Tip of the Iceberg: Low Associativity Paging

Virtual address translation is a growing bottleneck for large data applications, such as machine learning and graph analytics, as each virtual address translation requires multiple, slow, memory accesses. TLBs attempt to solve this problem, but have limited coverage and cannot keep up with the rate at which virtual memory scales. This paper argues that a path forward is to reconsider the necessity of fully associative page translations, where any virtual page can map to any physical location. Reduced associativity, as in hashed page tables, can improve TLB coverage by reducing the number of bits needed to cache one translation. Unfortunately, reduced associativity is often implemented using techniques that cause well-sung usability problems that stem from associativity conflicts. We summarize a new hashing scheme, iceberg hashing, that is suitable for virtual memory, as it successfully addresses flaws in prior hashing schemes. This thesis describes how iceberg hashing can be implemented and used in a virtual memory system. Finally, the thesis presents a study and preliminary data using xv6, which indicates iceberg hashing can be configured such that associativity conflicts only manifest when memory is more than 95% full, and can be resolved using standard swapping techniques.