Stanford’s AHA (Agile Hardware) Center

Rick Bahr*, Clark Barrett, Joel Emer, Kayvon Fatahalian, Pat Hanrahan, Mark Horowitz, Fred Kjoelstad, Priyanka Raina, Chris Torng*

9/7/21
Welcome To Our Retreat!

It is great to see everyone again
• We have all been through a very “interesting” year
• I think we all will enjoy talking with each other

Let’s be safe
• Masks are required inside; they are optional outside
• Talks will be inside. Discussion sections will be outside
A Long Time Ago …

(ok, it was only 5 years ago)
Stanford’s AHA (Agile Hardware) Design Center

Clark Barrett, Pat Hanrahan, Mark Horowitz

10/24/16
Our Silicon Future

Will see tremendous innovative uses of computation
- Capability of today’s technology is incredible
- Can add computing and communication for nearly $0
- Key challenge is to allow more users access to hardware design

Most performance system will be energy limited
- These systems will be optimized for energy efficiency
- At fixed power, more ops/sec requires lower energy/op
  - Technology is no longer providing the needed reductions
  - We will need to tailor hardware toward applications
Making Hardware Design More Friendly
Addresses Both Issues
Project Goal Is to Reduce Design Cost

Admit that complex systems are, well, complex
• They take a long time to debug
• Often don’t serve the function as well as you expect

Leverage software approaches to deal with these issues
• Rapid prototyping/deployment
• AKA Agile Design
Philosophy: Generalize from Examples

Initial driving application: Programmable Image/Vision System
- Working with Halide DSL
- Generate complete working hardware/software system
- Worry about usability (debugging, performance tuning)

Initial hardware approach: Coarse Grain Reconfigurable Array
- Create a SoC with this CGRA array attached
- Evolve the compiler, tools, and CGRA as we learn
We Have Come A Long Way …
Leadership Team

Rick Bahr*, Clark Barrett, Joel Emer, Kayvon Fatahalian, Pat Hanrahan, Mark Horowitz, Fred Kjoelstad, Priyanka Raina, Chris Torng*

Meet weekly to talk about research ideas and program
Created Applications to Silicon Flow
Created Application to Silicon Flow

Actually created two interconnected flows:
• Hardware generation flow
• Application mapping to generated hardware

Key insight:
• Generators with parameters isn’t enough
• Need to “understand” parameters to tie systems together

DSLs provide single source of truth (SST)
Amber

ARM M3 CPU

Global Buffer
16 256kB Banks

CGRA 16x32
128 MEM, 384 PE
We Are Still Far From Desired Goal

Democratize chip design
• Allow almost any programmer / designer to create a chip
• Allow experts to be more efficient in chip creation

Enable continued innovation as technology progress slows
• Otherwise only companies like you will create chips
In Agile Design, One Continuously Learns

Great news:
• The problem we are working on is more important than every

New Issues:
• Broaden application set
• SoC construction issues
Many Interesting Problems Ahead

Prove our continual improvement strategy
• Use DSE to create an improved design

Expand application domain
• Look at sparse application

Improve SoC generation
• Smart Components
Meeting Organization

We tried to leave a lot of time for discussions
• Breaks / food / discussion sections

The talks will be on zoom and recorded
• If you dialed in to zoom, please be muted unless you want speak
• If you want to speak raise your hand, and we will call on you

The discussion sections will not be on zoom
## Agenda - Today

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Wednesday</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4:00 - 4:15</td>
<td>Welcome</td>
<td>Mark Horowitz</td>
</tr>
<tr>
<td>4:15 - 5:00</td>
<td>Reflections on 4 generations of Google TPUs</td>
<td>Dave Patterson</td>
</tr>
<tr>
<td>5:00 - 5:30</td>
<td>Lightning Talks</td>
<td>Students</td>
</tr>
<tr>
<td>5:30 - 6:30</td>
<td>Poster Session</td>
<td>Students</td>
</tr>
<tr>
<td>6:30 - 8:00</td>
<td>Dinner</td>
<td>ALL</td>
</tr>
<tr>
<td></td>
<td>Walk to Beach</td>
<td>ALL</td>
</tr>
</tbody>
</table>
## Tomorrow Morning

<table>
<thead>
<tr>
<th>Time</th>
<th>Topic</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00 - 9:10</td>
<td>Day 2 Introduction</td>
<td>Priyanka Raina</td>
</tr>
<tr>
<td>9:10 - 9:40</td>
<td>Overview of Amber Chip</td>
<td>Alex Carsello</td>
</tr>
<tr>
<td>9:40 - 10:10</td>
<td>Tool Chain / Design Space Exploration</td>
<td>Jack Melchert</td>
</tr>
<tr>
<td>10:00 - 10:10</td>
<td>Break</td>
<td></td>
</tr>
<tr>
<td>10:30 - 10:45</td>
<td>Virtualization</td>
<td>Priyanka Raina</td>
</tr>
<tr>
<td>10:45 - 11:00</td>
<td>Validation</td>
<td>Clark Barrett</td>
</tr>
<tr>
<td>11:00 - 11:45</td>
<td>How to Democratize Design:</td>
<td></td>
</tr>
<tr>
<td>11:45 - 12:10</td>
<td>Breakout Session Reports</td>
<td></td>
</tr>
<tr>
<td>12:10 - 1:30</td>
<td>Lunch</td>
<td></td>
</tr>
</tbody>
</table>
## Tomorrow Afternoon

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:30-2:00</td>
<td>Smart Components / Use of Session Types</td>
<td>Lenny Truong</td>
</tr>
<tr>
<td>2:00-2:30</td>
<td>Accelerating Sparse Applications</td>
<td>Fred Kjolstad</td>
</tr>
<tr>
<td>3:00 - 3:45</td>
<td>Breakout Sessions: AHA Research Agenda Feedback</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>DSE Feedback</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>Virtualization Feedback</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>Smart component feedback</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>Sparse application acceleration feedback</td>
<td>--</td>
</tr>
<tr>
<td>3:45 - 4:10</td>
<td>Breakout Session Reports</td>
<td>Mark Horowitz</td>
</tr>
<tr>
<td>4:10- 4:40</td>
<td>Closing Thoughts &amp; AHA's Research Intentions</td>
<td></td>
</tr>
</tbody>
</table>