ISTC for Cloud Computing: Center Overview

Greg Ganger & Phil Gibbons
ISTC-CC Retreat 2013
November 7, 2013

http://www.istc-cc.cmu.edu/

Intel Science & Technology
Center for Cloud Computing

Why ISTC for Cloud Computing?

- **Cloud computing has exploded onto the scenes**
  - By 2020, most computing may be cloud-based

- **Cloud computing promises huge benefits in...**
  - resource efficiency and utilization
  - agility and productivity of application dev+deploy
  - system and service robustness
  - data sharing and end user access

- **But, hype far ahead of capability & knowledge**
  - huge need for research at many levels

ISTC-CC: Institutions & Faculty

- **Carnegie Mellon University**
  - Greg Ganger (PI), Dave Andersen, Guy Blelloch, Garth Gibson, Mor Harchol-Balter, Todd Mowry, Onur Mutlu, Priya Narasimhan, M. Satyanarayanan, Dan Siewiorek, Alex Smola, Eric Xing

- **Georgia Tech**
  - Greg Eisenhower, Ada Gavrilovska, Ling Liu, Calton Pu, Karsten Schwan, Matthew Wolf, Sudha Yalamanchili

- **Princeton University**
  - Mike Freedman, Margaret Martonosi

- **University of California at Berkeley**
  - Anthony Joseph, Randy Katz, Ion Stoica

- **University of Washington**
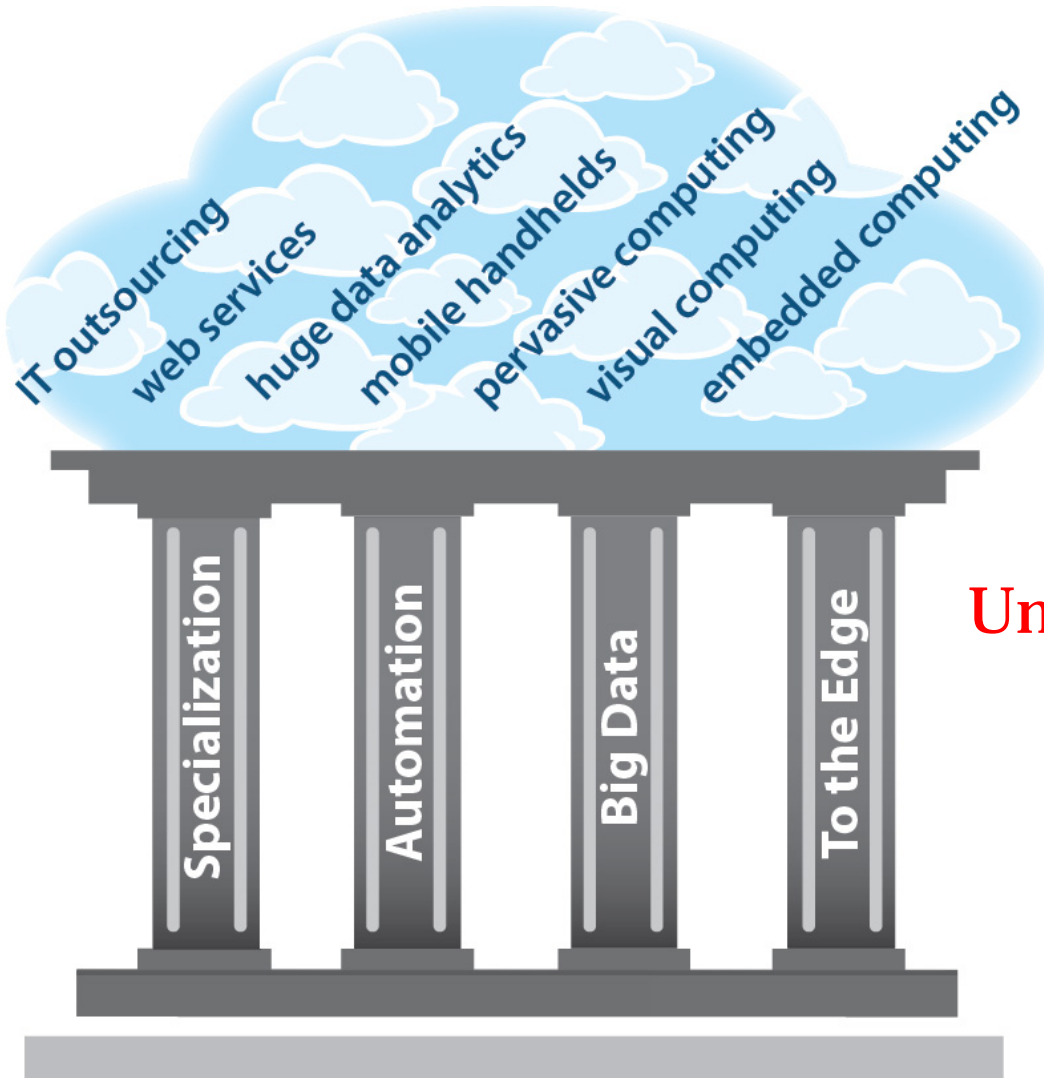 - Carlos Guestrin

- **Intel Labs**
  - Phil Gibbons (PI), Michael Kaminsky, Mike Kozuch, Babu Pillai

ISTC-CC: Intel Oversight & BoA

- **Intel oversight**
  - Rich Uhlig (Executive Sponsor)
  - Scott Hahn (Managing Director)
  - Chris Ramming (Director of UCO)
  - Jeff Parkhurst (Program Director)

- **Board of Advisors (including Rich and Scott)**
  - Randy Bryant (Dean of School of CS, CMU)
  - Jeff Chase (Professor of CS, Duke)
  - Balint Fleischer (Gen. Mgr. Data Center Group, Intel)
  - Frans Kaashoek (Professor of CS&Eng, MIT)
  - Pradeep Khosla (Chancellor, UC San Diego)
  - Jason Waxman (Gen. Mgr. Cloud Infra. Group, Intel)

ISTC-CC's Research Pillars

IT outsourcing · web services · huge data analytics · mobile handhelds · pervasive computing · visual computing · embedded computing

Specialization · Automation · Big Data · To the Edge

Underlying Infrastructure enabling the future of cloud computing

www.istc-cc.cmu.edu

Automation Pillar

- Automation is crucial to cloud reaching potential
  - We suspect that no one here needs to be convinced of this…

- Management is very hard, but cloud makes it worse
  - Much larger scale
  - Much more varied mix of applications/activities
  - Much less pre-knowledge of applications
  - And, we're adding in platform specialization ☺

- Leaps forward needed on many fronts…
  - Diagnosis, scheduling, instrumentation, isolation, tuning, …

Specialization	Automation	Big Data	To the Edge

Automation Projects

- **A1: Resource Scheduling for Heterogeneous Cloud Infrastructures**
  - maximizing the effectiveness of a cloud composed of diverse specialized platforms servicing diverse app types
  - enabling software framework specialization via hierarchical scheduling

- **A2: Problem Diagnosis and Mitigation**
  - new tools and techniques for rapid, robust diagnosis of failures and performance problems
  - automated mitigation based on "quick and dirty" online diagnoses

Specialization	Automation	Big Data	To the Edge

Automation at Retreat 2013

- Talks
  - **"Better Problem Diagnosis in the Cloud"** – Priya Narasimhan [Thurs am]
  - **"Hierarchical Scheduling for Diverse Datacenter Workloads"** – Ali Ghodsi [Fri am]

- Posters -- first session
  - Per-Application Server Specialization in Data Centers
  - Enabling End-to-End Latency & Throughput SLOs on Shared Storage
  - SpringFS: Bridging Agility and Performance in Elastic Distributed Storage
  - Dataflow-Analysis-Based Dynamic Parallel Monitoring
  - Tetrisched: Space-Time Scheduling for Heterogeneous Datacenters
  - Automated performance problem mitigation with efficient resource allocations
  - Tachyon: Reliable File Sharing at Memory-speed Across Cluster Frameworks
  - Oncilla: A GAS Run-time for Efficient Resource Partitioning in Accelerated Clusters
  - Experiences with Fault-Injection in a Byzantine Fault-Tolerant Protocol

Big Data Pillar

Customer Database

~600 TB

HD Internet Video

12 EB/yr

Particle Physics

300 EB/yr

- **Extracting insights from large datasets**
  - "Analytics" or "Data-intensive computing"
  - Becoming critical in nearly every domain
    - likely to dominate future cloud data centers

- **Need right programming/execution models**
  - For productivity, efficiency, and agility
  - Resource efficient operation on shared, specialized infrastructures

10^6 MEGA 10^9 GIGA 10^{12} TERA 10^{15} PETA 10^{18} EXA

9

Big Data Projects

- **B1: Big Learning Systems**
  - new programming abstractions and execution frameworks enabling efficiency and productivity for large-scale Machine Learning on Big Data

- **B2: Big Data Storage**
  - exploring trade-offs and new approaches in Big Data storage, including support for high ingress and multi-framework sharing of data

Specialization	Automation	Big Data	To the Edge

Big Data at Retreat 2013

- Talks -- Thurs late afternoon (+ 1 Thurs morning)
  - **"Discretized Streams: Fault-Tolerant Streaming Computation at Scale"** – Ion Stoica
  - **"More Effective Distributed ML via a Stale Synchronous Parallel Parameter Server"** – Eric Xing
  - **"Scaling Big Data Processing with Utility-aware Distributed Data Partitioning"** – Ling Liu
  - **"Parrot + dBug: Fast and Reliable Multithreading"** – Jiri Simsa
  - **"What's New with GraphLab"** – Carlos Guestrin

Big Data at Retreat 2013

- **Posters -- first session**
  - Big Machine Learning: Needs and Directions
  - LazyTable: Distributed Machine Learning with the Stale Synchronous Parallel Model
  - Exploiting Bounded Staleness To Speed Up Big Data Analytics
  - STRADS: A Distributed ML Task Scheduler For Large-Scale Convex Optimization Problems
  - FastHASH: A New Algorithm for Fast and Comprehensive Next-generation Sequence Mapping
  - Discretized Streams: Fault-Tolerant Streaming Computation at Scale
  - Approximating User-Defined Functions in BlinkDB
  - What Test Should I Run? Improving Concurrency Tests using State Space Estimation
  - Scaling Distributed File System Metadata Throughput using IndexFS
  - ShardFS: Scalable Metadata for HDFS
  - Out-of-core Metadata for HDFS
  - DSS Gen-2:Enabling Storage QoS for Virtual Machines
  - Reducing Contention Through Priority Updates

Specialization Pillar

Low power nodes

Many-core

Phase-change memory (PCM)

- **Specialization is fundamental to efficiency**
  - No single platform best for all application types
  - Called **division of labor** in sociology

- **Cloud computing must embrace specialization**
  - As well as consequent heterogeneity and change-over-time
  - Stark contrast to common cloud thinking

- **New approaches needed to enable...**
  - Effective mixes of targeted and general platform types
  - Nimble incorporation of new technologies and accelerators

Specialization Projects

- **S1: Specialized Platforms of Wimpy Nodes**
  - exploring + extending range of apps that run (most) efficiently on such platforms by overcoming OS limits, memory limits, and scalability issues

- **S2: Specialized Platforms of Heterogeneous Multi-Cores**
  - exploring best ways to devise and use heterogeneity on multi-core nodes, considering core types, accelerators, and DRAM/NVM memory, with a focus on cloud's virtualized, multi-tenancy workloads

Specialization	Automation	Big Data	To the Edge

Specialization at Retreat 2013

- Talks -- Friday morning

  - "**NVM and Hetero Systems**" – Karsten Schwan

  - "**Rethinking Memory System Design for Data-Intensive Computing**" – Onur Mutlu

 - "**Fast Networked Key-Value Stores and Cuckoo Hashing**" – Dave Andersen

Specialization at Retreat 2013

Posters   -- second session

- Scalable, High Performance Ethernet Forwarding on x86 Platforms
- MICA: A Holistic Approach to Near-Line-Rate In-Memory Key-Value Caching on General-Purpose Hardware
- Hardware + Algorithms = Seriously Concurrent Hash Tables
- Design Space Analysis for Heterogeneous Systems
- Exploration of Data Warehousing and Graph Applications with GPUs
- ANUMA: Asymmetry-aware Execution Placement on Manycore Chips
- Tiered-Latency DRAM: A Low Latency and Low Cost DRAM Architecture
- Base-Delta-Immediate Compression: Practical Data Compression for On-Chip Caches
- Linearly Compressed Pages: A Low Complexity, Low Latency Main Memory Compression Framework
- RowClone: Fast and Energy-efficient In-DRAM Bulk Data Copy and Initialization
- A Case for Subarray-Level Parallelism (SALP) in DRAM
- MISE: Providing Performance Predictability in Shared Main Memory Systems
- Improving DRAM Performance by Parallelizing Refreshes with Accesses
- A Case for Hardware/Software Cooperative Management of Storage and Memory

To the Edge Pillar

- **Edge devices will participate in cloud activities**
  - Serving as bridge to physical world (sense/actuate)
  - Enhancing interactivity despite location / connectivity

- **Need new programming/ execution models**
  - For adaptive cloud
    + edge cooperation
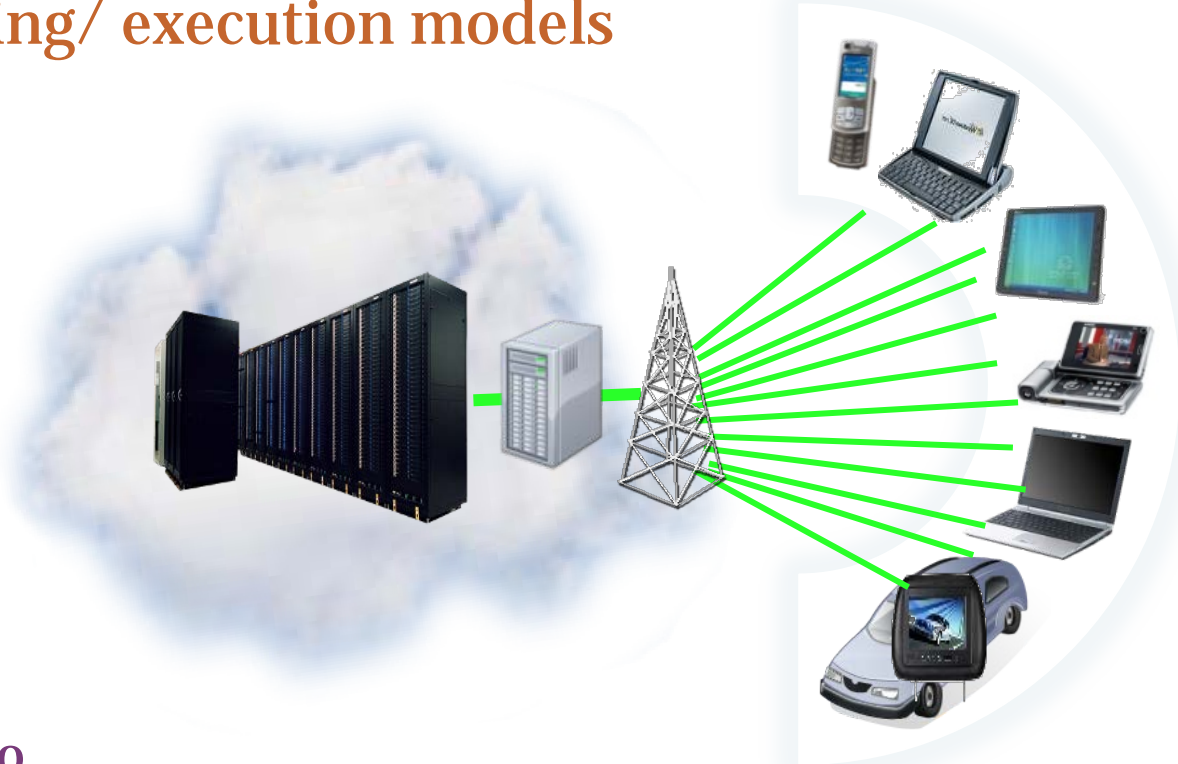
Cloudlet demo

To the Edge Projects

- **E1: Cloud-Assisted Mobile Client Computations**
  - new abstractions and system architectures for dynamic exploitation of edge-local cloud resources to enable rich edge device experiences

- **E2: Geographically Distributed Data Storage**
  - new techniques for geographically distributed data storage/caching that reduce both access latency & reliance on expensive WAN-uplink bandwidth, while providing the desired scalability, fault tolerance, consistency & findability

Specialization	Automation	Big Data	To the Edge

To the Edge at Retreat 2013

- Talks -- Thursday afternoon
  - **"What's New with Cloudlets"** - Satya
  - **"Stronger Semantics for Low-Latency Geo-Replicated Storage"** - Wyatt Lloyd
  - **"There Is More Consensus in Egalitarian Parliaments"** - Iulian Moraru

- Posters -- second session
  - OpenStack++ for Cloudlets
  - Agentless Cloud-wide Monitoring of Virtual Disk State
  - vTube: Efficient Streaming of Virtual Appliances Over Last-Mile Networks
  - Scalable Crowd-Sourcing of Video from Mobile Devices
  - QuiltView: a Crowd-Sourced Video Response System
  - Hyrax: Edge Clouds of Mobile Devices
  - Leveraging Embedded-to-Cloud Framework for Retail Environments
  - Egalitarian Paxos

ISTC-CC: Cross-Institution Collaboration

	Project	Personnel
S1	Specialized Platforms of Wimpy Nodes	Andersen[C], Schwan[G], Freedman[P], Kaminsky[I], Kozuch[I], Pillai[I]
S2	Specialized Platforms of Heterogeneous Many-Cores	Mowry[C], Mutlu[C], Gavrilovska[G], Schwan[G], Yalamanchili[G], Martonosi[P], Gibbons[I], Kozuch[I]
A1	Resource Scheduling for Heterogeneous Cloud Infrastructures	Joseph[B], Katz[B], Stoica[B], Ganger[C], Harchol-Balter[C], Kozuch[I]
A2	Problem Diagnosis and Mitigation	Ganger[C], Narasimhan[C], Eisenhauer[G], Liu[G], Schwan[G], Wolf[G]
B1	Big Learning Systems	Stoica[B], Andersen[C], Blelloch[C], Ganger[C], Gibson[C], Smola[C], Xing[C], Guestrin[W], Gibbons[I]
B2	Big Data Storage	Andersen[C], Ganger[C], Gibson[C], Xing[C], Pu[G], Schwan[G]
E1	Cloud-Assisted Mobile Client Computations	Satya[C], Siewiorek[C], Gavrilovska[G], Liu[G], Schwan[G], Martonosi[P], Pillai[I]
E2	Geographically Distributed Data Storage	Andersen[C], Satya[C], Siewiorek[C], Freedman[P], Kaminsky[I], Pillai[I]

ISTC-CC: Year 2 Publication Highlights

- PACT (Sep'12) – 3 papers
- ICAC (Sep'12) – 3
- ICCD (Sep'12) – 2
- OSDI (Oct'12) – 3
- SOCC (Oct'12) – 6
- Middleware (Dec'12) – 2
- HPCA (Feb'13) – 4
- PPoPP (Feb'13) – 2
- DATE (Mar'13) – 2
- NSDI (Apr'13) – 3
- IPDPS (May'13) – 2
- HotOS (May'13) – 5

- ISCA (Jun'13) – 3
- SIGMOD (Jun'13) – 3
- Mobisys (Jun'13) – 2
- ICWS (Jun'13) – 2
- ICDCS (Jul'13) – 2
- KDD (Aug'13) – 3
- SIGCOMM (Aug'13) – 2
- IRI (Aug'13) – 2

Year 3 (so far):
- ICCD (Sep'13) – 2
- SOCC (Oct'13) – 3
- SOSP (Nov'13) – 4
- MICRO (Dec'13) – 2

85 published papers highlighted in
ISTC-CC Newsletter for Year 2 (July'12- Sept'13)

Open Source Code Releases in Year 2

Open Source page: www.istc-cc.cmu.edu/research/ossr/

- **GraphBuilder 1.0** released open source in Jun'13
- **GraphLab 2.2** released open source in Jul'13

- **Spark 0.8** release Sep'13 – Apache incubator

- **Mesos 0.14** released Oct'13 – Apache

- Other open source releases on github include:
  Eiger, EPaxos, Parrot, Cloudlet OpenStack++, CuckooFilter, RankSelect, MemC3, NVMalloc, etc.

Also Benchmarks page: www.istc-cc.cmu.edu/research/benchmarks/

ISTC-CC: Year 2 Honors

Highlights of Year 2 Honors

- **ACM Fellows –** Gibson, Stoica [now 8 in all + 5 IEEE Fellows]
- **ABIE Technical Leadership Award –** Martonosi
- **2013 R&D100 award** – Schwan, Wolf, others
- **Founding Editor-in-Chief of ACM TOPC** – Gibbons
- **LANL Outstanding Innovation Award** – Gibson
- **Endowed Early Career Professor** – Mutlu
- **Allen Newell Award** – Andersen, Kaminsky
- **2012 SPEC Distinguished Dissertation Award** – Meng
- Multiple best paper awards, Program chairs, Conference Chairs, Outstanding dissertation honors, etc

ISTC-CC: Amplifying Funding

New Amplifying Funding awarded in Year 2

- Guy Blelloch (CMU) & Phil Gibbons (IL) – NSF
- David Andersen (CMU) & Michael Kaminsky (IL) – NSF
- Mor Harchol-Balter (CMU) – NSF
- Sudha Yalamanchili (GA Tech) – NSF
- Karsten Schwan, Matt Wolf, Ada Gavrilovska (GA Tech)
- Satya (CMU) – Inst. For Museum & Library Sciences
- Satya (CMU) – Sloan Foundation

- Consortia: AMPlab (Berkeley), CERCS (GA Tech), PDL (CMU)
- Fellowships: Intel (2), Facebook, IBM, Microsoft, NVIDIA, Qualcomm, VMWare

+ Considerable Amplifying Funding carrying over into Year 2

=> Intel support for 25 students, yet working with 87

ISTC-CC Students @ Intel

Full time Employees:
- Soila Kavulya – CMU
- Min Lee – GA Tech

Interns (2012 or 2013):
- Hrishikesh Amur (2012) – GA Tech*
- Chris Fallin (2012) – CMU
- Liting Hu (2013) – GA Tech*
- Sudarsun Kannan (2013) – GA Tech
- Yoongu Kim (2012) – CMU
- Xiaozhou Li (2013) – Princeton*
- Jamie Liu (2012) – CMU
- Dan Lustig (2013) – Princeton
- Alex Merritt (2012,2013) – GA Tech
- Dipanjan Sengupta (2013) – GA Tech
- Vivek Seshadri (2012) – CMU (1 month visit)
- Priyanka Tembey (2012) – GA Tech

* With embedded
ISTC-CC researchers

Intel Fellowship winners (2013):
- Dan Lustig – Princeton
- Yoongu Kim – CMU

Agenda for Rest of Today

[10:15-11:05] Research Talks session #1

- Ion Stoica (UC Berkeley), Priya Narasimhan (CMU)

[11:05-11:30]   Break

[11:30-noon]   Poster previews (30 secs each) - Babu Pillai (Intel Labs)

[noon-1:20]   Lunch /Poster session #1

[1:20-2:40]   Poster session #2

[2:40-4:20]   Research Talks session #2

- Mahadev Satyanarayanan (CMU), Wyatt Loyd (ex-Princeton), Iulian Moraru (CMU), Eric Xing (CMU)

[4:20-4:40]   Break

[4:40-5:55]   Research Talks session #3

- Ling Liu (GA Tech), Jiri Simsa (CMU), Carlos Guestrin (UW)

[5:55-6:15]   Minutes of Madness - Michael Kaminsky (Intel Labs)

[6:15-6:30]   Group Photo   then Transportation to dinner

[6:30-9:00]   Reception / Dinner at Pittsburgh Athletic Club

Concluding Thoughts

Retreat Goal: Benefit the research projects

- Community building
- Brainstorming/feedback (but not ARs ☺) on:
  - ideas & approaches for tackling the research challenges
  - jump-start collaborations
  - synergies & connections
- Learn, share & have fun

Who to see about...
- Poster Previews/Sessions: see Babu Pillai
- Minutes of Madness: see Michael Kaminsky
- Birds-of-a-Feather Breakouts: see Mike Kozuch
- Logistics questions/issues: see Jennifer Gabig