Performance and Cost Evaluation of Public Cloud Cold Storage Services for Astronomy Data Archive and Analysis

March 2021

Hiroshi Yoshida, Kento Aida

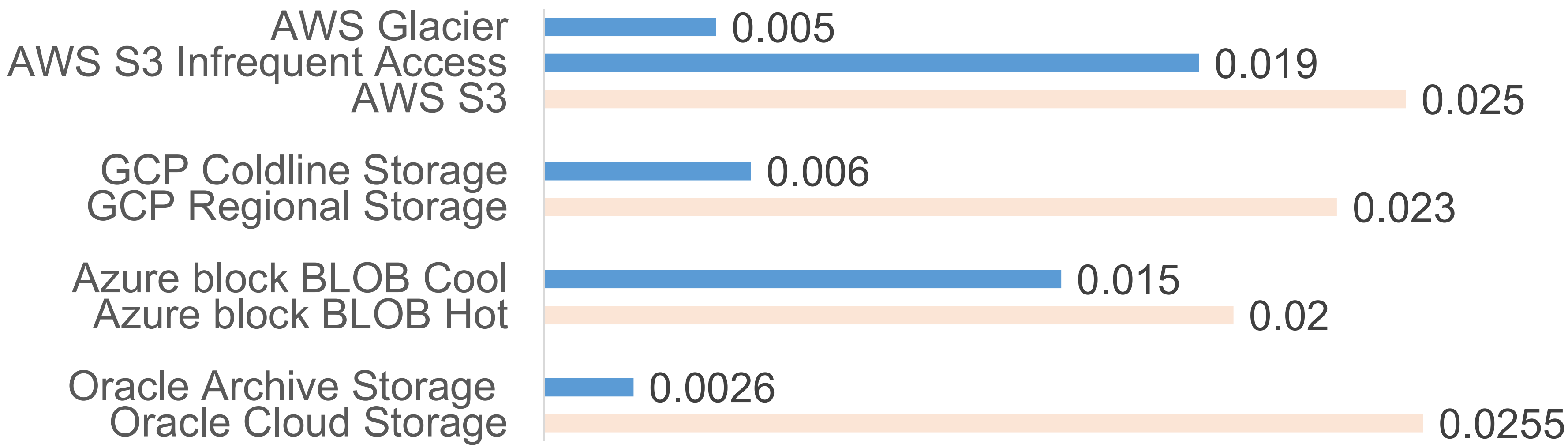
Center for Cloud Research and Development,
National Institute of Informatics

George Kosugi, Eisuke Morita, Takeshi Nakazato, Renaud Miel, Yohei Hayashi

National Astronomical Observatory of Japan

Cloud Cold Storage Services

- Major public IaaS providers provide cold storage services.



Price unit: USD/(GB*month)   [as of September 2020, **Japan region**

- Data store charge is relatively inexpensive compared to standard object storage services (2/3 - 1/10).
- Drawbacks
  - Time consuming restoration process (hours)
  - Extra charge for data retrieval
  - Minimal retention period (30 – 90 days)
  - Limited performance, or extra charge for additional performance
  - Reduced availability

Experiment in Cloud Cold Storage Services

- **Is it possible to adopt cloud cold storage to store a large amount of scientific research data for a long time?**
  - Expectation
    - Data store in clouds: Reduction of TCO and storage management labor
    - Data analyses in clouds: Flexibility to adapt to resource requirement

- Performance and cost of public clouds in scientific applications have not been well studied.
  →Difficult to determine whether cold storage services meet the performance requirements of research applications
  →Difficult to assess the feasibility of storing and accessing research data in cold storage services in terms of cost

- Evaluation using astronomical research data and applications
  - Store the observation and analysis data of the ALMA telescope project
  - Port the data archive system “NGAS” to AWS
  - Run common analysis applications “CASA” on a variety of instances to analyze the retrieved observation data inside cloud

Data and Application Used in the Experiment

■ Archive data

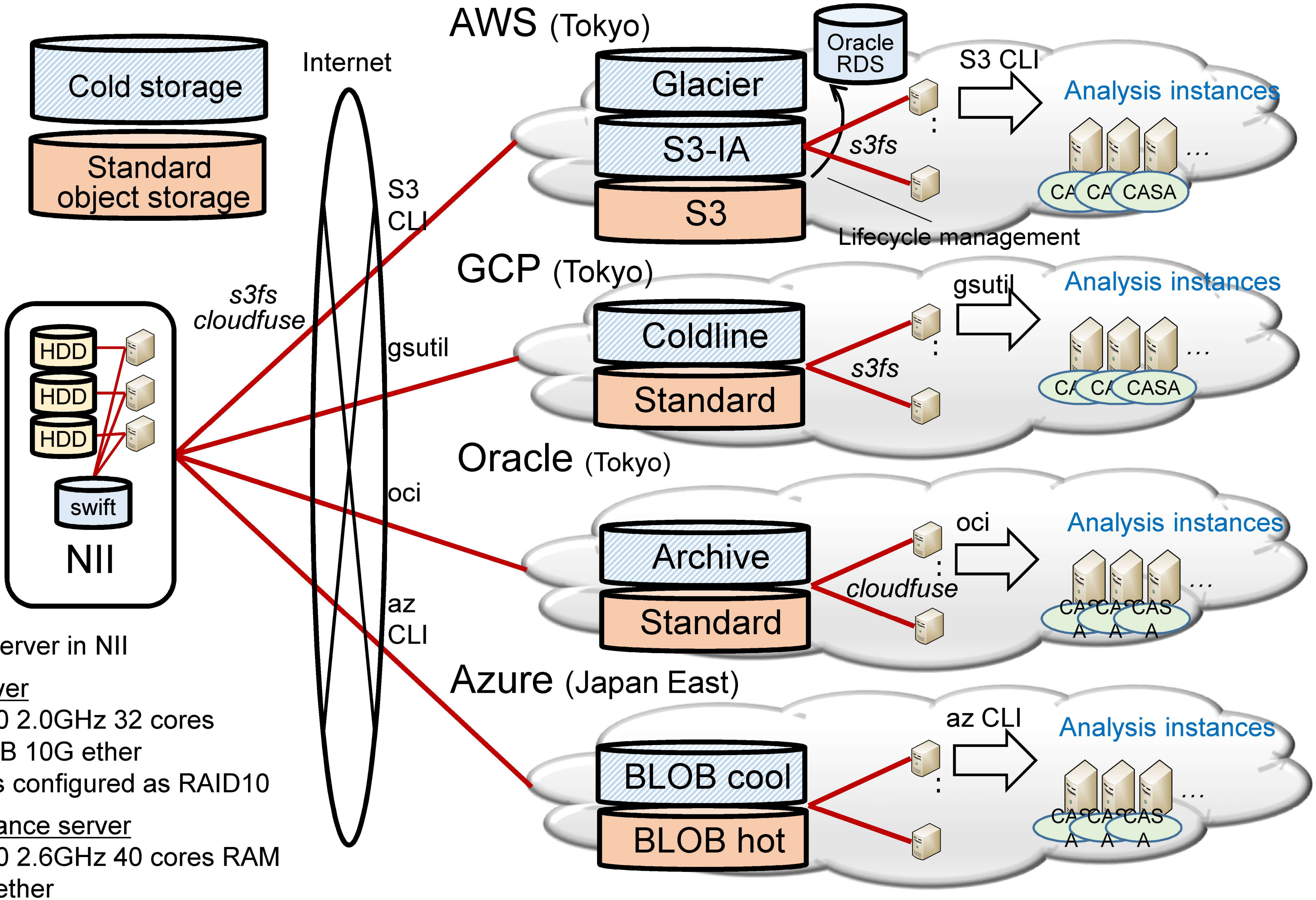
Data contents	ALMA radio telescope Observation/analysis data provided by National Astronomical Observatory of Japan (NAOJ)
Quantity	**58.5TiB, 1,380,000 files**
Size	Average 44MiB (falls between smaller than 1MiB and larger than 100GiB)
Application	Archive management: NGAS (Next Generation Archive System)

■ Analyzed data

Category	Number of datasets	Estimated analysis time	Size (GiB)	Number of included files
Small	3	≃ 1 hour	0.4～ 0.6	99～ 267
Medium	3	≃5 hours	2.2～ 3.9	240～4,000
Large	3	≃1 day	9.0～26.1	2,421～3,879
XLarge	1	≫ 1 day	87.3	456

■ Application: CASA (Common Astronomy Software Applications)

Experiment Environment



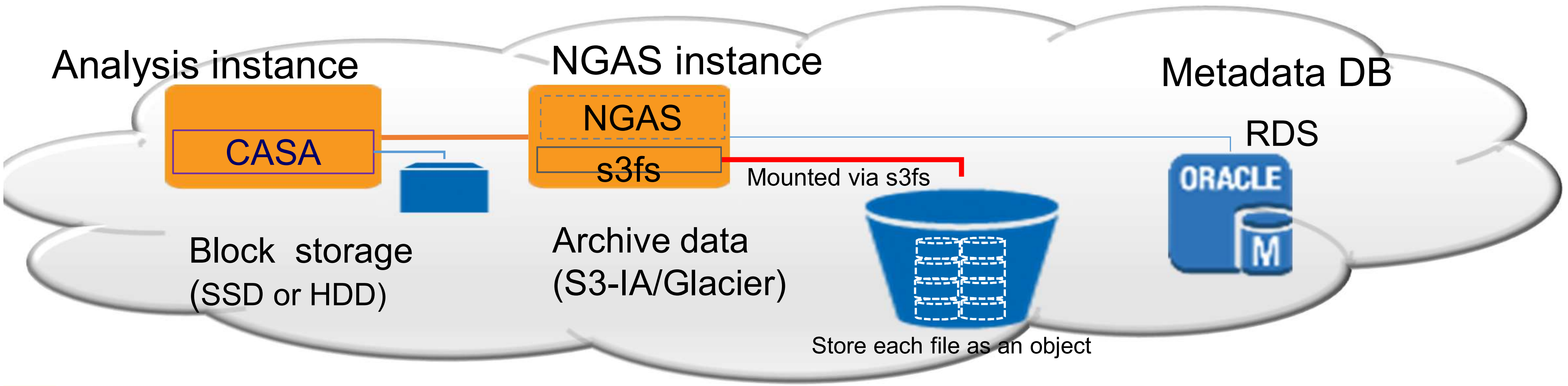
On-premise server in NII

Standard server
Xeon E5-2650 2.0GHz 32 cores
Memory 64GiB 10G ether
External disks configured as RAID10

High performance server
Xeon E5-2660 2.6GHz 40 cores RAM
128GiB 10G ether
External RAID connected via iSCSI

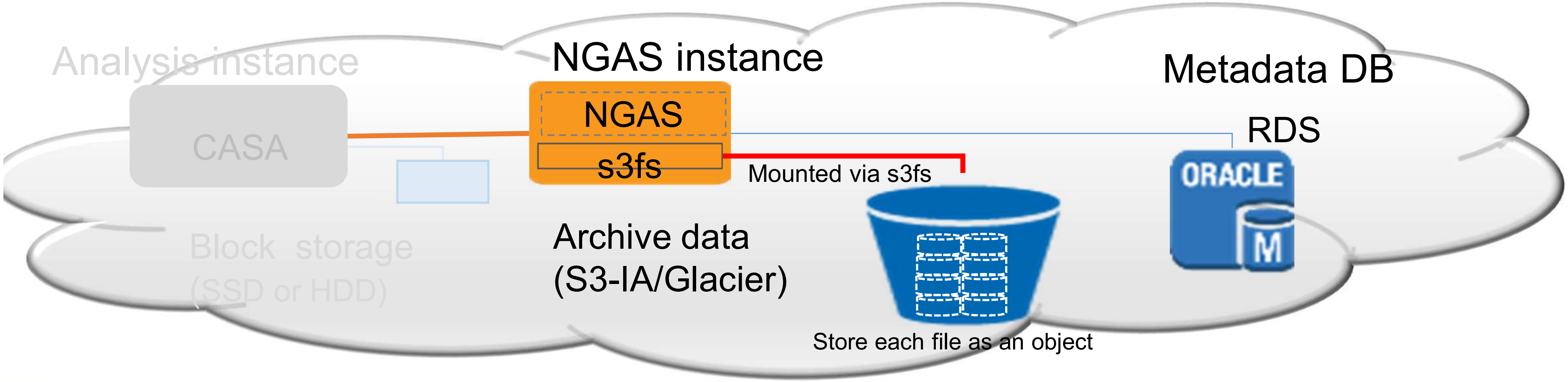
Target of experiments and System Configuration

- First step (FY2017 – FY2018)
  - Evaluate performance, cost, and manageability by porting archive management system NGAS to AWS and storing archive data in S3-IA and Glacier
- Second step (FY2019-)
  - Analyze observation data on public cloud instances to evaluate performance and cost
  - Investigate optimal selection and usage of instances

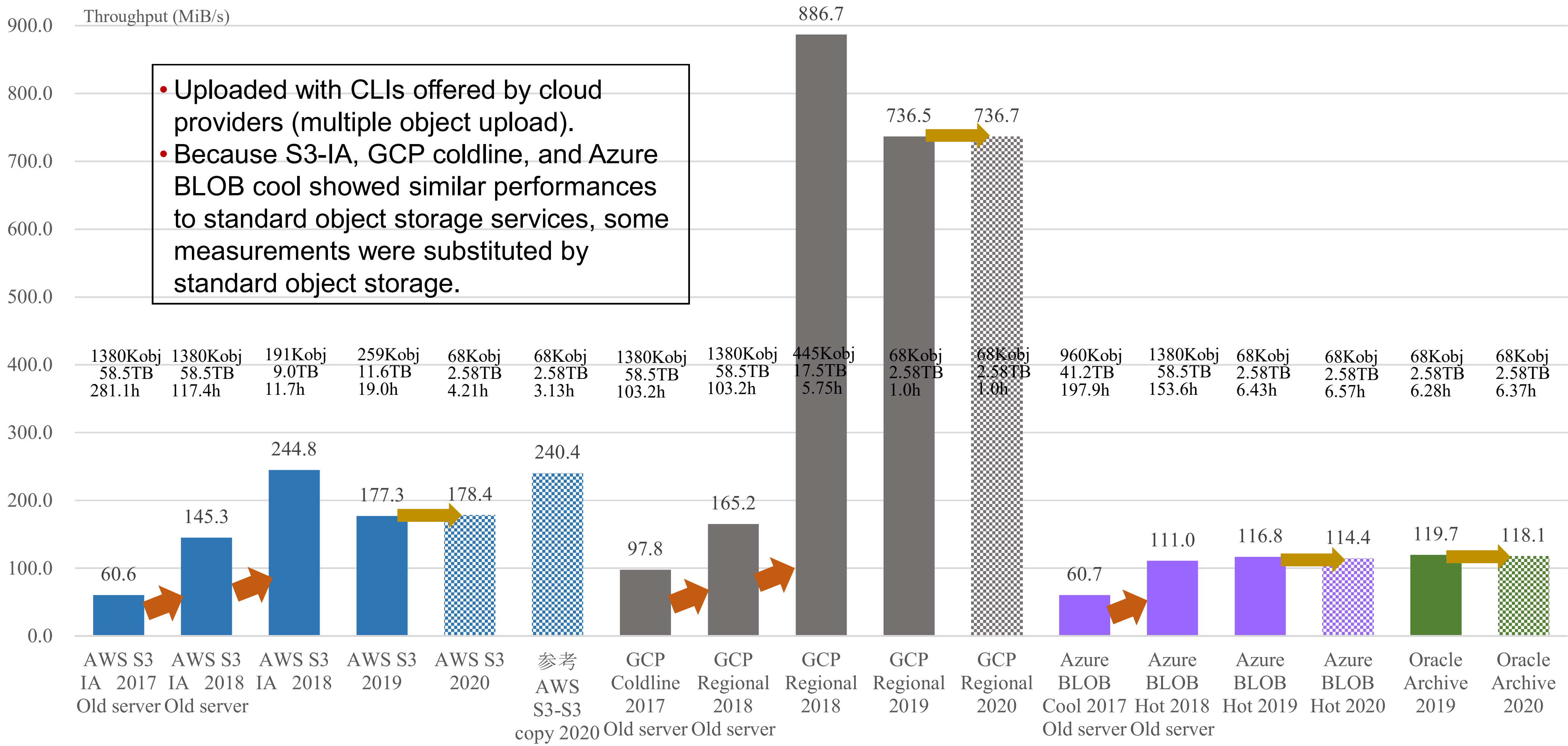


First Step Experiment

- First step (FY2017 – FY2018)
  - Evaluate performance, cost, and manageability by porting archive management system NGAS to AWS and storing archive data in S3-IA and Glacier
- Second step (FY2019-)
  - Analyze observation data on cloud instances on the instances of public cloud services to evaluate performance and cost
  - Investigate optimal selection and usage of instances

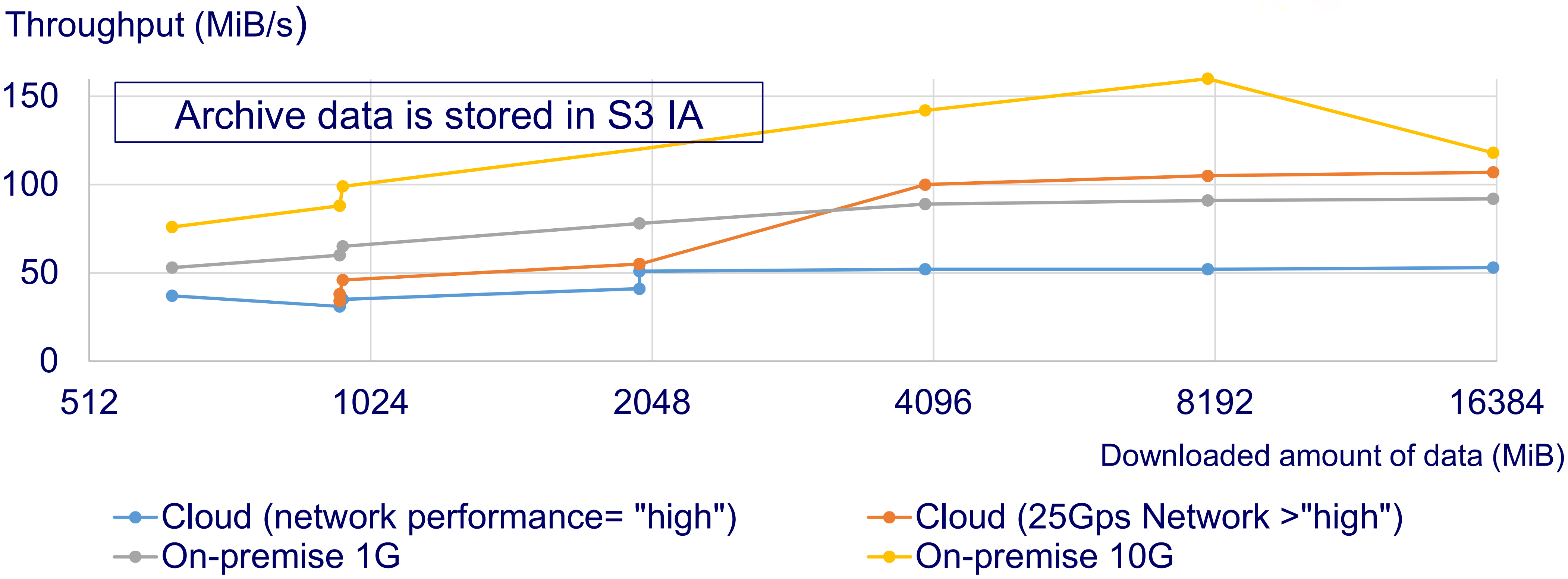


Performance of Uploading Archive Data via Internet



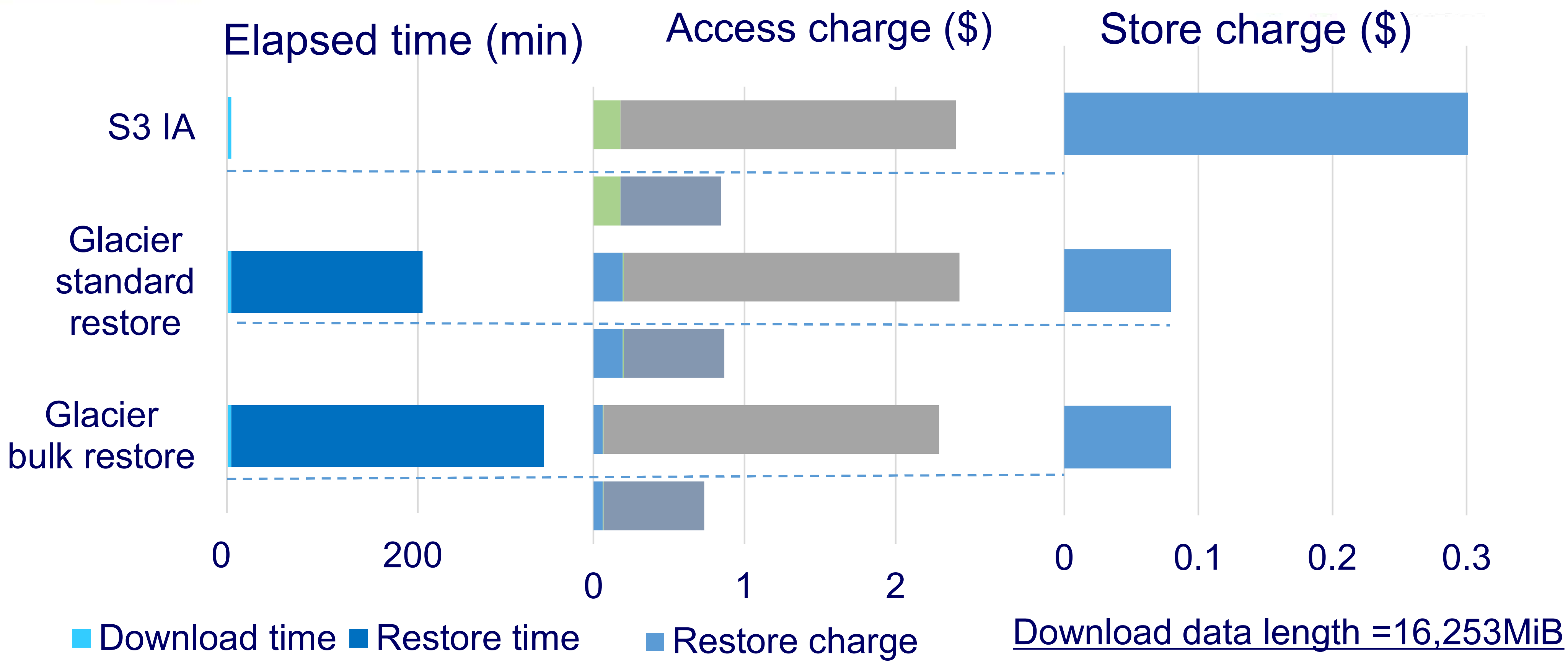
- The performance requirements can be generally fulfilled.
- Significant Performance improvement between 2017 and 2018
  - No performance increase after 2018

Data Download Performance of NGAS



- Performance of cloud NGAS is lower than that of on-premise NGAS
  - Possible causes
    - Performance difference between S3 and on-premise storage
    - Overhead of s3fs
    - Too small RDS instance
- Throughputs can be increased by increasing network performance
  → Practical performance can be achieved with appropriate sizing.

Download Performance and Cost of NGAS on AWS

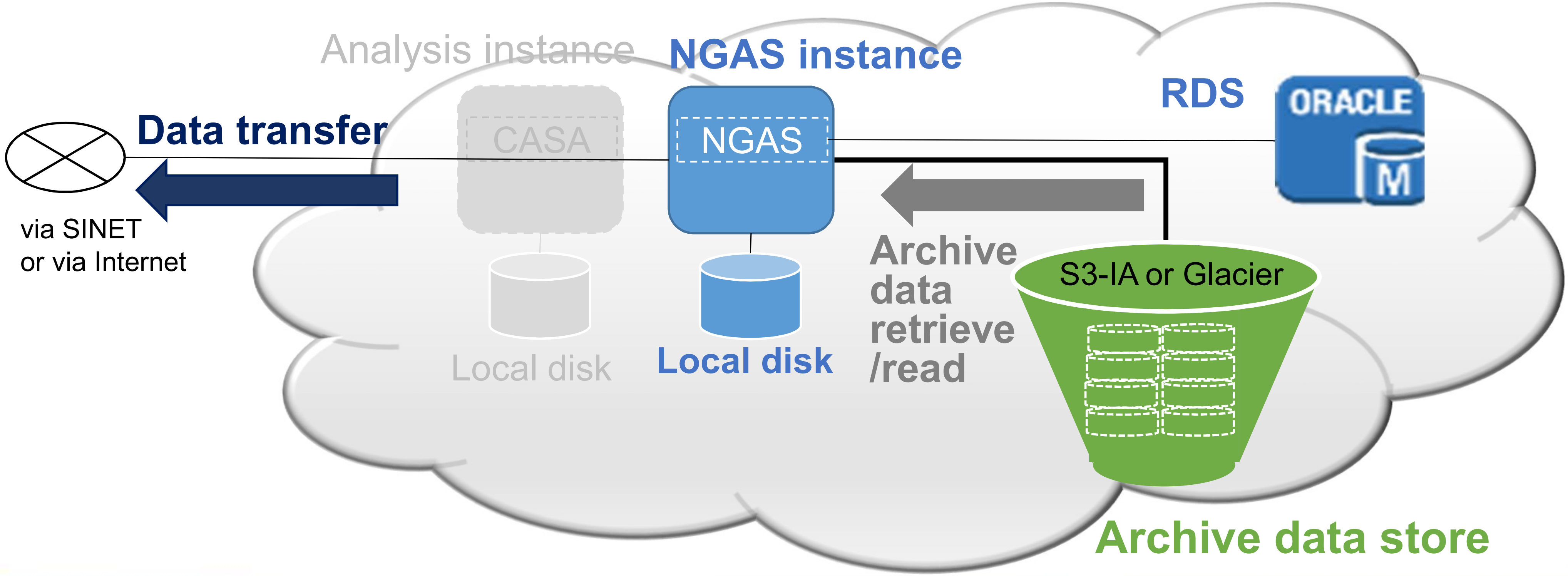


✓ S3 IA: actual measurement
✓ Glacier performances are estimated from manual restore operations

- A retrieval from cold storage requires read cost and retrieve cost.
- Restore time and cost are required in Glacier, instead of low store cost.
- Large egress transfer costs are required.

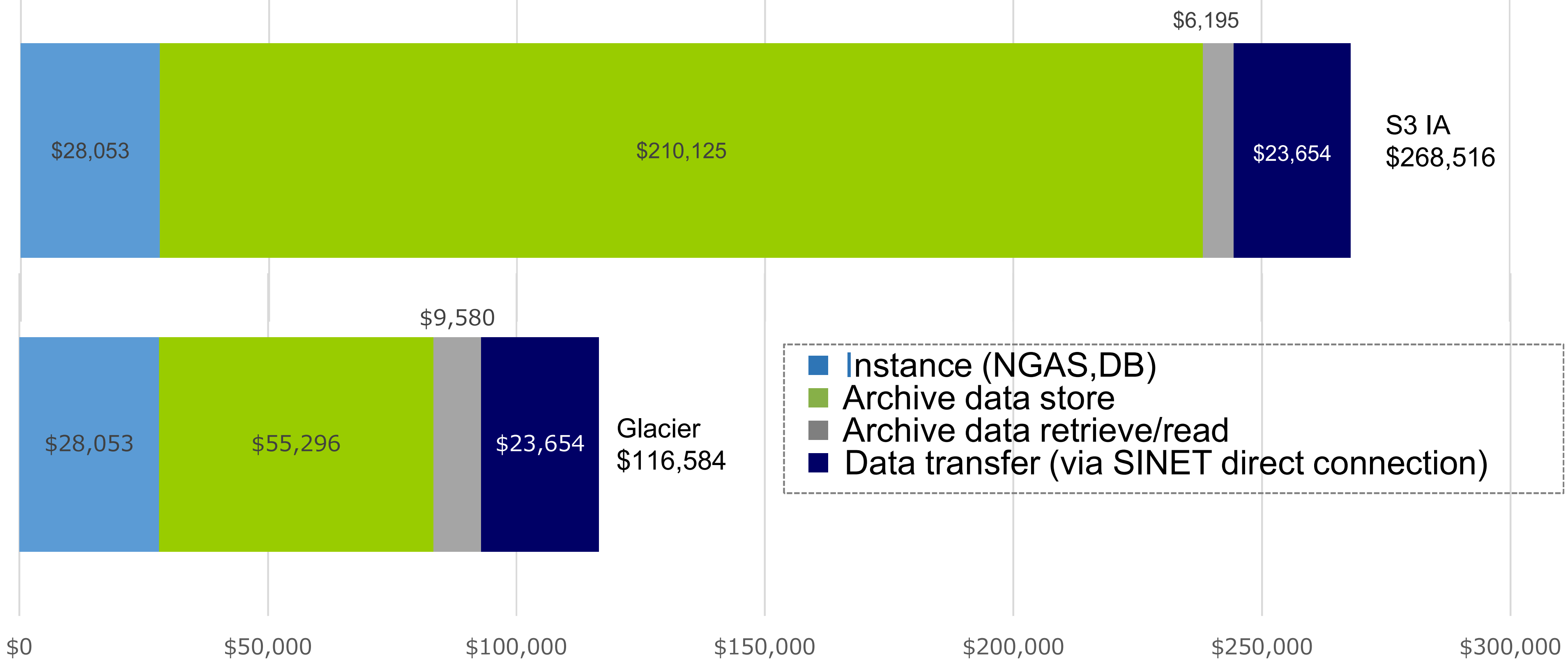
Cost Estimation Model of NGAS on AWS

- Parameters
  - Charge for NGAS and RDS instances
    based on instance types and local disk capacity
  - Capacity of archive data store
  - Amount of archive data retrieval/read
    - Read request charge can be converted
      to the proportional charge to the retrieval amount approximately .
  - Amount of egress data transfer



Yearly Cost Estimation of NGAS on AWS

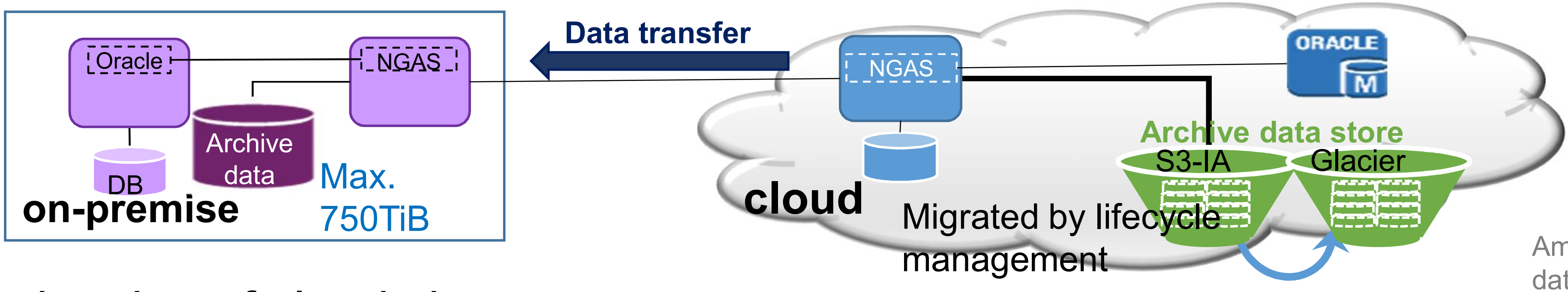
- Store 900TiB archive data in S3- IA / Glacier and Retrieve 550TiB/year



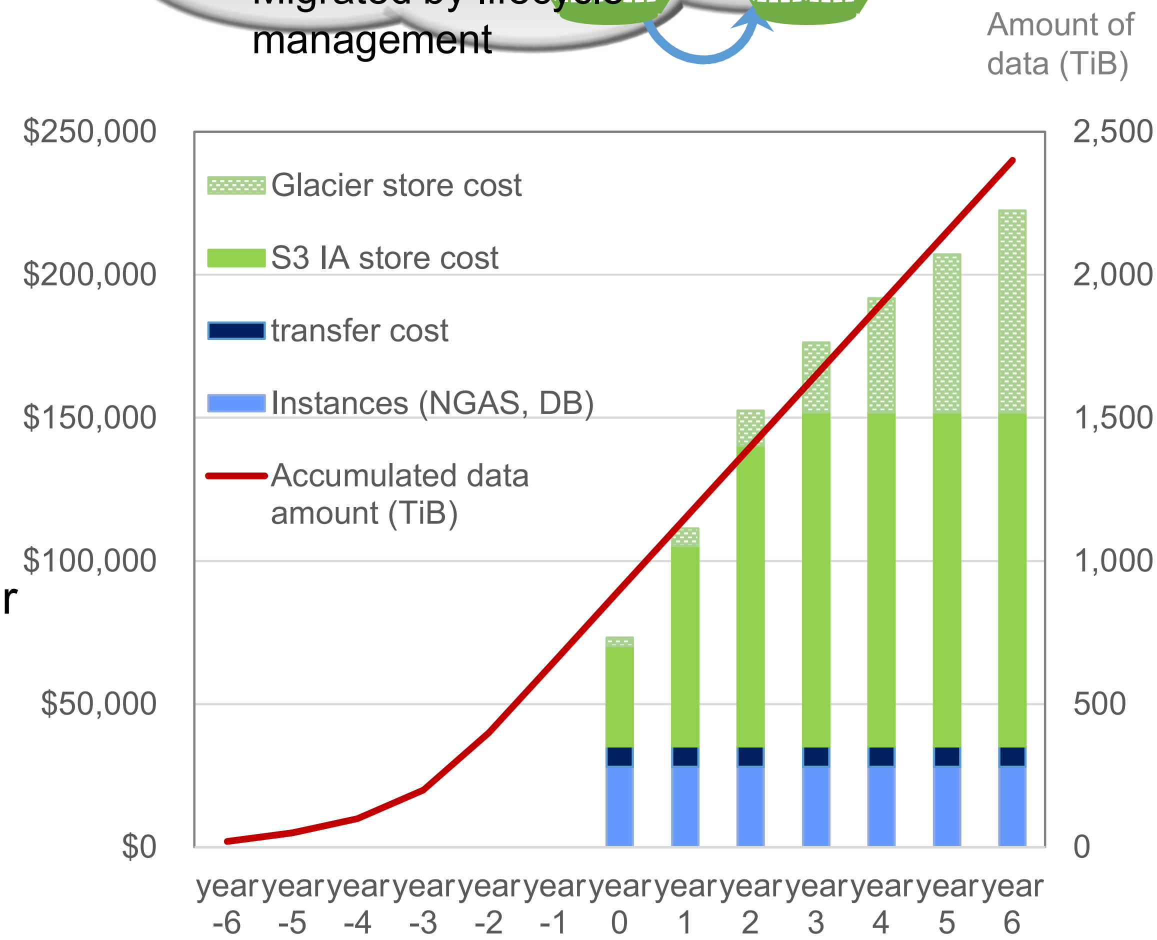
- Store cost of Glacier is less than one third of that of S3 IA
  - Tiered storage including Glacier, S3 IA, and on-premise archive can be a solution to balance cost saving with the disadvantages of retrieval time.
- Relatively high egress transfer cost
  → Motivation for the second step experiment (analyses in cloud)

Hybrid Configuration of Archive System

- Tiered storage including Glacier, S3 IA, and on-premise storage

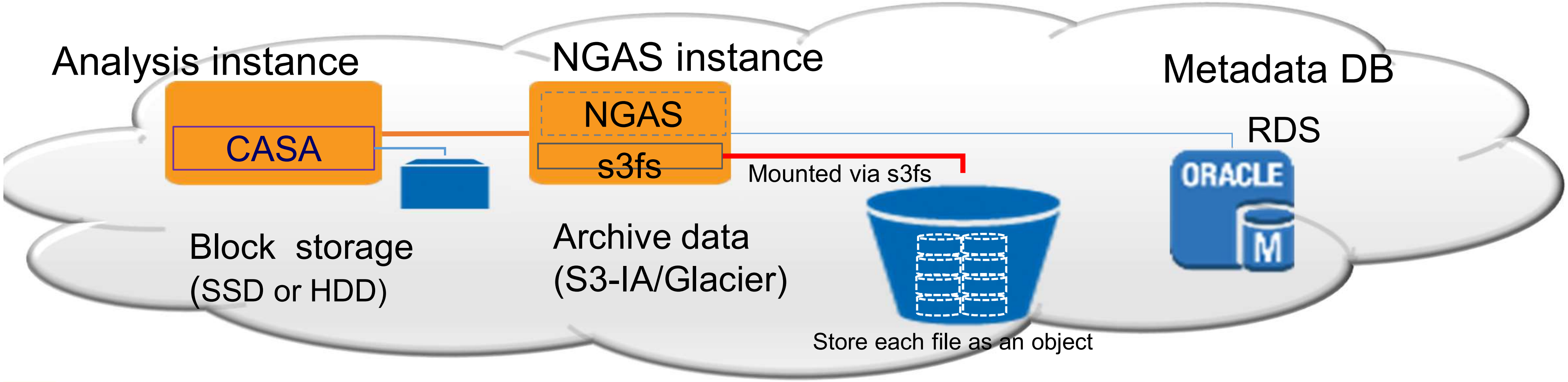


- Estimation of cloud charge in a hybrid configuration
  - Amount of archive data is 900TiB in year 0; increases 250TiB/year
  - Datasets are stored in on-premise storage for 3 years, S3 IA for 2 years, and then migrated to Glacier.
  - Amount of downloads is 550TiB/year; 20% from S3 IA and 10% from Glacier (**i.e. only 10% of downloads require additional 200 minute restore time**).



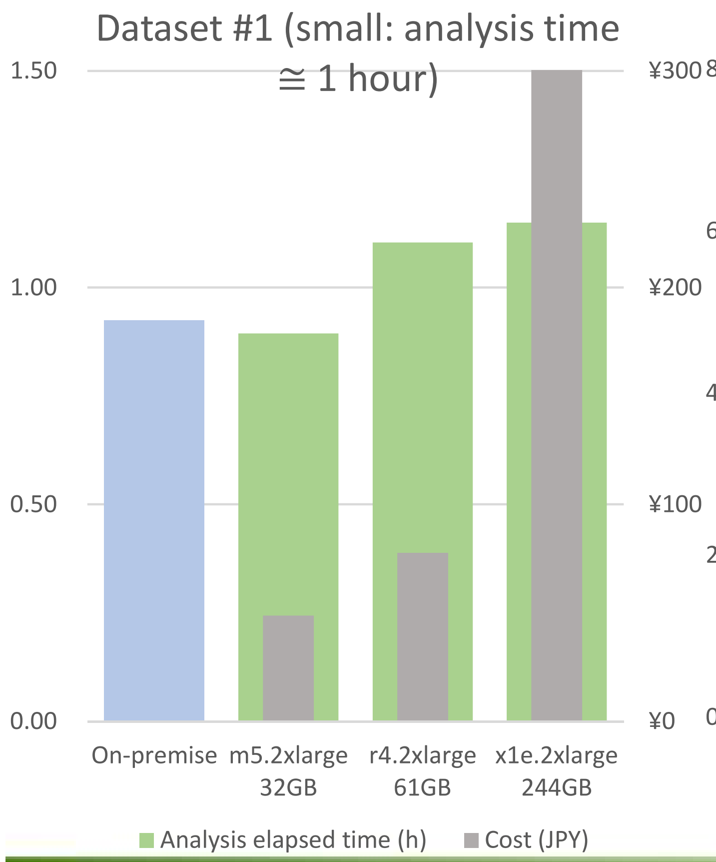
Second Step Experiment

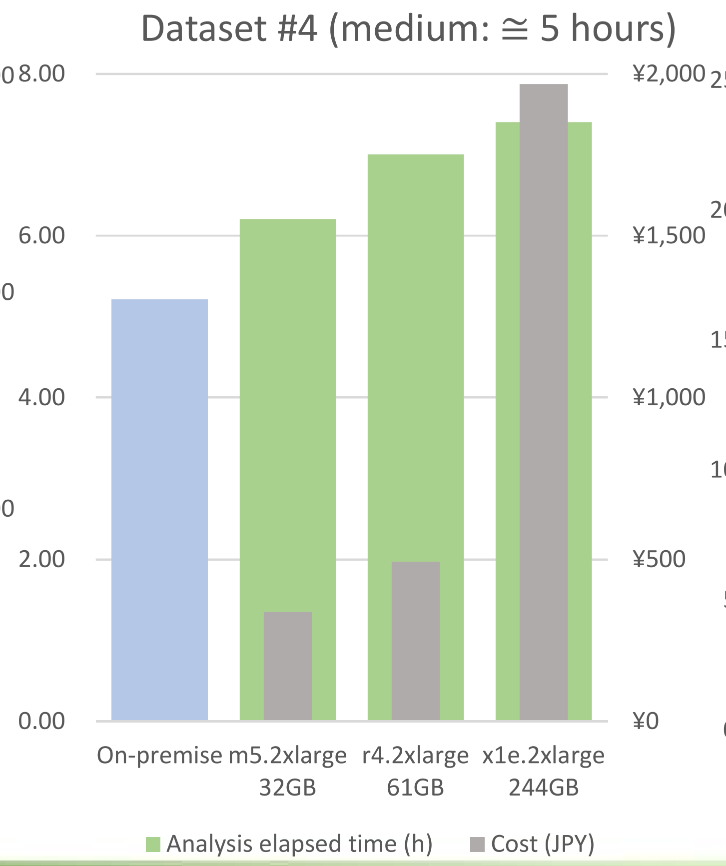
- First step (FY2017 – FY2018)
  - Evaluate performance, cost, and manageability by porting archive management system ”NGAS” to AWS and storing archive data in S3-IA and Glacier
- Second step (FY2019-)
  - Analyze observation data on public cloud instances to evaluate performance and cost
  - Investigate optimal selection and usage of instances

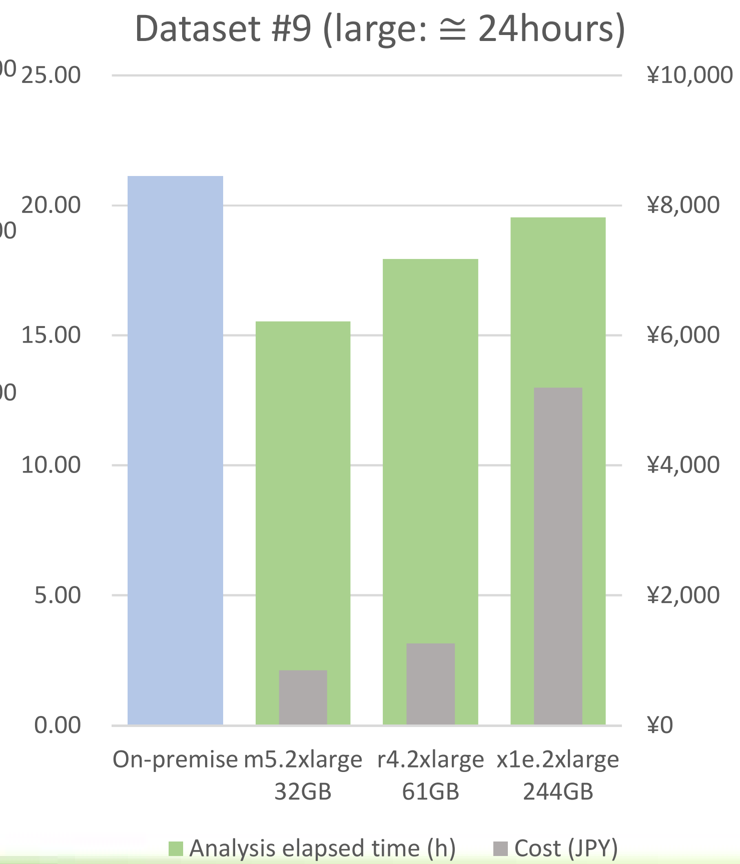


Elapsed Time and Cost of Data Analyses in AWS

- Enough performance for practical use
- Performance differences are not caused by the differences of memory capacity but by the differences of instance generation.
  - The analyses of chosen datasets don't require large memory capacity.
  - 61GB and 244GB memory instances belong to the older generations.



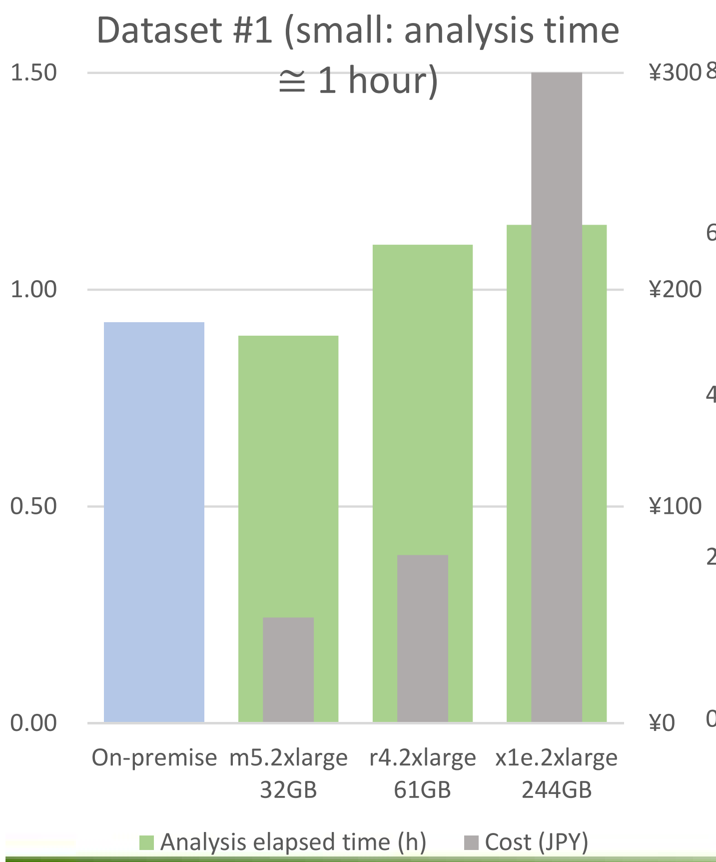


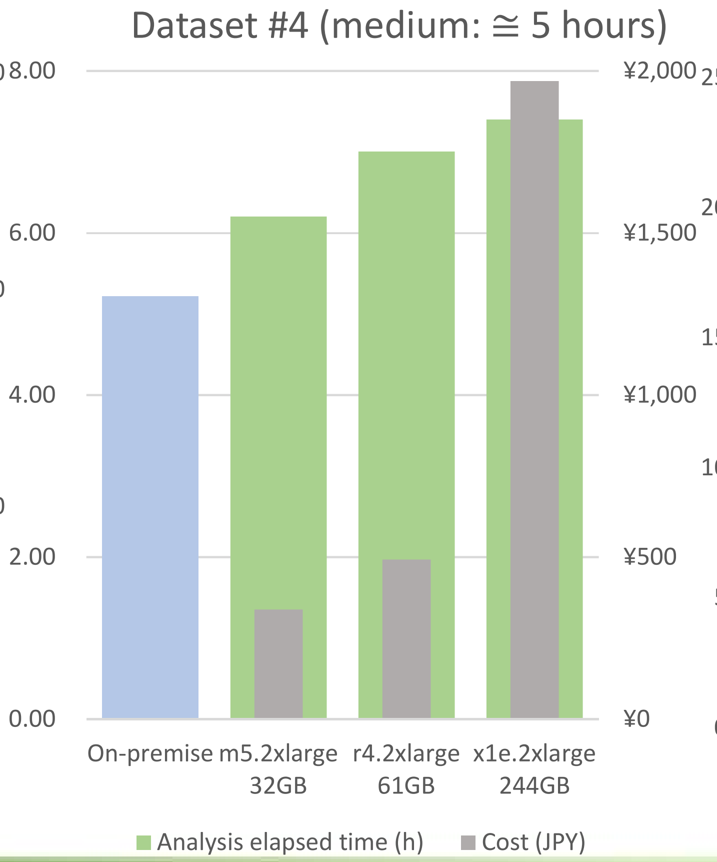


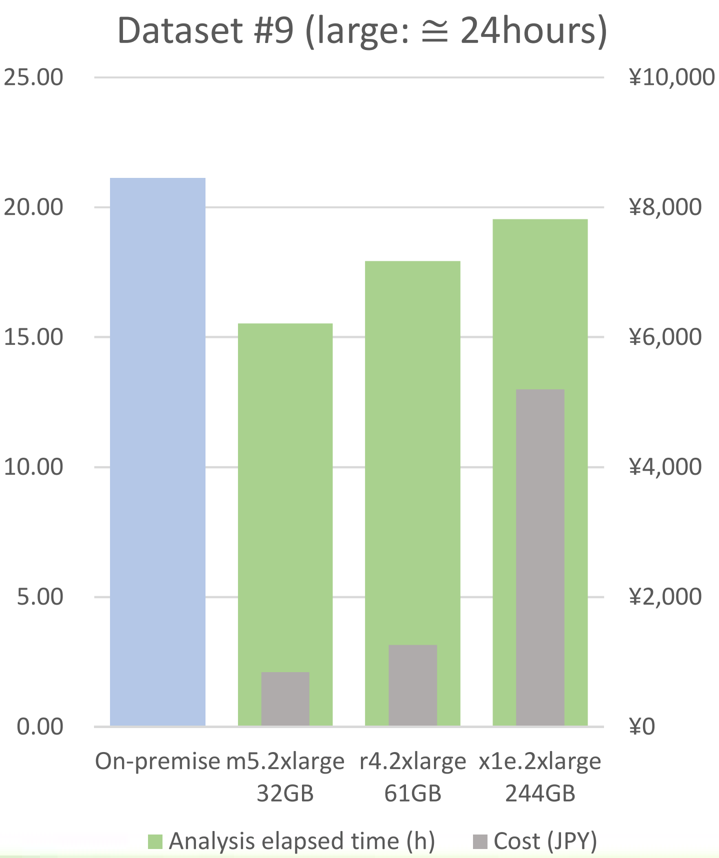
Elapsed Time and Cost of Data Analyses in AWS

- Instances with large capacity memories are expensive.
- A new generation instance is usually less expensive than old generation one with the same specification.

→Adopt new generation instances whenever possible
Select instances with appropriate memory capacity



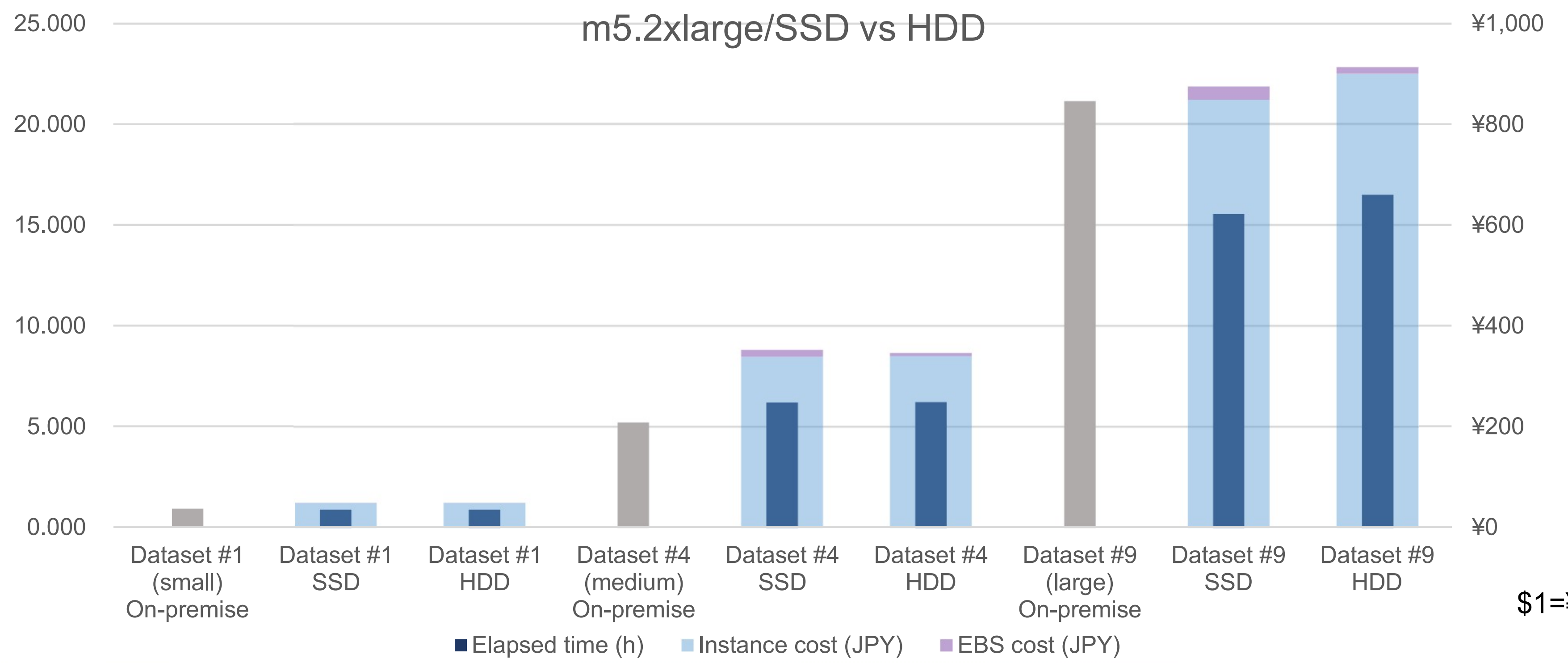




Cost Optimization Points (1): SSD vs HDD

- SSD and HDD can be chosen as block storage media.
  - Analysis data, working data, final results are stored in block storage.
- Up to 20% reduction of elapsed time by using SSD
- The increase of instance charge caused by the increase of elapsed time is higher than the cost difference between SSD and HDD.

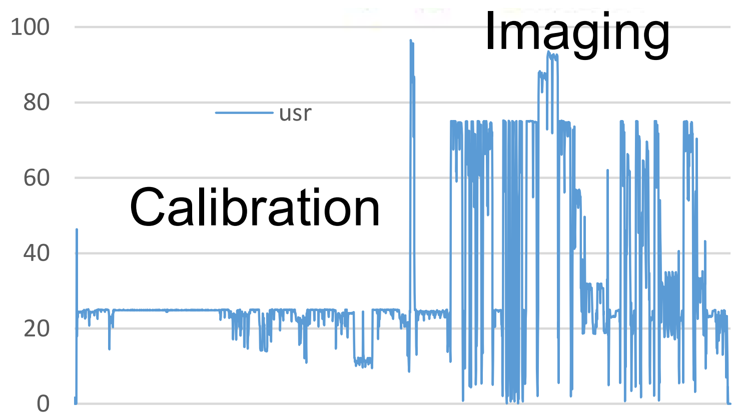
  →In CASA cases, SSD is always better.



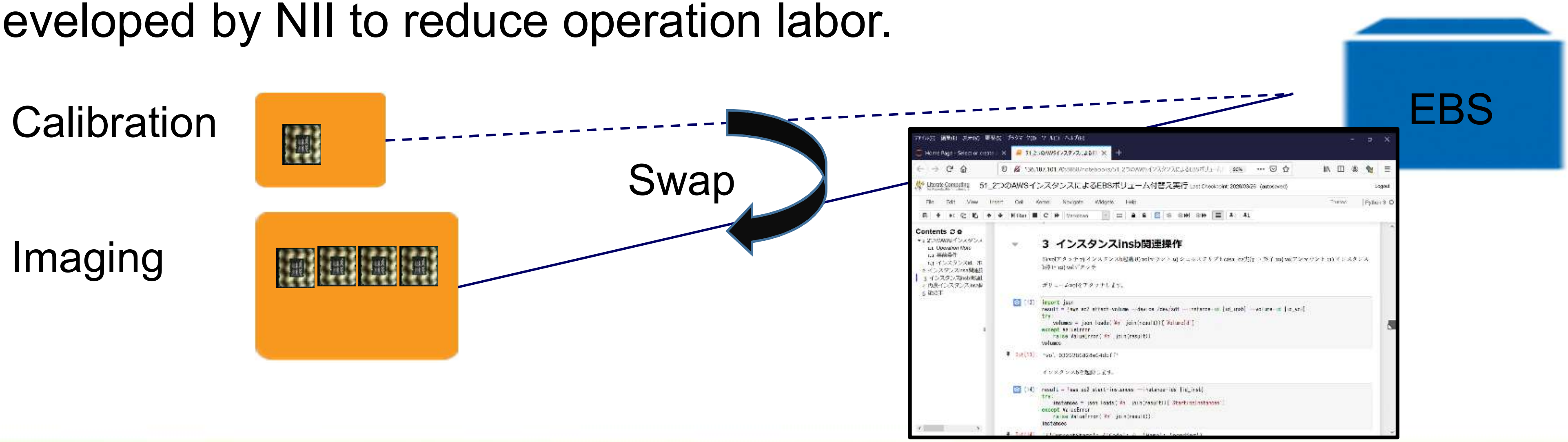
$1=¥110

Cost Optimization Points (2): Instance Swap

- Observation
  - Calibration is performed mostly on 1 core.
  - Imaging is performed on more cores.
  - Cloud block storage volumes can be dynamically attached to/detached from an instance using CLI and/or API

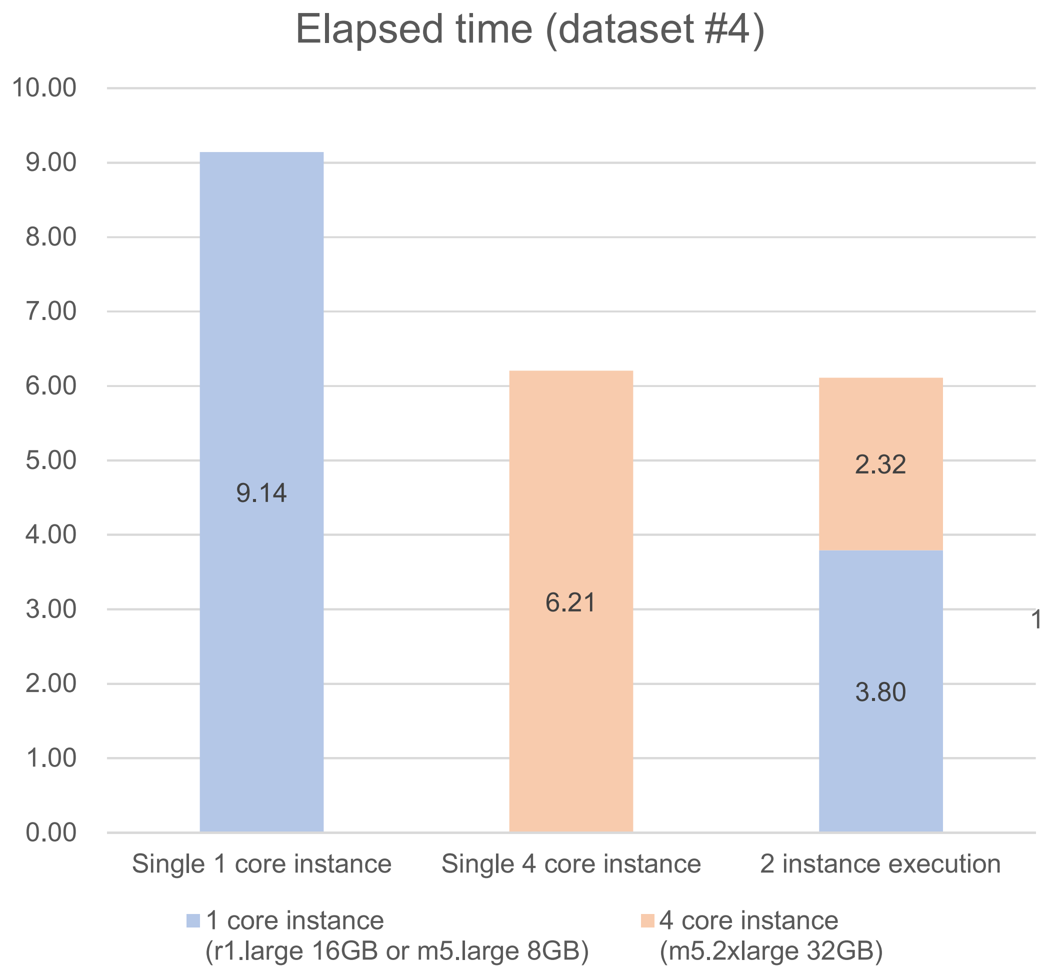


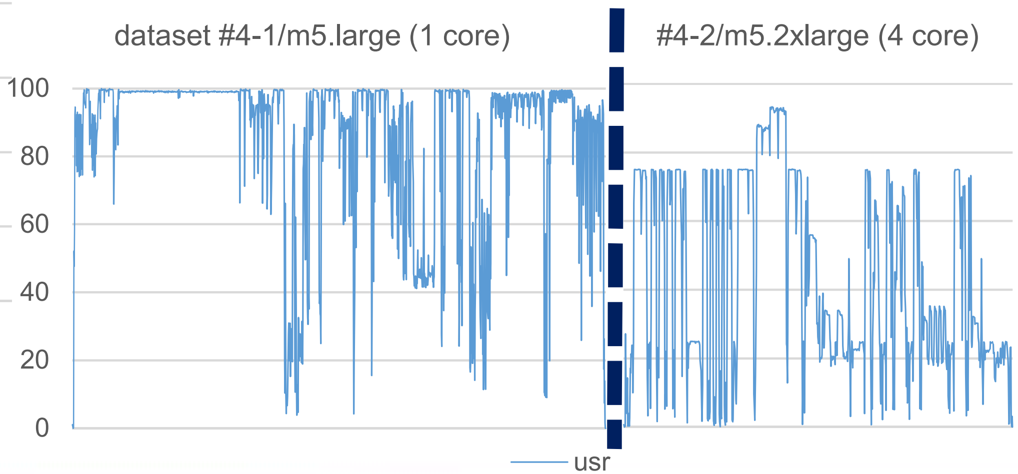
- Instance swap
  1) Perform calibration on a 1-core instance
  2) Attach the volume including working data to a multi-core instance and perform imaging
- The workflow is automated by using literate computing technology developed by NII to reduce operation labor.



Advantage of Instance Swap

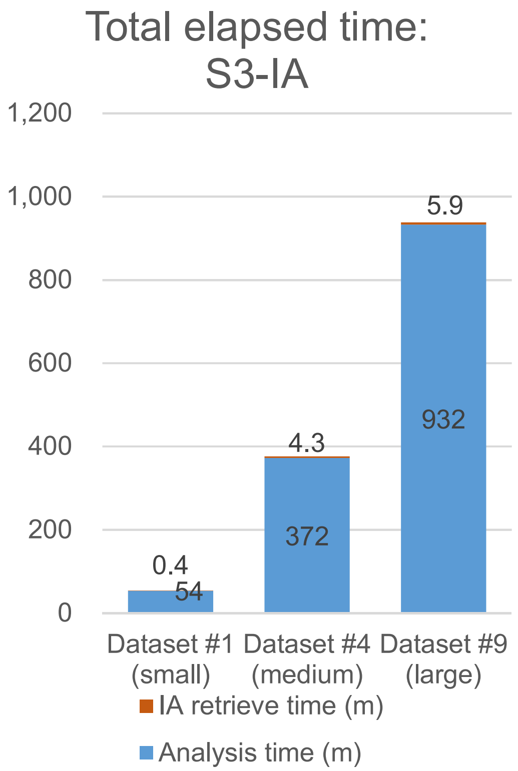
- Example: dataset#4
  46% cost reduction as the elapsed time is retained
  or 49% higher performance with additional 12% cost

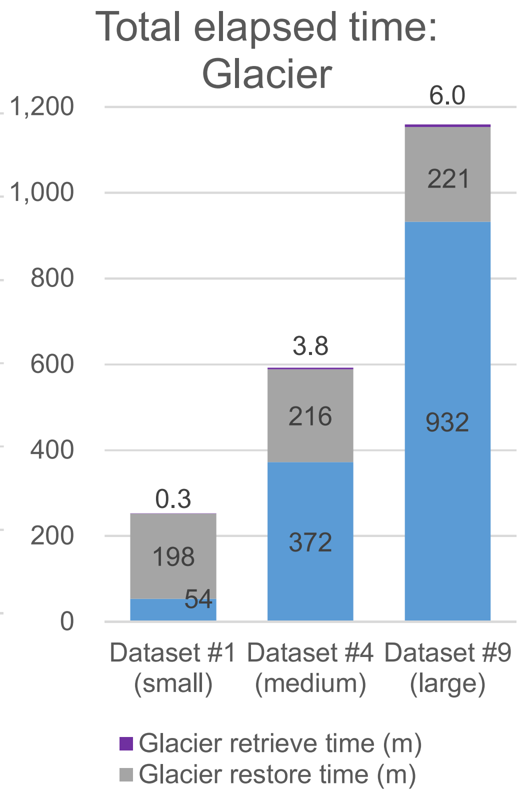


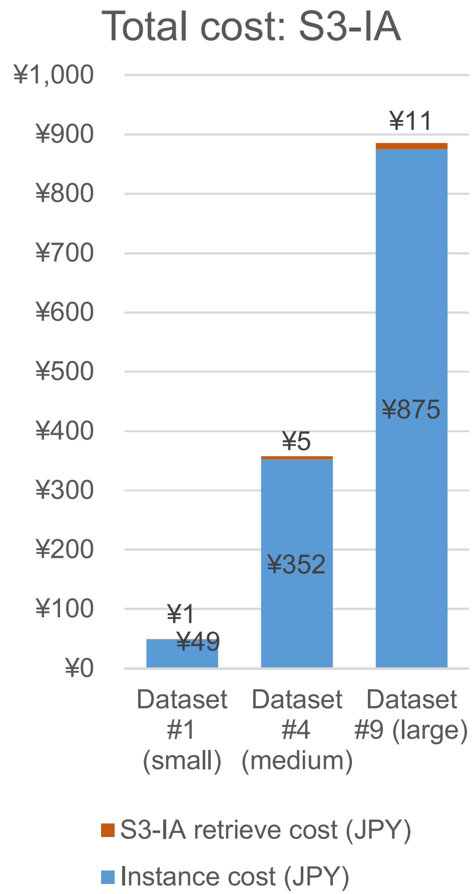


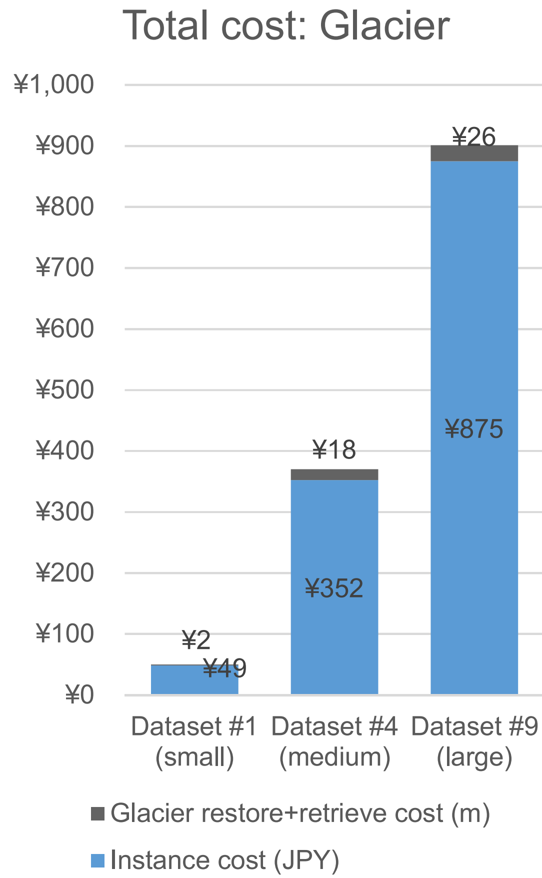
Performance and Cost from Archive to Analysis

- Retrieving and analyzing datasets archived in S3 IA and Glacier
  - Total elapsed time
  - Total cost







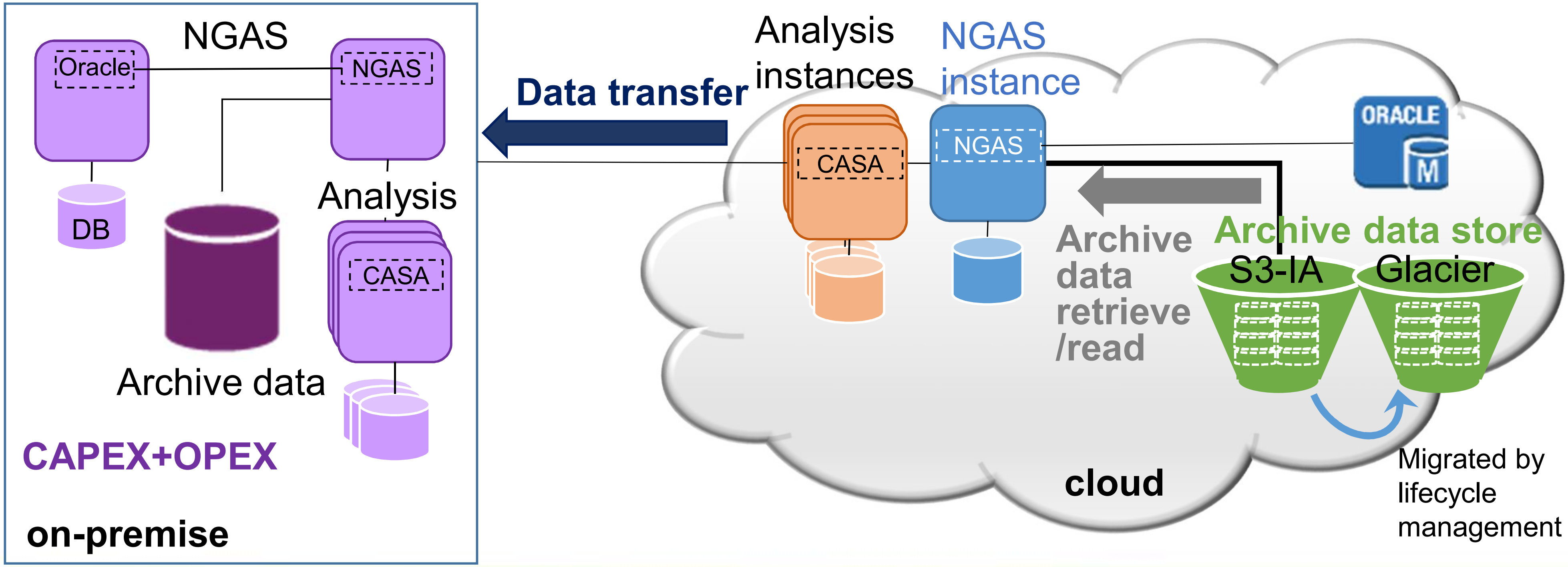


$1=¥110

- Glacier is advantageous to store cost in spite of long elapsed time and extra charge
  - 200 minute restore time before analysis
  - Restore operation charge

Hybrid Configuration Including Analysis System

- Optimization points of archive location and analysis location
  - Access frequency of datasets based on the age of data
  - Performance (1) data retrieval (2) data transfer (3) analysis
  - Cost ▪ Cloud charge (1) data store (2) data retrieve/read (3)egress data transfer (4) analysis instances
    ▪ CAPEX+OPEX of the on-premise system
  - Resource deployment flexibility in cloud
  - Usability and automatability of workflow



Summary

- Practical results on storing and analyzing scientific research data in cloud are acquired.

- Practically acceptable performance of data access can be achieved with appropriate sizing of resources and tuning of the system.
  - Inexpensive cold storage services (such as Glacier) are significant options in terms of store cost, although they require hourly restore time before accessing data.
    - Can be mitigated with an appropriate tiered storage architecture

- The cost estimation model enables to estimate total cost in cloud.
    - Data store and egress transfer costs are major parts of the total cost of the archive system.
    - Data retrieval costs of cold storage services have little effect.
  - The model is also capable to estimate cost on a hybrid system organized by clouds and on-premise systems.

Next Steps

- Establish methodology to estimate required resources such as number of cores, CPU usage patterns, memory capacity, and block storage capacity based on the dataset characteristics to choose optimal instance
- Optimization of the hybrid system configuration and archive/analysis locations
- Investigate optimizations of applications and usages of cloud services
  - Improve mapping between files and objects to accelerate restore operation
    e.g., 1 file to 1 object → multiple files to 1 object
  - Adopt cloud-native object storage API to improve performance
  - Instance swap
- Share the practical information and the best practices of cloud usage with researchers of other scientific field

Inter-University Research Institute Corporation / Research Organization of Information and Systems

National Institute of Informatics