Jim Larus

• Dean of IC (since October 2013)

• Assistant/Associate Professor: University of Wisconsin – Madison
• Researcher, Manager, Director: Microsoft Research (16 years)

• Published 100+ papers and book:
  • programming languages
  • compilers
  • software development
  • systems
  • computer architecture
  • parallel programming
  • cloud computing

• ACM Fellow
• 30+ US Patents
• Persona – High Performance Genomic Computation
• Non-Volatile Memory
• Automated Debugging
• Machine Learning Applied to Defect Detection
Persona: A High-Performance Bioinformatics Framework

Stuart Byma\textsuperscript{1}, Sam Whitlock\textsuperscript{1}, Laura Flueratoru\textsuperscript{2}, Ethan Tseng\textsuperscript{3}, Christos Kozyrakis\textsuperscript{4}, Edouard Bugnion\textsuperscript{1}, James Larus\textsuperscript{1}

EPFL\textsuperscript{1}, U. Polytehnica of Bucharest\textsuperscript{2}, CMU\textsuperscript{3}, Stanford\textsuperscript{4}
Sequencing cost

Not a wet lab problem anymore → IT / Systems problem
Implications

Need efficient systems that scale well
Persona – Bioinformatics, Unified
Aggregate Genomic Data

- Bases
- Q-Scores
- Metadata

Storage Subsystem

Manifest

- Header
- Index
- Data

compressed
Dataflow

• Modularity
• Balance/tuning
• (bounded) Queueing
Distributed Computation

Client

$ persona client bwa-align

Queue Service

Server 0

Server 1

Server 1

Server N

Storage Subsystem
Using Non-Volatile Memory
NVM (Non-Volatile Memory)

- 3D XPoint
- PCM
- STT-RAM
- ReRAM

Cache-line granularity

Persistent

DRAM-level access time
How to use NVM?

Problem:

Can’t just put data in NVM and use loads and stores to write to it.

Data must be in consistent state to resume computation after failure.

```
begin transaction
    from_acct.debit(amount);
    to_acct.cred(amount);
end transaction
```
Projects?

I’m looking for a small number of new students this year

Please come and talk with me if you are interested in doing a project