package com.twitter.search.earlybird.common.userupdates;

import java.util.Date;

import java.util.concurrent.TimeUnit;

import com.twitter.common.util.Clock;

import com.twitter.decider.Decider;

import com.twitter.search.common.indexing.thriftjava.UserUpdateType;

import com.twitter.search.common.schema.earlybird.EarlybirdCluster;

import com.twitter.search.earlybird.common.config.EarlybirdConfig;

/\*\*

\* Contains logic for deciding whether to apply a certain user update to the {@link UserTable}.

\*/

public class UserUpdatesChecker {

private final Date antisocialStartDate;

private final Decider decider;

private final boolean isFullArchiveCluster;

public UserUpdatesChecker(Clock clock, Decider decider, EarlybirdCluster cluster) {

// How many days of antisocial users to keep. A value of -1 means keeping all user updates.

long antisocialRecordDays =

EarlybirdConfig.getLong("keep\_recent\_antisocial\_user\_updates\_days", 30);

this.antisocialStartDate = antisocialRecordDays > 0

? new Date(clock.nowMillis() - TimeUnit.DAYS.toMillis(antisocialRecordDays)) : null;

this.decider = decider;

this.isFullArchiveCluster = cluster == EarlybirdCluster.FULL\_ARCHIVE;

}

/\*\*

\* Decides whether to skip the given UserInfoUpdate.

\*/

public boolean skipUserUpdate(UserUpdate userUpdate) {

if (userUpdate == null) { // always skip null updates

return true;

}

UserUpdateType type = userUpdate.updateType;

if (type == UserUpdateType.PROTECTED && skipProtectedUserUpdate()) {

return true;

}

if (type == UserUpdateType.ANTISOCIAL && skipAntisocialUserUpdate(userUpdate)) {

return true;

}

// NSFW users can continue to tweet even after they are marked as NSFW. That means

// that the snapshot needs to have all NSFW users from the beginning of time. Hence, no NSFW

// users updates check here.

// pass all checks, do not skip this user update

return false;

}

// Antisocial/suspended users can't tweet after they are suspended. Thus if our index stores

// tweets from the last 10 days, and they were suspended 60 days ago, we don't need them since

// there will be no tweets from them. We can save space by not storing info about those users.

// (For archive, at rebuild time we filter out all suspended users tweets, so for a user that

// was suspended before a rebuild, no need to use space to store that the user is suspended)

private boolean skipAntisocialUserUpdate(UserUpdate userUpdate) {

return antisocialStartDate != null && userUpdate.getUpdatedAt().before(antisocialStartDate);

}

// skip protected user updates for realtime and protected clusters

private boolean skipProtectedUserUpdate() {

return !isFullArchiveCluster;

}

}