TAGE Branch Predictor on Sniper Simulator

CS5222 – Project 1

Shen Jiamin
A0209166A
shen_jiamin@u.nus.edu

[7]

Implementing the branch predictor in Sniper will allow us to learn the details of a novel branch predictor. What I feel is important for this project is the quality of the comparison against existing branch predictors, as well as whether the benchmarks meet the expected performance from the reviewed paper. In addition, describing the key features of the branch predictor chosen is an important part of this project. Why did you choose this predictor? Below are some examples of some questions that I'd like to see answers to in your report (adding more background / description / results are a plus and will help your grade).

(1) Why did you choose this branch predictor? What are interesting components of this branch predictor? What was novel about this predictor when it was published? While a critical review is not necessary for this report, an in-depth overview is needed.

(2) What branch predictors were this one compared to in the original paper? Did your implementation perform as expected? How did you debug your implementation?

(3) Please compare this branch predictor against the two other branch predictors in Sniper, and plot as an MPKI miss chart (misses per 1K instructions) for each of the 5 benchmarks and 3 branch predictors.

(4) Bonus, implement the branch predictor that was compared to in the paper (the one that was beaten) so that we can compare performance numbers against the original (20% bonus).

In the paper "A case for (partially) TAgged GEometric history length branch prediction" [7] by André Seznec (2006), the author presented his design of a new branch predictor named TAGE predictor. The paper was produced in the background that (1) exploiting several different history lengths is acknowledged, and (2) hybrid predictors combining the prediction result from multiple meta predictors have been developed as well as prediction combination functions such as majority vote, predictor fusion and partial tagging and the adder tree seems to outstanding among them.

The TAGE predictor users a geometric series as the list of history length, which is the main characteristic of O-GEHL[2], while it derives the main structure from PPM-like predictor[1]. The significant difference of TAGE from O-GEHL is the use of partial tags instead of adder tree, which the author believes provides better accuracy and storage efficiency and enables the adaptation of the design for conditional branch prediction to the indirection target branch prediction. Compared with the PPM-like predictor, TAGE proposed a new predictor update algorithm that minimizes the perturbation induced by a single occurrence of a branch. Based on the structure of TAGE, the author also proposed an indirect branch target predictor ITTAGE and a unified predictor COTTAGE,

which supports both conditional branch and indirect branch target prediction.

The author evaluated the accuracy of the TAGE predictor with the parameters listed in Table 2 in the paper. He compared the MPKI (mispredictions per kilo instructions) of TAGE, O-GEHL and PPM-like predictors with different prediction tables sizes. The result shows that the TAGE predictor outperforms O-GEHL and PPM-like in all tested conditions, indicating the correctness of the author's proposal. He also evaluated the accuracy regarding the variation of history length, which shows that with the maximum history length growing, the accuracy is better at first but no improvement when the length is too long. The evaluation plot for the impact of the tag width is not as clear as the ones ahead. The variation of interest is the tag width, but the plot's x-axis varies the prediction tables size, which makes it difficult to compare between different tag widths. The result shows that increasing the tag width over a threshold provides little return for accuracy improvement. Besides the explicit evaluations discussed above, the experiment results show that the mispredictions decrease with larger prediction tables but diminishing returns for accuracy.

The work is auspicious in that many variations of TAGE predictor are proposed after it and lead nearly all the Championship Branch Prediction afterwards, for example, L-TAGE[3] in 2007, ISL-TAGE[4] in 2011 and TAGE-SC-L[5, 6] in 2014 and 2016.

However, these following works also indicate the limitations of the TAGE predictor. In [3], the author found that using the alternate prediction, instead of the prediction using the longest history, for newly allocated entries is more efficient in some applications. He also identified that TAGE fails to predict when the control flow path inside the loop body is irregular and augmented TAGE with loop predictor. In [4], it is identified that TAGE sometimes fails to predict branches that are not strongly biased but that are only statistically biased and thus the Statistical Corrector Predictor is introduced. He also identified that the number of access to the predictor can be reduced and had an Immediate Update Mimicker as an add-on.

REFERENCES

[1] Pierre Michaud. 2005. A PPM-like, Tag-based Predictor. *The Journal of Instruction-Level Parallelism* 7 (2005), 10 pages. https://jilp.org/vol7/

[2] André Seznec. 2005. Analysis of the O-GEometric History Length Branch Predictor. In *32nd International Symposium on Computer Architecture (ISCA'05)*. IEEE Computer Society, Madison, WI, USA, 394–405. https://doi.org/10.1109/ISCA.2005.13

[3] André Seznec. 2007. The L-TAGE Branch Predictor. *The Journal of Instruction-Level Parallelism* 9 (2007), 13 pages. https://jilp.org/vol9/index.html

[4] André Seznec. 2011. A 64 Kbytes ISL-TAGE branch predictor. In *JWAC-2: Championship Branch Prediction*. JILP, San Jose, United States, 4 pages. https://hal.inria.fr/hal-00639040

[5] André Seznec. 2014. TAGE-SC-L Branch Predictors. In *JILP - Championship Branch Prediction*. Minneapolis, United States, 8 pages. https://hal.inria.fr/hal-01086920

[6] André Seznec. 2016. TAGE-SC-L Branch Predictors Again. In *5th JILP Workshop on Computer Architecture Competitions (JWAC-5): Championship Branch Prediction (CBP-5)*. Seoul, South Korea, 4 pages. https://hal.inria.fr/hal-01354253

[7] André Seznec and Pierre Michaud. 2006. A case for (partially) TAgged GEometric history length branch prediction. *The Journal of Instruction-Level Parallelism* 8 (2006), 23 pages. https://jilp.org/vol8/