How to Install and Set Up the Anaconda arm64 Version on AWS Graviton2 Processors

What's Inside

4..... AWS: From Custom Chips to Graviton Processors

5..... Anaconda on AWS Graviton2 Processors

6..... How to Access Anaconda with the AWS Marketplace Image

10..... Using Anaconda on AWS Graviton2 with the Miniconda installer

14..... Links to More Information

Abstract

This white paper outlines the benefits of using AWS Graviton2-based instances for more cost-effective workloads and demonstrates how to use the arm64 version of Anaconda to run data science workloads on AWS Graviton2 processors.

Anaconda's availability for arm64 processors is a win/win for performance as well as savings: the Anaconda/Graviton2 combination enables data scientists to run their workloads on AWS Graviton2-based instances and offers significant price/performance improvement.

Read on for more information and a step-by-step setup guide.

AWS: From Custom Chips to Graviton Processors

For years, AWS has been designing custom chips that enable faster innovation, deliver increased security, improve performance by offloading virtual functions, and reduce cost.

These custom chips have led to innovations in customer workload security (Nitro Security Chip and Nitro Enclaves), throughput and latency enhancements for networking and storage I/O (Nitro Card with high IOPS EBS storage and up to 100Gbps networking) and virtualization technology (Nitro Hypervisor). The purpose-built, modular Nitro System protects hardware resources, improves monitoring and security posture, and benefits from better memory and CPU allocation – delivering bare metal-like performance.

AWS Graviton2 processors continue the tradition of silicon innovation and are custom-built by Amazon Web Services using 64-bit ARM Neoverse cores.

Read more about the benefits of AWS Graviton2 and the Amazon EC2 instances powered by AWS Graviton processors at AWS Graviton.

Optimized across AWS, Graviton processors deliver the best price-performance for cloud workloads running in Amazon EC2.

AWS Graviton2 processors provide even more choice by powering multiple instance types:*

* As of December 31, 2021.

- **M6g** – powers general purpose workloads such as application servers, microservices, gaming servers, mid-size data stores, and caching fleets.
- **C6g** – optimized for compute-bound applications that benefit from high performance such as high-performance computing (HPC), batch processing, media encoding, and CPU-based machine learning (ML).
- **R6g** – offers a higher memory footprint for applications that process large data sets in memory including databases, in-memory caches, and real-time big data analytics.
- **X2gd** – built for memory-intensive workloads such as open-source databases (MySQL, MariaDB, and PostgreSQL) and in-memory caches (Redis, KeyDB, Memcached). X2gd offers the lowest cost per GiB of memory in Amazon EC2.
- **Im4gn** and **Is4gen** are the AWS Graviton2 storage optimized instances, built for SQL databases (MySQL, MariaDB, PostgreSQL), NoSQL databases (Cassandra, ScyllaDB, MongoDB), search engines, analytics, streaming, and large distributed file systems.
- **G5g** offers the best price performance for Android game streaming. It is built for graphics applications including Android game streaming and ML inference.
- **T4g** is ideal for low-cost, burstable general-purpose workloads.

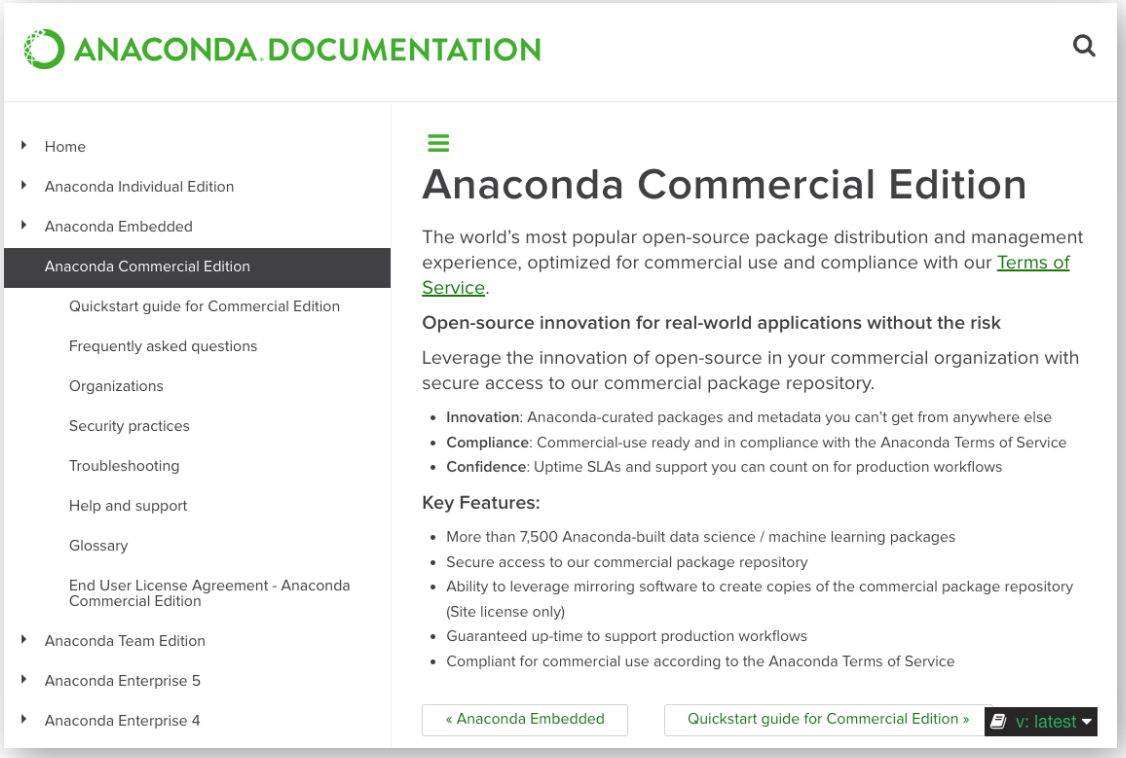
Anaconda on AWS Graviton2 Processors

Anaconda is now available for Linux on the aarch64 (arm64) platform optimized for AWS Graviton2 processors.

This innovation means that data scientists who run workloads with Anaconda packages on x86-based machines can leverage the power and savings of Graviton2 while still using favorite tools and frameworks such as Conda, SciKit-Learn, and XGBoost. Running these frameworks on AWS enables end-to-end data science in the cloud, including development, training, testing, and production.

As with the current Linux-64 (x86-based) conda packages, Linux-aarch64 packages are supported in Anaconda's "defaults" channel. Anaconda regularly updates these packages as the open-source community publishes new releases.

Download Miniconda, Anaconda Individual Edition, or Anaconda Commercial Edition[1] to start using linux-aarch64 packages on AWS Graviton2.



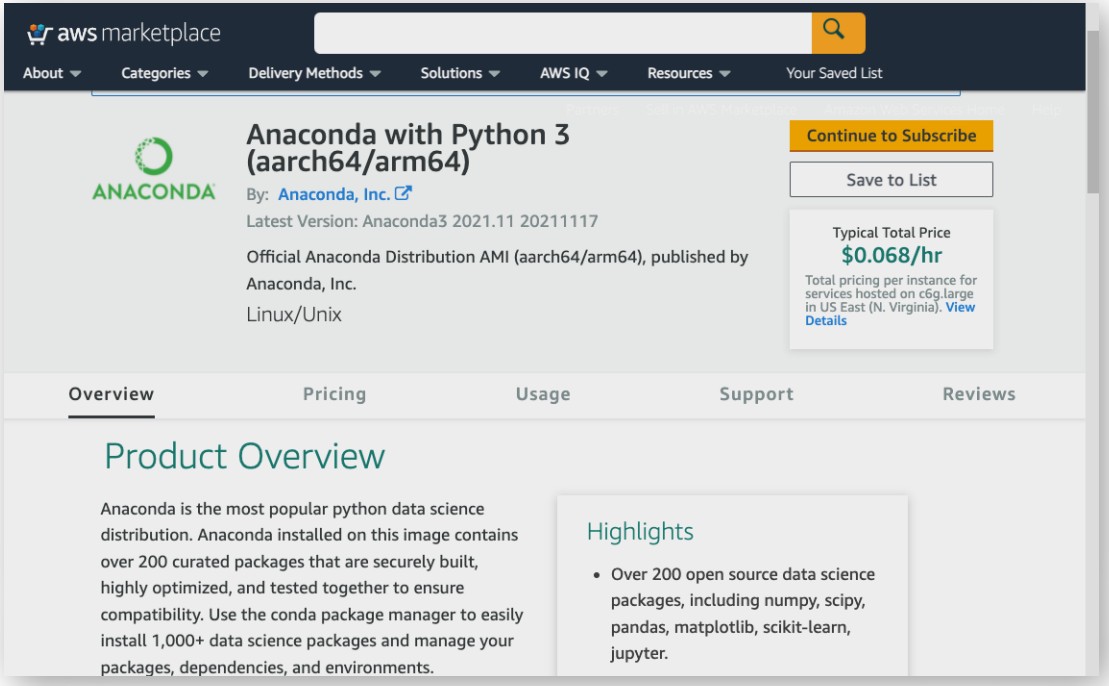
How to Access Anaconda with the AWS Marketplace Image

Let's take a look at how to access Anaconda with the Python 3 (aarch64/arm64) AWS Marketplace Image.

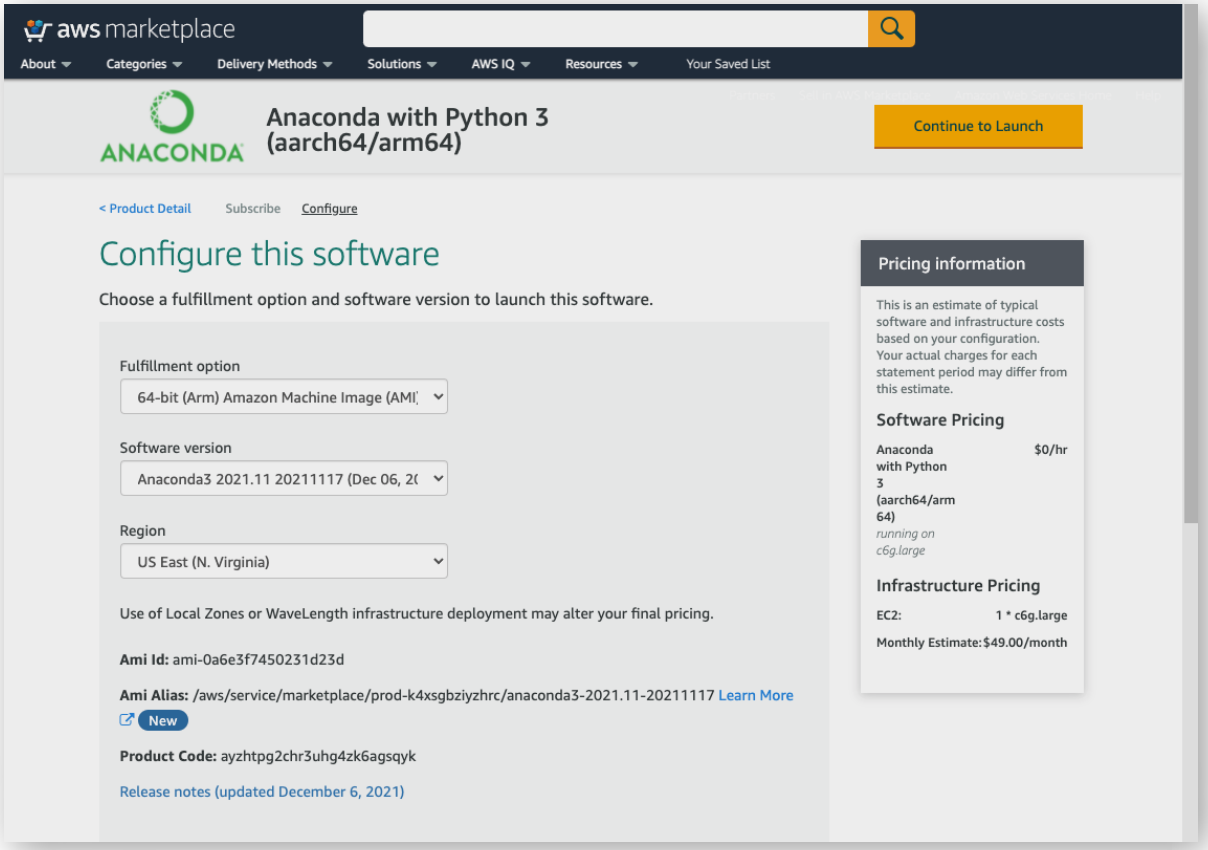
To make it easy for data scientists to use Anaconda, Anaconda is distributed via official images published on the AWS MarketPlace.

First, go to the Anaconda arm64 AMI page on the AWS MarketPlace.

Click Subscribe and then accept the terms and conditions.

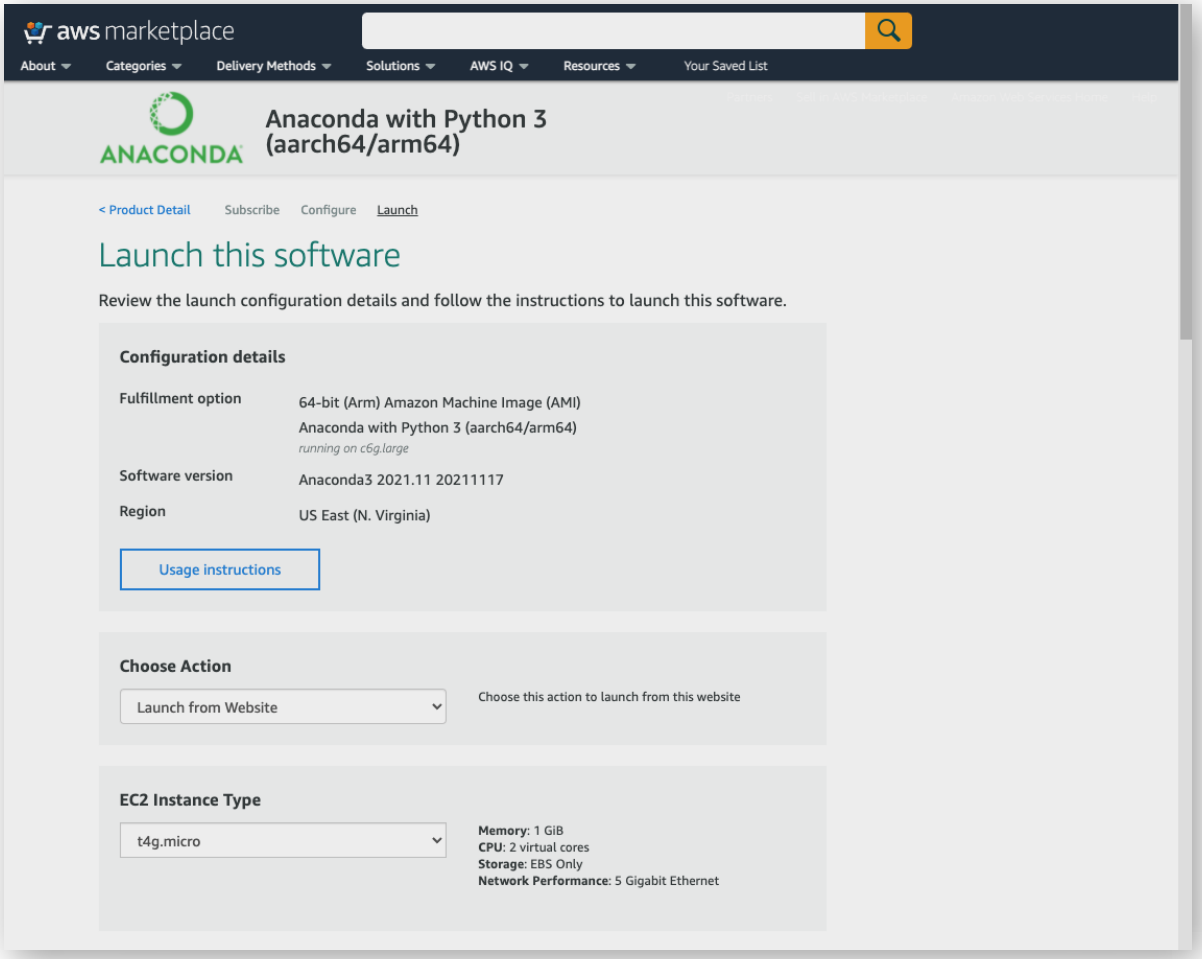


Second, configure the version and AWS region.



Third, you will be guided through a wizard to launch the Anaconda Amazon Machine Image on the AWS Graviton2 based instance of your choice.

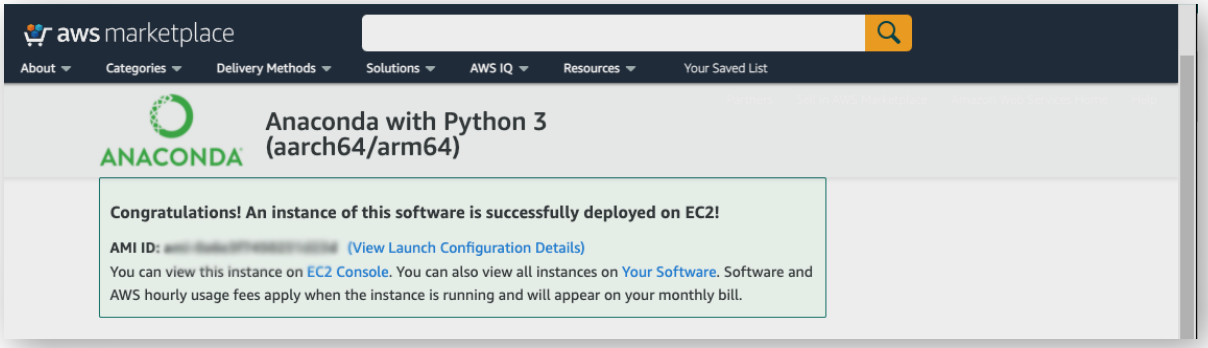
To get started, an Amazon T4g instance is a good choice for experimenting.



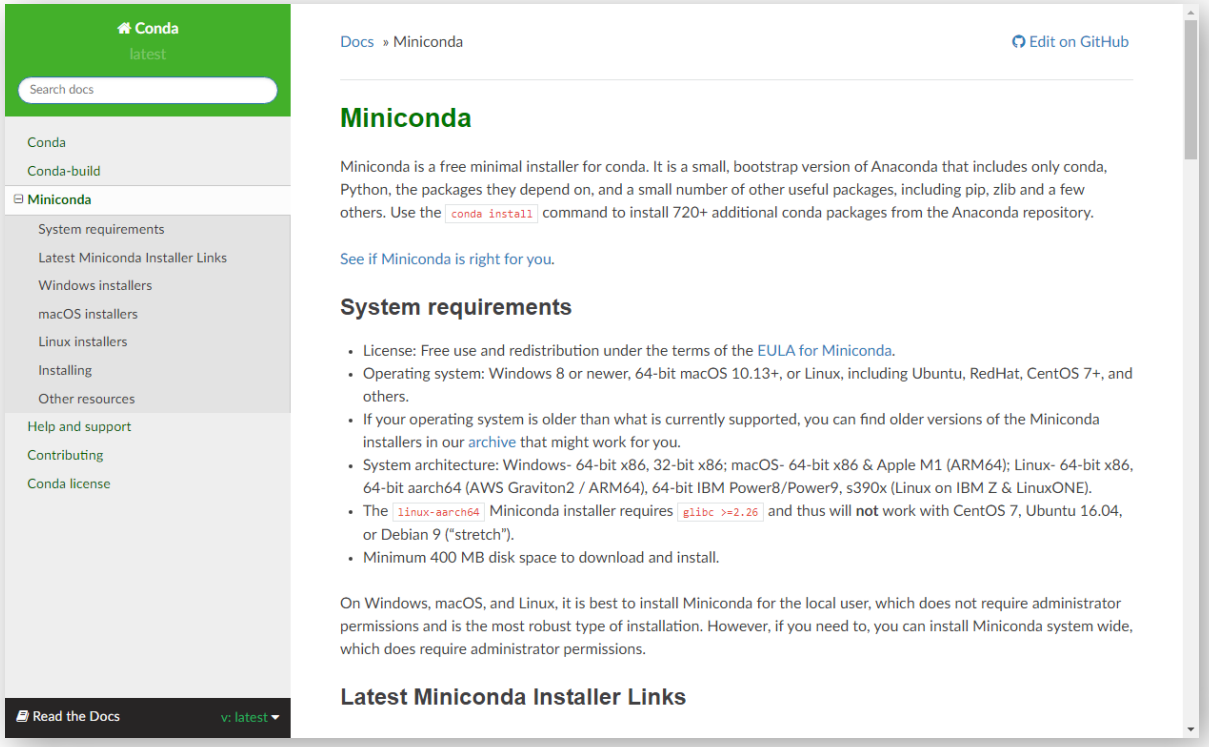
You're almost done!

Last, after selecting the launch button, you will be able to access your AWS Graviton2-based instance via the Amazon EC2 Console.

From there, you will be able to run data science tools on an AWS Graviton2 powered Amazon EC2 instance, using the Anaconda environment.



Using Anaconda on AWS Graviton2 with the Miniconda Installer



Instead of using a fully pre-installed and pre-configured environment, it is sometimes convenient to add just what you need and start with a minimalistic environment on top of a pre-existing, already launched Amazon EC2 instance using the Linux distribution of your choice.

This is the ideal use case for the Miniconda installer: a stripped-down version of the Anaconda installer with minimal dependencies.

Lets go through the steps of installing Miniconda on top of an existing Linux distribution running on an AWS Graviton2-based instance.

Using Anaconda on AWS Graviton2 with the Miniconda Installer

First, download and launch the installer:

```
$ wget https://repo.anaconda.com/miniconda/Miniconda3-latest-Linux-aarch64.sh
$ chmod a+x Miniconda3-latest-Linux-aarch64.sh
$ ./Miniconda3-latest-Linux-aarch64.sh
```

The installer will prompt you to accept the Anaconda License Agreement, then asks for an installation directory. You can leave the default values here.

Once the installation is completed, you'll have to logout/login or reload your shell configuration file for the Anaconda environment to be initialized.

You can now test the benefit of using Anaconda on Graviton2 by installing a custom environment (`demo.yml`):

```
name: demo
channels:
  - conda-forge
dependencies:
  - numpy
  - hdbscan
  - joblib
  - kiwisolver
  - matplotlib
  - numba
  - pandas
  - pillow
  - pyarrow
  - scikit-learn
  - scipy
```

In order to install the environment, you need to call the conda env create command:

```
(base) $ conda env create -f demo.yml
```

Conda will install all the requested packages so that you'll be able to run your data science workloads.

As an example, determine that `numpy` is now available by launching a Python interpreter and importing the `numpy` module:

```
(base) $ conda activate demo
(demo) $ python
Python 3.9.7 | packaged by conda-forge | (default, Oct 10 2021, 15:08:54)
[GCC 9.4.0] on linux
Type "help", "copyright", "credits" or "license" for more information.
>>> import numpy
>>>
```

Want more information?

Watch a demo and check out these resources:

- Announcement: Anaconda Support on AWS Graviton2
- More info on AWS Graviton
- How to get started on AWS Graviton based instances
- Install Anaconda on AWS Graviton2 (arm64)

Amazon Web Services (AWS) provides a broad and deep choice of Amazon Elastic Compute Cloud (Amazon EC2) instances to match the wide spectrum of computing needs of our customers such as general purpose, compute-optimized, memory-optimized, storage-optimized, and accelerated computing workloads. This enables customers to choose the most cost-effective instance type suitable for their data science workload.

With more than 25 million users, Anaconda is the world's most popular data science platform and the foundation of modern machine learning. Pioneering the use of Python for data science and being a champion of its vibrant community, it continues to steward open-source projects that make tomorrow's innovations possible. Anaconda enterprise-grade solutions enable corporate, research, and academic institutions around the world to harness the power of open-source for competitive advantage, groundbreaking research, and a better world.

© 2022 Anaconda, Inc.