package com.twitter.search.core.earlybird.index.util;

import java.io.IOException;

import org.apache.lucene.index.LeafReader;

import org.apache.lucene.index.Terms;

import org.apache.lucene.index.TermsEnum;

import org.apache.lucene.search.DocIdSetIterator;

import org.apache.lucene.util.BytesRef;

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

import com.twitter.search.core.earlybird.index.EarlybirdRealtimeIndexSegmentAtomicReader;

/\*\*

\* Used to iterate through all of the documents in an Earlybird segment. This is necessary so that

\* we can ensure all of the documents we are reading have been published to the readers. If we used

\* the doc ID mapper to iterate through documents, it would return documents that have been only

\* partially added to the index, and could return bogus search results (SEARCH-27711).

\*/

public class AllDocsIterator extends DocIdSetIterator {

public static final String ALL\_DOCS\_TERM = "\_\_all\_docs";

private final DocIdSetIterator delegate;

public AllDocsIterator(LeafReader reader) throws IOException {

delegate = buildDISI(reader);

}

private static DocIdSetIterator buildDISI(LeafReader reader) throws IOException {

if (!isRealtimeUnoptimizedSegment(reader)) {

return all(reader.maxDoc());

}

Terms terms =

reader.terms(EarlybirdFieldConstants.EarlybirdFieldConstant.INTERNAL\_FIELD.getFieldName());

if (terms == null) {

return all(reader.maxDoc());

}

TermsEnum termsEnum = terms.iterator();

boolean hasTerm = termsEnum.seekExact(new BytesRef(ALL\_DOCS\_TERM));

if (hasTerm) {

return termsEnum.postings(null);

}

return empty();

}

@Override

public int docID() {

return delegate.docID();

}

@Override

public int nextDoc() throws IOException {

return delegate.nextDoc();

}

@Override

public int advance(int target) throws IOException {

return delegate.advance(target);

}

@Override

public long cost() {

return delegate.cost();

}

/\*\*

\* Returns whether this is a realtime segment in the realtime index that is still unoptimized and

\* mutable.

\*/

private static boolean isRealtimeUnoptimizedSegment(LeafReader reader) {

if (reader instanceof EarlybirdRealtimeIndexSegmentAtomicReader) {

EarlybirdRealtimeIndexSegmentAtomicReader realtimeReader =

(EarlybirdRealtimeIndexSegmentAtomicReader) reader;

return !realtimeReader.getSegmentData().isOptimized();

}

return false;

}

}