Computer Architecture Reading Group Notes

Date: 2/12/04
Discussion Leader: Mattan
Notes: Vicky
Topic: Reliability
Papers:

1. Todd Austin, DIVA: A Reliable Substrate for Deep Submicron Microarchitecture Design, ACM/IEEE 32nd Annual Symposium on Microarchitecture (MICRO-32), November 1999, received Best Paper Award.
2. L. Spainhower and T. A. Gregg, IBM S/390 Parallel Enterprise Server G5 fault tolerance: A historical perspective, IBM Journal of Research and Development, Vol. 43, Nos. 5/6, 1999, pp. 863-874.

DIVA: A Reliable Substrate for Deep Submicron Microarchitecture Design

Summary:
This paper proposes a mechanism that is tagged on a processor to increase reliability. It is designed to hide design flaws and transient errors. The basic idea is to attach a checker to the processor core. The checker re-computes each instruction and compares the result against that from the core. The proposed checker has been implemented in a simulator and simulation results show that the checker adds little overhead

Discussion:

1. Checker becomes a single point of failure. But then the paper is trying to argue that the checker can be verified formally, thus simplifying verification. (There is still the problem of testing though.)
2. There is a lack of analysis on transient errors and size of devices to be used in the checker. The frequency of transient errors caused by alpha particles depends on the size and density of devices (e.g. any decent size cache would need ECC). Large device -- smaller effect from alpha particles but there is a higher probability it will be hit.
3. Missing in the paper: how branch instructions are checked?
4. Q: Why deeper pipeline? A: Checker is supplied with digested data and it does not need to do any decoding, prediction, etc.
5. Q: After adding an extra cache port to remove resource contention, how can the checker keep up with the core? A: Core IPC is not 4 but more like 2.
6. Checking communication cannot be done in parallel with the computation check of the instruction that produces the data communicated, according to the way it is described in the paper (input values provided by core are compared against architectural states). It is unclear what it means for the checker to be a 4 instructions wide pipeline. Possible solutions:
    a. Let CHKcomp run ahead of CHKcomm.
    b. Compare input values against output values supplied to the checker (but the checker has to be supplied with the dependency information).

7. The claim that DIVA removes burden of correctness does not seem to hold considering that performance suffers at 1 error/cycle. However, DIVA may allow a design to be released before it is 100% verified (isn't this the way it is done now? :)
8. With large designs, DIVA's area overhead can be big (not counting cache). With small designs, replicating the whole core may not be that much bigger.

IBM S/390 Parallel Enterprise Server G5 Fault Tolerance: A Historical Perspective

Summary:
This paper is a description of a complete system. While it does not present any research ideas, it is valuable because it is a real system. There is a historical thread on the transition from the TCM technology to CMOS. The other thread is G5's design for fault tolerance – they chose to replicate execution units and only matched results are committed. If a unit fails more times than some threshold, it is replaced by a spare.

Discussion:
1. There is another paper that focuses on the processor ("IBM's S/390 G5 microprocessor design" in IEEE Micro, vol. 19, no. 2, pages 12-23, Mar/Apr 1999).
2. There is no single point of failure throughout the system. The design strives for high reliability and availability without performance impact. MTTF is 45 years (official guarantee is like 30 yrs).
3. In one MCM, there are 12 processor chips, with each chip having at least one working processor. (The number you can use depend on how much you pay.)
4. A reference to Horst's paper: one error/yr in 75 unchecked processors – it makes you wonder how many errors go undetected.
5. In the other paper, it is shown how errors can be tracked to their sources so that they can be improved in the next generation.
6. Software also plays a role. The OS has direct access to all the error registers and keeps track of the error history.