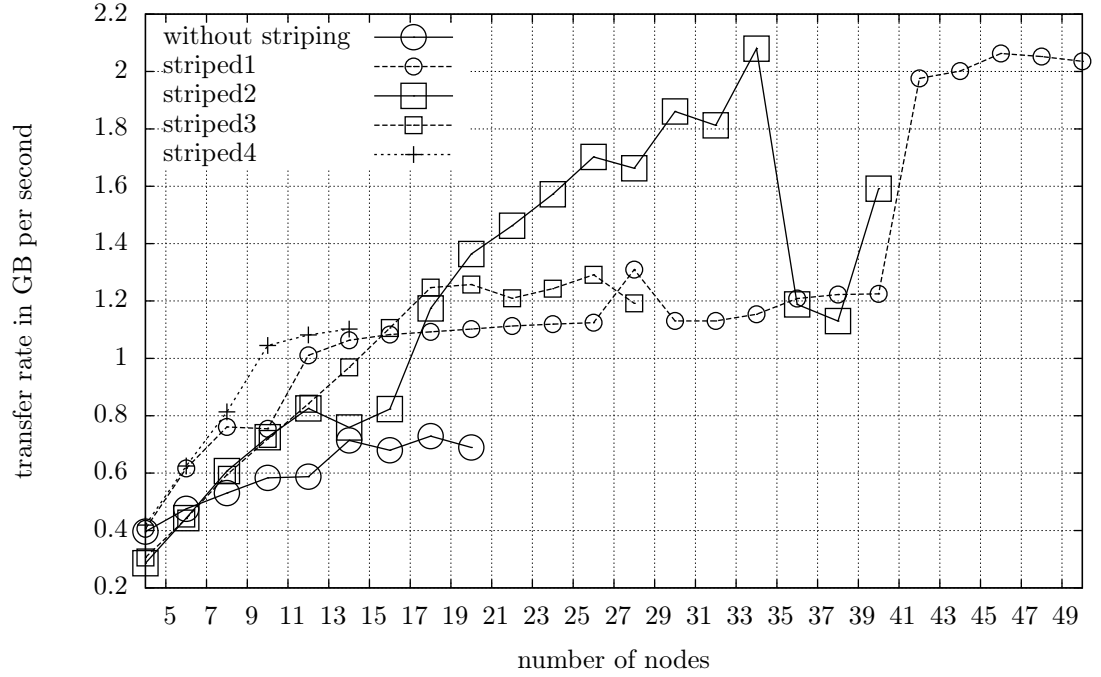


Data distribution, HPC cluster. 224 GB of pair-ended FASTQ files



The efficiency ratios?? E_p/E_{56} for different Hadoop on the Cloud cluster sizes 112 and 224 cores. The data for Shared TM and SSH TM deployment scenarios are shown. The efficiency ratios are shown for different pipeline stages: reads pre-processing (1-3), alingment (4), and SNP calling (5). The values above 1 shows that the cluster of current size is more efficient than the $p = 56$ cores cluster, and vice versa.