



IBS 리서치솔루션센터 HPC 사용자간담회

2025/11/26

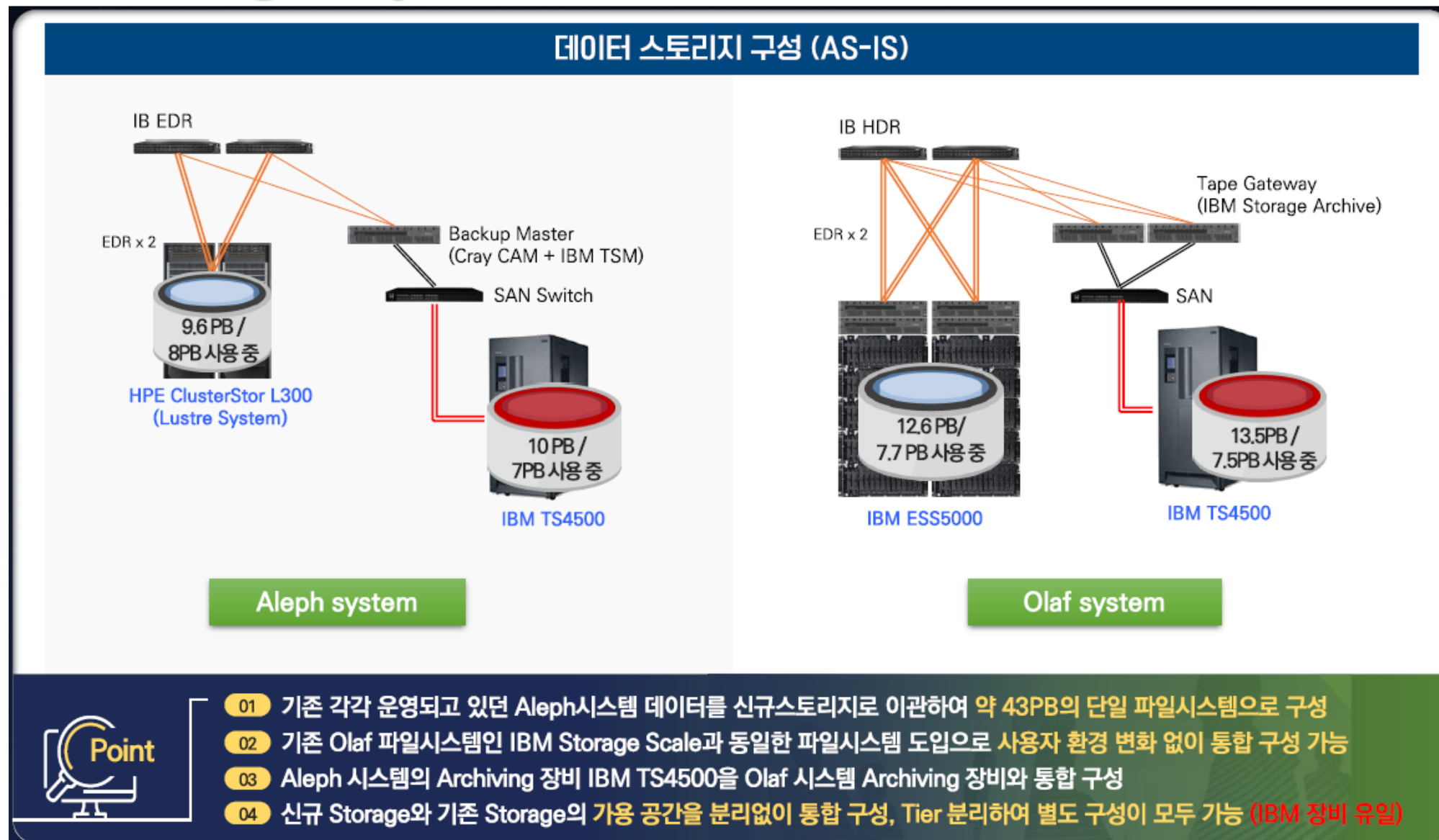
허무영

RSC 인프라운영팀

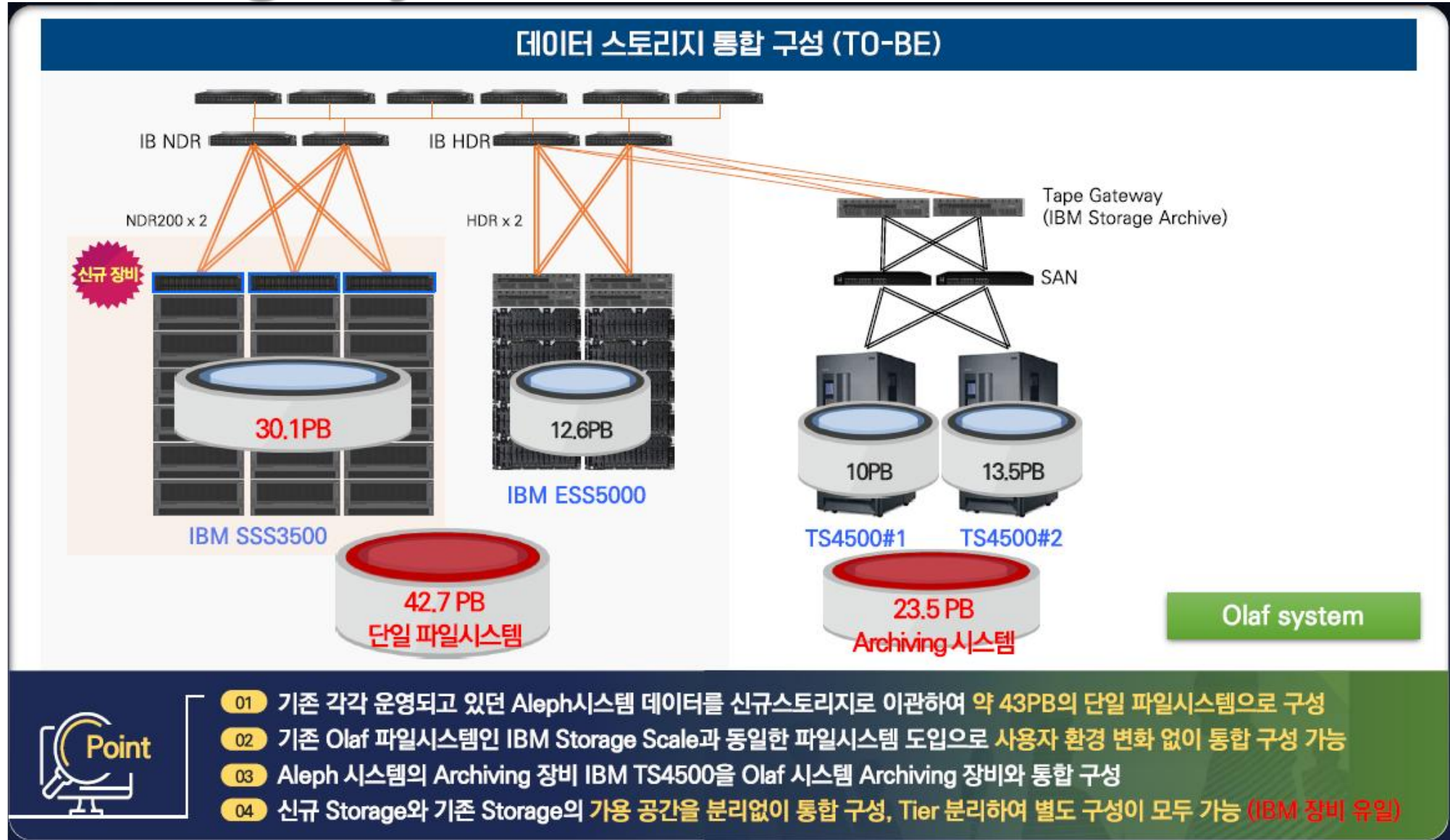
논의사항

- 신규 구축 데이터 스토리지 시스템 안내
- LLM API 활용 및 Chatbot 서비스 안내
- HPC 홈페이지 활용 안내
- Globus 활용 데이터 전송 안내
- SLURM 스케줄러 자원 할당 정책 (**활용률 및 활용수요에 따른 탄력적 배분**)
 - Olaf-cu : normal/long
 - Olaf-c: normal_cpu/large_cpu/core_s/core_m/core_l
 - Olaf-g: AIP/mig_1g.10gb/mig_3g.40gb
- 기타 초고성능컴퓨팅 자원 활용 방안 논의 및 질의응답

New Storage System



New Storage System



LLM Chatbot Service

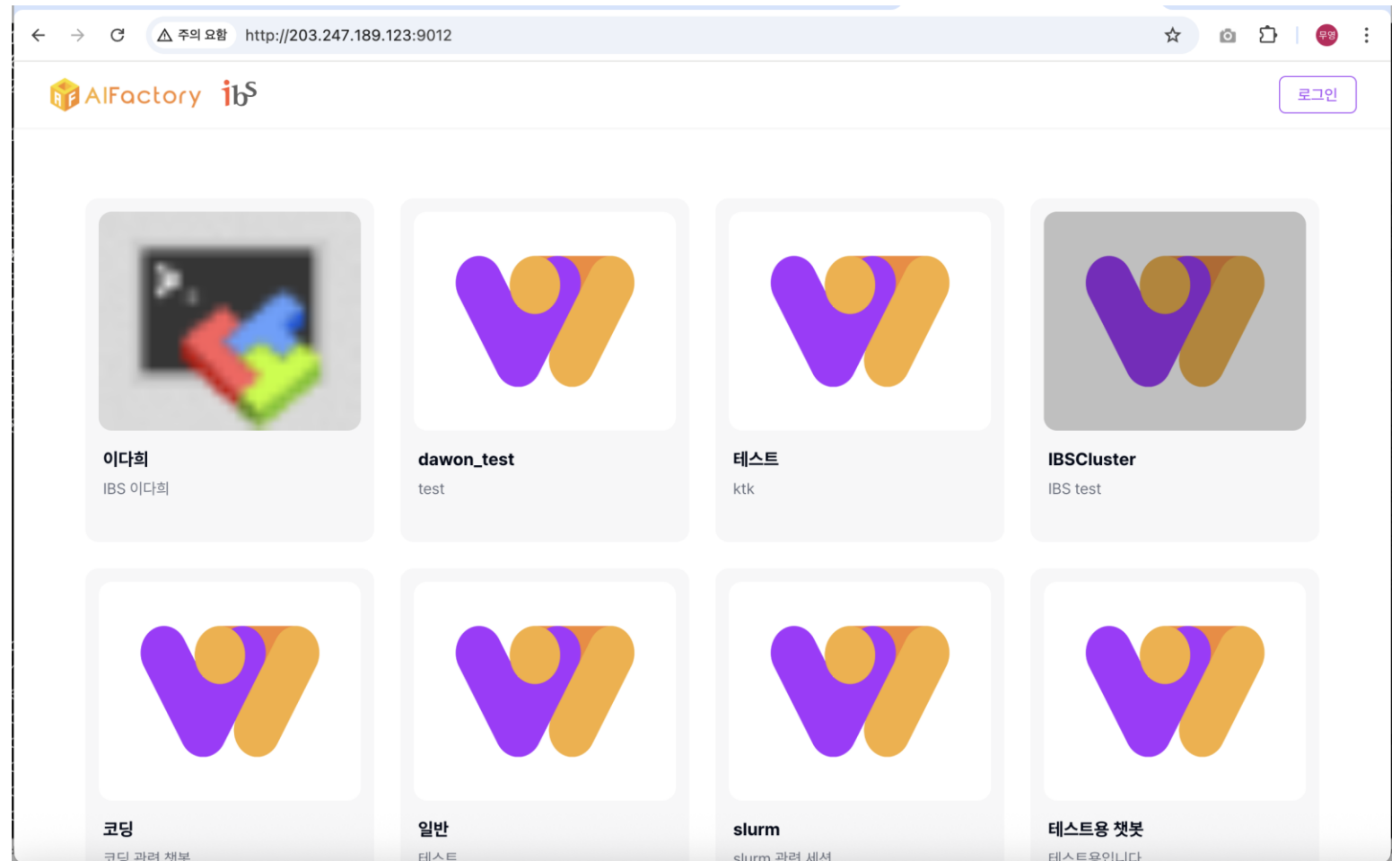
LLM: **Meta Llama-3.1-405b**

LLM inferencing: vllm 0.5.5

Chatbot: AIFactory WERT

Address: <http://203.247.189.123:9012>

In 2025, we are considering to replace LLM to **OpenAI gpt-oss-120b**.



HPC ITSM Webpage

서비스 요청관리

서비스 요청을 관리하고 처리 상태를 확인하세요.

서비스 요청 등록

요청제목 *

요청 제목을 입력하세요

요청구분 *

요청구분 선택

요청자

시스템관리자

요청 내용 *

요청 내용을 상세히 입력하세요

첨부파일

파일 추가 (0/5개 파일 선택됨)

취소

등록

서비스 요청관리

서비스 요청을 관리하고 처리 상태를 확인하세요.

현재 권한: 관리자

+ 새 요청 등록

총 진행 현황
1
건

요청
1
건

접수
0
건

진행중
0
건

완료
0
건

보류
0
건

검색 조건

요청기간

년-월-일

~ 년-월-일

요청제목

요청제목 입력

요청자

요청자 입력

요청구분

요청구분 선택

완료일시

년-월-일

~ 년-월-일

완료예정일시

년-월-일

~ 년-월-일

진행상태

진행상태 선택

진행률

최소

최대

담당자

담당자 입력

초기화

검색

검색 결과

전체 조회 건수: 1건

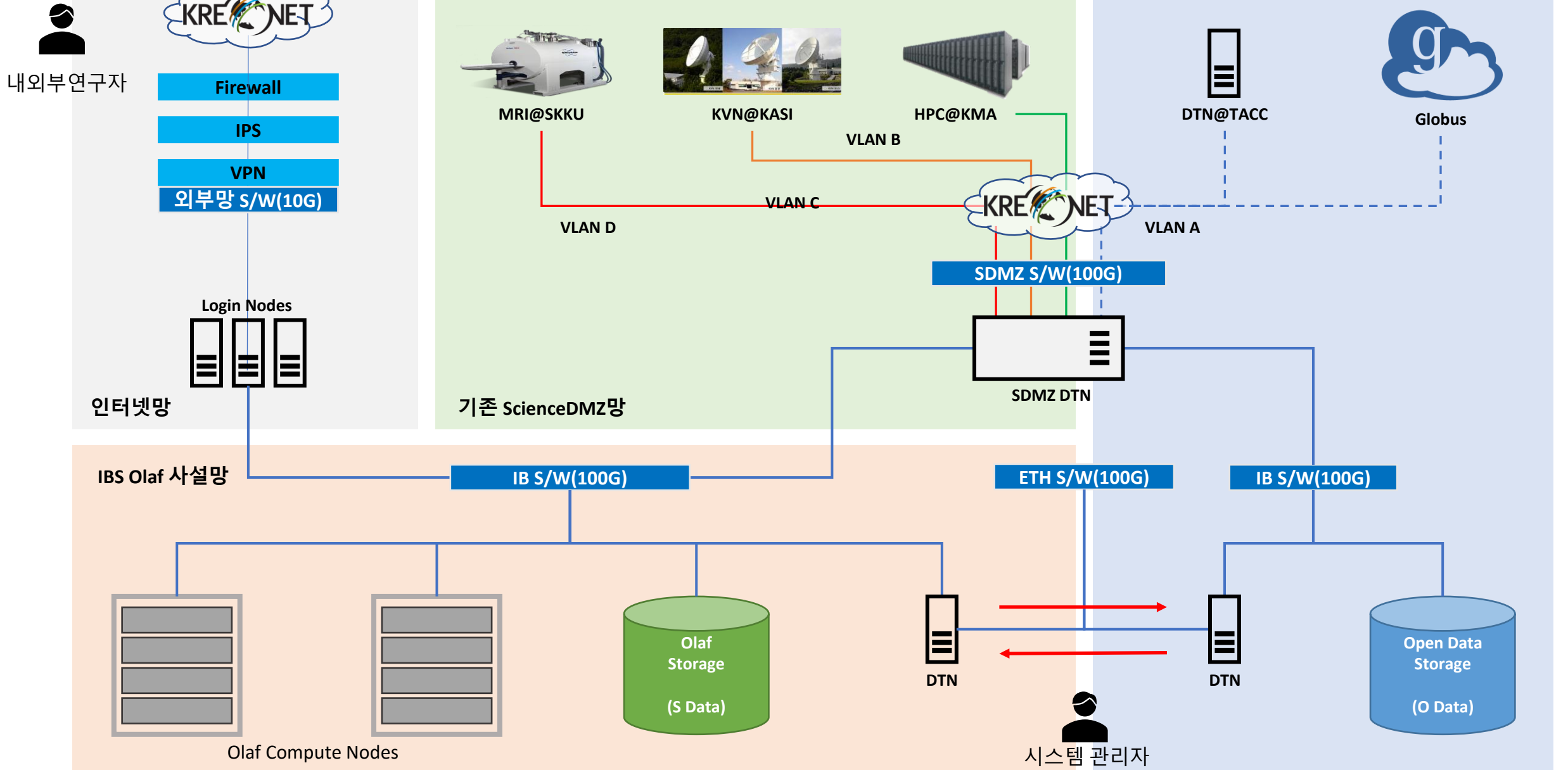
10

요청제목	요청구분	진행상태	진행률	요청자	요청일시	담당자	완료예정일시	완료일시
test	오피스인프라	요청	0%	시스템관리자	2025-10-23 14:52:29	-	-	-

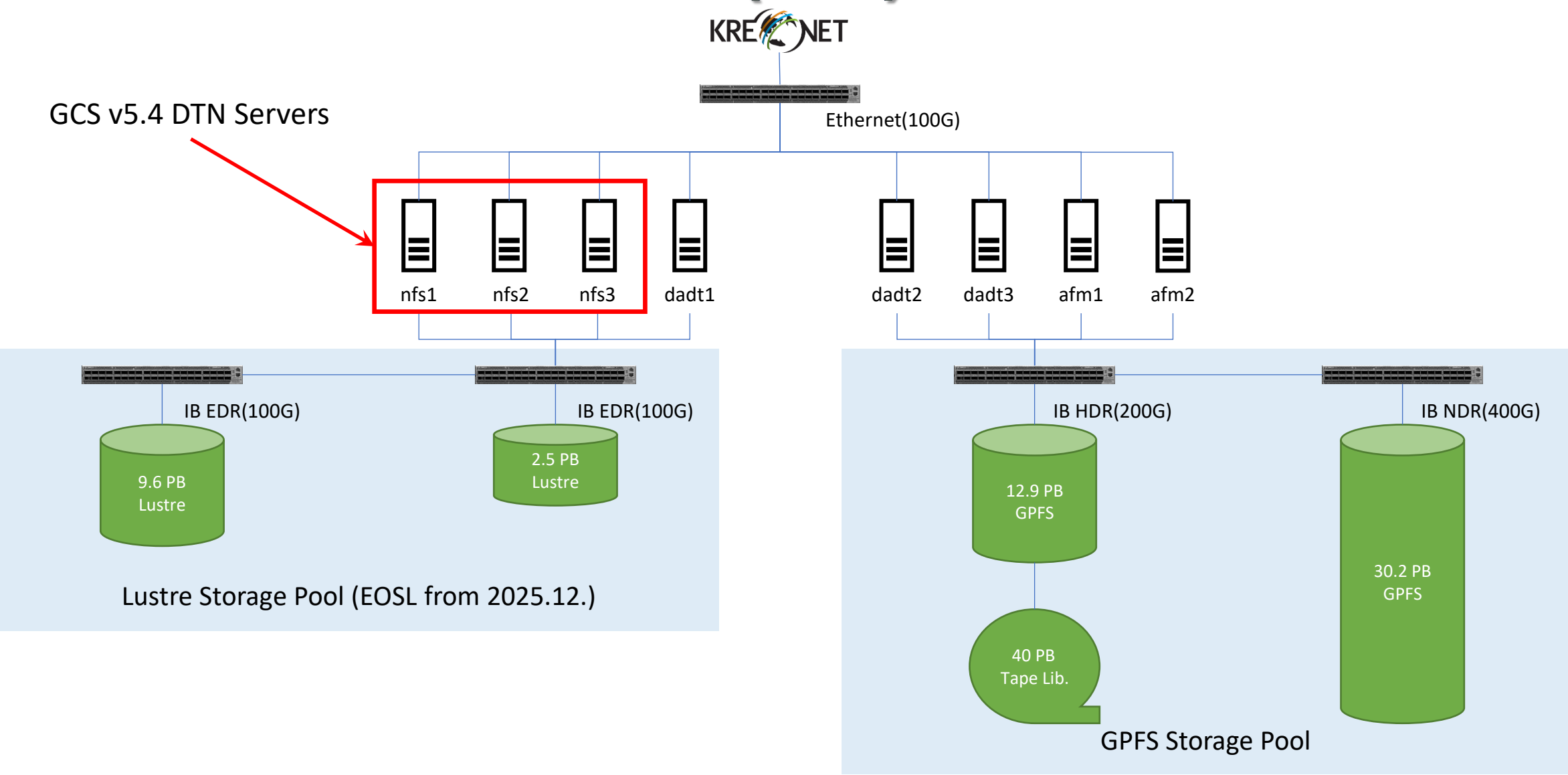
< 이전 1 다음 >

1 / 1 페이지 (총 1건)

Globus 활용 데이터 전송



Globus-Connect-Server (GCS)



Globus.org -> Olaf OpenDTN Collection

The screenshot displays the Globus.org web interface. The top section, titled 'Collections', features a search bar with the text 'olaf' and a search icon. Below the search bar are several filter buttons: 'Recent Tasks', 'Administered By You', 'In Use', 'Shareable By You', and 'Shared With You'. A 'TYPE' dropdown menu is set to 'All'. The main content area shows a table of collections:

COLLECTION	HA	SUBSCRIBED	STATUS	ROLE
Olaf DTN Collection: globusid.org Mapped Collection (GCS) on IBS Olaf DTN			ready	
Olaf OpenDTN Collection: globusid.org Mapped Collection (GCS) on Olaf OpenDTN			has active tasks	

The bottom section, titled 'Settings', shows the 'Account' tab. It displays a list of identities linked to the Globus account, with the primary identity being 'lbs@globusid.org'. Below this, the identity details are shown:

- ID: 19c29[redacted]
- Identity Provider: Globus ID
- Name: Muyoung Heo
- Email: mh@ib.s.re.kr
- Organization: Institute for Basic Science
- Use It To: ☒ log into Globus, ☒ access data, ☒ share data with you

On the right side of the settings page, there are buttons for 'Manage Identities', 'Link Another Identity', and 'Manage Your Consents'. A link to 'Learn more about your Globus account and why you may want to link more of your identities to it.' is also present. At the bottom, it states 'This id is managed by Globus ID, where you can change your password, and update your profile.'

- Create an account for globus.org
- For ID mapping, please send us Globus ID key
- Install Globus-Connect-Personal (GCP, <https://www.globus.org/globus-connect-personal>) in your laptop or server.
- Find “Olaf OpenDTN Collection” for access Olaf Open Storage
- Transfer your data to Olaf Open Storage using globus.org.
- Contact Olaf administrator to approve the data move.
- Find your data in /proj/lustre_approved

Active File Management(AFM) 설정

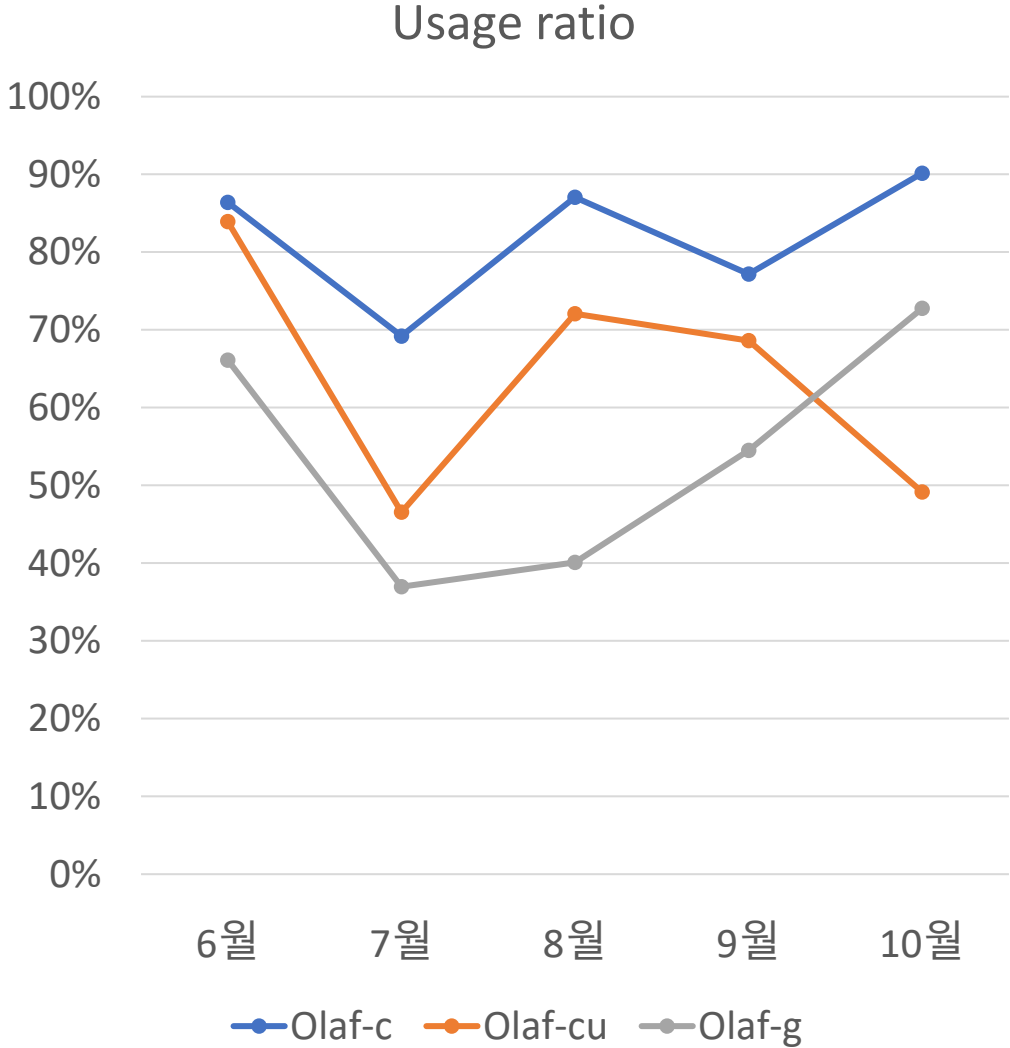
- AFM: NFS(Network File System)의 meta data 및 file을 caching
- /mnt/lustre는 lustre client를 통해 HQ/HQ2/Jepyc 등 구형 클러스터에서 접속 가능하며,
- Olaf 계산노드(olaf-cu, olaf-c, olaf-g)에서는 **AFM**을 통해 접근 가능
- AFM
 - aleph:/mnt/lustre -> /proj/aleph_cache
 - olaf:/mnt/lustre/ibs -> /proj/lustre_ibs
 - olaf:/mnt/lustre/internal_users -> /proj/lustre_internal_users
 - olaf:/mnt/lustre/external_users -> /proj/lustre_external_users
 - olaf:/mnt/lustre/approved -> /proj/lustre_approved

HPC 계산노드 정보

Name	구축연도	CPU	GPU	RAM/Node	Network (B/W)	Network Fabric	# of Nodes	Peak Performance
Olaf_cu	2020	CCLx2 (2.1GHz, 26c)	NVidia V100SXMx8	768GB	IB EDRx2 (200 Gbps)	Single S/W	5	313 TF
Olaf_c	2022	ICLx2 (2.4GHz, 36c)	-	256GB	IB HDR (200 Gbps)	Nonblocking Fat-Tree	210	1,161 TF
Olaf_g	2023	Genoax2 (2.7GHz, 32c)	NVidia H100SXMx4	1TB	IB HDRx4 (1,600 Gbps)	Single S/W	12	3,216 TF

Monthly HPC Usage Ratio

Olaf-c monthly usage ratio in Oct.: 90.2 %



Partition	Node Alloc.	Usage ratio
core_s/m/l	40 노드 [121-144, 195-210] Shared mode	79.52%
normal_cpu/ long_cpu/ express_cpu	50 노드 [145-194] Exclusive mode	96.96%
large_cpu/ express_lcpu	120 노드 [1-120] Exclusive + min. 8 nodes	90.86%

Usage Statistics in the 1st week of Nov.

Partition	Assigned node numbers	Max. walltime	Priority	Total usage time (node*sec.)	Abs. usage ratio (%)	Partition usage ratio (%)	Waiting time (hr)
core_s	40	2 hr	200	5,409,615	4.26	24.34	6.1
core_m		3 days	20	477,613	0.38		122.6
core_l		14 days	10	1,677	0.00		3.8
normal_cpu	50	3 days	20	19,277,676	15.18	99.26	24.8
long_cpu		14 days	2	9,805,940	7.72		38.2
express_cpu		14 days	220	932,943	0.73		4.9
large_cpu	120	3 days	20	42,183,460	33.21	91.64	6.1
express_lcpu		14 days	220	24,326,456	19.15		5.0
	210			102,415,380	80.64		

Usage Statistics in the 2nd week of Nov.

Partition	Assigned node numbers	Max. walltime	Priority	Total usage time (node*sec.)	Abs. usage ratio (%)	Partition usage ratio (%)	Waiting time (hr)
core_s	40	2 hr	200	8,714,282	6.9	49.5	13.6
core_m		3 days	20	3,240,173	2.6		90.0
core_l		14 days	10	16,659	0.0		39.0
normal_cpu	50	3 days	20	28,908,894	22.8	98.2	11.0
long_cpu		14 days	2	801,664	0.6		26.6
express_cpu		14 days	220	-	0.0		-
large_cpu	120	3 days	20	41,407,879	32.6	92.6	6.3
express_lcpu		14 days	220	25,783,941	20.3		5.0
	210			108,873,493	85.7		

Slurm QoS Policy Update (suggestion?)

- core_s/m/l partions has 5,760 (=144x40) cores

Partition	Assigned node numbers	Max. walltime	Priority	QoS (Max Cores Per Group)
core_s	40	2 hr	200	2,000
core_m		3 days	20	1,000
core_l		14 days	10	500

Summary

- 신규 스토리지 구축 및 DTN 노드 증설
- core_s/m/l 파티션 사용자들의 long waiting time 문제는 좀 더 시간을 갖고 들여다 볼 계획이며, 현 시점에서 normal_cpu, large_cpu 와 자원 재조정은 불필요한 것으로 판단
- core_s/m/l 파티션 long waiting time 문제 해결을 위한 그룹 당 최대 코어 개수를 QoS로 설정
- 추후 core_s/m/l 파티션의 대기시간이 지속적으로 길어질 경우, HQ, HQ2, Jepyc 등 구형 노드 활용
- Globus(GCS)를 활용한 데이터 전송을 원하는 경우 hpcservice@ibs.re.kr로 문의