

DDP Communication Hooks

DDP communication hook is a generic interface to control how to communicate gradients across workers by overriding the vanilla allreduce in [DistributedDataParallel](#). A few built-in communication hooks are provided, and users can easily apply any of these hooks to optimize communication. Besides, the hook interface can also support user-defined communication strategies for more advanced use cases.

How to Use a Communication Hook?

To use a communication hook, the user just needs to let the DDP model register the hook before the training loop as below.

```
:func:`torch.nn.parallel.DistributedDataParallel.register_comm_hook`
```

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master] [docs] [source] ddp_comm_hooks.rst, line 18); [backlink](#)

Unknown interpreted text role "func".

What Does a Communication Hook Operate On?

Communication hook provides a flexible way to allreduce gradients. Therefore, it mainly operates on the gradients on each replica before allreduce, which are bucketized to increase the overlap between communication and computation. Particularly, `:class:`torch.distributed.GradBucket`` represents a bucket of gradient tensors to be allreduced.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master] [docs] [source] ddp_comm_hooks.rst, line 23); [backlink](#)

Unknown interpreted text role "class".

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master] [docs] [source] ddp_comm_hooks.rst, line 28)

Unknown directive type "autoclass".

```
.. autoclass:: torch.distributed.GradBucket
```

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master] [docs] [source] ddp_comm_hooks.rst, line 30)

Unknown directive type "autofunction".

```
.. autofunction:: torch.distributed.GradBucket.index
```

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master] [docs] [source] ddp_comm_hooks.rst, line 31)

Unknown directive type "autofunction".

```
.. autofunction:: torch.distributed.GradBucket.buffer
```

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master] [docs] [source] ddp_comm_hooks.rst, line 32)

Unknown directive type "autofunction".

```
.. autofunction:: torch.distributed.GradBucket.gradients
```

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master] [docs] [source] ddp_comm_hooks.rst, line 33)

Unknown directive type "autofunction".

```
.. autofunction:: torch.distributed.GradBucket.is_last
```

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master][docs][source]ddp_comm_hooks.rst, line 34)

Unknown directive type "autofunction".

```
.. autofunction:: torch.distributed.GradBucket.set_buffer
```

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master][docs][source]ddp_comm_hooks.rst, line 35)

Unknown directive type "autofunction".

```
.. autofunction:: torch.distributed.GradBucket.parameters
```

Default Communication Hooks

Default communication hooks are simple **stateless** hooks, so the input state in `register_comm_hook` is either a process group or None. The input bucket is a `:class:`torch.distributed.GradBucket`` object.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master][docs][source]ddp_comm_hooks.rst, line 40); [backlink](#)

Unknown interpreted text role "class".

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master][docs][source]ddp_comm_hooks.rst, line 44)

Unknown directive type "currentmodule".

```
.. currentmodule:: torch.distributed.algorithms.ddp_comm_hooks.default_hooks
```

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master][docs][source]ddp_comm_hooks.rst, line 45)

Unknown directive type "autofunction".

```
.. autofunction:: allreduce_hook
```

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master][docs][source]ddp_comm_hooks.rst, line 46)

Unknown directive type "autofunction".

```
.. autofunction:: fp16_compress_hook
```

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master][docs][source]ddp_comm_hooks.rst, line 47)

Unknown directive type "autofunction".

```
.. autofunction:: bf16_compress_hook
```

Additionally, a communication hook wrapper is provided to support `:meth:`~fp16_compress_hook`` or `:meth:`~bf16_compress_hook`` as a wrapper, which can be combined with other communication hooks.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master][docs][source]ddp_comm_hooks.rst, line 49); [backlink](#)

Unknown interpreted text role "meth".

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master][docs][source]ddp_comm_hooks.rst, line 49); [backlink](#)

Unknown interpreted text role "meth".

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master] [docs] [source] ddp_comm_hooks.rst, line 52)

Unknown directive type "autofunction".

```
.. autofunction:: fp16_compress_wrapper
```

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master] [docs] [source] ddp_comm_hooks.rst, line 53)

Unknown directive type "autofunction".

```
.. autofunction:: bf16_compress_wrapper
```

PowerSGD Communication Hook

PowerSGD (Vogels et al, [NeurIPS 2019](#)) is a gradient compression algorithm, which can provide very high compression rates and accelerate bandwidth-bound distributed training. This algorithm needs to maintain both some hyperparameters and the internal state. Therefore, PowerSGD communication hook is a **stateful** hook, and the user needs to provide a state object defined as below.

PowerSGD State

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master] [docs] [source] ddp_comm_hooks.rst, line 68)

Unknown directive type "currentmodule".

```
.. currentmodule:: torch.distributed.algorithms.ddp_comm_hooks.powerSGD_hook
```

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master] [docs] [source] ddp_comm_hooks.rst, line 69)

Unknown directive type "autoclass".

```
.. autoclass:: PowerSGDState
```

PowerSGD Hooks

Warning

PowerSGD typically requires extra memory of the same size as the model's gradients to enable error feedback, which can compensate for biased compressed communication and improve accuracy.

Warning

PowerSGD hooks may conflict with [Apex automatic mixed precision package](#). Please use PyTorch [native automatic mixed precision package](#) instead.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master] [docs] [source] ddp_comm_hooks.rst, line 84)

Unknown directive type "autofunction".

```
.. autofunction:: powerSGD_hook
```

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master] [docs] [source] ddp_comm_hooks.rst, line 85)

Unknown directive type "autofunction".

```
.. autofunction:: batched_powerSGD_hook
```

Debugging Communication Hooks

As the name implies, debugging communication hooks are **only** used for debugging and performance optimization purpose.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master][docs][source]ddp_comm_hooks.rst, line 92)

Unknown directive type "currentmodule".

```
.. currentmodule:: torch.distributed.algorithms.ddp_comm_hooks.debugging_hooks
```

Warning

Debugging communication hooks do not necessarily output the correct results.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master][docs][source]ddp_comm_hooks.rst, line 97)

Unknown directive type "autofunction".

```
.. autofunction:: noop_hook
```

Acknowledgements

Many thanks to PowerSGD paper author **Thijs Vogels** for the code review on PowerSGD communication hook, as well as the [comparison experiments](#), which show that the performance of PowerSGD communication hook is on par with the implementation in the original [paper](#).