## Performance Report Jiang Feiyu 3035770800

| Thread Numbers | Speed (tok/s) | User Time (s) | System Time (s) | Use Time / System |
|----------------|---------------|---------------|-----------------|-------------------|
|                |               |               |                 | Time              |
| 0 (Sequential) | 61.288006     | 4.179295      | 0.048166        | 86.76857          |
| 1 (1 (child)   | 46.051448     | 5.456793      | 0.239314        | 22.80181          |
| Thread)        |               |               |                 | 22.00101          |
| 2              | 78.963603     | 5.542089      | 0.267504        | 20.71778          |
| 4              | 114.592659    | 6.173410      | 0.415057        | 14.87364          |
| 6              | 129.032258    | 6.817340      | 0.594142        | 11.47426          |
| 8              | 141.514649    | 7.235925      | 0.869799        | 8.319077          |
| 10             | 143.096702    | 7.863737      | 1.010648        | 7.780886          |
| 12             | 148.837209    | 8.521305      | 1.613329        | 5.281815          |
| 16             | 149.358226    | 9.447389      | 1.866307        | 5.062077          |

As the number of threads increases, the speed (tok/s) initially shows an increasing trend and eventually stabilizes. It gradually improves from an initial value of 61.288006 tok/s to 149.358226 tok/s. This is because with more threads, it becomes possible to simultaneously process more tasks, thus enhancing processing speed. However, after reaching a certain point, the performance gains start diminishing due to increased competition and scheduling overhead among threads, which can create performance bottlenecks.

Both user time and system time increase as the number of threads increases. This is because with more threads, the overall process requires more time to execute the tasks in both user mode and kernel mode. However, the growth rate of user time and system time is relatively slow, which may be attributed to the efficiency of parallel execution and task partitioning among threads.

The "Use Time / System Time" ratio decreases as the number of threads increases. This is because, with more threads, the proportion of user time decreases relative to system time. This can be attributed to the effectiveness of parallel execution and task partitioning among threads, resulting in a more efficient utilization of user time compared to system time.