UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

# 4CeeD: Real-Time Data Acquisition and Analysis Framework for Material-related Cyber-Physical Environments

Phuong Nguyen, Steven Konstanty, Todd Nicholson, Thomas O'brien, Aaron Schwartz-Duval, Timothy Spila, Klara Nahrstedt, Roy H. Campbell, Indranil Gupta, Michael Chan, Kenton McHenry, Normand Paquin

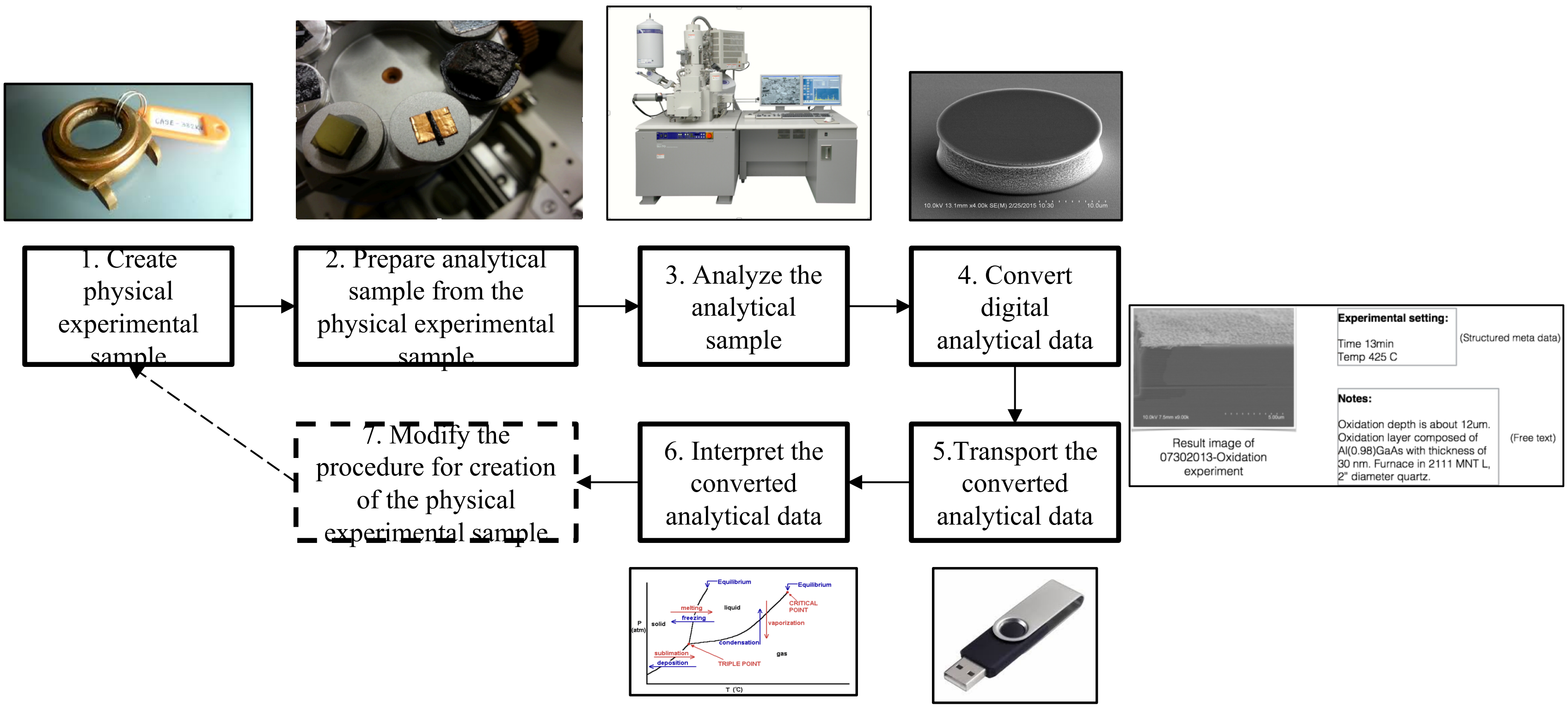
University of Illinois at Urbana-Champaign

# Outline

- **Background & Challenges**
- Our solution: 4CeeD
  - Curation Service
  - Coordination Service
- Evaluation
- Conclusions

# Background: Long process from material discovery to device fabrication

- It typically takes 20 years to go from the discovery of new materials to fabrication of new and next-generation devices*




* Holdren, J.P. "*Materials genome initiative for global competitiveness*". National Science and Technology Council OSTP 2011.

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

# Background: Issues with current experimental cycle

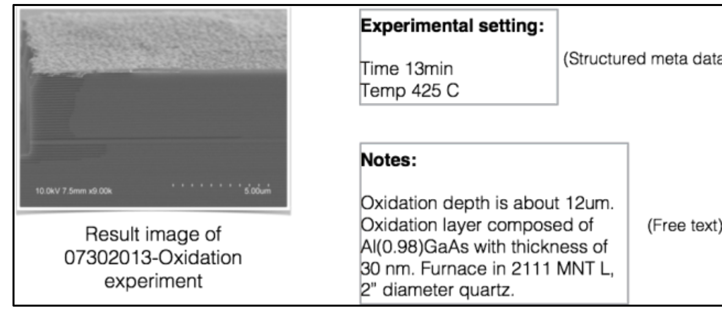
- Data acquisition & transfer
  - *"Sneaker-net"* data transfer
  - No data conversion is available during data acquisition

- Data management & processing
  - Using local file system, or general-purpose cloud storage services (e.g., DropBox, Google Drive, etc.)

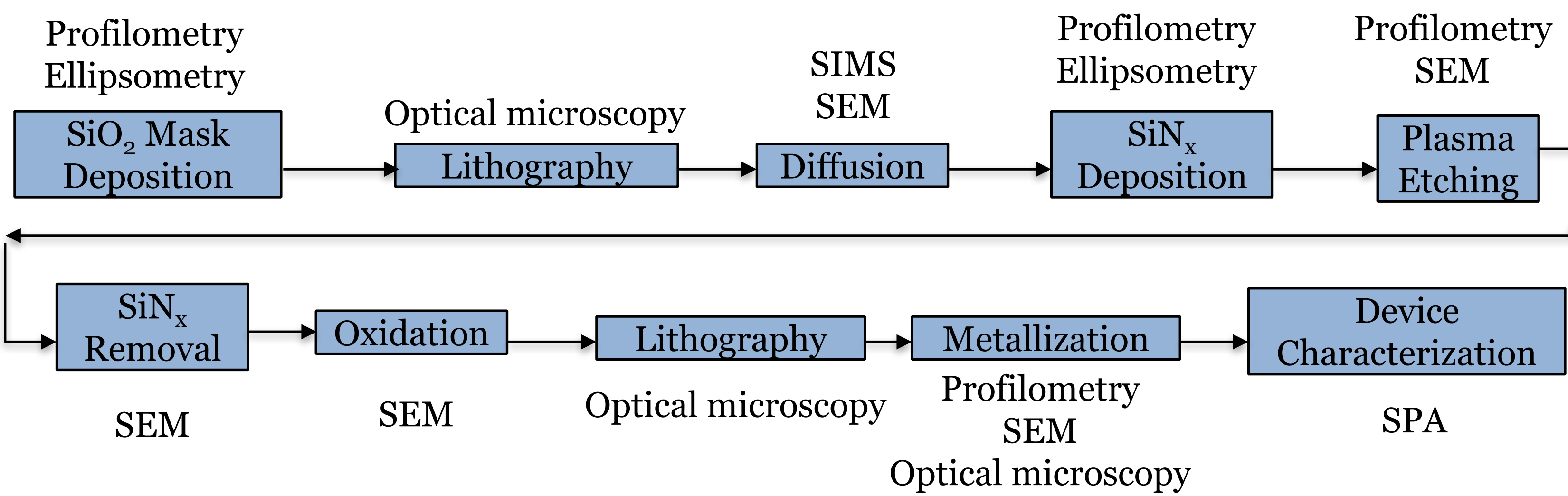
- Data access & sharing
  - Materials science & semiconductor fabrication areas have never been digitally connected
  - Only *"best"* results and data are kept for publishing

# Challenge: Data of multimodality and of heterogeneous types

- Data come of various types and formats
  - Each type requires a different data processing workflow




Sample output data from SEM microscopy

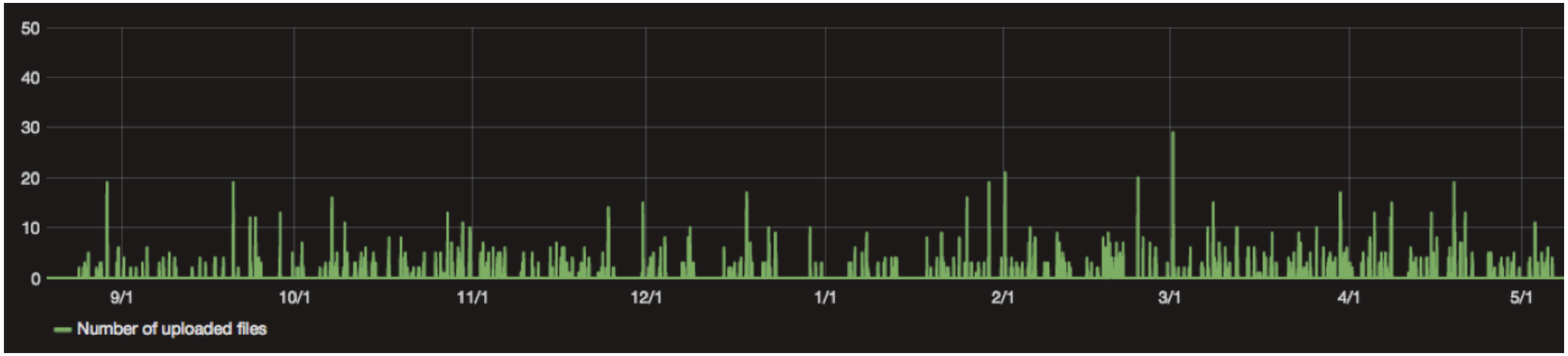



Heterogeneous types of experiments often done on a material sample

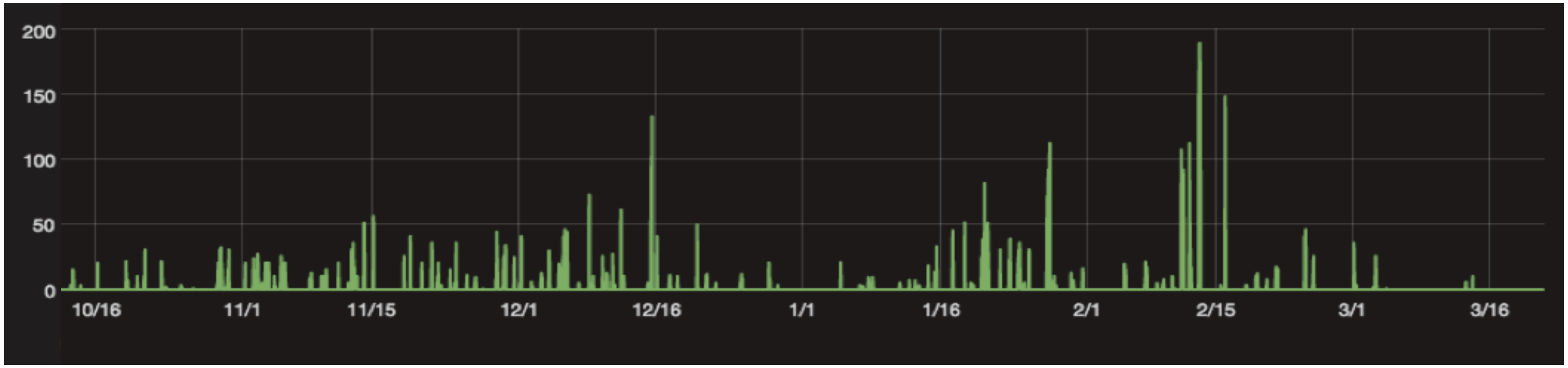
ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

# Challenge: Shorten time from digital capture to interpretation & insights

- Dynamism of workload: Data generated from scientific instruments are highly variable and often bursty




Number of files generated on a JOEL instrument from Sep 2015 to May 2016




Data (in MB) generated on a HeliosFIB instrument from Oct 2015 to March 2016

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

# Outline

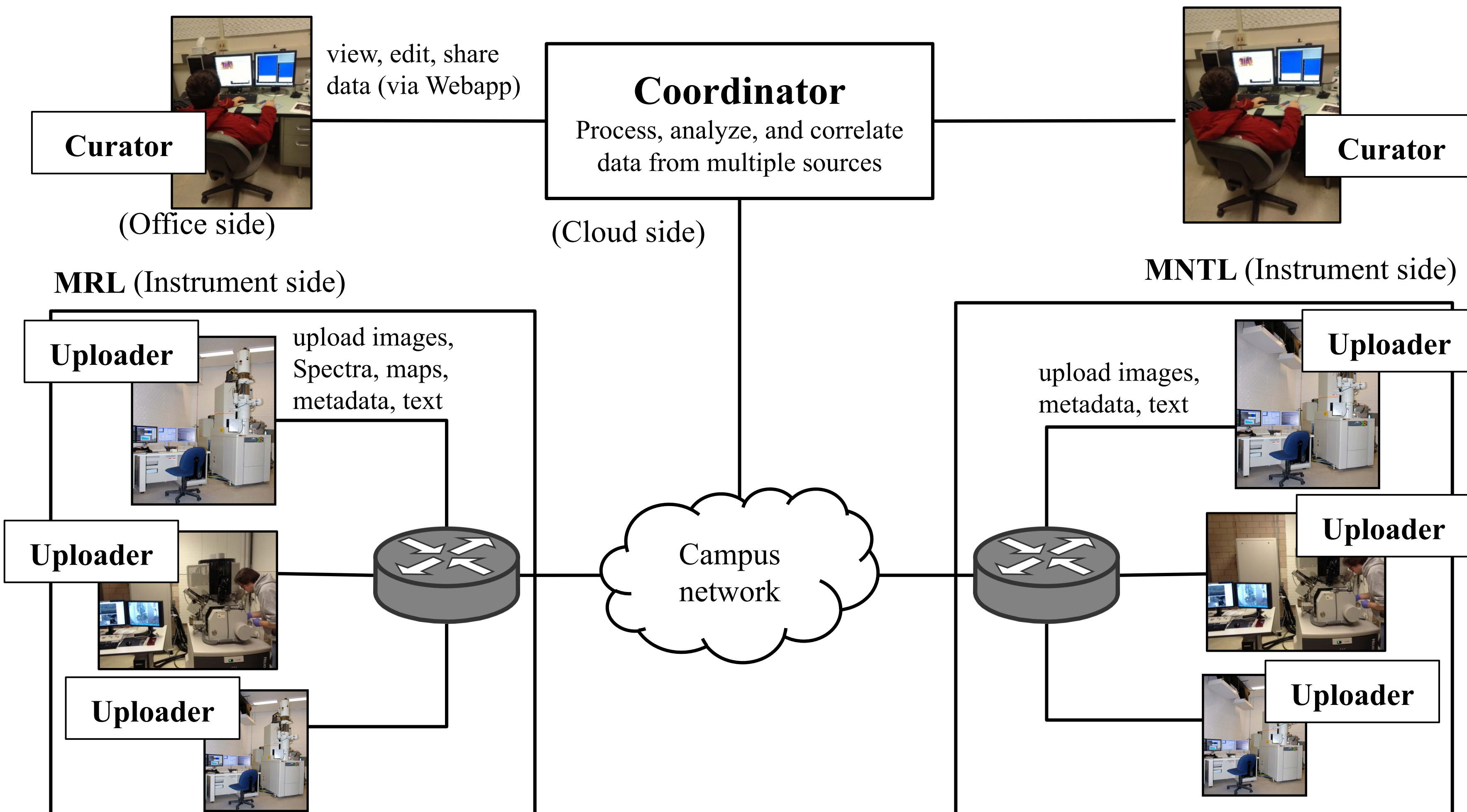
- Background & Challenges
- **Our solution: 4CeeD**
  - Curation Service
  - Coordination Service
- Evaluation
- Conclusions

# Our approach

- Existing systems focusing on:
  - Making *existing datasets* more accessible & sharable (e.g.: DataUp, SkyServer)
  - *Long-term preservation* of data (e.g., SEAD)
  - *Easy access* to grid/cloud infrastructure & collaboration (e.g.: HubZero, NanoHub, BrownDog, Data Conservancy Instance)

Our approach is to provide an expedient mean to capture, transfer, and process the digital data in real-time and in trusted manner ***before*** archiving, further analysis, visualization and sharing of the experimental results

# **4CeeD**: Real-time data acquisition and analysis framework for material-related environment




* **4CeeD** stands for **C**apture, **C**urate, **C**oordinate, **C**orrelate, and **D**istribute material-related experimental data.

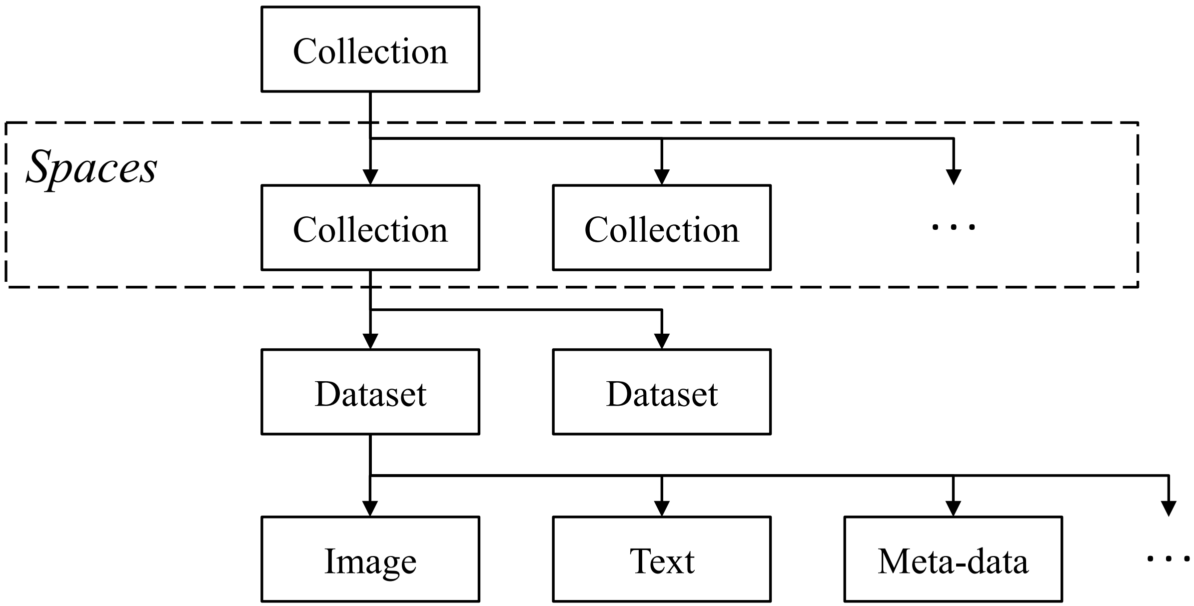
ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

# Outline

- Background & Challenges
- Our solution: 4CeeD
  - **Curation Service**
  - Coordination Service
- Evaluation
- Conclusions

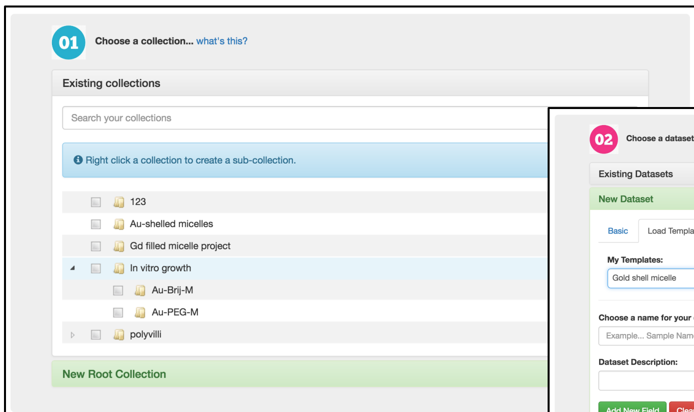
# 4CeeD's curation service: Data model

- Extendable data model
  - ***Vertical*** organization: Using concepts of *collections*, *datasets*, and *files* to provide flexible data organization
  - ***Horizontal*** organization: Using *spaces* to support sharing of data

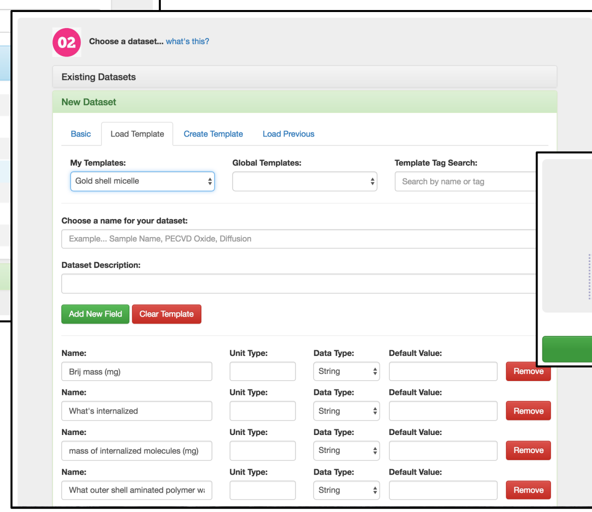


# 4CeeD's curation service: Data uploader

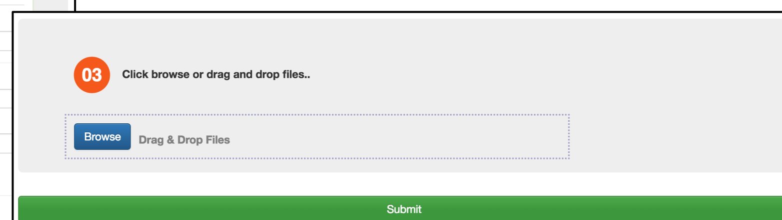
- Uploader offers a simple 3-step interface, following the data model




Create collection



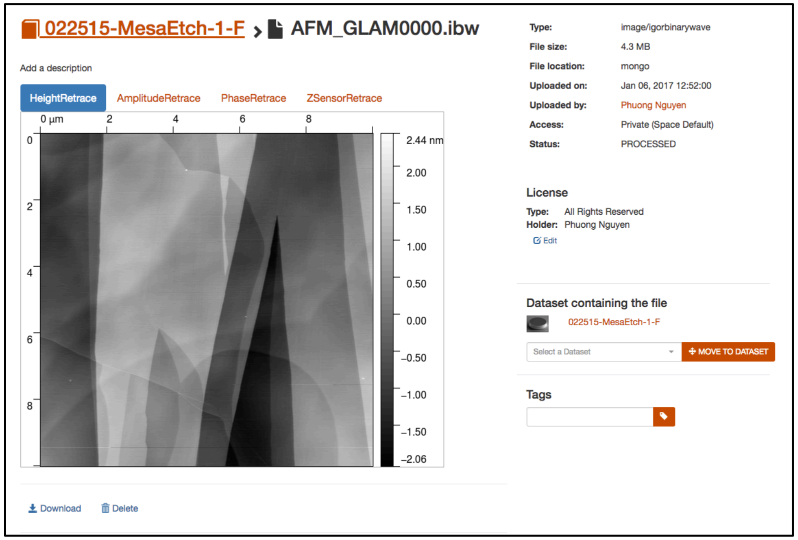

Create dataset &
associated meta-data




Select raw data files
to upload

# 4CeeD's curation service: Data curator

- Uploaded raw data is processed and available for curation (tagging, annotation) and sharing
- Processed data is indexed and available for search





Example of processed AFM experiment data with gray-shade preview & meta-data

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

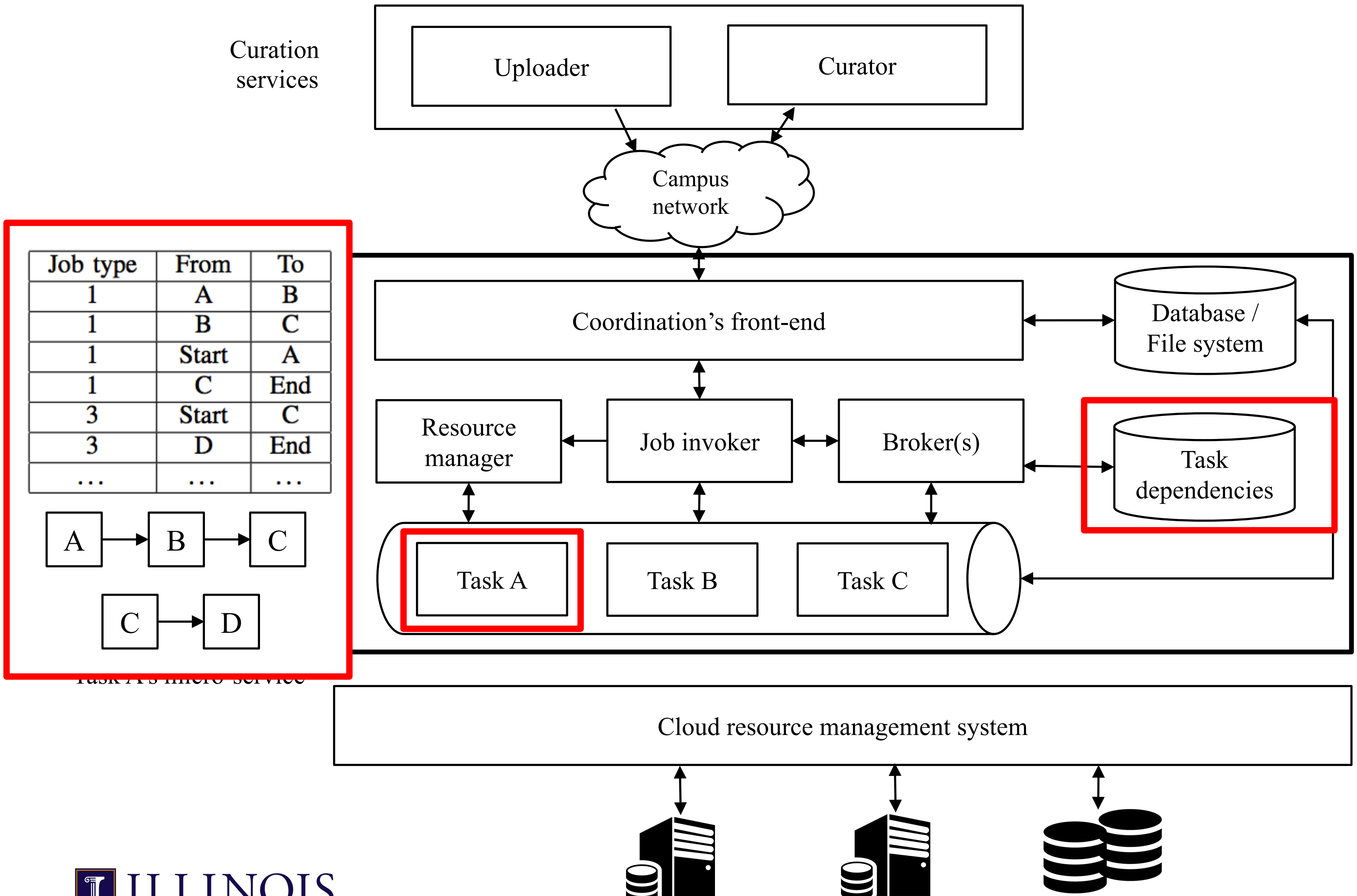
# Outline

- Background & Challenges
- Related work
- Our solution: 4CeeD
  - Curation Service
  - **Coordination Service**
- Evaluation
- Conclusions

# 4CeeD's coordination service design methodology

- Challenges/requirements (revisited)
  - Support heterogeneous data processing workflows
  - Handle variable & bursty workload
- Our approach:
  - Leverage resource abstraction provided by cloud resource management system (e.g., Kubernetes, YARN, Mesos)
  - Micro-service execution model:
    - Separate task dependencies from task implementation & deployment
  - Dynamic resource scheduling using system's performance model

# 4CeeD's coordination service's architecture

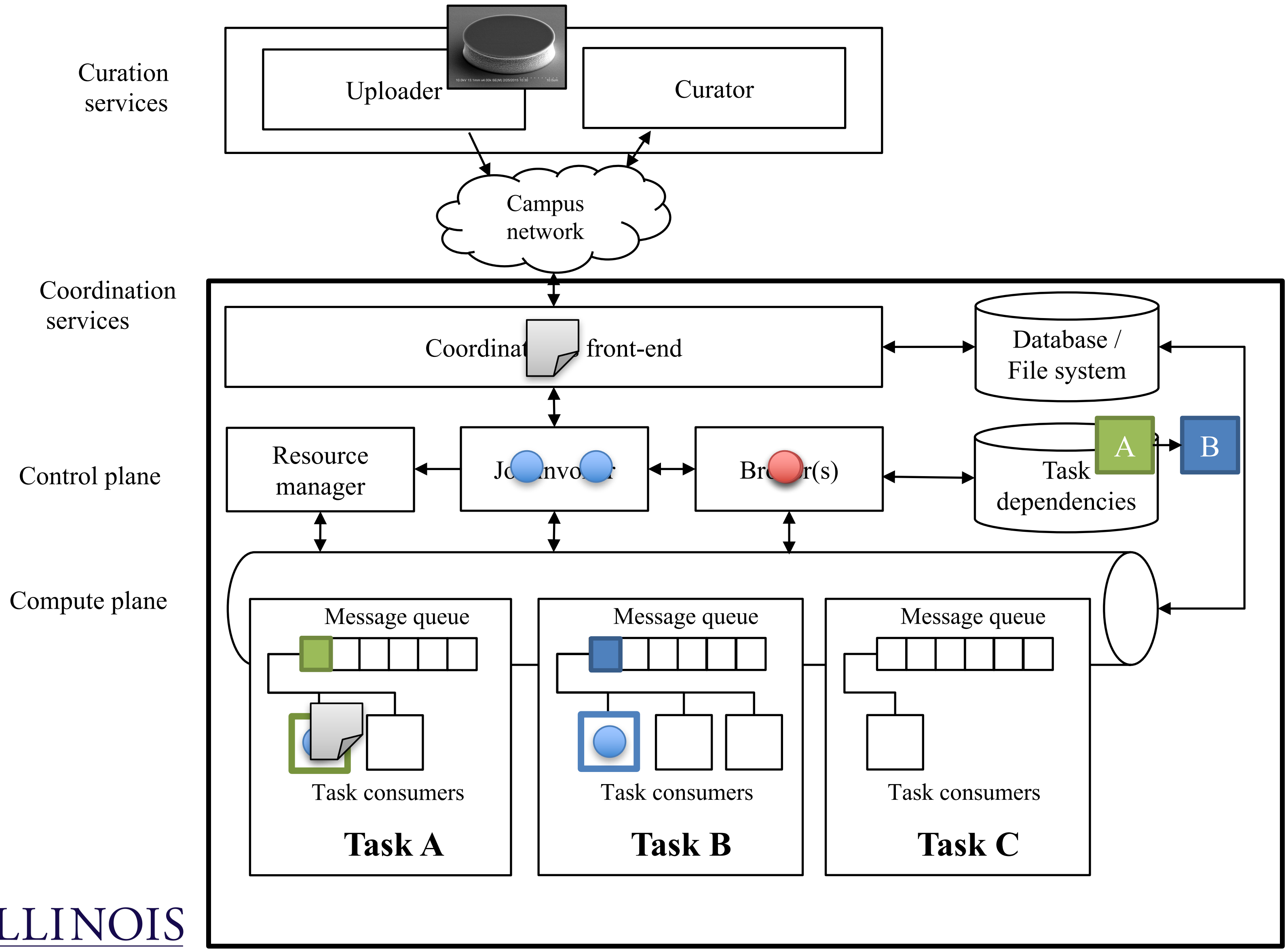



| Job type | From | To |
|---|---|---|
| 1 | A | B |
| 1 | B | C |
| 1 | Start | A |
| 1 | C | End |
| 3 | Start | C |
| 3 | D | End |
| … | … | … |

# 4CeeD's coordination protocol

- Coordination front-end:
  - Receives data, stores it in database, and create a job request to process the data.
  - Forwards the request to Job invoker
- Job invoker:
  - Notifies Resource manager about the new arrival request
  - Asks Brokers about the **first** task micro-service to handle the request
  - Forwards the request to the corresponding first task micro-service
- Task micro-service:
  - Pick-ups new request from its request queue & process the request
  - Asks Brokers about the **next** task micro-service(s) to handle the request
  - Forwards the request to the corresponding next task micro-service(s)

# 4CeeD's coordination service's architecture




ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

# Coordination service's resource scheduling

- Objective:
  - To decide how many consumers should be allocated to each task's micro-service to optimize the performance, subjected to resource constraint
- We use *work-in-progress* as system performance metric:
  - $WIP = \sum_j WIP_j$ ($WIP$ is proportional to response time, via Little's law)
  - Leverage our previous work* on performance model of elastic pub/sub system to represent $WIP_j$ as a function of number of consumers of task $j$-th: $WIP_j(m_j)$

$$argmin_{\{m_j\}} \sum_j WIP_j(m_j)$$

$$\sum_j m_j \leq \mathcal{C}$$

* Nguyen, P. and Nahrstedt, K., *"Resource Management for Elastic Publish Subscribe Systems: A Performance Modeling-based Approach"*, IEEE CLOUD 2016

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

# Dynamic resource scheduling algorithm (GRESMAN)

- Greedily allocate additional consumer (one for each iteration) to the task that maximize the benefit of reducing system's $WIP$, until the performance threshold ($\mathcal{T}$) or resource constraint ($\mathcal{C}$) is met.

---

**Algorithm 1** Dynamic Greedy Elastic Scaling Algorithm (GRESMAN)

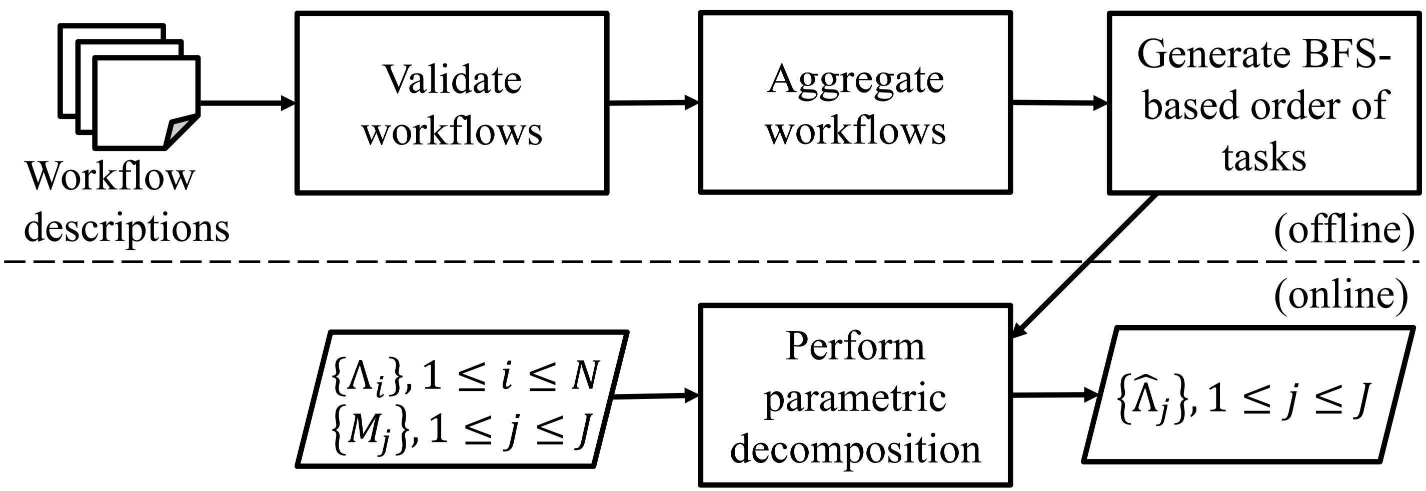
---

**procedure** `GRESMAN`
    Define $\mathbf{m}^0$ as the current configuration of consumers
    Initialize allocation plan $\mathcal{A} = []$
    Initialize iteration count $i = 1$
    $\{\hat{\Lambda}_j\} = \texttt{DECOMPOSE}(\{\Lambda_j\}, \{M_j\}, \mathbf{m}^0)$ ▷ Initial decomposition
    **while** $WIP(\mathbf{m}^0) > \mathcal{T}$ and $F(\mathbf{m}) < \mathcal{C}$ **do**
        $\hat{j} = argmax_{1 \leq j \leq J} \Delta(m_j^i, m_j^i + 1)$ ▷ Find the most beneficial PC
        $m_{\hat{j}}^i = m_{\hat{j}}^i + 1$ ▷ Add one consumer to that PC
        $\mathcal{A}.append(\hat{j})$ ▷ Update allocation plan
        $\{\hat{\Lambda}_j\} = \texttt{DECOMPOSE}(\{\Lambda_j\}, \{M_j\}, \mathbf{m}^i)$ ▷ Update $\{\hat{\Lambda}_j\}$
        $i = i + 1$ ▷ Update iteration count
    Return $\mathcal{A}$, $\mathbf{m}^i$

---

# Coordination service's resource scheduling

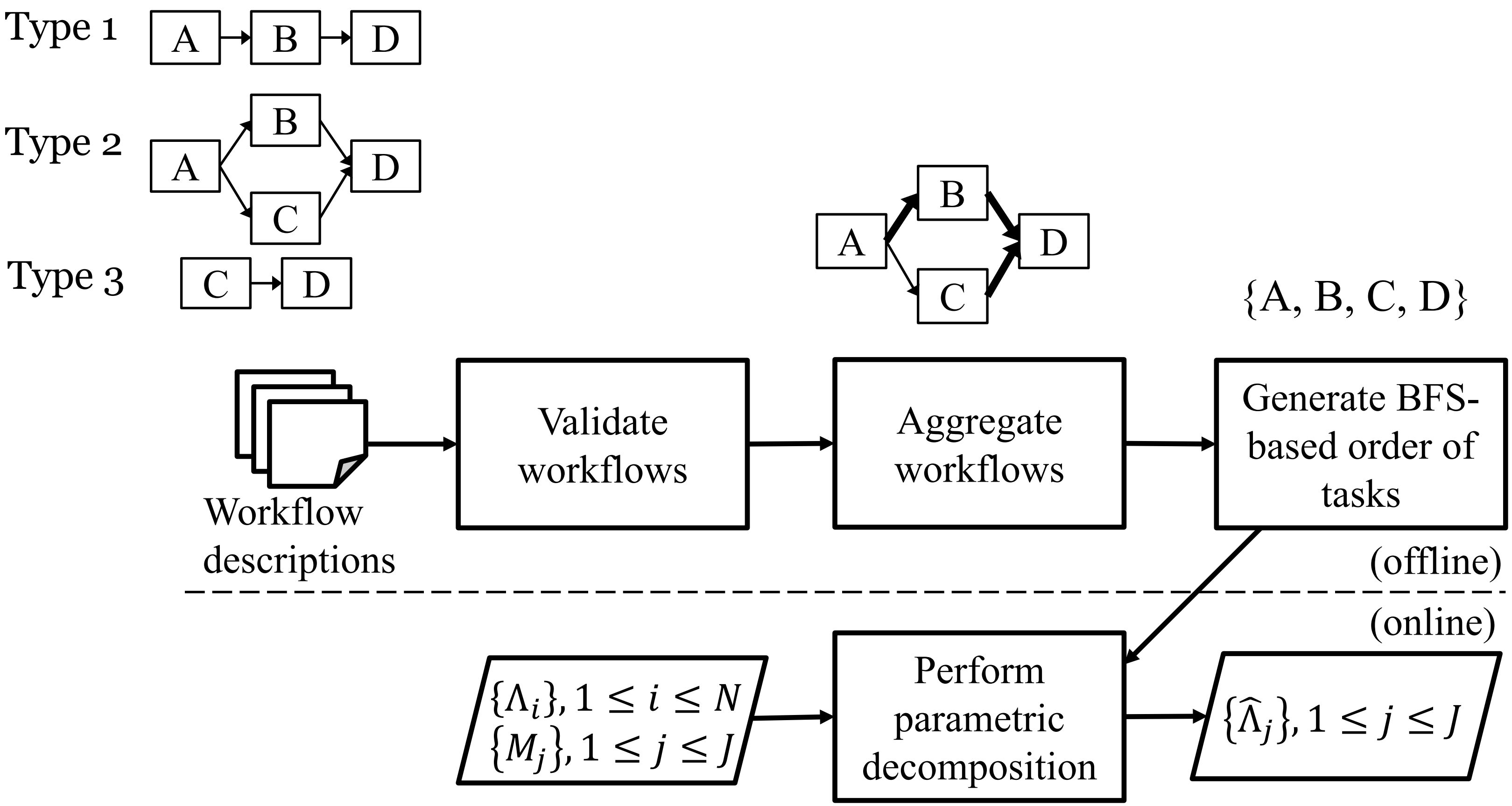
- Under micro-service execution model, calculating $WIP_j(m_j)$ requires **decomposition** of system-specific workload into task-specific's workloads
  - From $\{\boldsymbol{\Lambda_i}\}, \mathbf{1} \leq \boldsymbol{i} \leq \boldsymbol{N}$ (and known $\{M_j\}, 1 \leq j \leq J$) to $\{\boldsymbol{\widehat{\Lambda}_j}\}, \mathbf{1} \leq \boldsymbol{j} \leq \boldsymbol{J}$




Workload parameters decomposition procedure

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

# Coordination service's resource scheduling

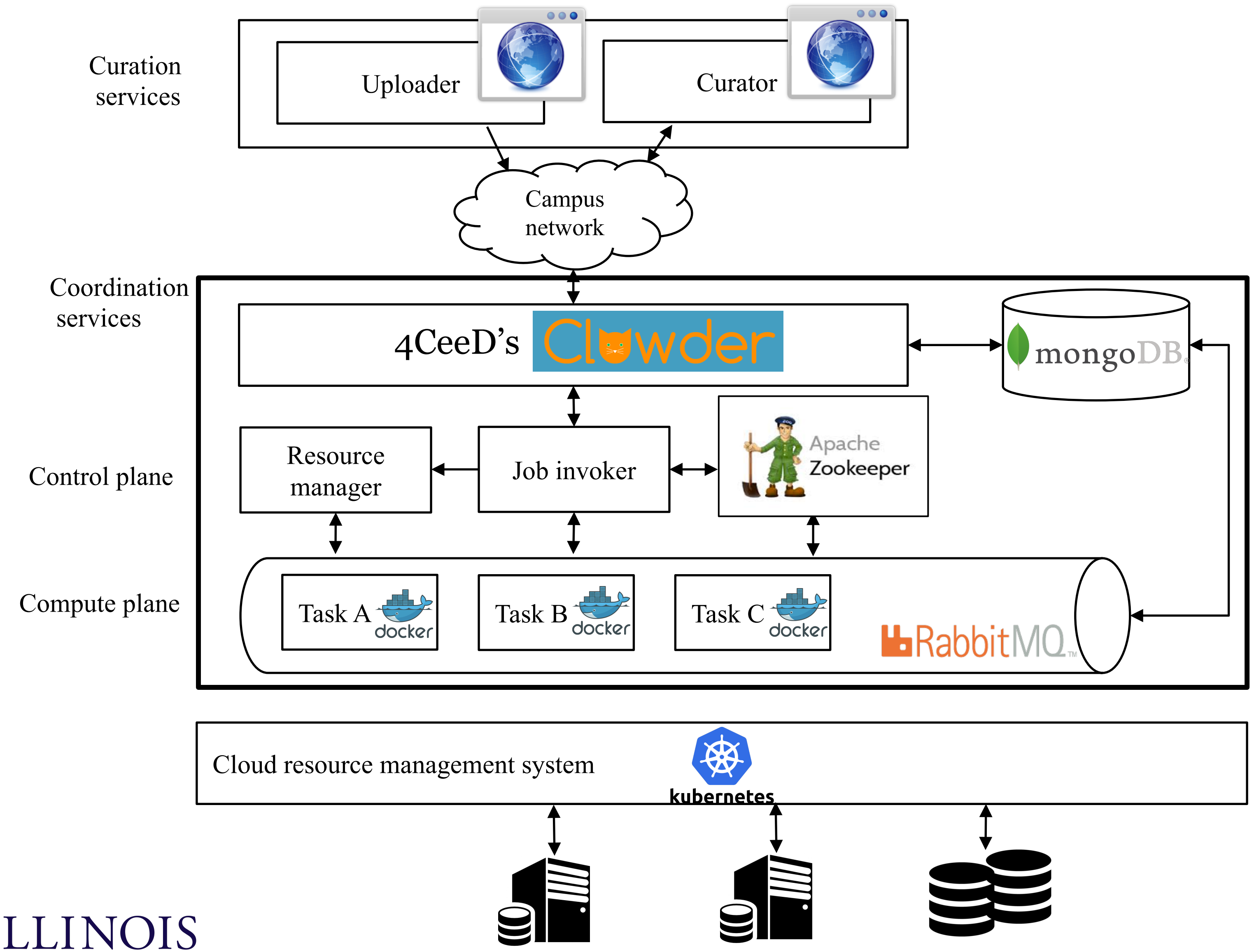



E.g.: Type 1 & type 2 workloads merge at A, and then split to B & C

# Outline

- Background & Challenges
- Our solution: 4CeeD
  - Curation Service
  - Coordination Service
- **Evaluation**
- Conclusions

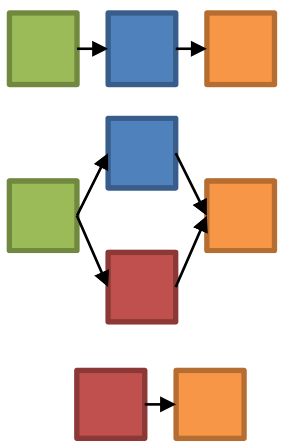
# System implementation



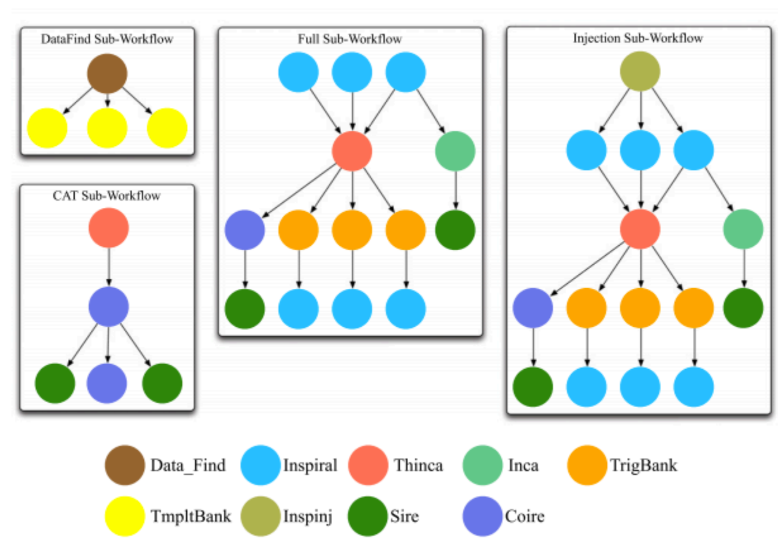

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

# Evaluation settings

- Data processing workflows:
  - MDP: material data processing workflows (to process DM3, AFM, etc.)
  - LIGO: analyze data to study stars and black holes



MDP workflows

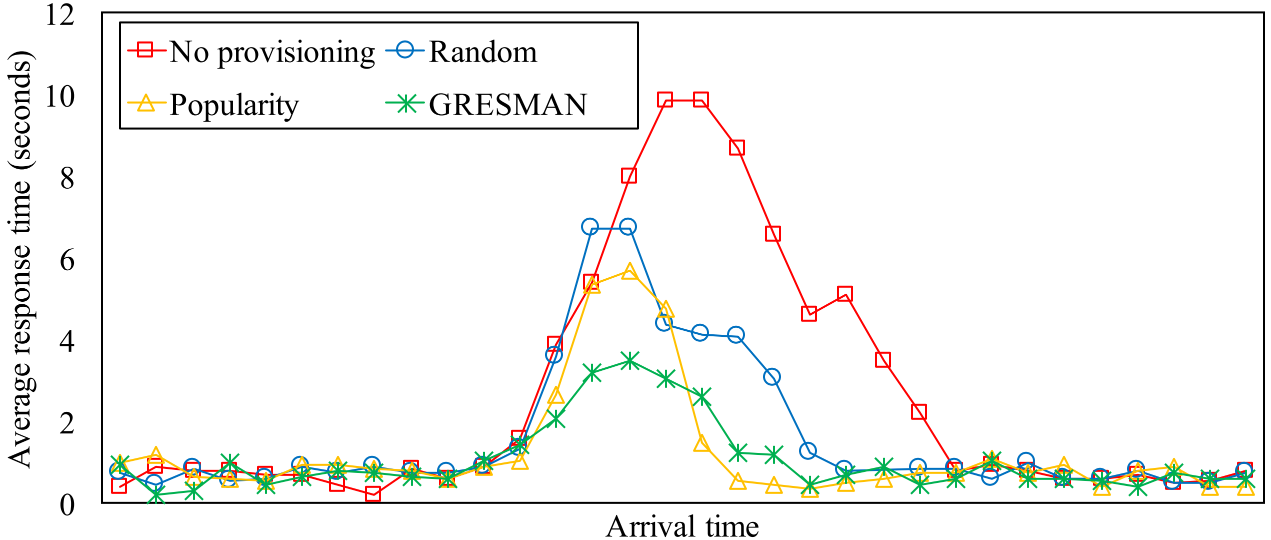



LIGO workflows

# Evaluation settings (cont.)

- Emulate the bursty workload situation by abnormally increasing the arrival rates of requests to up to **10x higher** on MDP and LIGO workflows
- Performance threshold ($\mathcal{T}$): Absolute delay guarantee on average processing time
  - Set to 2s for the MDP and 10s for LIGO workflows
- Resource constraint ($\mathcal{C}$): Maximum number of (homogeneous) consumers that can be allocated by the system
  - Set to 10 and 60 consumers for the MDP and LIGO workflows respectively
- 4CeeD coordination service is deployed on a cluster of three nodes, each node is equipped with an Intel Xeon quad core processor (1.2Ghz for each) and 16GB of RAM.

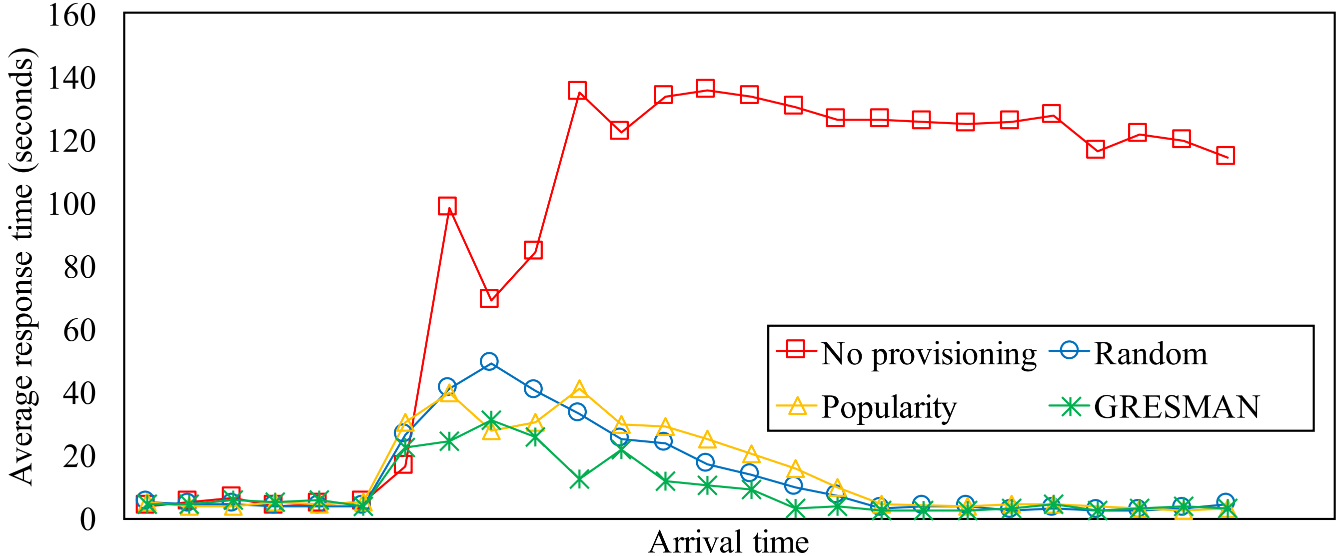
# Effectiveness of coordination service in handling bursty workload




MDP workflows

Our resource scheduling strategy (GRESMAN) demonstrates effectiveness in bursty workload situation.

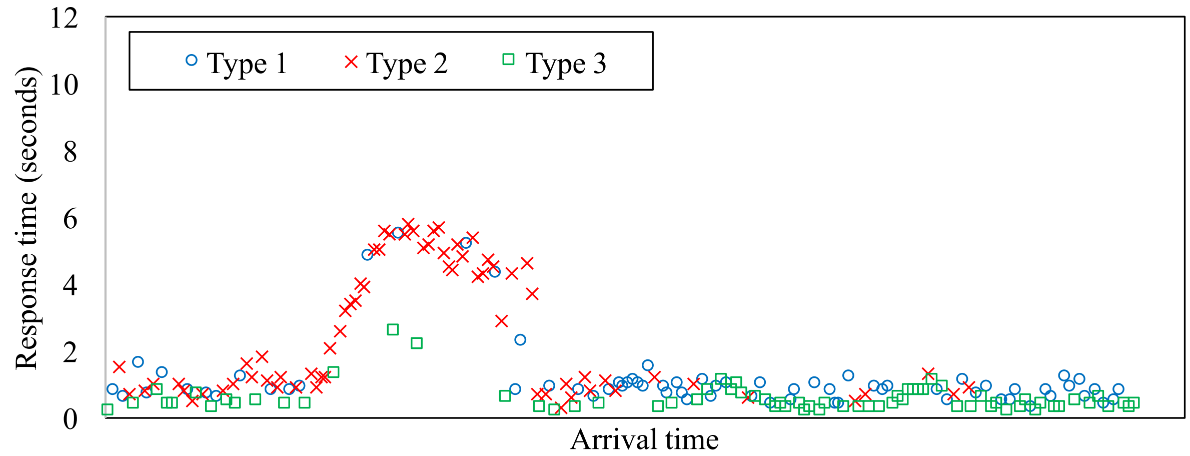
# Effectiveness of coordination service in handling bursty workload



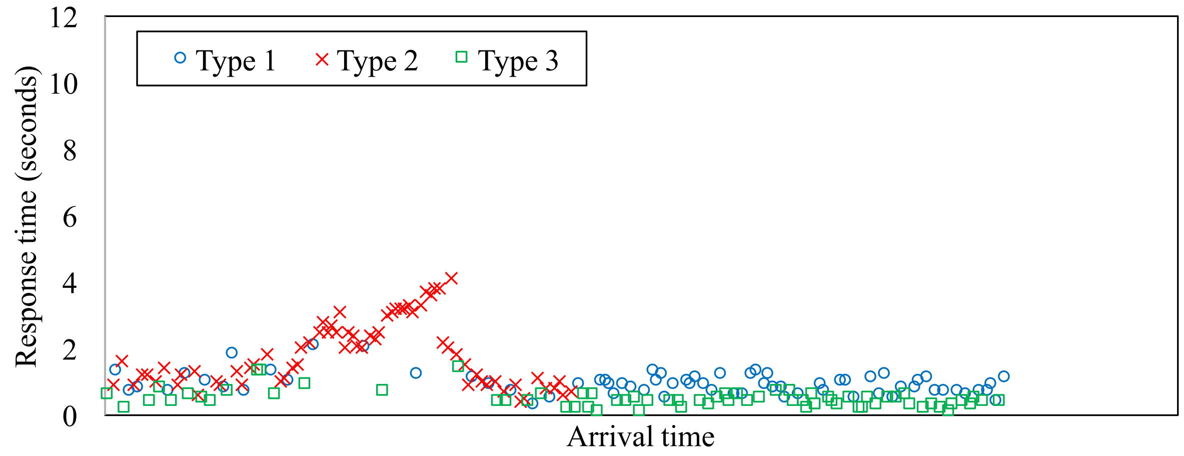

MDP workflows

Our resource scheduling strategy (GRESMAN) demonstrates effectiveness in bursty workload situation.

# Impact of order of scheduled resource in handling bursty workload with MDP workflows



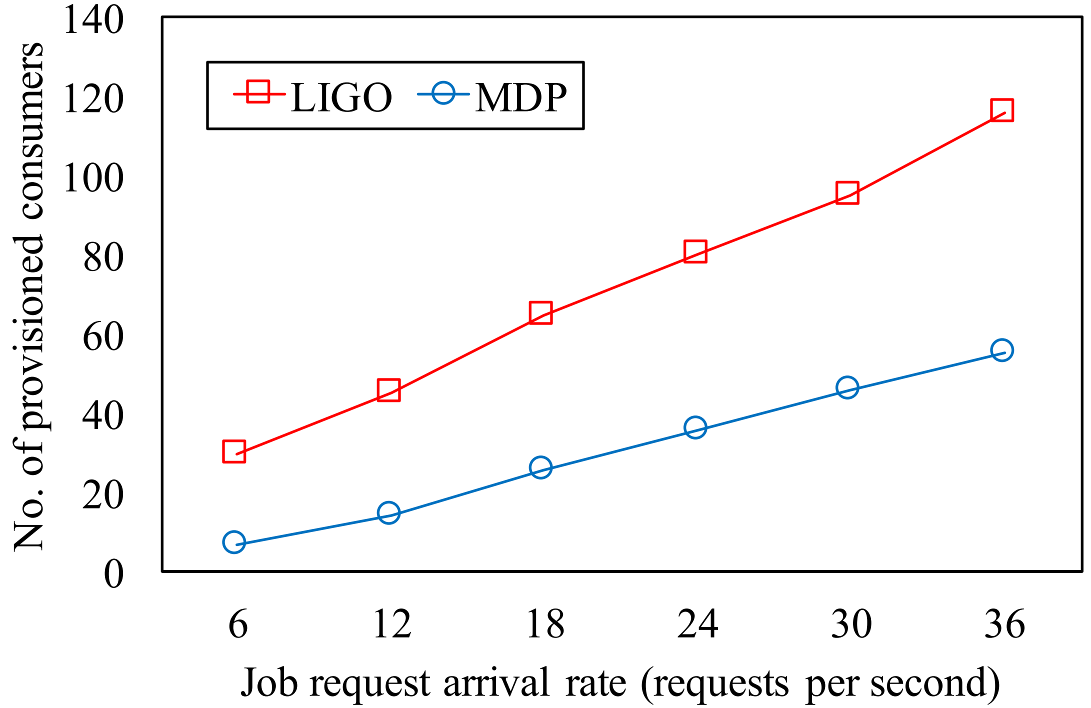

Non-optimal order




Optimal order

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

# Scalability of coordination service




Our resource scheduling strategy (GRESMAN) demonstrates linear scalability in terms of required number of consumers to meet performance constraint

# Outline

- Background & Challenges
- Our solution: 4CeeD
  - Curation Service
  - Coordination Service
- Evaluation
- **Conclusions**

# Conclusions

- We present the *first* real-time data acquisition and analysis framework for materials-related and semiconductor fabrication areas
  - Novel curation service for nimble, comprehensive data collection and management
  - Novel scalable cloud-based system architecture for processing of heterogeneous multi-modal material-related data.

# 4CeeD's research impacts

- Implemented and deployed the 4CeeD system at the University of Illinois at Urbana-Champaign (UIUC)
  - ~30 test users from Micro and Nano Technology Lab (MNTL) and Material Research Lab (MRL)
- 4CeeD is open-sourced and is being evaluated at NIST
  - https://4ceed.github.io/
- Feedbacks for users:
  - Using 4CeeD achieves about **30% time saving** during experimental sessions (e.g., file transfer, save metadata, export previews, etc.)
    - This time saving translates into saving $25 to $30 each hour of lab session cost
  - Shorten time from digital capture to curation, interpretation & insights
    - Traditionally takes hours
  - Other benefits include better data preservation, exploration, and security

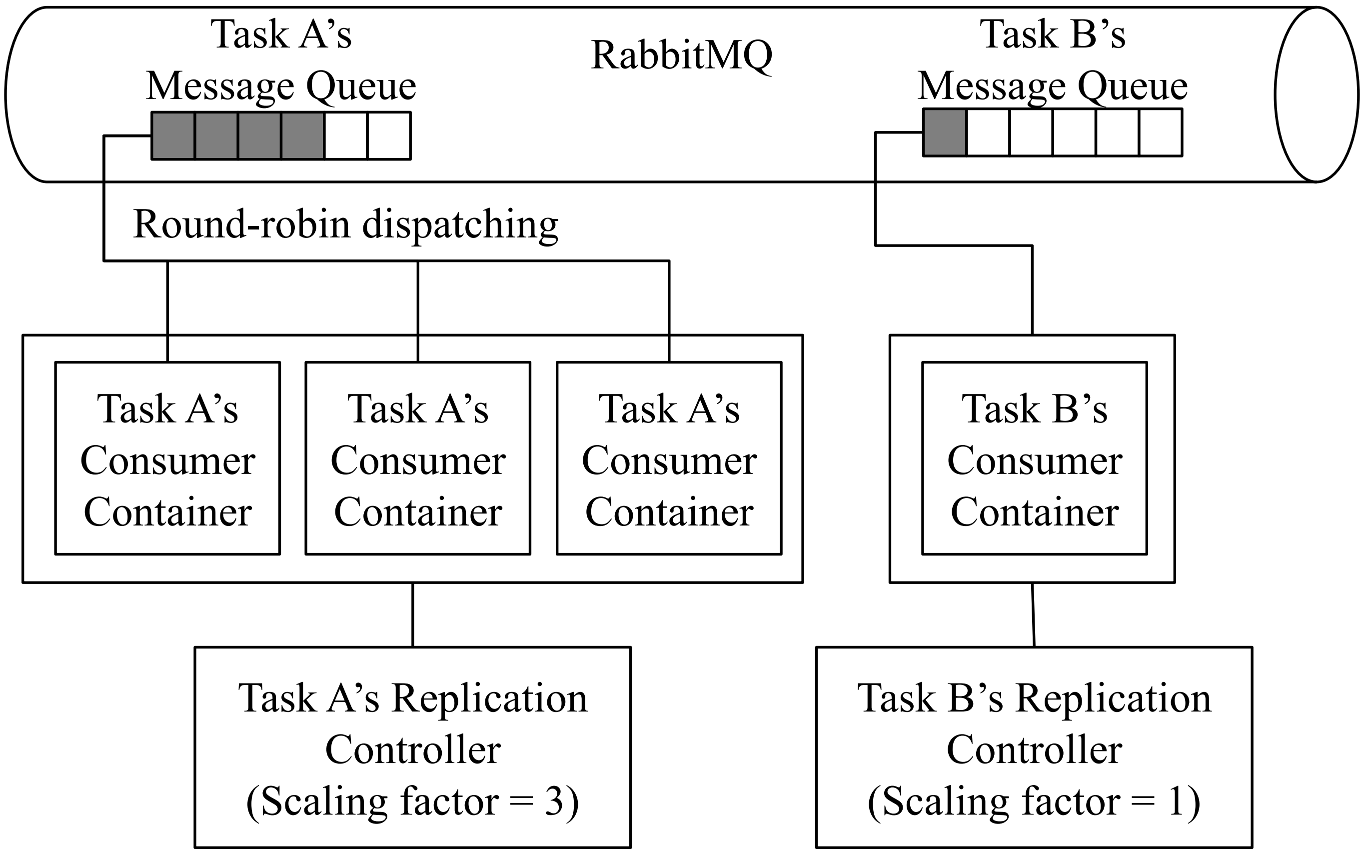
# Thank you!

# Back-up slides

# Related work

- Systems focusing on making *existing datasets* more accessible & sharable
  - E.g.: DataUp, SkyServer
- Systems focusing on *long-term preservation* of data
  - E.g.: SEAD
- Systems providing *easy access* to grid/cloud infrastructure & focus on collaboration
  - E.g.: HubZero, NanoHub (for nanotechnology simulation); BrownDog, Data Conservancy Instance (for cloud-based data curation)

# Compute layer implementation




ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN