ALTO Use Case: Resource Orchestration for Multi-Domain, Geo-Distributed Data Analytics

draft-xiang-alto-multidomain-analytics-01

Qiao Xiang[1,2], Franck Le[3], Y. Richard Yang[1,2],
Harvey Newman[4], Haizhou Du[1], J. Jensen Zhang[1]

[1] Tongji University, [2] Yale University,
[3] IBM Watson Research Center,
[4] California Institute of Technology

March 19, 2018, IETF 101 ALTO

Takeaway from IETF 100 Interim

- Unicorn design.
  - Three-phase resource discovery, i.e., storage/computation resource discovery, path discovery and networking resource discovery.
- Unicorn development.
  - Demonstrated at SuperComputing 2017 in Nov. 2017.

Updates for IETF 101

- Substantial updates since version-00.

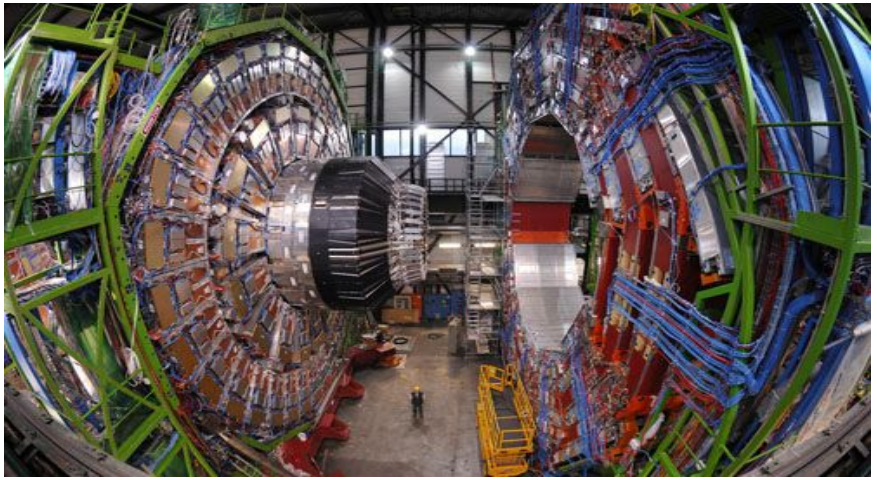
- Goals:
  - Better prepare for document review.
  - Propose an ALTO extension to support accurate, privacy-preserving resource discovery across multiple domains.

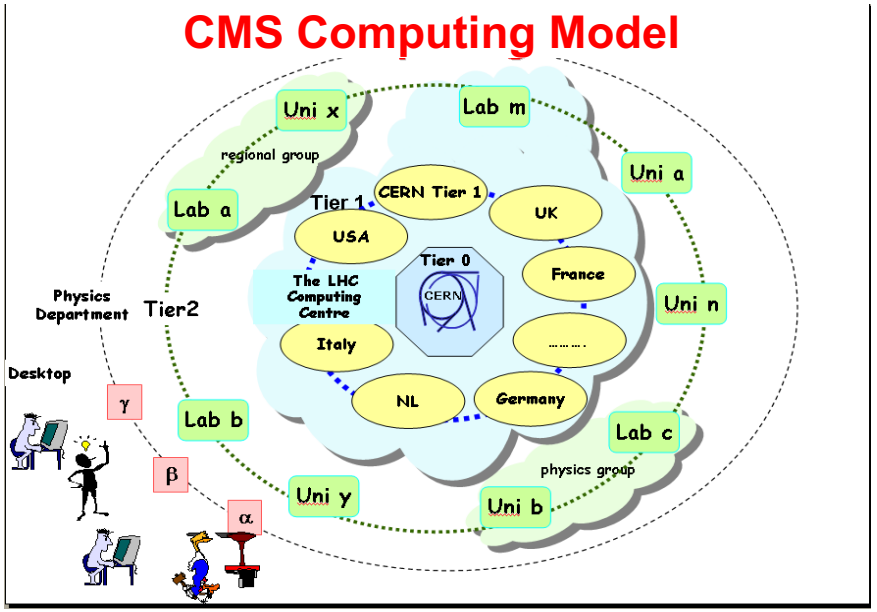
Updates for IETF 101

- Key updates:
  - Add an overview of the characteristics of multi-domain, geo-distributed data analytics.
  - Add the design requirements of the resource orchestration for multi-domain, geo-distributed data analytics.
  - Add a review of existing resource orchestration system designs for data analytics systems.
  - Update the motivation of using ALTO as the key information model in Unicorn.
  - Update the architecture of the Unicorn and the three-phase resource discovery in Unicorn.
  - Design an ALTO extension for privacy-preserving interdomain resource information aggregation.

Recap: Multi-Domain, Geo-Distributed Data Analytics

- **Settings**: Different organizations contribute various resources (e.g. , sensing, computation, storage and networking resources) to collaboratively collect, share and analyze extremely large amounts of data.
  - Example: the CMS experiment, coalitions between different organizations, cloud exchange, etc.





Characteristics of Multi-Domain, Geo-Distributed Data Analytics

- Dynamic Data Analytics Workload
  - Highly dynamic, in terms of the number of users, the types of applications, the number of jobs, the decomposition of jobs and the resource requirements of tasks.
- Dynamic Resource Availability
  - Each member network provides different types of resources with different amounts.
  - Many member networks are interconnected with high bandwidth-delay-product links.

Design Requirements

- Users' perspective
  - REQ1: Provide performance predictability for data analytics jobs.
  - REQ2: Achieve the efficient resource sharing among data analytics jobs.
- Member networks' perspective
  - REQ3: Achieve the high utilization of different types of resources in member networks.
  - REQ4: Maintain the autonomy and privacy of member networks.
  - REQ5: Provide compatibility with different data analytics applications and resource management systems to maximize the deployment.

Review of Existing Designs

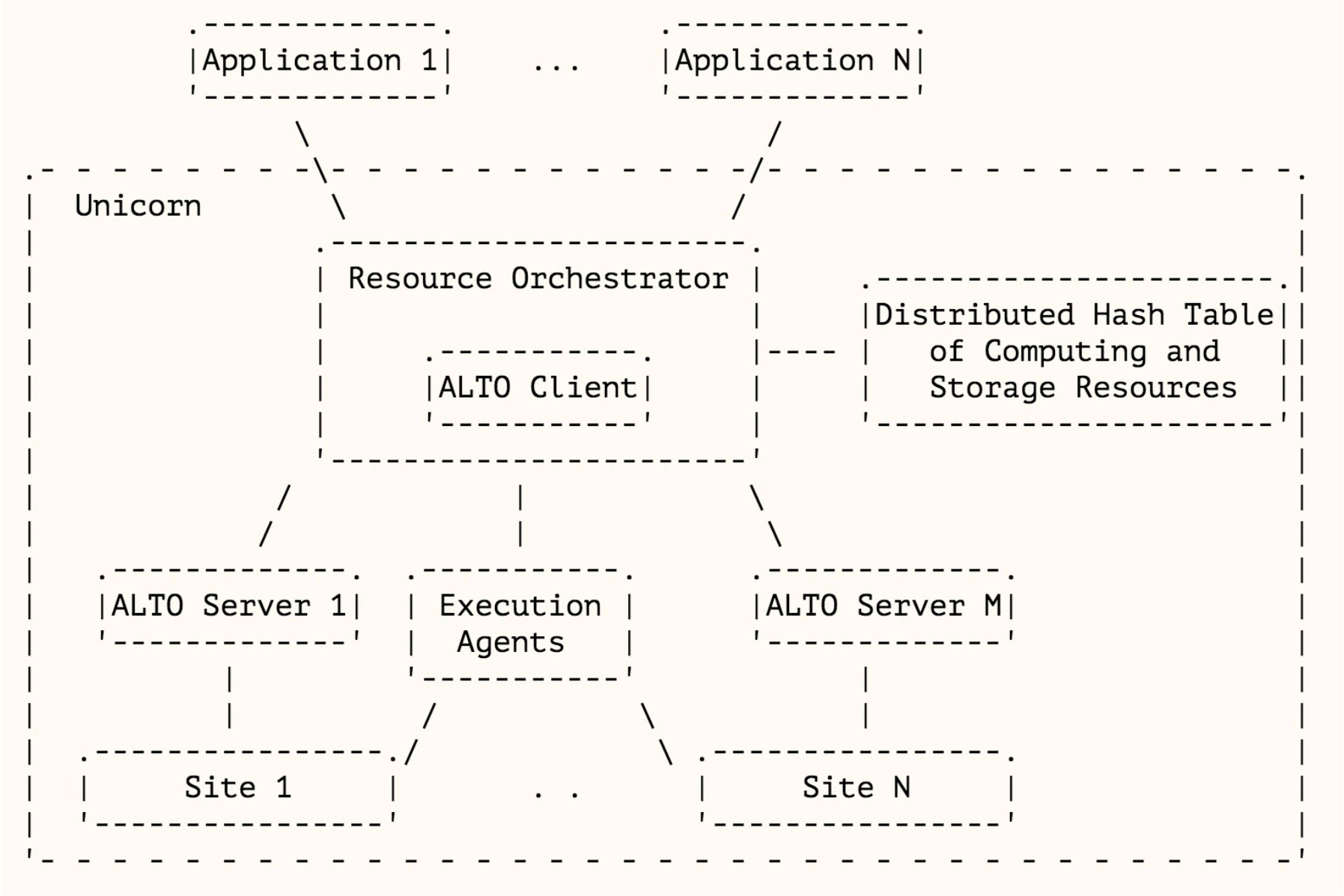
- Centralized resource-graph-based orchestration
  - Examples: Mesos, Borg and etc.
- Centralized ClassAds-based orchestration
  - Examples: HTCondor.
- Distributed opportunistic orchestration
  - Examples: Sparrow, Apollo and etc.

- Existing designs do not satisfy REQ1-4 due to the lack of an information model to support the accurate, yet privacy-preserving resource discovery across different member networks.

Fundamental Design Decision: Choose ALTO as the Resource Information Model

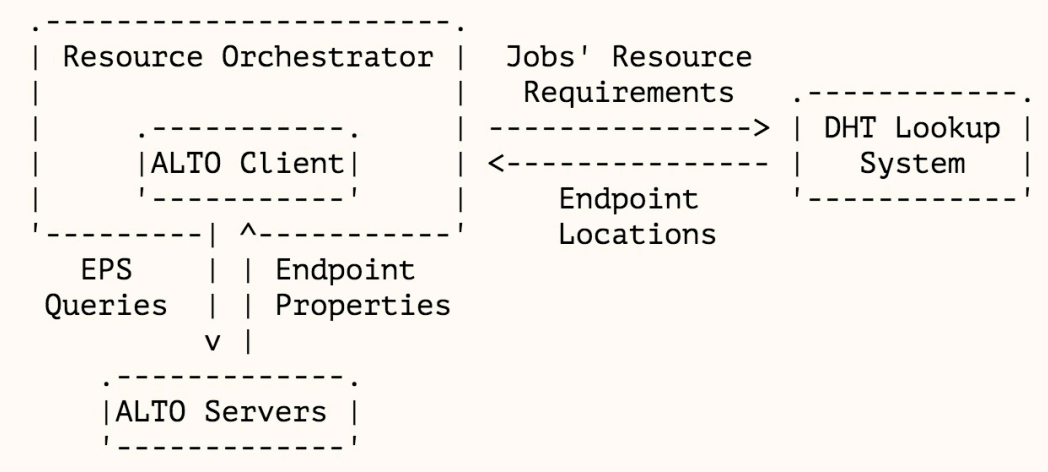
- **Reasons**
  - The use of different abstract maps in ALTO supports member networks to provide accurate information on different types of resources, e.g., the endpoint-property service.
  - The ALTO abstract maps provide a simplified view of member networks' resources, protecting the private information of networks, e.g., the network map.
  - Applications can use ALTO clients to accurately describe their requirements of different types of resources, e.g., the multi-cost service.

Architecture of Unicorn

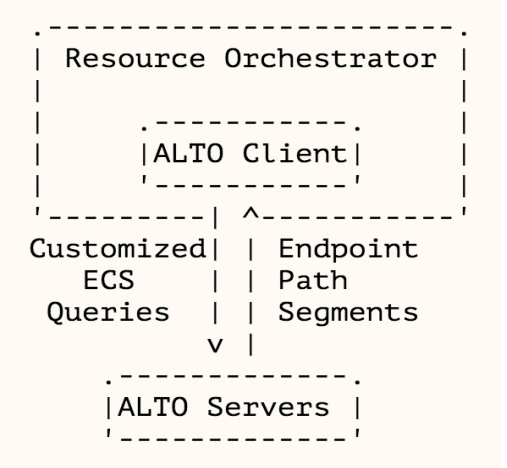


Three-Phase Resource Discovery

- Phase 1: Endpoint Property Discovery
  - Discover the locations and properties of computing and storage resources via ALTO EPS service.

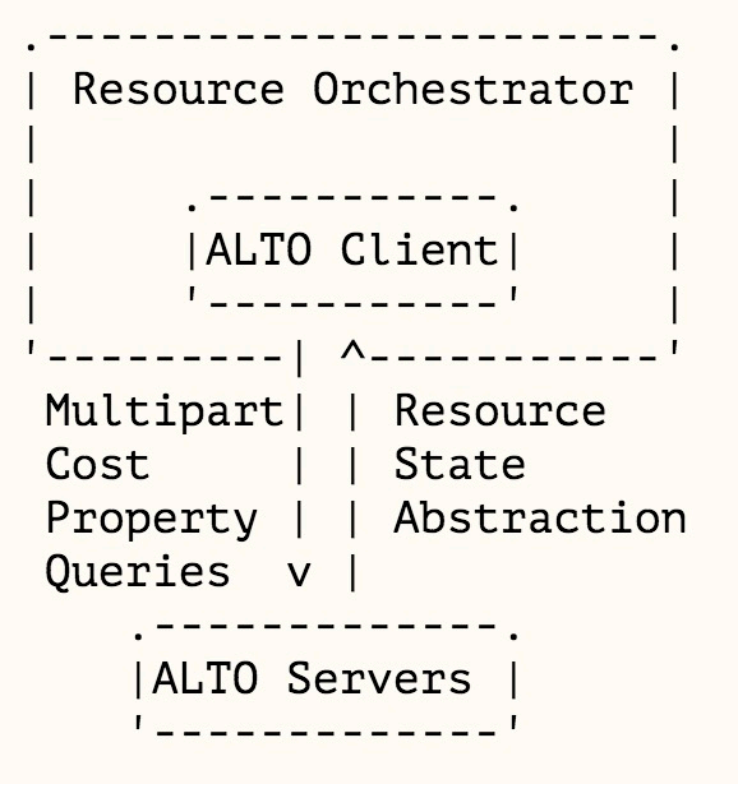


- Phase 2: Endpoint Path Discovery
  - Discover the connectivity between computing and storage resources via network map and ECS service.



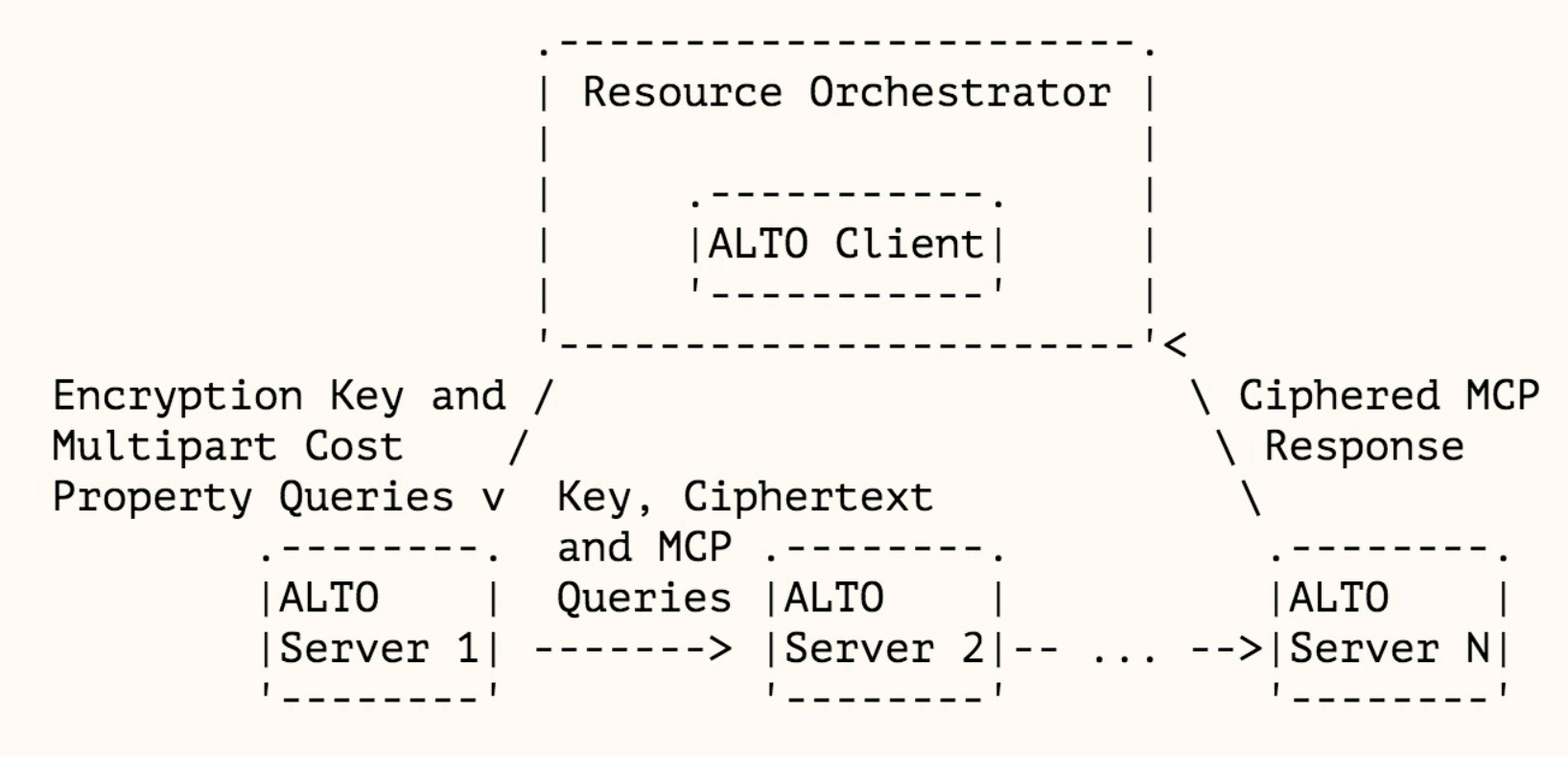
Three-Phase Resource Discovery

- Phase 3: Resource State Abstraction Discovery
  - Discover the networking resource sharing between flows via ALTO multipart cost property (MCP) service.
  - **Option 1**: Each ATLO server independently sends the responses to the ALTO client.
    - **Drawback**: expose the private capacity region of each network.



Three-Phase Resource Discovery

- Phase 3: Resource State Abstraction Discovery
  - Discover the networking resource sharing between flows via multipart cost property service.
  - **Option 2**: an ALTO-extension for privacy-preserving interdomain resource information aggregation (see the detailed algorithm in the draft), which returns the **intersected** capacity region of all networks.



Summary

- **Importance to the ALTO WG**:
  - Unicorn provides a generic design for large-scale, multi-domain data center resource optimization, a major use case of ALTO listed in the WG Charter.
  - The implementation and deployment experience of Unicorn provides practice guidelines for the use of multiple ALTO services.
- **Next steps**:
  - Full paper submission to SuperComputing 2018 by March 28.
  - Large-scale demonstration and deployment trials by IETF 103.