Mitigating the Impact of Tail Latency of Storage Systems on Scalable Deep Learning Applications

Hiroki Ohtsuji\textsuperscript{1}, Erika Hayashi\textsuperscript{1}, Naoto Fukumoto\textsuperscript{1}, Eiji Yoshida\textsuperscript{1} \
Takuya Okamoto\textsuperscript{2}, Takeru Kuramoto\textsuperscript{3}, Osamu Tatebe\textsuperscript{3} \
\textsuperscript{1}Fujitsu Laboratories Ltd. \
\textsuperscript{2}Fujitsu Ltd. \
\textsuperscript{3}University of Tsukuba
Massive-scale Fast Distributed Deep Learning

- **ResNet-50/ImageNet (1000x1000 Images, 120 GB) training in a minute**
  - All training processes read the data sets during the training
  - Can shared storage systems provide enough access performance to those workloads?

![Graph showing training time over dates](image)

- 29 hours for training
- 1.2 minutes (by Fujitsu Lab. [1], Mar 2019)

Data parallel Distributed Deep Learning

- More than 1K processes read training data sets from a shared storage

<table>
<thead>
<tr>
<th># of Files</th>
<th>1M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size:</td>
<td>100GB+</td>
</tr>
</tbody>
</table>

- Training data

## Key Points
- **Shared storage**
- **Node #0**
  - Weight \( \Delta w_0 \)
- **Node #1**
  - Weight \( \Delta w_1 \)
- **Node #N-1**
  - Weight \( \Delta w_{N-1} \)
- **All Reduce**
  - Size: 100GB+
  - # of Files: 1M

Sharing the new weight values.
The Impact of Delayed I/O Requests

- A single delayed I/O can block the entire training process
- A few seconds of tail latency is harmful due to the shortened training time
  - If the training takes **26 hours**, a 2 seconds delay does not matter.
  - However, a 2 seconds delay in a **1-minute** training process is **significant**.

![Diagram showing synchronized processing sensitivity to delayed I/Os]

Synchronized processing makes the workload more sensitive to delayed I/Os.
Observation: Tail Latency of Storage Systems

Latency distribution when 16 processes read the ImageNet data set from the shared storage system

Measured on Cygnus Supercomputer at Univ. of Tsukuba

There are very rare (<0.01%) long-delayed I/O requests on a shared storage system

>99.99% of requests complete in 0.5 sec
<0.01% of requests take more than 1.5 sec
I/O Latency Aware Dynamic I/O Replacement

- Monitoring all read I/O operations and...
  - Discarding/replacing the delayed I/O with alternative data sets
  - Managing the history of data replacement to minimize the impact on the training.

Replacing the training data set when the measured latency exceeds the threshold.

Avoiding the impact of (rare) long-delayed I/O with the negligible drawback in precision of the trained model.

< 0.01 % of total I/Os
Conclusion and Future work

- I/O tail latency has a significant impact on the performance of massive scale fast deep learning applications.

- Shared storage systems have very rare long tail latency, that affects to the performance of distributed deep learning workloads.

- We are developing a method to avoid the impact of delayed I/O requests on training processes.

  Dynamic I/O Replacement
  - Replacing less than 0.01 % of total I/O operations.
  - Adaptive data replacement to minimize the drawbacks in the precision of the trained model.
Fujitsu

shaping tomorrow with you