

Introduction to Cloud Computing

B. Ramamurthy
Bina@buffalo.edu

CSE Department, University at Buffalo

This work is partially supported by the following grants from
National Science Foundation:

NSF-TUES-0920335, NSF-OCI-1041280

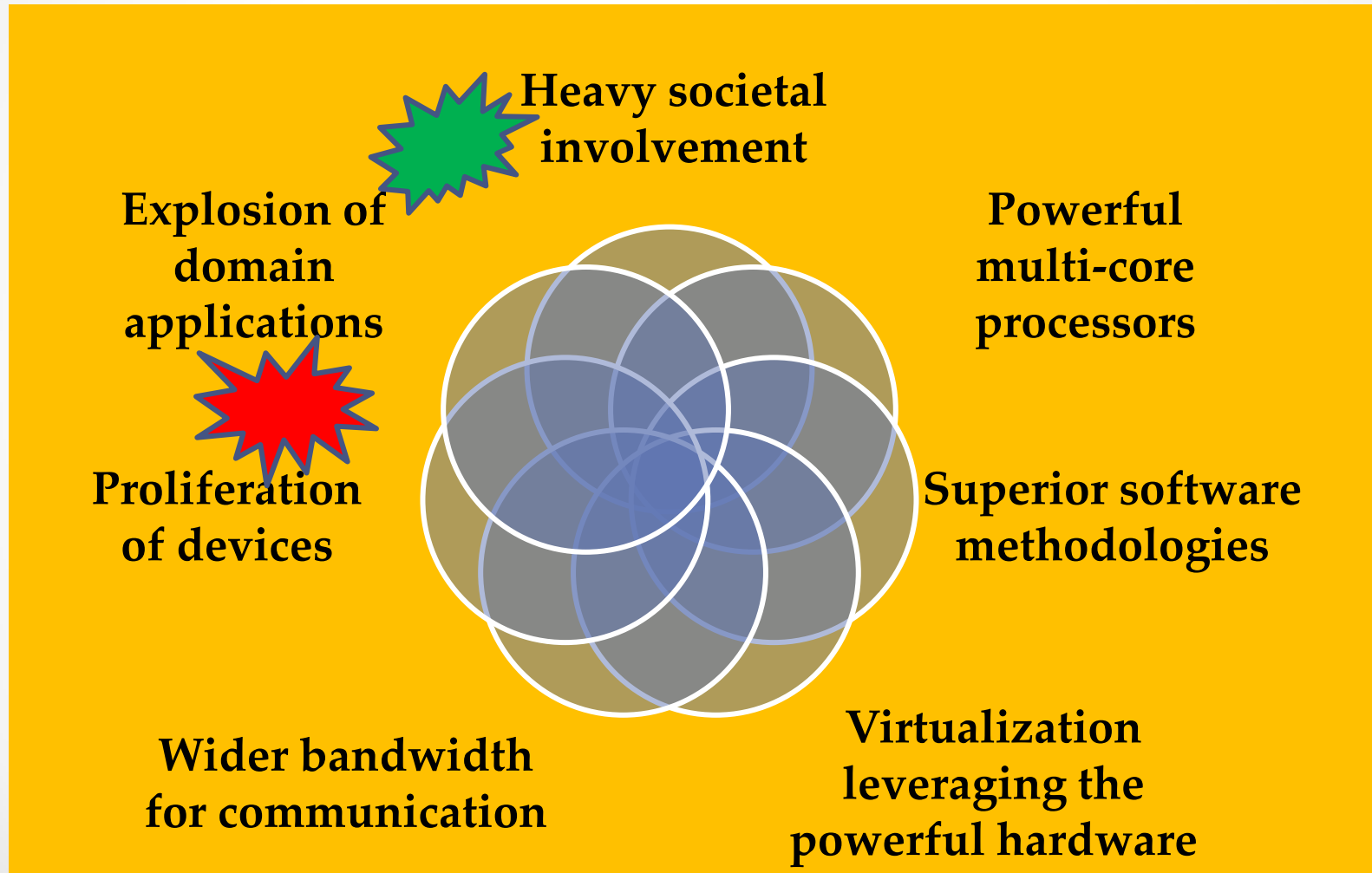
Presenter's Background in cloud computing

- Bina
 - Is a PI on two current NSF grants that uses cloud computing:
 - 2009-2012: Data-Intensive computing education: CCLI Phase 2: \$250K
 - 2010-2012: Cloud-enabled [Evolutionary Genetics Testbed](#): OCI-CI-TEAM: \$250K (Collaborative with Dr. Poulin and Dr. Ditmarr of Biology department)
 - Faculty at the CSE department at University at Buffalo.

Outline of the talk

- **Golden Era in Computing**
- **Data and Computing challenges**
- **Cloud Computing**
- **Popular Cloud Providers**
- **Our experience with Cloud hosting**
- **Summary**
- **References**
- **Questions and Answers**

A Golden Era in Computing



Computing Challenges

- Scalability issue: large scale data, high performance computing, automation, response time, rapid prototyping, and rapid time to production
- Need to effectively address (i) ever shortening cycle of obsolescence, (ii) heterogeneity and (iii) rapid changes in requirements
- Transform data from diverse sources into intelligence and deliver intelligence to right people/user/systems
- How to store the big-data? What new computing models are needed?
- What about providing all this in a cost-effective manner?
- How to make computing available and accessible as a public resource?

Enter the cloud

- **Cloud computing** is Internet-based computing, whereby shared resources, software and information are provided to computers and other devices on-demand, like the electricity grid.
- The cloud computing is a culmination of numerous attempts at large scale computing with seamless access to virtually limitless resources.
 - on-demand computing, utility computing, ubiquitous computing, autonomic computing, platform computing, edge computing, elastic computing, **grid computing**, ...

The Cloud Computing

- Cloud provides processor, software, operating systems, storage, monitoring, load balancing, clusters and other requirements as a service
- Pay as you go model of business
- When using a public cloud the model is similar to renting a property than owning one.
- An organization could also maintain a private cloud and/or use both.
- Cloud computing models:
 - platform (PaaS), Eg., Windows Azure
 - software (SaaS), Eg., Google App Engine
 - infrastructure (IaaS), Eg., Amazon AWS
 - Services-based application programming interface (API)



Google App Engine

- This is more a web interface for a development environment that offers a one stop facility for design, development and deployment Java and Python-based applications in Java, Go and Python.
- Google offers the same reliability, availability and scalability at par with Google's own applications
- Interface is software programming based
- Comprehensive programming platform irrespective of the size (small or large)
- Signature features: templates and appspot, excellent monitoring and management console;
- Free version to explore at: <http://code.google.com/appengine/>
- Software as a service: [Evolutionary Genetics Testbed](#)



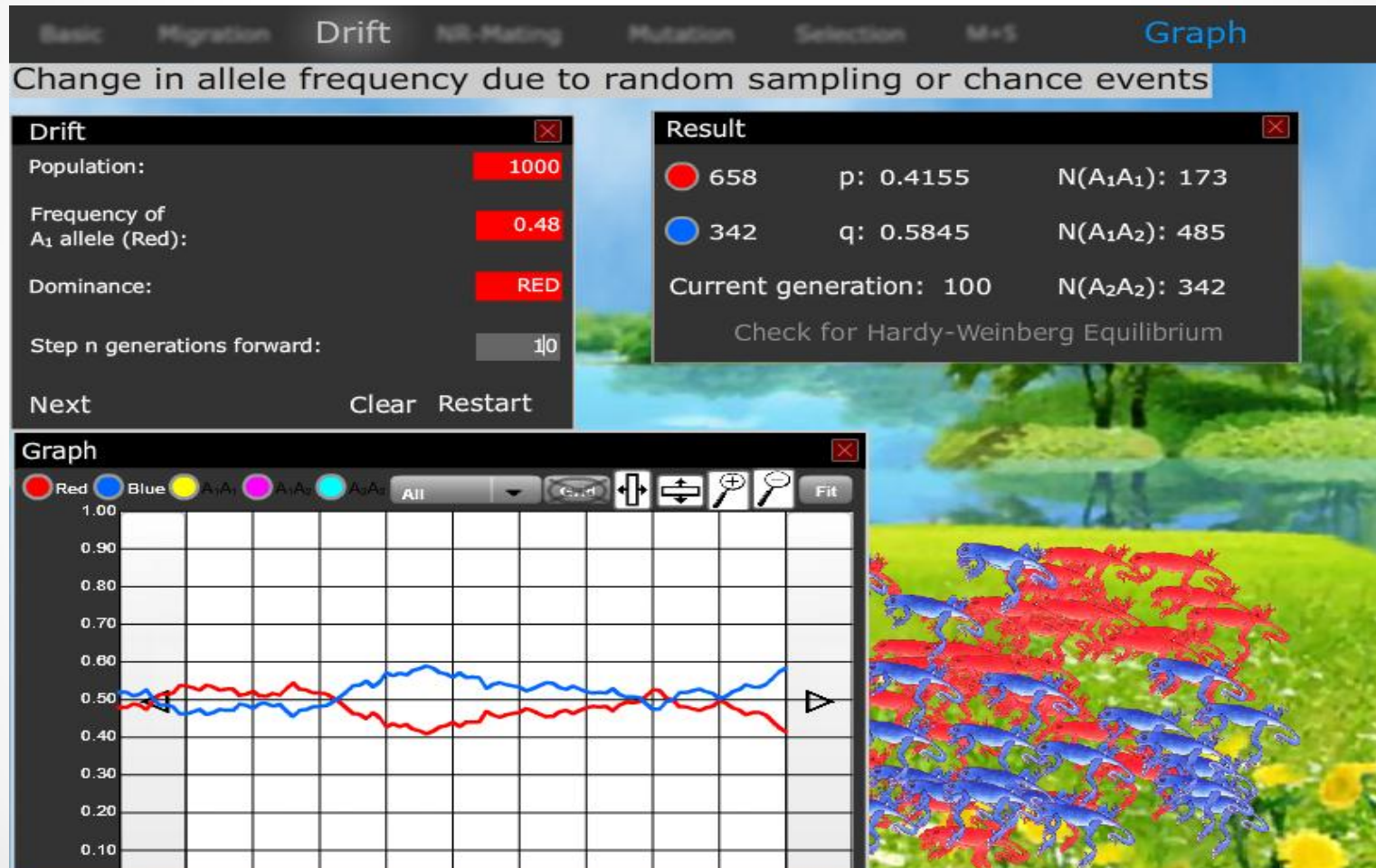
Amazon EC2

- Amazon EC2 is one large complex web service.
- EC2 provides an API for instantiating computing instances with any of the operating systems supported.
- It can facilitate computations through Amazon Machine Images (AMIs) for various other models.
- Signature features: S3, Cloud Management Console, MapReduce Cloud, Amazon Machine Image (AMI)
- Excellent distribution, load balancing, cloud monitoring tools
- You can explore amazon using the free account at:
- <http://aws.amazon.com/free/>

Pop!World

- Collaborative with BIO
- Dr. Jessica Poulin and Dr. Katharina Dittmar of Biology
- NSF supported
- Used by 1000+ entry level students in Biology
- Other upper level students also use it
- K-12 component is also included (Pop!World Gateway)
- Cloud deployed on Google App Engine
- Monitoring student learning behavior
- <http://popworld15.appspot.com>

Pop!World



Google App Engine Load Monitoring

Dashboard - Evolutionary Biology App - Mozilla Firefox

File Edit View History Bookmarks Tools Help

google.com https://appengine.google.com/dashboard?app_id=popworld10

Most Visited Getting Started Latest Headlines

Google App Engine - Google Code Dashboard - Evolutionary Biology ...

binakumar82@gmail.com | My Account | Help | Sign out

popworld10 1 < My Applications

Main

- Dashboard
- Quota Details
- Logs
- Cron Jobs
- Task Queues
- Blacklist

Data

- Datastore Indexes
- Datastore Viewer
- Datastore Statistics
- Blob Viewer

Administration

- Application Settings
- Permissions
- Versions
- Admin Logs

Billing

- Billing Settings
- Billing History

Resources

- Documentation
- FAQ

Charts

Requests/Second

6 hrs 12 hrs 24 hrs 2 days 4 days 7 days 14 days 30 days

Billing Status: Enabled (Daily budget: \$2.00) - Settings

Quotas reset every 24 hours. Next reset: 5 hrs

Resource	Usage	Cost / Budget
CPU Time	0% 0.00 of 18.50 CPU hours	\$0.00 / \$1.20
Outgoing Bandwidth	53% 1.95 of 3.67 GBytes	\$0.12 / \$0.32
Incoming Bandwidth	0% 0.00 of 1.80 GBytes	\$0.00 / \$0.08
Total Stored Data	0% 0.00 of 81.00 GBytes	\$0.00 / \$0.40
Recipients Emailed	0% 0 of 2,000	\$0.00 / \$0.00

Estimated cost for the last 19 hours: **\$0.12 / \$2.00**

Current Load

URI	Requests last 19 hrs	Avg CPU (API) last hr	% CPU last 19 hrs
/favicon.ico	536	0 (0)	0%
/templates/main.html	287	0 (0)	0%
/js/swfobject.js	29	0 (0)	0%

Errors

URI	Count	% Errors last 19 hrs
/favicon.ico	536	100%
/templates/main.html	258	90%

https://appengine.google.com/

S3 Fox

9:58 PM 10/11/2010

MemCache on GoogleAppEngine (2010)

10/17/2013

Memcache Status [System Status](#)

Today October 11, 2010

Get

- Latency
- Throughput
- Error Rate

Set

- Latency
- Throughput
- Error Rate

Get multi

- Latency
- Throughput
- Error Rate

Set multi

- Latency
- Throughput
- Error Rate

Delete multi

- Latency
- Error Rate

Get: Latency
Measures the latency, in milliseconds, of call to memcache.get() for a single key referencing 10kB of data.

Time (US/Pacific)	Scope	Description
06:06 PM - 06:27 PM	Memcache	Memcache partial unavailability 15% of applications experienced Memcache unavailability. All get calls would have returned a cache miss during this period.

Java is a trademark or registered trademark of Sun Microsystems, Inc. in the United States and other countries.
© 2009 Google | [Terms of Service](#) | [Privacy Policy](#) | [Blog](#) | [Discussion Forums](#)

Memcache partial unavailability

Summary

- We are entering a watershed moment in the internet era.
- This involves in its core and center, big data analytics and tools that provide intelligence in a timely manner to support decision making.
- Newer storage models, processing models, and approaches have emerged.
- Among these cloud computing has the potential to significantly improve accessibility to computing
- See: UB-implemented a SUNY-wide a [Certificate Program in Data-intensive Computing](#)

References & useful links

- Amazon AWS: <http://aws.amazon.com/free/>
- AWS Cost Calculator:
<http://calculator.s3.amazonaws.com/calc5.html>
- Windows Azure: <http://www.azurepilot.com/>
- Google App Engine (GAE):
<http://code.google.com/appengine/docs/whatisgoogleappengine.html>
- For miscellaneous information:
<http://www.cse.buffalo.edu/~bina>
- <http://www.cse.buffalo.edu/~bina/DataIntensive>