Meet up

Snap ML Session - Logistic Regression in Tensor Flow (CPU) Vs Snap ML (GPU's)

- Thursday, February 28, 2019
  5:00 PM to 6:00 PM

(IBM Building atrium)

590 Madison/57th · New York, NY

Agenda

- Introduction
- Current Affairs
- Generalized Linear Models – 5 min
- Getting started with Snap ML– 5 min
- IBM Spectrum Conductor-5 min
- Python Modules -5 min
- Model Building Example – 10 min
- Training a multi-class logistic regression classifier — 30 min

Introductions

Network with IBM team present at the event

Discuss a 'Use Case'
Extend a LinkedIn Invite
Connect on twitter

- Technical team Leader

John Ernst
Technical BUE COM/CSI Market
IBM Global Markets - Systems HW Sales

Indianapolis, IN, United States 2:05 PM | 14:05
Building: 0876 Office: MOBILE

- Project head

Joy Lee
Systems Architect and Systems Integrator Architect Manager, US Comm/CSI Market
IBM Global Markets - Systems HW Sales

Coppell, TX, United States 1:05 PM | 13:05
Building: DFW Office: MOBILE

- Speaker

Gautam Siwach
Comms/Csi, Cognitive solutions leader - North America
IBM Global Markets - Systems HW Sales

Hartford, CT, United States 2:03 PM | 14:03
Building: 9232 Office: MOBILE

Robert Green
Systems Architect - Communications and CSI accounts
IBM Global Markets - Systems HW Sales

Poughkeepsie, NY, United States 10:53 AM | 10:53
Building: 416 Office: MOBILE

- Co-Organizer

Diane Britton
IBM North America Cognitive/Power TSE
IBM Global Markets - Systems HW Sales

Poughkeepsie, NY, United States 10:53 AM | 10:53
Building: 707 Floor: 0L Office: N02

Current Affairs

- PowerAI: The World's Fastest Deep Learning Solution Among Leading Enterprise Servers
- Deep Learning Software Distro
- Evaluating Julia for Deep Learning on Power + NVIDIA Tesla

IBM Power Systems AC922: The best server for enterprise AI

Announcing the next generation in accelerated computing

Read the press release

Trending in Tech

1 TECH NEWS
Huawei breaks price ceiling with US$2,600 folding 5G smartphone

2 TECH NEWS
China's technology challenge is bigger than just Huawei, British

4 TECH NEWS
Let Google Assistant help you with your Messages

5 TECH NEWS
Samsung Galaxy S10: Ever evolving

6 TECH NEWS
China's Xiaomi unveils US$680 smartphone, sees growth in Africa

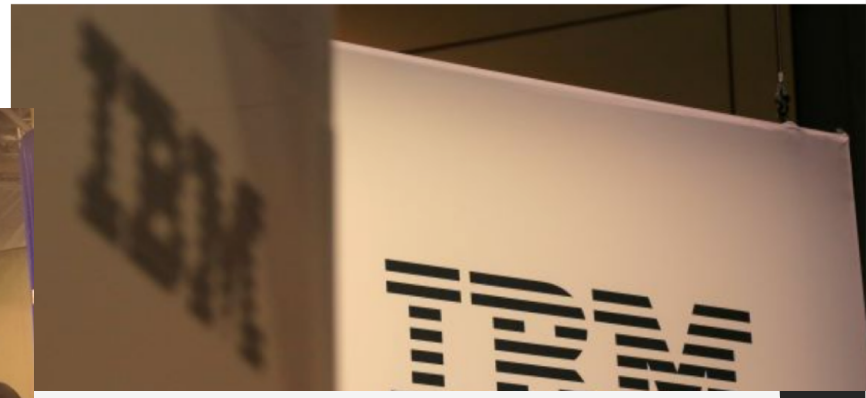
Star ONLINE | News | Business | Sport | Metro | Tech | Lifestyle | Opinion | Videos | Property | Jobs

TOPICS ▸ Asean+ | True or Not | Do You Know | Star Golden Hearts Award | SOBA 2018

IBM develops new technology to help prevent power outages

TECH NEWS

Thursday, 7 Feb 2019



Trending topics about technology

Top 10 Trending Technologies

- Artificial Intelligence.
- Blockchain.
- Augmented Reality and Virtual Reality.
- Cloud Computing.
- Angular and React.
- DevOps.
- Internet of Things (IoT)
- Intelligent Apps (I – Apps)

Future is here ! Welcome to the world of Accelerated Computing.

Watson Machine Learning Accelerator for Enterprise AI

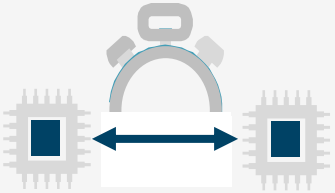
flattening the deep learning time to value curve

- A single binary download of the most ubiquitous open source AI and major deep learning frameworks packages that are all precompiled with of the supporting software that they require to run.

Enterprise Deep Learning Distribution

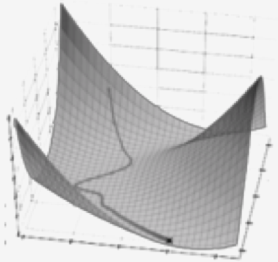
Enterprise Software Distribution

- PowerAI S/W
- Support



Performance (Faster Training Times for Data Scientists)

- Distributed Deep Learning
- Large Model Support



Tools for Ease of Development

- PowerAI VISION
- Apache-Spark Data ETL and Preparation Tool
- DL Insight

PowerAI is the "Red Hat" of Deep Learning

- Simplified deployment, integrated support, excellent performance

Key Benefits

- Offers simplified maintenance
  - Provides enterprise multi-tenancy and dependability
  - Protected by IBM support
  - Efficient and cost-effective
  - Provides flexibility to run additional applications
- Incorporates the most popular deep learning frameworks

Enables rapid deep learning deployment

Accelerate time to deploy deep learning workloads and get results quicker from initial prototype to enterprise-wide deployment.

Delivers faster insights at scale

Exploit unique data architecture and turn deep learning performance into value with IBM Power Systems and other available hardware.

Supports larger, more accurate models

Leverage Large Model Support and Distributed Deep Learning to support models at higher resolution, up to almost any scale.

Business-ready and dependable

Expand your deep learning deployment to multiple users and lines of business with multi-tenancy and role-based access controls. Protect them with end-to-end security.

Snap Machine Learning

Generalized Linear Models

IBM PowerAI 1.5.3

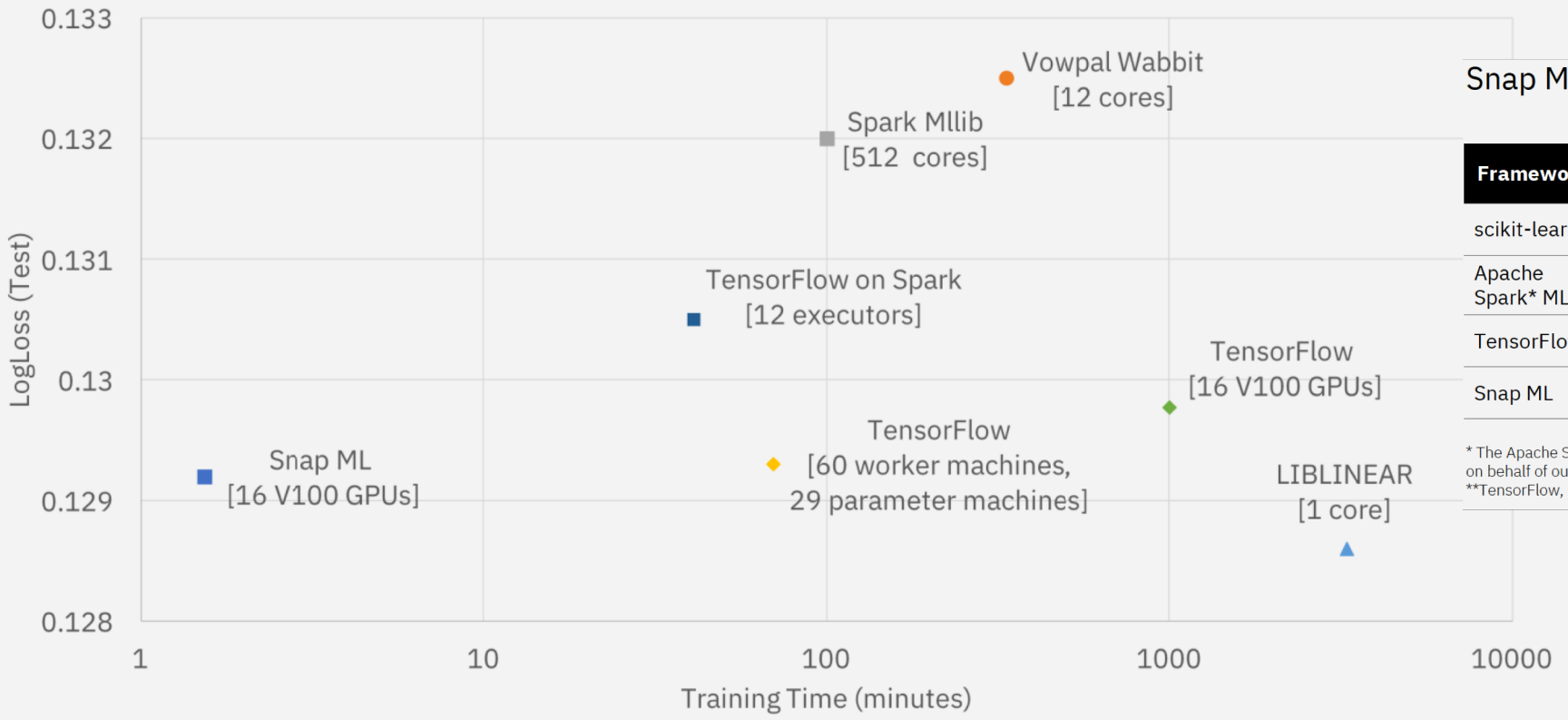
GLM
Regression and Classification

1. Snap Machine Learning (*Snap ML*) is a library for training generalized linear models. It is being developed at the IBM Research - Zürich laboratory
2. A highly accelerated, scalable software library for Machine Learning
3. Snap ML Bests Tensor flow in benchmarks

snap-ml-local **snap-ml-mpi** **snap-ml-spark**

- snap-ml-local is used for machine learning on a single machine
- snap-ml-mpi is used for distributed training of machine learning models across a cluster of machines.
- Similar to snap-ml-mpi, the snap-ml-spark package offers distributed training of models across a cluster of machines. The library is exposed to the user through a spark.ml-like interface and can seamlessly be integrated into existing pySpark application.

Dataset	Examples	Nodes	Total GPUs	Network	CPU-GPU interface
criteo-tb	4.2 billion	4x Power AC922 server	16x V100	InfiniBand	NVLINK 2.0



Snap ML: A new framework for fast training of GLMs

Framework	Models	GPU Acceleration	Distributed Training	Sparse Data Support
scikit-learn	ML/{DL}	No	No	Yes
Apache Spark* MLlib	ML/{DL}	No	Yes	Yes
TensorFlow**	ML	Yes	Yes	Limited
Snap ML	GLMs	Yes	Yes	Yes

* The Apache Software Foundation (ASF) owns all Apache-related trademarks, service marks, and graphic logos on behalf of our Apache project communities, and the names of all Apache projects are trademarks of the ASF.
**TensorFlow, the TensorFlow logo and any related marks are trademarks of Google Inc.

Terabyte Scale Benchmark

EXAMPLE – CLICK THROUGH RATE (CTR) PREDICTION

Comparison of Tensorflow on Google Cloud with Snap ML on POWER9 (AC922) cluster

Workload:	Click-through-rate prediction for computational advertising, using Logistic Regression
Dataset:	Criteo Terabyte Click Logs, 4.2 billion training examples, 1 million features
Model:	Logistic Regression

Test LogLoss: 0.1293 (Tensorflow), 0.1292 (Snap ML)

Platform:	89 machines (Tensorflow) compared to 8 Power9 CPUs + 16 NVIDIA Tesla V100 GPUs (Snap ML)

IBM **Research**

CASE STUDY- SNAP ML PERFORMANCE EVALUATION ON MARKETING & ADVERTISEMENT DATASET

Case Study Question

Do the people visiting a web page corresponds to click an advertisement (calculate CTR)?

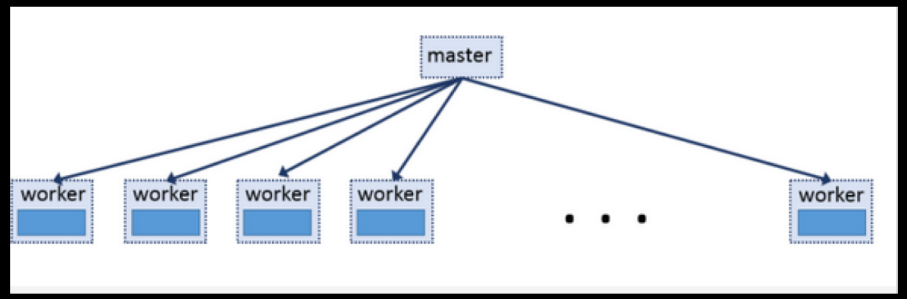
- Data Source : **Kaggle Display Advertising Challenge Dataset**

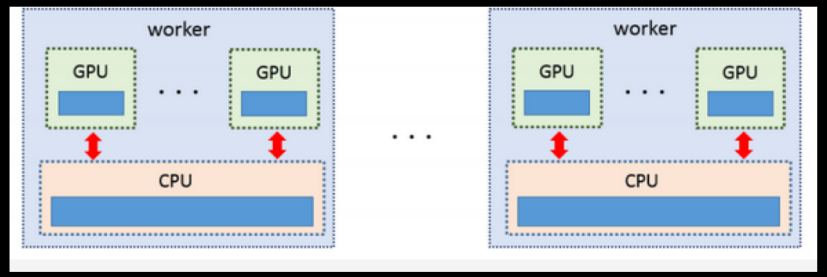
Each row corresponds to a display ad served by Criteo. Positive (clicked) and negatives (non-clicked) examples have both been subsampled at different rates in order to reduce the dataset size. The examples are chronologically ordered.

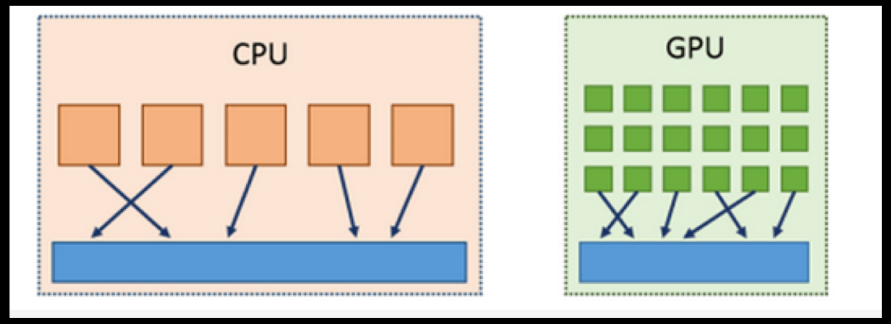
- Website : http://labs.criteo.com/2014/02/download-kaggle-display-advertising-challenge-dataset/
- **Label** - Target variable that indicates if an ad was clicked (1) or not (0); 13 columns of integer features ; A total of 26 columns of categorical features.

System description

- Level – 1 : Parallelism across worker nodes connected via network Interface
- Level -2 : Parallelism across GPUs (heterogeneous compute units), connected within the same (worker) node connected through NV Link
- Level 3 : Multi –core Parallelism within individual Compute units (GPU's)

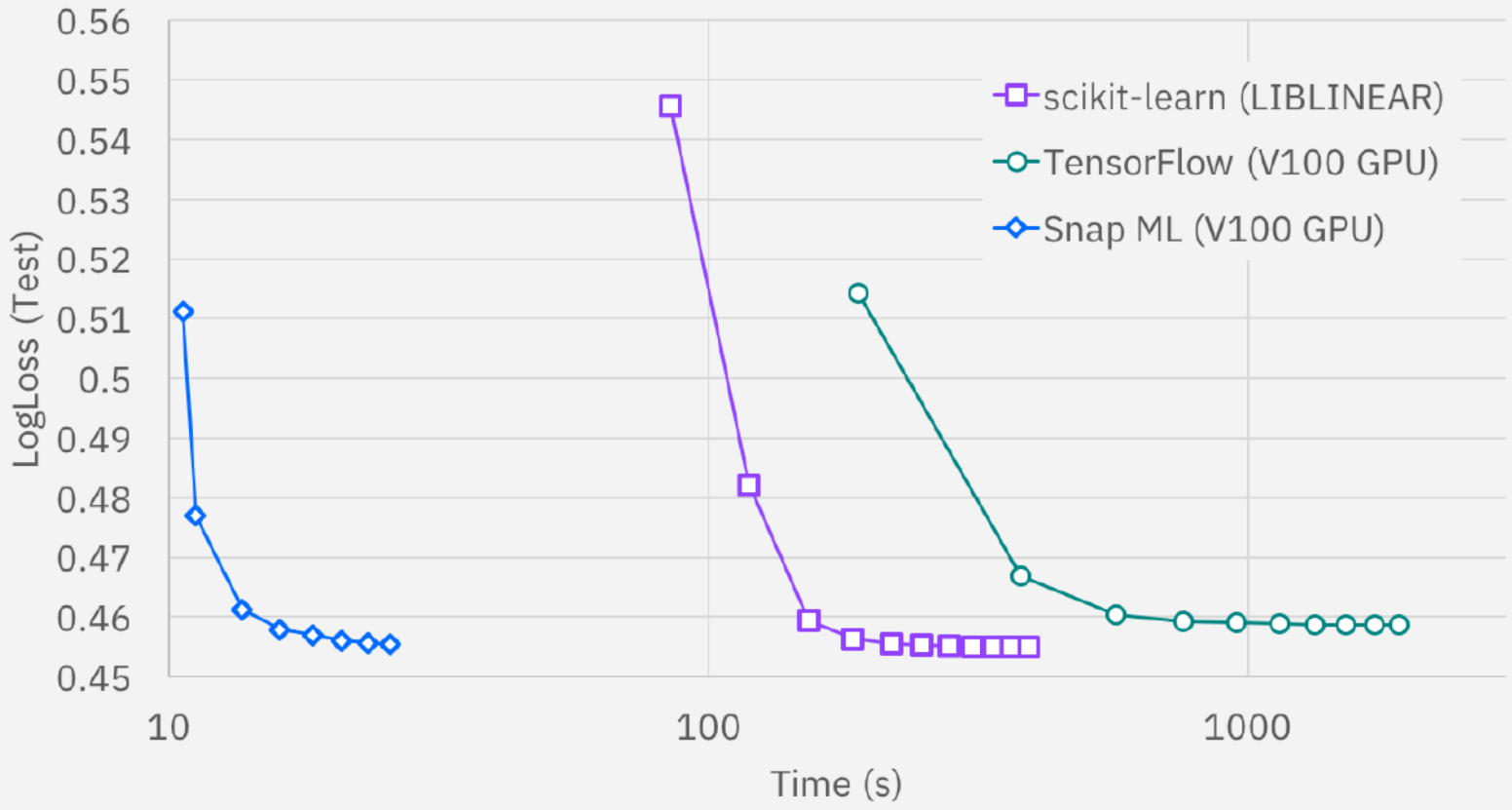






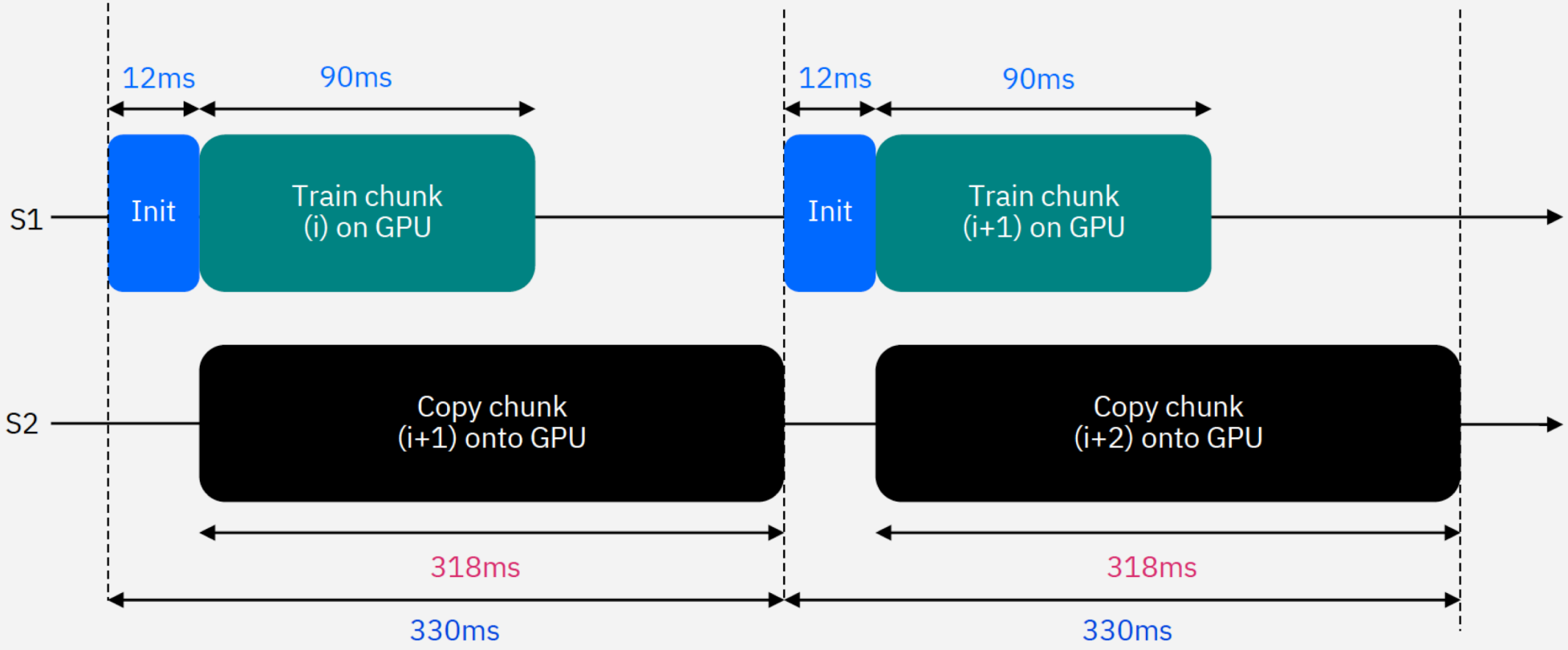
Single node experiment – Performance evaluation

Dataset	Examples	Nodes	Total GPUs	Network	CPU-GPU interface
criteo-kaggle	45 million	1x Power AC922 server	1x V100	n/a	NVLINK 2.0



Out-of-core performance (PCIe Gen3)

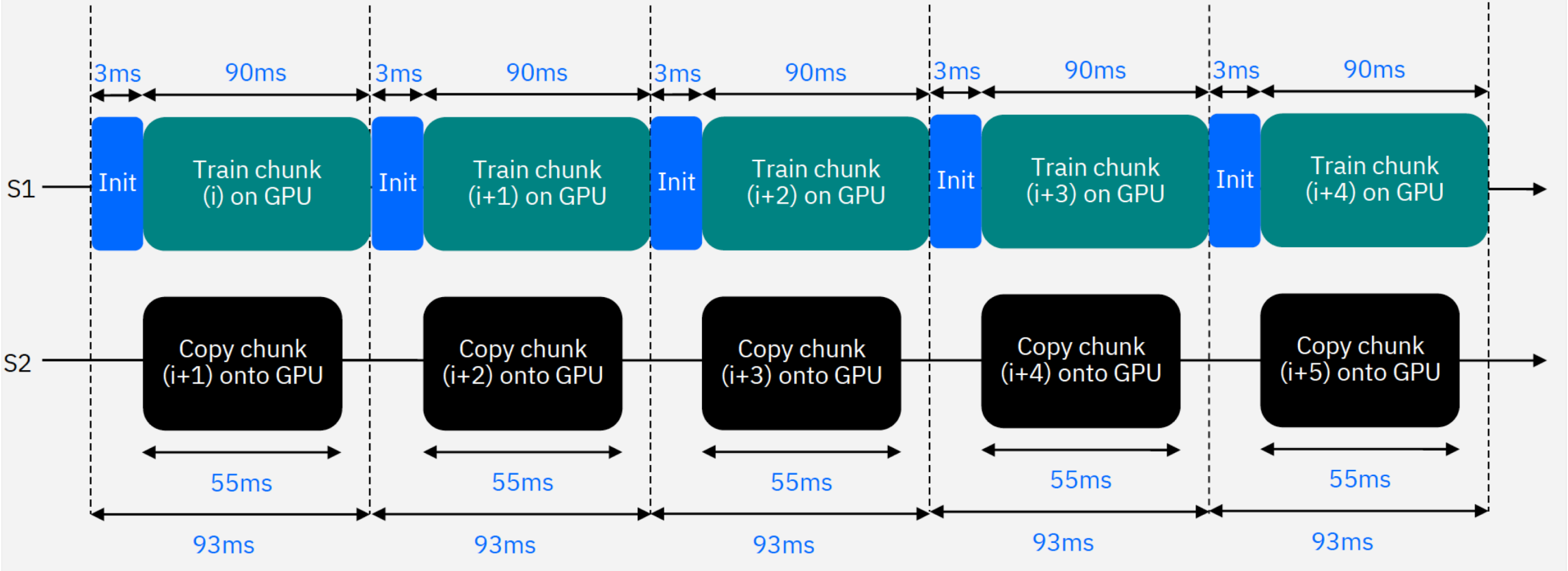
Dataset	Examples	Nodes	Total GPUs	Network	CPU-GPU interface
criteo-tb	200 million	1x Intel Xeon* Gold 6150	1x V100	n/a	PCI Gen3



*Intel Xeon is a trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

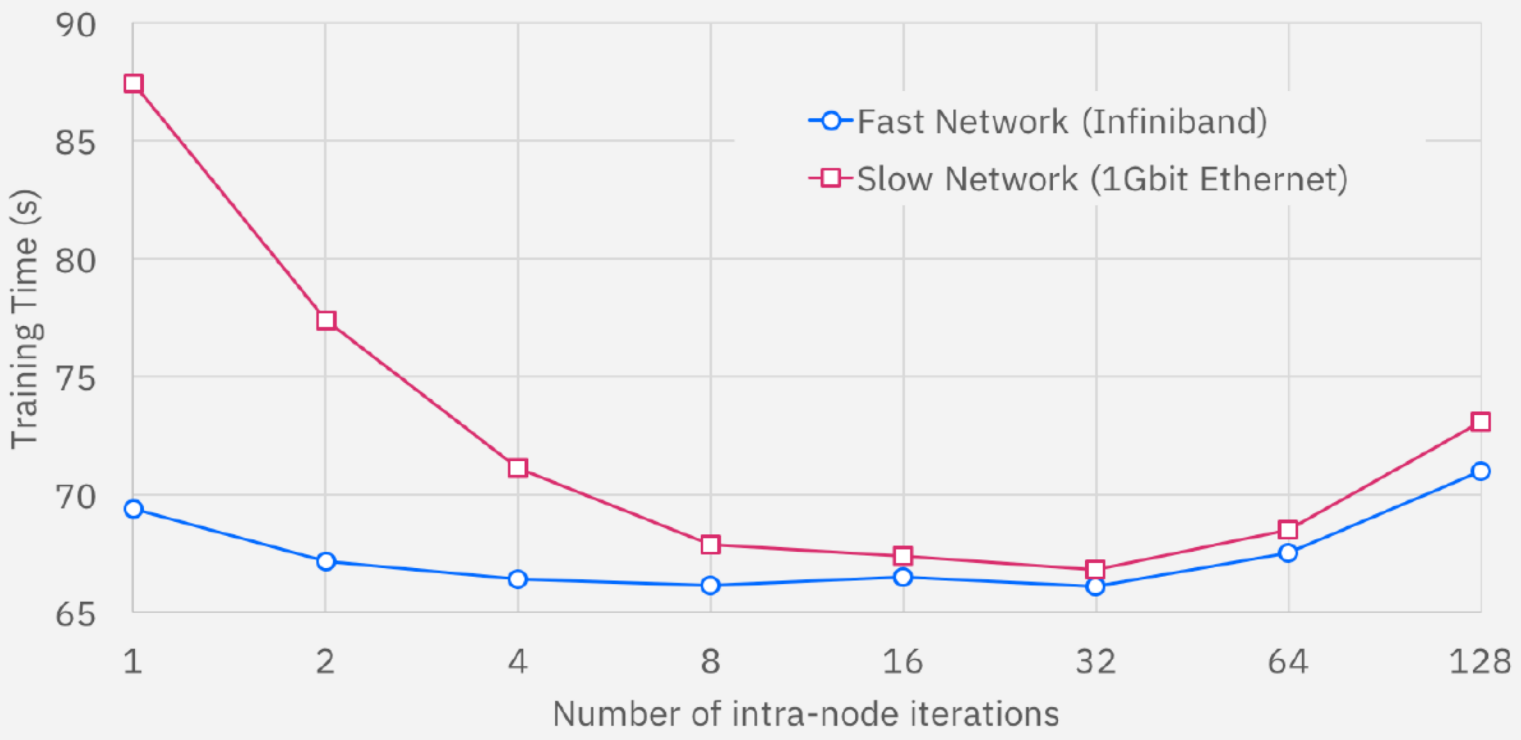
Out-of-core performance (NVLINK 2.0)

Dataset	Examples	Nodes	Total GPUs	Network	CPU-GPU interface
criteo-tb	200 million	1x Power AC922 server	1x V100	n/a	NVLINK 2.0

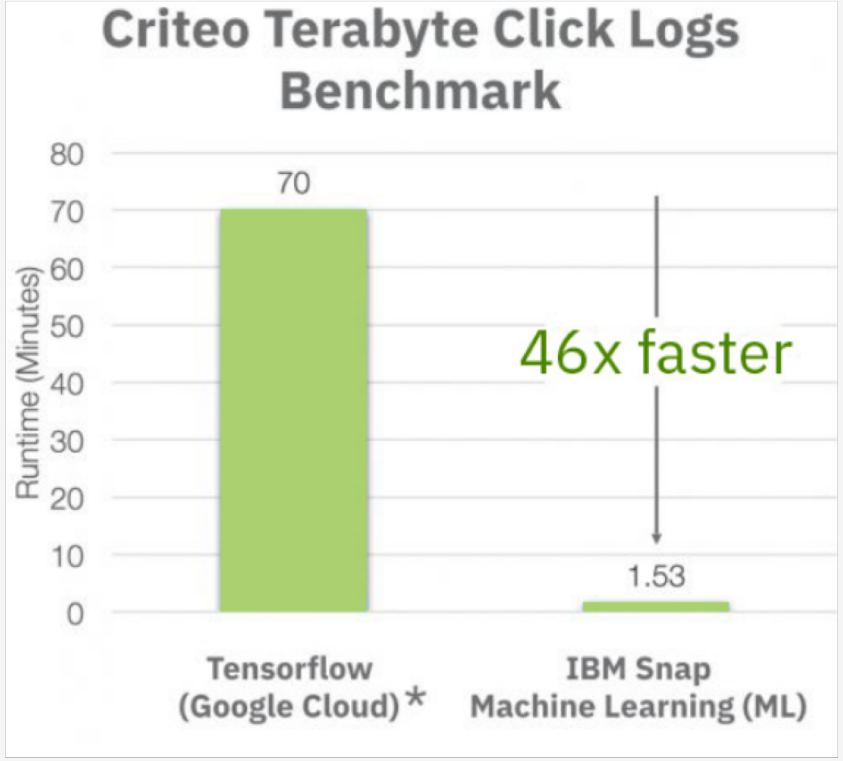


Dealing with slow networks

Dataset	Examples	Nodes	Total GPUs	Network	CPU-GPU interface
criteo-tb	1 billion	4x Power AC922 server	16x V100	1Gbit Ethernet	NVLINK 2.0
criteo-tb	1 billion	4x Power AC922 server	16x V100	InfiniBand	NVLINK 2.0



CONCLUSION



Logistic regression classifier trained in 91.5 seconds, 46× faster than the best result reported so far

Topics for discussion

- **Neural Networks:** Artificial Intelligence, Leverage Neural Networks to solve subsets of general Intelligence problems that includes perception, computer vision, prediction analysis, knowledge representation, Intelligent Agents, NLP, Automated reasoning, AI Planning (DFS is one approach to solve this), Interface for control systems.

- **Model fit /Optimization /Analytic and Simulation:** Real time decision making, autonomous agents to make real time in less time, and provide support order Actions, Hard to put a boxer on any particular problem, need to have individuals for the areas that cannot be modeled accurately that area would be useful to know.

- **Discussions on deployment and Continuous learning:** How to monitor and manage models that could go to production. Underlying the learning track for modern Data platforms, Big Data Analytic and Open Source.

- **Illustrating descriptive and diagnostic approach:** Value within a data science not just building models, but supporting decision making and automating them by linear programming type approaches; discussing initial descriptive and diagnostic effort that is tremendously valuable

Neural networks
|
Capsnet

Capsule Network : When multiple predictions agree, a higher level *capsule* becomes active.

Thank you

28
FEB

Thursday, February 28, 2019

Snap ML Session - Logistic Regression in Tensor Flow (CPU) Vs Snap ML (GPU's)

Attendees (26)

Manage

Bob Green
Event organizer
8 shared groups

Diane Britton
Co-organizer
4 shared groups

Ajit Gangaram
Member
4 shared groups

Amit Mohan Redkar
Member
3 shared groups

Tim Fiedler
Member
3 shared groups

Akshay Kadechkar
Member
2 shared groups

Bob Chambers
Member
2 shared groups

Luis Armenta
Member
2 shared groups