

Portability Across Cloud Computing Services

Andres Riofrio, Chris Bunch, Chandra Krintz

Department of Computer Science, University of California, Santa Barbara

Background

Cloud computing providers rent out virtual machines by the hour or minute. Each public cloud provider has a different pricing and execution model.

Problems

Unfamiliarity for non-Computer Science users

Automatically generate cloud applications: Oration.
Automatically deploy applications: Neptune.

Difficult to switch: "lock in"

Task-based abstraction layer: Cicero API.

Data

Compare programming languages

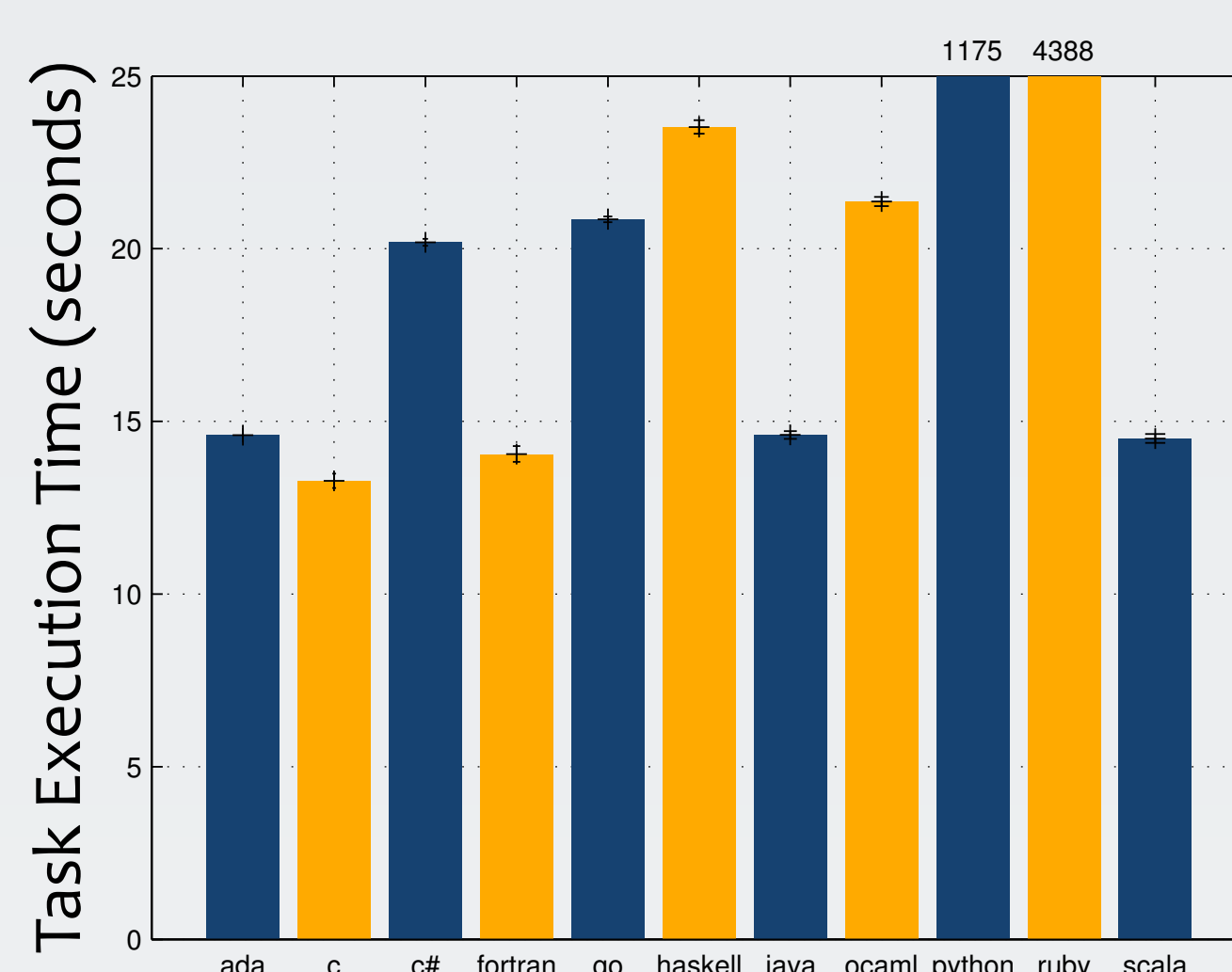


Figure 1 Running time for implementations of the n-body benchmark in different programming languages. Average over ten runs.

Language	Cost to Run Task
Ada	\$0.0076 ± \$0.0002
C	\$0.0069 ± \$0.0002
C#	\$0.0105 ± \$0.0000
Fortran	\$0.0073 ± \$0.0003
Go	\$0.0105 ± \$0.0000
Haskell	\$0.0120 ± \$0.0000
Java	\$0.0075 ± \$0.0000
OCaml	\$0.0110 ± \$0.0000
Python	\$0.5876 ± \$0.0057
Ruby	\$2.1944 ± \$0.0198
Scala	\$0.0075 ± \$0.0000

Figure 2 Cost for programs in Figure 1, using hypothetical per-second pricing model on Amazon EC2. Average over ten runs.

Compare cloud computing providers

Cloud Service	Cost
Amazon EC2	\$0.3200 ± \$0.0000
Google App Engine (Java)	\$0.0013 ± \$0.0000
Google App Engine (Python)	\$0.0049 ± \$0.0006
Microsoft Azure Worker Roles	\$0.0200 ± \$0.0000

Figure 3 Cost for n-body benchmark ran on different cloud computing services. Using per-hour pricing for EC2 and Azure, per-minute pricing for AppEngine.

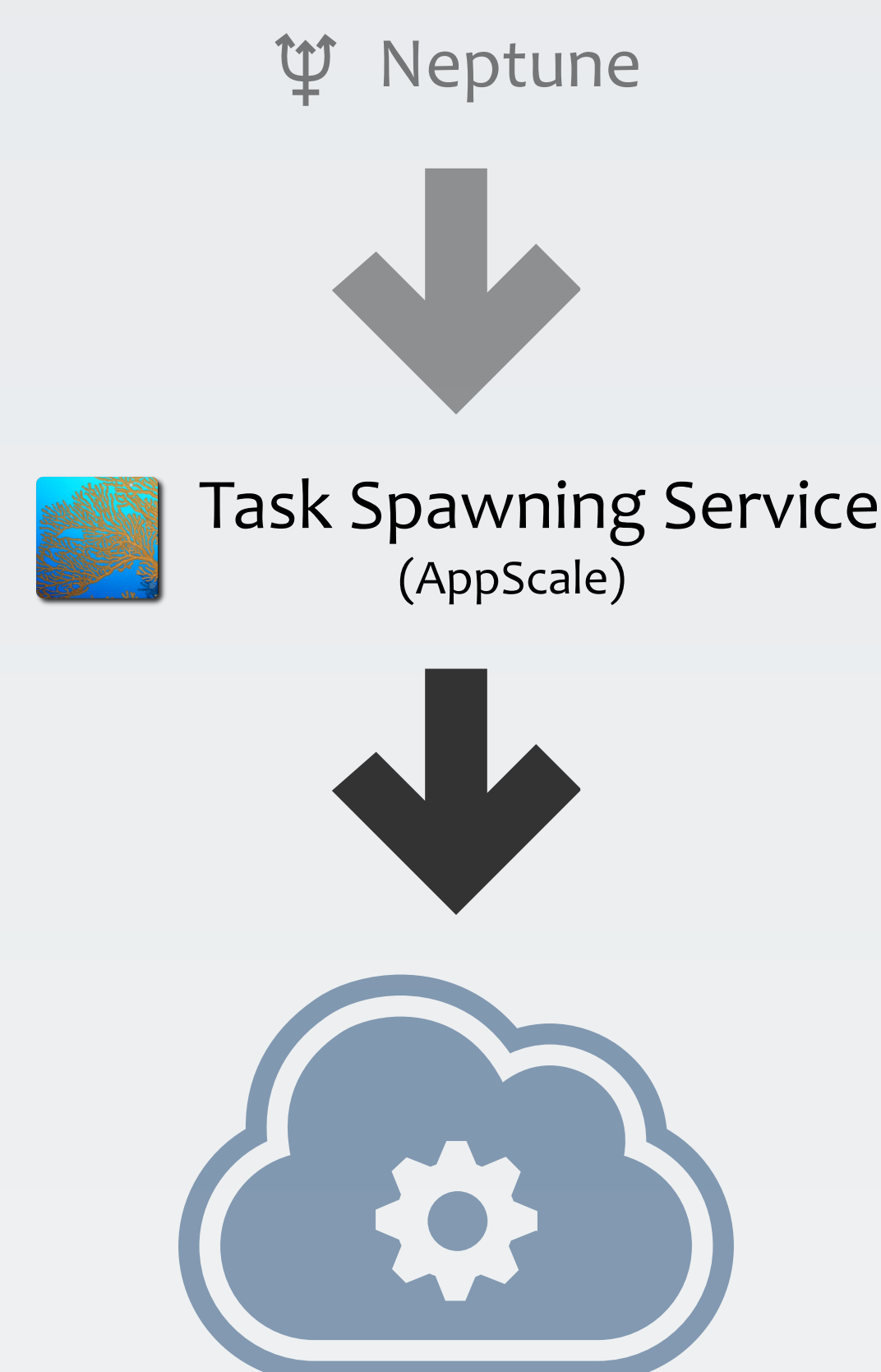
Architecture

Neptune language

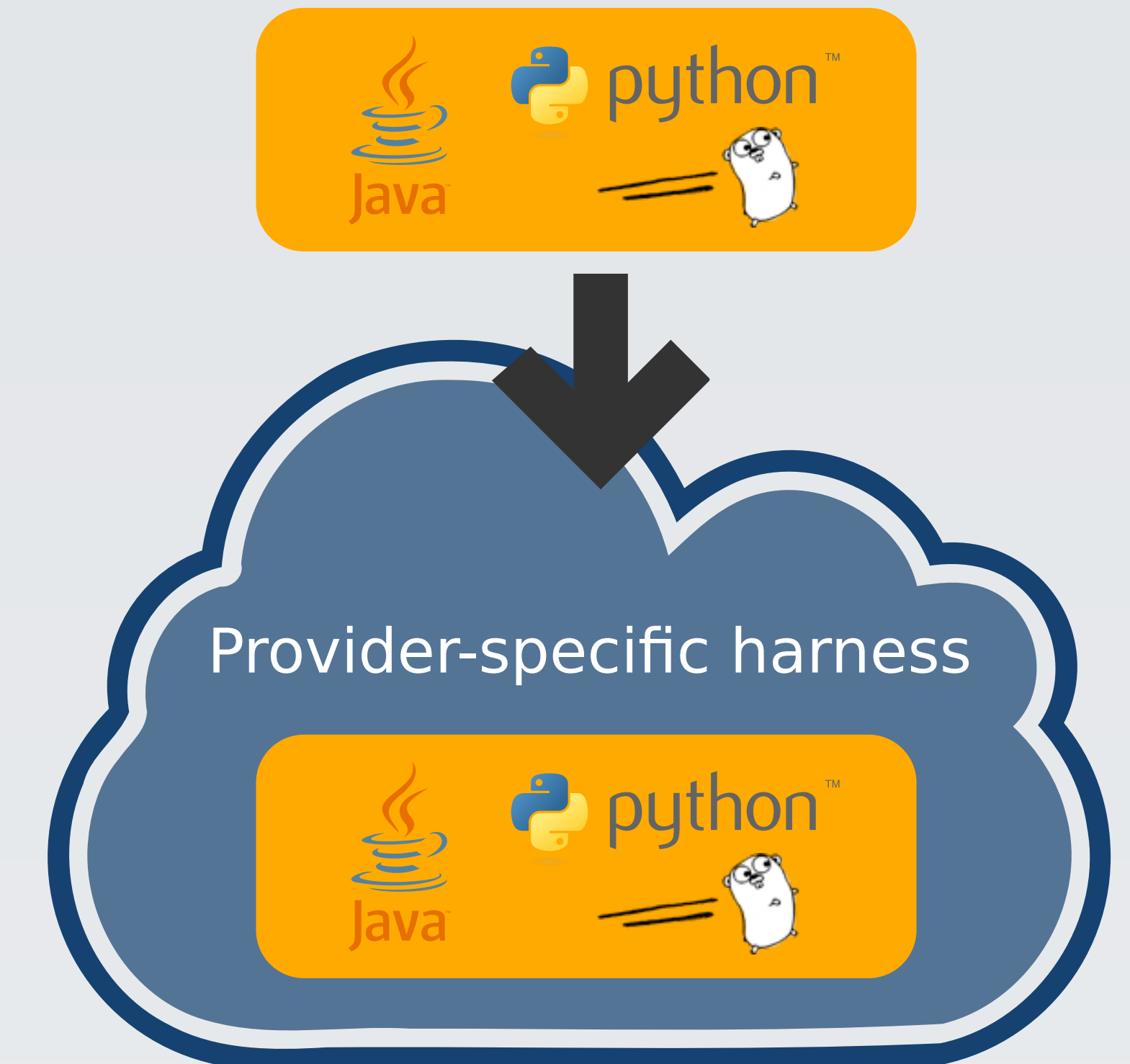
```
babel :file => "~/Code/dfsp/dfsp.py",
      :storage => "s3",
      :worker => "medea",
      :queue => "sqs",
      ...
```



Cicero interface



Oration generator



Cloud computing services

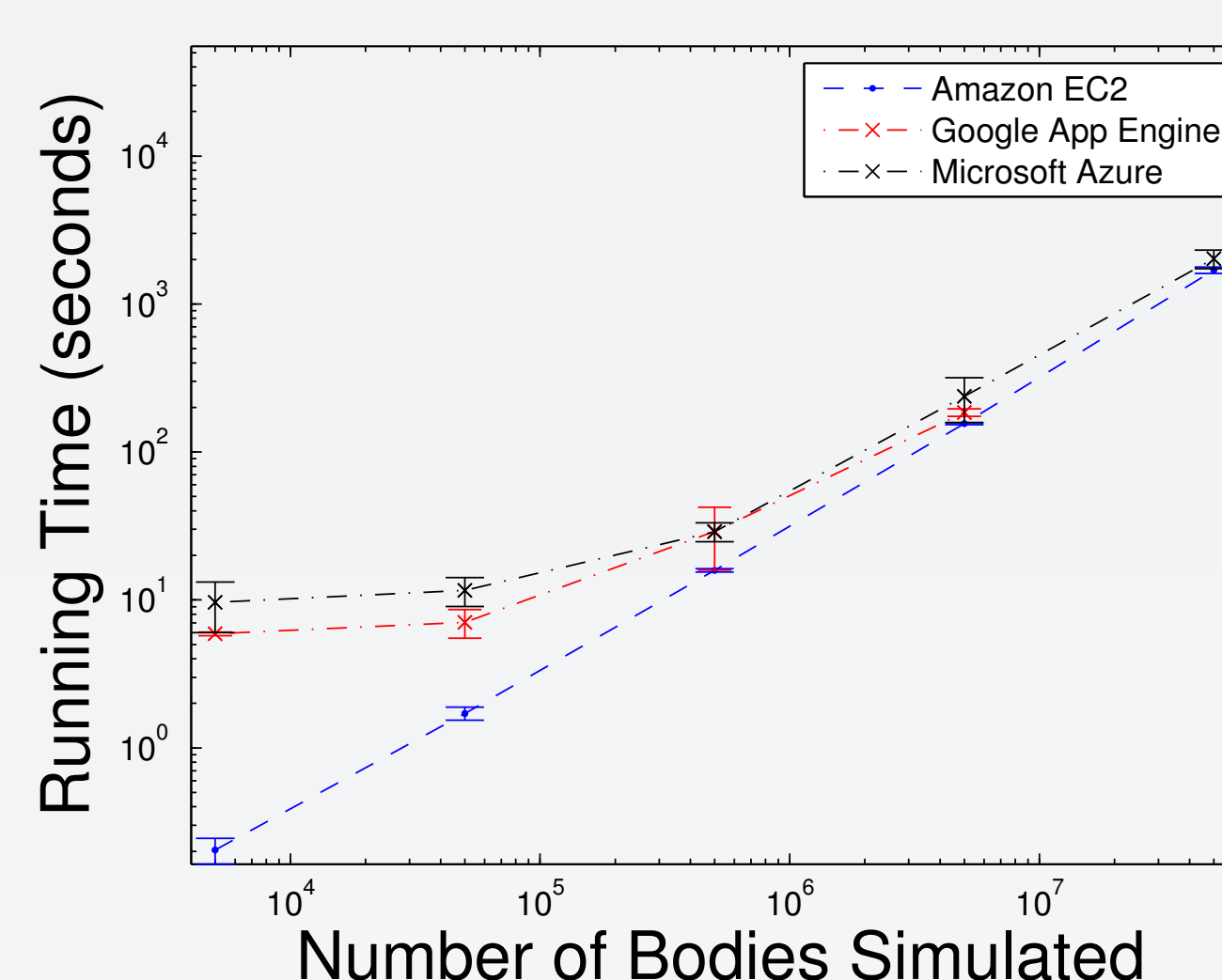


Figure 4 Running time for benchmarks in Figure 4, using a Java implementation of the n-body benchmark.

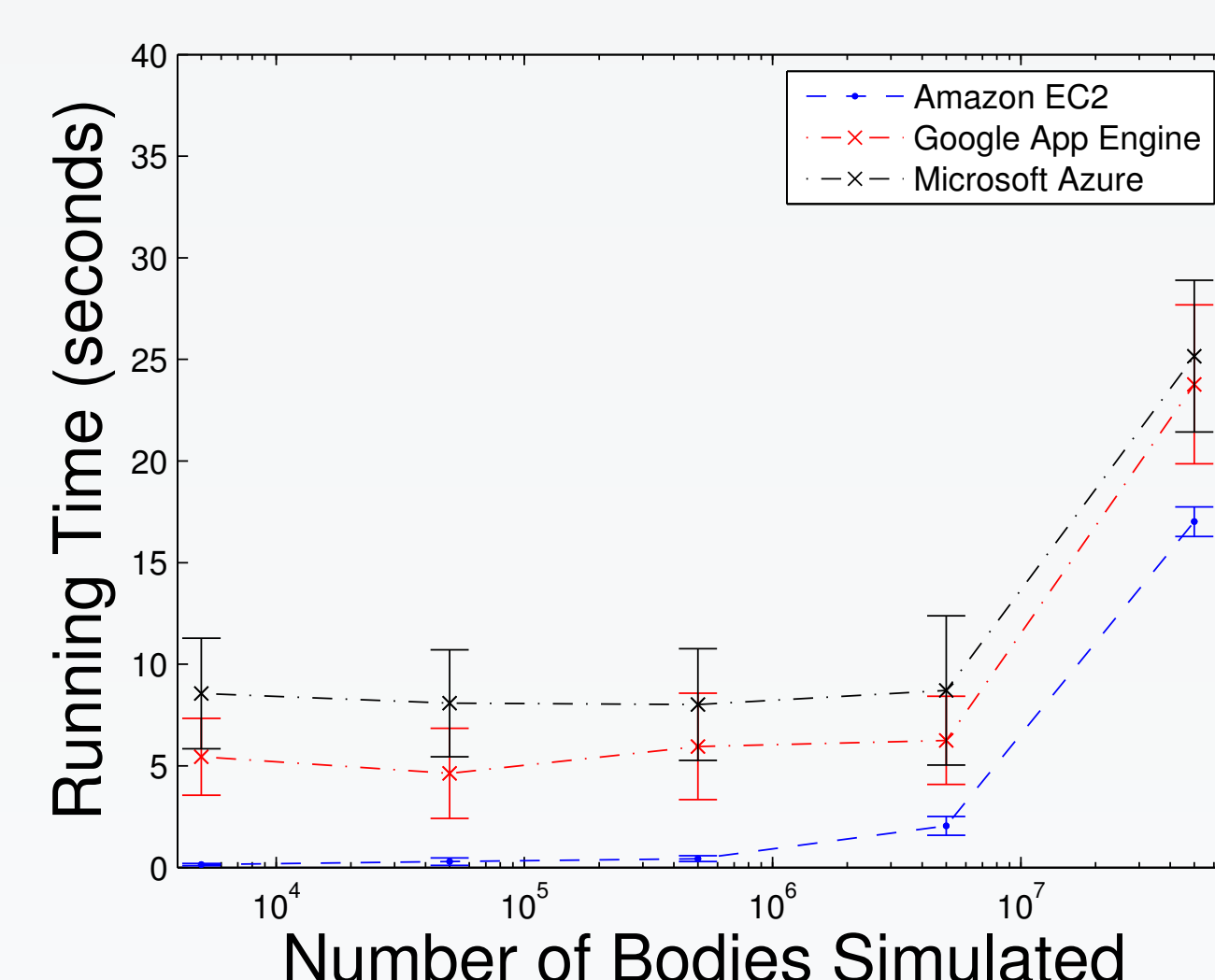


Figure 5 Running time for benchmarks in Figure 4, using a Python implementation of the n-body benchmark.

Current and Future Work

Since some cloud computing providers are faster than others depending on the specific application, the choice of provider is important to minimize execution time or the cost of execution. Scientists generally don't want to learn the nuances of each provider; instead, we are working on automatically choosing for them based on empirical performance measurements.

