Algorithms have two costs: arithmetic and communication, i.e. moving data between levels of a memory hierarchy or processors over a network. Communication costs (measured in time or energy per operation) already greatly exceed arithmetic costs, and the gap is growing over time following technological trends. Thus our goal is to design algorithms that minimize communication. We present algorithms that attain provable lower bounds on communication, and show large speedups compared to their conventional counterparts. These algorithms are for direct and iterative linear algebra, for dense and sparse matrices, as well as direct n-body simulations. Several of these algorithms exhibit perfect strong scaling, in both time and energy: run time (resp. energy) for a fixed problem size drops proportionally to p (resp. is independent of p). Finally, we describe extensions to algorithms involving arbitrary loop nests and array accesses, assuming only that array subscripts are linear functions of the loop indices.