Dealing with complex test suites

LPC 2019 - *Testing and fuzzing micro-conference*  
10\textsuperscript{th} September 2019  
Guillaume Tucker  
gtucker@collabora.com
Once upon a time...
There was a little project

- kernelci.org started in 2015 around ARM ecosystem
- Doing only kernel builds initially
- Then gradually adding boot tests
- ...and many more boot tests...
But things got complicated

- In 2018 started to run automated boot bisections
  - Effective reports to patch authors & maintainers
- Now starting to run functional tests: v4l2, igt, suspend
  - Extending test coverage to more subsystems
- How do we cope with the results? Bisect them?
  - Help!
Test suites went running wild

- Some have a large number of test cases: IGT, LTP
- Testing every commit in linux-next is not practical
- When something fails, a bisection is needed
- Test cases can all start passing/failing independently
- `git bisect` only tracks one result per revision
- Reporting regressions can become messy
Meanwhile, at Intel...
EzBench and the origin of gfx-ci

- Started in 2015 by Martin Peres
  https://gitlab.freedesktop.org/ezbench/ezbench.git
- Used in automated testing between 2016 and 2018
- Has many powerful features:
  - Test scheduler
  - Report generator
  - **Bisection for test suites**
    as well as benchmarks and rendering
The challenge was clear

- Core bisection logic embedded within EzBench
- Similar problem to solve in KernelCI
- Needed to prove it could be done
- It’s only typing!
Many keystrokes later...
A new tool was made: scalpel

- Proof-of-concept tool inspired by EzBench
  https://gitlab.collabora.com/gtucker/scalpel.git

- Aiming to be generic better-than-git bisect

- Ready to start becoming part of KernelCI
With a running demo – start

History:
- a35e81a85f4c master.9
- f2d9fdce68bc master.8
- 1dd19aae11a master.7
- 4397a95c6a51 master.6
- cb8c76131b27 Merge branch 'issue'
- 1913f2fbb1b0 master.5
- 7526332c820f master.4
- e9732769d500 master.3
- 6a02f9c86120 master.issue.3
- 0e6ef1c559af master.issue.2
- 85a814796514 master.issue.1
- 3b6875a6d0a0 master.2
- cf41748b3d2f master.1

Next revisions and tests to run:
* cb8c76131b27564ae0dc5e5b7db075c19afe5620 'foo'
* 3b6875a6d0a02e305381dbbccbd747bf298f6d2b 'foo' (merge base for branch 'issue')
* 3b6875a6d0a02e305381dbbccbd747bf298f6d2b 'bar' (merge base for branch 'issue')
With a running demo - step 1

History:

a35e81a85f4c master.9 | foo 5.300 5.349 5.262 | bar 1.830 1.846 1.846 Reported
f2d9fdce68bc master.8
1dd19aaea11a master.7
4397a95c6a51 master.6
cb8c76131b27 Merge branch 'issue' | foo 6.200 6.200 6.200
1913f2fbb1b0 master.5
7526332c820f master.4 | foo 6.200 6.197 6.220 | bar 1.150 1.144 1.154 Reported
e9732769d500 master.3
6a02f9c86120 master.issue.3
0e6ef1c559af master.issue.2 | foo 5.100 5.130 5.133 | bar 1.830 1.824 1.817 Reported
85a814796514 master.issue.1
3b6875a6d0a0 master.2 | foo 5.100 5.147 5.146 | bar 1.150 1.144 1.151
cf41748b3d2f master.1

Next revisions and tests to run:
* 1dd19aaea11a494427175f02e0a47c3e8c4a7251 'foo'
* 85a814796514a18a5bd8803aa0016f406f3bb132 'bar'
With a running demo - step 2

History:

<table>
<thead>
<tr>
<th>Commit SHA</th>
<th>Branch</th>
<th>S 1</th>
<th>S 2</th>
<th>S 3</th>
<th>S 4</th>
<th>B 1</th>
<th>B 2</th>
<th>B 3</th>
<th>Reported</th>
</tr>
</thead>
<tbody>
<tr>
<td>a35e81a85f4c</td>
<td>master.9</td>
<td>5.300</td>
<td>5.349</td>
<td>5.262</td>
<td></td>
<td>1.830</td>
<td>1.846</td>
<td>1.846</td>
<td>Reported</td>
</tr>
<tr>
<td>f2d9fdce68bc</td>
<td>master.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1dd19aeea11a</td>
<td>master.7</td>
<td>5.300</td>
<td>5.266</td>
<td>5.317</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4397a95c6a51</td>
<td>master.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cb8c76131b27</td>
<td>master.5</td>
<td>6.200</td>
<td>6.200</td>
<td>6.200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1913f2fbb1b0</td>
<td>master.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7526332c820f</td>
<td>master.3</td>
<td>6.200</td>
<td>6.197</td>
<td>6.220</td>
<td></td>
<td>1.150</td>
<td>1.144</td>
<td>1.154</td>
<td></td>
</tr>
<tr>
<td>0e6ef1c559af</td>
<td>master.issue.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e9732769d500</td>
<td>master.issue.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6a02f9c86120</td>
<td>master.issue.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>85a814796514</td>
<td>master.issue.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3b6875a6d0a0</td>
<td>master.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cf41748b3d2f</td>
<td>master.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Next revisions and tests to run:
* 4397a95c6a5154705154823f56bb09e314d48353 'foo'
* e9732769d5001091930da32d25739079c1dd3cdf 'foo'
With a running demo - step 3

History:

<table>
<thead>
<tr>
<th>Commit</th>
<th>Author</th>
<th>Message</th>
<th>Date</th>
<th>Time</th>
<th>Commit ID</th>
<th>Status</th>
<th>Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>a35e81a85f4c</td>
<td>master.9</td>
<td>foo 5.300 5.349 5.262</td>
<td>bar 1.830 1.846 1.846</td>
<td>Reported</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f2d9fdce68bc</td>
<td>master.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1dd19aae11a</td>
<td>master.7</td>
<td>foo 5.300 5.266 5.317</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4397a95c6a51</td>
<td>master.6</td>
<td>foo 6.200 6.183 6.198</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cb8c76131b27</td>
<td>Merge branch 'issue'</td>
<td>foo 6.200 6.200 6.200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1913f2fbb1b0</td>
<td>master.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7526332c820f</td>
<td>master.4</td>
<td>foo 6.200 6.197 6.220</td>
<td>bar 1.150 1.144 1.154</td>
<td>Reported</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e9732769d500</td>
<td>master.3</td>
<td>foo 5.100 5.107 5.093</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6a02f9c86120</td>
<td>master.issue.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0e6ef1c559af</td>
<td>master.issue.2</td>
<td>foo 5.100 5.130 5.133</td>
<td>bar 1.830 1.824 1.817</td>
<td>Reported</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>85a814796514</td>
<td>master.issue.1</td>
<td></td>
<td>bar 1.830 1.834 1.839</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3b6875a6d0a0</td>
<td>master.2</td>
<td>foo 5.100 5.147 5.146</td>
<td>bar 1.150 1.144 1.151</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cf41748b3d2f</td>
<td>master.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Commits found:
* ('7526332c820fe3c221169cc7a44c6c65ce2284d5', 'foo')
* ('1dd19aae11a494427175f02e0a47c3e8c4a7251', 'foo')
* ('85a814796514a18a5bd8803aa0016f406f3bb132', 'bar')
With a running demo – history

<table>
<thead>
<tr>
<th>Commit</th>
<th>Branch</th>
<th>foo</th>
<th>bar</th>
<th>Reported</th>
</tr>
</thead>
<tbody>
<tr>
<td>a35e81a85f4c</td>
<td>master.9</td>
<td>5.300 5.349 5.262</td>
<td>1.830 1.846 1.846</td>
<td></td>
</tr>
<tr>
<td>f2d9fdce68bc</td>
<td>master.8</td>
<td>5.300 5.343 5.320</td>
<td>1.830 1.818 1.837</td>
<td></td>
</tr>
<tr>
<td>1dd19aae11a</td>
<td>master.7</td>
<td>5.300 5.266 5.317</td>
<td>1.830 1.836 1.840</td>
<td></td>
</tr>
<tr>
<td>cb8c76131b27</td>
<td>master.6</td>
<td>6.200 6.183 6.198</td>
<td>1.830 1.831 1.841</td>
<td></td>
</tr>
<tr>
<td>1913f2fbb1b0</td>
<td>master.5</td>
<td>6.200 6.151 6.225</td>
<td>1.150 1.161 1.139</td>
<td></td>
</tr>
<tr>
<td>7526332c820f</td>
<td>master.4</td>
<td>6.200 6.197 6.220</td>
<td>1.150 1.144 1.154</td>
<td></td>
</tr>
<tr>
<td>e9732769d500</td>
<td>master.3</td>
<td>5.100 5.107 5.093</td>
<td>1.150 1.149 1.151</td>
<td></td>
</tr>
<tr>
<td>6a02f9c86120</td>
<td>master.issue.3</td>
<td>5.100 5.072 5.147</td>
<td>1.830 1.829 1.825</td>
<td></td>
</tr>
<tr>
<td>85a814796514</td>
<td>master.issue.1</td>
<td>5.100 5.130 5.133</td>
<td>1.830 1.824 1.817</td>
<td></td>
</tr>
<tr>
<td>3b6875a6d0a0</td>
<td>master.2</td>
<td>5.100 5.147 5.146</td>
<td>1.150 1.144 1.151</td>
<td></td>
</tr>
<tr>
<td>cf41748b3d2f</td>
<td>master.1</td>
<td>5.100 5.135 5.126</td>
<td>1.150 1.142 1.158</td>
<td></td>
</tr>
</tbody>
</table>

Commits found:
* ('7526332c820fe3c221169cc7a44c6c65ce2284d5', 'foo')
* ('1dd19aae11a494427175f02e0a47c3e8c4a7251', 'foo')
* ('85a814796514a18a5bd8803aa0016f406f3bb132', 'bar')
This is not the end
Next steps

• Python package to use in multiple projects
  - Initially in EzBench and KernelCI
• New KernelCI bisection implementation
  - Portable kci_bisect tool rather than pure Jenkins job
• See how this evolves
  - More shared tools, maybe improve git bisect?
Also needed

- Email reports
  - Details of issues found, but also low noise
- Tracking of known issues across revisions
  - Bug trackers, dashboards with test results
- Debugging
  - Facilitate reproducing the issue, debug information
Stay tuned!