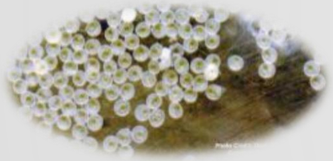
Integrating Cloud Computing Capability to ECO-PTM Simulations

May 2025

Neutrally Buoyant Particles



Particles with behavior

Position Oriented Particles

Xiaochun Wang, Nicky Sandhu

Modeling Support Office

GOAL:
How ECO-PTM uses cloud computing to conduct studies

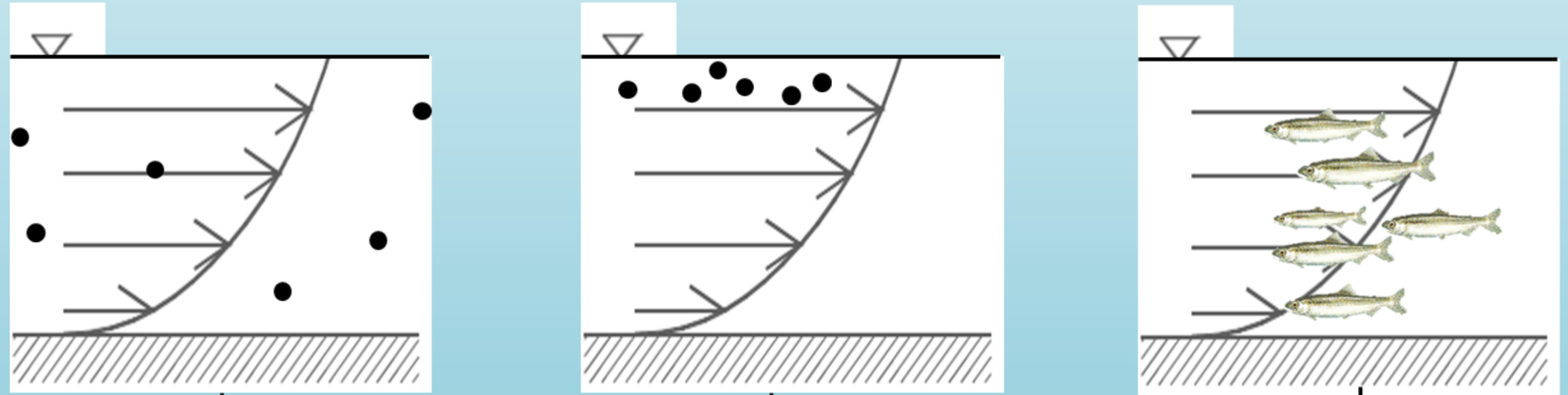
- Overview of ECO-PTM
  - What is the model
  - Typical ECO-PTM studies
  - Why we need cloud computing
- ECO-PTM cloud computing
  - How to use batch and storage accounts
  - ECO-PTM batch job processing

WHAT IS ECO-PTM?

Neutrally Buoyant
(Delta Smelt Larvae)

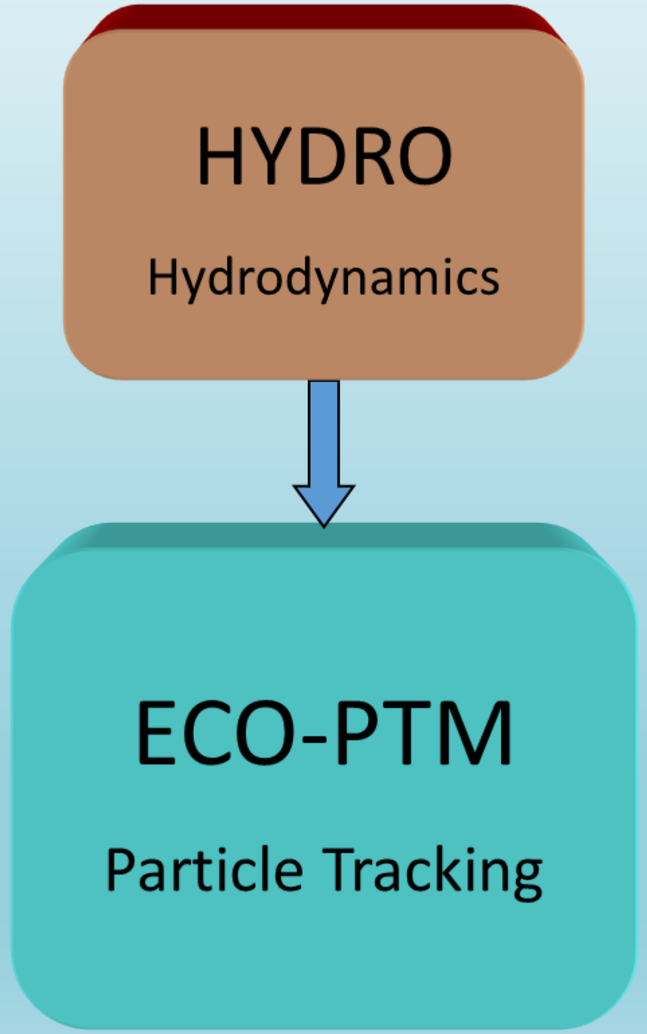
Position Oriented
(Longfin Smelt Larvae)

Particles with Behavior
(Chinook Salmon Smolts)



Particle Flux

Fish Survival Rate



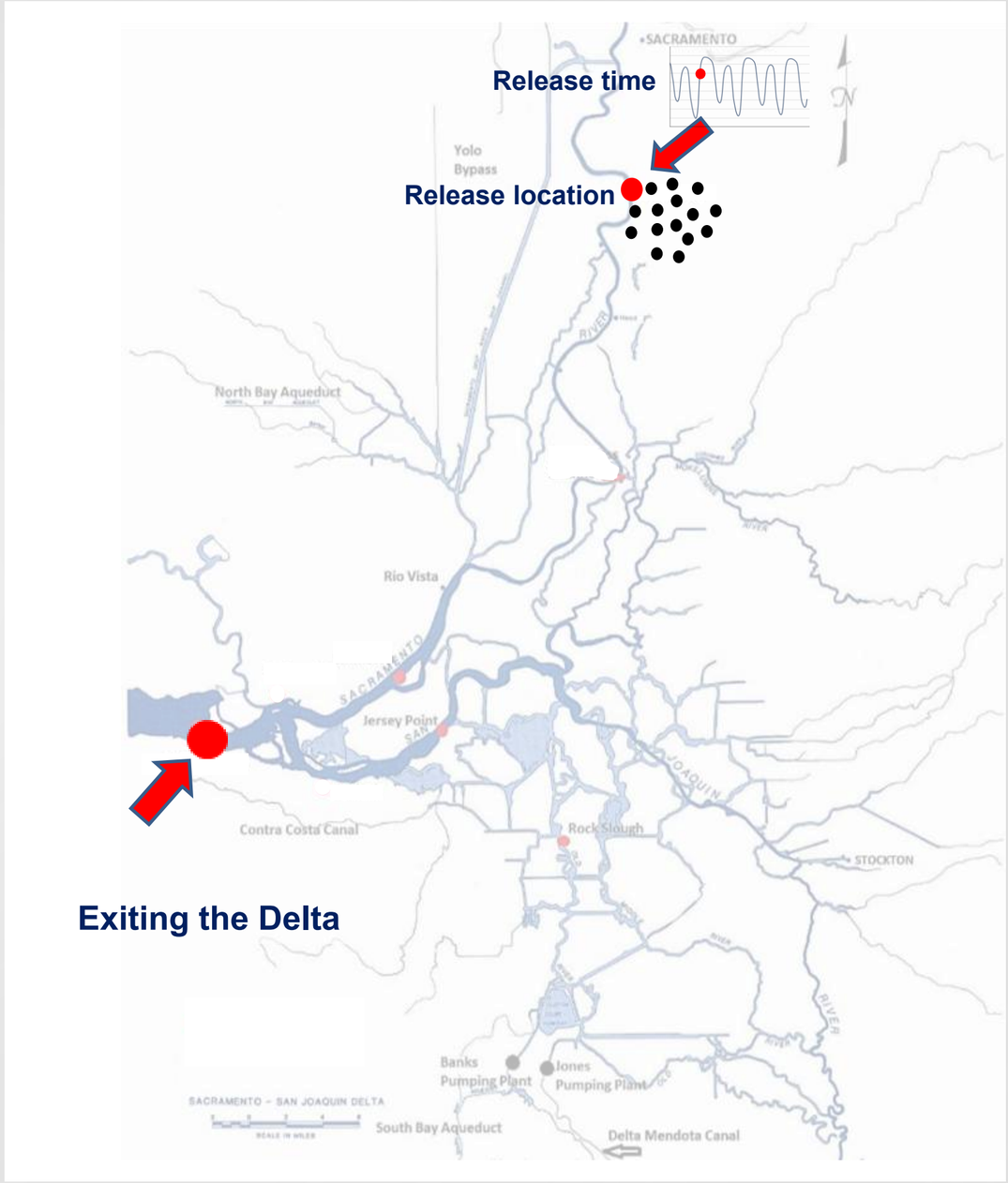
Projects that have used ECO-PTM

- Long Term Operation (LTO)
- Delta Conveyance Program (DCP)
- Incidental Take Permit (ITP)
- Temporary Urgency Change Petitions (TUCP)
- Drought Barriers
- Fish Barriers
- ……

Typical ECO-PTM analysis

A typical simulation:

- A group of particles (4000-10000), released at a specific location on a specific date
- 90+ days simulation period
- Calculating flux/survival at the end
- Different release locations and dates - different simulations

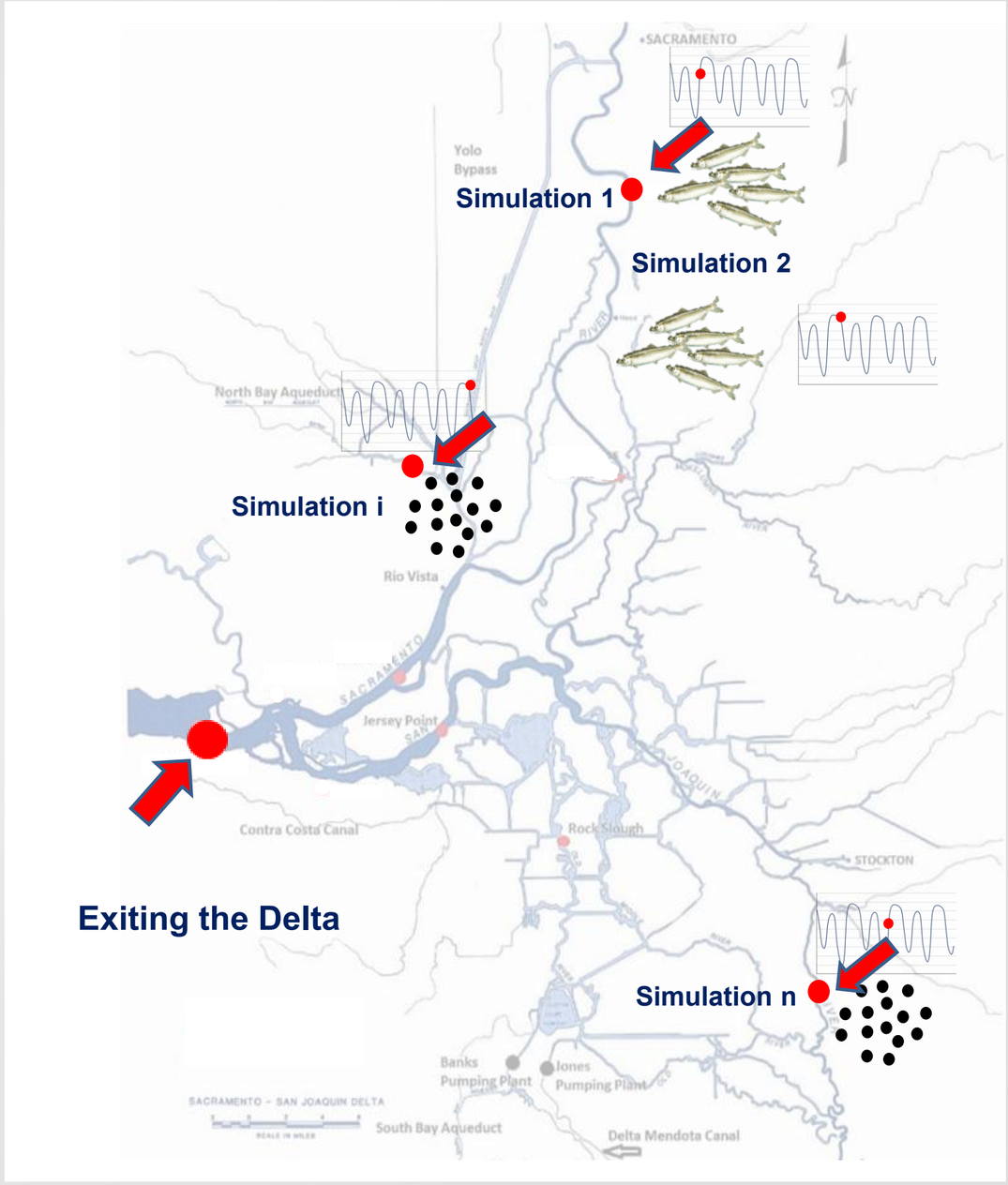


Typical ECO-PTM analysis

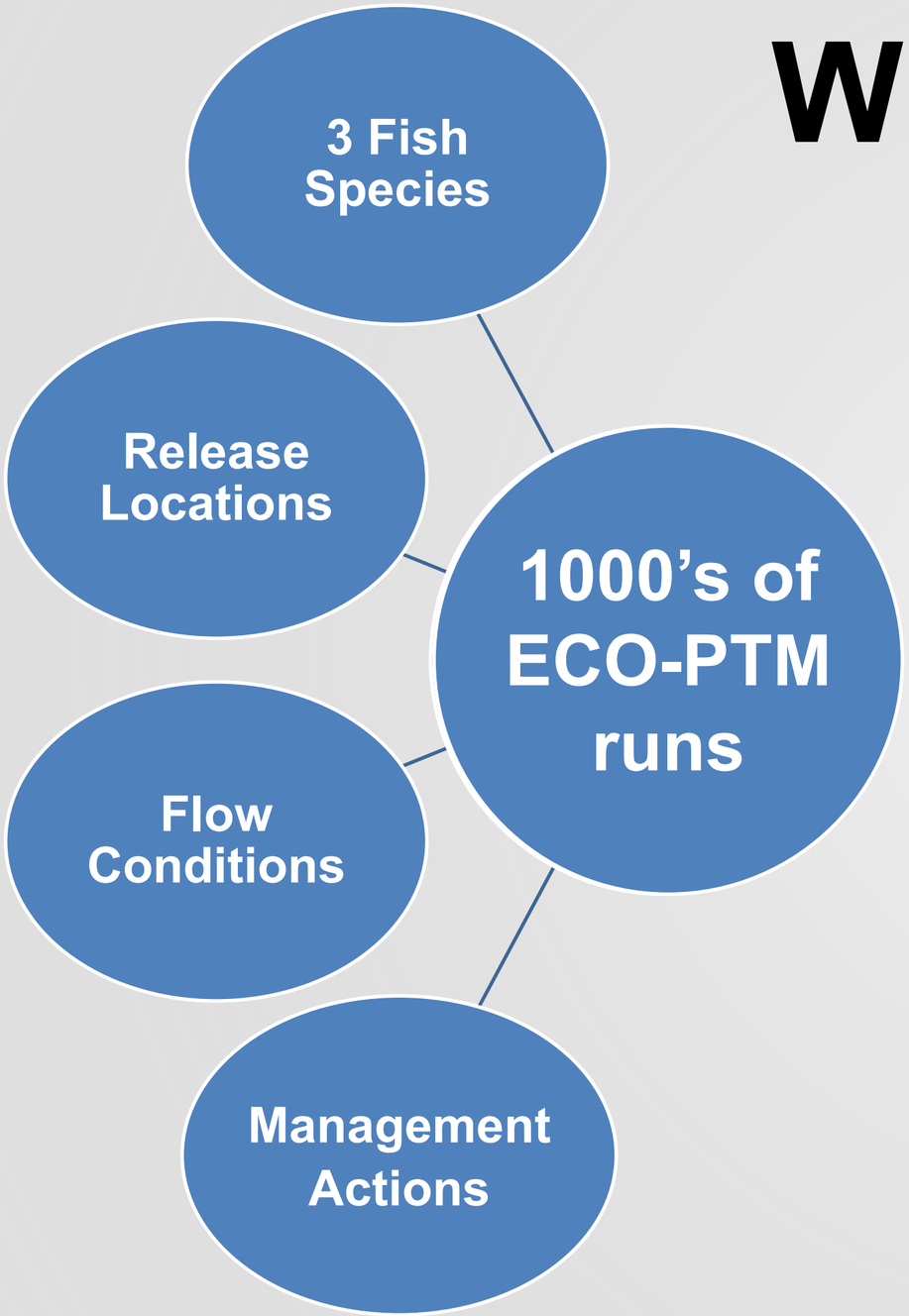
A typical study:

- Simulations:
  - 3 types of particles,
  - Large number of particles (e.g., 1000s)
  - Multiple release locations (e.g., 39)
  - Numerous release dates (e.g., release particles daily during the fish season over 100 years)
  - Many management scenarios

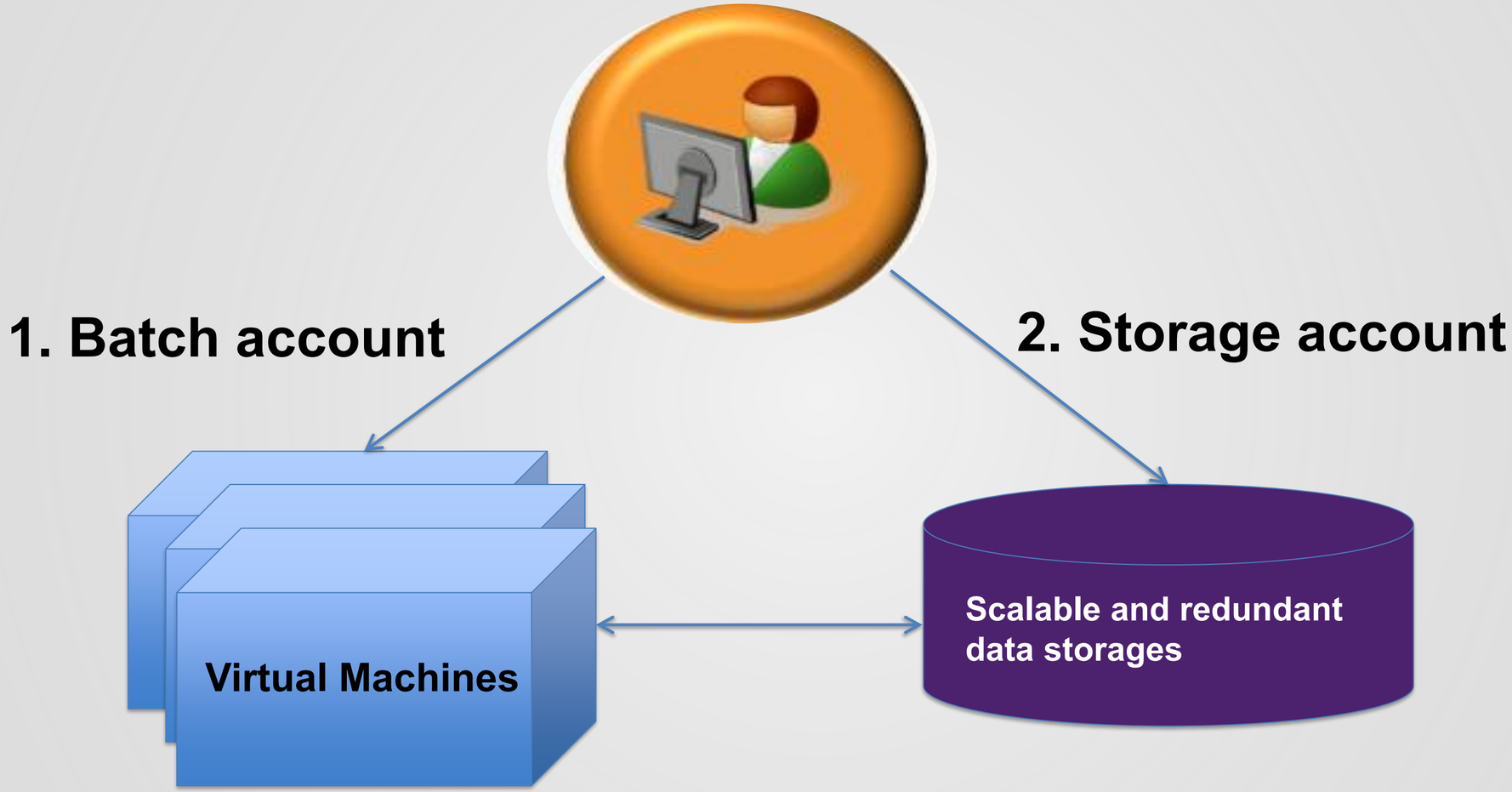
→ Large number of simulations



Why Cloud Computing

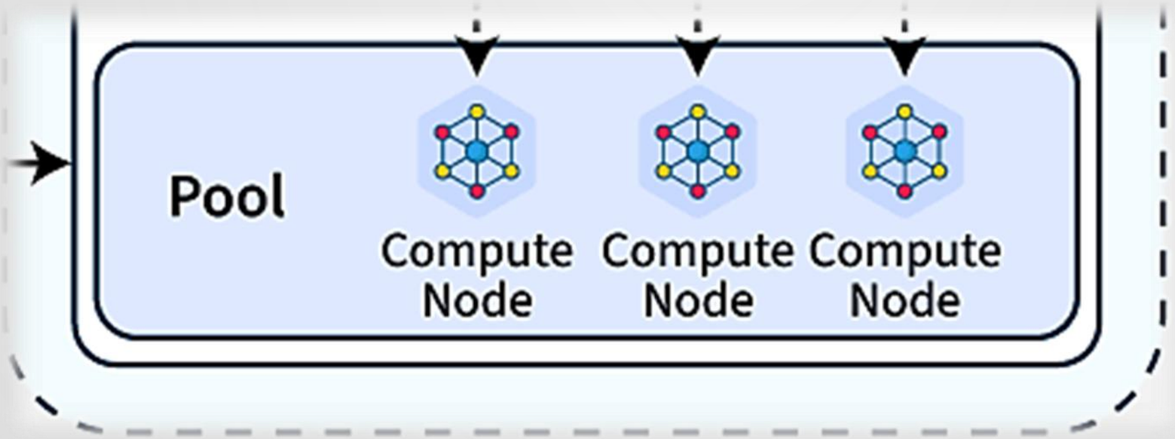


ECO-PTM Cloud Computing



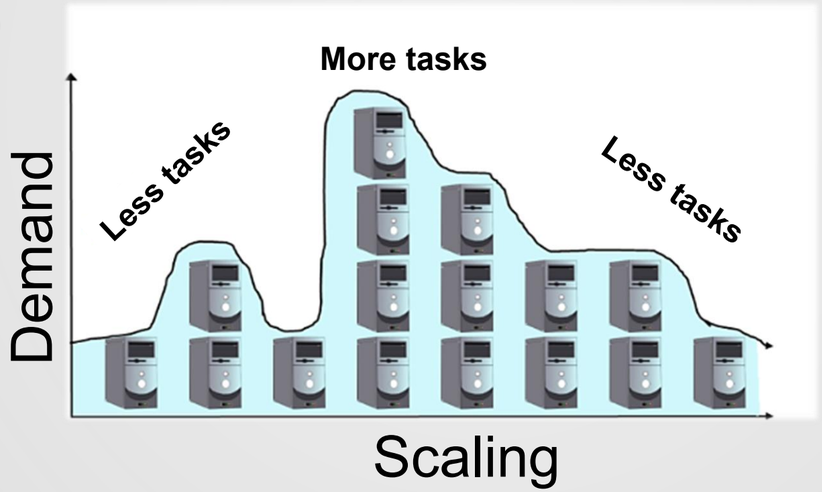
Cloud Computing: Batch Account

- Virtual machine pool:
  - Virtual machine (node):
    - Azure Standard_DS5_v2:
      - » *16 vCPUs and 56 GiB of memory*



Cloud Computing: Batch Account

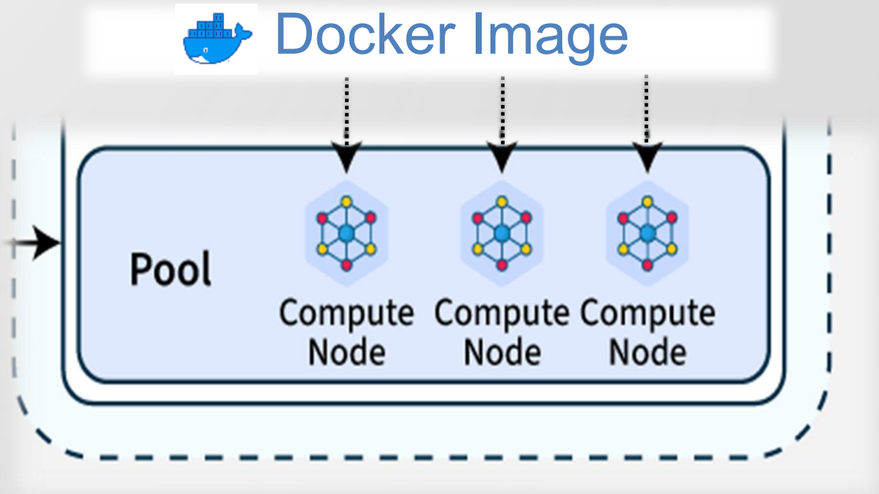
- Virtual machine pool:
  - Auto scaling formula:
    - Scaling up and down nodes:
      - » *Set 1 node initially, expand out if necessary, 0 nodes If all job done*



    - Number of dedicated and low priority nodes: *only 1 dedicated*
    - Maximum number of nodes: *2000*

Cloud Computing: Batch Account

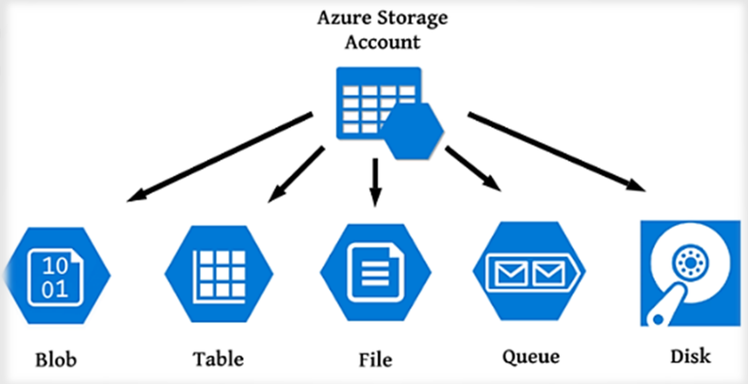
- Software Environment:
  - Application packages
    - DSM2 executables
    - Linux operation system (OS)
    - Libraries needed to run ECO-PTM
    - Script to install libraries, etc.
  - Docker Images
    - Package of executables, OS and libraries
    - Two docker images available: DSM2 v8.2.2 and DSM2 v8.4.1



Cloud Computing: Storage Account

Provides scalable and redundant data storage

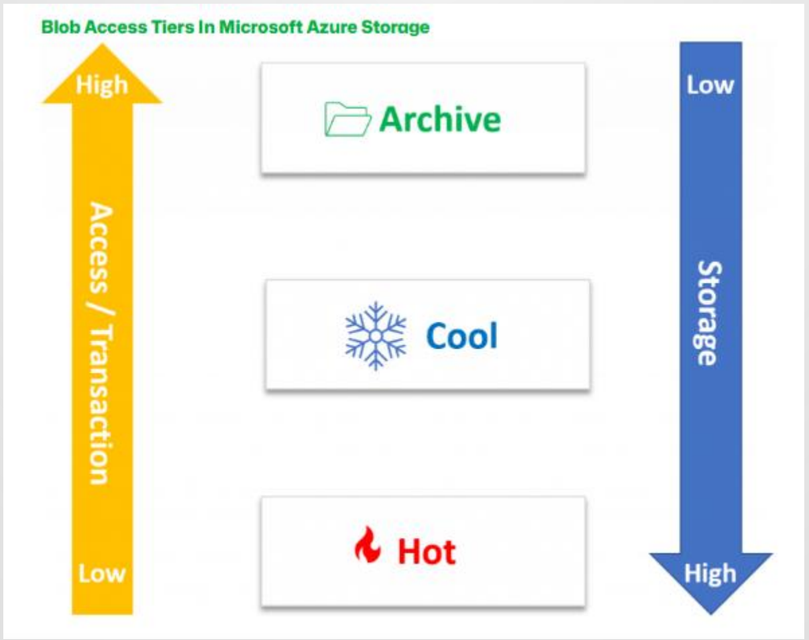
– Storage types:



For all ECO-PTM studies:

- Blob containers: StorageV2

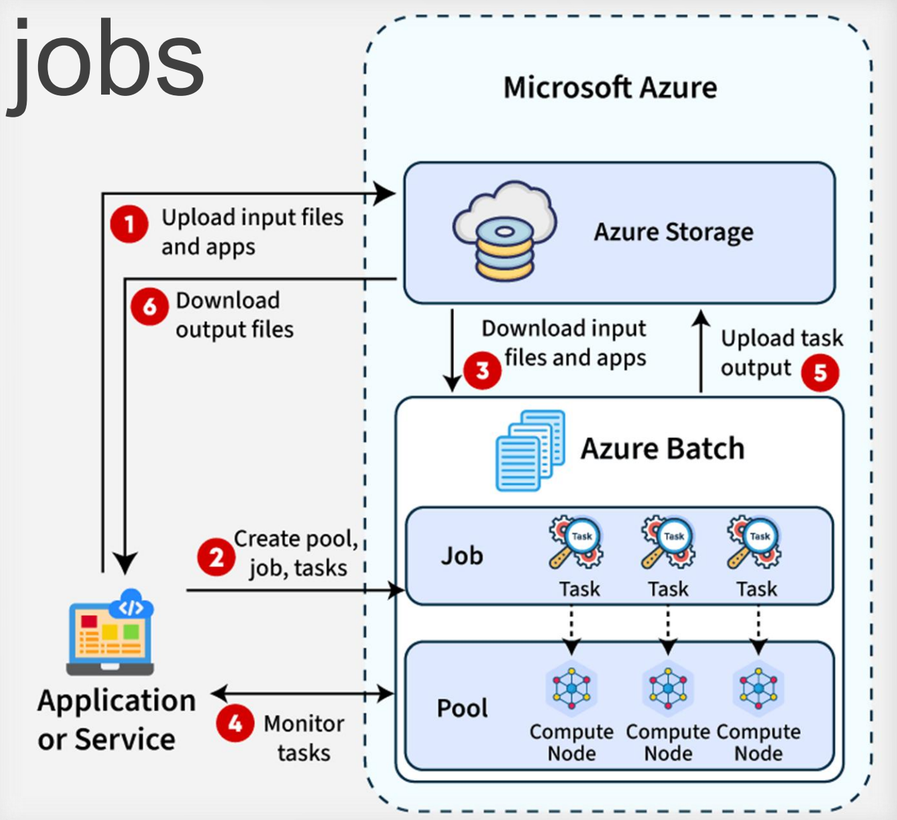
– Access tiers:



- Hot tier

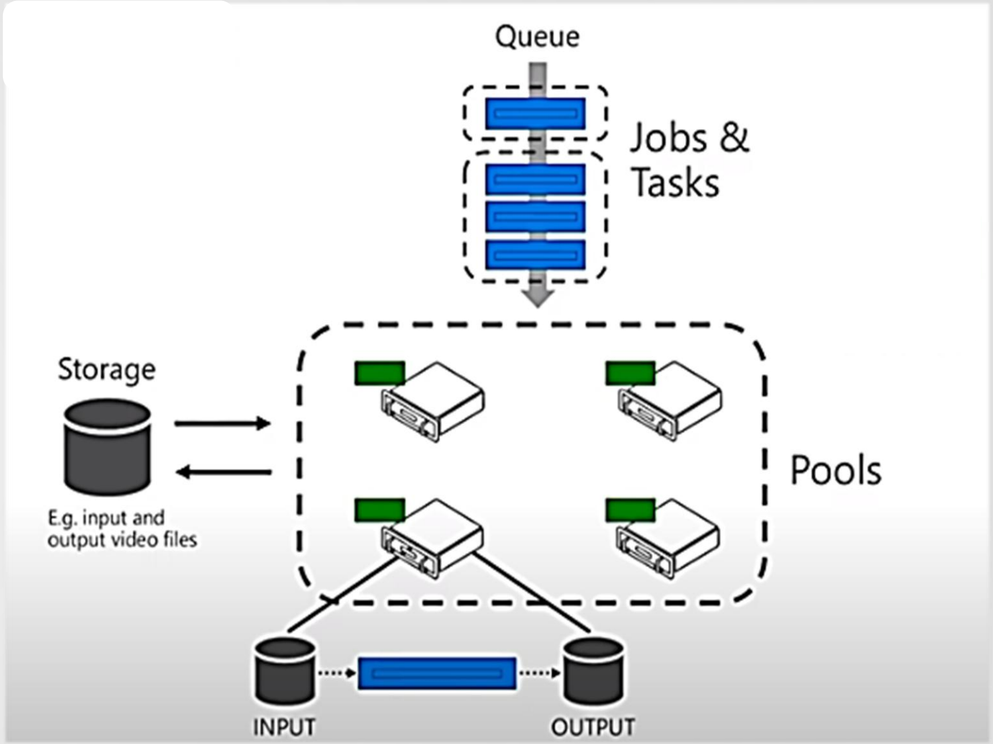
ECO-PTM Cloud Computing

- Processing ECO-PTM jobs
  - preparing
  - submitting
  - processing
  - post processing



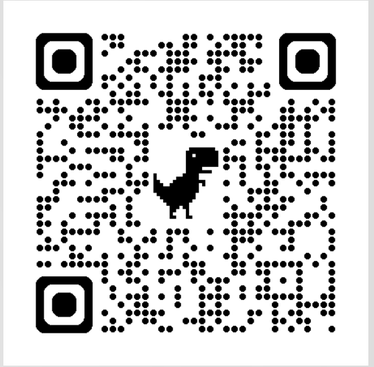
Cloud Computing: Processing ECO-PTM Jobs

- Prepare batch jobs
  - Task:
    - One simulation: tracking a group of particles released at a specific location on a specific date
  - Batch (of tasks):
    - Simulations: a collection of simulations for various release locations and dates
    - One batch per particle type in a scenario
  - Shared tide file
    - One flow field for all simulations in a batch



Cloud Computing: Processing ECO-PTM Jobs

- Submit a job:
  - azure_dsm_batch package:
    - Create batch and storage account clients
    - Manage batch pools
    - Create tasks and batch jobs
    - Submit batch jobs
    - Manage communication between batch and storage accounts
    - Manage storage spaces
  - Configuration file:
    - Override default values, e.g., virtual machines, docker images
    - Specify job processing actions: python script to run simulations



Azure_dsm_batch GitHub site

Cloud Computing: Processing ECO-PTM Jobs

- Post processing:
  - Extract data from DSS file:
    - Extract data after each simulation has been completed
    - Perform within the same simulation batch
  - Assemble all simulation data to a single file:
    - Run the assembly in a separate batch
  - Only download the file assembled

Conclusion

- ECO-PTM studies need a large number of simulations
- Cloud computing makes the studies possible:
  - Offer computing power with flexible scaling capabilities
  - Provide scalable and redundant data storage
  - Use a docker image for all simulations
  - Support libraries to streamline submission processes
  - Deliver cost/benefit options

Acknowledgement

- Nicky Sandhu: DWR Modeling Support Office
- Kijin Nam: DWR Modeling Support Office

DWR Modeling Support Office DSM2 GitHub site:

https://cadwrdeltamodeling.github.io/dsm2/