

Activator Library

Focus on maximizing the value of your data, gain business insights, increase your team's productivity, and achieve success.



ACTIVATORS

Designed to give your team assistance when you need it most without requiring a major time commitment or lengthy scope of work. Activators let you focus on maximizing the value of your data, gain business insights, increase your team’s productivity, and achieve success. These iterative remote sessions provide you with expert guidance in one-time configurations, optimizations, troubleshooting and transfers of knowledge. Although not all Activators are the same where some are quick one session solutions and others will require a series of meetings, the outcomes remain the same – deliver maximum ROI.

How it Works

Activators are a convenient way to assist you with moving your big data projects forward and fit into your day when you’re available. Simply review your business objectives, review our library and request one or a series of Activators and one of our experts will schedule a call to determine your requirements and match you with an expert(s) with the best skills.

Discovery



The first step in completing each Activator is a discovery session to determine your business requirements and align on a successful outcome.

Web Conference



Through a series of collaborative sessions and remote services our team of experts will iterate through a series of tasks to define, complete and validate the outcome.



Remote Services

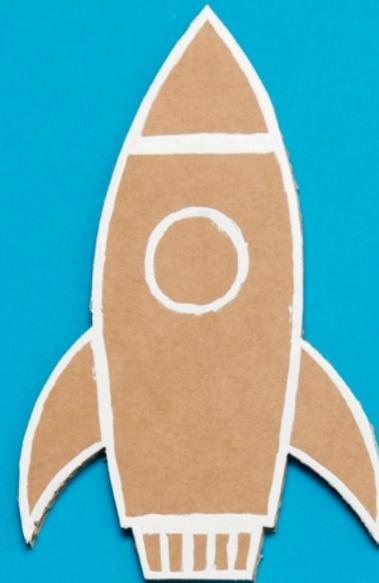


Review



The final step to completing each Activator is agreement that the outcome satisfies your business requirements and next steps, if required, are discussed.

Don't see what you're looking for. Ask us how we can create a custom Activator that meets your business needs.



ACTIVATORS: ACCOUNT CONFIGURATION

Account Configuration Activators focus on cloud account access, permissions, policies and storage configuration options. Select any one of these Activators to ensure your Qubole account is configured properly.

REF	TITLE	OVERVIEW	CLOUD PLATFORM	OUTCOMES
ACT01	Configure Qubole Account with IAM Role	Learn how to configure Qubole with an IAM cross-account role.	AWS	Qubole account configured with an IAM Cross-Account Role
ACT02	Configure Qubole Account with Dual IAM Roles	Learn how to configure Qubole with a dual IAM role.	AWS	Qubole account configured with dual IAM role
ACT03	Configure Qubole Account with Service Principal	Learn how to configure Qubole Account with service principal.	Azure	Configured Qubole Azure account
ACT04	Configure Shared S3	Learn how to configure an S3 bucket to be shared across multiple AWS accounts.	AWS	Configured S3 bucket accessible by more than one account
ACT05	Configure Access to Multiple Blob Storage Accounts	Learn how to configure a QDS cluster to securely access multiple Blob storage accounts.	Azure	Configured access to storage
ACT06	Configure Amazon Virtual Private Cloud (VPC)	Learn how to configure a Qubole cluster in an Amazon VPC with public and private subnets.	AWS	Qubole cluster running within a VPC
ACT07	Configure an Azure Virtual Network	Learn how to configure a Qubole cluster in an Azure Virtual Network.	Azure	Configured virtual network
ACT08	Migrate to Custom Metastore	Gain flexibility with a custom metastore. Get expert help with migrating an existing account to a custom Hive metastore.	AWS and Azure	Custom metastore
ACT09	Troubleshoot Account Configuration	Get expert advice on troubleshooting account connectivity and access configuration issues.	AWS and Azure	Functioning Qubole AWS account

ACTIVATORS: SECURITY

Security Activators provide ease of access to your data in a structured and controlled manner. Select any one of these Activators to further secure your data and Qubole environment.

REF	TITLE	OVERVIEW	OUTCOMES
SEC01	Manage Access to QDS	Develop a data governance strategy for your organization.	Data security strategy Groups, users and roles
SEC02	Configure Single Sign- on (SSO)	Create a seamless secure login process for your organization by enabling SSO with SAML or Google Authorization Service.	Configured SSO access to Qubole
SEC03	Configure Encryption for Data at Rest and in Transit	Provide secure access to your organization's data.	Encrypted data
SEC04	Security Review on Data Requests	Learn security best practices when accessing data through queries and commands.	Security review on one data request
SEC05	Configure Hive Authorization	Enhance security on data residing in Qubole HDFS with Hive Authorization.	Secured Hive tables

ACTIVATORS: CLUSTER MANAGEMENT

Cluster Management Activators provide best practices around deploying and managing clusters. Select any one of these Activators to learn more about proper cluster configuration and run your workloads at optimal performance.

REF	TITLE	OVERVIEW	CLUSTER	OUTCOMES
CLM01	Configure a QDS Cluster	Learn best practices and appropriately size clusters to run at peak performance and optimal cost.	Hive, Spark and Presto	Create an optimally configured Spark cluster and obtain knowledge around Spark cluster sizing best practices and cost optimization
CLM02	Create Custom Bootstrap Script	Customize the installation, management and configuration of the tools useful for cluster monitoring and data loading.	Hive, Spark, Presto and Airflow	Enable additional functionality on the Hive cluster
CLM03	Add User Defined Function (UDF)	Expand the functionality of one of Qubole's data engines with a custom UDF.	Hive, Spark and Presto	Configure and deploy a custom UDF
CLM04	Add Custom JAR File	Expand the functionality of one of Qubole's data engines with a custom JAR.	Hive, Spark and Presto	Enable a JAR file for cluster
CLM05	Configure New Qubole Environment	Expand the functionality of one of Qubole's data engines with a custom JAR.	Hive, Spark and Presto	Create a New Qubole Environment
CLM06	Configure RubiX	Improve the performance of your Presto queries and share the cache data across multiple jobs.	Presto	Presto configured to run with RubiX
CLM07	Configure an Airflow Cluster	Learn the fundamentals of configuring an Airflow cluster.	Airflow	Configured Airflow cluster

ACTIVATORS: OPTIMIZATION

Optimization Activators assist you with tuning and troubleshooting your jobs. Select any on of these Activators to learn more about getting peak performance and reduction in costs.

REF	TITLE	OVERVIEW	CLUSTER	OUTCOMES
OPM01	Optimize a Qubole Job	Be it increased performance or cost reduction get the most out of each job through optimization best practices.	Hive, Spark and Presto	Workload optimized for performance and cost
OPM02	Troubleshoot a Qubole Job	Unblock dependent operations and learn the fundamentals of troubleshooting a job running on a Qubole cluster.	Hive, Spark and Presto	Restored dependent operations
OPM03	Enable Spark Streaming	Process real-time data from various data pipelines while reusing the same code for batch processing, join historical data, or run ad-hoc queries on stream state.	Spark	Functionality to leverage streaming use cases
OPM04	Tune Spark Application with Sparklens	Learn how to tune Spark applications with Sparklens.	Spark	Knowledge on leveraging Sparklens to optimize Spark workloads

ACTIVATORS: MIGRATION

Migration Activators give you the flexibility to use the right data engine solution for the challenge at hand. Select any one of these Activators to learn more about a new data engine and steps required for migration.

REF	TITLE	OVERVIEW	CLUSTER	OUTCOMES
MIG01	Migrate SQL Job to Hive SQL	Realize the benefits of Apache Hive and migrate a SQL job to Hive SQL.	Hive	Migrated and tested Hive QL
MIG02	Migrate Hive MapReduce Job to TEZ	Build high performance batch and interactive data processing native YARN applications.	Hive	Tez job that Integrates natively with Apache Hadoop YARN and performs well within mixed workload clusters
MIG03	Migrate Hadoop 1 Job to Hadoop 2	Realize the benefits in scalability, performance and reliability with Hadoop 2.	Hadoop	Updated code running on Hadoop 2
MIG04	Migrate Hadoop Job to Hive	Realize the benefits of Apache Hive and migrate a Hadoop job to Hive SQL.	Hive	Migrated and tested Hive job
MIG05	Migrate Hadoop Job to Presto	Provide analysts with low latency ad-hoc SQL query capabilities across HDFS and other data sources.	Presto	Migrated and tested Presto job
MIG06	Migrate Hive Job to Spark	Expand the capabilities of your data engineering and data scientists teams with Apache Spark.	Spark	Migrated and tested Spark job
MIG07	Migrate Hive Job to Presto	Migrate one Hive query to Presto.	Presto	Migrated and tested Presto job

ACTIVATORS: SCHEDULERS

Scheduler Activators allow your workloads to run on time and error free. Select any one of these Activators to learn more about Qubole's built in Scheduler or Airflow for advanced workflow scheduling.

REF	TITLE	OVERVIEW	SCHEDULER	OUTCOMES
SCH01	Schedule a Job with the Qubole Scheduler	Configure jobs, specify intervals at which they run, and learn more about the additional capabilities that make the Qubole Scheduler a powerful tool for automating your workflow.	Qubole	Scheduled workload
SCH02	Create and Deploy DAG to an Airflow Cluster	Learn the fundamentals of developing and deploying DAGs to an Airflow cluster.	Airflow	Validated and deployed DAG in Airflow
SCH03	Implement Airflow Connections and Variables	Implement Airflow Connections and Variables for a single use case or DAG.	Airflow	Implemented connection or variable for Airflow
SCH04	Create DAG to Schedule an External Job	Manage workloads external to Qubole with an Airflow DAG.	Airflow	Qubole Airflow DAG running an external process
SCH05	Troubleshoot a DAG	Learn the fundamentals of troubleshooting a DAG.	Airflow	Tested running Airflow DAG
SCH06	Migrate an Existing DAG to Qubole Airflow	Migrate an existing schedule DAG to Qubole Airflow.	Airflow	Operating Airflow DAG on a Qubole Airflow cluster
SCH07	Integrate with External Airflow Schedule	Schedule a Qubole Command using Qubole Operator and an externally managed Airflow.	Airflow	Operating Airflow DAG on an external Airflow cluster

ACTIVATORS: DATA MANAGEMENT

Data Management Activators are key to the ingestion and exploration of your datasets. Select any one of these Activators to start importing, exporting and exploring your data.

REF	TITLE	OVERVIEW	OUTCOMES
DMT01	Configure a Datastore	Manage data sets through import and export procedures in Qubole.	Configured datastore
DMT02	Migrate Data to a Different Format	Prepare data for consumption by multiple users and applications.	Tested Qubole Hive code required to convert data to a new format
DMT03	Data Migration Across Cloud Providers	Work with an expert to migrate data from one cloud platform to another for Qubole consumption.	Data migrated to new cloud platform
DMT04	Import, Export and Query Data with Analyze Commands	Learn how to use Analyze to import, export and run queries against internal and external datasets.	Learn how to use Analyze and successfully import, export and query data
DMT05	Troubleshoot Analyze Commands	Troubleshoot issues with the import, export or query of data.	Restored dependent operations

ACTIVATORS: NOTEBOOKS AND DASHBOARDS

Notebook and Dashboard Activators provide an easy way to analyze, visualize and share data. Select any one of these Activators to explore this functionality and leverage the benefits of Notebooks and Dashboards.

REF	TITLE	OVERVIEW	OUTCOMES
NDS01	Use Qubole Notebooks	Ingest, analyze, visualize and share data through the use of notebooks.	Gain insights into your data with a configured Notebook
NDS02	Configure a Notebook Interpreter	Use a specific language/data processing backend in the Notebook.	Provide enhanced functionality with a Spark Notebook interpreter configured for a specific use case
NDS03	Migrate a Jupyter Notebook	Assist with the migration of a local Jupyter notebook to a Qubole Spark notebook.	Configured Spark Notebook running code from a local Jupyter Notebook
NDS04	Implement a Local Jupyter Notebook	Configure a local Jupyter notebook with Livy and Sparkmagic.	Access to Qubole Spark clusters through a local Jupyter Notebook
NDS06	Use Qubole Dashboards	Share data and visualizations with non-Qubole and Qubole users.	Shared access to data and visualizations from a Notebook

ACTIVATORS: DATA SCIENCE

Data Science Activators provide access to the latest machine learning features and functionality within Qubole. Select one of these Activators to supercharge your data analytics.

REF	TITLE	OVERVIEW	OUTCOMES
DSC01	Compile a Machine Learning Framework (non-GPU)	Enable a single non-GPU machine learning framework on one cluster.	Machine Learning code running in a Notebook on a Spark Cluster without GPU
DSC02	Enable GPU Support for a Machine Learning Framework	Enable GPU support for a single machine learning framework on a single cluster.	Machine Learning code running in a Notebook on a Spark Cluster with GPU
DSC03	Optimize Machine Learning Training Spark Application for Performance	Optimize Machine Learning Training Spark Application for Performance.	Tested and optimized Spark code (performance improvement not guaranteed)
DSC04	Select and Train Machine Learning Models	Validate existing machine learning model was properly selected and trained	Recommend model or training procedures for selected Machine Learning Notebook
DSC05	Migrate Local Machine Learning Model to a Qubole Spark Cluster	Migrate a ML Model and provide best practices on ML in a distributed compute environment.	Migrated local model into Qubole environment
DSC06	Migrate Local Data Preparation to a Qubole Spark Cluster	Familiarize with using Spark and distributed computing.	Local data preparation process running in the cloud
DSC07	Select and Configure Access to a Supplementary Dataset	Select and configure access to an open source supplementary dataset to assist in a data science project.	Access to open source dataset from within a Machine Learning model running in Qubole

ACTIVATORS: INTEGRATIONS

Integration Activators offers expanded functionality for big data users by partnering with other solutions. Select any one of these Activators to speed up data access or run advanced data transformation and machine learning on data warehouses.

REF	TITLE	OVERVIEW	PARTNER	OUTCOMES
INT01	Integrate with Snowflake	Use the Qubole UI, API, and notebooks to read and write data to the Snowflake data warehouse by using the Qubole Dataframe API for Apache Spark.	Snowflake	Configured Snowflake Data Store and tested Spark code utilizing Snowflake integration
INT02	Integrate with Datadog	Monitor, troubleshoot, and optimize QDS clusters with Datadog.	Datadog	Datadog access to Qubole cluster metrics
INT03	Integrate with Tableau	Deliver faster analytics, visualization and business intelligence leveraging Big Data stored in the cloud.	Tableau	Expanded analytic capabilities and Tableau access to data through a Qubole cluster
INT04	Integrate with Looker	Implement self-service data exploration and visualization without having to rely on a Data Science team. Looker makes the data in Qubole available to everyone across an organization regardless of technical ability for data exploration, visualization, or sharing.	Looker	Expanded analytic capabilities and Looker access to data through a Qubole cluster
INT05	Integrate with Apache Kafka	This Activator will assist with integrating Spark Qubole Cluster with Kafka.	-	Real-time ingestion of data from Kafka.
INT06	Integrate with Amazon Kinesis	This Activator will assist with integrating Spark Qubole Cluster with Kinesis.	-	Real-time ingestion of data from Kinesis.
INT07	Integrate with Talend	Learn how to use Talend Studio and Qubole to integrate data from various sources into the cloud data lake, and build the data quality workflows to clean, mask, and transform data as per the business requirements.	Talend	Talend configured to utilize Qubole clusters and access data

ACTIVATORS: MISCELLANEOUS

Miscellaneous Activators cover a broad range of topics and are designed to cover our ad-hoc user needs.

REF	TITLE	OVERVIEW	OUTCOMES
MSC01	Qubole Ask Me Anything (AMA)	Gain access to Qubole subject matter experts over the course of four two-hour sessions.	Four completed AMA sessions.



Ready to get started with the activation of your data.

Please contact your account team or email us at sales@qubole.com to learn more and get started with Activators.

Corporate Headquarters

469 El Camino Real
Suite 205
Santa Clara, CA 95050
(855) 423-6674
www.qubole.com

About Qubole

Qubole is revolutionizing the way companies activate their data—the process of putting data into active use across their organizations. With Qubole’s cloud-native Big Data Activation Platform, companies exponentially activate petabytes of data faster, for everyone and any use case, while continuously lowering costs. Qubole overcomes the challenges of expanding users, use cases, and variety and volume of data while constrained by limited budgets and a global shortage of big data skills. Qubole’s intelligent automation and self-service supercharge productivity, while workload-aware auto-scaling and real-time spot buying drive down compute costs dramatically. Qubole’s platform delivers freedom of choice, eliminating legacy lock in—use any engine and any tool to match your company’s needs.

