

## Unravel Newsletter: April 2019

Greetings from Unravel,

Hope you had a fun April Fools' Day this past Monday! What we are going to discuss next is some serious business (in the cloud).

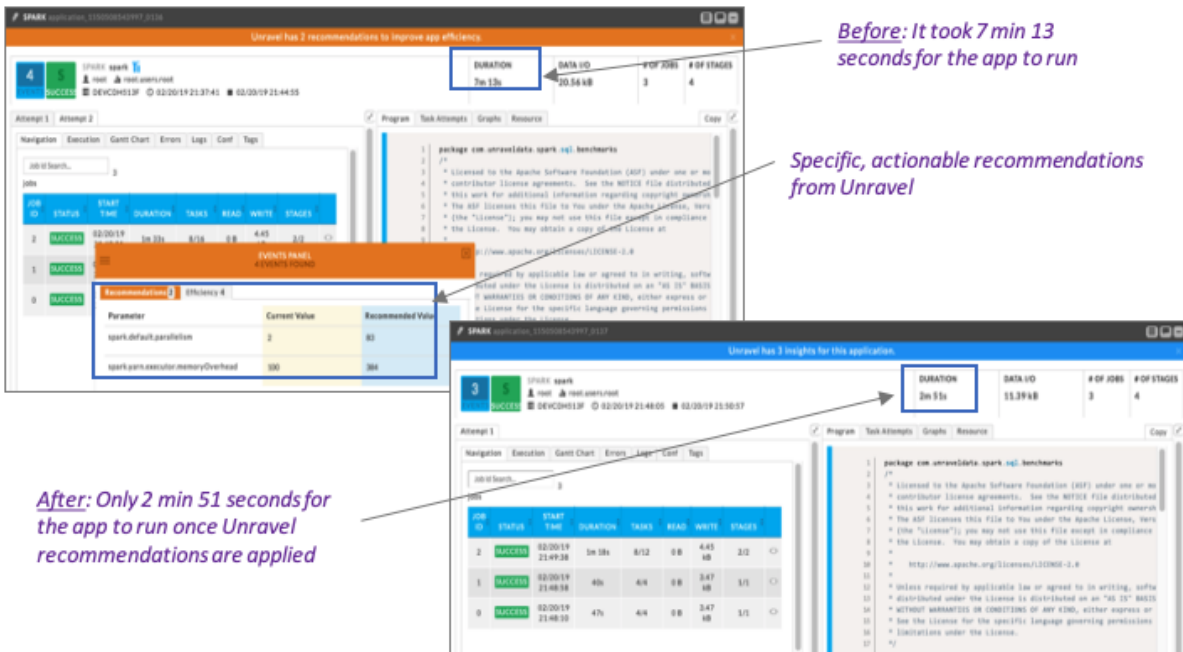
The results from [this survey](#) we carried out (79% of the respondents are already running big data apps in cloud, 83% have strategy to), quote from [Forrester Research](#) (global spending on big data solutions via cloud subscriptions will grow almost 7.5 times faster than on-premise subscriptions) and many other sources depict a clear trend of migration of big data apps to the cloud.

Like it or not, the cloud wave is here! We have heard from some of you already about your plans and Unravel fully intends to help you through your journey to Plan, Migrate and Manage your modern data applications in the cloud. Our goal is to cover the entire range of big data platforms and services in the cloud: From IaaS (CDH/HDP/MapR on cloud instances) to PaaS (e.g. AWS EMR, Azure HDInsight etc.) to Cloud-Native (e.g. Redshift, Databricks etc.) to Serverless (e.g. Athena, Big Query etc.). So, whatever the choice you make, you can enjoy all the same Unravel value props!

### Running Big Data Workloads in the cloud?

Currently, if you are running your modern data applications on Cloud IaaS or AWS EMR or Azure HDInsight, Unravel can help with you in the following ways:

- **Maximize Performance:** Unravel provides AI powered insights and recommendations to speed up your applications in the cloud while using the same amount of resources.



Unravel Insights and Recommendations help speed-up your big data apps running in the cloud

- **Reduce Cloud Expenses:** Unravel provides insights and recommendations for running apps more efficiently and optimizing storage (e.g. remove/tier unused or less frequently used data). Unravel Auto-Actions feature enables you to automatically kill rogue apps that could cause cost overruns. You can optimize your cluster-size by auto-scaling or decide to change to less expensive instance types by utilizing the usage information made available by Unravel. These are just a few ways Unravel can help you minimize costs in the cloud.
- **Fast Troubleshooting & Root Cause Analysis:** Unravel automatically detects Root Cause of the app's failure and also provides specific recommendations to prevent such failures in future. Also, makes all the information needed for troubleshooting available at your fingertips – metrics, errors, logs, configuration parameters and more. As a result, you can achieve better MTTR, reduction in number of tickets and improved productivity.
- **Operational Insights:** Learn who all are using the cluster and how. See which apps are running and resources used per user, queue, department, project etc. View Chargeback/show back reports and so on.

More details coming soon - check our [blog](#) for updates.

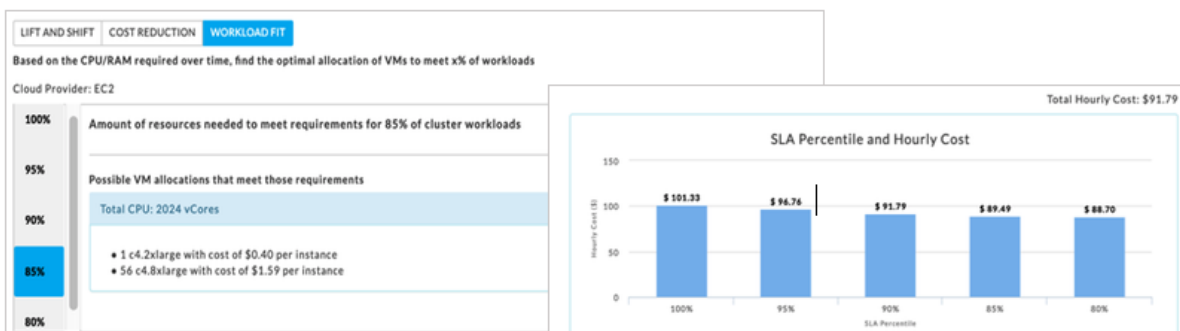
## Looking to Migrate your Big Data Workloads to the Cloud?

This is no easy task and Unravel aims to help. We built a Goal-Driven and Adaptive Solution for helping Migrating Big Data Applications to the Cloud. We provide the following:

**Pre-migration - Planning** that starts with a “Cluster Discovery”: Understanding your cluster - topology, services, apps running etc. Then identifying apps suitable for the Cloud, e.g. ones that are bursty in nature or apps belonging to tenants specific to the Business – whatever strategy you choose.

The next part of planning is **Mapping your On-Premise Cluster to a Deployment Topology in the Cloud** based on your goal/strategy for the migration. The first two methods below provide a one to one mapping of each existing on-premise host to the most suitable instance type in Cloud. The third one provides a list of instances that match the entire on-premise cluster’s resource usage.

- *Lift and Shift*: Least Risky. Ensures that your cloud deployment will have the same (or more) amount of resources available as your current on-premise environment.
- *Cost Reduction*: Topology in cloud that matches your on-premise usage. If your on-premise hosts are underutilized this method will always be less expensive than lift and shift.
- *“Workload Fit”*: Determines optimal assignment of cloud instance types to meet requirements while minimizing cost. Typically, the most cost-effective, also offers price-performance trade-off.



*Unravel provides you a mapping of your current on-premise environments to a cloud based one. Strategy: “Workload Fit”*

Next, Unravel can help you **track the migration and its success** as you move your big data apps from on-premise to cloud. Unravel’s insights and recommendations can also help bring your migration back on track in case the app’s performance is not up to par in its new home.

More details in this blog: [Getting the Most From Modern Data Applications in the Cloud - Part 1, Planning and Migration](#) and also check out the [Unravel Cloud Operations Guide](#).

## Webinars

- Register for a webinar with Unravel Software Engineer, Alejandro Fernandez, on April 10 at 1PM ET and learn how to assess, plan, execute, and validate a successful migration of data workloads to the cloud: [Using AI-powered Automation for High Performance Data Pipelines in the Cloud](#)
  - Check out this Unravel demo: [Simplifying DataOps: Unravel demo with Kunal Agarwal and Henry Eckerson](#) that provides a glimpse into how Unravel makes sure your Big Data apps are fast and reliable, and your entire Big Data infrastructure and setup is cost-efficient and highly utilized.
- 

## Upcoming Events

- Join us at the [#SparkAISummit](#) in SF for a session on April 25 at 11:50 AM PT where our CTO, Shivnath Babu, will discuss how organizations can leverage [#AI](#) chatbots to enhance their [#Spark](#) performance mgmt. Details [here](#).
  - Join us at the [Strata Data Conference](#) in London for a session on May 1 at 16:35 where our CTO, Shivnath Babu, will discuss how to build a decomposable time-series model, for optimal cost and resource allocation for the big data stack. Details [here](#).
- 

## Community Highlights

Keeping with the Cloud theme this month

- Microsoft Software Engineers share their experience and learnings from running one of world's largest Kafka deployments - this is the main component of a service that achieves near real-time data transfer to handle up to 30 million events per second. Running on *Azure cloud with Apache Kafka on HDInsight*: Read the blog from Microsoft, [Processing trillions of events per day with Apache Kafka on Azure](#).
  - In this [post](#), AWS Solutions Architects focus on cost-optimizing and efficiently running Spark applications on Amazon EMR by using Spot Instances.
  - This [article](#) describes a number of options for copying data between S3 and HDFS in context of an EMR cluster, and it shows how (based on a couple of simple Presto queries) querying data in HDFS can be much faster.
- 

## Resources

- [Learn more](#) about Unravel.
- [Online Product Demo](#)

- [Unravel Partners](#)
- [Unravel Product Releases and Documentation](#)
- [Unravel Datasheet](#)
- [More Unravel News](#)



[Contact Us](#). [Sign Up for 30-day Trial](#).

2019 Unravel. All Rights Reserved. 2 Palo Alto Sq, Suite 120, Palo Alto, CA 94306