

System : DGX A100

CPU : (2) Dual AMD Rome 7742 128 cores total, 2.25 GHz(base), 3.4 GHz (max boost)

Mem : 1TB

GPU : (8) NVIDIA A100 Tensor Core 40 GB

OS : DGX OS 5.0

Driver : 450.80.02 / CUDA : 11.2

Anaconda 4.9.2

tensorflow-gpu 2.4.1

Date : 2021-03

model	Batch Size	GPUs				Speed Up Ratio			
		1	2	4	8	1gpu/1gpu	2gpu/1gpu	4gpu/1gpu	8gpu/1gpu
inception4	32	252	447	761	1224	1.0	1.8	3.0	4.9
inception4	64	286	539	978	1625	1.0	1.9	3.4	5.7
inception3	32	474	786	1321	2089	1.0	1.7	2.8	4.4
inception3	64	535	993	1826	3169	1.0	1.9	3.4	5.9
alexnet	32	4644	8798	16578	22505	1.0	1.9	3.6	4.8
alexnet	64	6537	12799	24830	39477	1.0	2.0	3.8	6.0
resnet50	32	678	1226	2121	3213	1.0	1.8	3.1	4.7
resnet50	64	786	1492	2726	4491	1.0	1.9	3.5	5.7
resnet50_v2	32	698	1249	2219	3462	1.0	1.8	3.2	5.0
resnet50_v2	64	805	1528	2819	4957	1.0	1.9	3.5	6.2
resnet101	32	414	724	1151	1701	1.0	1.8	2.8	4.1
resnet101	64	488	901	1596	2660	1.0	1.8	3.3	5.4
resnet101_v2	32	423	739	1200	1836	1.0	1.7	2.8	4.3
resnet101_v2	64	498	917	1638	2712	1.0	1.8	3.3	5.4
resnet152	32	279	486	781	1172	1.0	1.7	2.8	4.2
resnet152	64	340	627	1066	1831	1.0	1.8	3.1	5.4
resnet152_v2	32	295	502	820	1271	1.0	1.7	2.8	4.3
resnet152_v2	64	346	633	1115	1883	1.0	1.8	3.2	5.4
vgg16	32	566	1094	2172	3824	1.0	1.9	3.8	6.8
vgg16	64	609	1196	2343	4407	1.0	2.0	3.8	7.2
vgg19	32	480	955	1770	3055	1.0	2.0	3.7	6.4
vgg19	64	515	1021	1960	3640	1.0	2.0	3.8	7.1
trivial	32	34166	50915	83643	125229	1.0	1.5	2.4	3.7
trivial	64	57025	89881	152530	236302	1.0	1.6	2.7	4.1