Scientific Computing using AWS

Sanjay Padhi, Ph.D
Amazon Web Services
sanpadhi@amazon.com
Amazon Web Services with HTCondor

Version 8.7.8

Release Notes:

- HTCondor version 8.7.8 released on May 10, 2018.

New Features:

- `condor_annex` may now be setup in multiple regions simultaneously. Use the `-aws-region` flag with `-setup` to add new regions. Use the `-aws-region` flag with other `condor_annex` commands to choose which region to operate in. You may change the default region by setting `ANNEX_DEFAULT_AWS_REGION`. (Ticket #6632).
- Added default AMIs for all four US regions to simplify using `condor_annex` in those regions. (Ticket #6633).
- HTCondor will no longer mangle `CUDA_VISIBLE_DEVICES` or `GPU_DEVICE_ORDINAL` if those environment variables are set when it starts up. As a result, HTCondor will report GPU usage with the original device index (rather than starting over at 0). (Ticket #6584).
- When reporting `GPUsUsage`, HTCondor now also reports `GPUsMemoryUsage`. This is like `MemoryUsage`, except it is the peak amount of GPU memory used by the job. This feature only works for nVidia GPUs. (Ticket #6544).
- Improved error messages when delegation of an X.509 proxy fails. (Ticket #6575).
- `condor_restart` will no longer limit the width of the output to 80 columns when it outputs to a file or pipe. (Ticket #6643).
- Submission of jobs via the Python bindings `Submit` class will now attempt to put all jobs submitted in a single transaction under the same ClusterId. (Ticket #6649).
- Added support for `condor_schedd` option in the Python bindings. (Ticket #6619).
- Added DAP support. (Ticket #6648).

Bugs Fixed:

- Fixed a problem where, when starting enough `condor_annex` instances simultaneously, some (approximately 1 in 100) instances would neither join the pool nor terminate themselves. (Ticket #6638).
- When using HAD, the job scheduler for the pool was not always started. This has been fixed. (Ticket #6637).
The AWS Secret Region is designed and built to meet the regulatory and compliance requirements of DOD, IC, etc.
Evolution in Compute Services

Virtual Server Hosting, Container management, and Serverless Computing

Amazon EC2
Provides resizable cloud-based compute capacity in the form of EC2 instances, which are equivalent to virtual servers

Amazon EC2 Container Service
A highly scalable, high performance container management service

AWS Lambda
Run code without thinking about servers. Serverless compute for stateless code execution in response to triggers
Heterogeneity in Compute Resource Instance Types & CPU

General purpose
- M5
- C5
- M4
- C4

Compute optimized
- M5.24xlarge
  - 96 vCPU
  - 384GB RAM
  - Up to 25Gps n/w
- C5.18xlarge
  - 72 vCPU
  - 144GB RAM
  - EBS only

Storage and IO optimized
- I3, H1
  - I3.16xlarge
    - 64 vCPU
    - 488GB RAM
    - 8 x 2TB NVMe SDD
    - 25 Gbps
  - H1.16xlarge
    - 64 vCPU
    - 256GB RAM
    - 8 x 2TB HDD
    - 25 Gbps

Accelerated Computing
- D2
  - P3
  - R3
- D2.8xlarge
  - 36 vCPU
  - 256GB RAM
  - 8 x 2TB HDD
  - 25 Gbps

Memory optimized
- X1/e
- X1e.32xlarge
  - 128 vCPU
  - 4TB RAM
  - 2 x 1.9TB SSD
  - 14k EBS Mbps
- R4
- R4.16xlarge
  - 64 vCPU
  - 488GB RAM
  - SSD EBS
  - 25 Gbps

Selection of different Intel Xeon processors
- 2.3/2.4 GHz Intel Broadwell/Haswell CPUs: M4, I3, H1, D2, G3, P3/2 instance types
- 2.9 GHz Intel Haswell CPUs: C4
- 2.5 GHz Intel Platinum CPUs: w/AVX-512 instruction set: M5
- 3.0 GHz Intel Platinum CPUs: w/AVX-512 instruction &set Turbo up to 3.5Ghz: C5
CPUs vs GPUs vs FPGA for Compute

**CPU**
- 10s-100s of processing cores
- Pre-defined instruction set & datapath widths
- Optimized for general-purpose computing

**GPU**
- 1,000s of processing cores
- Pre-defined instruction set and data path widths
- Highly effective at parallel execution

**FPGA**
- Millions of programmable digital logic cells
- No predefined instruction set or datapath widths
- Hardware timed execution
GPU Performance Comparison

ResNet-50 Training Performance
(Using Synthetic Data, TensorFlow 1.5)

- P2 Instances use K80 Accelerator (Kepler Architecture)
- P3 Instances use V100 Accelerator (Volta Architecture)
# Security Requirements

<table>
<thead>
<tr>
<th>Standard</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISO 9001</td>
<td>Global Quality Standard</td>
</tr>
<tr>
<td>ISO 27001</td>
<td>Security Management Standard</td>
</tr>
<tr>
<td>ISO 27017</td>
<td>Cloud Specific Controls</td>
</tr>
<tr>
<td>ISO 27018</td>
<td>Personal Data Protection</td>
</tr>
<tr>
<td>PCI DSS Level 1</td>
<td>Payment Card Standards</td>
</tr>
</tbody>
</table>

**Participating Organizations**

- SOC 1 Audit Controls Report
- SOC 2 Compliance Controls Report
- SOC 3 General Controls Report

**Regulatory Compliance**

- CJIS
- DoD SRG
- FDA
- FedRAMP
- DoD Data Processing
- Food and Drug Administration
- FedRAMP TIC
- FedRAMP-Trusteed Internet Connection
- FERPA
- Educational Privacy Act
- FIPS
- Government Security Standards
- FISMA
- Federal Information Security Management Act
- HIPAA
- Protected Health Information
- SEC Rule 17a-4(f)
- Financial Data Standards
- ITAR
- International Arms Regulations

**Other Standards**

- FISC [Japan]
- IRAP [Australia]
- MLPS Level 3 [China]
- MTCS Tier 3 [Singapore]
- My Number Act [Japan]
- DNB [Netherlands]
- EU Data Protection Framework
- G-Cloud [UK]
- IT-Grundschutz [Germany]
- Privacy Shield
- UK Cyber Essentials
- Plus
- Cyber Threat Protection

[https://aws.amazon.com/compliance/](https://aws.amazon.com/compliance/)
Predictive Analytics in Scientific Computing
The Large Hadron Collider @ CERN includes 6,000+ researchers from over 40 countries and produces approximately 25PB of data each year.

The ATLAS and CMS experiments are using AWS for Monte Carlo simulations, processing, and analysis of LHC data.
Analytics at the LHC

Production rate

# / sec

QCD

Level 1

Hardwired processors (ASIC, FPGA)

Level 2

Standard CPUs farms & Networks

Level 3

HLT

Recorded Events

25 ns few µsec ~ms > Sec

W,Z

Top

Z’

Higgs
Elasticity in Computing: On demand auto-expansion to AWS – HTCondor based

~60,000 slots using AWS spot instances. A factor of 5 larger than Fermilab capacity!
Peter Shanahan (Co-spokesperson of the NOvA experiment): “Our experience with Amazon Web Services shows its potential as a reliable way to meet our peak data processing needs at times of high demand”


Neutrinos are ghost like particles → Needed advanced ML analytics to detect
Elasticity – Machine Learning & Natural Language Processing at Clemson University, 1.1 Million vCPUs with EC2 Spot Instances

HTCondor 8.7.2 on Amazon Linux 2017.03 (GPU-ready)

Sold by: Center for High Throughput Computing  Latest Version: v8.7.2

The HTCondor high-throughput computing system, version 8.7.2, installed on Amazon Linux 2017.03.

Linux/Unix  ★★★★★  (0)  Free Tier

Overview  Pricing  Usage  Support  Reviews

Product Overview

The HTCondor high-throughput computing system is a workload management system for compute-intensive jobs.

Highlights

- Complete single-node HTCondor pool, ready to run
- Use the condor_annex command to easily add more instances to this HTCondor pool
• DNNs have the potential to greatly improve physics performance in the trigger system

• In order to implement an algorithm, need to ensure inference latencies of μs (ms) for L1 (HLT)
  - For L1, this means we must use FPGAs

• How can we run neural network inference quickly on an FPGA?
FPGA Acceleration

FPGA handles compute-intensive, deeply pipelined, hardware-accelerated operations

```verilog
module filter1 (clock, rst, strm_in, strm_out)

for (i=0; i<NUMUNITS; i=i+1)
    always@(posedge clock)
        integer i,j; //index for loops
        tmp_kernel[j] = k[i*OFFSETX];

CPU handles the rest
```
Acceleration with AWS

- Development for FPGA kernel and CPU host code is done with SDAccel environment
  - Invokes Vivado HLS under the hood, produces traditional synthesis reports etc.
- Run host code on CPU, manages data transfer and FPGA kernel execution
- hls4ml project only needs to be wrapped to provide specific inputs/outputs for SDAccel to interface properly
  - Can be done generically
  - Have accelerated variety of hls4ml projects on AWS F1
- Limited in speed by I/O bandwidth
An Acceleration Case Study (2)

- Have successfully implemented/run the network inference on AWS using his4ml/SDAccel
- [https://github.com/drankincms/AcelFPGA](https://github.com/drankincms/AcelFPGA)

- Including data transfer to/from CPU, whole **FPGA inference process takes 2 ms** for all 16k HCAL channels
  - Inference alone takes 80 us (70 ns for one inference)

- Has been tested inside standard CMS software code environment, using high-level trigger job
  - Every event sends input features to FPGA, waits for callback
- **Iterative fit procedure takes 50 ms** for same inputs

- FPGA inference is a fixed-latency procedure, iterative fit is not
- Inference on **CPU or GPU** also significantly faster than iterative fit
  - FPGA inference fastest

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Architecture</th>
<th>Time/event (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iterative fit</td>
<td>CPU</td>
<td>50</td>
</tr>
<tr>
<td>NN Inference</td>
<td>CPU</td>
<td>15</td>
</tr>
<tr>
<td>NN Inference</td>
<td>GPU</td>
<td>12</td>
</tr>
<tr>
<td>NN Inference</td>
<td>FPGA</td>
<td>2</td>
</tr>
</tbody>
</table>
A key component in autonomous vehicle and fast triggering systems, learn how FPGAs do real-time DNN inference in this hands-on course. Topics include:

- Model compression and quantization
- High-level synthesis
- Firmware implementation
- Model acceleration on cloud FPGAs

The class is given by Dr. Jennifer Ngadiuba (CERN) and Dr. Dylan Rankin (MIT) and consists of half a day of lectures as well as a hands-on sessions.

You'll learn how to compress and synthesise your own TensorFlow model, as well as implement it on a Xilinx FPGA on the Amazon cloud.

The course is targeted at PhD, Postdocs and Professors, but others will be allowed to participate if there are available places.

The lectures and hands-on session will take place at the UZH Irchel Campus in the Physik Institut (building 36).

https://indico.cern.ch/event/769727/

How to do ultrafast Deep Neural Network inference on FPGAs

5. February 2019
Physik Institut - Universität Zürich

Organizers:
Thea Aarrestad (UZH)
Jennifer Ngadiuba (CERN)
Dylan Rankin (MIT)
Maurizio Pierini (CERN)
Ben Kilminster (UZH)

All course material can be found as attachments to the timetable, or at

https://github.com/FPGA4HEP/course_material
Methods to use FPGA (AWS F1 Instance)

1. **Hardware Engineers/Developers**
   - Developers who are comfortable programming FPGA
   - Use F1 Hardware Development Kit (HDK) to develop and deploying custom FPGA accelerations using Verilog and VHDL

2. **Software Engineers/Developers**
   - Developers who are not proficient in FPGA design
   - Use OpenCL to create custom accelerations

3. **Software Engineers/Developers**
   - Developers who are not proficient in FPGA design
   - Use pre-build and ready to use accelerations available in AWS Marketplace
Children's Hospital of Philadelphia and Edico Genome Achieve Fastest-Ever Analysis Of 1,000 Genomes

GUINNESS WORLD RECORDS title for Fastest time to analyze 1,000 human genomes

The Amazon EC2 F1 instances, with Xilinx Virtex UltraScale+ field programmable gate arrays (FPGAs) was used for 1,000 diverse pediatric genomes.

The study was completed in two hours and twenty-five minutes.

Hubble Space Imagery on AWS: 28 Years of Data Now Available in the Cloud

Research Collaborations
Research Collaboration: AMPLab

AMP stands for “Algorithms, Machines, and People”

Research Outcomes

http://mesos.apache.org
http://www.mlibase.org
http://spark.apache.org
http://spark.apache.org
https://databricks.com
http://snap.cs.berkeley.edu
Research Collaboration: RISELab (Real-time Intelligent Secure Execution)

Collaborative 5-year effort between UC Berkeley, National Science Foundation and Industry

AWS as a founding partner - https://rise.cs.berkeley.edu/

Goal: Develop open source platforms, tools and algorithms for intelligent real-time decisions on live-data

- Researchers at RISELab use AWS to rapidly prototype and develop systems at scale
- Resulted in Apache Spark, developed on AWS and integrated with core services

From live data to real-time decisions
**Research Data Management**

**AWS part of Nationwide Open Storage Network (Funded by NSF)**

Collaboration includes:

- JHU, UCSD, MGHPCC, AWS, Globus, Internet2, 4 NSF Big Data Hub, etc
Collaborative programs with the National Science Foundation (NSF)

- 2017/2018: The NSF Big Data program supported by multiple directorates at NSF
- 2019: Collaboration with NSF CISE/OAC on Campus Cyberinfrastructure
- 2019: NSF CISE - Exploring Cloud for Acceleration of Science (E-CAS)
- 2019: NSF Cloud Access Model - All
Examples of Research Collaboration with NSF (2018)

- **Automating Analysis and Feedback to Improve Mathematics Teachers' Classroom**
  
  *University of Colorado, Boulder*

- **Collaborative Research: Protecting Yourself from Wildfire Smoke: Big Data Driven Adaptive Air Quality Prediction Methodologies**
  
  *University of Nevada, Reno*

- **Collaborative Research: TIMES: A tensor factorization platform for spatio-temporal data**
  
  *Emory University*

- **Collaborative Research: Optimizing Log-Structured-Merge-Based Big Data Management Systems**
  
  *University of California, Riverside*

- **Collaborative Research: Intelligent Solutions for Navigating Big Data from the Arctic and Antarctic**
  
  *Texas A&M University Corpus Christi*

“This NSF big data award, coupled with AWS’s advanced computational and analytic services, is expected to help unlock the secrets of interactions among biomolecules that drive human and animal biological processes.”

*Dr. Bin Yu, Chancellor’s Professor at University of California,*
Examples of Research Collaboration with NSF (2017)

- **Detecting Financial Market Manipulation: An Integrated Data- and Model-Driven Approach**
  University of Michigan, Georgia Tech

- **Scalable and Interpretable machine learning: bridging mechanistic and data-driven modeling in the biological sciences**
  University of California, Berkeley

- **Taming Big Networks via Embedding**
  University of Virginia, University of Illinois at Urbana-Champaign

- **Domain Adaptation Approaches for Classifying Crisis Related Data on Social Media**
  Kansas State University, University of North Texas and Pennsylvania State University

- **Distributed Semi-Supervised Training of Deep Models and Its Applications in Video Understanding**
  University of Central Florida

"In today's era of data-driven science and engineering, we are pleased to work with the AWS Research Initiative via the NSF BIGDATA program, to provide cloud resources for our Nation's researchers to foster and accelerate discovery and innovation."

Dr. Jim Kurose, Assistant Director, CISE, National Science Foundation (NSF)
Natural Language Processing is part of Artificial Intelligence (AI):

- Intersects with Computers and human (natural) languages
- Involves Speech recognition, natural language understanding & natural language generation
Domain Adaptation Approaches for Classifying Crisis Related Data on Social Media

Billion-dollar disasters of 2017 in the US

- California wildfires, March
  - $1.5B
- Northern California fires, October
  - $9.4B
- Western wildfires, Summer-fall
  - $2.5B
- ND, SD, and MT drought, Spring-fall
- Upper Midwest severe weather, June
  - $2.5B
- Midwest severe weather, June
  - $2.5B
- Midwest tornado outbreak, March
  - $2.1B
- Missouri and Arkansas flooding, April
  - $2.2B
- Southern tornado outbreak, January
  - $1.7B
- Southeast severe weather, March
  - $2.1B
- Southeast severe weather, March
  - $2.1B
- Hurricane Harvey, August
  - $198B
- Hurricane Maria, September
  - $65B
- Hurricane Irma, September
  - $102B

Sources: NOAA, Ball State University Center for Business and Economic Research (for Harvey), Reuters (for Maria), and CoreLogic (for Irma)
Domain Adaptation Approaches for Classifying Crisis Related Data on Social Media

Methodology

Data Collection
- JSON tweets

Data Extraction
- Tweet id, create time, text

Data Processing
- Stop words, special characters, URLs, Emails

Topics Modeling
- Streaming Corpus
- Latent Dirichlet Allocation

Analysis
- Preparedness, During Hurricane, Aftermath
- Hurricane timeline
Classes of machine learning algorithms

Supervised learning [Imran et al., 2013; Ashktorab et al., 2014; Caragea et al., 2014; Imran et al., 2018]
- Labeled tweets needed, but not readily available for an emergent disaster

Domain adaptation [Li et al., 2015; Li et al., 2017, Alam et al., 2018, Mazloom et al., 2018]
- Knowledge from a prior source disaster is transferred to a target disaster

Unsupervised learning, e.g., topic modeling [Resch et al., 2017]
- Topic modeling can help associate topics/categories with tweets
Using Amazon Comprehend results to get aggregate statistics
Using Comprehend results to determine frequent entities
Using locations identified by Comprehend to track hurricane path

Hurricane Irma predicted path

Hurricane Irma real path

Source: Weather Channel
Domain Adaptation Approaches for Classifying Crisis Related Data on Social Media

Doina Caragea, KSU

1. Damage
2. Damage
3. Damage
4. Damage

1. DAV = 0.413
2. DAV = 0.453
3. DAV = 0.423
4. DAV = 0.385

We thank the National Science Foundation and Amazon Web Services for support from the grant IIS-1741345.

IEEE/ACM ASONAM 2018, August 28-31, 2018, Barcelona, Spain
978-1-5386-6051-5/18/$31.00 © 2018 IEEE
BigData Market Manipulation Project
Michael Wellman, University of Michigan

• Collaboration between U.Michigan and Georgia Tech
• Sponsored by NSF BIGDATA program, computational support from AWS
• Interdisciplinary: Computer Science (AI/ML), Finance, Law & Public Policy

Goal: New techniques for detecting and mitigating manipulation
As “Spoof” Trading Persists, Regulators Clamp Down
Bluffing Tactic That Dodd-Frank Banned in 2010 Can Distort Markets

CHICAGO—One June morning in 2012, a college dropout whom securities traders call “TheRussian” logged on to his computer and began trading Brent-crude futures on a London exchange from his skyscraper office here.

Over six hours, (pseudonym)

Traders call the illegal bluffing tactic “spoofing” and say it has long been used to manipulate prices of anything from stocks to bonds to futures.

Prosecutors said Michael Coscia wanted to lure other traders to markets by creating an illusion of demand so that he could make money on smaller trades.

A US jury has found high-frequency trader Michael Coscia guilty of commodities fraud and “spoofing” in the US government’s first criminal
Spoofing

Definition: Practice of submitting large spurious orders to buy or sell some security to mislead other traders about market state.

Source: UK Financial Conduct Authority
Amazon SageMaker

**Build**
- Pre-built notebook instances
- Highly-optimized machine learning algorithms

**Deploy**
- Fully-managed hosting at scale
- Deployment without engineering effort

**Train**
- One-click training for ML, DL, and custom algorithms
- Easier training with hyperparameter optimization

© 2017, Amazon Web Services, Inc. or its Affiliates. All rights reserved.
Collaboration with NSF and Internet2: E-CAS

Internet2 and @NSF_CISE Announce Selection of First-Phase Research Proposals for Exploring Clouds for Acceleration of Science (E-CAS) Project https://bit.ly/2U56d1R #NSFfunded #research #cloud

tweet this

The successful proposals for the year-long first phase of the E-CAS project are:

Development of BioCompute Objects for Integration into Galaxy in a Cloud Computing Environment

Raja Mazumder, George Washington University

BioCompute objects allow researchers to describe bioinformatic analyses comprised of any number of algorithmic steps and variables to make computational experimental results clearly understandable and easier to repeat. This project will create a library of BioCompute objects that describe bioinformatic workflows on AWS, which can be accessed and contributed to by users of the widely used bioinformatics platform, Galaxy.

IceCube computing in the cloud

Benedikt Riedel, University of Wisconsin

The IceCube Neutrino Observatory located at the South Pole supports science from a number of disciplines including astrophysics, particle physics, and geographical sciences operating continuously being simultaneously sensitive to the whole sky. This project aims to burst into cloud to support follow-up computations of observed events, as well as alerts to and from the research community, such as other telescopes and LIGO.
Baylor College of Medicine CHARGE

Baylor College of Medicine Human Genome Sequencing Center and DNAnexus using the Mercury Pipeline for the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium

Supports 300+ researchers around the world

Analyzed the genomes of over 14,000 individuals, encompassing 3,751 whole genomes and 10,940 whole exomes (~1PB of data)

Used 3.3 million core hours over 4 weeks to complete the job 5.7x faster than what could have been accomplished on-premise

The outcomes

- Easier collaboration
- Faster time to science
- Cost-effective: On-premise was prohibitively expensive
- No longer constrained by on-premise capacity
- Scientists focusing on Science as opposed to infrastructure

Access to shared data repositories

https://aws.amazon.com/solutions/case-studies/baylor/
AWS signs on with NIH cloud precision medicine project

Researchers associated with the National Institutes of Health (NIH) will have more resources to access and analyze data thanks to a new partnership with Amazon Web Services (AWS).

The newly-formed partnership with NIH’s Science and Technology Research Infrastructure for Discovery, Experimentation, and Sustainability (STRIDES) Initiative will give NIH-associated biomedical researchers access to AWS technologies. AWS is Amazon’s cloud-based computing platform subsidiary.


Stanford: Automatic Grading of Diabetic Retinopathy through Deep Learning using AWS

Early Detection of Diabetic Complications

Skin Cancer Detection
At Physician-Levels (or better)
FDA-Approved Medical Imaging
De-identify medical images with the help of Amazon Comprehend Medical and Amazon Rekognition

Language Support:

- English (Australia)
- English (Canada)
- English (India)
- English (UK)
- English (US)
- German
- Japanese
Speaker Adaptation


5-7% relative reduction in word error rate compared to speaker independent model
Transfer Learning from English to German
Ohio Health Automates Patient Interactions

“Amazon Lex represents a great opportunity for us to deliver a better experience to our patients. Everything we do at OhioHealth is ultimately about providing the right care to our patients at the right time and in the right place. Amazon Lex’s next generation technology and the innovative applications we are developing using it will help provide an improved customer experience. We are just scratching the surface of what is possible,”

Michael Krouse
Senior Vice President and CIO – Ohio Health

• Delivers Personalized Care Recommendations
• Makes Customer Appointments
• Drives Urgent Care Referrals
Incorporates ~47 different voices and fully managed services

https://aws.amazon.com/polly/
Deep Learning

Significantly improve many applications on multiple domains

image understanding  speech recognition  natural language processing  autonomy

“deep learning” trend in the past 10 years
Autonomous Driving Systems
Centimeter-accurate positioning
AWS DeepRacer

A fully autonomous 1/18th-scale race car designed to help you learn about reinforcement learning through autonomous driving

- Build machine learning models in Amazon SageMaker
- Train, test, and iterate on the track using the AWS DeepRacer 3D racing simulator
- Compete in the world’s first global autonomous racing league, to race for prizes and a chance to advance to win the coveted AWS DeepRacer Cup
What is Amazon Managed Blockchain?

Amazon Managed Blockchain is a fully managed service that makes it easy to create and manage scalable blockchain networks using popular open source frameworks:

**Hyperledger Fabric and Ethereum**

1. **Create a network**: Choose an open source blockchain framework, set up a new blockchain network, and your membership in your AWS account with just a few clicks.
2. **Invite members**: Invite other AWS accounts to join the network.
3. **Add nodes**: Create and configure blockchain peer nodes that store a copy of the distributed ledger.
4. **Deploy applications**: Create and deploy decentralized applications to your network through your per nodes. Transact with other members on the network.
Applications in various domains

Proof of Ownership
Documents/Contracts
Digital Security Trading
Enterprise Platforms
Mortgage Loans
Voting Mechanisms
Patient Records
Corporate Governance
Financial
Insurance

Capital Markets
HCLS
Real Estate
Legal
Agriculture
Gaming
Transportation
M & E
Digital Advertising
Power/Utilities
Retail
Cloud
Singapore Exchange: Project Ubin’s blockchain use case

Challenges with existing financial systems:
- Lack of trust
- Inefficient processes for sending data across borders
- API divergence is expensive and cumbersome to maintain

Benefits of implementing a blockchain:
- Distributed application provides trust
- Provides reliability and resiliency
- Easy to add new participating members
- Efficient transfer of data and transactions without intermediaries
AWS RoboMaker - Robotics

- Cloud Extensions for ROS
- Development Environment
- Simulation
- Fleet Management
AWS Cloud Credits for Research

1. Build cloud-hosted publicly available science-as-a-service applications, software, or tools to facilitate their future research and the research of their community.
2. Perform proof of concept or benchmark tests evaluating the efficacy of moving research workloads or open data sets to the cloud.
3. Train a broader community on the usage of cloud for research workloads via workshops or tutorials.

AWS Machine Learning Research Awards

Funding academic research at the forefront of machine learning.

https://aws.amazon.com/research-credits/
https://aws.amazon.com/aws-ml-research-awards/
https://ara.amazon-ml.com/proposals/#apply
https://aws.amazon.com/opendata/
# AWS Research Workshops

This repo provides a managed SageMaker jupyter notebook with a number of notebooks for hands on workshops in data lakes, AI/ML, Batch, IoT, and Genomics.

## Workshops

Please review and complete all prerequisites before attempting these workshops.

<table>
<thead>
<tr>
<th>Title</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Introduction to AWS Basics</strong></td>
<td>Learn about core AWS services for compute, storage, database and networking. This workshop has a hands-on lab where you will be able to launch an auto-scaled Apache web server behind an ALB, S3 bucket hosting content of the home page, and how to define the appropriate roles for each resource.</td>
</tr>
<tr>
<td><strong>Building a Data Lakes</strong></td>
<td>In this series of hands-on workshops, you will learn how to understand what data you have, how to drive insights, and how to make predictions using purpose-built AWS services. Learn about the common pitfalls of building data lakes, and discover how to successfully drive analytics and insights from your data. Also learn how services such as Amazon S3, AWS Glue, Amazon Athena, and Amazon AI/ML services work together to build a serverless data lake for various roles, including data scientists and business users.</td>
</tr>
<tr>
<td><strong>Build Serverless Applications in Python with AWS SAM CLI Coming Soon</strong></td>
<td>AWS Serverless Applications in Python: With AWS Serverless computing you can run applications and services without having to provision, scale, and manage any servers. In this workshop, we will introduce the basics of building serverless applications and microservices using services like AWS Lambda, Amazon API Gateway, Amazon DynamoDB, and Amazon S3. You’ll learn to build and deploy your own serverless application using these services for common use cases like web applications, analytics, and more.</td>
</tr>
<tr>
<td><strong>Tensorflow with Amazon SageMaker</strong></td>
<td>Amazon SageMaker is a fully-managed platform that enables developers and data scientists to quickly and easily build, train, and deploy machine learning models at any scale. Amazon SageMaker removes all the barriers that typically slow down developers who want to use machine learning. We will show you how to train and build a ML model on SageMaker then how to deploy the inference end points on tools like AWS Greengrass or Serverless applications.</td>
</tr>
</tbody>
</table>

[https://github.com/aws-samples/aws-research-workshops](https://github.com/aws-samples/aws-research-workshops)
Amazon SageMaker Examples

This repository contains example notebooks that show how to apply machine learning and deep learning in Amazon SageMaker

Examples

Introduction to Ground Truth Labeling Jobs

These examples provide quick walkthroughs to get you up and running with the labeling job workflow for Amazon SageMaker Ground Truth.

- From Unlabeled Data to a Deployed Machine Learning Model: A SageMaker Ground Truth Demonstration for Image Classification is an end-to-end example that starts with an unlabeled dataset, labels it using the Ground Truth API, analyzes the results, trains an image classification neural net using the annotated dataset, and finally uses the trained model to perform batch and online inference.
- Ground Truth Object Detection Tutorial is a similar end-to-end example but for an object detection task.
- Basic Data Analysis of an Image Classification Output Manifest presents charts to visualize the number of annotations for each class, differentiating between human annotations and automatic labels (if your job used auto-labeling). It also displays sample images in each class, and creates a pdf which concisely displays the full results.
- Training a Machine Learning Model Using an Output Manifest introduces the concept of an "augmented manifest" and demonstrates that the output file of a labeling job can be immediately used as the input file to train a SageMaker machine learning model.
- Annotation Consolidation demonstrates Amazon SageMaker Ground Truth annotation consolidation techniques for image classification for a completed labeling job.

Introduction to Applying Machine Learning

These examples provide a gentle introduction to machine learning concepts as they are applied in practical use cases across a variety of sectors.

https://github.com/awslabs/amazon-sagemaker-examples
Thank you!

Sanjay Padhi
sanpadhi@amazon.com