

FOEDUS: OLTP Engine for 1,000 Cores & NVRAM



Hideaki Kimura

<hideaki.kimura@hpe.com>

HP Labs

HPTS 2015 Lightning Talk

Next-Generation Server Hardware



HP
The
Machine



UC
Berkeley
Firebox



Intel
Rack Scale
Architecture

...

Differences

- HP Photonics? Infiniband? QPI?
- HP Memristor? PCM? Intel 3D XP?
- ...

Commonalities

- 1,000s of Cores
- TB/PBs of NVRAM
- Fast Interconnect

FOEDUS: OLTP Engine for 1,000 Cores & NVRAM

- ✓ From-Scratch, Open-Source, **Fully ACID/Serializable** Database Kernel in C++
- ✓ **Orders of Magnitude** Faster on *Next-Gen. H/W*
- ✓ Interested? *Read paper (*)! Talk to us! Join us (#)!*

(*) [Kimura,
SIGMOD'15]

(#) <http://github.com/hkimura/foedus>

FOEDUS Key Principles

1. Extreme Decentralization

Lock-free Techniques and SILO-like Lightweight OCC

2. Dual-Page: physically independent, logically equivalent

- ✓ Mutable Volatile Pages in DRAM
- ✓ Immutable Snapshot Pages in NVRAM
- ✓ Log Gleaner to Keep them in sync

3. Master-Tree: Simple and Scalable OCC for NVRAM

***Talk to Me or Read Paper,
OK?***

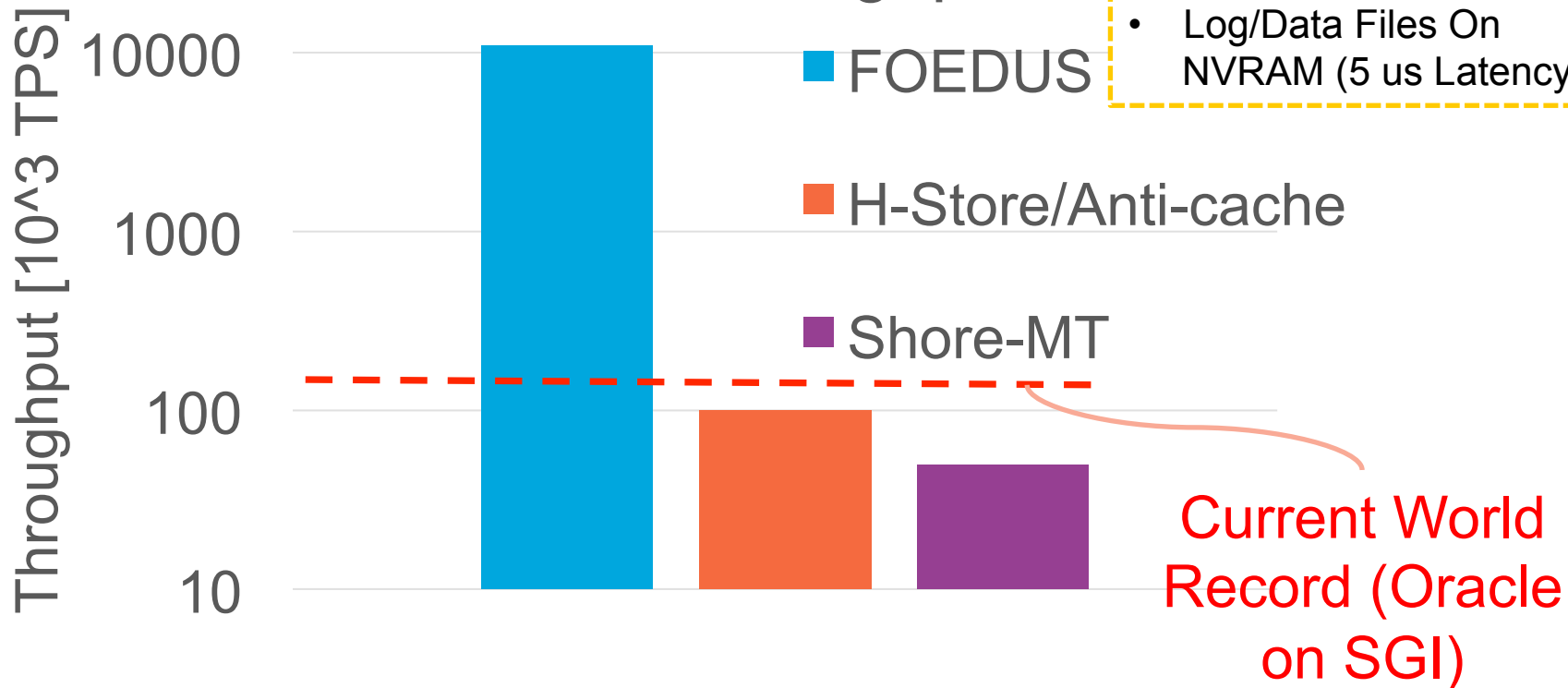
Masstree + Foster B-Tree + Foster-Twin

How Scalable is FOEDUS?

Environment

- 16 Sockets 240 Cores
- 12 TB DRAM (DDR3)
- Log/Data Files On NVRAM (5 us Latency)

TPC-C Throughput



Even Faster than pure In-memory DBMS!

Observations

1. SILO-style more resilient content
2. 100x~ faster than Store

Hideaki's Challenge/Litmus test

TPC-C 17 Million TPS

No Plateauing on 288 Core

Dear OLTP Folks,

Could you PLEASE BEAT US?

or at least get to the same orders of magnitude.

Cores

16

60

240

2

FOEDU
S

H-Store

Remote Transactions [%]

