Summary:
The paper discussed presents several mechanisms for reducing off-chip access in multicore architectures. The first employs the use of another cache level L3 in conjunction with a write-back-history-table to reduce the number of write-backs to the L3 layer and off-chip memory. The second mechanism further reduces hits to the L3 layer by using L2-L2 transfer of data to peer caches so the data is stored in neighboring caches instead of the lower layers. The results by the researchers claim that the number of clean write-backs is reduced by up to 50%.

Summary of strengths:

1. The majority of noted strengths in the paper are that it clearly results in increased performance of various applications by reducing the off-chip access rate.
2. Each mechanism has a positive contributing effect on the reduction of the miss rates. Although the combined effect of both mechanisms is not additive, together, the two allow TP to perform better than either would alone.

Summary of weaknesses:

1. The scheme generally only improves applications where overall throughput is a factor in performance. In general, the benchmarks show a greater impact on performance when the outstanding memory load increases (i.e. “memory pressure”). An example of this was the greater performance increase in Transaction Processing as opposed to the more marginable increase in performance for the CPW2, NotesBench, etc. applications. The 13% increase in application performance might not justify the cost of implementing the mechanisms.
2. Implementation details and complexity along with increased costs associated with introducing another level of cache have not been considered. Such considerations might prevent the actual use of the mechanisms presented in this paper as it might reduce the overall number of cores available on-chip.
3. Power issues that might be introduced by the sophisticated cache mechanisms were not considered in the paper and might be an issue in a many-core architecture.