

# LIDC: A Location Independent Multi-Cluster Computing Framework for Data Intensive Science

Sankalpa Timilsina, Susmit Shannigrahi  
Tennessee Technological University

# Introduction

**Context:** Increasing use of distributed computing platforms in scientific research

**Challenges:** Current compute placement relies on centralized controllers like Kubernetes (K8s), which aren't ideal for multi-organization collaborations

- Requires manual configurations
- Relies on a single control entity
- Difficulty adapting to dynamic, distributed infrastructure

**Goal:** Overcome limitations of centralized control, adapt to dynamic infrastructure changes

# Proposed Solution - LIDC Framework

**Overview:** LIDC (Location Independent Data and Compute) introduces a **decentralized control plane in the network layer**

- Named Data Networking (NDN) names assigned to data and services to direct computations across clusters

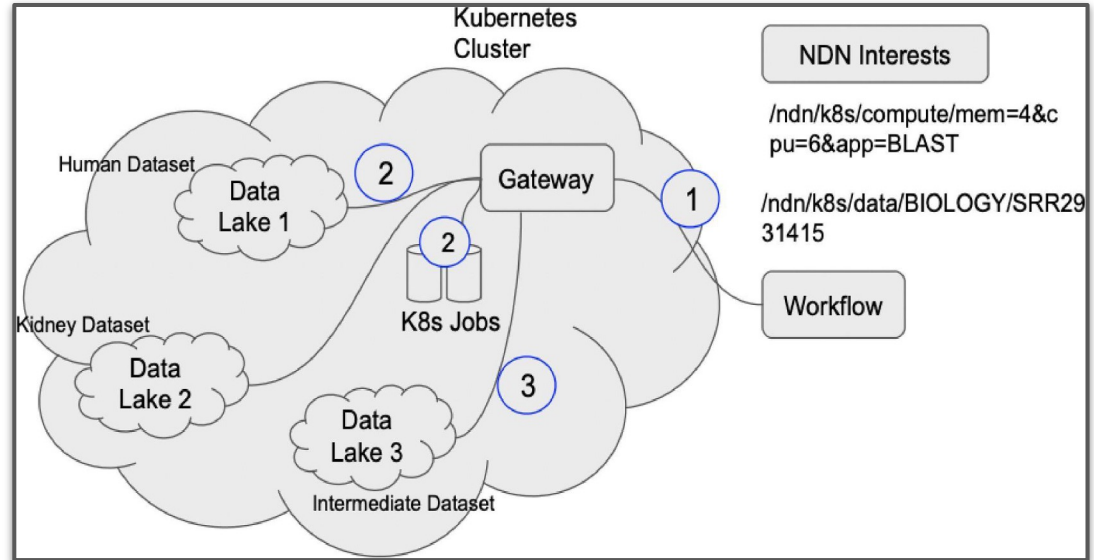
## **Benefits:**

- Location-independent compute placement
- Dynamic, real-time adaptability

# Methodology

**Concept:** Assign semantic names that describe job types and resource requirements

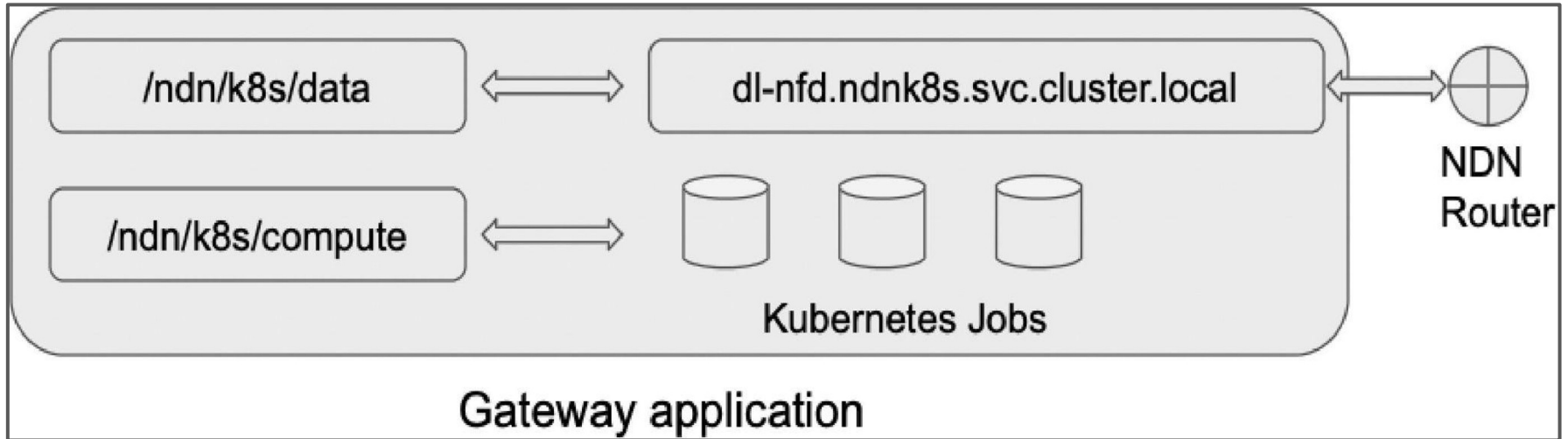
**Function:** Matches requests to named Kubernetes service endpoints, enabling cluster-agnostic compute



## Methodology (Contd..)

**Integration:** Combines NDN's routing with Kubernetes' orchestration capabilities

**Outcome:** Flexible, seamless end-to-end job placement across clusters



# Results and Conclusion

TABLE I  
COMPUTATION PERFORMANCE

SRR_ID	Ref. Database	Genome Type	Memory (GB)	CPU	Run Time	Output Size
SRR2931415	HUMAN	RICE	4	2	8h9m50s	941MB
SRR2931415	HUMAN	RICE	4	4	8h7m10s	941MB
SRR5139395	HUMAN	KIDNEY	4	2	24h16m12s	2.71GB
SRR5139395	HUMAN	KIDNEY	6	2	24h2m47s	2.71GB

**Decentralized control plane** with semantic names (NDN) with K8s.

Enables **dynamic, location-independent compute placement** across distributed clusters