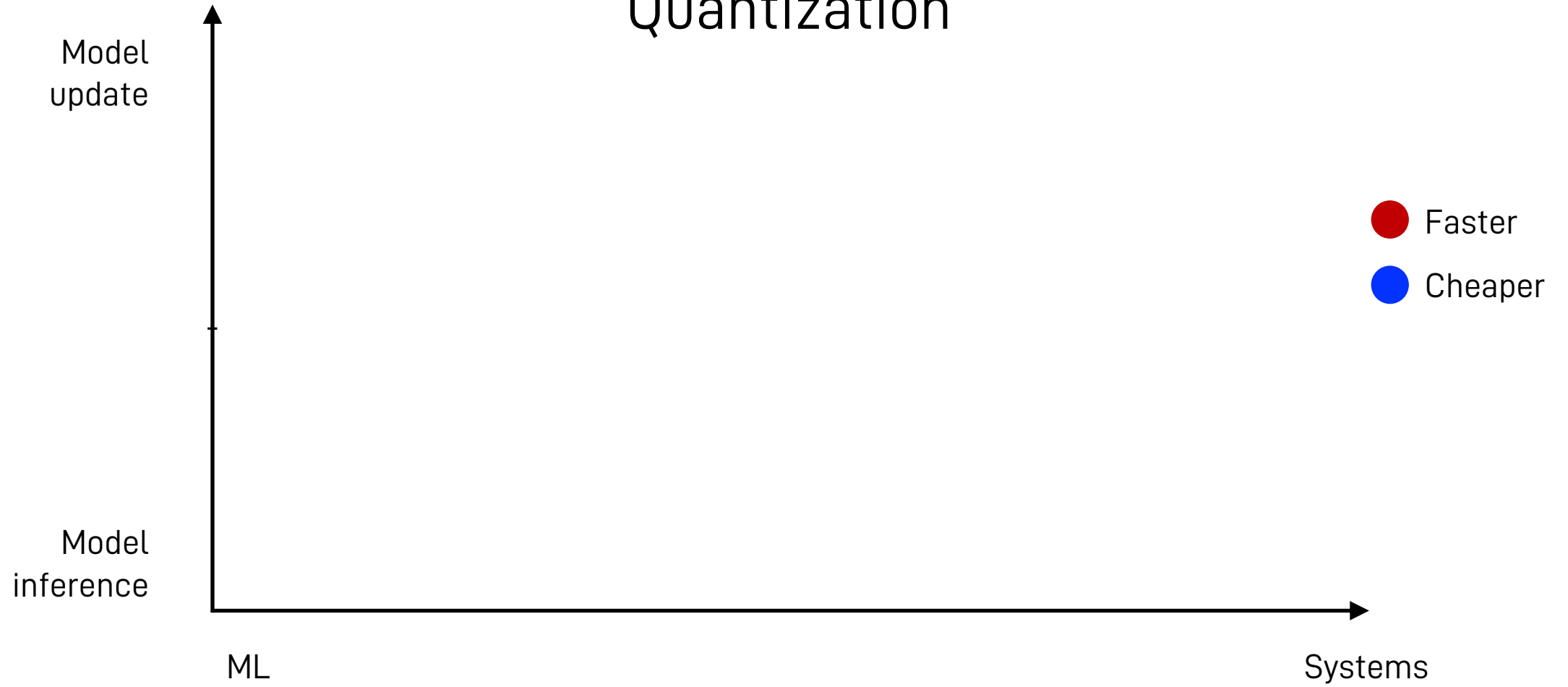
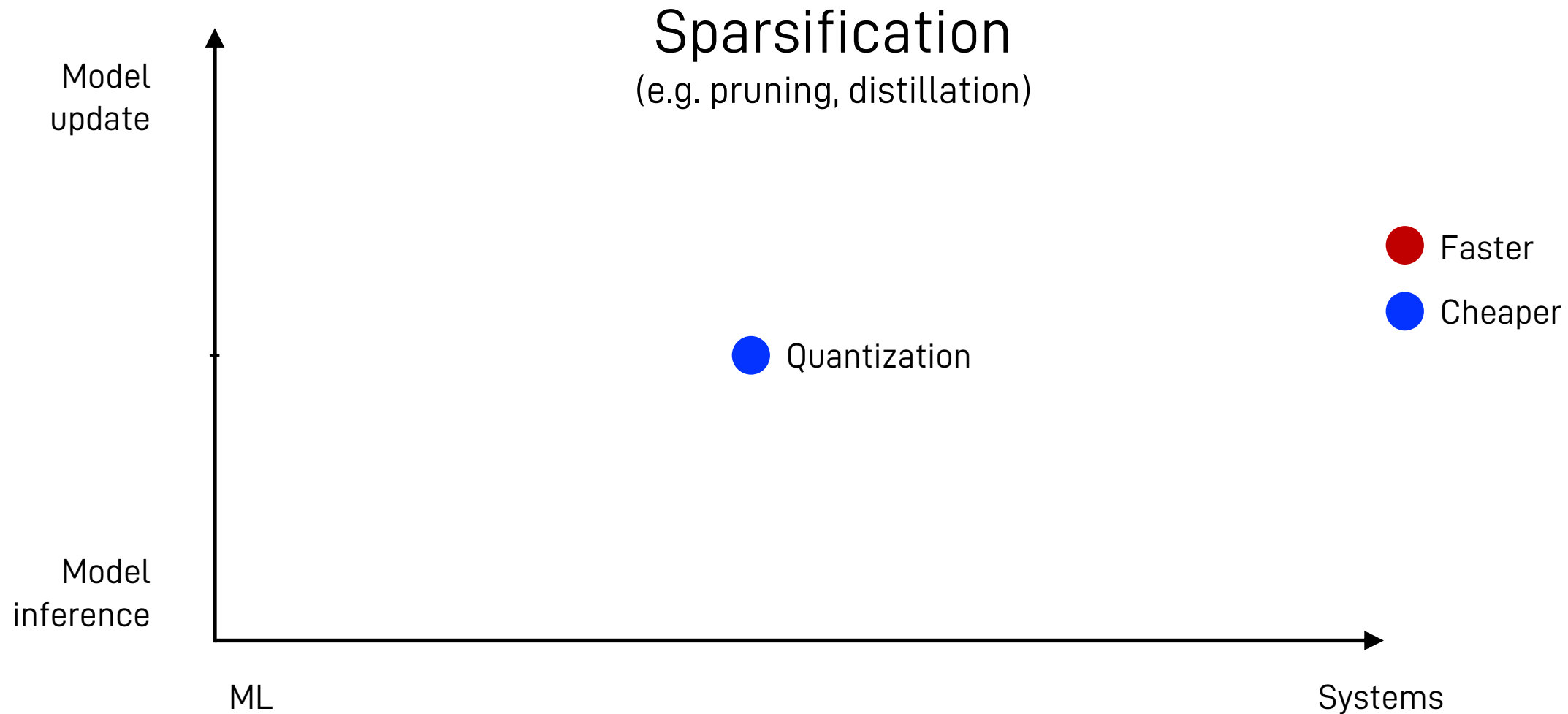


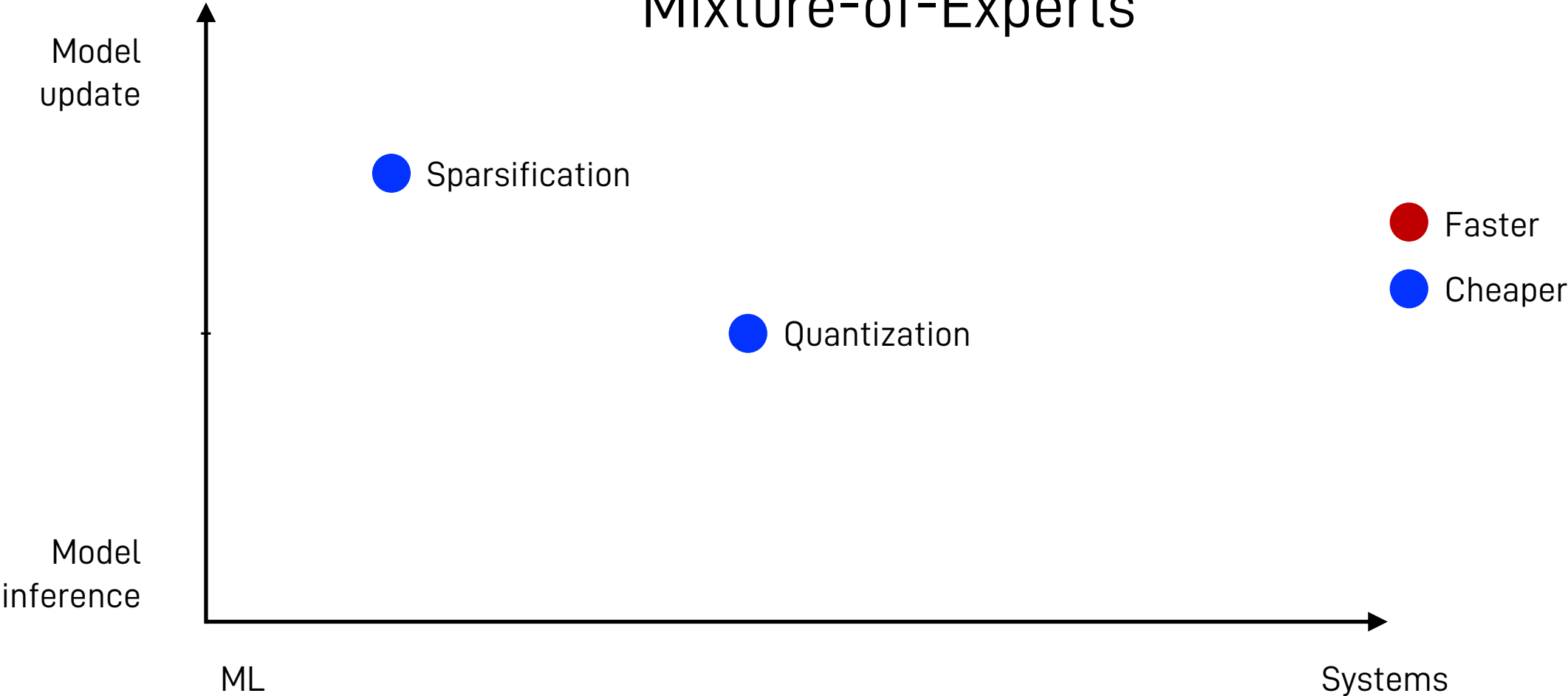


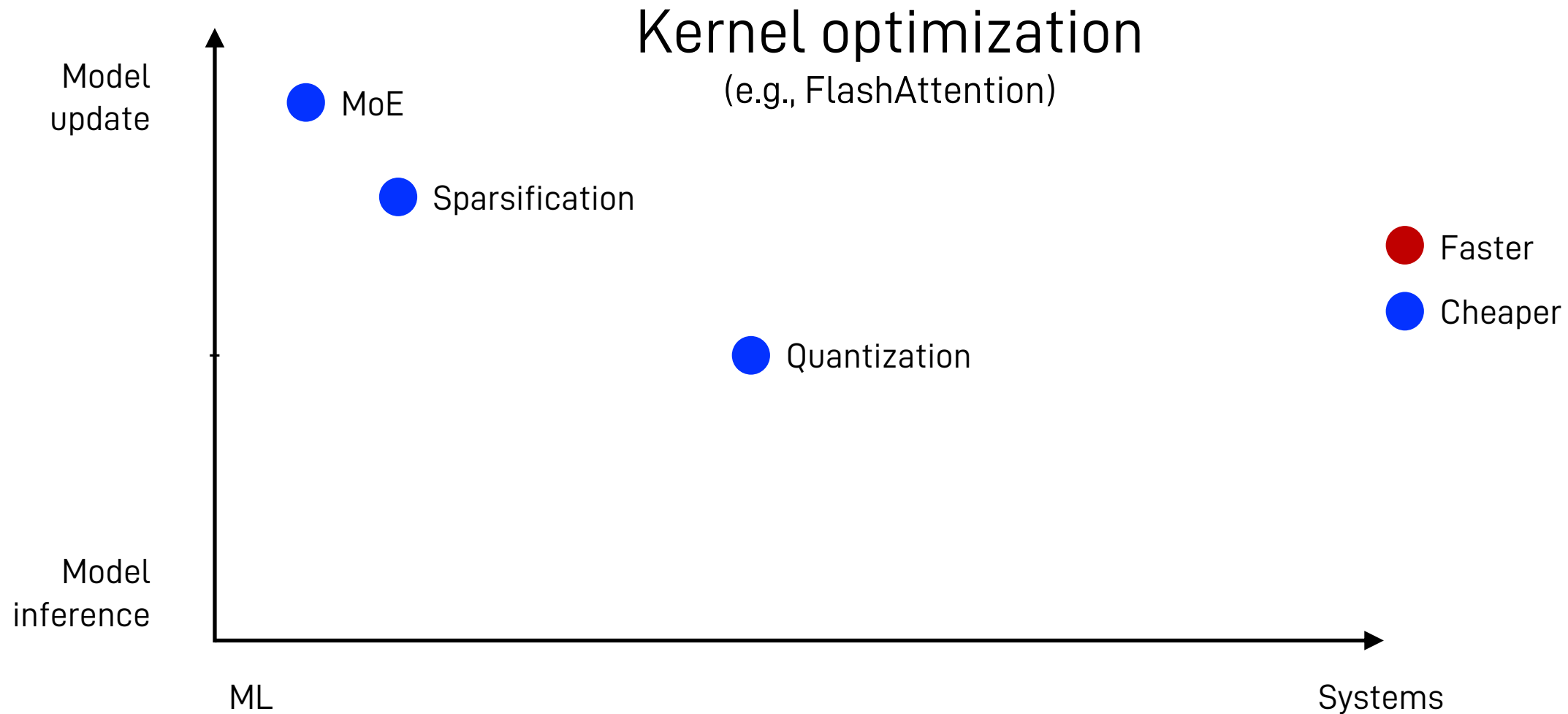
Quantization

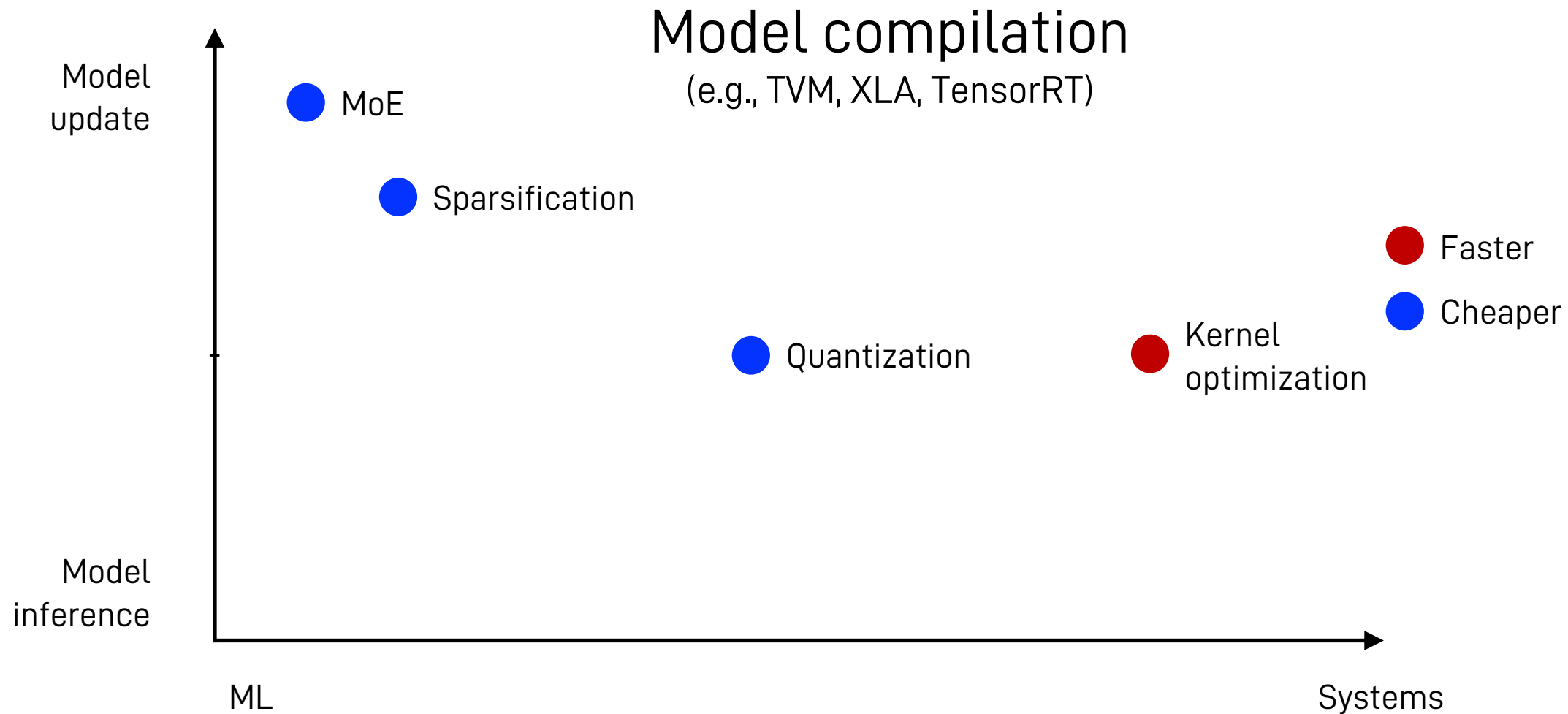




Mixture-of-Experts







Superoptimization

