Towards Automated RISC-V Microarchitecture Design with Reinforcement Learning

Chen Bai\textsuperscript{1}, Jianwang Zhai\textsuperscript{2}, Yuzhe Ma\textsuperscript{3}, Bei Yu\textsuperscript{1}, Martin D.F. Wong\textsuperscript{4}

\textsuperscript{1}The Chinese University of Hong Kong
\textsuperscript{2}Beijing University of Posts and Telecommunications
\textsuperscript{3}The Hong Kong University of Science and Technology (Guangzhou)
\textsuperscript{4}Hong Kong Baptist University

January 6, 2024
Outline

1 Introduction
2 Preliminaries
3 Reinforcement Learning Methodology
4 Experiments
5 Conclusion & Discussions
Introduction
Problem Formulation

Given the microarchitecture design space and target workloads, how do we efficiently search for optimal microarchitectures that can satisfy the pre-determined performance, power, and area (PPA) design targets?

An example microprocessor’s design space for SPEC2K6

An example of the microarchitecture design space exploration.
An overview of the example microprocessor microarchitecture, including different components.
Related Work

- Industry: computer architects’ expertise.
- Academia:
  - Analytical methodology: interpretable PPA models, explainable search strategy, etc.\textsuperscript{12}
  - Black-box methodology: machine-learning-based PPA modeling and search strategy.\textsuperscript{345}

Limitations:

- Industry solution: architects’ personal bias can yield sub-optimal solutions.
- Academic solution: not tightly coupled with expert knowledge & mathematical limitation in the Gaussian process modeling.


\textsuperscript{5} Dandan Li et al. (2016). “Efficient design space exploration via statistical sampling and AdaBoost learning”. In: ACM/IEEE Design Automation Conference (DAC), pp. 1–6.
The kernel function of the Gaussian process mathematically attributes the PPA differences between two microarchitectures to the microarchitecture embedding distances.

An example of different BOOM microarchitectures to demonstrate the claim.

---

Highlights of our new black-box methodology:

- Remove mathematical limitation in the Gaussian process modeling (i.e., free of unrealistic assumptions).
- Tightly coupled with expert knowledge: *microarchitecture scaling graph*.
- PPA design preference-driven exploration.
- Lightweight agent training environment design to accelerate the learning process.
Preliminaries
Table 1: RISC-V Microarchitecture Design Space

<table>
<thead>
<tr>
<th>Design</th>
<th>Component</th>
<th>Parameters</th>
<th>Candidate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rocket</td>
<td>Branch predictor</td>
<td>RAS: 0:12:3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BTB.nEntries: 0:56:14</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BHT.nEntries: 0:1024:256</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I-cache</td>
<td>nWays: 1,2,4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>nTLBWay: 4:32:4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Functional unit</td>
<td>FPU: 1,2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>mulDiv: 1,2,3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>VM: 1,2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D-cache</td>
<td>nSets: 32,64</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>nWays: 1,2,4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>nTLBWay: 4:32:4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>nMSHRs: 1,2,3</td>
<td></td>
</tr>
<tr>
<td>Small/Medium</td>
<td>Branch predictor</td>
<td>maxBrCount: 4:22:2</td>
<td></td>
</tr>
<tr>
<td>Large/Mega</td>
<td>IFU</td>
<td>numFetchBufferEntries: 6:46:2</td>
<td></td>
</tr>
<tr>
<td>Giga SonicBOOM</td>
<td></td>
<td>fetchWidth: 4,8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ftq.nEntries: 12:64:4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>pipelineWidth: 1:5:1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ROB</td>
<td>24:160:4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PRF</td>
<td>numIntPhysRegisters: 40:176:8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>numFpPRF: 34:132:6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ISU</td>
<td>numPhysRegisters: 1:5:1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>numEntries: 6:52:2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>dispatchWidth: 1:5:1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LSU</td>
<td>LDQ: 6:32:2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>STQ: 6:36:2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I-cache</td>
<td>nWays: 4,8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>nSets: 32,64</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D-cache</td>
<td>nWays: 4,8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>nSets: 64,128</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>nMSHRs: 2:10:2</td>
<td></td>
</tr>
</tbody>
</table>

* The values are start number:end number:stride, e.g., 0:12:3 denotes the entries of RAS can be 0, 3, 6, etc., until 12.
Microarchitecture PPA Modeling

- Pre-RTL PPA modeling infrastructure:
  - GEM5\textsuperscript{7}, McPAT\textsuperscript{9}, etc.

- RTL PPA modeling infrastructure:
  - Synopsys VCS\textsuperscript{10}, Synopsys PrimeTime PX\textsuperscript{11}, etc.


A microarchitecture scaling graph of an example out-of-order microprocessor.

- The microarchitecture scaling graph is a directed graph. Nodes are components. Edges are components’ decision precedences\(^\text{12}\).
- The graph is derived from the interval analysis. And the conclusion from the graph is general for Von Neumann architecture.

Reinforcement Learning Methodology
An overview of our reinforcement learning methodology.

- The state space is the microarchitecture design space.
- The action space is the candidate set of components’ types of corresponding hardware resources.
- The state transition is defined based on the microarchitecture scaling graph.
- Embed the architect’s PPA design preference in the methodology.
Reward scalarization:

\[ r = r(\text{Perf, Power, Area}) \cdot (\alpha, \beta, \gamma)^\top \]  

(1)

\(\alpha, \beta, \) and \(\gamma\) are weights controlling the PPA trade-off.

Simplex constraints for PPA design preference \(\phi\):

\[ r = r(\text{Perf, Power, Area}) \cdot \phi^\top \]
\[ \forall i, \ \phi_i \geq 0, \ \sum_i \phi_i = 1 \]  

(2)

\(^{13}\)Perf, Power, Area are normalized values.
Generalized Bellman optimality equality:

\[ Q(s, a, \phi) = \frac{1}{\text{Power}} \]

\[ r(s, a) + \zeta \mathbb{E}_{s' \sim \mathcal{P}(.|s, a)} \mathcal{T}(Q(s', a, \phi)), \]

\[ \mathcal{T}(Q(s', a, \phi)) = \arg \max_{Q} \max_{s' \in A, \phi' \in \Phi} Q(s', a', \phi') \phi'^\top, \]

(3)

\( \zeta \) is the discount factor, \( Q(s, a, \phi) \) is the state-action vector, and \( \phi \) is the PPA design preference.

\[ \mathcal{T}(Q(s, a, \phi)) = Q(s, a_2, \phi) \]

\[ = \arg \max_{Q} \left\{ Q(s, a_1, \phi) \phi^\top_1, Q(s, a_1, \phi) \phi^\top_2 \right\} \]

Optimization procedure with the generalized Bellman optimality equality.
We adopt the asynchronous advantage actor-critic (A3C)\textsuperscript{15}.

We utilize the conditioned neural network design\textsuperscript{16}.

Gradients of the actor weights $\theta_a$:

$$\nabla \theta_a = \kappa \nabla \theta_a H(\pi(s_t; \theta_a)) +$$

$$\mathbb{E}_{\xi \sim \pi}[\sum_{t=0}^{\infty} \nabla \theta_a \log \pi_{\theta_a}(a_t | s_t)A(s_t, a_t, \phi')\phi^\top],$$

(4)

Loss function of the critic:

$$L_c = \rho \|(Q^* - Q(s, a, \phi'; \theta_c))\phi^\top\|_2^2 +$$

$$(1 - \rho)\|Q^* - Q(s, a, \phi'; \theta_c)\|_2^2,$$

(5)


An overview of the PPA calibration.

- PPA models calibration flow\textsuperscript{17}.
- PPA models update policy.

Experiments
Experiments Environment

- RTL implementation: Rocket & BOOM\textsuperscript{18,19}.
- Technology nodd: 7-nm ASAP7 PDK\textsuperscript{20}.
- EDA tools: Cadence Genus 18.12-e012_1, Synopsys VCS M-2017.03, PrimeTime PX R-2020.09-SP1, etc.
- Server: 80 × Intel(R) Xeon(R) CPU e7-4803 v2 @ 2.20GHz, 1TB main memory.

Baselines

- ISCA’14: ArchRanker\textsuperscript{21}
- DAC’16: AdaBoost\textsuperscript{22}
- ICCAD’21: BOOM-Explorer\textsuperscript{23}

\textsuperscript{22}Dandan Li et al. (2016). “Efficient design space exploration via statistical sampling and AdaBoost learning”. In: ACM/IEEE Design Automation Conference (DAC), pp. 1–6.
The accuracy of lightweight PPA models, and MAPE and Kendall $\tau$ curves w.r.t. the calibration data size.
(a) Rocket

(b) Small-scale BOOM

(c) Medium-scale BOOM
(a) Large-scale BOOM

(b) Mega-scale BOOM

(c) Giga-scale BOOM
## Table 2: Comparison w. Human Efforts & Prior Arts

<table>
<thead>
<tr>
<th>Design</th>
<th>Method</th>
<th>Performance IPC</th>
<th>Power W</th>
<th>Area mm²</th>
<th>Perf / Power Val.</th>
<th>Perf / Power Ratio</th>
<th>Perf / Area Val.</th>
<th>Perf / Area Ratio</th>
<th>(Perf x Perf) / (Power x Area) Val.</th>
<th>Runtime</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rocket</strong></td>
<td>Human Efforts</td>
<td>0.7338</td>
<td>0.0027</td>
<td>0.9082</td>
<td>267.4708</td>
<td>1.0000</td>
<td>1.0000</td>
<td>216.1090</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ISCA'14</td>
<td>0.8157</td>
<td>0.0023</td>
<td>0.7943</td>
<td>359.3222</td>
<td>1.3434 ×</td>
<td>1.0270</td>
<td>369.0075</td>
<td>1.7075 ×</td>
<td>8.6111 ×</td>
</tr>
<tr>
<td></td>
<td>DAC'16</td>
<td>0.5485</td>
<td>0.0018</td>
<td>0.5337</td>
<td>305.3090</td>
<td>1.1415 ×</td>
<td>1.0278</td>
<td>313.8042</td>
<td>1.4527 ×</td>
<td>5.8961 ×</td>
</tr>
<tr>
<td></td>
<td>ICCAD'21</td>
<td>0.7278</td>
<td>0.0021</td>
<td>0.7448</td>
<td>352.7717</td>
<td>1.3187 ×</td>
<td>0.9797</td>
<td>344.6327</td>
<td>1.5947 ×</td>
<td>1.5011 ×</td>
</tr>
<tr>
<td></td>
<td>Ours</td>
<td>0.7278</td>
<td>0.0023</td>
<td>0.5762</td>
<td>313.6958</td>
<td>1.1728 ×</td>
<td>1.2631</td>
<td>396.2335</td>
<td>1.8335 ×</td>
<td>1.0000</td>
</tr>
<tr>
<td><strong>Small SonicBOOM</strong></td>
<td>Human Efforts</td>
<td>0.7837</td>
<td>0.0203</td>
<td>1.5048</td>
<td>38.6057</td>
<td>1.0520</td>
<td>20.1062</td>
<td>34.9710</td>
<td>1.7393 ×</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ISCA'14</td>
<td>0.8197</td>
<td>0.0150</td>
<td>1.2838</td>
<td>54.7692</td>
<td>1.4187 ×</td>
<td>0.6385</td>
<td>35.3765</td>
<td>1.7954 ×</td>
<td>4.7918 ×</td>
</tr>
<tr>
<td></td>
<td>DAC'16</td>
<td>0.8076</td>
<td>0.0147</td>
<td>1.2152</td>
<td>54.8119</td>
<td>1.4198 ×</td>
<td>0.6454</td>
<td>35.3765</td>
<td>1.7954 ×</td>
<td>4.7918 ×</td>
</tr>
<tr>
<td></td>
<td>ICCAD'21</td>
<td>0.8469</td>
<td>0.0200</td>
<td>1.5026</td>
<td>42.3436</td>
<td>1.0968 ×</td>
<td>0.5636</td>
<td>23.8645</td>
<td>1.1869 ×</td>
<td>1.3053 ×</td>
</tr>
<tr>
<td></td>
<td>Ours</td>
<td>0.8403</td>
<td>0.0152</td>
<td>1.2538</td>
<td>55.2813</td>
<td>1.4320 ×</td>
<td>0.6702</td>
<td>37.0491</td>
<td>1.8427 ×</td>
<td>1.0000</td>
</tr>
<tr>
<td><strong>Medium SonicBOOM</strong></td>
<td>Human Efforts</td>
<td>1.1938</td>
<td>0.0256</td>
<td>1.9332</td>
<td>46.6952</td>
<td>1.6175</td>
<td>28.8363</td>
<td>26.5292</td>
<td>1.2817 ×</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ISCA'14</td>
<td>1.2362</td>
<td>0.0196</td>
<td>1.6242</td>
<td>62.9622</td>
<td>1.3484 ×</td>
<td>0.7611</td>
<td>47.9192</td>
<td>1.6618 ×</td>
<td>5.6879 ×</td>
</tr>
<tr>
<td></td>
<td>DAC'16</td>
<td>1.3757</td>
<td>0.0254</td>
<td>1.9247</td>
<td>54.0894</td>
<td>1.1584 ×</td>
<td>0.7148</td>
<td>38.6609</td>
<td>1.3407 ×</td>
<td>4.6966 ×</td>
</tr>
<tr>
<td></td>
<td>ICCAD'21</td>
<td>1.4454</td>
<td>0.0271</td>
<td>2.1583</td>
<td>53.3342</td>
<td>1.1422 ×</td>
<td>0.6697</td>
<td>35.7170</td>
<td>1.2386 ×</td>
<td>1.2793 ×</td>
</tr>
<tr>
<td></td>
<td>Ours</td>
<td>1.2872</td>
<td>0.0206</td>
<td>1.7351</td>
<td>62.5886</td>
<td>1.3404 ×</td>
<td>0.7419</td>
<td>46.4339</td>
<td>1.6103 ×</td>
<td>1.0000</td>
</tr>
<tr>
<td><strong>Large SonicBOOM</strong></td>
<td>Human Efforts</td>
<td>1.4871</td>
<td>0.0446</td>
<td>3.2055</td>
<td>33.3430</td>
<td>0.4639</td>
<td>15.4686</td>
<td>28.2626</td>
<td>1.8271 ×</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ISCA'14</td>
<td>1.4900</td>
<td>0.0309</td>
<td>2.5240</td>
<td>48.2184</td>
<td>1.4461 ×</td>
<td>0.5861</td>
<td>28.2626</td>
<td>1.8271 ×</td>
<td>5.8920 ×</td>
</tr>
<tr>
<td></td>
<td>DAC'16</td>
<td>1.4919</td>
<td>0.0324</td>
<td>2.6744</td>
<td>45.9976</td>
<td>1.3795 ×</td>
<td>0.5578</td>
<td>25.6592</td>
<td>1.6588 ×</td>
<td>4.8651 ×</td>
</tr>
<tr>
<td></td>
<td>ICCAD'21</td>
<td>1.9162</td>
<td>0.0409</td>
<td>3.6715</td>
<td>46.8507</td>
<td>1.4051 ×</td>
<td>0.5219</td>
<td>24.4520</td>
<td>1.5808 ×</td>
<td>1.3252 ×</td>
</tr>
<tr>
<td></td>
<td>Ours</td>
<td>1.5882</td>
<td>0.0314</td>
<td>2.5634</td>
<td>50.6324</td>
<td>1.5185 ×</td>
<td>0.6193</td>
<td>31.3580</td>
<td>2.0272 ×</td>
<td>1.0000</td>
</tr>
<tr>
<td><strong>Mega SonicBOOM</strong></td>
<td>Human Efforts</td>
<td>1.9500</td>
<td>0.0578</td>
<td>4.8059</td>
<td>33.7571</td>
<td>0.4058</td>
<td>13.6972</td>
<td>20.5020</td>
<td>1.4968 ×</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ISCA'14</td>
<td>2.4957</td>
<td>0.0566</td>
<td>5.3676</td>
<td>44.0942</td>
<td>1.3062 ×</td>
<td>0.4650</td>
<td>20.6513</td>
<td>1.5077 ×</td>
<td>4.5780 ×</td>
</tr>
<tr>
<td></td>
<td>DAC'16</td>
<td>2.4995</td>
<td>0.0562</td>
<td>5.3979</td>
<td>44.4843</td>
<td>1.3167 ×</td>
<td>0.4646</td>
<td>20.6513</td>
<td>1.5077 ×</td>
<td>4.5780 ×</td>
</tr>
<tr>
<td></td>
<td>ICCAD'21</td>
<td>2.4823</td>
<td>0.0607</td>
<td>4.7008</td>
<td>40.9170</td>
<td>1.2121 ×</td>
<td>0.5281</td>
<td>21.6066</td>
<td>1.5774 ×</td>
<td>1.2470 ×</td>
</tr>
<tr>
<td></td>
<td>Ours</td>
<td>2.5232</td>
<td>0.0557</td>
<td>5.2512</td>
<td>45.3005</td>
<td>1.3420 ×</td>
<td>0.4805</td>
<td>21.7674</td>
<td>1.5892 ×</td>
<td>1.0000</td>
</tr>
<tr>
<td><strong>Giga SonicBOOM</strong></td>
<td>Human Efforts</td>
<td>1.8717</td>
<td>0.0716</td>
<td>5.0691</td>
<td>26.1538</td>
<td>0.3692</td>
<td>9.6572</td>
<td>13.5970</td>
<td>1.4080 ×</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ISCA'14</td>
<td>2.2528</td>
<td>0.0622</td>
<td>6.0010</td>
<td>36.2192</td>
<td>1.3849 ×</td>
<td>0.3754</td>
<td>13.5970</td>
<td>1.4080 ×</td>
<td>5.6321 ×</td>
</tr>
<tr>
<td></td>
<td>DAC'16</td>
<td>2.2522</td>
<td>0.0733</td>
<td>5.9995</td>
<td>29.1480</td>
<td>1.1145 ×</td>
<td>0.4022</td>
<td>11.7236</td>
<td>1.2140 ×</td>
<td>4.6505 ×</td>
</tr>
<tr>
<td></td>
<td>ICCAD'21</td>
<td>2.2650</td>
<td>0.0745</td>
<td>5.8862</td>
<td>30.4162</td>
<td>1.1630 ×</td>
<td>0.3862</td>
<td>11.7460</td>
<td>1.2163 ×</td>
<td>1.2668 ×</td>
</tr>
<tr>
<td></td>
<td>Ours</td>
<td>2.2692</td>
<td>0.0595</td>
<td>5.7459</td>
<td>38.1587</td>
<td>1.4590 ×</td>
<td>0.3949</td>
<td>15.0696</td>
<td>1.5605 ×</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

1. “−” denotes not applicable.
Analysis w. More Workloads

(a) Rocket

(b) Small-scale BOOM

(c) Medium-scale BOOM

(d) Large-scale BOOM
Summary:

- Our solutions achieve an average of 24.64%, 17.13%, and 6.33% than ICCAD’21, DAC’16, and ISCA’14 in PPA trade-off, respectively with $4.07 \times$ higher efficiency than baselines.

- For large-scale BOOM, compared to human implementations, our solution demonstrates significant improvements in three metrics by factors of $1.35 \times$, $1.23 \times$, and $1.66 \times$, respectively.
Conclusion
• Remove mathematical limitation in the Gaussian process modeling (i.e., free of unrealistic assumptions).

• Tightly coupled with expert knowledge: microarchitecture scaling graph.

• PPA design preference-driven exploration.

• Lightweight agent training environment design to accelerate the learning process.

• Experiments show that our method achieves an average of 16.03% improvement in PPA trade-off with 4.07× higher efficiency than previous state-of-the-art approaches.

Discussion:
• Why do we choose A3C instead of PPO or SAC?
• Why does our method not perform very well on medium-scale SonicBOOM?
• Is the RL method the sole remedy to resolve the mathematical limitation (unrealistic assumption)?
• Remove mathematical limitation in the Gaussian process modeling (i.e., free of unrealistic assumptions).

• Tightly coupled with expert knowledge: microarchitecture scaling graph.

• PPA design preference-driven exploration.

• Lightweight agent training environment design to accelerate the learning process.

• Experiments show that our method achieves an average of 16.03% improvement in PPA trade-off with $4.07 \times$ higher efficiency than previous state-of-the-art approaches.

Discussion:

• Why do we choose A3C instead of PPO or SAC?

• Why does our method not perform very well on medium-scale SonicBOOM?

• Is the RL method the sole remedy to resolve the mathematical limitation (unrealistic assumption)?
THANK YOU!