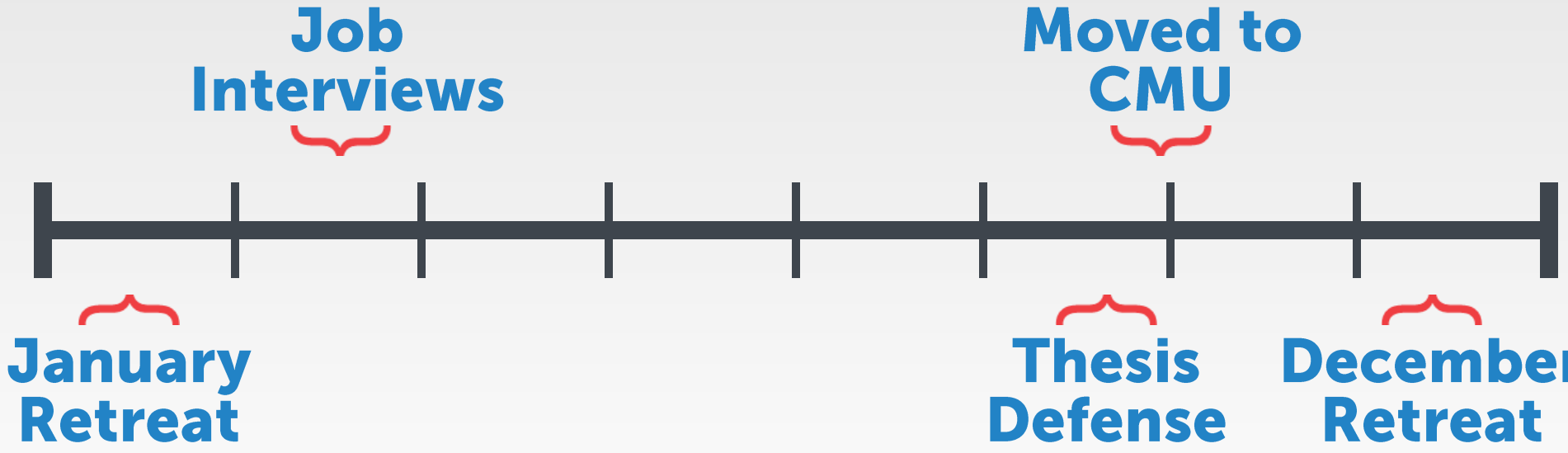
OLTP on the NVM SDV:
YMMV

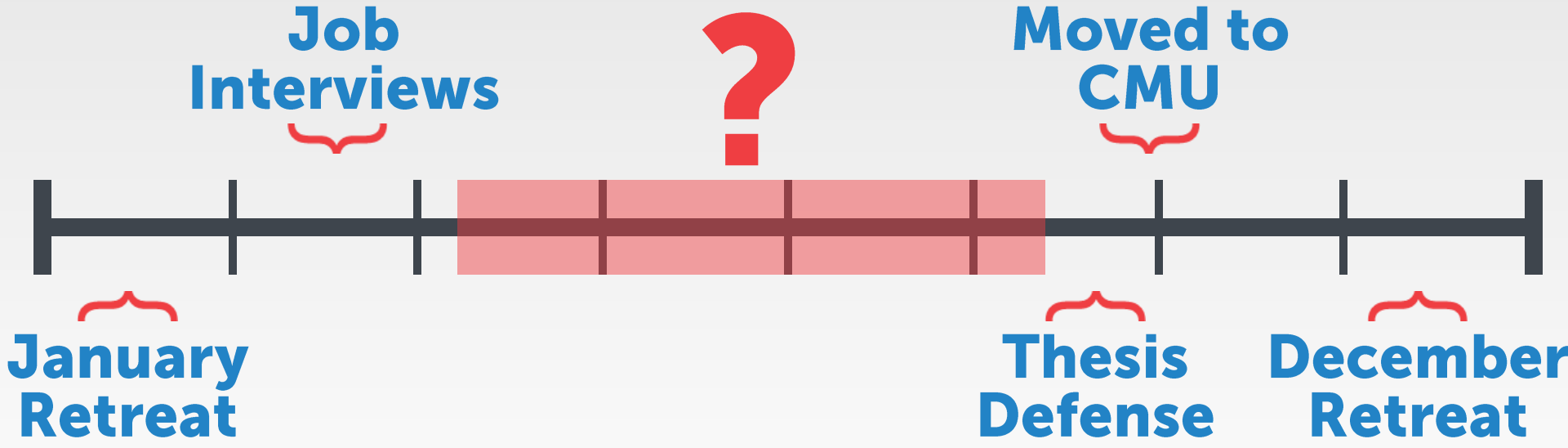
ISTC
BIG DATA

@andy_pavlo

2013


Job Interviews
Moved to CMU
January Retreat
Thesis Defense
December Retreat

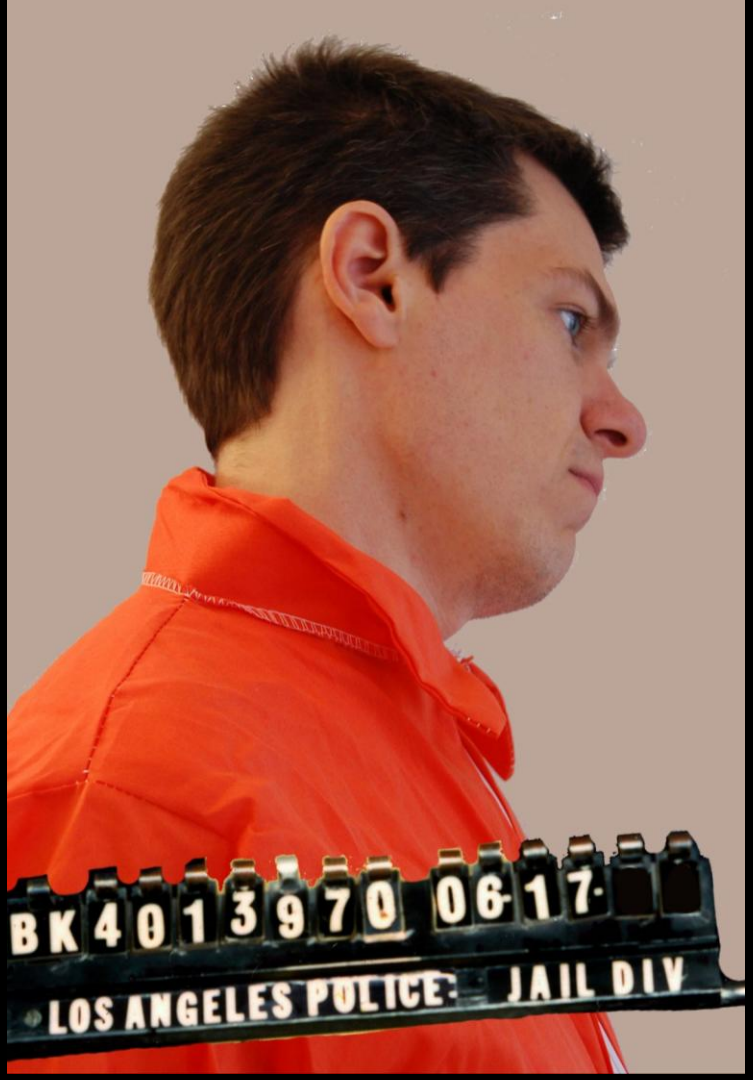
2013



ISTC BIG DATA


ROVIDENCE
POLICE
DEPARTMENT


BK4013970 06-17-94
LOS ANGELES POLICE JAIL DIV


BK4013970 06-17-
LOS ANGELES POLICE JAIL DIV

Prison Life

Washing Dishes
Not Fighting
Repentant

Cafeteria Thievery
Shankings
Making Pruno

NVM OLTP

Lightweight CC
Logical Logging
Snapshots

Heavyweight CC
ARIES Logging
Making Pruno

Project Overview

- Understand the performance characteristics of NVM to develop an optimal DBMS architecture for it.

Current Research Status

- Compare DBMS architectures:
  - *H-Store with Logging → NVM*
  - *Anti-Caching → NVM (block)*
  - *Anti-Caching → NVM (tuple)*
  - *H-Store with MMAP → NVM*

H-Store – NVM Edition

- Converted storage manager to use NVM through MMAP:
  - *Table tuple storage.*
  - *VARCHAR/BLOB data pool.*
  - *Table indexes.*

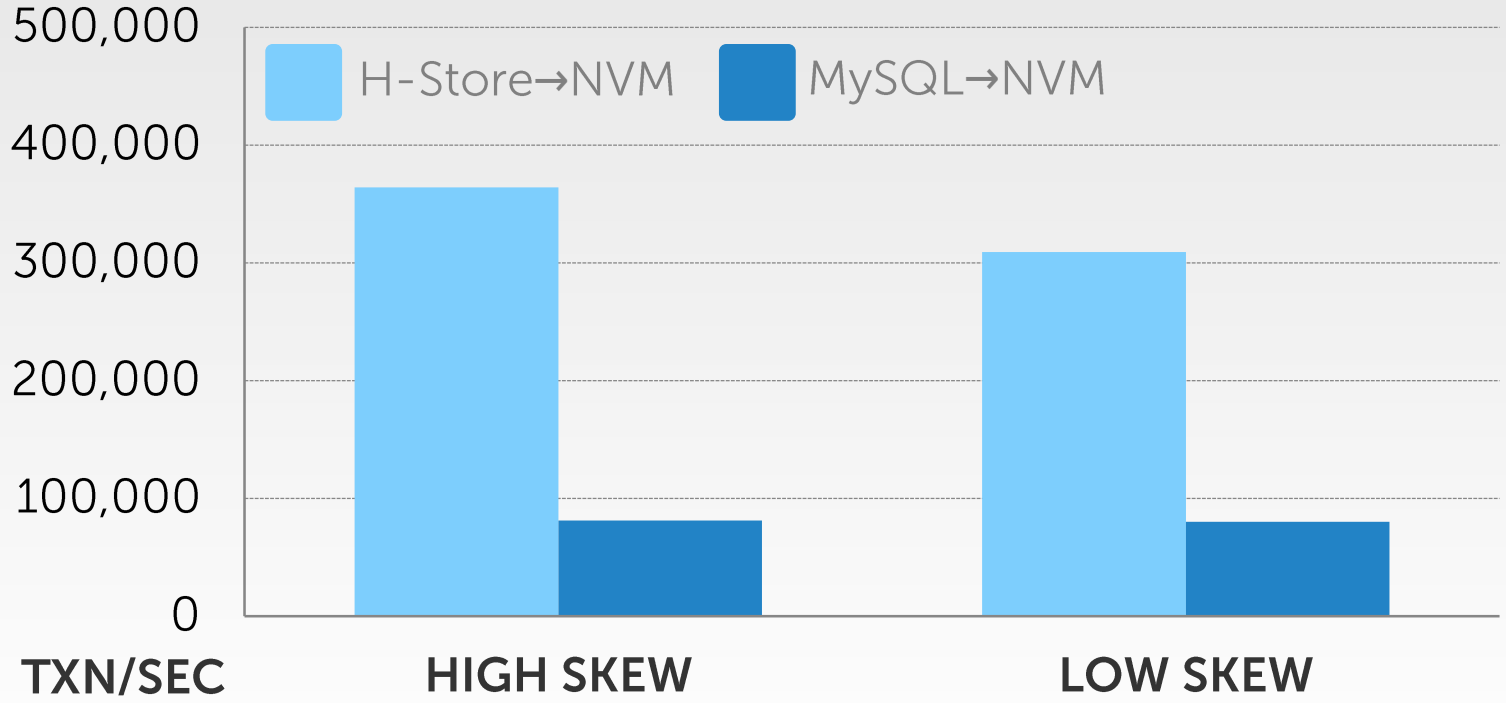
H-Store – NVM Edition

- Minor changes to exec engine:
  - *Group commit (msync)*
- Current implementation is not transactional fail-safe.

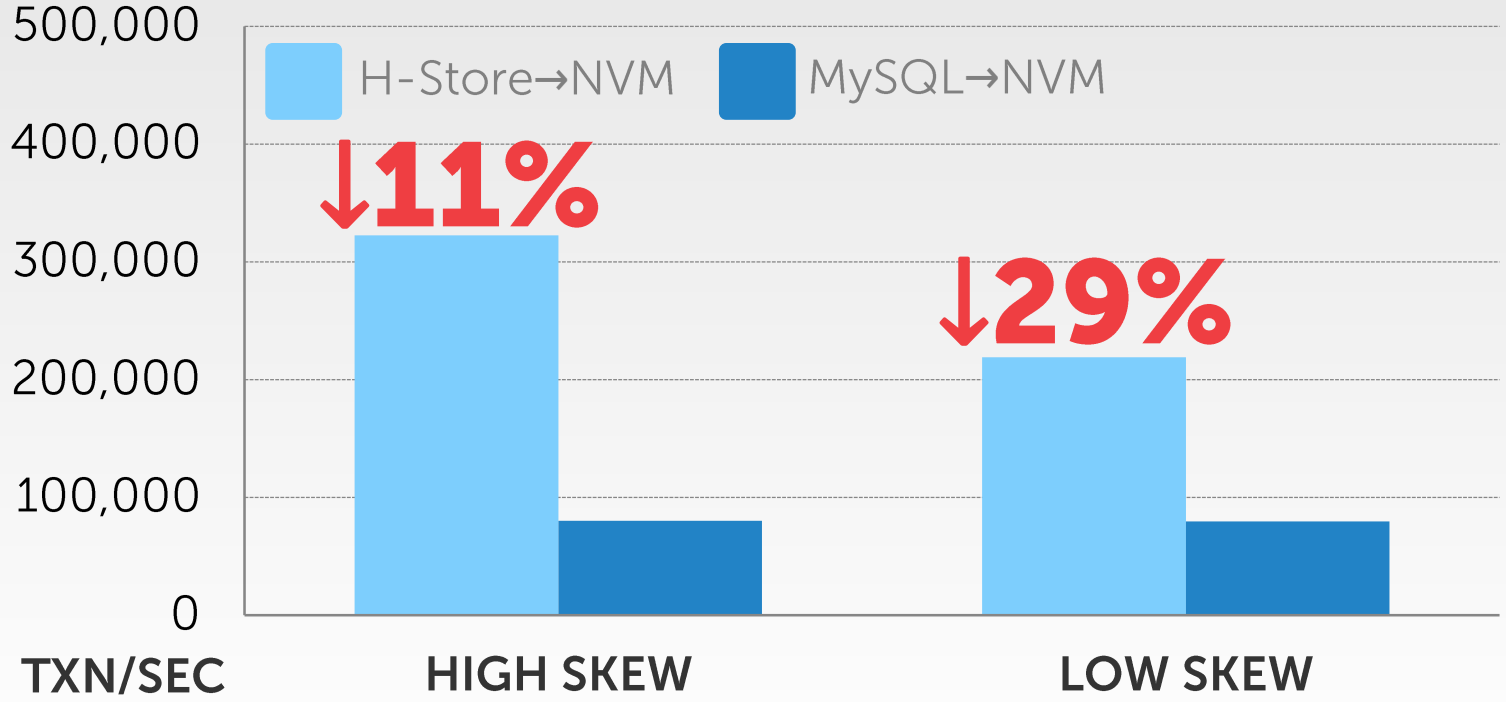
Experimental Evaluation

- Compare two architectures:
  - *H-Store with MMAP → NVM*
  - *MySQL → NVM*
- Yahoo! Cloud Serving Benchmark:
  - *10 million records (~10GB)*
  - *8x database / memory MySQL*

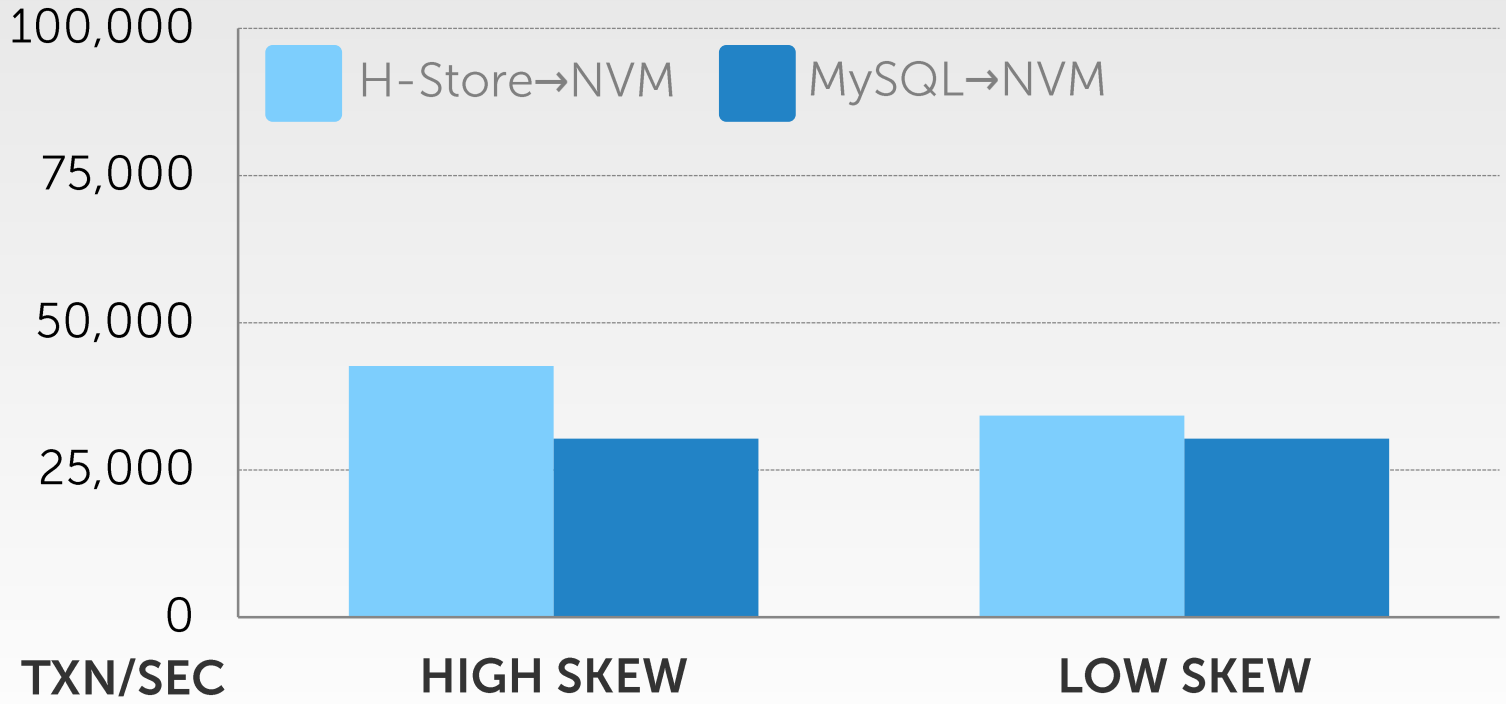
YCSB // *Read-Only Workload*
2x Latency Relative to DRAM



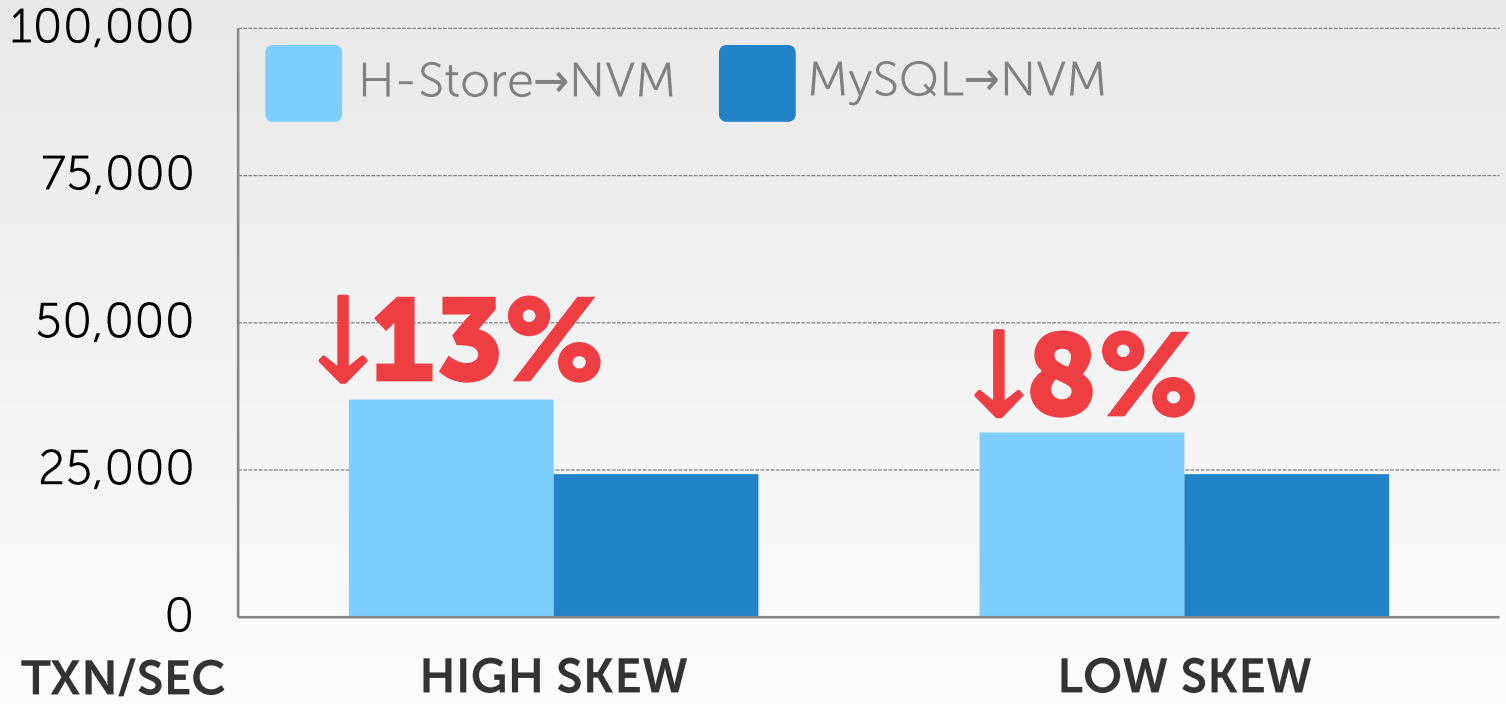
YCSB // *Read-Only Workload*
16x Latency Relative to DRAM



YCSB // *50% Reads / 50% Writes Workload*
2x Latency Relative to DRAM



YCSB // *50% Reads / 50% Writes Workload*
16x Latency Relative to DRAM



Discussion

- Overhead is only from reads and msync syscalls.
- We're not doing anything special with the NVM.
  - *Contents of NVM are ignored for crash recovery.*

What's Next?

- Port snapshot code to use NVM.
- Improve MMAP storage manager.
- Full comparison + experiments.
- Start building the next system…

N-STORE

N-Store – Possibility #1

- Keep DRAM.
- Hybrid OLTP/OLAP architecture:
  - *Row Store → DRAM*
  - *Column Store → NVM*
- Multi-faceted indexes.

N-Store — Possibility #2

- Ditch DRAM.
- Instant recovery:
  - *No snapshots.*
  - *No logging.*
- NVM-friendly concurrency control scheme.

What do we need?

- Be able to tune the msync latency of the NVM SDV.
- Custom malloc implementation that runs entirely of PMFS.

Justin
DeBrabant

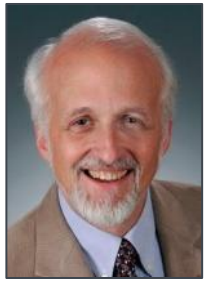
Joy
Arulraj

Rajesh
Sankaran

Subramanya
Dulloor

Andy
Pavlo

Mike
Stonebraker



Col. Stan
Zdonik

Jeff
Parkhurst

intel

MIT

BROWN

UNIVERSITY OF CALIFORNIA

Carnegie Mellon University

W

Portland State UNIVERSITY

END

@ANDY_PAVLO