Bandwidth-Based Page Allocation for NUMA Architectures

Martin I. Oliveira Universidade Estadual de Campinas
Emilio C. Francesquini Universidade Estadual de Campinas
Alexandro Baldassin Universidade Estadual Paulista
Guido Araujo Universidade Estadual de Campinas
J. Nelson Amaral University of Alberta

Assigning memory pages to nodes of a NUMA architecture in a way that matches the demand for remote data transfers with the available communication bandwidths and memory-controller capacity in the machine can have a significant impact in the execution time of programs. Such assignment typically requires dealing with four simultaneous goals: (a) keep threads close to the memory pages they access; (b) evenly distribute the workload among nodes; (c) maintain memory demand below memory controllers' bandwidth; and (d) reassign threads and pages to follow changes in the memory access pattern of the program. However, most solutions to this problem address only a subset of these goals, mainly because they seek to avoid complex solutions or expensive implementation overheads. This talk will discuss a heuristic-based algorithm that simultaneously allocates (P)ages, (T)hreads and (B)andwidth for each node of a NUMA architecture. In contrast to alternative approaches, PTB integrated solution seeks both to uniformly distribute workload and to limit memory demand to the controllers' bandwidth while also addressing asymmetry issues found in the communication paths of modern NUMA architectures. This talk will discuss experimental results using Parsec, NAS and Metis benchmarks comparing this new heuristic algorithm with the Linux's default scheduler.