SO YOU WANT TO FORK POSTGRESQL

Joy Arulraj, Carnegie Mellon University
HPTS 2015

(Disclaimer: Advised by Andy Pavlo)
Guide for forking PostgreSQL

- Difficult process
- Not much documentation available
Option 1: The Pat Helland Way

- Hire core PostgreSQL developer
- Team of 30 engineers
Option 2: Poor Man’s Postgres Fork

- This summer with 3 interns
- Goal: Easier to develop
  - Multi-threaded model
  - Port to C++
  - Integrate custom engine
Multi-threaded Model

- Multi-process to multi-threaded model

**Problem:**
No Copy-on-Write Semantics

**Solution:**
Use Postgres support for Win32

<table>
<thead>
<tr>
<th>Process 1</th>
<th>Process 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>x = 5</td>
<td>x = 5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Thread 1</th>
<th>Thread 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>x = 5</td>
<td>x = 5</td>
</tr>
</tbody>
</table>
• Port to C++

Problem: C-style casting

Solution: Inheritance
Integrate Custom Engine

- Optimized for main-memory
- Hybrid workloads
- Machine learning model

**Problem:** Integrate engine

**Solution:** Plan Transformation
Peloton

- Multi-threaded C++ Postgres fork
- Distributed in-memory HTAP DBMS
- Check out Peloton @ [http://pelotondb.org](http://pelotondb.org)
- Testbed for NVM research
- Storage and recovery methods [SIGMOD’15]
Thanks to Sponsors

Heaven Hill Distilleries, Inc.