AI HARDWARE INFRASTRUCTURE AT FACEBOOK

Xiaodong Wang
Kevin Lee
Today’s AI Hardware Infrastructure

New Hardware Announcement

What’s Next
Applied Machine Learning  AI Research  AI Infrastructure

OPEN Compute Project

Facebook Open Source
Artificial Intelligence

- Program that can sense, reason, act, and adapt like humans

Machine Learning

- Program that can perform actions without explicitly being programmed

Deep Learning

- Subset of machine learning in which multilayered neural networks learn
Services and Use Cases

- Face tagging
- Search
- Translation
- Ads
- News Feed
Why Now?

Research

Data

Compute
<table>
<thead>
<tr>
<th>Internal Platforms &amp; Toolkits</th>
<th>Frameworks &amp; Infrastructure</th>
</tr>
</thead>
<tbody>
<tr>
<td>FBLearner Feature Store</td>
<td>Caffe2</td>
</tr>
<tr>
<td>FBLearner Flow</td>
<td>PyTorch</td>
</tr>
<tr>
<td>FBLearner Predictor</td>
<td>ONNX</td>
</tr>
</tbody>
</table>
How Are We Leveraging Machine Learning at FB?

- Bryce Canyon
- Big Basin
- Tioga Pass
- Twin Lakes

Data | Features | Training | Evaluation | Inference

- FBLearner Feature Store
- FBLearner Flow
- FBLearner Predictor
2012 Experimentation
2013 HP SL270s G8
2015 Big Sur
2016 Big Basin Pascal
2018 Announcing Today
Today’s AI Hardware Infrastructure

Announcing
Big Basin Volta

What’s Next
Big Basin Volta at OCP 2018
Big Basin Volta at OCP 2018

- Paired with OCP Tioga Pass platform
- Increased GPU performance
- Improved airflow and serviceability
Computer Vision at FB
Computer Vision Performance

Multi-GPU speedup

FP-32

1-V100  2-V100  4-V100  8-V100
Computer Vision Performance

High-bandwidth FP-16 TensorCore
Machine Translation

Better Translation Quality

Phrase-based statistical approach

Onların, İzmir’ın neden hayır dediğini anlamalarını beklemiyoruz.

Their, İzmir’s why you said no we don’t expect them to understand.

Necip Fazıl Ayan

1 hr •

• Rate this translation

Neural network approach

Onların, İzmir’ın neden hayır dediğini anlamalarını beklemiyoruz.

We don’t expect them to understand why İzmir said no.

Necip Fazıl Ayan

1 hr •

• Rate this translation
Today’s AI Hardware Infrastructure

Announcing Big Basin Volta

What’s Next
Consolidated Hardware Design
Max Operating Efficiency
Disaggregated Design
Vendor Agnostic Platform
Please join us Wednesday at 11:30am for Big Basin Volta workshop!