Sunrise or Sunset: Exploring the Design Space of Big Data Software Stack

HPBDC 2017 panel

Panel moderator: Dr. Jianfeng Zhan

Professor, ICT, Chinese Academy of Sciences and University of Chinese Academy of Sciences

May 29. 2017
Orlando , USA

中科院计算所
INSTITUTE OF COMPUTING TECHNOLOGY

The past panel (2015)

- **Wide Adoption of HPC Techniques in Big Data: Hype or Reality?**
- Panel Moderator: **Jianfeng Zhan**
- Panellists:
  - D. K. Panda, The Ohio State University
  - Dan Stanzione, Texas Advanced Computing Center
  - Zhiwei Xu, Institute of Computing Technology, Chinese Academy of Sciences, China
  - Xiaodong Zhang, The Ohio State University

The past panel (2016)

- **Merge or Split: Mutual Influence between Big Data and HPC Techniques**
- Panel Moderator: **Jianfeng Zhan**
- Panellists:
  - **Chaitanya Baru**, San Diego Supercomputer Center **Slides**
  - **Pete Beckman**, Argonne National Laboratory, The University of Chicago
  - **Andrew A. Chien**, The University of Chicago, Argonne National Laboratory **Slides**
  - **Geoffrey C. Fox**, Indiana University Bloomington **Slides**
  - **D. K. Panda**, The Ohio State University **Slides**

This year's panel

- **Sunrise or Sunset: Exploring the Design Space of Big Data Software Stack**
- Panel Moderator: **Jianfeng Zhan**
- Panellists:
  - Prof. Geoffrey C. Fox, Indiana University Bloomington
  - Prof. Satoshi Matsuoka, Tokyo Institute of Technology
  - Dr. Ren Wu, NovuMind
  - Prof. D. K. Panda, The Ohio State University

Topics

- Are big data software stacks mature or not?
  - If yes, what are the new technology challenge?
  - If not, what are the main driving forces for new-generation big data software stacks?
- What chances are provided for the academia communities in exploring the design spaces of big data software stacks?

Two driving forces

- Application-driven
  - One-size-fits-a-bunch
    - SQL, NoSQL, DW
  - Need Benchmarks
- Technology-driven
  - Super computers and Big Data enable deep learning.

BigDataBench summary

- An open-source Big Data Benchmark suite
  - http://prof.ict.ac.cn/BigDataBench

  - L. Wang, J. Zhan and etc. BigDataBench: a Big Data Benchmark Suite from Internet Services. HPCA' 14, February 15-19, 2014, Orlando, Florida, USA.
  - Gao, W., Luo, C., Zhan, J., Ye, H., He, X., Wang, L., ... & Tian, X. (2015). Identifying Dwarfs Workloads in Big Data Analytics. *arXiv preprint arXiv:1505.06872*.

BigDataBench 3.2

BDGS(Big Data Generator Suite) for scalable data

Wikipedia Entries	Amazon Movie Reviews	Google Web Graph
Facebook Social Network	E-commerce Transaction	ProfSearch Resumes
ImageNet	English broadcasting audio	DVD Input Streams
Image scene	Genome sequence data	Assembly of the human genome
SoGou Data	MNIST	MovieLens Dataset

15 个真实数据集

Search Engine | Social Network | E-commerce

Multimedia | Bioinformatics

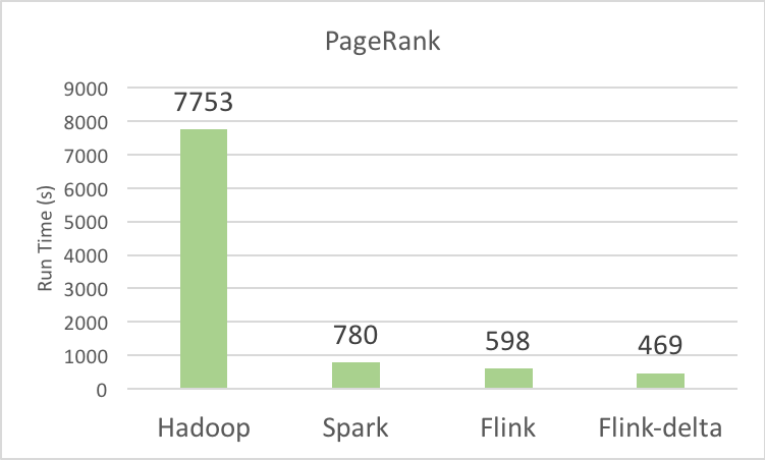
37 个负载

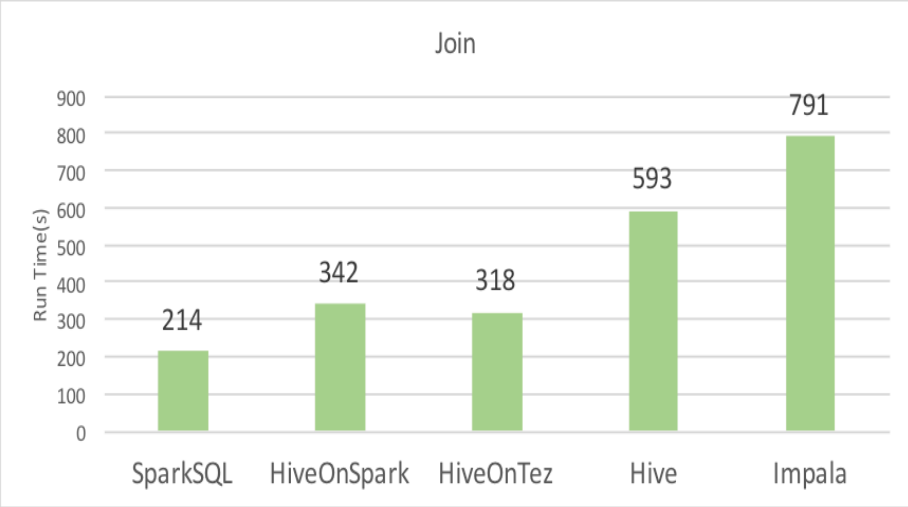
Impala, Flink, JStorm (Alibaba), NoSql, GraphX, Shark, GraphLab (Carnegie Mellon), Hive, Oracle, MySQL, Hadoop RDMA, MVAPICH, MPI, Spark (Lightning-Fast Cluster Comp), hadoop, DataMPI

软件栈

The BigData 100 project

- http://www.bafst.com/items/top100/index.html
- Using BigDataBench data sets and workloads





Requirement-Driven

- Human activities in terms of hundreds milliseconds
  - Nature of our nervous and motor systems
- Computers or smart devices consistently provide information and knowledge to human being in the order of a few tens milliseconds.
  - We coin a new term **10-ms computing** to call attention to this class of workloads
- Lu, G., Zhan, J., Hao, T., & Wang, L. (2016). 10-millisecond Computing. arXiv preprint arXiv: 1610.01267.

Millisecond-scale computing

- Grand Challenges to both big data software stack and hardware stack
- Go game is only one of benchmarks
- **Energy efficiency** of human brain!

Cost of deep learning

- https://www.reddit.com/r/MachineLearning/comments/6b64u4/d_nvidia_k80_training_time_performance/
- Imagenet 120 epochs 256 batch_size (~4k batches per epoch) ~ 3 or 4 gpus for a Resnet 50 (batchsize 64 per gpu).
  - With mxnet this setting it takes around 1.1 sec per batch 4.4K sec per epoch (say 1.2 hours) x120 epoch ---> 144hours *4 *1$ **~ 570$$ of training time** !!!
- Human being is still much expensive!

Workload driven

- Internet services have much simple workloads
  - Select, aggregation and etc.
- Hive, SparkSQL and etc
- We need consider more broader workloads

Big data dwarfs in BigDataBench

- Linear Algebra
- Sampling
- Transform operation
- Graph operation
- Logic operation
- Set operation
- Statistic operation
- Sort

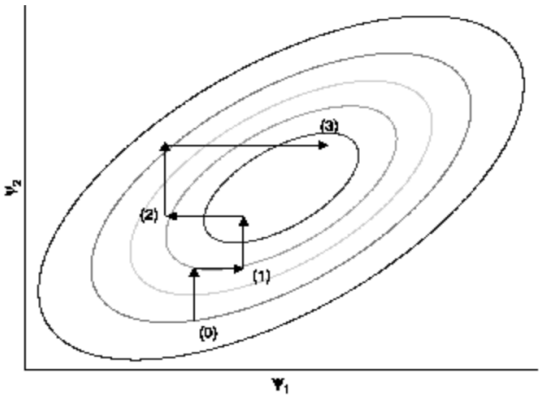
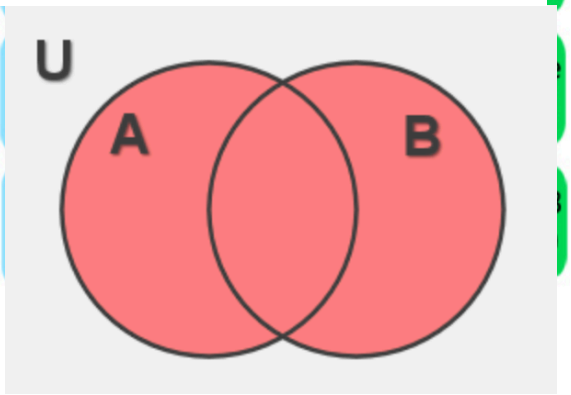


Figure 3.4: Gibbs sampling algorithm in two dimensions starting from an initial point and then completing three iterations



Paradigm change

- SMB just rent infrastructure
- Big data in cloud
- Alibaba
  - Can not tolerate 100 ms service interruption

Schedules

- Positions from the panelists (each one has 10 minutes)
- First round of rebuttals (each one has 4 minutes)
- Second round of rebuttals (each one has 4 minutes)
- Questions from the audience


QUESTIONS
And
Answers