Unravel Newsletter: January 2020

Greetings from Unravel,

Hope you have had a restful holiday and wish you very Happy New Year!

In this month's newsletter, we start with discussing a much awaited and requested feature: multi-cluster support in Unravel. Then, we will discuss two promotions for getting started on Unravel cloud offerings, followed by an exciting announcement about Unravel joining the AWS Partner Network Global Startup Program. We will continue with some community highlights in the area of Data and Analytics and include links to recent recordings and upcoming events that you may find interesting.

Multi-Cluster Support in Unravel

For the past few months, we have discussed with you and understood that for a variety of reasons including the following, there is an urgent need for adding multicluster support for Unravel:

- Analyze workloads from Test, UAT and Prod clusters in one place
- Comparing performance of apps as it goes through SDLC
- Single pane of glass into your business across all clusters
- Ease of management of one Unravel instance

Our engineering team has been hard at work to come up with the architecture and implementation for this. With this in place, here are some of the features and functionality you can expect:

 On various operational screens, you can choose to view metrics aggregated across clusters or choose to see the details of one of the selected clusters only. For example:

Select "All Clusters" to see aggregate metrics or select any one of the clusters being monitored to see its details

OPERATIONS APPLICATIONS	TIONS REPORTS API				Bearch by app name, us	er, table er Dost. Q.	🖈 Docs 🙀 Support	🛔 admin 🛛 🔻
OPERATIONS								
USINGE DETAILS					🗊 REFRESH	etember 28, 2019 - De	cember 27, 2019 - Daily	AllChatters
FINISHED YARN APPLICATIONS	2 OPEN SECTION	RECENT EVENTS & ALERTS						
Θ_{α}	success 🛞		FAILED	0	KILLED	-		narketin 8
\sim	4		0		0	■ 11/22/2019		cluster1- finance
	~~~~1_					10:24	Long running YARN application	sales- cluster0
RUNNING YARN APPLICATIONS Last checked	at 11/22/19 13:27:37				2 OPEN SECTION	06:12	Long running YARN application	finance- cluster-1
4	RUNNING		$\ominus$	PENDING		₩ 11/3D/3D19		finance-
	0			0		01:51	Long running YARN application	
						■ 11/07/2019		
RESOURCES					2 OPEN SECTION	23:04	Long running YARN opplication	
VCores-Total		2 SHOW MORE	Memory-Total		2 SHOW MORE	23:03	Regue user allocated memory	
10		=	38.08		=			
			17.18			22:44	Long running YARN application	
5	. \A		1 21 GB			22:41	Long running YARN application	

 On the Applications View, you can see the Cluster on which any given application ran by looking at the "Cluster ID" column, or you can choose to see the list of applications running on a particular cluster only by filtering on the corresponding Cluster ID.

In the "Applications View", Cluster ID column shows which cluster the Application ran on

Unravel OPERATIONS	APPLICATIONS	5 REPORTS	API						earch by app name	e, user, table or clust	0, 🖈 Docs 🙀 Supp
PPLICATIONS											
SHOW Applications Worldlows									$\mathbf{N}$	September 28, 201	7 - December 27, 2019 +
Showing 28 results RESET	TALE 1	STATUS	USER	1 APP NAME / ID	A (	START TIME	DURATION	I READ	¢ WRITE	CLUSTER ID	1 QUEUE
FILTER BY APP NAME	SPARK	SUCCESS	tepcat	org.apache.apark.examples.SparkPl application_1574446543531_0001	Yi 🖷	11/22/19 10:22:46	4n 29s	08	0B	democluster-2	default
Search by app name	57A1X	SUCCESS	tepcat	org.apache.spark.examples.SparkPi application_1574431300239_0003	¥i #	11/22/19 06:11:11	4m 45s	08	08	finance-cluster-1	default
APP TYPE Cascading[0]	MR	SUCCESS	h dfs	QuasiMonteCarlo Job_1574431300237_0002	ñ	11/22/19 D6:09:46	264	2698	2158	finance-cluster-1	default
2 Catcading(0) 2 Impala(0)	SPARK	SUCCESS	root	hello-world.py application_1574431300239_0001		11/22/19 D6:09:18	285	08	08	finance-cluster-1	default
MapReduce(5)	\$PARK	SUCCESS	reat	hello-world.py application_1574277983179_0001		11/20/19 11:33:31	11s	08	OB	sales-cluster1	default
2 Spark(23) 2 Tex00	SPARK	SUCCESS	root	hello-world.py application_1574254216284_0001		11/20/19 D5:23:29	10s	08	08	finance-cluster-0	default

In the "Applications View", select only the applications that ran on specific cluster(s) by filtering on those

	MR	SUCCESS	hdfa	QuesiMenteCarlo job_1574431300239_0002	т	11/22/19 06:09:46	269	2698	2155	finance-cluster-1	default	
	37476	5000235	reet	helio-world.py application_1574431300239_0001		11/22/19 06:09:18	285	08	08	finance-cluster-1	default	
iz Impolaj0) iz MapReduce(5) iz Pigi0)	3844K	500CE85	reet	helio-world.py application_1574277983179_0001		\$1/20/19 \$1:33:31	115	08	08	sales-cluster1	default	
<ul> <li>Spark(23)</li> <li>Tep30</li> </ul>	SPARK	SUCCESS	reet	helio-world.py application_1574254216284_0001		11/20/19 05:23:29	10s	08	08	finance-cluster-0	default	
2 Athenal 01 ESSANEW	SPARK	SUCCESS	topcat	org.apache.apark.examples.SparkPl application_1574242972025_0008	¥6-4	11/20/19 01:50:29	4m 51s	08	08	sales-cluster0	default	
STATUS	MR	SUCCESS	hdts	QuasiMonteCarlo joit_1574242972005_0002	Ti	11/20/19 01:49:31	254	2668	215B	sales-cluster0	default	
R Success (28) R Fulled (20)	SPARK	SUCCESS	rest	helio-world.py application_1574242972035_0001		11/20/19 01/47/08	241	00	00	sales-cluster0	default	
E Killed (0) E Ranning (0)	SPARK	SUCCESS	rest	helio-world.py application_1574072432777_0003		11/18/19 02:26:55	141	00	03	cluster-7317	default	
it Welling (0) it Pending (0)	SPARK	SUCCESS	rest	helio-world.py application_1574072432777_0002		11/18/19 02:27:47	151	DB	05	claster-7317	default	
H Unknown (0)	SPARK	5000835	rest	helio-world.py application_1574072432777_0001		11/18/19 02:25:29	21s	CB	05	claster-7317	default	
	SPARK	\$UDCE88	reet	helio-world.py application_1573183590415_0007		11/08/19 12:27:37	95	08	08	cluster1-finance	default	
	SPARK	NUCCESS	topcat	example0-before application_1573168147795_0009	¥6.4	11/07/19 23:02:48	8m 295	10.208	154.98B	daster1-marketing	default	
USER	SPARK	SUCCESS	tapcat	example3-after application,1573168147795,0008		11/07/19 22:43:45	8m 22s	10.26B	154.9KB	claster1-marketing	default	
CLUSTER												Page 1 of 2 IP IN
1												
ckoter1-finance												
cluster-7317 finance-cluster-1												
mance-cluster-1 sales-cluster0												
democlaster-2												
finance-cluster-0												

• Various Reports support the same functionality as well, you can select "All Clusters" to see aggregate metrics across clusters.

Chargeback Report: Select "All Clusters" to see aggregate metrics or select any one of the clusters being monitored to see its details separately

C unravel OPERATIONS	APPLICATIONS REPORTS API				Search by app name, user, table or clust	Q # Docs	wil Support	🛔 admin 🛛 👻
REPORTS							<u> </u>	
OPERATIONAL INSIGHTS DATA INSIGHTS					₩ 5e	plexber 28, 2019 - Dec	ember 27, 2019	Al Clasters +
CHARGEBACK YANN CHARGEBACK I	IMPALA CLUSTER SUMMARY CLUSTER	COMPARE CLUSTER WOR	KLOAD					All Clusters
GROUP BY								cluster1- marketin
Est Application Type User Que	c.c				VCore / Hour (\$) Amount P	Memory MB / Hour (5)	Amount	e ckster1- finance
Showing Top Results in Application Ty	199							sales- cluster0
								finance- cluster-1
	App Count			Cpu Hours		Memory Hours		finance-
(	)	1paris 82.34%	C	€ spirkPR.04		)		• 1904:193.42% m=0.58%
	pe (September 28, 2019 - December 27, 201	191						
CHARGE BACK							±0	OWNLOAD CSV
APPLICATION	TYPE: APP COUNT)	CPU HOURS1	SLCPU HOURS	MEMORYHOURS	% HEHORY HOURS	CPU COST0		HEHORY COST:
spark	23	4.03	90.84%	2Y 1M 1D 4h 24m 49s	59.42N			
er .	5	0.05	1.16%	4D 9h 47in 52s	0.58%			
Showing YARN Applications: Septem								DOWNLOAD CSV
TYPE 1 STATUS 1 USER	APRINAME/ID	QUEUE 1 STA	RTTIME	E PINISHED TIME E	MEMORY ME SECONDS	1 VCORE SE	CONDS	1
SPARK SUCCESS Topical	org.apacke.spark.examples.SparkPl application_1573168147795_0001	default 11/0	7/19 15:17:46	11/07/19 19 29:22	10.478,246	1.780		
	new score he smark, su seen les Smark Pi							

#### Cluster Compare Report: Select "All Clusters "or a specific cluster only



#### **Promotions for getting started with Unravel Cloud Offerings**

Unravel is offering some exciting promotions for you to try out two of our Cloud offerings:

- Try Unravel on Amazon EMR for 14 days with no obligation to buy, and we will credit you \$2000 in AWS credits which you may redeem for your AWS account. Start here.
- Try Unravel on Azure HDInsight for 14 days with no obligation to buy, and we will credit you \$2000 in Azure credits which you may redeem for your Azure account. Start here.

#### **Unravel Joins AWS Partner Network Global Startup Program**

We are excited to announce our membership in the Amazon Web Services (AWS) Partner Network (APN) Global Startup Program.

The APN Global Startup Program is a unique "white glove" support and go-tomarket program for selected startup APN Partners, allowing members to build on their AWS expertise, better serve shared customers, and accelerate their growth. To be selected for the APN Global Startup Program, Unravel had to meet predefined criteria, including a clear, demonstrated product market fit for an innovative enterprise tech product, be backed and recommended by a top-tier venture capital firm, and demonstrate a strategic commitment to building their AWS and cloud expertise. Check out the complete blog here.

Keep subscribing to these newsletters for upcoming exciting Unravel features and functionality or **get in touch** with us directly to get your feature requests in!

# **Conference Sessions and Recordings**

- Watch the recording of this webinar by Unravel and AWS: "Application Performance Management & Operational Intelligence for Amazon EMR"
- Watch the recording of this webinar by Unravel and AWS: Look Before You Leap: Migrating On-Premises Hadoop to AWS
- Check out this webinar Recording: Best Practices: Troubleshoot and Optimize Spark Data Pipelines with Unravel
- Check out this webinar Recording: Unravel Demo Big Data Application
   Performance Management
- Check out How to get started with Unravel for Amazon EMR from the AWS Marketplace.
- Check out How to get started with Unravel for Azure HDInsight from the Azure Marketplace.

## **Upcoming Events & Webinars**

- Join us at Hy's Steakhouse in Toronto for Toronto IT Leaders & "Big Data" Professionals Luncheon (sponsored by Microsoft Azure & Unravel Data), in Toronto, Canada, January 30, 2020.
- Join us at the AWS Spheres for Dinner with AWS and Unravel + Amazon Spheres Tour, in Seattle, February 6, 2020.
- Debugging and Demystifying Spark with Unravel Join Dave Goodhand, Head of Customer Success EMEA, Unravel Data on 30 Jan 7:30 pm PT to learn about uncovering and understanding the key datasets, metrics, and best practices needed to develop mastery with Spark performance management with Unravel Data.

# **Community Highlights**

- Big Data Quality at Criteo: Criteo writes about how they enforce data quality for the 450 PB that they have in their Hadoop clusters. Since Hadoop is very flexible/lenient in the data it supports, they run over 7,000 data quality checks per day using statistics from Hive tables and custom queries.
- Dynein: Building an Open-source Distributed Delayed Job Queueing System: Airbnb has open sourced Dynein, their job queuing system that

they use for offloading tasks from the main request path and performing other asynchronous operations. It uses DynamoDB as a scheduler for future jobs and SQS for queuing—the post describes how this is built in a highly scalable way.

- Running and Deploying Elasticsearch on Kubernetes: An overview of Kubernetes and how to deploy Elasticsearch on Kubernetes. This is a great introduction to many of the core concepts of Kubernetes (e.g. Deployment, Pod), including those that are important for running a stateful service (e.g. StatefulSet, PersistentVolumeClaim). It also shows how to configure accounts for ES using Kubernetes RBAC and use the helm package manager for deploying to Kubernetes.
- A Billion Rows A Second: Wouldn't it be great if you could load a 1 Terabyte data file in Python instantly, reading only the parts you need, with a strategy that smart kernel programmers have spent decades optimizing? The answer is Vaex.

### 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. Take Unravel for a Test Drive. 2020 Unravel. All Rights Reserved. 2 Palo Alto Sq, Suite 120, Palo Alto, CA 94306