When a new peer zone is added, rgw starts in 'full sync'. This starts by listing all of that zone's bucket instances, in order to build a list of all of the bucket index shards we need to sync.

```c
/* FIXME: need a better scaling solution here, requires streaming output */
call(new RGWReadRESTResourceCR<list<string>>(store->ctx(), sync_env->conn, sync_env->http_manager,
         entrypoint, NULL, &result));
```

This attempts to read the entire list of bucket instance ids from the remote zone’s “/admin/metadata/bucket.instance” admin api. RGWOp_Metadata_List is the op for this admin api, and it will return a maximum of 1000 entries - so if there are more bucket instances than that, data sync won’t see them. RGWOp_Metadata_List accepts a “marker” parameter that allows you to resume a listing from where the last request left off. RGWListBucketIndexesCR should use this marker in order to continue listing keys until the end (where truncated = false).

**Related issues:**
- Copied to rgw - Backport #40352: nautilus: multisite: RGWListBucketIndexesCR ... Resolved
- Copied to rgw - Backport #40353: luminous mimic nautilus: RGWListBucketIndexesCR ... Rejected
- Copied to rgw - Backport #40354: mimic: multisite: RGWListBucketIndexesCR for... Resolved
#4 - 06/13/2019 09:02 PM - Nathan Cutler
- Copied to Backport #40352: nautilus: multisite: RGWListBucketIndexesCR for data full sync needs pagination added

#5 - 06/13/2019 09:02 PM - Nathan Cutler
- Copied to Backport #40353: luminous: multisite: RGWListBucketIndexesCR for data full sync needs pagination added

#6 - 06/13/2019 09:02 PM - Nathan Cutler
- Copied to Backport #40354: mimic: multisite: RGWListBucketIndexesCR for data full sync needs pagination added

#7 - 01/27/2021 07:11 PM - Nathan Cutler
- Status changed from Pending Backport to Resolved

While running with --resolve-parent, the script "backport-create-issue" noticed that all backports of this issue are in status "Resolved" or "Rejected".