Metadata Architecture for ESG, Version 3

Ann Chervenak

1. Underlying Assumptions for the Metadata Architecture

- Gateways: there will be three Gateway located at NCAR, LLNL and ORNL, that will provide:
  - rich ESG services
  - sophisticated administrative support
  - We may deploy additional gateways in the future
- Current architecture: We agreed that metadata will initially be published to a tier1/global master metadata catalog service; this catalog may then be replicated to each gateway
- Gateways must continue to function even when other gateways fail.
- Metadata queries made at any gateway should provide the ability to discover ALL datasets available in ESG
  - We also require the ability to restrict searches at a gateway
- We will provide highly-available, high-level services for:
  - Querying ESG data sets (highest availability requirement)
  - Harvesting ESG metadata (publishing operation needs to be atomic)
- Where these services are deployed (tier1/Global or tier 2/gateway) remains to be determined
- Synchronization requirements for distributed databases are less than one day (one hour, if possible)
- Data nodes will produce data sets, will be responsible for authoring at least some of the metadata for those data sets, and they may use ESG services to assist with this authoring.
- Data nodes are considered the authoritative source of the data and the associated metadata
- Metadata produced at data nodes will need to be copied to the metadata catalogs hosted at tier 1 metadata service.
- The group decided against using an architecture that included dynamic referrals (where metadata queries at one gateway are forwarded to other gateways).
  - Reasons: performance may suffer; if any gateway is down, the query results will be incomplete, which is not acceptable.

2. Open Issues for the Metadata Architecture

A number of other issues have yet to be decided:

- We don't know whether metadata harvesting will use push, pull or subscription/notification mechanisms to move metadata between data nodes and the tier1 metadata service.
- We don't know whether the data nodes will host local metadata catalogs. The data nodes may just author metadata and store it temporarily until the metadata is harvested.
- We will handle the different types of metadata differently (discovery-level, user-level, security-level, etc.), but the details of this are yet to be determined.
- We are still discussing whether a centralized metadata service would be sufficient.

3. Centralized and partitioned metadata catalogs

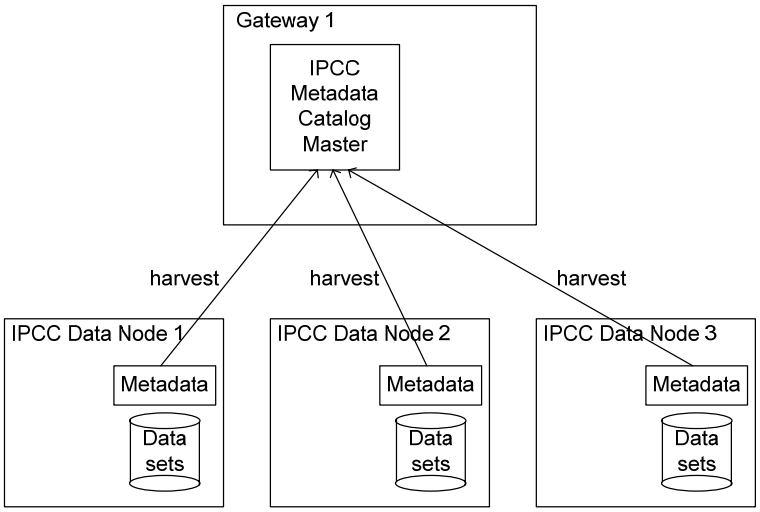
One option for ESG is to simply have a centralized metadata catalog at the tier1/global services level that can provide a high level of availability for the catalog. All metadata updates would go to the central catalog, and all metadata queries would be directed to the central catalog.

A centralized metadata catalog would be the simplest to implement. However, the central catalog represents a single point of failure. If the site hosting the centralized catalog becomes unavailable, then users are unable to query or update the metadata catalog at other gateways.

Based on the requirements we specified in the ESG proposal, a centralized implementation is probably not desirable.

4. Metadata Harvesting

We have not yet determined how metadata will be generated at the data nodes and then sent to the Gateway metadata catalogs. The figure below shows one option for metadata harvesting. In this scenario, the data nodes generate metadata for the data sets that they publish and store this metadata in a local catalog. Then the data node periodically sends, or "pushes", its metadata to a gateway node.



Other options include the possibility of the gateway node "pulling" metadata in from the data nodes or the use of a publish/subscribe mechanism for harvesting.

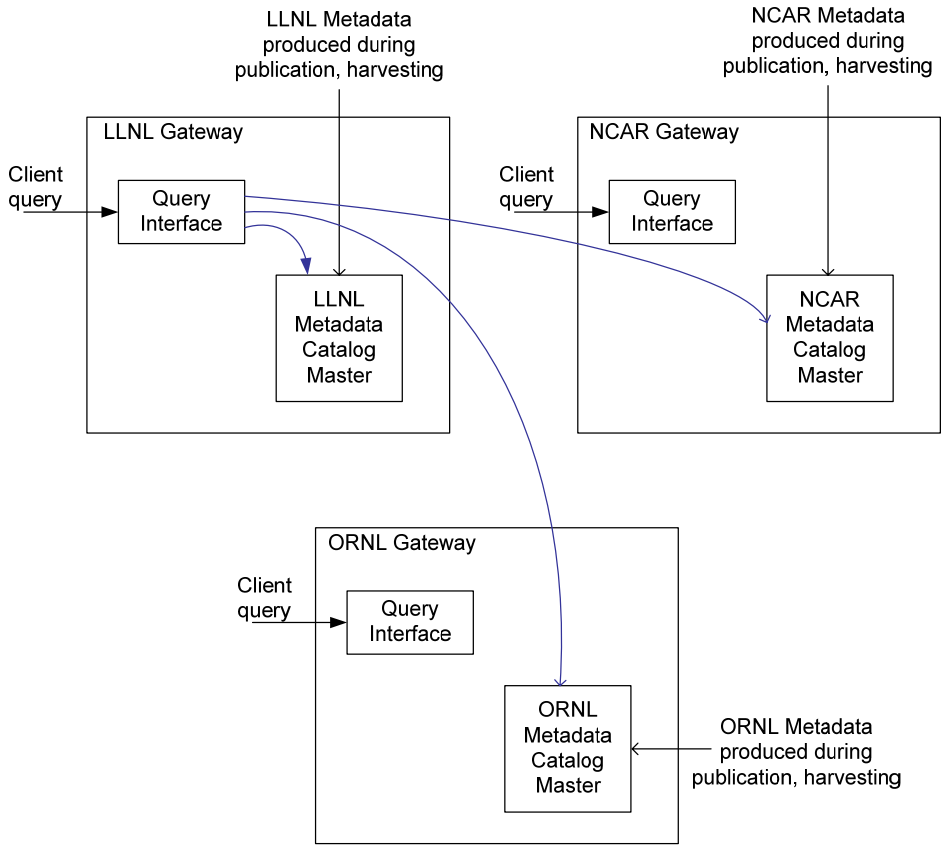
5. Rejected Architecture: A Partitioned Catalog with Dynamic Referrals

An architecture that we rejected for ESG is a simple partitioned deployment that would allow each of the Gateway nodes to maintain a separate Metadata catalog. A query arriving at a particular gateway could either be answered by the local metadata catalog or referred dynamically to another gateway's catalog. This would be simple to implement. One tradeoff of this architecture is that, since the metadata is not replicated at each gateway, whenever a gateway Metadata catalog is unavailable, it will not be possible to forward a query to that catalog, so the user will not be able to issue queries for those ESG datasets.

Because incomplete results for an ESG metadata query are unacceptable, we rejected this architecture option. Another disadvantage of an architecture using dynamic referrals is that queries could potentially be much slower than in a system where metadata catalogs are replicated at each gateway,

since queries and results would have to be sent across wide area links and results from multiple gateways would have to be combined.

The figure below shows how queries could be dynamically forwarded to metadata catalogs at other gateways.



6. A Replicated Metadata Catalog using Master-Slave Replication

The architecture that we agreed upon envisions a single master metadata catalog that is hosted at the Tier1 or Global services layer of the ESG architecture. In practice, this service may be deployed at a particular gateway, but logically, the metadata master catalog is independent of any gateway. All updates to metadata must be performed on this master catalog.

In addition, the contents of the master metadata catalog will be replicated periodically to each of the ESG gateway nodes. This allows users to issue metadata queries at any gateway.

There will be some delay between when updates are made at the master metadata catalog and when these updates are propagated to the metadata catalog replicas at each gateway. If users require certainty that they are querying the most up-to-date metadata information, they can query the Master catalog. However, having multiple replicated catalogs that can answer user queries can provide load balancing so that the Master catalog does not become a performance bottleneck. The application

requirement for ESG is that updates at the master catalog are propagated to the replicated catalogs within one day. Our goal is that these updates happen sooner, typically within one hour.

If the master catalog is unavailable, then no metadata update operations can be performed. However, metadata queries can continue to be satisfied at any gateway using the replicated metadata catalogs.

Below, we illustrate the architecture in which a master metadata catalog at the tier1/global services level is periodically replicated to metadata catalogs at each gateway.