**Review of Project Adam**

# **Summary**

Adam achieves high efficiency and scalability through whole system co-design that optimizes and balances workload computation and communication.

By using Adam to train a large DNN model, large-scale commodity distributed systems can be used to efficiently train very large DNNs to world-record accuracy on hard vision tasks using current training algorithms.

**Contributions**

**①**Optimizing and balancing both computation and communication for the application through whole system co-design. (Restructure the computation across machines to reduce communication requirements.)

**②**By exploiting the ability of machine learning training to tolerate inconsistencies well, Adam achieve high performance and scalability.

**③**Demonstrating that system efficiency, scaling, and asynchrony all contribute to improvements in trained model accuracy.

**Comments**

**Flash Points:** ①Large-scale commodity distributed systems can be used to efficiently train very large DNNs. ② Adam is significantly more efficient and scalable.

**Limitations:**  ①They did not implement is to have the faster machines steal work from the slower ones. ② large scale systems need much larger DNNs.