Table 1: Prefill and decode performance for Ring vs Regular attention. Normalized TTFT is represented as ms per prompt token.

Strategy	**Prompt**	**TTFT (ms)**	**TTFT/token**	**Decode (ms)**	**Total (ms)**	**Comm (ms)**	**Comp/Comm**
Ring	256	1576.04	6.16	558.09	18318.77	12823.92	0.43
Regular	256	66.58	0.26	57.10	1779.55	0.00	N/A
Ring	512	1513.53	2.96	565.80	18487.42	12773.48	0.45
Regular	512	71.20	0.14	55.99	1751.03	0.00	N/A
Ring	1024	1667.12	1.63	566.92	18674.79	12690.79	0.47
Regular	1024	108.38	0.11	56.90	1815.44	0.00	N/A
Ring	4096	3220.51	0.79	573.09	20413.19	13001.37	0.57
Regular	4096	453.37	0.11	47.90	1890.37	0.00	N/A