2024 Design, Automation & Test in Europe Conference (DATE 2024)

RLPlanner: Reinforcement Learning based Floorplanning for Chiplets with Fast Thermal Analysis

Yuanyuan Duan[1], Xingchen Liu[1], Zhiping Yu[2], Hanming Wu[1], Leilai Shao[3]*, Xiaolei Zhu[1]*
[1]School of Micro-Nano Electronics, Zhejiang University, Hangzhou, China
[2]School of Integrated Circuits, Tsinghua University, Beijing, China
[3]School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China

***Abstract*—Chiplet-based systems have gained significant attention in recent years due to their low cost and competitive performance. As the complexity and compactness of a chiplet-based system increase, careful consideration must be given to microbump assignments, interconnect delays, and thermal limitations during the floorplanning stage. This paper introduces *RLPlanner*, an efficient early-stage floorplanning tool for chiplet-based systems with a novel fast thermal evaluation method. *RLPlanner* employs advanced reinforcement learning to jointly minimize total wirelength and temperature. To alleviate the time-consuming thermal calculations, *RLPlanner* incorporates the developed fast thermal evaluation method to expedite the iterations and optimizations. Comprehensive experiments demonstrate that our proposed fast thermal evaluation method achieves a mean absolute error (MAE) of ±0.25 K and delivers over 120x speed-up compared to the open-source thermal solver *HotSpot*. When integrated with our fast thermal evaluation method, *RLPlanner* achieves an average improvement of 20.28% in minimizing the target objective (a combination of wirelength and temperature), within a similar running time, compared to the classic simulated annealing method with *HotSpot*.**

***Index Terms*—reinforcement learning, fast thermal evaluation, chiplet floorplanning**

I. Introduction

To address the increasing cost of large Systems-on-Chip (SoCs) on advanced technology nodes, chiplet-based design or 2.5D integration has emerged as a solution. However, as the complexity and compactness of chiplet-based systems increase, it becomes critical to address issues such as microbump assignments, interconnect delays, and thermomechanical stress during the initial floorplanning stage.

Traditional physical floorplanning of monolithic chips primarily focuses on reducing total wirelength and minimizing area [1], which results in compact floorplans that may lead to potential thermal-induced failures. Recent floorplanning works have started considering the thermal aspect in chiplet-based systems [2]. They employ simple greedy strategies or simulated annealing (SA) to handle chiplet system floorplanning, lacking the flexibility and transferability required for complex 2.5D integrations. Moreover, these optimization methods are constrained by time-consuming thermal evaluations, which hinder fast and efficient exploration of thermal-aware floorplanning.

This work was supported by Pre-research project of ministry foundation (Grant No.31513010501).(* *Corresponding Authors: Leilai Shao (leilaishao@sjtu.edu.cn) and Xiaolei Zhu (xl_zhu@zju.edu.cn)*)

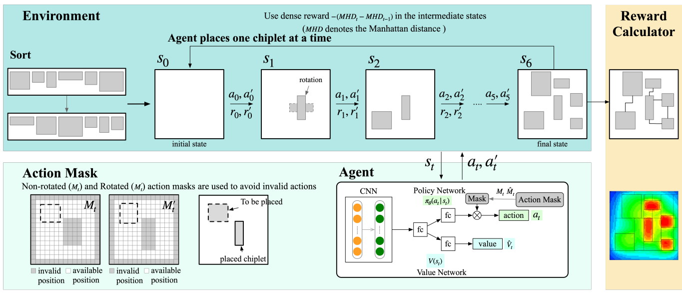

Fig. 1. Overview of *RLPlanner*: The implementation of chiplet floorplanning using reinforcement learning.

Some attempts have been made to use convolutional neural networks (CNN) or graph convolutional networks (GCN) to build surrogate models for accelerated thermal evaluations [3]. However, these models often rely on empirical parameters, such as tile and window sizes, requiring domain knowledge for determination, or show limited speed-ups (<3x), making them impractical for real designs.

In this work, we present *RLPlanner*, a floorplanning tool based on reinforcement learning (RL) for early-stage chiplet-based systems. *RLPlanner* utilizes advanced RL techniques and incorporates a novel fast thermal evaluation method to optimize both the maximum operating temperature and total wirelength of the chiplet system.

II. Methodology

A. Overall Architecture of RLPlanner

As shown in Fig. 1, the overall architecture of *RLPlanner* consists of three main parts: the floorplanning environment for chiplets, the RL-based agent, and a thermal-aware reward calculator. The environment updates the action mask M_t and state s_t based on previously placed chiplets at each time step t. The agent generates the probability matrix of actions $\pi_\theta(a_t|s_t)$ and the expected reward $V(s_t)$ by the policy and value networks, correspondingly. Before sampling the action a_t that places chiplets sequentially, the probability of infeasible actions will set to be '0' based on M_t. Once all chiplets have been placed, the reward calculator will first perform microbump assignments to minimize the total wirelength by allocating pin locations for each inter-chiplet connection, and then conduct the thermal evaluation.

979-8-3503-4859-0/DATE24/© 2024 EDAA

TABLE I

COMPARISONS AGAINST BASELINES ON BENCHMARK SYSTEMS

Design	**Multi-GPU System** [4]				**CPU-DRAM System** [5]				**Ascend 910 System** [6]			
Method	**Reward**	**Wirelength** (mm)	**Temperature** (°C)	**Runtime** (s)	**Reward**	**Wirelength** (mm)	**Temperature** (°C)	**Runtime** (s)	**Reward**	**Wirelength** (mm)	**Temperature** (°C)	**Runtime** (s)
RLPlanner	-37.1263	97742	91.15	20910	-44.9467	176246	92.88	8925	-7.4063	18130	77.12	7773
RLPlanner(RND)	-40.2777	104636	91.85	20380	-41.7496	164460	92.15	8993	-7.4433	18221	76.84	9991
TAP-2.5D(*HotSpot*)	-42.4572	124639	91.68	35211	-60.3570	181269	97.94	15056	-8.7651	21456	74.94	15731
TAP-2.5D*(Fast Thermal Model)	-41.3358	111545	91.97	20782	-50.2010	231859	92.82	9172	-7.7890	19067	76.16	9984

* takes a similar amount of time as training *RLPlanner* for 600 epochs.

TABLE II

ACCURACY AND SPEED COMPARISON DURING THERMAL EVALUATION

Metrics	**Fast Thermal Model**	***Hotspot***
MSE	0.1732 K	Ground Truth
RMSE	0.4162 K	
MAE	0.2523 K	
MAPE	0.0726%	
Inference Speed	0.1012 s (127×)	12.8976 s

B. Reinforcement Learning and Agent Architecture

In our agent architecture, the policy network and the value network share the same feature encoding CNN layers and two separate fully connected layers are used to get the probability matrix and expected reward. We employ the proximal policy optimization (PPO) [7] to train the networks. Moreover, we employ a random network distillation (RND) [8] bonus to encourage the agent to explore novel states. It involves two neural networks: a fixed and randomly initialized target network, and a predictor network trained on data collected by the agent.

C. Reward Calculations

The target of chiplet floorplanning is to minimize the total wirelength and maximum operating temperature. We customize the reward function as $R = -\lambda \times W - \mu \times \frac{(max(T-T_0,0))^\alpha}{1+e^{-(T-T_0)}}$ where W and T are the wirelength and temperature, λ and μ are the weights, and T_0 is the temperature limitation and α is a hyperparameter to avoid gradient non-smoothness at $T = T_0$. The microbump assignments and wirelength optimization can be found in [4]. Traditional thermal analysis simulator *HotSpot* [9] is CPU-intensive since it involves constructing a thermal model for the entire system and solving large linear equations repeatedly. To address the computational burden while maintaining accurate thermal analysis, a physical-informed fast thermal model is proposed. It simplifies the thermal resistance network structure by treating it as an linear and time-invariant (LTI) system. By calculating the self-thermal and mutual-thermal resistance in the thermal resistance network, the chiplets' temperatures can be obtained. Thus, we characterize the self-thermal resistance by setting a chiplet's power to a non-zero value and run *HotSpot* to create a 2D self-thermal resistance table, and characterize the mutual-thermal resistance by a 1D table with respect to the distance between power source and grid location. More details can be found at the github.

III. EXPERIMENTAL RESULTS

A. Quantitative evaluations of the fast thermal model

A dataset comprising 2,000 synthetic chiplet systems are used to conduct thorough comparisons between the proposed fast thermal model and *Hotspot*. The results are summarized in TABLE II, where mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) are calculated. The comparisons demonstrate that the proposed fast thermal model can accurately estimate the maximum temperature of a chiplet-based system and achieve a speedup over 127x during thermal evaluations compared to *Hotspot*.

TABLE III

COMPARISONS OF REWARD ON 5 SYNTHETIC SYSTEMS

	Reward				
Method	**Case1**	**Case2**	**Case3**	**Case4**	**Case5**
RLPlanner	-5.8288	**-6.3236**	-10.0058	-8.4076	-8.6193
RLPlanner(RND)	**-5.1062**	-6.7848	**-9.9335**	**-8.3903**	**-8.2049**
TAP-2.5D(*HotSpot*)	-6.6439	-8.9846	-12.3946	-10.5525	-10.6965
TAP-2.5D*(Fast Thermal Model)	-6.3627	-7.1250	-10.7151	-9.8286	-8.5189

* takes a similar amount of time as training *RLPlanner* for 600 epochs.

B. Performance evaluations of RLPlanner

Three open-source benchmarks and five synthetic systems are used to evaluate the performance of *RLPlanner*. Comparisons and analysis between the developed *RLPlanner* and TAP-2.5D [4], an SA-based thermal-aware floorplanning algorithm, are presented in TABLE I and TABLE III. Across all eight cases, *RLPlanner* with RND achieves an average improvement of 20.28% in optimization goals compared to TAP-2.5D with *HotSpot*, and an average improvement of 9.25% compared to TAP-2.5D with the fast thermal model within similar or less running times. It confirms our intuition that by conducting end-to-end co-optimizations, *RLPlanner* shows better efficiency and quality in optimizing chiplet-based systems.

REFERENCES

[1] T.-C. Chen *et al.*, "Modern floorplanning based on b/sup*/-tree and fast simulated annealing," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 25, no. 4, pp. 637–650, 2006.

[2] A. Coskun *et al.*, "A cross-layer methodology for design and optimization of networks in 2.5 d systems," in *2018 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*. IEEE, 2018, pp. 1–8.

[3] L. Chen *et al.*, "Fast thermal analysis for chiplet design based on graph convolution networks," in *2022 27th Asia and South Pacific Design Automation Conference (ASP-DAC)*. IEEE, 2022, pp. 485–492.

[4] Y. Ma *et al.*, "Tap-2.5 d: A thermally-aware chiplet placement methodology for 2.5 d systems," in *2021 Design, Automation & Test in Europe Conference & Exhibition (DATE)*. IEEE, 2021, pp. 1246–1251.

[5] A. Kannan *et al.*, "Enabling interposer-based disintegration of multi-core processors," in *Proceedings of the 48th international symposium on Microarchitecture*, 2015, pp. 546–558.

[6] Huawei ascend 910 provides a nvidia ai training alternative. https://www.servethehome.com/huawei-ascend-910-provides-a-nvidia-aitraining-alternative/.

[7] J. Schulman *et al.*, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[8] Y. Burda *et al.*, "Exploration by random network distillation," *arXiv preprint arXiv:1810.12894*, 2018.

[9] W. Huang *et al.*, "Hotspot: A compact thermal modeling methodology for early-stage vlsi design," *IEEE Transactions on very large scale integration (VLSI) systems*, vol. 14, no. 5, pp. 501–513, 2006.