KIP-61: Add a log retention parameter for maximum disk space usage percentage

- Status
- Motivation
- Public Interfaces
- Proposed Changes
- Compatibility, Deprecation, and Migration Plan
- Rejected Alternatives

Status

Current state: Under Discussion

Discussion thread: *here [Change the link from the KIP proposal email archive to your own email thread]*

JIRA: KAFKA-1489

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

Motivation

This KIP is related to KAFKA-1489: Global threshold on data retention size.

In dynamic situations where topics are added in unpredictable ways, the existing log retention parameters are not sufficient to prevent out-of-disk conditions from occurring. Consider, for example, a Kafka cluster that serves the needs of a team of developers working on new Kafka-based services. Such users may begin pushing arbitrary amounts of content into new topics at any time.

The existing log retention parameters reflect a frame of reference in which topic usage is predictable, and an administrator can make reasonable projections and choose configuration values based on provisioned disk capacity.

The alternative frame of reference reflected in this proposal includes the following assumptions and priorities:

- Topic usage (number of topics, topic volume) is unpredictable.
- Availability of the latest content for all topics is more important than specific time- and size-based limits per topic.
- The disks should never ever run out of free space.
- Periodically deleting the globally oldest segments is a reasonable strategy to prevent the disks from running out of free space.

Public Interfaces

We propose adding a new broker configuration option:

```
val LogRetentionDiskUsagePercentProp = "log.retention.disk.usage.percent"
val LogRetentionDiskUsagePercentDoc = "The maximum percentage of disk space allowed to be in use
    (per-disk). Deletes oldest segments (across all topics) to maintain this usage ceiling."
```

The default value is 100, which effectively disables the feature.

Proposed Changes

Add a log retention parameter that sets a soft upper limit on the percentage of disk space that can be in use. After the other log retention policies have been applied in each cleanup cycle, compute the amount of disk space in use on each physical device. For each device that is over the configured usage limit, compute the number of "excess bytes" to free and delete as many of the globally oldest segments as is necessary to reach this goal.

Compatibility, Deprecation, and Migration Plan

There are no known migration issues.

The feature is disabled by default. If enabled, it may supersede the following time-based log retention parameters (just like the existing size-based parameters):

- log.retention.hours
- log.retention.minutes
- log.retention.ms

Rejected Alternatives

The parameter could be expressed as "minimum percentage of disk space to keep free". It is expressed instead as "maximum percentage of disk space to allow to be used" because that seems more congruent with the other size-based log retention parameters, which are expressed as maximums.

The parameter could be expressed as a global maximum byte count. Ops monitoring tools are typically configured to trigger alerts based on disk usage percentages, however. Adopting the same units makes configuration slightly easier. It also allows for disk capacity to be increased without the need for reconfiguration, in some cases.

The parameter is defined as "percentage of disk space in use", not "percentage of disk space in use *by Kafka*". The latter definition would be somewhat more expensive to compute each cycle. Perhaps more importantly, it would weaken the guarantee that the parameter is designed to provide - that the disks will never ever run out of free space.