KIP-827: Expose log dirs total and usable space via Kafka API

Status

Current state: Accepted

Discussion thread: here

Voting thread: here

JIRA: KAFKA-13958

Please keep the discussion on the mailing list rather than commenting on the wiki (wiki discussions get unwieldy fast).

Motivation

Storage is one of the key resources in a Kafka cluster. Administrators typically monitor the disk usage of each log directory via metrics to enable them to properly manage the storage attached to brokers. Metrics provide an easy way to see trends and set alerts and administrators should always use them to monitor disk usage.

There are also use cases where metrics are not a good way to retrieve the disk usage. For example, in tooling and automation, it would be useful to also be able to retrieve disk capacity and usable space directly. That would allow to easily validate whether disk operations (like a resize), or topic deletion (log deletion only happen after a short delay) have completed. For that reason this KIP proposes exposing disk total and usable sizes via the Kafka API.

Public Interfaces

We already have the DescribeLogDirs API that returns logdirs and details about the replicas they contain. To expose logdirs total and usable space, this KIP proposes adding 2 new fields to the DescribeLogDirsResponse message and bumping its protocol version to 4. The LogDirDescription class will also be updated to expose these 2 new fields to the Admin API.

Proposed Changes

DescribeLogDirs v4

No changes in the Request. Two new fields are added to the Response: TotalBytes and UsableBytes

```
{
  "apiKey": 35,
  "type": "response",
  "name": "DescribeLogDirsResponse",
  // Starting in version 1, on quota violation, brokers send out responses before throttling.
  "validVersions": "0-4",
  // Version 2 is the first flexible version.
  // Version 3 adds the top-level ErrorCode field
  // Version 4 adds the TotalSpace and UsableSpace fields
  "flexibleVersions": "2+",
  "fields": [
    { "name": "ThrottleTimeMs", "type": "int32", "versions": "0+",
      "about": "The duration in milliseconds for which the request was throttled due to a quota violation, or
zero if the request did not violate any quota." },
    { "name": "ErrorCode", "type": "int16", "versions": "3+", "about": "The error code, or 0 if there was no
error." },
    { "name": "Results", "type": "[]DescribeLogDirsResult", "versions": "0+",
      "about": "The log directories.", "fields": [
      { "name": "ErrorCode", "type": "int16", "versions": "0+",
        "about": "The error code, or 0 if there was no error." },
      { "name": "LogDir", "type": "string", "versions": "0+",
        "about": "The absolute log directory path." },
      { "name": "Topics", "type": "[]DescribeLogDirsTopic", "versions": "0+",
        "about": "Each topic.", "fields": [
          ...
        ]}
      ]},
      { "name": "TotalBytes", "type": "int64", "versions": "4+", "ignorable": true, "default": "-1",
        "about": "The total size in bytes of the volume the log directory is in."
      },
      { "name": "UsableBytes", "type": "int64", "versions": "4+", "ignorable": true, "default": "-1",
        "about": "The usable size in bytes of the volume the log directory is in."
      }
    ]}
  ]
}
```

ReplicaManager

When handling a DescribeLogDirs request, ReplicaManager will retrieve the total and usable space in bytes of the volume each logdir is into. In case these sizes are larger than Long.MAX_VALUE, (see https://bugs.openjdk.java.net/browse/JDK-8162520), brokers will return Long.MAX_VALUE.

LogDirDescription

LogDirDescription is used by the Admin API to represent the results from describeLogDirs(). Two new methods are added to this type:

```
/**
 * Returns the total size in bytes of the volume the log directory is into. The optional will be empty if the
broker does not support this feature or if an error happened accessing the log directory (see the error field).
 */
public OptionalLong totalBytes()

/**
 * Returns the currently usable size in bytes of volume the log directory is into. The optional will be empty
if the broker does not support this feature or if an error happened accessing the log directory (see the error
field).
 */
public OptionalLong usableBytes()
```

If multiple log directories are on the same actual volume, they will both return the sizes of that volume.

Compatibility, Deprecation, and Migration Plan

Only new clients will use the new version, this does not change the behavior of existing clients.

Rejected Alternatives

None