- ➢ **Distributed Training**

- ➢ **Multiple GPUs and Machines**

- ➢ **Introduction**

- ➢ **DataParallelSGD**

- ➢ **BlockMomentumSGD**

- ➢ **ModelAveragingSGD**

- ➢ **DataParallelASGD**

- ➢ **Prerequisites**

- ➢ **Configuring Parallel Training in CNTK in Python**

- ➢ **Configuring Parallel Training in CNTK in BrainScript**

- ➢ **Running Parallel Training with CNTK**

- ➢ **Running Parallel Training with Python**

- ➢ **Data-Parallel Training with 1-bit SGD**

- ➢ **The "1-bit SGD" Algorithm**

- ➢ **Using 1-bit SGD in BrainScript**

- ➢ **Using 1-bit SGD in Python**