Katello - Bug #6005

Content reindexing Incredibly Inefficient

06/01/2014 03:26 PM - Justin Sherrill

Status:	Closed		
Priority:	Normal		
Assignee:	Justin Sherrill		
Category:	Repositories		
Target version:	Katello 2.0		
Difficulty:		**Fixed in Releases:**	
Triaged:	Yes	**Found in Releases:**	
Bugzilla link:	1112749	**Red Hat JIRA:**	
Pull request:			

Description
Currently when running 'rake katello:reindex' we loop over each repo and index packages, errata, etc.. on each repo. While this is simpler it is incredibly inefficient especially if you many content views with many versions. The same set of packages will end up being reindex over and over and over again. A more efficient way would be to just go through the entire $ITEM list in pulp and index that list once.

Associated revisions

Revision 2c7d6aae - 06/02/2014 04:35 PM - Justin Sherrill

Fixes #6005 - speeding up content reindexing

Instead of indexing per repo, we can index all of each content
and it will be much faster. Sample times with a content view with about
5 versions:

Old Reindex: 7m05s
New Reindex: 3m13s

the more content views and versions you have the greater the difference

Revision ede3a9b2 - 06/03/2014 01:28 PM - Justin Sherrill

Merge pull request #4190 from jlsherrill/index

Fixes #6005 - speeding up content reindexing

History

#1 - 06/02/2014 12:26 PM - Eric Helms

- Target version set to 45

- Triaged changed from No to Yes

#2 - 06/03/2014 01:52 PM - Justin Sherrill

- Status changed from Assigned to Closed

- % Done changed from 0 to 100

Applied in changeset katello|commit:2c7d6aaeff5fca84347bbd96fe9f8096eae29ef7.

#3 - 06/24/2014 03:41 PM - Eric Helms

- Bugzilla link set to https://bugzilla.redhat.com/show_bug.cgi?id=1112749

#4 - 08/22/2014 08:28 AM - Eric Helms

- translation missing: en.field_release set to 13