Critique of paper presented on 11/5/2008
Venkatraghavan Ramesh
1200072784


Presenter – Naresh Sukumar

The paper seeks to design and evaluate a general prediction logic that replaces optimizations for specific sharing patterns to reduce memory latency for misses to remotely cached blocks. Based on the discussion in the class, the following can be listed as the strengths and weaknesses of the paper –

Strengths –
1. Dynamic pattern instead of static pattern – cosmos can discover and track application-specific patterns not known a priori. This remains the chief contribution of the paper given that a general predictor overcome the dependence of predictor on application type making it adapt to new application patterns.
2. The prediction is immune to order of arrival of messages and has multiple predictions to choose from compared to PAp.
3. Cosmos tries to separates the predictor from the protocol which, if implemented correctly will enable predictor to be designed and evaluated fairly independent of coherence protocol

Weakness –
1. Since the paper deals with coherence protocol message prediction in isolation as opposed to integrating it into a coherence protocol, it may be making certain simplifying assumptions. While the justification for the same has been given in terms of readiness of tool and implementation concerns that might obscure initial studies, the efficacy of the scheme in its entirety compared to existing static pattern predictors cannot be fully understood.
2. While the paper repeatedly dwells on the accuracy of prediction, the discussion in section 4.4 on the performance benefits in the form of speedup are not adequate. While the paper mentions that its motive is not to investigate impact on runtime, this would be a vital factor to establish the success of cosmos irrespective of its high prediction accuracy.
3. While cosmos is an attempt at designing a general predictor scheme, its reliance on the slow changing patterns/signatures as illustrated by low accuracy of Cosmos for “barnes” and the authors’ confession that the cosmos may function worse than directed predictors do not make a strong case for adopting cosmos. This is all the more significant given that cosmos is likely to require more states than directed optimization. Hence, the argument validating higher investment based on the claim of higher prediction accuracy is not strong since the higher prediction accuracy is dependent on nature of application.
4. More memory to accommodate History tables for Message and History – the overhead discussed in the paper is for small cache blocks (128 bytes), so the extension to larger cache blocks in use today is not known.
5. Another simplifying assumption is that of full-map and write-invalidate directory protocol. This ignores the performance of cosmos for more complicated schemes.