Scaling Limits of the SOP Actor-Learner Framework for Large Robot Fleets

Anonymous Author(s)

ABSTRACT

We investigate whether the near-linear scaling of the Scalable On-line Post-training (SOP) framework persists for significantly larger robot fleets beyond the 1–4 actor regime demonstrated empirically. Through analytical modeling of the SOP actor-learner pipeline with parameterized communication overhead, we simulate scaling behavior for fleets of 1 to 128 robots under three communication models (linear, $\sqrt{N}$, and $\log N$). Under $\sqrt{N}$ communication overhead, the system achieves 57.8× speedup at 64 actors (90.4% efficiency) and maintains >70% efficiency up to 162 actors. With $\log N$ overhead, efficiency remains above 92.6% at 64 actors with a critical fleet size of 256+. Linear communication overhead limits efficient scaling to 46 actors. We identify learner throughput saturation and bandwidth constraints as the primary bottlenecks for large fleets and discuss architectural modifications to extend the near-linear scaling regime.

1 INTRODUCTION

The Scalable Online Post-training (SOP) framework [6] couples distributed real-world data collection from a robot fleet with centralized online learning for VLA models. Empirically, SOP demonstrates near-linear speedups when scaling from 1 to 4 actors. Whether this favorable scaling extends to significantly larger fleets remains an open question, as system bottlenecks in communication, learner throughput, and synchronization latency may emerge at scale [3].

Understanding scaling limits is critical for the practical deployment of distributed robotic learning systems [2, 5]. We develop an analytical model of the SOP pipeline to predict scaling behavior and identify bottleneck thresholds.

2 SCALING MODEL

2.1 SOP Architecture

The SOP system consists of N robot actors collecting experience in parallel and a central learner performing gradient updates. The wall-clock time per training cycle is:

$$T(N) = \max\left(\frac{T_{\text{collect}}}{N}, T_{\text{learn}}\right) + T_{\text{comm}}(N) \qquad (1)$$

where T_{collect} is the serial data collection time, T_{learn} is the per-step learning time, and $T_{\text{comm}}(N)$ is the communication overhead.

2.2 Communication Models

We consider three overhead models reflecting different system architectures:

- **Linear**: $T_{\text{comm}} = \alpha N$ (point-to-point communication)
- **Sublinear**: $T_{\text{comm}} = \alpha \sqrt{N}$ (aggregation tree)
- **Logarithmic**: $T_{\text{comm}} = \alpha \log_2 N$ (hierarchical protocols)

2.3 Scaling Metrics

The speedup $S(N) = T(1)/T(N)$ and scaling efficiency $E(N) = S(N)/N$ quantify how well the system utilizes additional actors.

Table 1: Speedup $S(N)$ and efficiency $E(N)$ at key fleet sizes.

Model	N=16		N=64		N=128	
	S	E	S	E	S	E
Linear	15.2	0.95	34.5	0.54	29.0	0.23
$\sqrt{N}$	15.8	0.99	57.8	0.90	98.4	0.77
$\log N$	15.8	0.99	59.3	0.93	107.9	0.84

3 RESULTS

3.1 Speedup Analysis

Table 1 shows that all models achieve near-linear scaling for $N \leq 16$, consistent with SOP's empirical results. Divergence occurs beyond $N = 32$: linear overhead leads to *decreasing* speedup beyond 64 actors, while $\sqrt{N}$ and $\log N$ models maintain useful scaling to 128+ actors.

3.2 Critical Fleet Size

The fleet size where efficiency drops below 70% is: 46 (linear), 162 ($\sqrt{N}$), and 256+ ($\log N$). This suggests that with appropriate communication infrastructure, SOP can scale to over 100 robots while maintaining practical efficiency.

3.3 Bottleneck Analysis

Two bottlenecks emerge at scale: (1) bandwidth saturation when aggregate data generation exceeds network capacity, and (2) learner throughput saturation when the GPU cannot process updates as fast as data arrives. Under our default parameters, the learner becomes the primary bottleneck at approximately 300 actors.

4 DISCUSSION

Our analysis indicates that near-linear scaling can persist well beyond 4 actors under favorable conditions. The key requirements are:

(1) Sublinear communication overhead via hierarchical aggregation
(2) Sufficient network bandwidth (100+ MB/s for 64 actors)
(3) Learner parallelism for $N > 100$ (multi-GPU or pipeline parallelism)

Amdahl's law [1] provides a lower bound assuming fixed serial fraction ($p = 0.95$ gives max speedup of 20), while Gustafson's law [4] predicts better scaling for SOP's growing workload model. Our simulation results fall between these bounds, consistent with SOP's moderate serial fraction.

5 CONCLUSION

We predict that SOP's near-linear scaling extends to 64+ actors under $\sqrt{N}$ communication overhead (90.4% efficiency) and to 128+

actors under log N overhead (84.3% efficiency). The critical requirements are hierarchical communication protocols and adequate bandwidth. These results encourage scaling SOP deployments to significantly larger robot fleets.

REFERENCES

[1] Gene M Amdahl. 1967. Validity of the single processor approach to achieving large scale computing capabilities. *AFIPS Conference Proceedings* 30 (1967), 483–485.

[2] Anthony Brohan et al. 2023. RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control. *arXiv preprint arXiv:2307.15818* (2023).

[3] Jeffrey Dean et al. 2012. Large scale distributed deep networks. *Advances in Neural Information Processing Systems* 25 (2012).

[4] John L Gustafson. 1988. Reevaluating Amdahl's law. *Commun. ACM* 31, 5 (1988), 532–533.

[5] Sergey Levine et al. 2016. End-to-end training of deep visuomotor policies. *JMLR* 17, 39 (2016), 1–40.

[6] Yifeng Pan et al. 2026. SOP: A Scalable Online Post-Training System for Vision-Language-Action Models. *arXiv preprint arXiv:2601.03044* (2026).