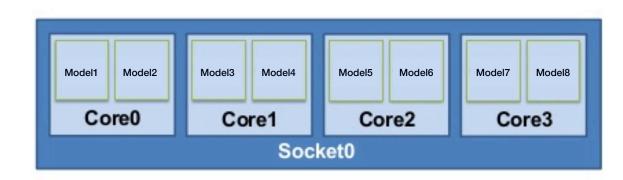
Parallel Training on CPU Threads

- Train one distinct model per thread
- Create as many Ray tasks as your CPU threads (24)
- 2hr scan evaluated 27 NNs

On atlaslogin02

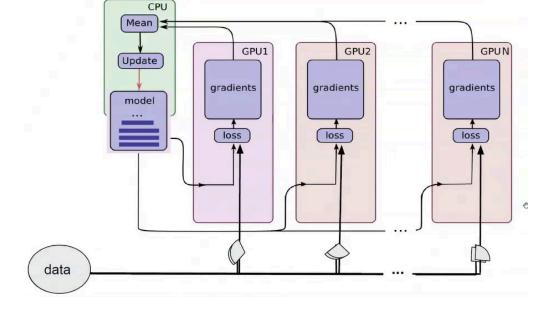
- CPU threads = 48
 Although not all of them were running:
 RuntimeError: Resource temporarily unavailable
- 2hr scan evaluated 67 NNs

ACHINE VIEW LOGICAL VIEW	RAY CONFIG								
Host	Workers	Uptime	CPU	RAM	Disk	Sent	Received	Logs	Errors
atlasdpb0.hep.anl.gov (130.202.17	3.9) 24 workers / 24 cores	5d 20h 29m 05s	99.2%	21.1 GB / 62.8 GB (34%)	26.7 GiB / 115.4 GiB (23%)	0.1 MiB/s	0.0 MiB/s	View all logs (818 lines)	No errors
	object_store_memory: 0 GiB /	12.74 GiB, memory:	0 G1B / 36.94	5 G1B, CPU: 0 / 24, GPU: 0 /	3, node:130.202.173.9: 0 / 1				
ray (PID: 30903)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	130.4%	1018.7 MiB	N/A	N/A	N/A	View log (42 lines)	No errors
ray (PID: 30904)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	117.5%	895.9 MiB	N/A	N/A	N/A	View log (40 lines)	No errors
ray (PID: 30905)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	29.4%	830.6 MiB	N/A	N/A	N/A	View log (24 lines)	No errors
ray (PID: 30906)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	76.7%	929.7 MiB	N/A	N/A	N/A	View log (30 lines)	No errors
ray (PID: 30907)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	132.0%	1677.3 MiB	N/A	N/A	N/A	View log (42 lines)	No errors
ray (PID: 30908)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	132.0%	909.6 MiB	N/A	N/A	N/A	View log (32 lines)	No errors
ray (PID: 30909)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	61.8%	906.5 MiB	N/A	N/A	N/A	View log (32 lines)	No errors
ray (PID: 30910)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	123.6%	929.4 MiB	N/A	N/A	N/A	View log (28 lines)	No errors
ray (PID: 30911)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	140.7%	1653.9 MiB	N/A	N/A	N/A	View log (42 lines)	No errors
ray (PID: 30912)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	49.6%	671.6 MiB	N/A	N/A	N/A	View log (30 lines)	No errors
ray (PID: 30913)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	125.3%	1279.6 MiB	N/A	N/A	N/A	View log (40 lines)	No errors
ray (PID: 30914)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	23.4%	753.3 MIB	N/A	N/A	N/A	View log (26 lines)	No errors
ray (PID: 30915)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	19.4%	501.3 MiB	N/A	N/A	N/A	View log (30 lines)	No errors
ray (PID: 30916)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	111.1%	833.4 MiB	N/A	N/A	N/A	View log (42 lines)	No errors
ray (PID: 30917)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	135.9%	972.5 MiB	N/A	N/A	N/A	View log (38 lines)	No errors
ray (PID: 30918)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	46.0%	823.2 MiB	N/A	N/A	N/A	View log (34 lines)	No errors
ray (PID: 30923)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	129.7%	1116.4 MiB	N/A	N/A	N/A	View log (34 lines)	No errors
ray (PID: 30924)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	136.7%	973.8 MiB	N/A	N/A	N/A	View log (34 lines)	No errors
ray (PID: 30925)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	92.1%	814.7 MiB	N/A	N/A	N/A	View log (32 lines)	No errors
ray (PID: 30927)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	31.8%	658.2 MIB	N/A	N/A	N/A	View log (24 lines)	No errors
ray (PID: 30928)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	134.1%	941.6 MiB	N/A	N/A	N/A	View log (40 lines)	No errors
ray (PID: 30929)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	148.4%	1104.8 MiB	N/A	N/A	N/A	View log (38 lines)	No errors
ray (PID: 30931)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	95.4%	1090.0 MiB	N/A	N/A	N/A	View log (26 lines)	No errors
ray (PID: 30932)	deephyper.evaluator.ray_evaluator.compute_objective()	00h 21m 16s	140.8%	1060.4 MiB	N/A	N/A	N/A	View log (38 lines)	No errors
Totals (1 host)	24 workers / 24 cores	N/A	99.2%	21.1 G8 / 62.8 G8 (34%)	26.7 Gi8 / 115.4 Gi8 (23%)	0.1 Mi8/s	0.0 MiB/s	818 lines	No errors



Distributed Training on GPUs

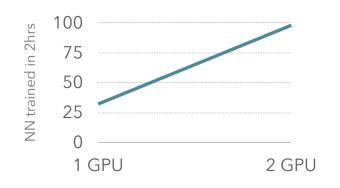
- Train one distinct model and split the data (batch) per GPU
- Using <u>TF MirroredStrategy</u>
- Advantage: process larger batch sizes, i.e. larger *effective* GPU memory
- 2hr scan evaluated 24 NNs

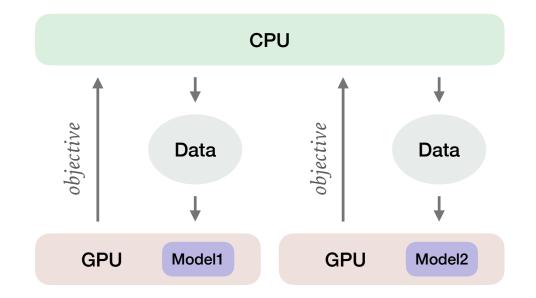


Problem: TF is not cleaning the GPU memory after each Model training

Parallel Training on GPUs

- Parallelize the HPS similar to the CPU threads not the training
- Create as many Ray tasks as your GPUs
- Each task sees only 1 GPU
- Results are aggregated in CPU
- 2hr scan evaluated **98 NNs**
- Scaling "linear":





8

Problem: TF is not cleaning the GPU memory after each Model training