



【招待講演(S-6)】

クライオ電子顕微鏡 IoT におけるAWSとHPCの活用： 現状の応用と今後の展望

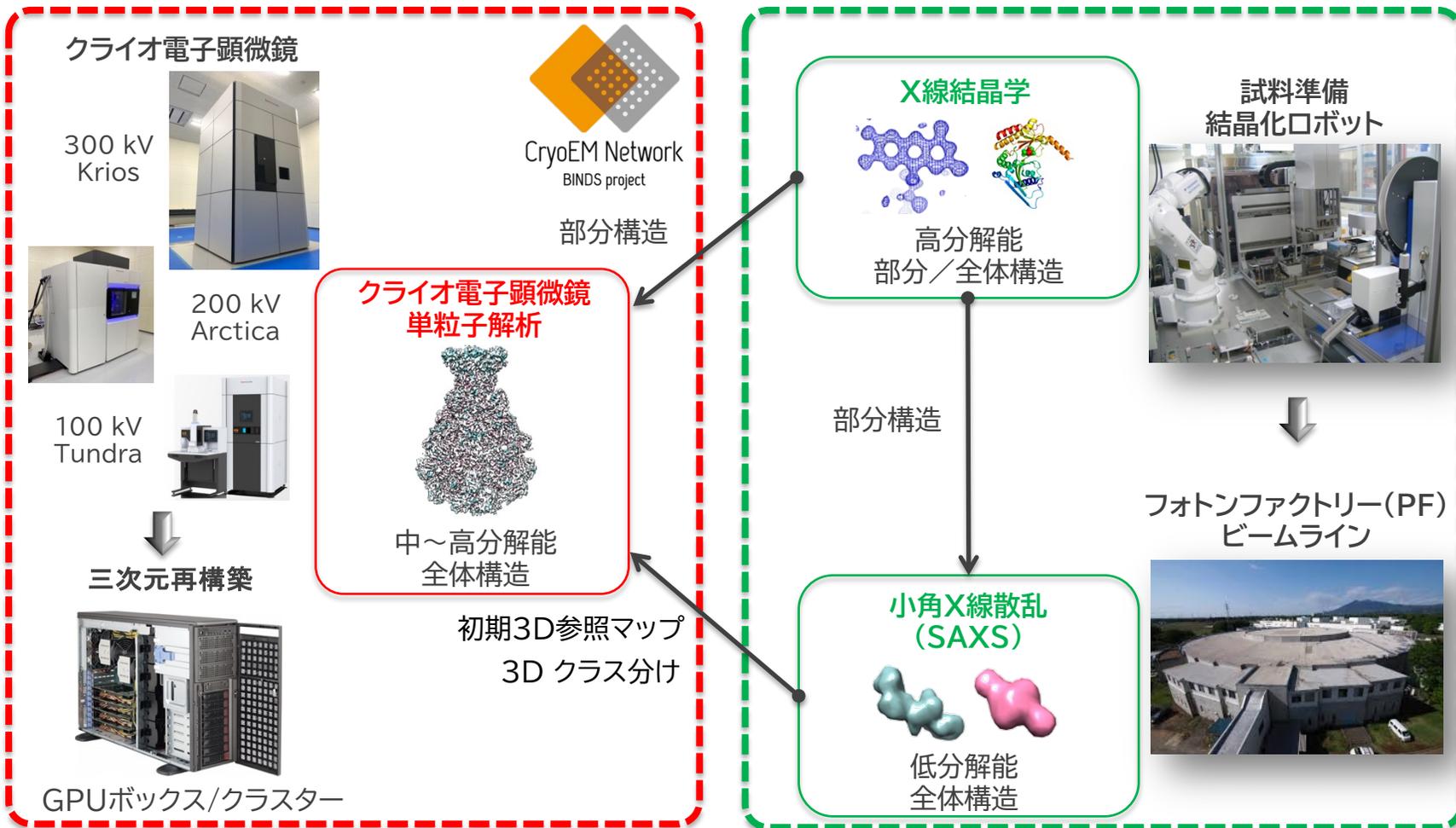
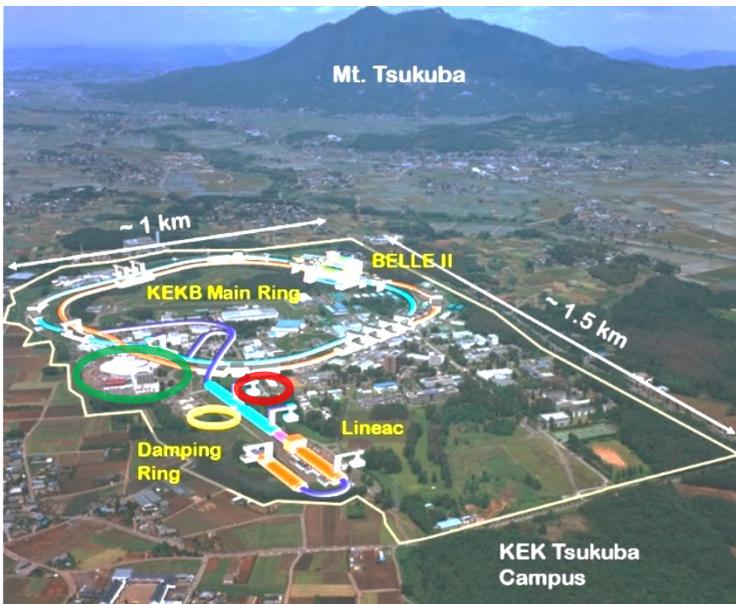
守屋 俊夫

高エネルギー加速器研究機構(KEK) 物質構造科学研究所(IMSS) 構造生物学研究センター(SBRC)

Supercomputing Japan 2026 @ タワーホール船堀 小ホール
2026年2月3日(火) 14:30~15:00

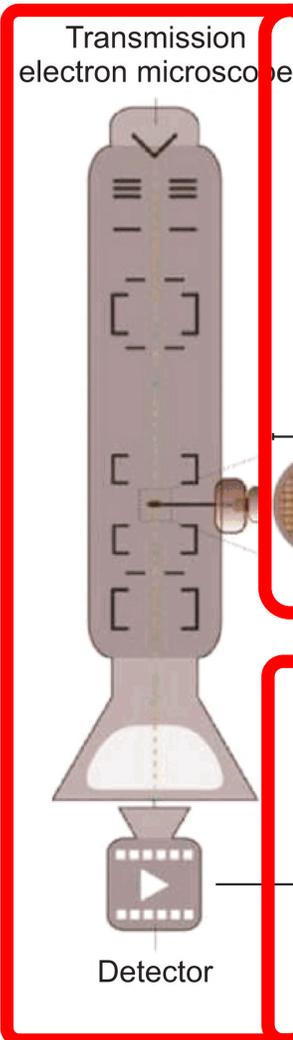
相関構造解析が可能な総合的**構造生物学**研究環境

KEK 構造生物学研究センター (KEK-SBRC)

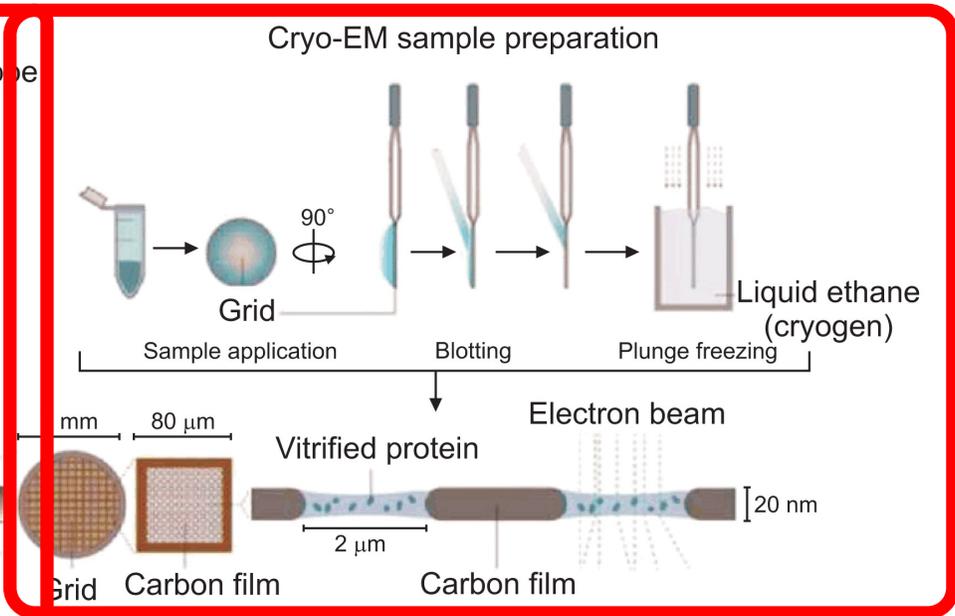


クライオ電子顕微鏡・単粒子解析の全体像

クライオ電顕撮影



試料準備(氷包埋法)

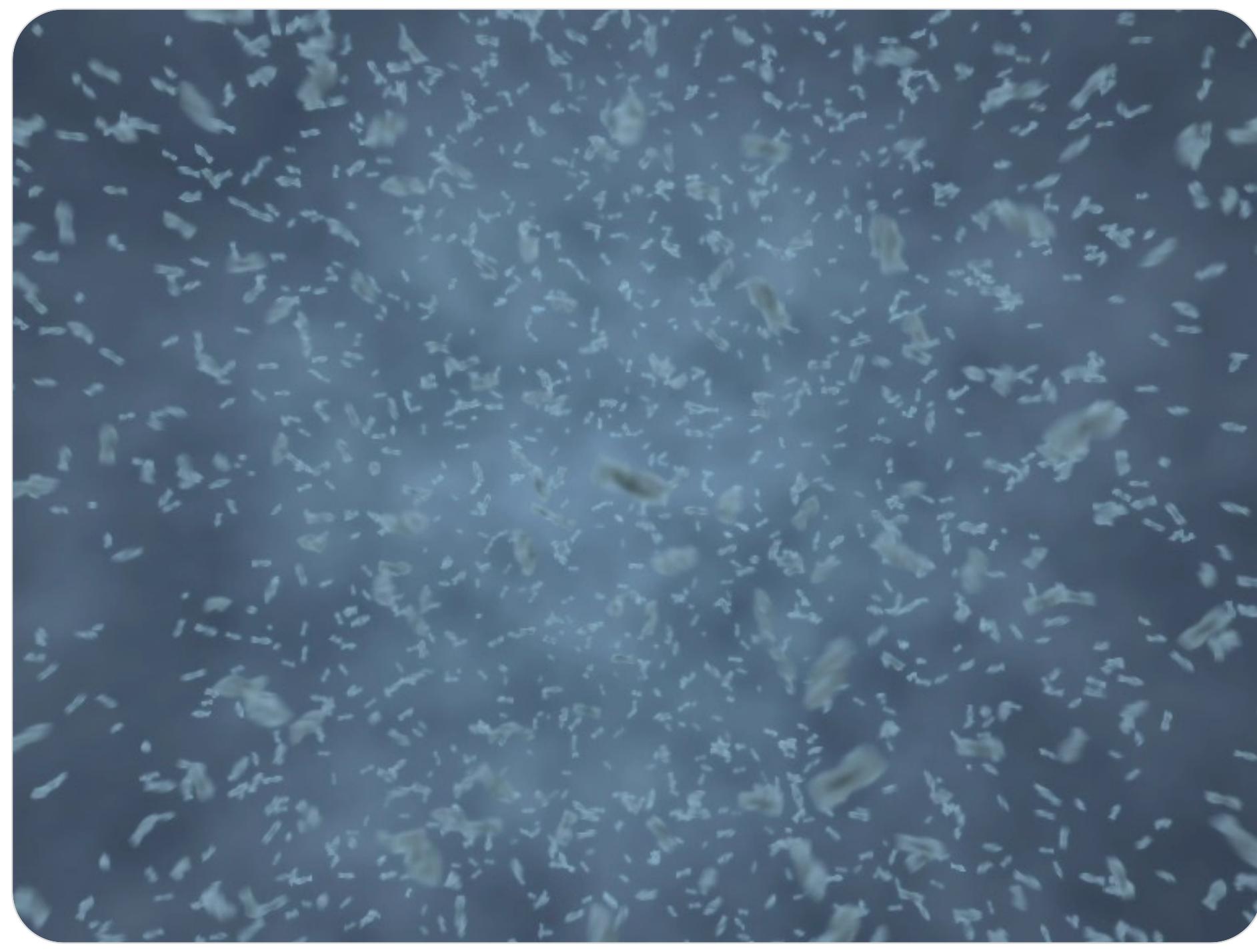


原子座標モデル構築



電顕像(マイクログラフ)

コンピューター画像処理(三次元構造再構築)



クライオ電子顕微鏡
単粒子解析

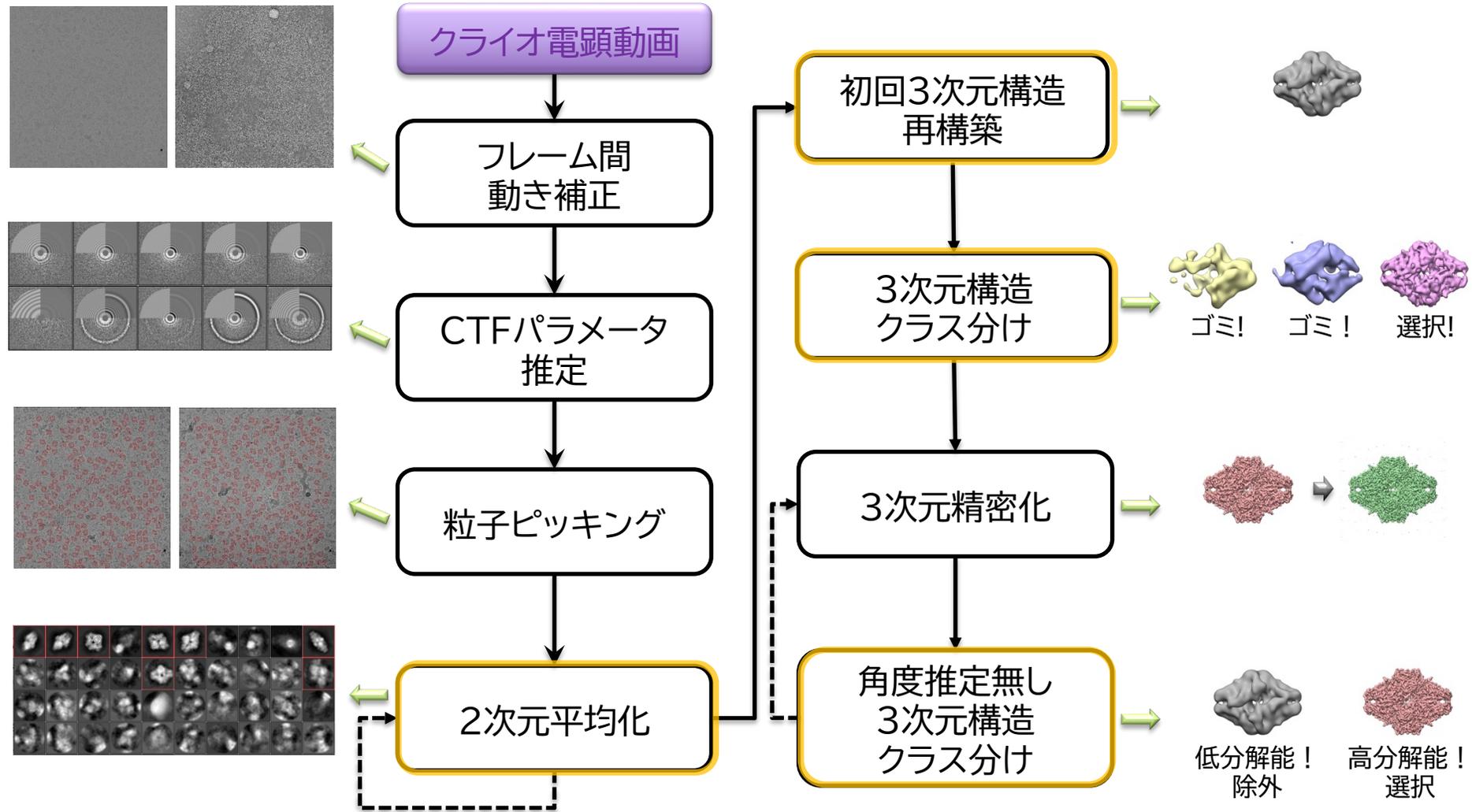
Cryo-EM SPA

Cryogenic Electron
Microscopy
Single Particle Analysis

ThermoFisher
SCIENTIFIC

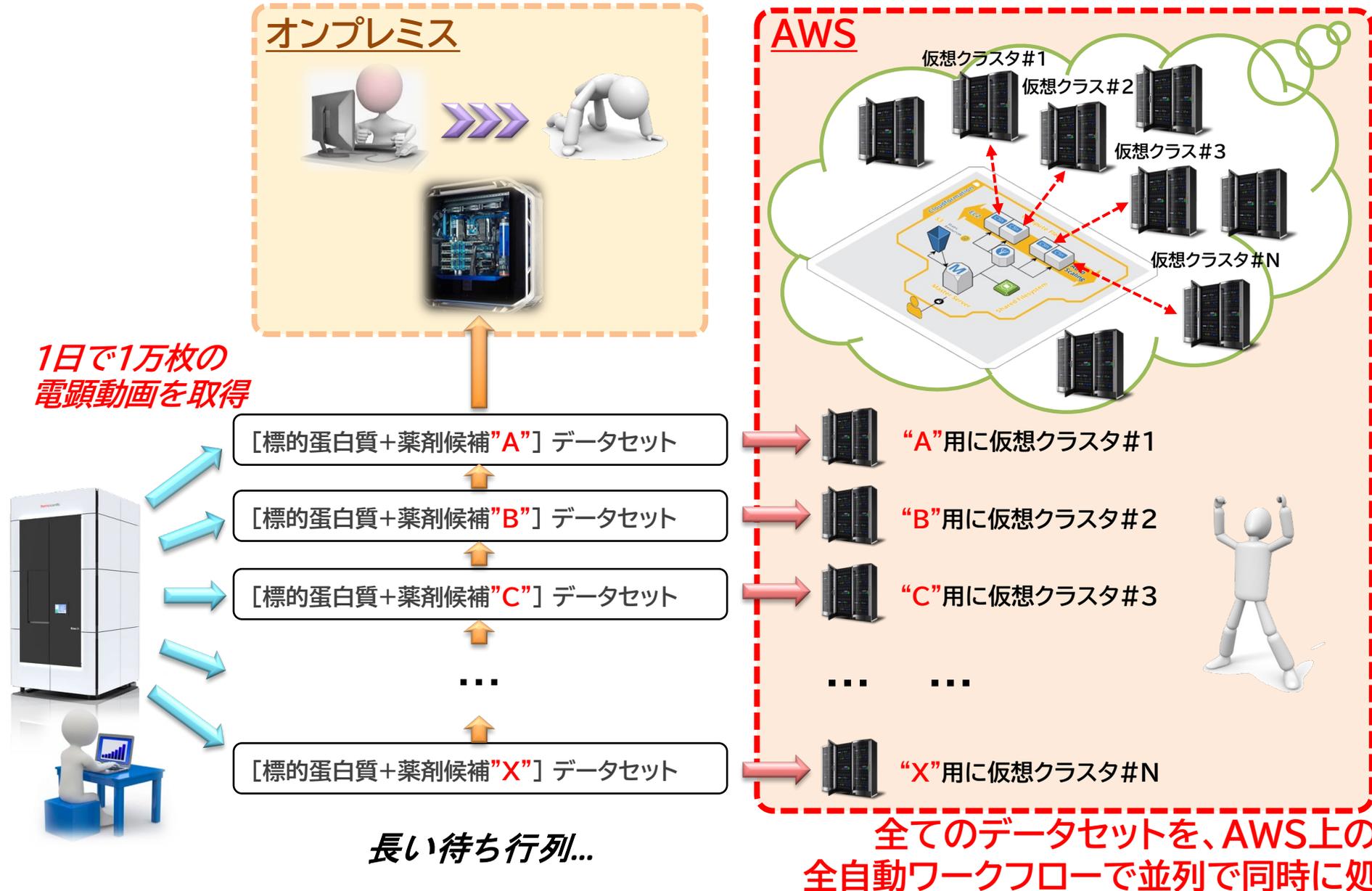
Proprietary & Confidential
shintaro.maeda@thermofisher.com
10-Sep -2024

クライオ電顕単粒子解析のデータ解析の概要とその課題



手動データ解析には平均的に少なくとも数日から一週間はかかる。

なぜクライオ電顕構造ベース創薬(SBDD)にAWSが必要？



ユーザー支援の高度化におけるこれまでの成果 ～単粒子解析ワークフローの全自動実行～

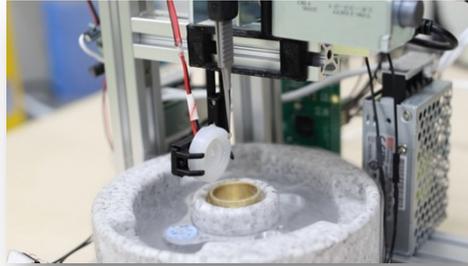
クライオ電顕



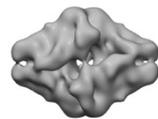
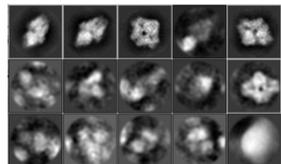
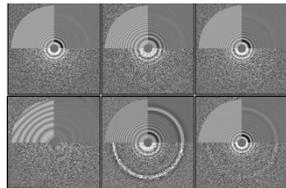
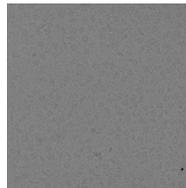
自動
データ転送



グリッド凍結

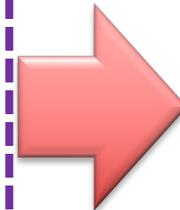


「GoToFly(JK)」@KEK-SBRC
独自開発の全自動On-The-Fly逐次処理

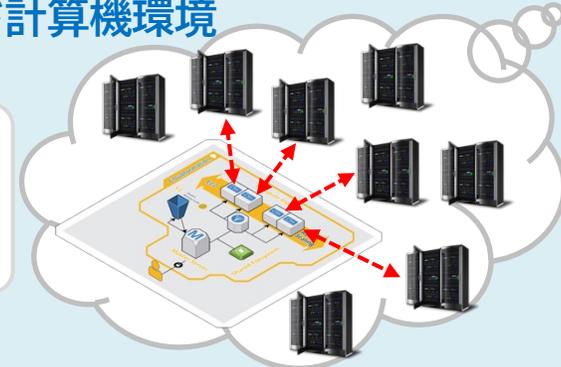


*Multi-MD: Multiple Mask Diameters

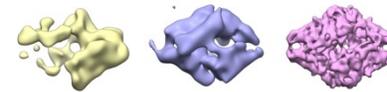
自動高速
データ転送



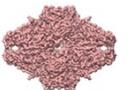
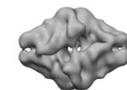
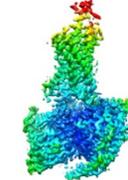
「GoToCloud」@AWS
独自開発のCryo-EM SPA向け
クラウド計算機環境



「CS-Schemes」@GoToCloud
独自開発の創薬向け
全自動Cryo-EM SPA解析処理



【処理結果例】
GPCR膜蛋白質



開始: 40万粒子
最終: 12万粒子
分解能: 2.96 Å
総時間: 9.0時間
総コスト: 187ドル

communications biology

Explore content ▾

About the journal ▾

Publish with us ▾

[nature](#) > [communications biology](#) > [articles](#) > article

Article | [Open access](#) | Published: 14 October 2024

GoToCloud optimization of cloud computing environment for accelerating cryo-EM structure-based drug design

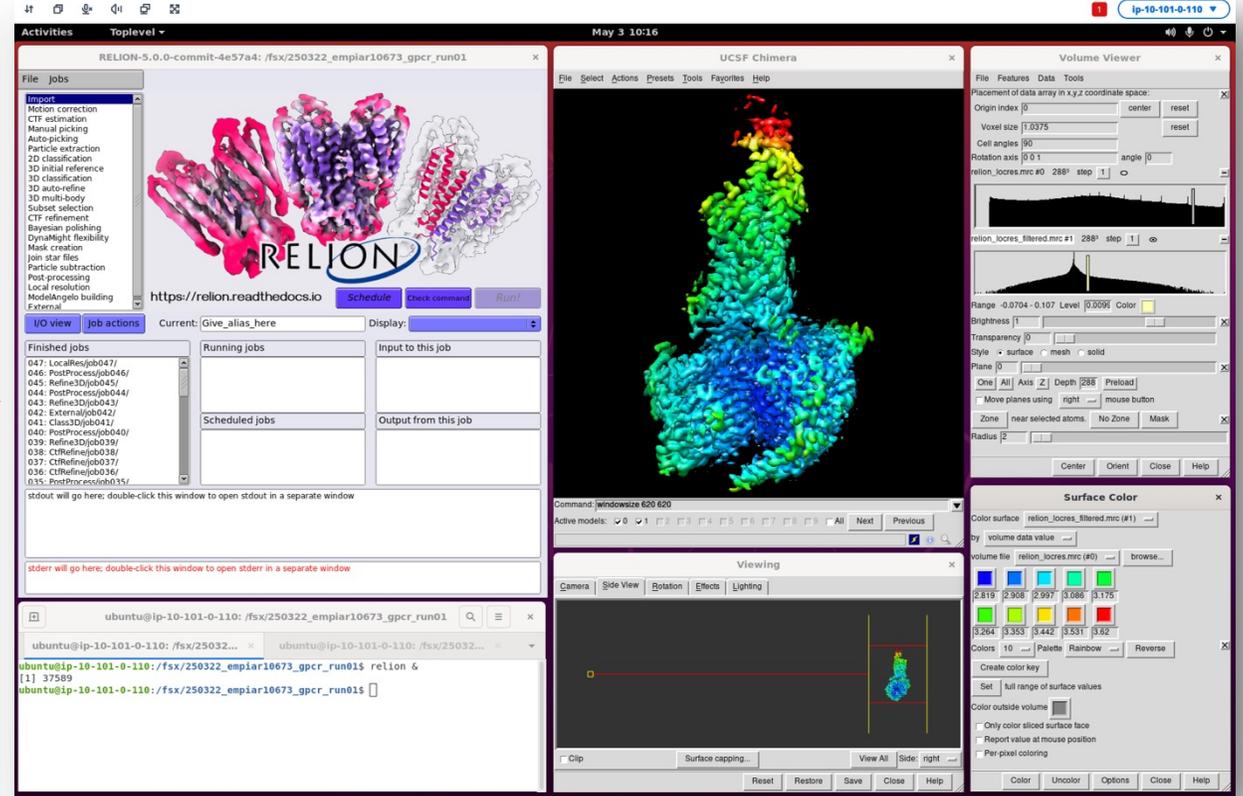
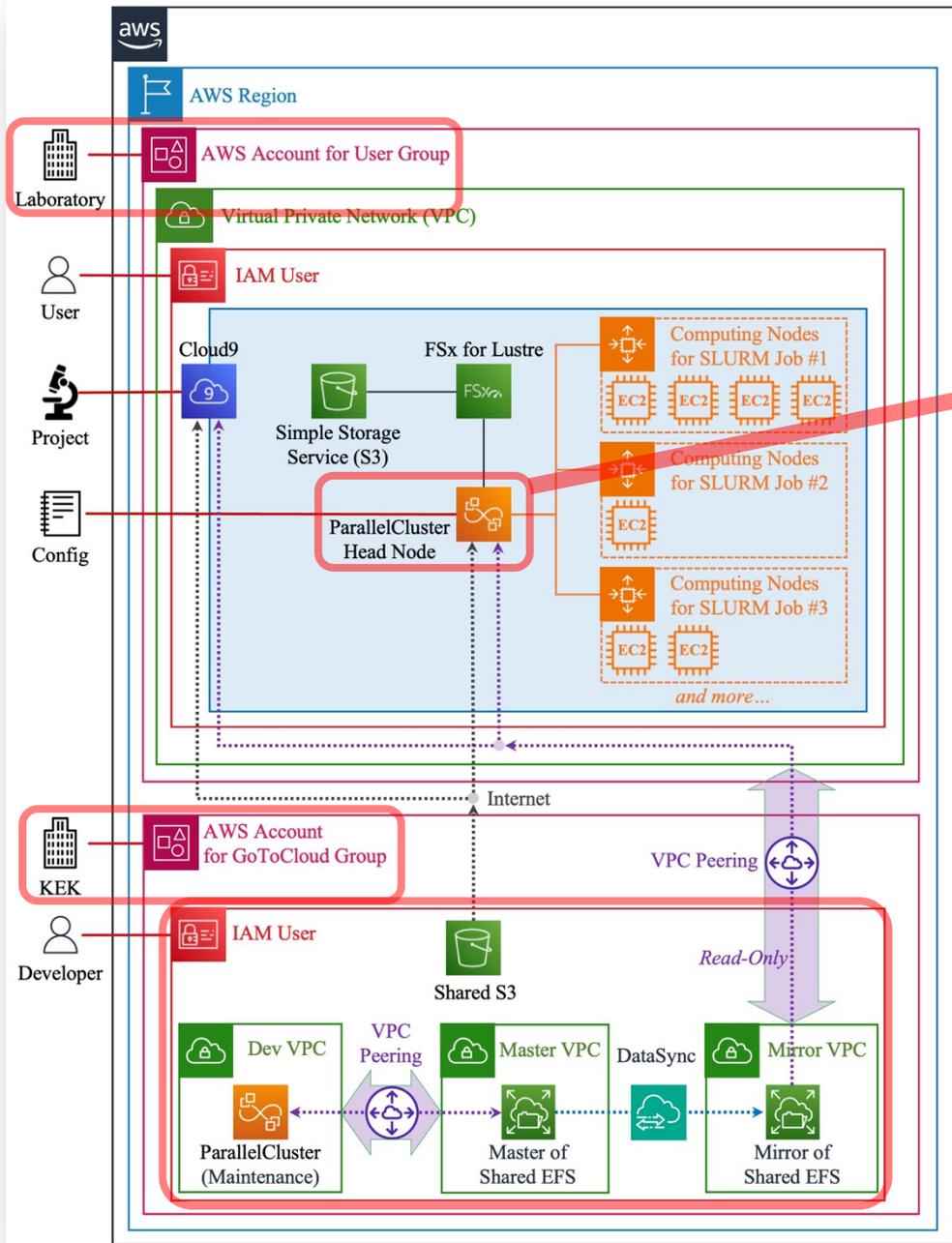
[Toshio Moriya](#) , [Yusuke Yamada](#), [Misato Yamamoto](#) & [Toshiya Senda](#) 

[Communications Biology](#) 7, Article number: 1320 (2024) | [Cite this article](#)

5232 Accesses | **6** Citations | 7 Altmetric | [Metrics](#)

2026年02月02日時点

GoToCloud@AWS ParallelCluster



キーポイント1

各研究グループごとに
AWSアカウントを分離

キーポイント2

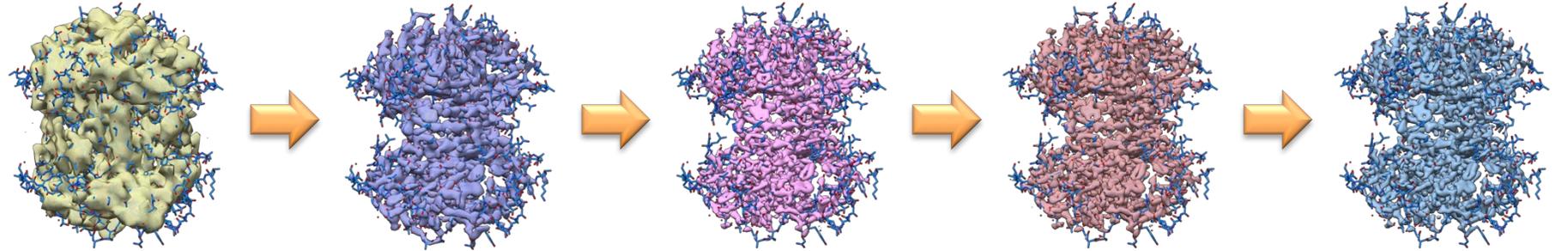
全てのAWSアカウントが
共有できるストレージ

AWSが既に提供している
堅牢なセキュリティ!

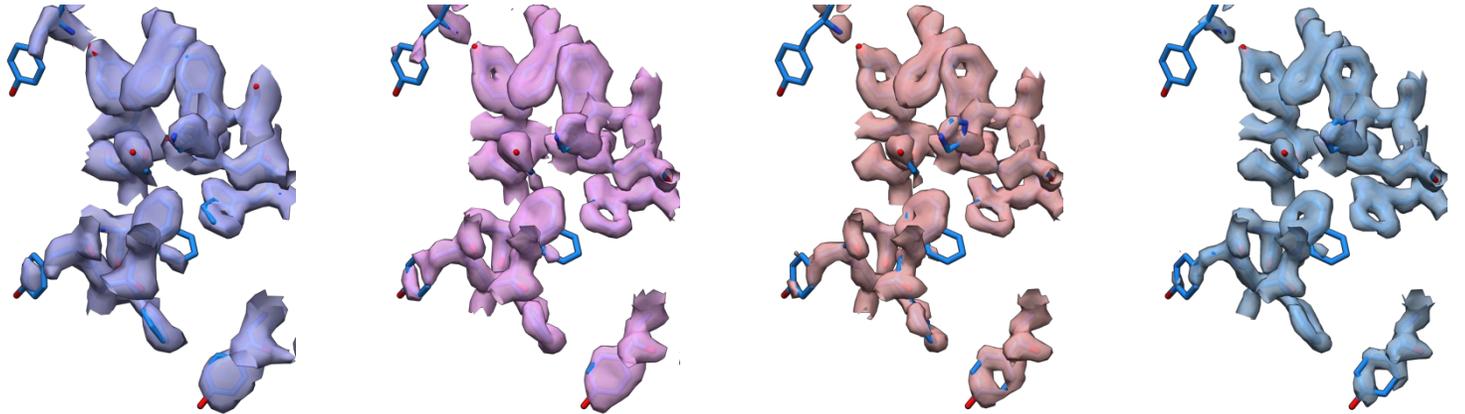
KEKの全てのノウハウを
ここに集約!

EMPIARデータセットを用いた 手動解析による高分解能構造解析@GoToCloud

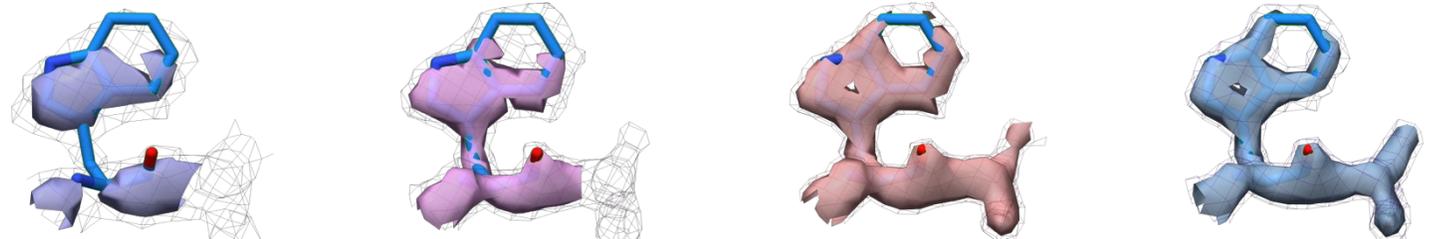
【EMPIAR-10641】
名称: ストレプトアビジン
分子サイズ: 53 kDa
分解能: 1.93 Å



芳香環



トリプトファン108



128 pix @1.60 Å
367,626 粒子
3.20 Å 分解能
956 ドル
34.41 時間

352 pix @0.80 Å
335,437 粒子
2.23 Å 分解能
1,128 ドル
42.06 時間

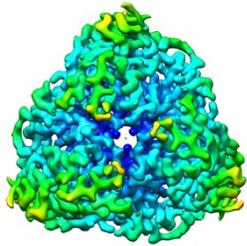
540 pix @0.60 Å
257,383 粒子
1.95 Å 分解能
2,204 ドル
77.65 時間

540 pix @0.60 Å
257,383 粒子
1.87 Å 分解能
4,545 ドル
136.55 時間

540[pix]@0.60[A]
122,011 粒子
1.83 Å 分解能
4,884 ドル
150.33 時間

亜硝酸還元酵素(NiR)ベンチマーク@GoToCloud

2021年4月以来



EMPIAR-10581: Nitrite Reductase at pH 8.1 (2.85Å)

Rescaled Box [Pix]	: 352	Particles	: 129,298
Rescaled Apix [A/Pix]	: 1.17	Symmetry	: C3
Original Box [Pix]	: 468	Mask Diameter	: 164
Original Apix [A/Pix]	: 0.88		

RELION version: 3.1.2, 4.0-beta, 4.0.0, 4.0.1, 5.0-beta, and 5.0.0

AWS ParallelCluster version: 3.0.3 and 3.7.0

AWS Region: us-east-1

OS: Ubuntu2004

With GPU

- g6 (NVIDIA L4 Tensor Core GPU)
- g5 (NVIDIA A10G Tensor Core)
- g4dn (NVIDIA T4)

3D Refinement

2D Classification

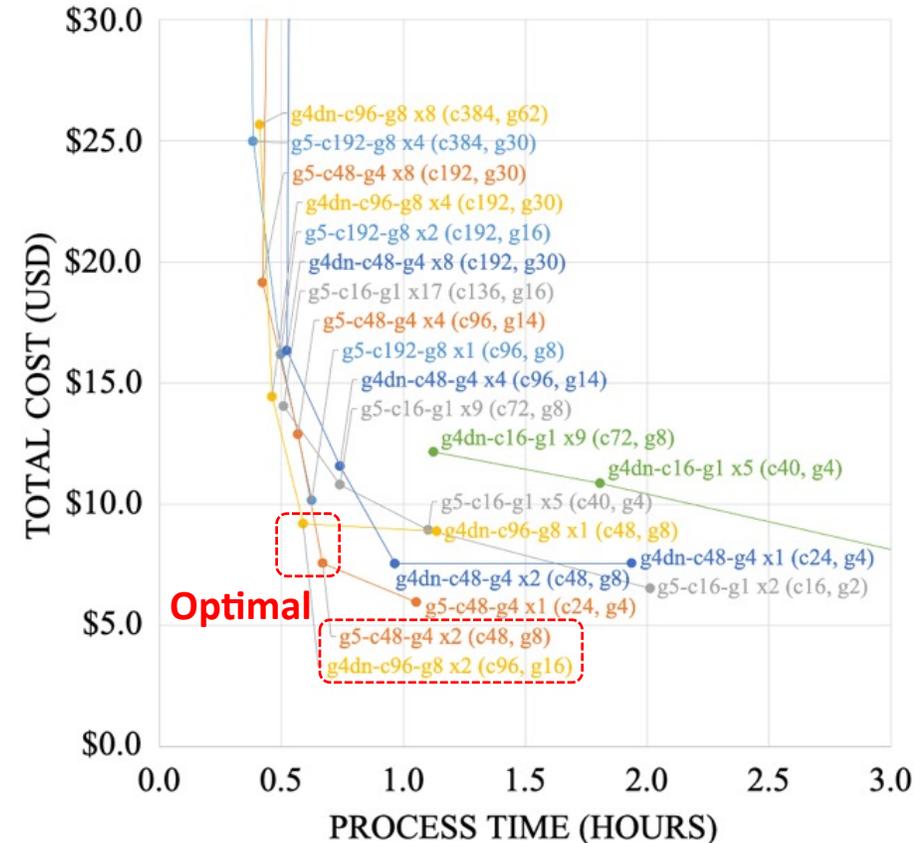
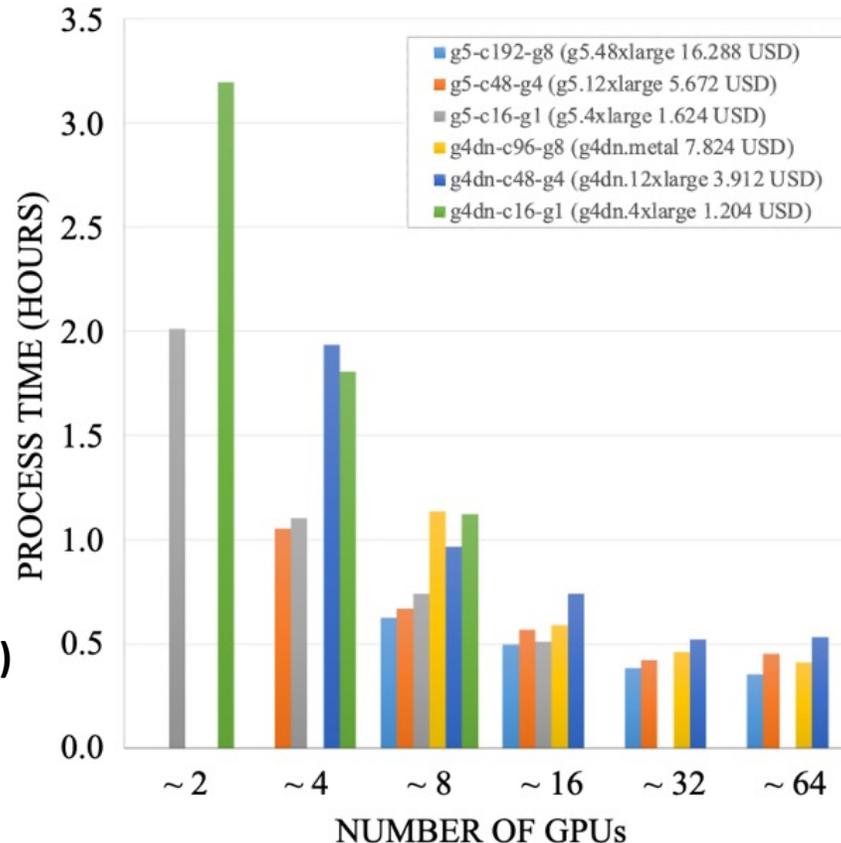
3D Classification

Only CPU (No GPU)

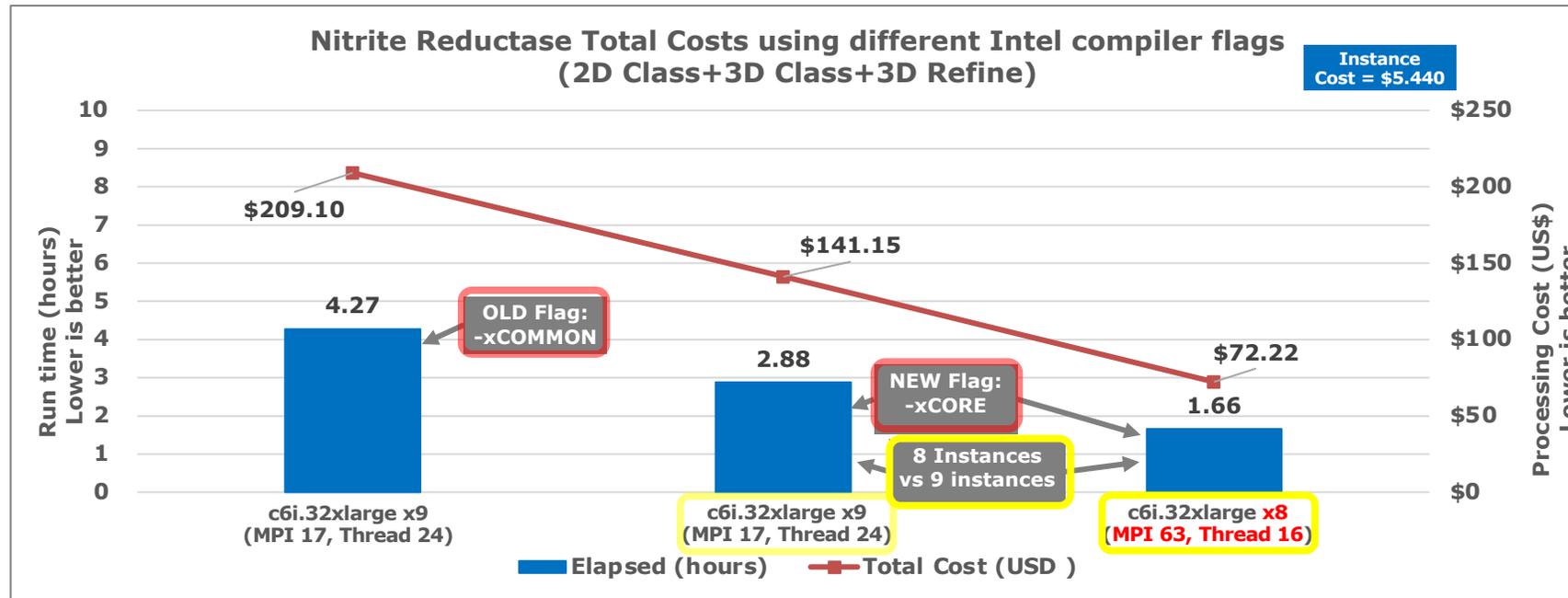
- C7i/M7i/R7i (INTEL Sapphire Rapids)
- C6i/M6i/R6i (INTEL Ice Lake)

Polishing

3D Refinement Benchmark Result



Time and Cost Savings for Cryo-EM Single Particle Analysis (SPA) on AWS



<OLD Compiler Flags>
 - Relion 4.0-beta-2
 - Compile option
 -DCMAKE_C_FLAGS="-O3 -ip -g -xCOMMON-AVX512 -restrict "
 -DCMAKE_CXX_FLAGS="-O3 -ip -g -xCOMMON-AVX512 -restrict "

<NEW Compiler Flags>
 - Relion 4.0.0
 - Compile option
 -DCMAKE_C_FLAGS="-O3 -ip -xCORE-AVX512 -qopt-zmm-usage=high -qoverride-limits"
 -DCMAKE_CXX_FLAGS="-O3 -ip -xCORE-AVX512 -qopt-zmm-usage=high -qoverride-limits"

Cost and time saving are between "Old Compiler Flags" using nine (9) c6i.32xlarge and the "New Compiler Flags" using eight (8) c6i.32xlarge

- reduced the total costs from \$209 to \$72 (a 65% cost reduction)
- shortened the execution time from 4.27h to 1.66h (2.57x faster).

Source: KEK | <https://www.kek.jp/en/>

Presented at AWS re:Invent 2022

AWSベンチマーク実行環境(2024年3月)

Relion			AWS		Applied benchmark tests
version	Acceleration type	Compiler	ParallelCluster version	EC2 Instance	
5.0-beta	CPU (-xCORE-AVX512)	Intel oneAPI 2023.2	3.7.0	c7i.48xlarge c6i.32xlarge	Refine3D Class3D Class2D
	GPU	GCC 9.4.0		g5.12xlarge	Refine3D Class3D (Class2D)
				g4dn.12xlarge	Class2D

各ジョブのOver-Subscribed設定およびノード数の最適化

Refine3D

- Fastest with G5 using over-subscribed **1 MPI/GPU (No-Split)**
- Best cost performance with G5 using **1 MPI/GPU, 4-nodes, and half vCPUs** (only physical cores)

Class3D

- Fastest with G5 using over-subscribed **2 MPI/GPU (2 Split)**
- Best cost performance with G5 using **2 MP/GPU, 4-node, and full vCPUs** (all virtual cores).

Class2D

- **Faster with C7i than G5**
- Fastest with G5 using over-subscribed **4 MPI/GPU (4 Split)**
- Best cost performance with G5 using **4 MPI/ GPU, 8-node, and full vCPUs** (all virtual cores).

最適化された各ジョブの並列計算パラメーター設定(2024年3月)

最適化前

Relion4.01 ap-northeast-1

GPUs	MPI	Thread	MPIs/Node	Nodes	EC2 Instance type	Cores	Cost/hour
	64	32	4	16	c6i.32xlarge	128	6.848
	64		4	16	c6i.32xlarge	128	6.848
	128		128	1	c6i.32xlarge	128	6.848
14	15	6	4	4	g4dn.12xlarge	48	5.281
4	1	4	1	1	g4dn.12xlarge	48	5.281
14	15	6	4	4	g4dn.12xlarge	48	5.281
	17	7	17	1	c6i.32xlarge	128	6.848
8	9	6	5	2	g4dn.12xlarge	48	5.281
	16	32	4	4	c6i.32xlarge	128	6.848
	4	128	1	4	c6i.32xlarge	128	6.848
	64		16	4	c6i.32xlarge	128	6.848
8		1	1	1	g4dn.metal	96	10.562
					m5.xlarge		
					m5.xlarge		
					m5.xlarge		
					m5.xlarge		

最適化後

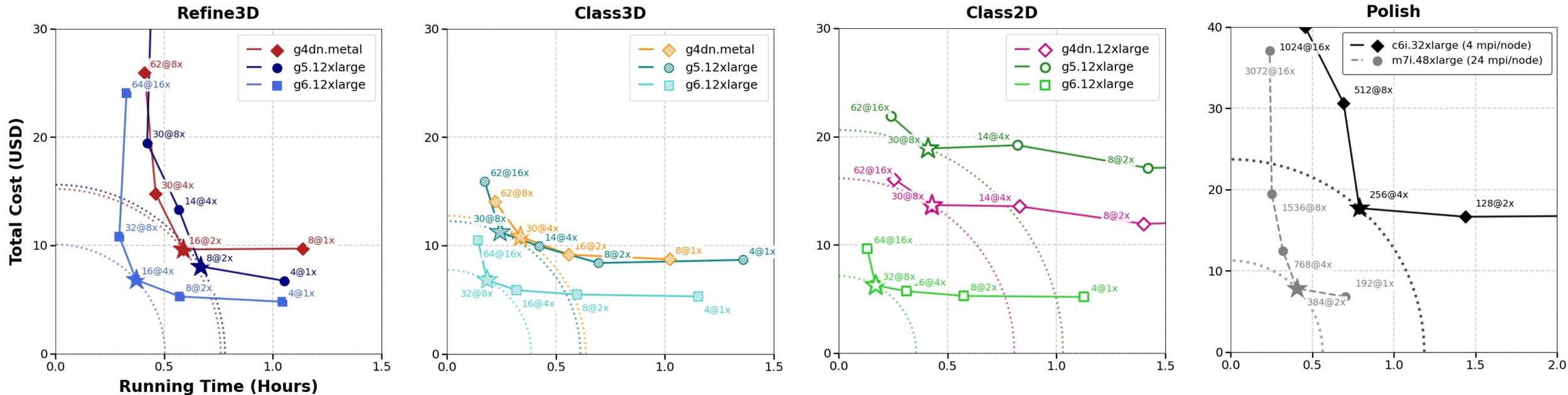
Relion5.0-beta us-east-1

GPUs	MPI	Thread	MPIs/Node	Nodes	EC2 Instance type	Cores	Cost/hour
	48	8	24	2	m7i.48xlarge	192	9.6768
	48		24	2	m7i.48xlarge	192	9.6768
	192		192	1	c7i.48xlarge	192	8.568
	63	12	16	4	c7i.48xlarge	192	8.568
4	1	4	1	1	g5.12xlarge	48	5.672
30	31	6	4	8	g5.12xlarge	48	5.672
	25	7	25	1	c7i.48xlarge	192	8.568
14	15	6	4	4	g5.12xlarge	48	5.672
	48	8	24	2	m7i.48xlarge	192	9.6768
	32	12	16	2	m7i.48xlarge	192	9.6768
	48		24	2	m7i.48xlarge	192	9.6768
8		1	1	1	g5.48xlarge	192	16.288
					m5.xlarge		
					m5.xlarge		
					m5.xlarge		
					m5.xlarge		

GPU 実行: G4dn から G5 (さらに G6)
CPU 実行: C6i から C7i/M7i

最新の並列計算パラメータ最適化の結果

G6(2025年2月)およびC7i, M7i, R7i(2024年1月)のベンチマーク結果
(最適な並列計算設定系列のみ抜粋)



現在、C8i, M8i, R8iのベンチマークを実施中。
次にG7eで実施予定。

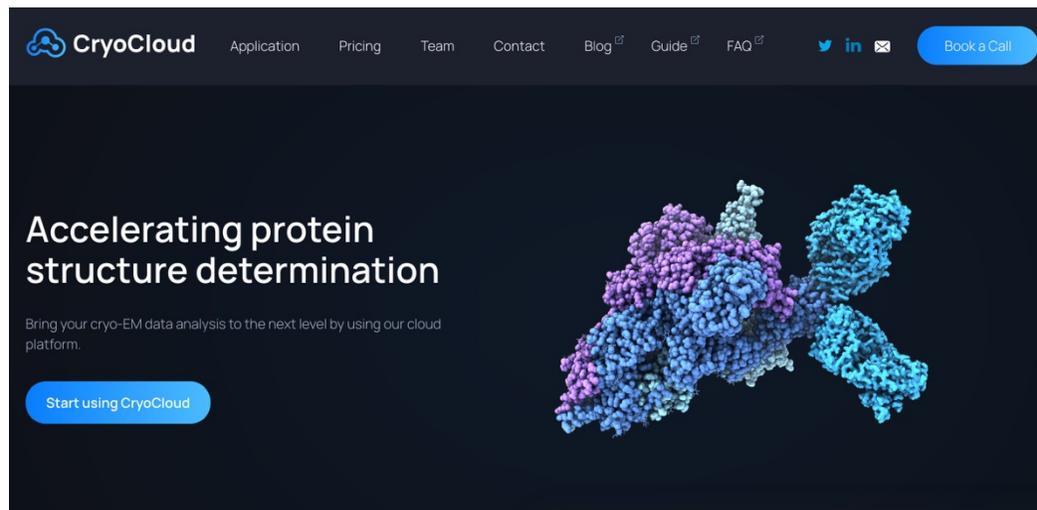
構造生物学向けクラウド計算環境ベースのプラットフォーム開発における近年の動向

GoToCloudにも ...

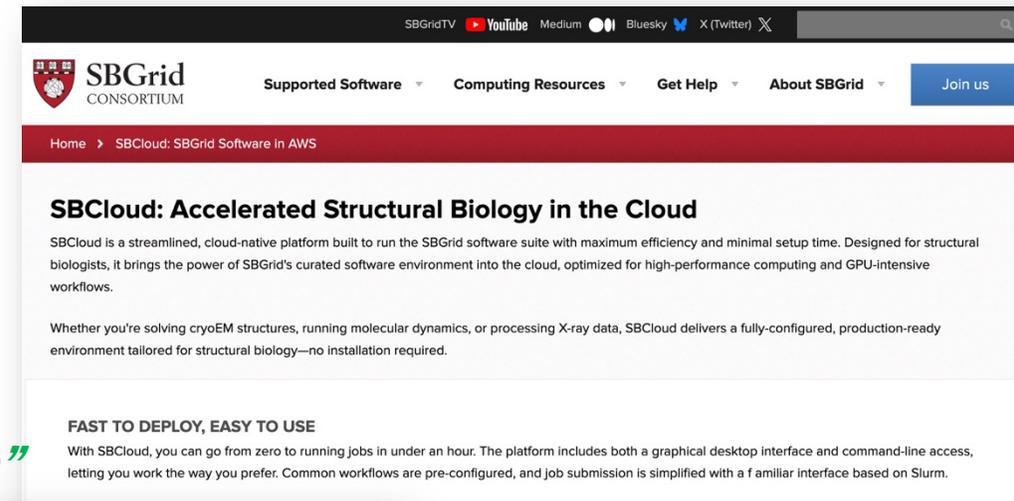
“CryoCloud”

“SBCloud”

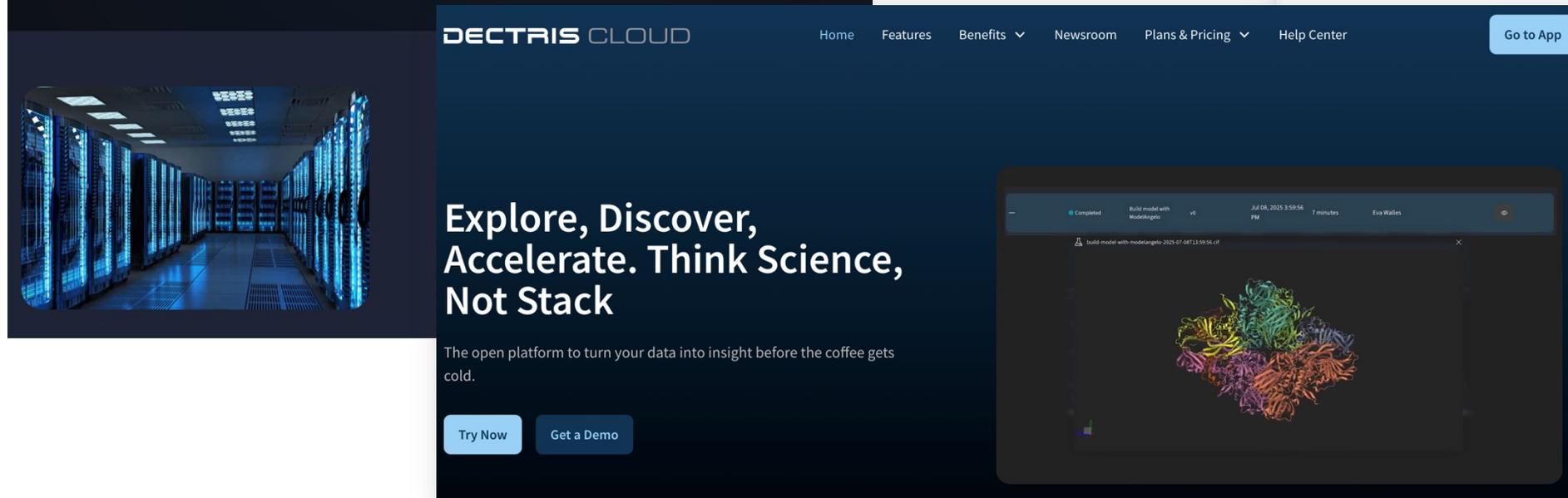
“DECTRIS CLOUD”



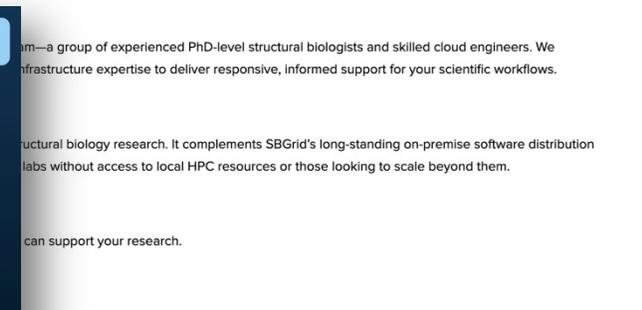
The screenshot shows the CryoCloud website homepage. At the top, there is a navigation bar with the CryoCloud logo and links for Application, Pricing, Team, Contact, Blog, Guide, and FAQ. A blue button labeled "Book a Call" is on the right. The main content area features the heading "Accelerating protein structure determination" and a sub-headline "Bring your cryo-EM data analysis to the next level by using our cloud platform." Below this is a large 3D protein structure model. A blue button labeled "Start using CryoCloud" is positioned at the bottom left.



The screenshot shows the SBGrid SBCloud website homepage. The top navigation bar includes the SBGrid Consortium logo and links for Supported Software, Computing Resources, Get Help, and About SBGrid. A blue button labeled "Join us" is on the right. The main content area features the heading "SBCloud: Accelerated Structural Biology in the Cloud" and a sub-headline "SBCloud is a streamlined, cloud-native platform built to run the SBGrid software suite with maximum efficiency and minimal setup time. Designed for structural biologists, it brings the power of SBGrid's curated software environment into the cloud, optimized for high-performance computing and GPU-intensive workflows." Below this is a paragraph describing the platform's capabilities. A blue button labeled "Go to App" is positioned at the bottom right.



The screenshot shows the DECTRIS CLOUD website homepage. The top navigation bar includes the DECTRIS CLOUD logo and links for Home, Features, Benefits, Newsroom, Plans & Pricing, and Help Center. A blue button labeled "Go to App" is on the right. The main content area features the heading "Explore, Discover, Accelerate. Think Science, Not Stack" and a sub-headline "The open platform to turn your data into insight before the coffee gets cold." Below this is a large 3D protein structure model. A blue button labeled "Try Now" and a blue button labeled "Get a Demo" are positioned at the bottom left.



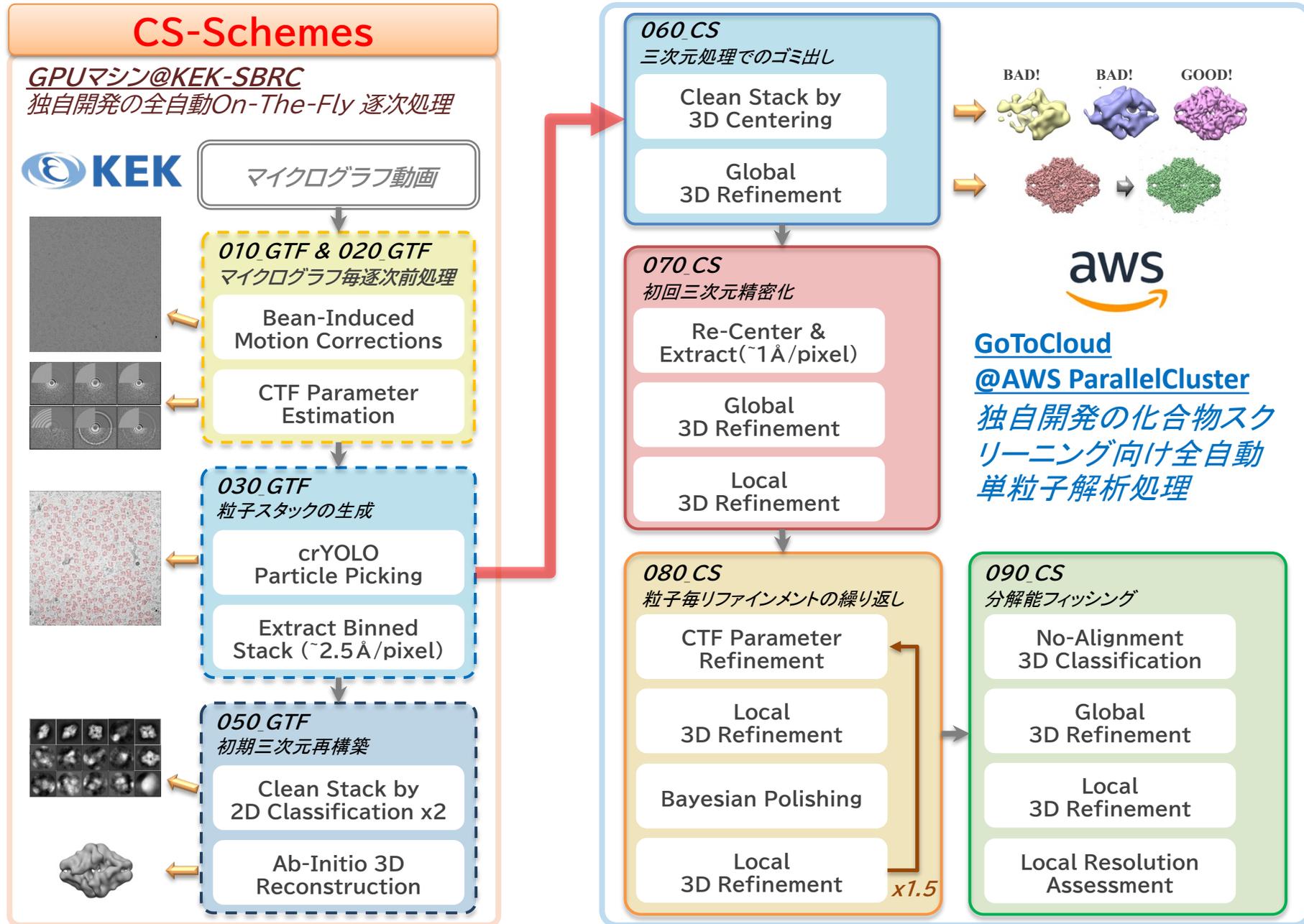
The screenshot shows a snippet of a webpage. The text is partially obscured but includes the following visible content: "am—a group of experienced PhD-level structural biologists and skilled cloud engineers. We infrastructure expertise to deliver responsive, informed support for your scientific workflows." Below this is another paragraph: "structural biology research. It complements SBGrid's long-standing on-premise software distribution labs without access to local HPC resources or those looking to scale beyond them." At the bottom, there is a line: "can support your research."

クラウドベースのプラットフォームは今後も多数登場しそう・・・

しかし、
ユーザーは依然としてクライオ電顕単粒子解析データセットの
データ解析方法を学ぶ必要がある・・・

**クライオ電顕単粒子解析ワークフローの
自動化が必須！**

KEK独自の全自動データ解析ワークフロー



処理時間およびAWSコスト算出用スクリプト

RELIONジョブ情報

並列計算

処理時間

EC2インスタンス情報

AWSコスト

JobType/JobID	Job Name	GPUs	MPI	Thread	MPIs/Node	Nodes	Process time (hh:mm:ss)	Process time (hours)	Running time (hh:mm:ss)	Running time (hours)	Time diff (minutes)	EC2 Instance Name	EC2 Instance Type	Cores/node	Cost/hours (USD)	Process Cost (USD)	Running Cost (USD)
Import/job001/	010010_Import_movies					0	0:00:00	0.0000	0:00:00	0.0000	0.00					\$0.0000	\$0.0000
MotionCorr/job002/	010020_MotionCorr		24	8	24	1	0:36:16	0.6040	0:30:04	0.5010	6.20	m7i-vcpu192-gpu0	m7i.48xlarge	192	9.6768	\$5.8490	\$4.8490
Select/job003/	010030_Select_mics					0	0:00:00	0.0000	0:00:00	0.0000	0.00					\$0.0000	\$0.0000
Select/job005/	020020_Select_mics					0	0:00:00	0.0000	0:00:00	0.0000	0.00					\$0.0000	\$0.0000
Select/job006/	020030_Select_icing					0	0:00:00	0.0000	0:00:00	0.0000	0.00					\$0.0000	\$0.0000
External/job007/	030020_External_cryolo	8		1	1	1	0:13:22	0.2230	0:05:58	0.0990	7.40	g6-vcpu192-gpu8	g6.48xlarge	192	13.3504	\$2.9740	\$1.3280
Extract/job008/	030040_Extract_picks		24		24	1	0:07:36	0.1270	0:01:36	0.0270	6.00	c7i-vcpu192-gpu0	c7i.48xlarge	192	8.568	\$1.0850	\$0.2280
Select/job009/	030050_Select_rm_bars_x					0	0:00:05	0.0010	0:00:05	0.0010	0.00					\$0.0000	\$0.0000
Select/job010/	030060_Select_rm_bars_xy					0	0:00:04	0.0010	0:00:04	0.0010	0.00					\$0.0000	\$0.0000
Import/job011/	060040_Import_ref3d					0	0:00:00	0.0000	0:00:00	0.0000	0.00					\$0.0000	\$0.0000
Import/job012/	060050_Import_mask3d					0	0:00:00	0.0000	0:00:00	0.0000	0.00					\$0.0000	\$0.0000
Class3D/job013/	060060_Class3D	8	33	12	16	2	1:04:04	1.0680	0:56:42	0.9450	7.40	g6-vcpu192-gpu8	g6.48xlarge	192	13.3504	\$28.5110	\$25.2320
External/job014/	060070_External_select3d			1		0	0:00:11	0.0030	0:00:11	0.0030	0.00					\$0.0000	\$0.0000
Refine3D/job015/	060080_Refine3D_global	8	17	12	8	2	0:08:20	0.1390	0:08:20	0.1390	0.00	g6-vcpu192-gpu8	g6.48xlarge	192	13.3504	\$3.7080	\$3.7080
PostProcess/job016/	060090_PostProcess_global					0	0:00:01	0.0000	0:00:01	0.0000	0.00					\$0.0000	\$0.0000
							⋮				⋮						
Polish/job032/	080060_Polish		16	12	16	1	0:46:40	0.7780	0:40:38	0.6770	6.00	m7i-vcpu192-gpu0	m7i.48xlarge	192	9.6768	\$7.5260	\$6.5530
Refine3D/job033/	080070_Refine3D_polish	8	17	12	8	2	0:56:00	0.9330	0:35:05	0.5850	20.90	g6-vcpu192-gpu8	g6.48xlarge	192	13.3504	\$24.9210	\$15.6130
PostProcess/job034/	080080_PostProcess_polish					0	0:00:38	0.0110	0:00:38	0.0110	0.00					\$0.0000	\$0.0000
CtfRefine/job035/	080010_CtfRefine_aberration		24	8	24	1	0:08:09	0.1360	0:02:24	0.0400	5.80	m7i-vcpu192-gpu0	m7i.48xlarge	192	9.6768	\$1.3140	\$0.3870
CtfRefine/job036/	080020_CtfRefine_aniso_mag		24	8	24	1	0:07:54	0.1320	0:02:19	0.0390	5.60	m7i-vcpu192-gpu0	m7i.48xlarge	192	9.6768	\$1.2740	\$0.3740
CtfRefine/job037/	080030_CtfRefine_ctf_params		24	8	24	1	0:02:26	0.0410	0:02:27	0.0410	0.00	m7i-vcpu192-gpu0	m7i.48xlarge	192	9.6768	\$0.3920	\$0.3950
Refine3D/job038/	080040_Refine3D_ctfrefine	8	17	12	8	2	1:32:56	1.5490	0:44:10	0.7360	48.80	g6-vcpu192-gpu8	g6.48xlarge	192	13.3504	\$41.3570	\$19.6550
PostProcess/job039/	080050_PostProcess_ctfrefine					0	0:00:40	0.0110	0:00:40	0.0110	0.00					\$0.0000	\$0.0000
Class3D/job040/	090010_Class3D_noalign		25	7	25	1	1:18:22	1.3060	0:39:58	0.6660	38.40	c7i-vcpu192-gpu0	c7i.48xlarge	192	8.568	\$11.1910	\$5.7070
External/job041/	090020_External_select3d			1		0	0:00:06	0.0020	0:00:06	0.0020	0.00					\$0.0000	\$0.0000
Refine3D/job042/	090030_Refine3D_global	8	17	12	8	2	0:42:43	0.7120	0:21:51	0.3640	20.90	g6-vcpu192-gpu8	g6.48xlarge	192	13.3504	\$19.0090	\$9.7240
PostProcess/job043/	090040_PostProcess_global					0	0:00:39	0.0110	0:00:39	0.0110	0.00					\$0.0000	\$0.0000
Refine3D/job044/	090050_Refine3D_local	8	17	12	8	2	0:20:53	0.3480	0:20:53	0.3480	0.00	g6-vcpu192-gpu8	g6.48xlarge	192	13.3504	\$9.2930	\$9.2930
PostProcess/job045/	090060_PostProcess_local					0	0:00:37	0.0100	0:00:37	0.0100	0.00					\$0.0000	\$0.0000
LocalRes/job046/	090070_LocalRes		24		24	1	1:58:48	1.9800	0:15:46	0.2630	103.00	m7i-vcpu192-gpu0	m7i.48xlarge	192	9.6768	\$19.1600	\$2.5430
							Total Time	13.4090			7.4200				Total Cost	\$245.1260	\$152.1980

合計処理時間

合計AWSコスト

CS-Schems実行結果: 亜硝酸還元酵素(NiR) データセット

EMPIAR-10581: Nitrite Reductase at pH 8.1 (110 kDa, C3)

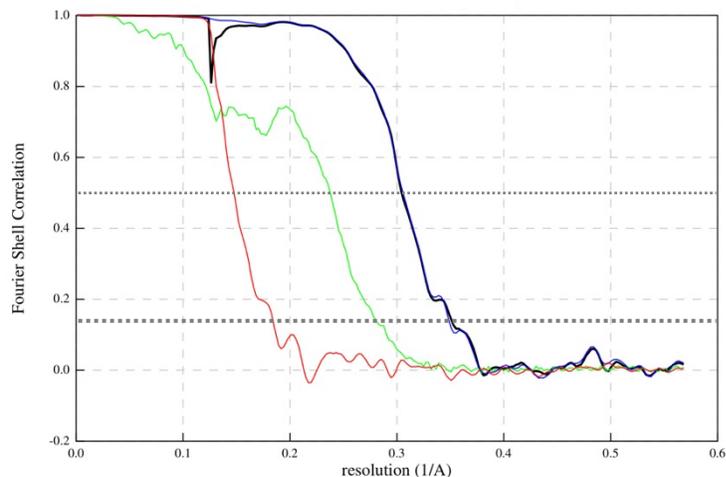
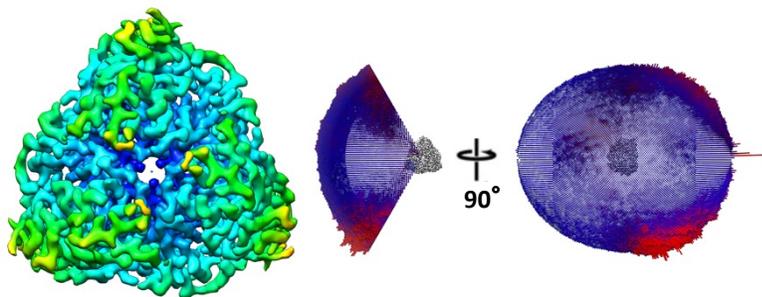
Imaging by Talos Arctica & Falcon 3EC

Voltage [kV]	: 200	Pixel size [$\text{\AA}/\text{pixels}$]	: 0.88	Exposure rate [$e/\text{\AA}^2/\text{frame}$]	: 1.02	Frames per image [frames]	: 49
Magnification	: 120K	Total exposure [$e/\text{\AA}$]	: 50	Defocus range [μm]	: -1.0 to -3.0	Number of micrographs	: 694

手動解析(我々の論文発表)

報告した分解能 2.85 \AA
合計処理時間 145.18 時間 (~6日)
89,513 粒子 (486 pix@0.88 \AA)

2018年12月

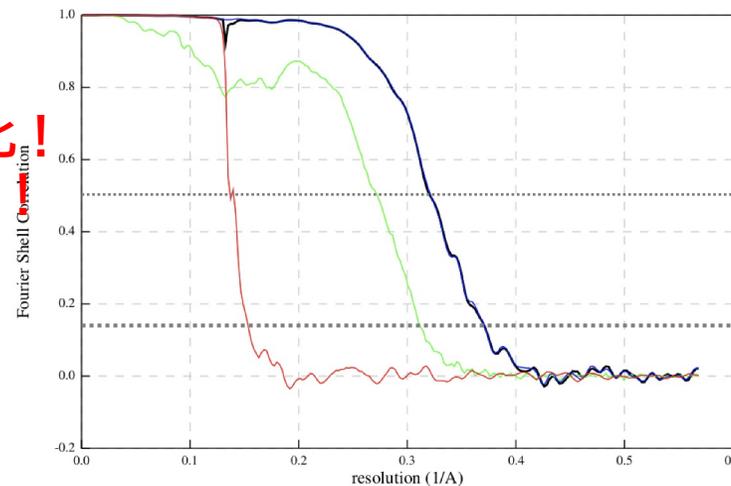
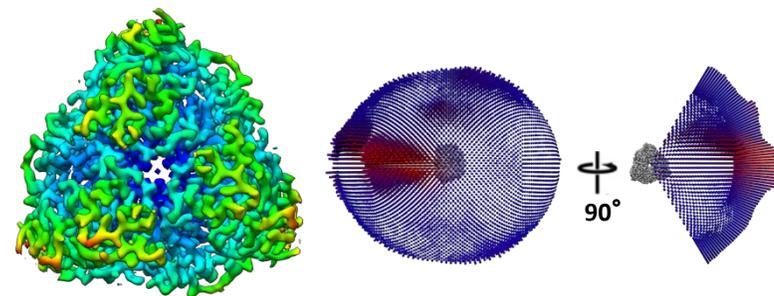


On-Premise GPU Box: RELION3, Ubuntu 18.04.1 LTS,
AMD Ryzen Threadripper 1950X x 1 (32 VCPUs), GeForce GTX 1080 Ti x 4

自動ワークフロー(CS-Schems)

ベスト分解能 2.69 \AA
合計処理時間 20.57 時間
74,609 粒子 (480 pix@0.88 \AA)

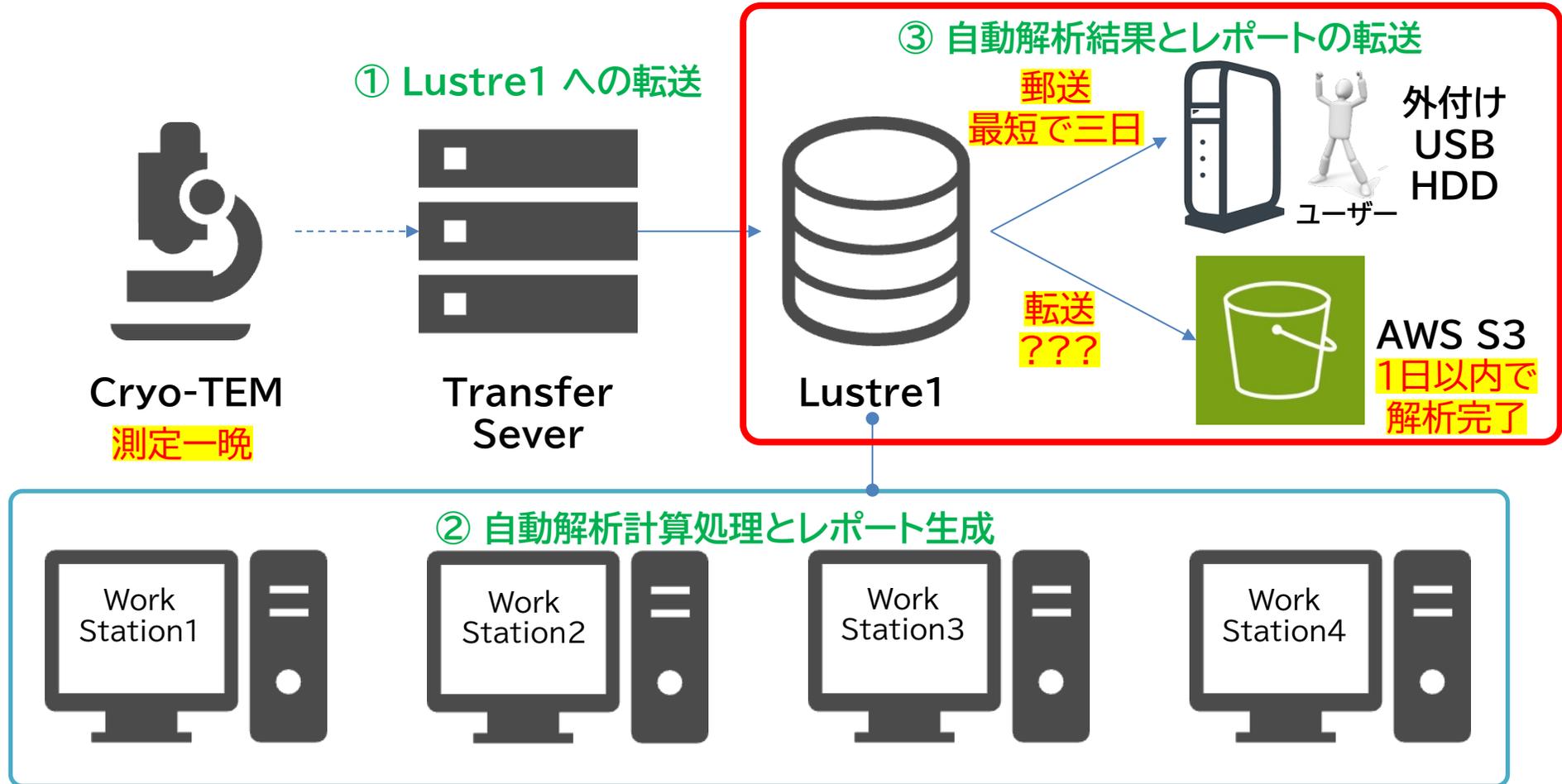
2023年05月
最適化前



GoToCloud: RELION4.0, Ubuntu2004,
AWS ParallelCluster ver. 3.0.3, AWS "us-east-1" region

7倍の高速化!
分解能向上

現行のデータ測定／解析支援のボトルネック ～データ転送の高速化～



【現状】測定データをHDDで受け渡し:データ測定終了から**「最短で三日後」**に解析開始可能

1. HDDにOn-The-Flyでコピー :9時半までには測定終了、お昼頃までにコピー完了
2. 発送:その日のうち、**忙しい時は翌日** ← Overhead !
3. 郵送:ユーザーのところに到着するのは、**【国内】通常は翌日、最悪で二日、【海外】1～2週間?!** ← Overhead !
4. ユーザーサイトでのHDDからコピー:**日中一杯(6～8時間)** ← Overhead !

高速データ転送技術 “Zettar zx”

intel®

Introduced by
Dr. Michael J Mcmanus

 Zettar Unmatched zx Transfer Speeds, Unaffected by File Size, Encryption, Latency

【技術の特徴】

- 大容量データ(TB/PB)を簡単かつ効率的に転送
- 内蔵バッファを使用し、通信距離のインパクトを取り除く
- 通信開始までの遅延時間(レイテンシー)の影響を受けない

Confidential Material

[ベンチマーク]

- 1 U.S. DOE Energy Sciences Network Q4 2020 Data Mover Evaluation
- 2 2500マイル 100G Esnet テスト環境
- 3 代表的な従来技術のGlobus GridFTPがベースライン

SCA19 DMC

Winner; Zettar zx

2020 ESnet Evaluation: Outperforms



2021 U.S. DOE Energy Sciences Network

Zettar zx

v.s. Globus-Connect-Server v5

v.s. MdtmFTP

Notes:

1. MdtmFTP is FermiLab's attempt to "improve" GridFTP. It didn't succeed so the project stopped.
2. Both MdtmFTP and Globus needs separate tuning for different data sets. Zettar zx uses just one for all.

Picture credits: 2021 U.S. DOE Energy Sciences Network

Confidential Material

高速データ転送技術「Zettar zx」のPoC



守屋 俊夫

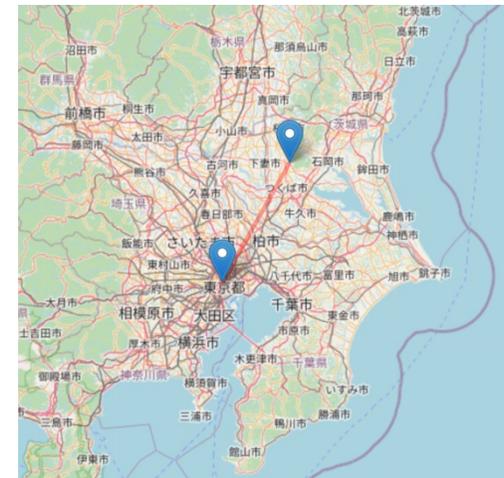


中村 司

KEK→AWS S3@東京

	aws-cli (AWS純正)		Zettar zx	
	転送レート (Gbps)	転送時間 (分)	転送レート (Gbps)	転送時間 (分)
1回目	2.02	75.13	6.71	22.67
2回目	1.99	76.33	6.72	22.62
3回目	1.99	76.07	6.70	22.7
平均	2.00	75.84	6.71	22.66
標準偏差	0.017	0.63	0.01	0.04

**3.34倍の
転送速度**



63キロ (39マイル)

KEKのアウトバウンド通信帯域 (定格): **10 Gbps**
NiR ベンチマークデータセットサイズ: **1.2 TiB**

KEK→AWS S3@北部バージニア

	aws-cli (AWS純正)		Zettar zx	
	転送レート (Gbps)	転送時間 (分)	転送レート (Gbps)	転送時間 (分)
1回目	-	238.55	4.10	44.57
2回目	-	234.35	4.74	38.52
3回目	-	232.65	4.96	36.83
平均	-	235.18	4.60	39.97
標準偏差	-	2.48	0.36	3.32

**5.88倍の
転送速度
Better!**



172倍の距離

3.1倍の転送時間



1.8倍の転送時間 Better!

10,851キロ(6,742マイル)

高速データ転送技術によるクライオ電顕施設間連携 ～公的データベース(DB)と共用大規模計算機環境～

【栗栖グループ】



大阪大学
OSAKA UNIVERSITY

蛋白質研究所
蛋白質構造
生物学研究部門



2025年7月から
運用開始！

The diagram illustrates the Protein Data Bank (PDB) network and data flow. At the top, the 'WORLDWIDE PDB PROTEIN DATA BANK' logo is shown with the website 'www.wwPDB.org' and email 'info@wwPDB.org'. Below this, a world map shows regional nodes: 'RCSB PDB PROTEIN DATA BANK' (Research Collaboratory for Structural Bioinformatics, www.pdb.org), 'PDBe Protein Data Bank in Europe' (Protein Data Bank in Europe, pdbe.org), and 'PDBj Protein Data Bank Japan' (Protein Data Bank Japan, www.pdbj.org). A 'BMRB BioMagResBank' (www.bmrwisc.edu) is also indicated. To the right, a vertical flow shows data processing: 'EMPIAR' (生データ画像(入力)) -> 'EMDB Electron Microscopy Data Bank' (最終出力密度マップ) -> 'WORLDWIDE PDB PROTEIN DATA BANK' (最終出力原子座標モデル).

公的データベース(DB)の「PDBj」を運用／高度化

The diagram shows the architecture of the 'PREWS' high-performance GPU cluster. It consists of three main components: '分散ストレージ' (Distributed Storage) with a physical capacity of 2,772 TB, connected via 25Gb links; 'GPUノードx30台' (30 GPU Nodes) with GPU: A5000 (24GB) x 4, 64 cores, and 512 GB RAM, connected via 40Gb links; and '管理サーバー' (Management Servers) including 'CPUノードx10台' (10 CPU Nodes) with 56 cores and 256 GB RAM, connected via 10Gb links. A photograph of server racks is shown on the left.

高性能GPUクラスター「PREWS」を構築／運営

データ解析支援の高度化 ～阪大・PREWS GPUクラスタの統合～

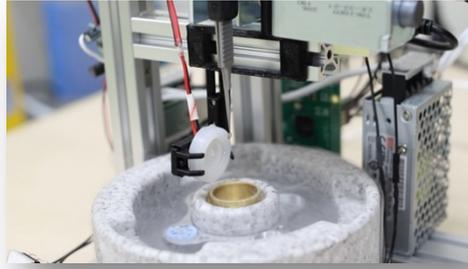
クライオ電顕



自動
データ転送

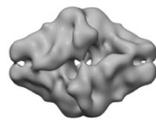
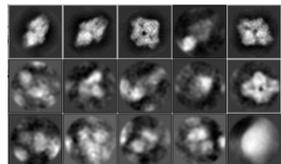
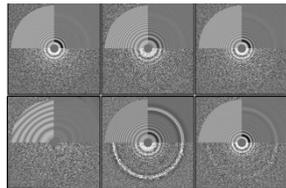
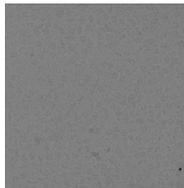


グリッド凍結



【目標】
PREWSクラスタでのデータ解析を
撮影終了後2～3時間以内に開始！
1.0～1.5日以内に計算完了！

「GoToFly(JK)」@KEK-SBRC
独自開発の全自動On-The-Fly逐次処理



*Multi-MD: Multiple Mask Diameters

自動高速
データ転送



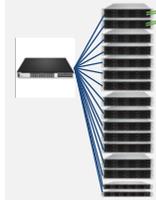
PREWSクラスタ@阪大
Cryo-EM SPA向け
高性能GPUクラスタ



分散ストレージ
2,772 TB

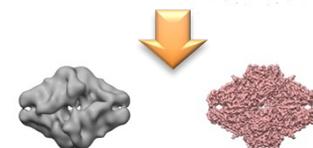
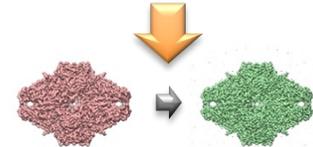
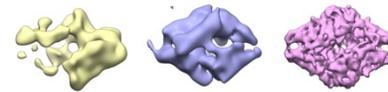
GPUノードx30台

管理サーバー



CPUノードx10台

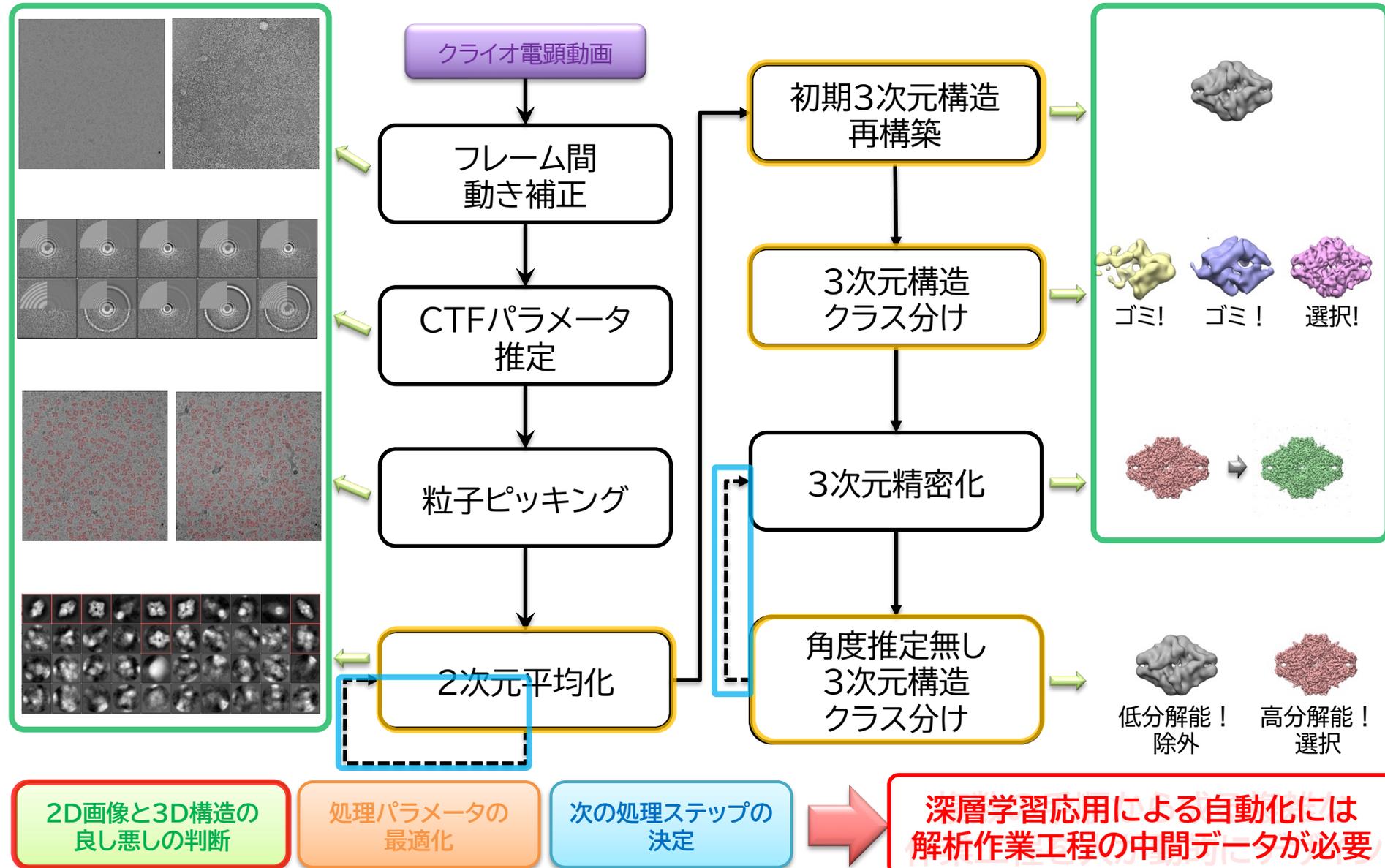
「CS-Schemes」@GoToCloud
独自開発の創薬向け
全自動Cryo-EM SPA解析処理



2025年度
開発

2026年度
運用予定

クライオ電顕単粒子解析のデータ解析の課題と自律実験



成長型知能システムの成長スパイラル

日々のクライオ電顕施設の運用を通じて
解析作業工程の中間データを蓄積し、
今後のAI応用開発の学習データとして活用！

統合

【既存の公的DB】

最終出力モデル

最終出力マップ

入力動画

【データ測定】

入力動画

自律実験による
動的全自動データ解析

【洗練期】

実践経験を蓄積する

【自立期】

独学で試行錯誤を行う

【模倣期】

人間から学ぶ

CS-Schemesによる
静的全自動データ解析

解析熟練者による
手動データ解析

【新規の公的DB】

解析工程記録

新規モデル

解析工程記録

新規モデル

解析工程記録

新規モデル

解析工程記録

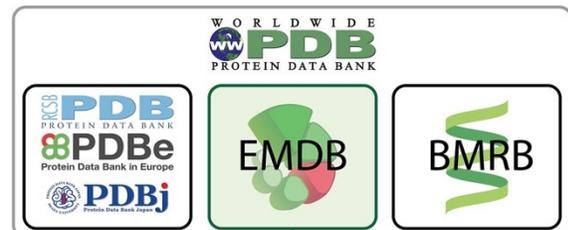
解析工程記録

知能システムの
トレーニング

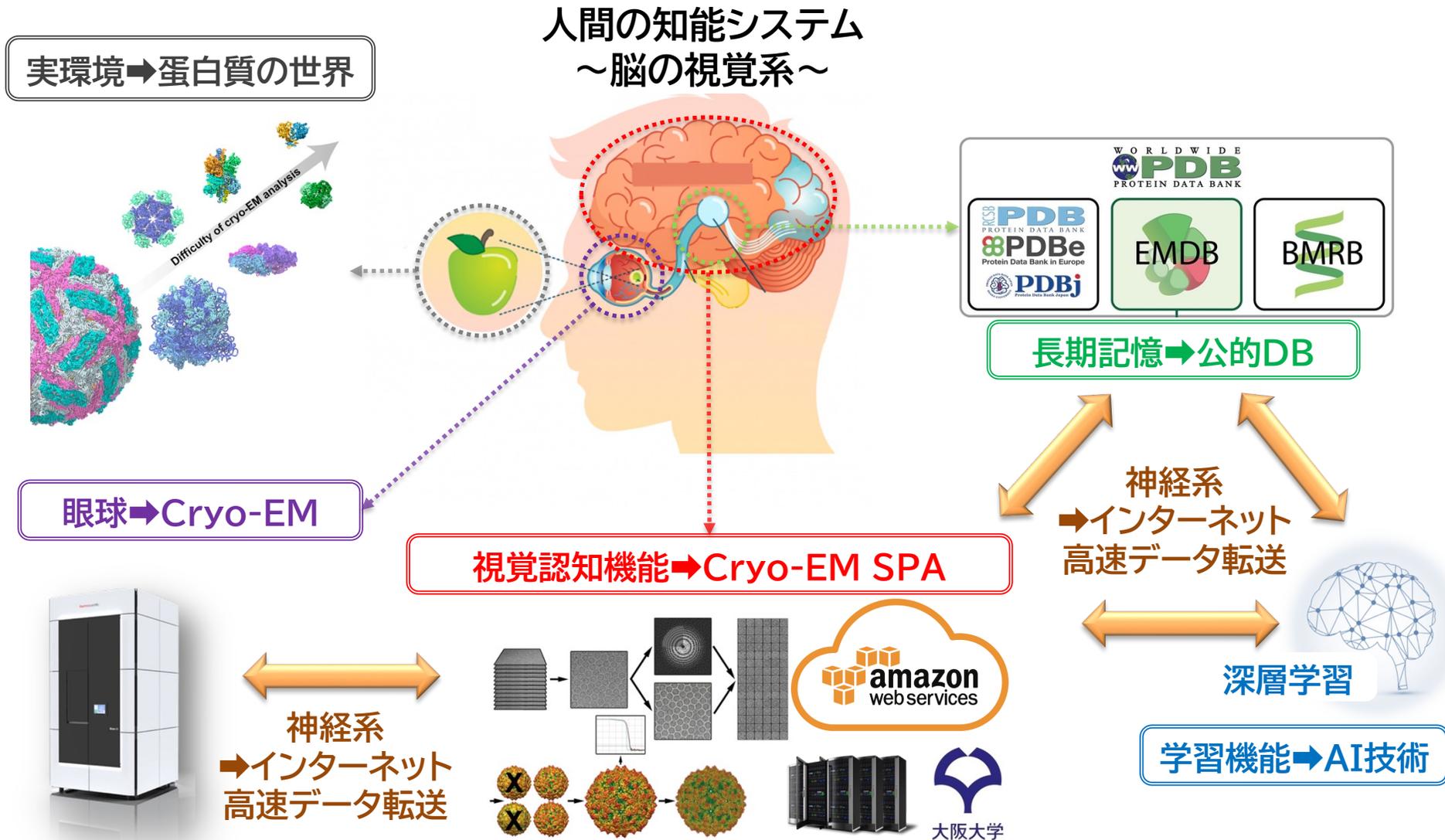
トレーニング(3)

トレーニング(2)

トレーニング(1)

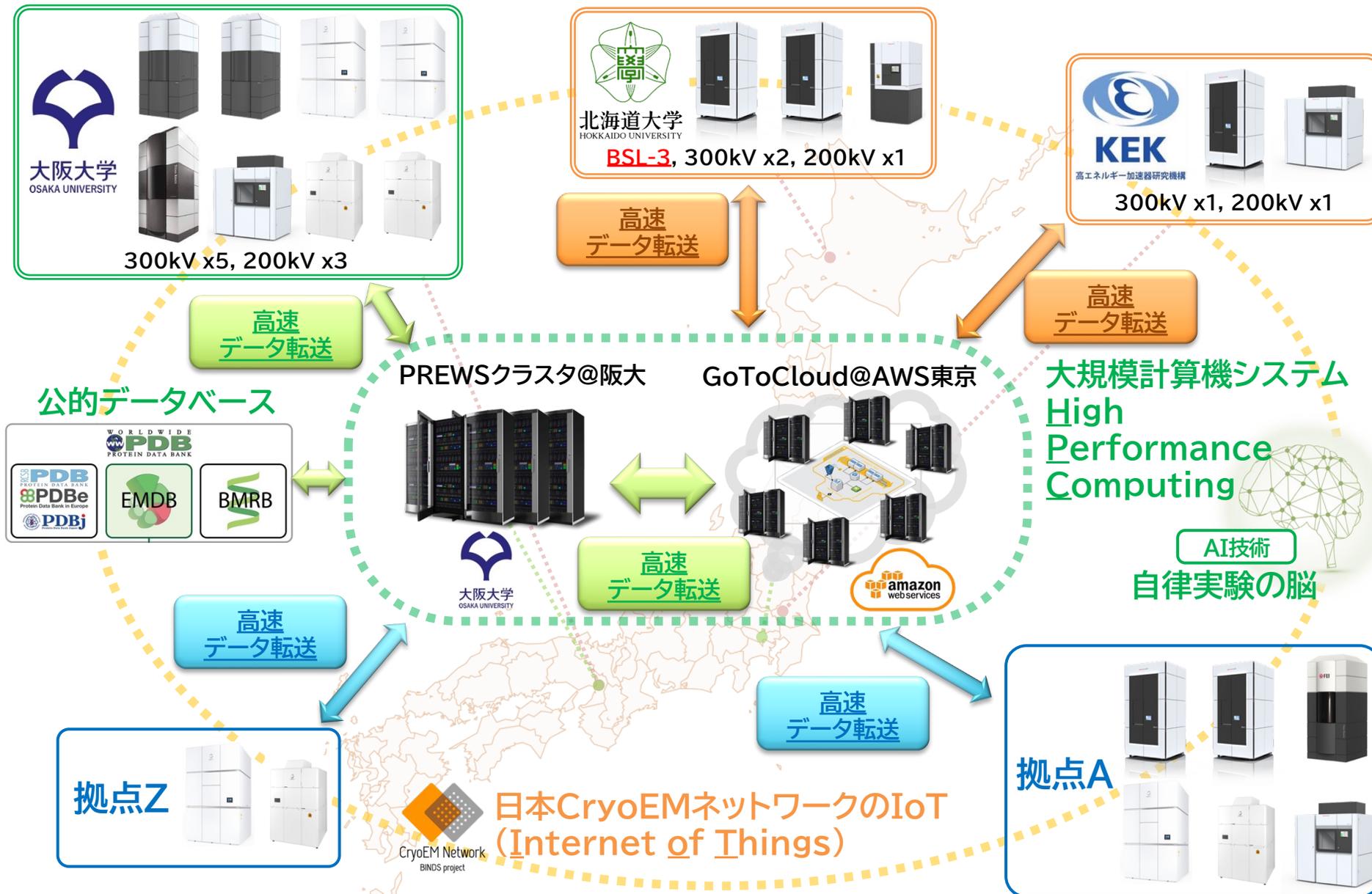


深層学習応用による自動化における 成長型知能システムのマッピング



今、AWSとHPCを利用するシステムの神経系である「高速データ転送」が重要課題！

自律実験によるCryo-EM IoTの全自動運営



国内外の産学連携



生存ダイナミクス研究センター (TARA)

岩崎 憲治



計算科学研究センター

重田 育照



大学院理学研究院
化学研究部門

村田 武士
宮下 靖臣



Department of Biological Sciences
Department of Computer Science

Daisuke Kihara (木原 大亮)
Han Zhu



Harvard Medical School
SBGrid Consortium

Piotr Sliz
Jason Key
Ben Eisenbraun



柳生 さおり
宮本 大輔
佐々木 けい



(元) 矢澤 克巳



田口 武慶



(元) Michael J Mcmanus



Stephen Litster
Jacob Mevorach



detecting the future
Max Burian



Mazdak Radjainia

❄️🔬 ご清聴ありがとうございました 🔬❄️

【KEK-SBRC クライオ電顕チーム】



WHAT IS LIFE?
STRUCTURAL BIOLOGY RESEARCH CENTER



千田 俊哉



川崎 政人



守屋 俊夫



田辺 幹雄



藤井 裕己



中村 司



高巢 晃



増田 千穂



坪松 卓耶
(技術支援員)



李 慶武
(技術支援員)



Jair Pereira J.
(RA 筑波大)



山田 悠介
(東北大)

【PI】
千田 俊哉

【Secretary】
増田 千穂

【Operation Team】

川崎 政人
藤井 裕己
坪松 卓耶
李 慶武

【Analysis Team】

守屋 俊夫
中村 司
Jair Pereira J.

【Support Team】

(Wet/HW/IT)
田辺 幹雄
高巢 晃
山田 悠介

【Cross-Team Activities】

【Analysis Support】
(As Research Collaboration)

守屋 俊夫
川崎 政人
藤井 裕己
中村 司

【GoToCloud@AWS Project】

守屋 俊夫
中村 司
Jair Pereira J.
山田 悠介(東北大)

【Grid Prep. Project】

川崎 政人 (Antifreeze Peptide)
高巢 晃 (Vitrification)
藤井 裕己 (Vitrification)

【Sample Prep. Project】

田辺 幹雄 (Nano-disc, Tomography etc.)