



Apache DistributedLog (incubating)

Apache DistributedLog (DL) is a *high-throughput, low-latency* replicated log service, offering *durability, replication* and *strong consistency* as essentials for building reliable *real-time* applications.

Status

Apache DistributedLog project graduated from Incubator at July 2017. It is now a sub-project of Apache BookKeeper.

The core components of *Apache DistributedLog* has been merged as part of *Apache BookKeeper*. The development of *Apache DistributedLog* has been moved under BookKeeper. See BP-26: Move distributedlog library as part of bookkeeper for more details.

Features

High Performance DL is able to provide *milliseconds* latency on *durable* writes with a large number of concurrent logs, and handle high volume reads and writes per second from thousands of clients.

Durable and Consistent Messages are persisted on disk and replicated to store multiple copies to prevent data loss. They are guaranteed to be consistent among writers and readers in terms of *strict ordering*.

Efficient Fan-in and Fan-out DL provides an efficient service layer that is optimized for running in a multi-tenant datacenter environment such as *Mesos* or *Yarn*. The service layer is able to support large scale writes (fan-in) and reads (fan-out).

Various Workloads DL supports various workloads from *latency-sensitive* online transaction processing (OLTP) applications (e.g. WAL for distributed database and in-memory replicated state machines), real-time stream ingestion and computing, to analytical processing.

Multi Tenant To support a large number of logs for multi-tenants, DL is designed for I/O isolation in real-world workloads.

Layered Architecture DL has a modern layered architecture design, which separates the *stateless service tier* from the *stateful storage tier*. To support large scale writes (fan-in) and reads (fan-out), DL allows scaling storage independent of scaling CPU and memory.

First Steps

- *Concepts*: Start with the basic concepts of DistributedLog. This will help you to fully understand the other parts of the documentation, including setup, integration and operation guide. It is highly recommended to read this first.
- *Quickstarts*: Run DistributedLog on your local machine or follow the tutorial to write a simple program to interact with *DistributedLog*.
- *Setup*: The docker and cluster setup guides show how to deploy DistributedLog stack.
- *User Guide*: You can checkout our guides about the basic concepts and the Core Library API or Proxy Client API to learn how to use DistributedLog to build your reliable real-time services.

Next Steps

- *Design Documents*: Learn about the architecture, design considerations and implementation details of DistributedLog.
- *Tutorials*: You can check out the tutorials on how to build real applications.
- *Admin Guide*: You can check out our guides about how to operate the DistributedLog Stack.

Get In Touch

Report a Bug

For filing bugs, suggesting improvements, or requesting new features, help us out by opening a jira.

Need Help?

Subscribe or mail the dev@distributedlog.incubator.apache.org list - Ask questions, find answers, join development discussions and also help other users.

Contributing

We feel that a welcoming open community is important and welcome contributions.

Contributing Code

1. See Developer Guide to get your local environment setup.
2. Take a look at our open issues.
3. Review our coding style and follow our code reviews to learn about our conventions.
4. Make your changes according to our code review workflow.
5. Checkout the list of project ideas to contribute more features or improvements.

Improving Website and Documentation

1. See [website/README.md](#) on how to build the website.
2. See [docs/README.md](#) on how to build the documentation.

About

Apache DistributedLog is an open source project of The Apache Software Foundation (ASF). The Apache DistributedLog project originated from Twitter.