



Deep Learning Framework Session Opening Intro

What we talk about
When we talk about
DL Framework

Biye Jiang Staff Engineer
Alimama Infar Efficiency team



Bio of myself



How to build interactive viz tools?

- imMens (Database in WebGL)
- Data illustrator @ Adobe Intern

How to make ML easy to use?

- ML Framework written in Scala
- Interactive ML @ Azure ML Intern

How to apply ML in production?

- Algorithm-System co-design
- Distributed training/serving
- Resource allocation under SLO

PhD in UC Berkeley



Alibaba Ads Team

BIDMach: An unusual ML framework

BIDMach: A Scala-written ML framework, developed by John Canny & Huasha Zhao in 2012

Scala: Great tools to make DSL

- Unified structure with compact codebase
- Flexible as C++, Maintainable as Java, Interactive as Python

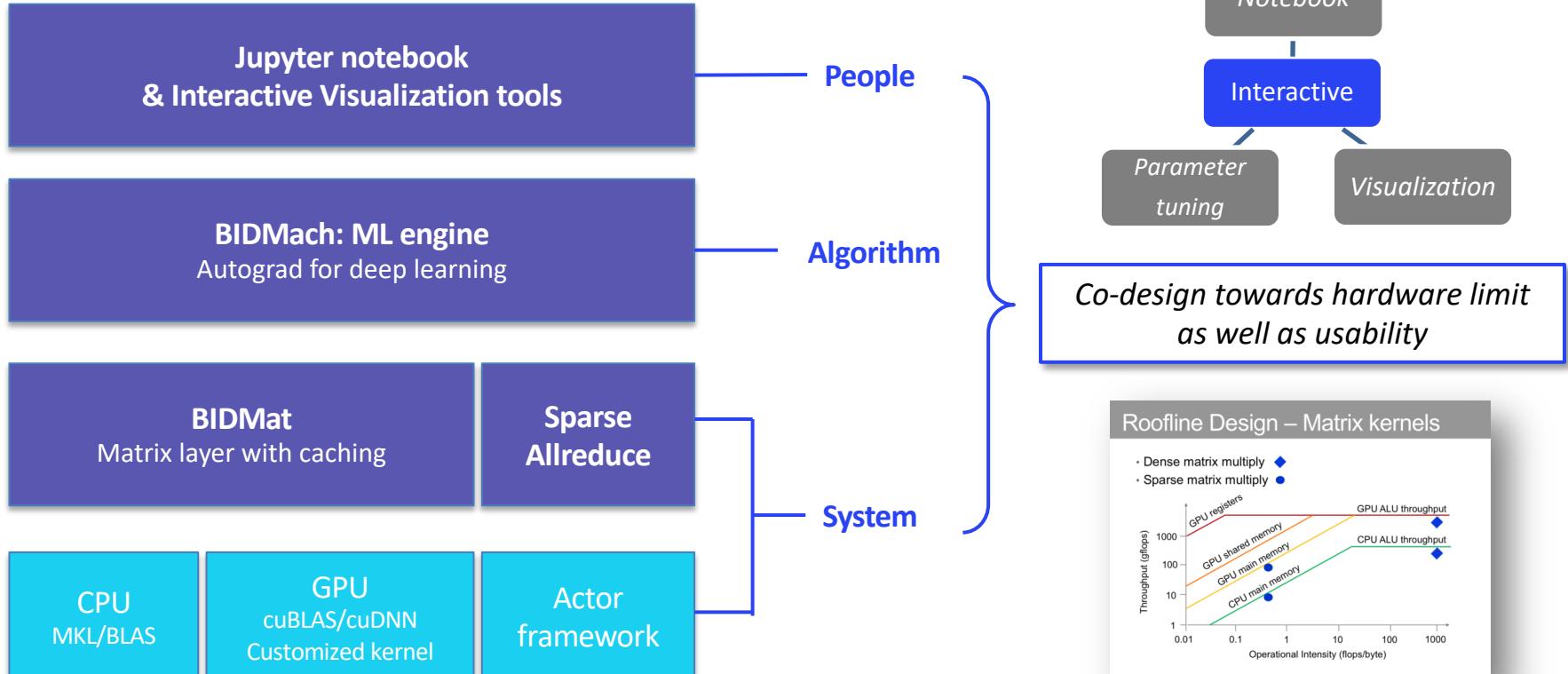


```
def eStep(sdata:Mat, user:Mat):Unit = {
    for (i <- 0 until opts.uiter) {
        val preds = SDDMM(mm, user, sdata)
        val unew = user o (mm * (sdata / preds)) + opts.alpha
        user <-> exppsi(unew)
    }
}
```

LDA e-step code example

BIDMACH

BIDMach Architecture

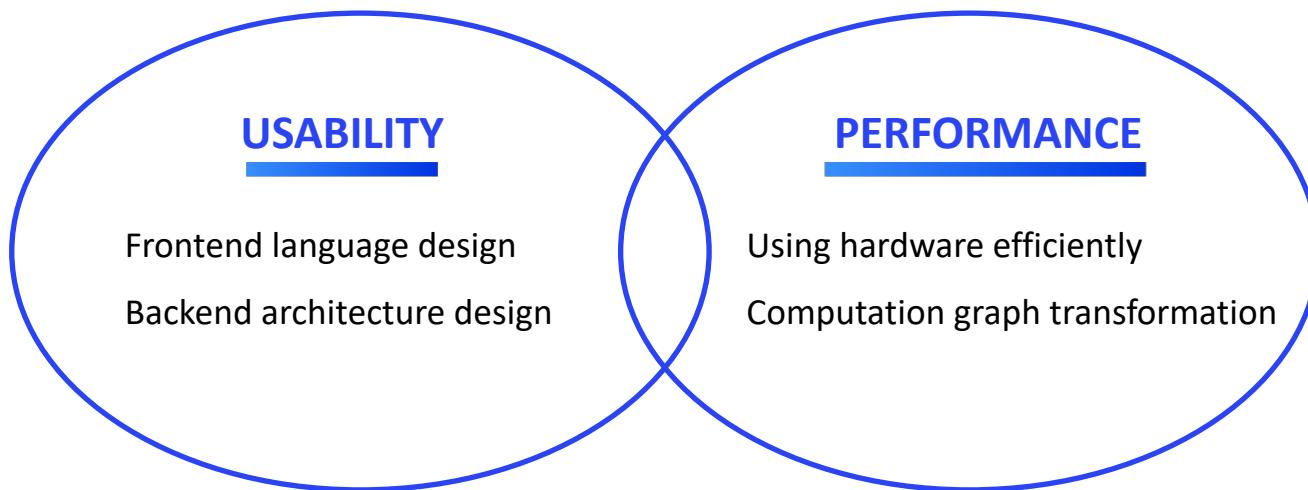


Kylix: A sparse allreduce for commodity clusters. Zhao et al. ICPP 2014
Extending the Limits of Machine Learning with GPUs @GTC2015
Interactive Machine Learning via a GPU-accelerated Toolkit. Jiang et al. IUI 2017

| So far

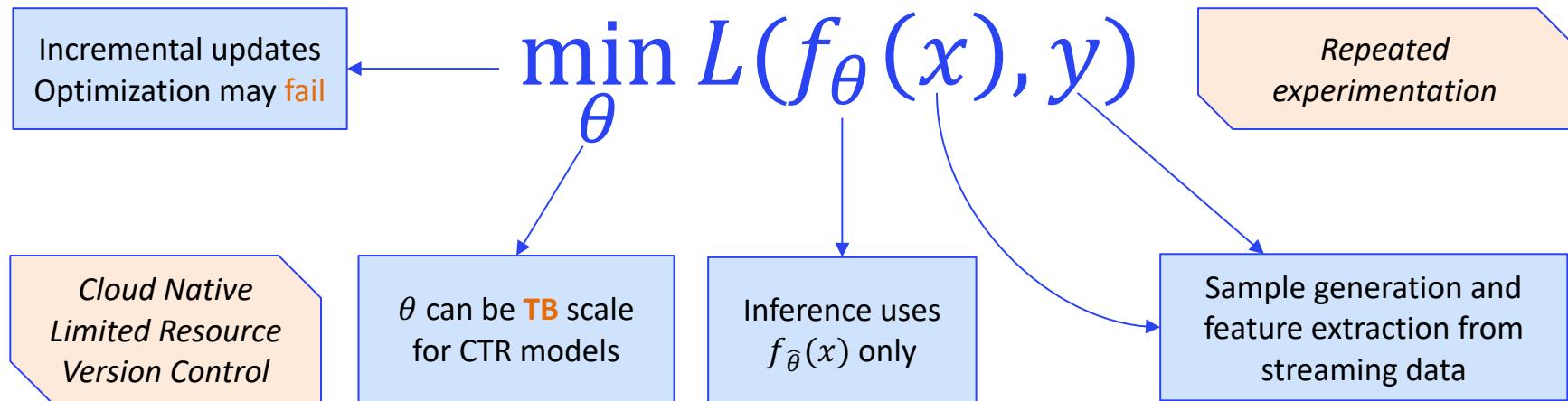
When we are talking about ML framework, we're talking about (at least):

“A piece of code to solve my problems!!!”

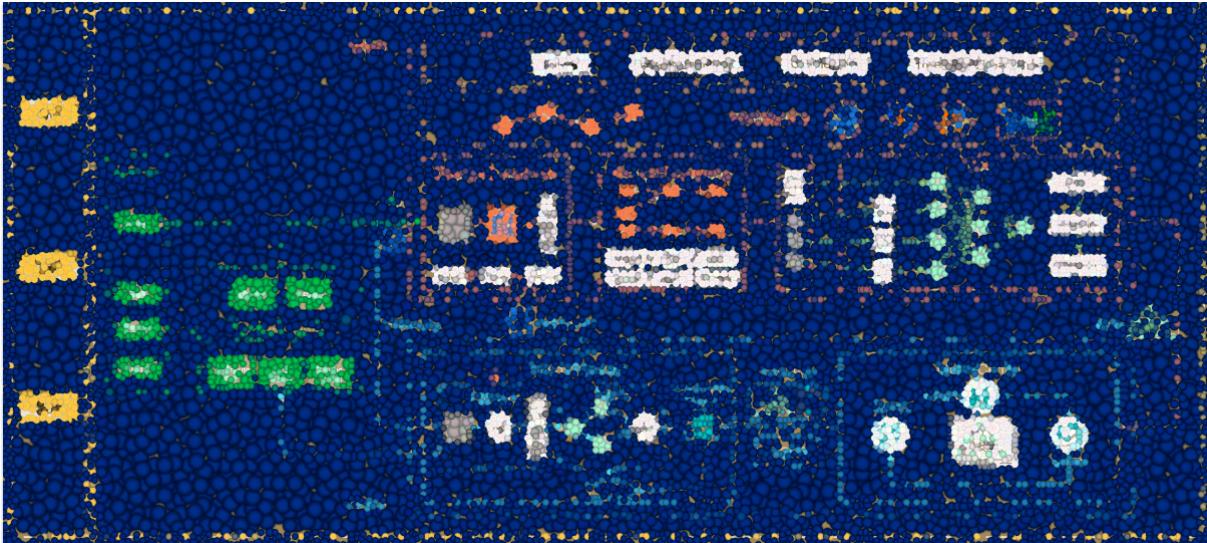


Moving into Production

- Frameworks like Tensorflow/PyTorch work very well in research
- But how does a “Framework” in Internet industry (Search Engine/Recommendation/Ads) look like?



Industrial MLSys@Alimama



How to maintain and optimize such system?

- Co-Design
- Evolution Cycles

Further Talks

- XDL -> *This afternoon*
- Bernoulli (Streaming ML) -> *Flink Forward Asia 2021*

<https://github.com/alibaba/x-deeplearning>

XDL: An Industrial Deep Learning Framework for High-dimensional Sparse Data. Jiang et al. DLP-KDD 2019

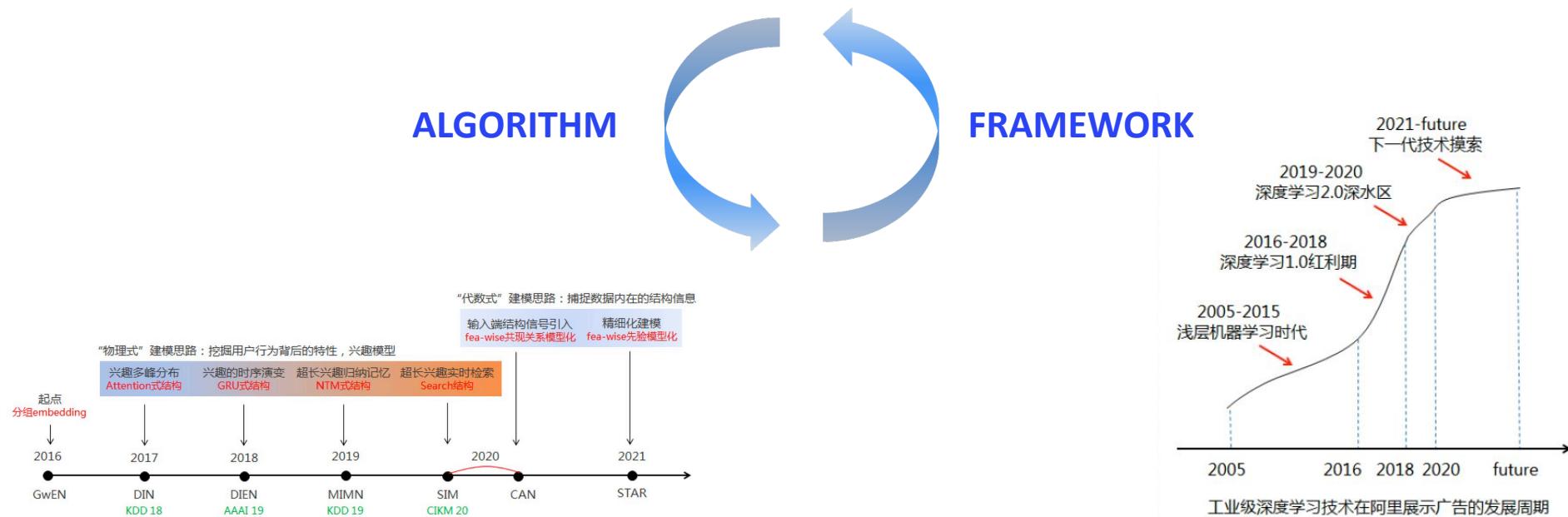
What Do We Need for Industrial Machine Learning Systems? Bernoulli, A Streaming System with Structured Designs. Luo et al. DLP-KDD 2021

《屠龙少年与龙：漫谈深度学习驱动的广告推荐技术发展周期》2021

Framework – Algorithm Co-Evolution

- “The Hardware Lottery”

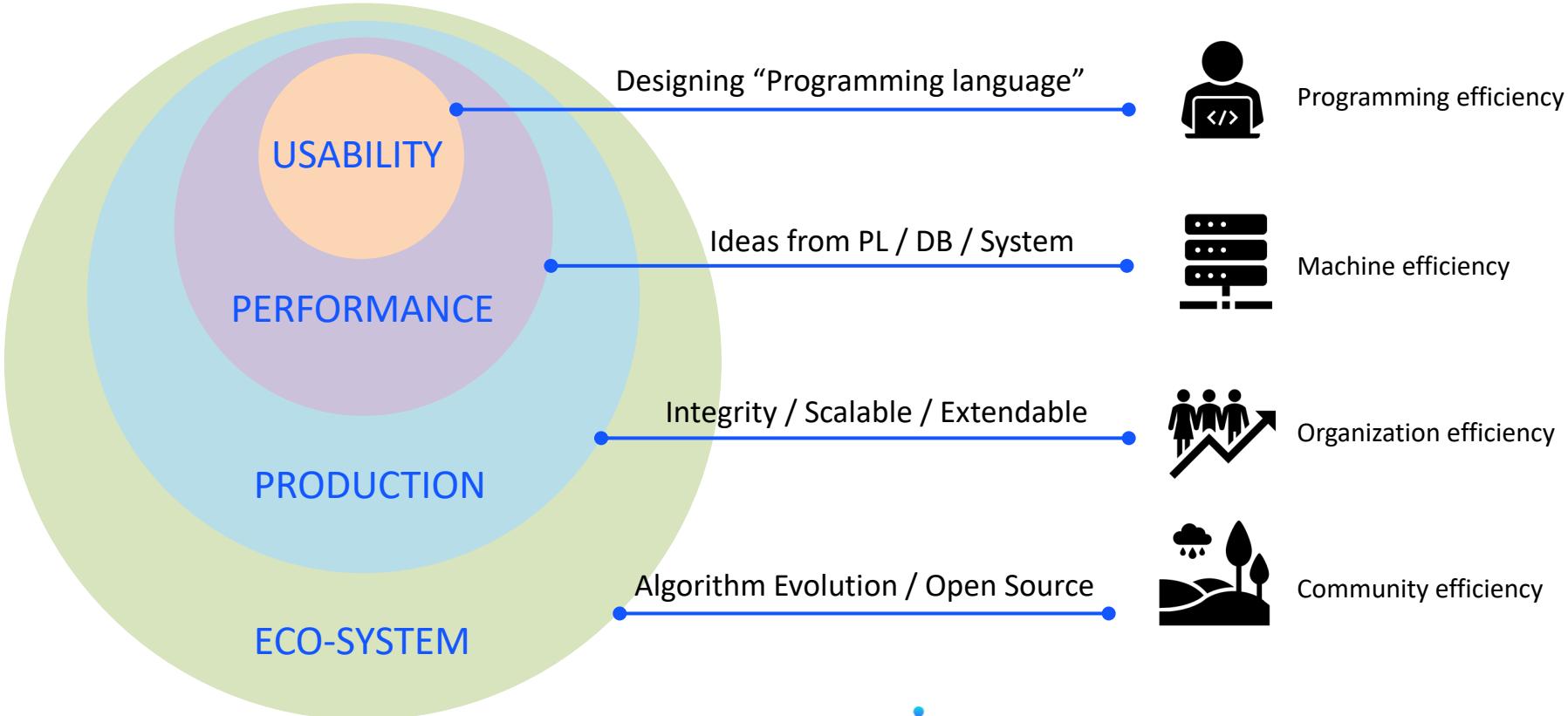
A research idea wins because it is suited to the available software and hardware



The Hardware Lottery. Sara Hooker 2020

《屠龙少年与龙：漫谈深度学习驱动的广告推荐技术发展周期》2021

Summary



|| Summary

- Programming languages are evolving and require designers to make tradeoffs, so do ML frameworks

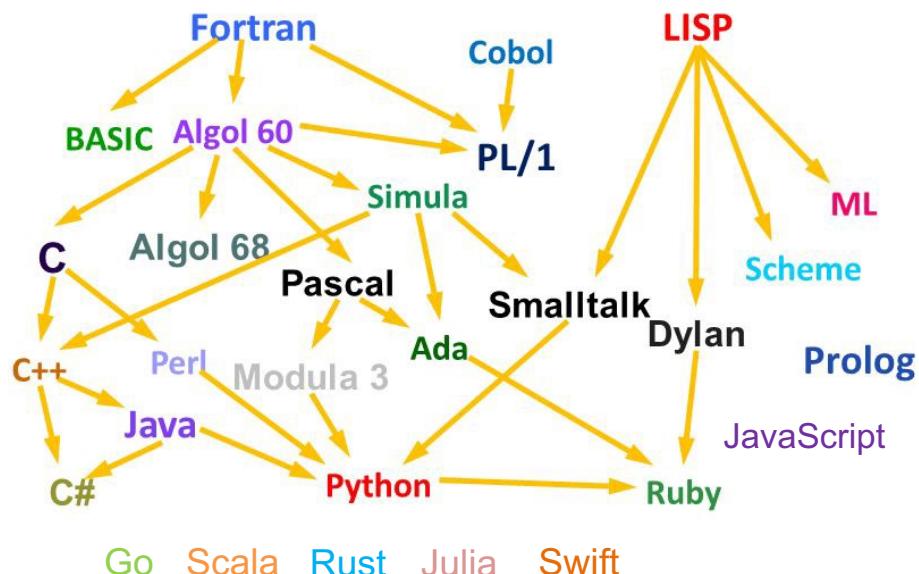


Image by Ana Harris



论坛议程



今天有6场精彩演讲，涵盖性能优化、框架整体设计、大模型训练、图神经网络等几大方向

主题	机构	演讲嘉宾
面向云计算的分布式机器学习优化实践	Google	蓝昶
超大模型的技术挑战与Mindspore的解决思路	华为	姚逸璠
超大模型高效训练的分布式框架Whale	阿里云	王林
飞桨：源于产业实践的开源深度学习平台	百度	蓝翔
MegEngine DTR 技术与训练框架技术创新	旷视科技	许欣然
构建高效易用的图深度学习平台	亚马逊	王敏捷

2021



THANKS!

Contact:
biye.jby@alibaba-inc.com

