Two Level Adaptive Training Branch Prediction by Tse-You Yeh and Yale N. Patt; Proc. 24th Annual International Symposium on Microarchitecture, 1991

Summary by Amit Goswami

Problem:
While deep pipelines help in improving performance, it can also waste a lot of cycles if it predicts wrong branches. Thus it is vital for these processors to have an effective branch predictor. In this paper the authors propose a new branch prediction scheme and compare its performance with other such schemes.

Proposed Solution:
The authors propose a two-level adaptive training branch prediction scheme that is based not only on the history of branches executed, but also on the occurrences of a particular branching pattern in the past. The history information is collected as the program executes and not in advance. This scheme is implemented by maintaining two tables: the history register table (HRT) and the history pattern table (PT). Branching predictions are made by indexing into the pattern table using the addresses in the HRT. After every prediction the PT is updated by a state transition function that is based on the old value in the PT and the branching decision made.

Two approaches have been suggested for the implementation of the HRT: Associative History Register Table (AHRT) and Hash History Register Table (HHRT). The AHRT works as a finite set that is replaced if necessary by the Least Recently Used algorithm. The HHRT works by using the branch address to hash into the table.

Research Methodology:
The authors use the Motorola ISIM simulator and nine benchmarks from the SPEC benchmark suite to generate instruction traces that are fed into the branch predictor simulator. This simulator executes the instructions and records the statistics regarding prediction accuracy. Three implementations were used for the HRT: AHRT and HHRT as described above and the IHR (Ideal HRT) which has a register for every conditional branch in the program. Five different automata are used in the analysis. For clarity purposes a naming convention was developed based on the different aspects of the proposed prediction method.

Results:
• Of the five automata used, the A2 automaton that uses a saturating up-down counter is the most efficient
• The best HRT implementations in order were: IHRT, 512-entry AHRT and 512-entry HHRT.
• Longer history register length improves accuracy
• The accuracy of the Two-Level Adaptive Training scheme is 97% which is 4% better than other schemes.
Comments & Critique:

1. The paper is clear, well-organized and an interesting read.
2. The simulation results are presented in a detailed manner which was helpful. However a more clear and detailed explanation of the proposed scheme is required, after all that is the essence of the paper.
3. The automata are well-chosen and reflect other branching schemes, hence appropriate.
4. Collisions in the hash implementation of the HRT can have a substantial effect on performance. Perhaps more analysis needs to be done in this area.
5. Since the standard SPEC benchmarks were used, it seems futile to simulate the other prediction schemes. Could they have used the existing benchmark data?