KIP-541: Create a fetch.max.bytes configuration for the broker

Status

Current state: Accepted

Discussion thread: *here*

JIRA: KAFKA-9101

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

Problem

Kafka consumers can choose the maximum number of bytes to fetch by setting the client-side configuration *fetch.max.bytes*. A high value for this configuration allows the client to fetch a lot of bytes at a time.

However, when this configuration value is too high, it may degrade performance on the broker for other consumers. The reason is because the broker will spend a lot of time on one very long fetch request, resulting in a situation that is less fair to the other consumers. Even worse, if the configuration value is set to an extremely high value, such as hundreds of megabytes, the client request may time out before being fulfilled.

Currently the Kafka broker has no way to put an upper limit on the maximum number of bytes that the client can choose to fetch. We would like to address this issue by adding a new configuration on the broker side to do just that.

Public Interfaces

There will be a new broker-side configuration, *fetch.max.bytes*. The effective maximum size of any fetch request will be the minimum of the maximum fetch size the client requests, and this value. The new value will be 55 megabytes by default.

Configuration Name	Type	Default Value	Importance
fetch.max.bytes	INT	55 * 1024 * 1024	HIGH

Fetch request from replicas will also be affected by the *fetch.max.bytes* limit.

Compatibility, Deprecation, and Migration Plan

Existing clients will continue to work, even if they have set a larger *fetch.max.bytes* than the one set on the server. They will simply receive a little less data than before. Clients must be prepared to handle receiving less than the maximum fetch size in any case.

Rejected Alternatives

Static fetch.max.bytes

We could put a static (unconfigurable) limit on *fetch.max.bytes* on the broker side. However, it's better to make this configuration, since system administrators may want to tune this based on their workloads and hardware.