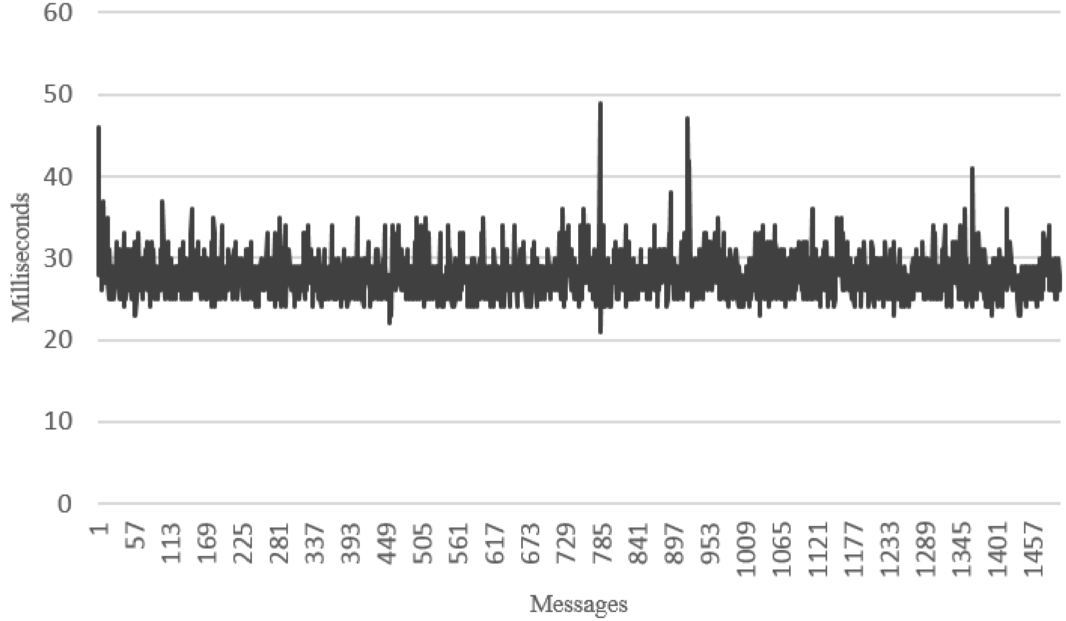
Varying number of Devices-RabbitMQ



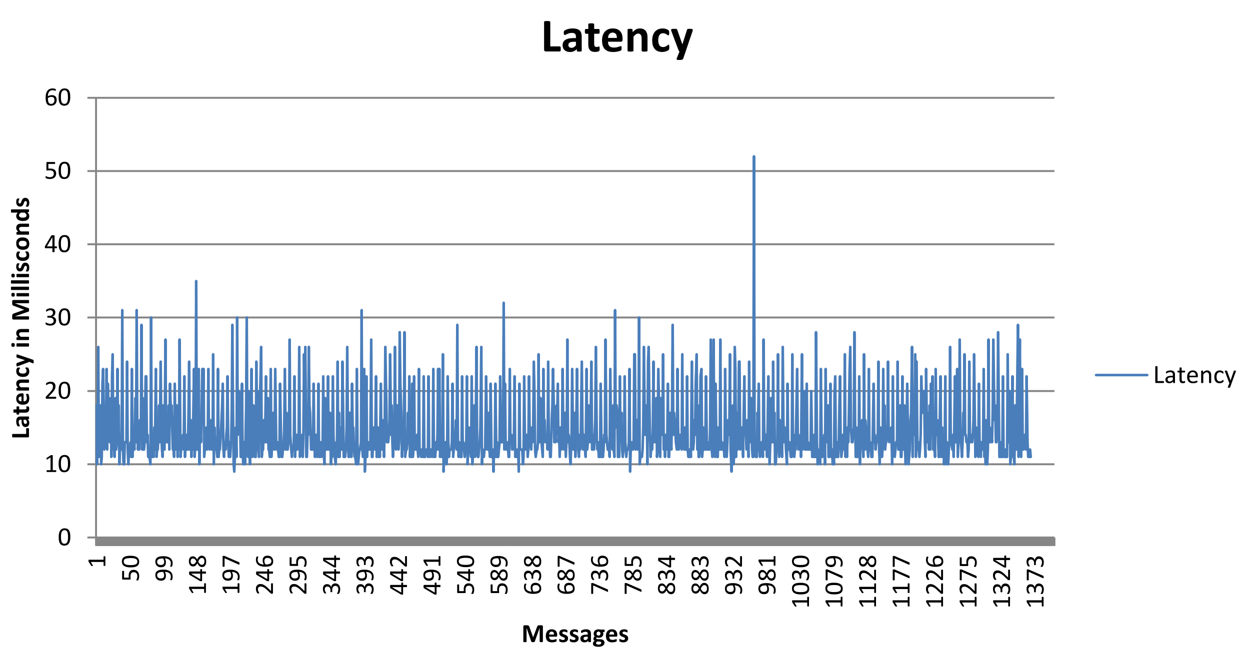
Varying number of Devices – Kafka

Robot Latency RabbitMQ

Turtlebot

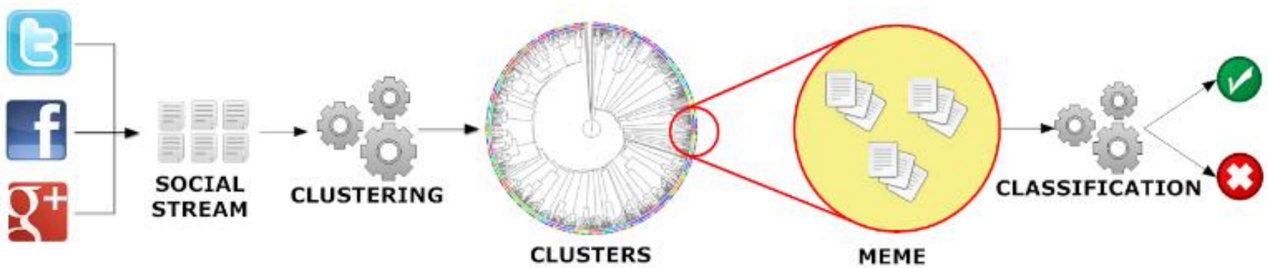


Drone

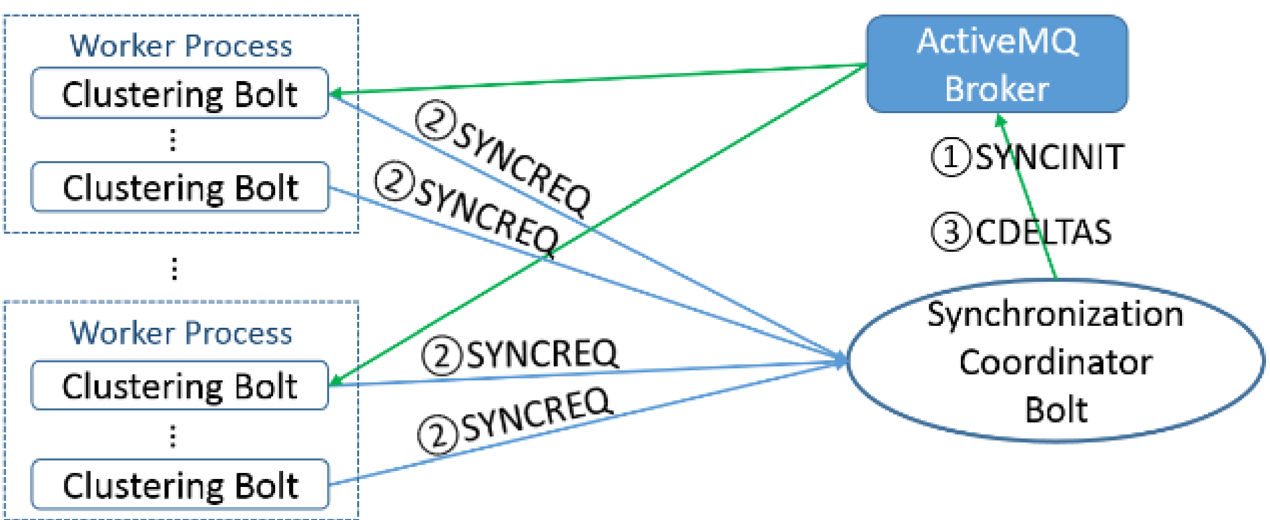


Parallel Tweet Clustering with Storm

- Judy Qiu and Xiaoming Gao
- Storm Bolts coordinated by ActiveMQ to synchronize parallel cluster center updates – add loops to Storm
- 2 million streaming tweets processed in 40 minutes; 35,000 clusters



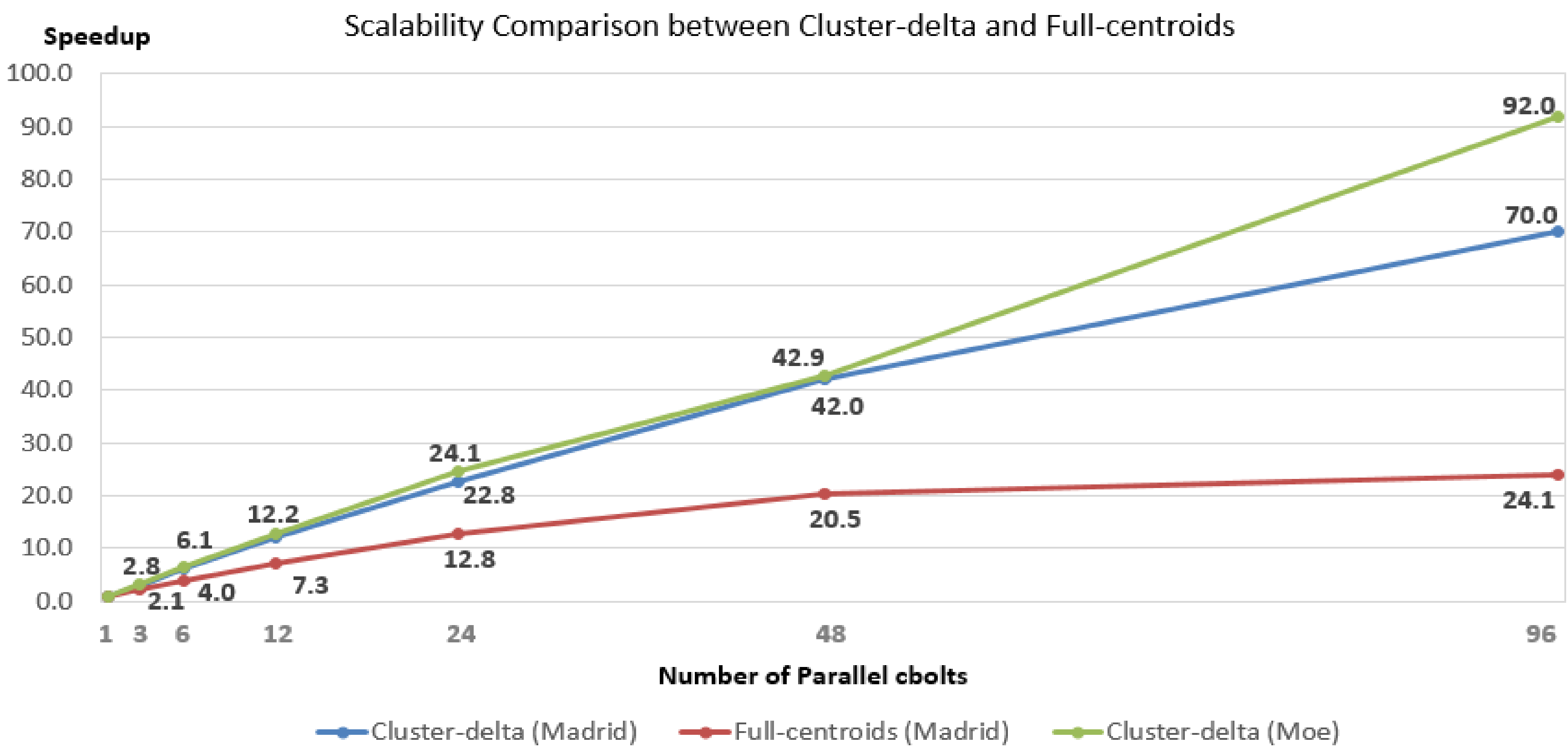
Sequential



Parallel – eventually 10,000 bolts

Parallel Tweet Clustering with Storm

- Speedup on up to 96 bolts on two clusters Moe and Madrid
- Red curve is old algorithm;
- green and blue new algorithm
- Full Twitter – 1000 way parallelism
- Full Everything – 10,000 way parallelism

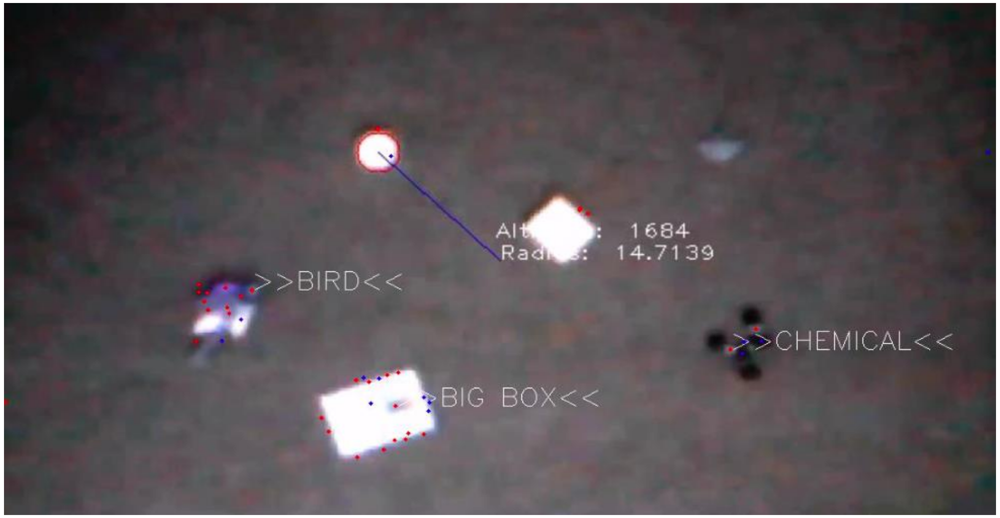


Proof-of-concept robot scenario

- Simulated search-and-rescue/-destroy
  - Aerial drone (Parrot AR.Drone 2) recognizes and tracks a moving ground-based robot (Sphero), while also recognizing other important objects and distractors
  - Drone needs to map unknown environment (Simultaneous Localization and Mapping)

Visual recognition and tracking

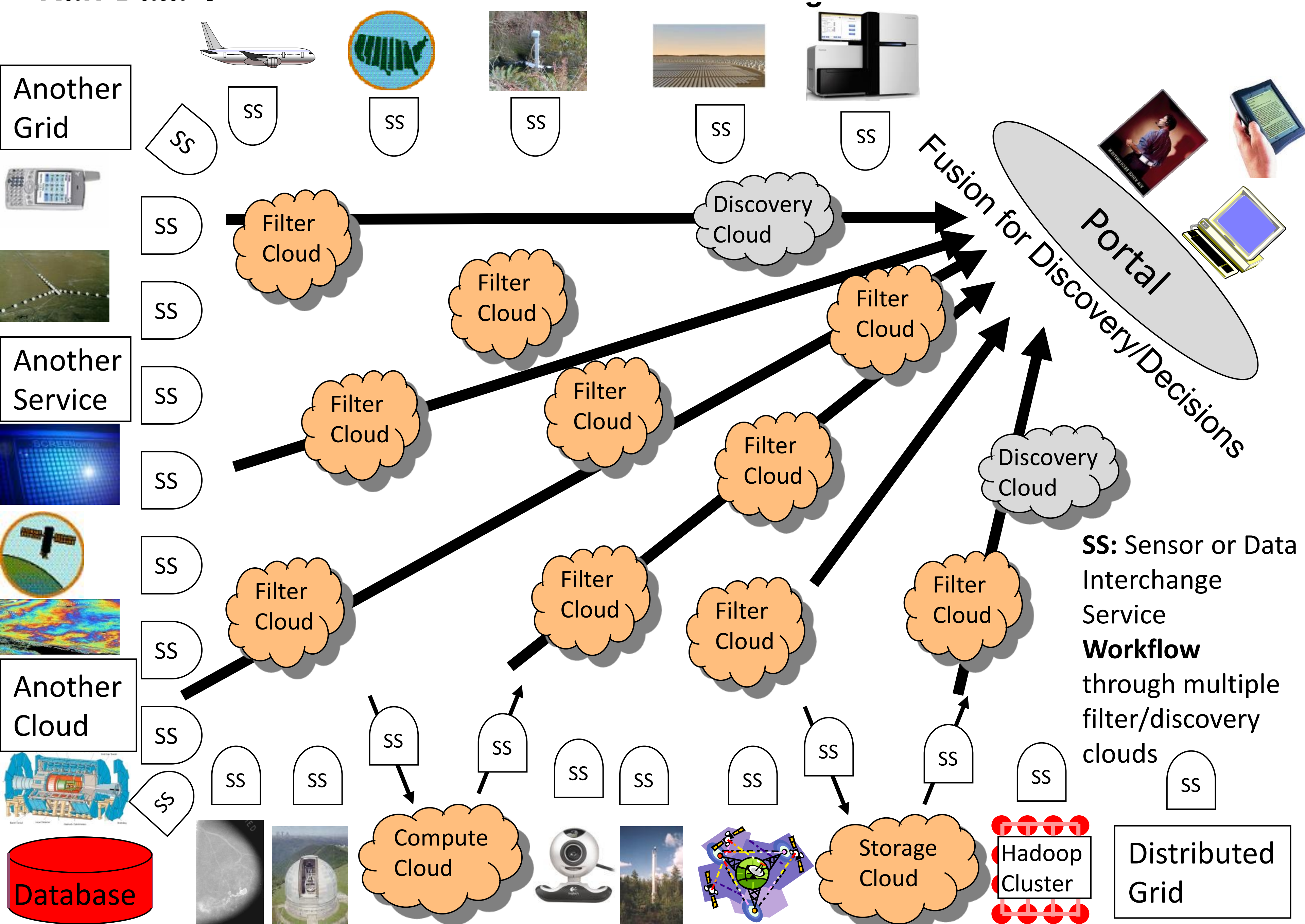
- Currently use shape, color, and structural features to recognize and track objects
  - Fast Sphero recognition/localization using generalized Hough transform and color matching, track with Kalman filters
  - Recognize other objects by matching **SURF** (Sped-Up Robust Features) and **SIFT** (Scale Invariant Feature Transform) points to a set of object models



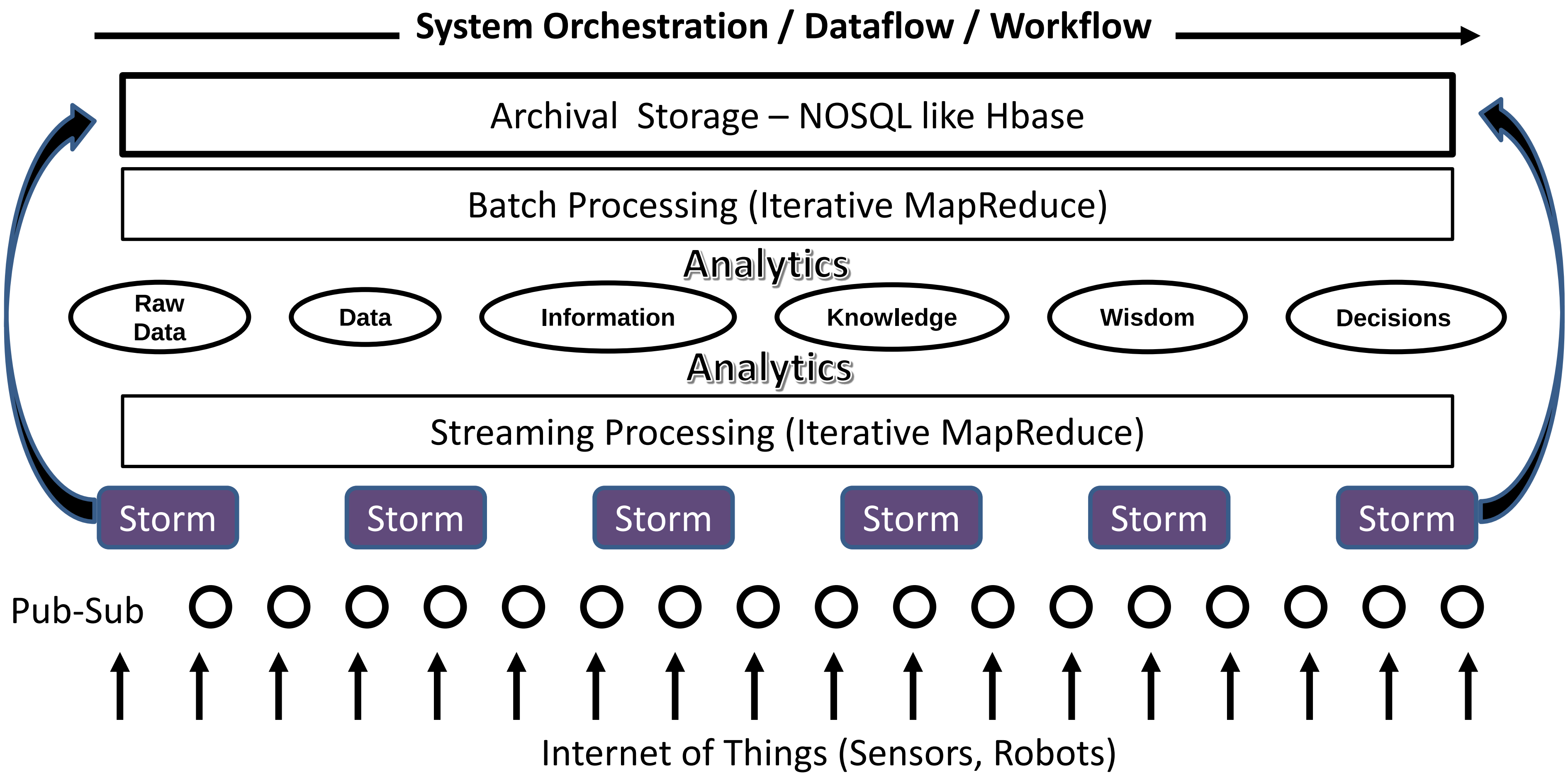
Role of cloud

- Cloud has latency but “unlimited” resources on demand for planning
- Drone visual processing done on **OpenStack** cloud (FutureSystems)
  - Raw images transmitted to cloud
  - Sphero recognition and tracking will be parallelized
  - Other object recognition pleasingly parallel (multiple object models can be considered in parallel)
- Turtlebot **SLAM** (Simultaneous Localization and Mapping) and **ORCA** (Optimal Reciprocal Collision Avoidance) algorithms will be supported in parallel
- Compare Storm and Harp (Hadoop with iterative high performance) parallel versions
- Local v. Cloud computing
- Security issues

Data Deluge is also Information/Knowledge/Wisdom/Decision Deluge?



Cloud DIKW based on HPC-ABDS to integrate streaming and batch Big Data



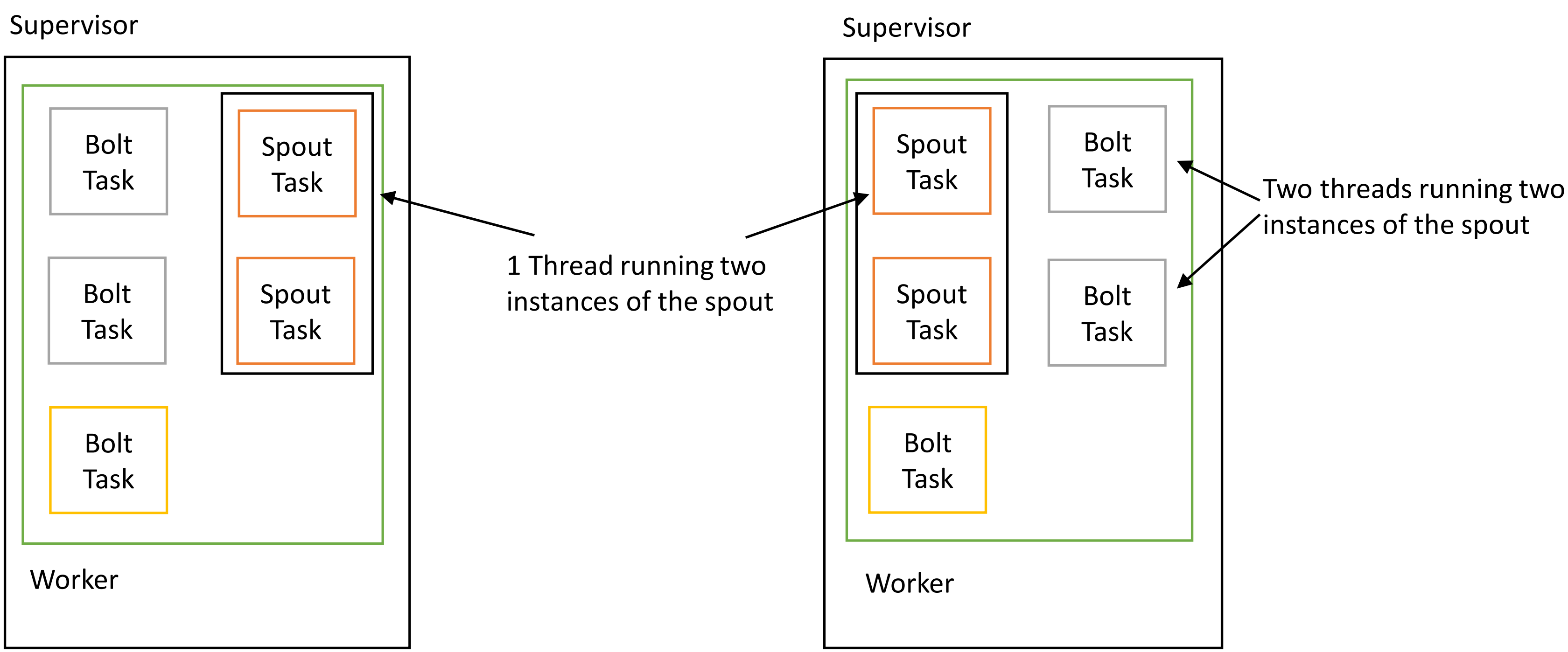
Conclusions

- All software as open source HPC enhancements to open source HPC-ABDS; sustainable!
- Reworked **DDDAS IoTCloud** using Apache **Storm** getting good results with **RabbitMQ** or ActiveMQ.
  - "official Kafka pub-sub system too heavy weight"
  - Latencies <~10ms upto 100K byte messages
- Batch + Streaming supported by Cloud DIKW
  - Integrated with DevOps dynamic interoperable deployment
- Tested with 10% Twitter feed on ~100 nodes parallel streaming clustering
- Drone and Turtlebot operational with scene analysis and planning being parallelized
- Applies to general DDDAS applications?

SPARE

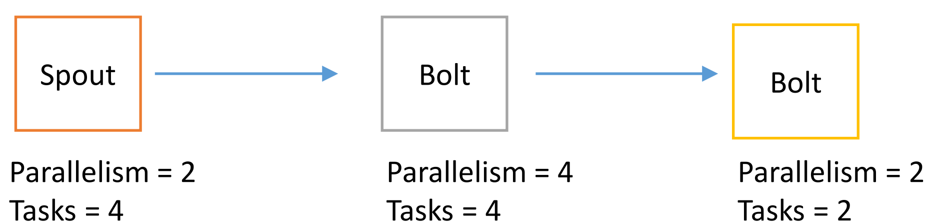
Runtime & Scalability

Components of a Topology runs as Tasks in Supervisor Nodes
Supervisor runs worker processes and these workers run the tasks
A user can run any number of instances of a Spout or Bolt in parallel

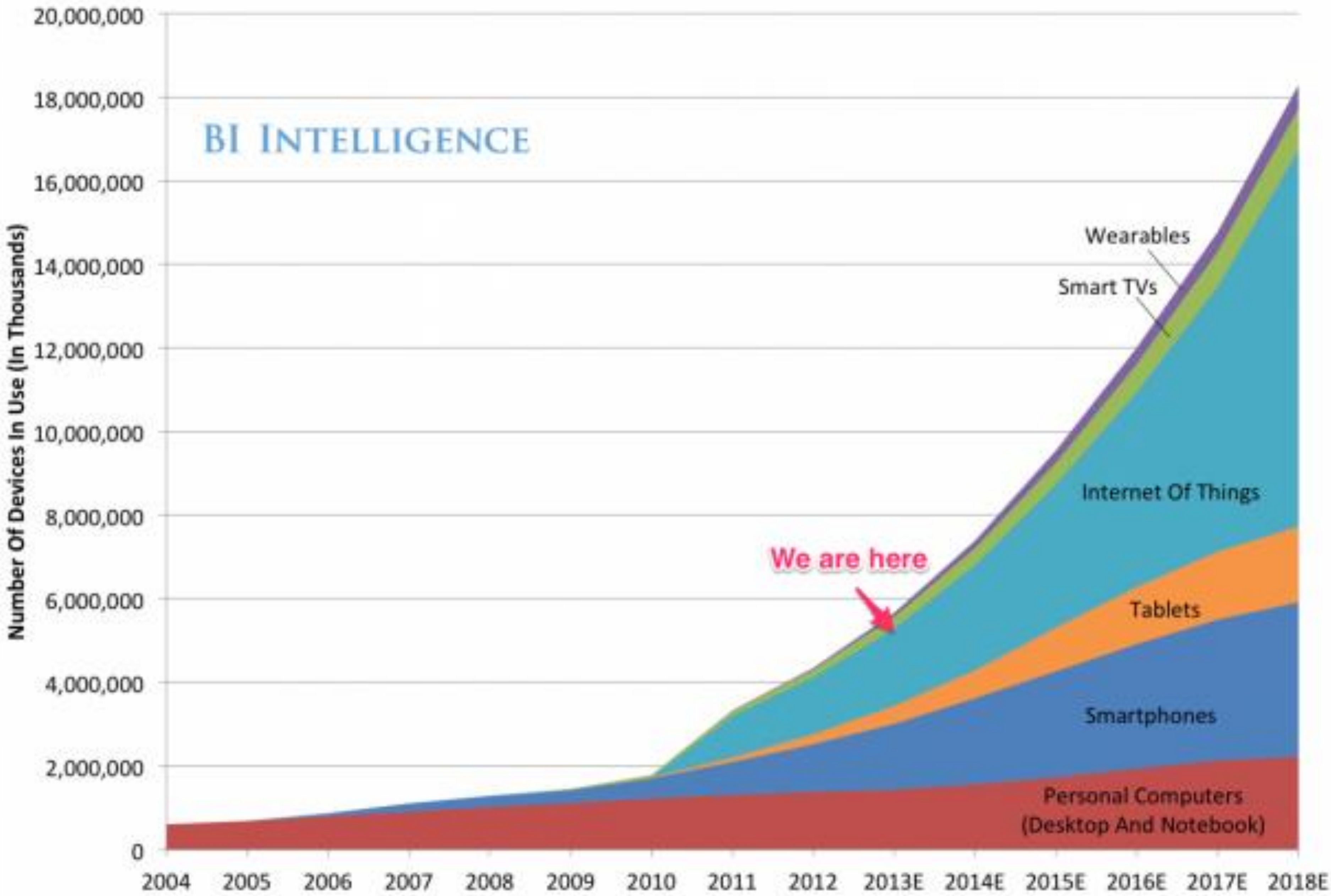


- Topology can specify how many workers it needs to run on
- Spout or Bolt can specify Parallelism and Tasks
  - Parallelism defines how many threads a spout or bolt runs on
  - Tasks defines how many instances of a spout or bolt to run
  - Parallelism ≤ Tasks

A Simple Topology running on two worker nodes



Global Internet Device Installed Base Forecast



Source: Gartner, IDC, Strategy Analytics, Machina Research, company filings, BII estimates

Scale of Industrial Internet

Social media versus electric generating power source

2012 Twitter Usage

vs.

Gas Turbine Compressor Blade Monitoring potential*

80 Gigabytes per day

enabling social connections

588 Gigabytes per day

enabling capital asset productivity

Data volume potential is 7x greater from a gas turbine than current Twitter usage

imagination at work

Ruh VP Software GE http://fisheritcenter.haas.berkeley.edu/Big_Data/index.html