

SPring-8/SACLAにおける データ基盤の開発

城地 保昌

JASRI/理研

大型放射光施設 SPring-8
Super Photon ring-8GeV



X線自由電子レーザー施設 SACLA
Spring-8 Angstrom Compact Free Electron LAser

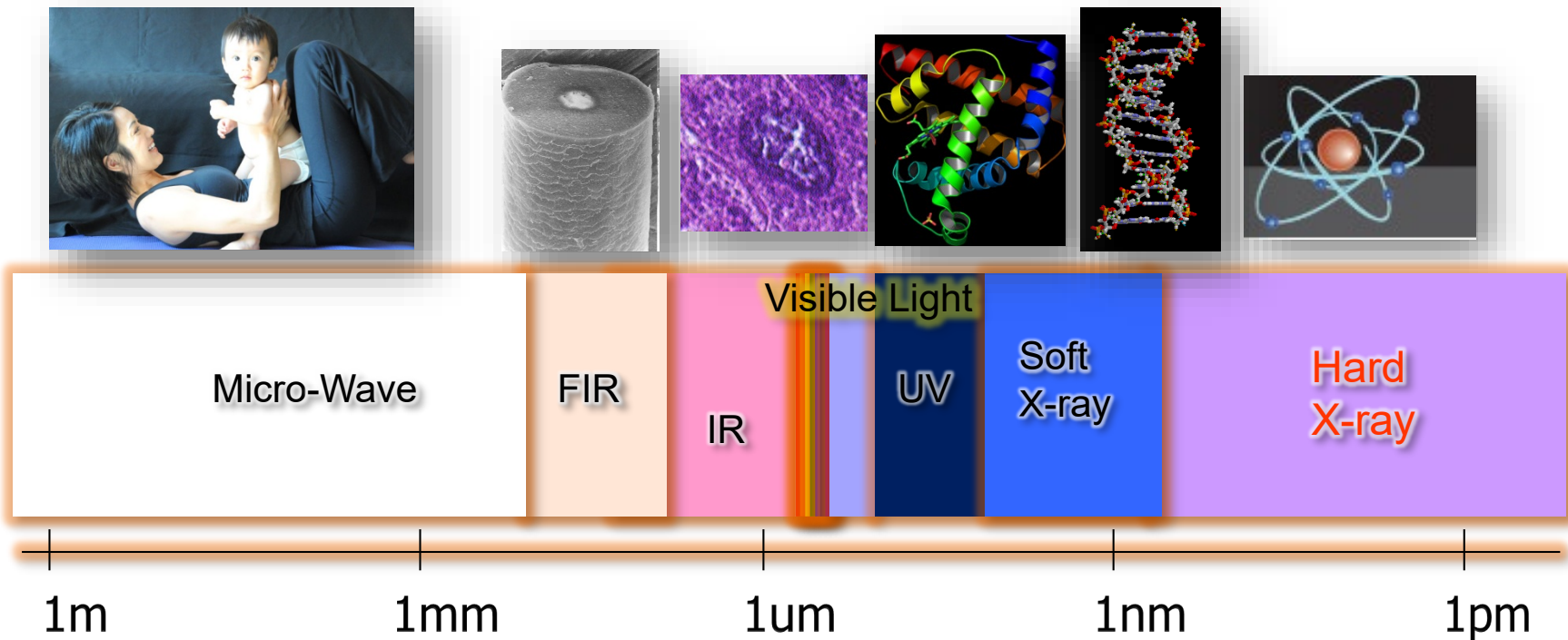
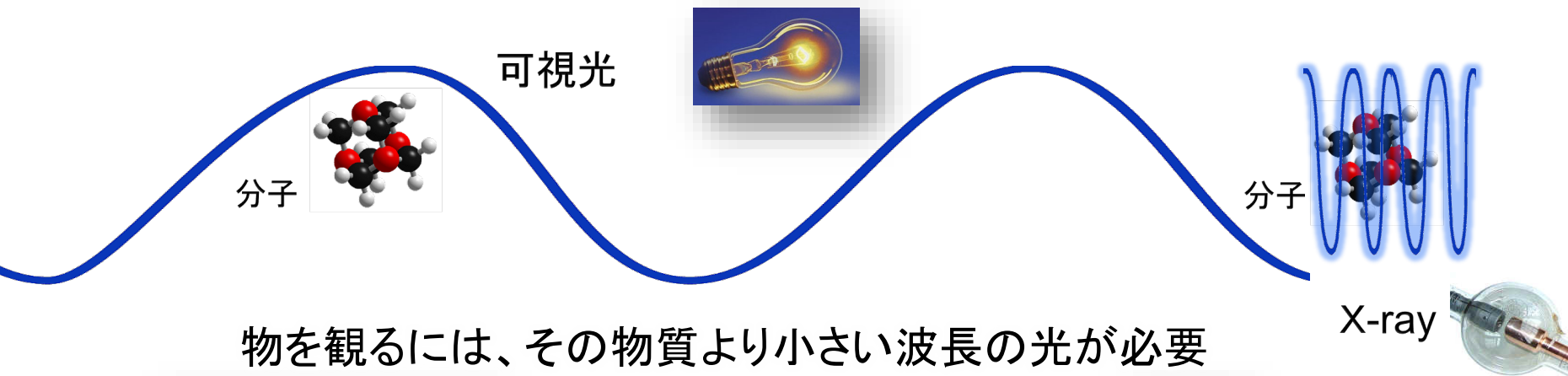


Contents

- SPring-8/SACLA
 - ✓ X線サイエンス
 - ✓ 大型放射光施設SPring-8
 - ✓ X線自由電子レーザー(XFEL)施設SACLA
- SACLAのデータ基盤開発
 - ✓ データ収集・解析基盤
 - ✓ データ管理
 - ✓ システム更新
- SPring-8高度化に向けたデータ基盤開発
 - ✓ 展望
 - ✓ 現状の取り組み

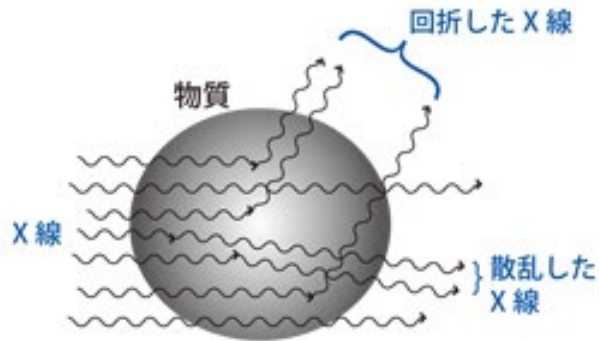
To see invisible

Why is wavelength important?

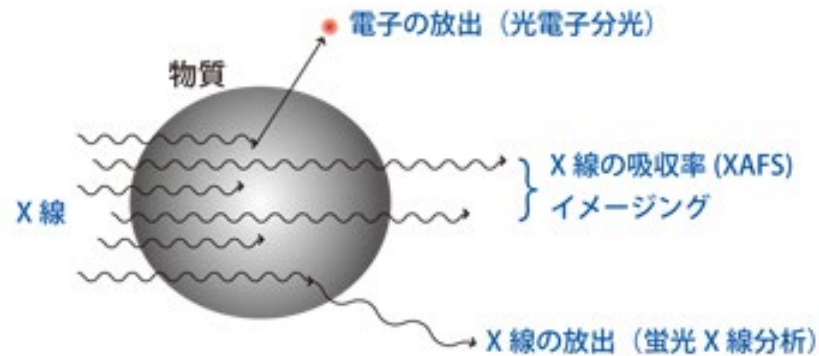


X線と物質の相互作用を利用して物質の微細構造を観る

(a) X線の回折・散乱



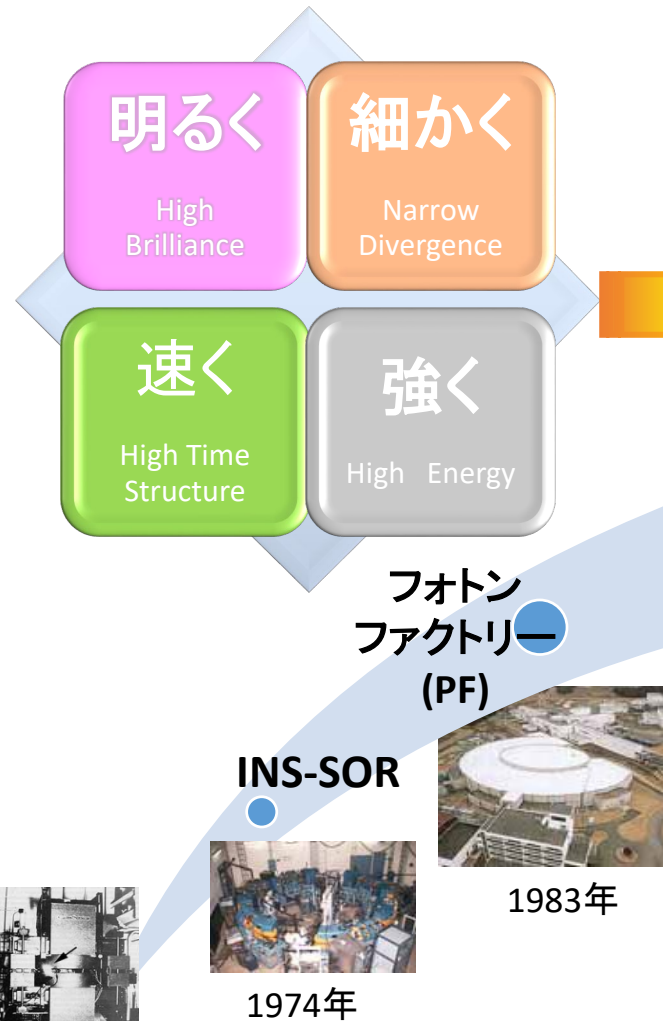
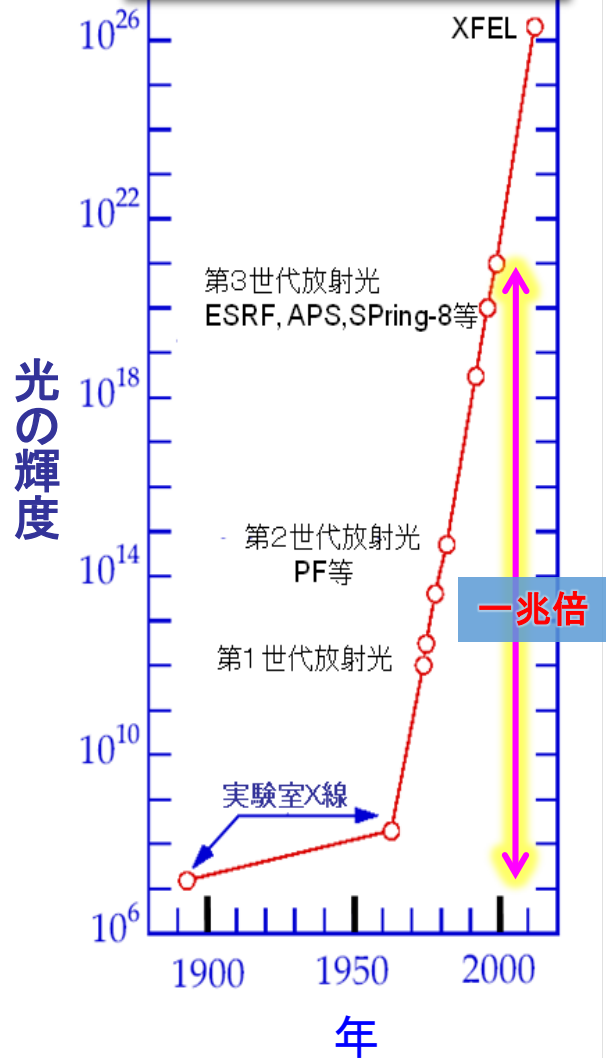
(b) X線エネルギーの吸収



- ライフサイエンス/生命の謎を分子レベルで解明する
- ソフトマター/ポリマーと有機物表面のナノ・メソ構造を見る
- 構造物性/原子と分子の並び方から物質の機能を知る
- 電子物性/機能発現に関わる電子の動きを見る
- 高圧地球科学/地球深部の物質の構造を探る
- 環境・エネルギー/人類生存の礎をミクロに解明する
- 核物理/高エネルギー光ビームで物質の究極を探る

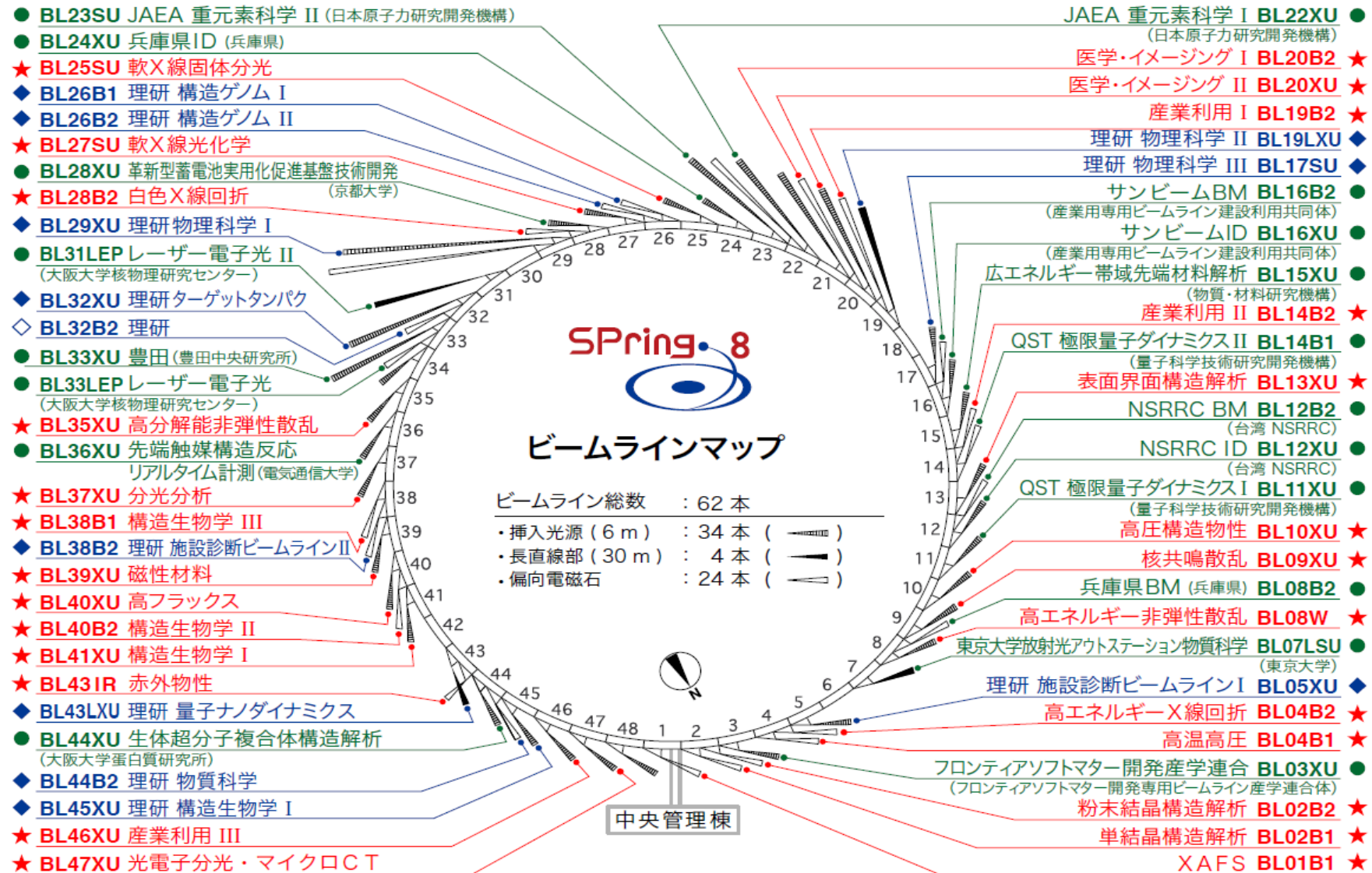
光の性能へさらに高まる要望と技術革新

より明るい光を
求めて



★：共用ビームライン
●：専用ビームライン
◆：理研ビームライン

(注) ☆、○、◇：計画・調整、建設中



SPring-8の製品開発利用例



新規液体燃料電池自動車
(ダイハツ工業株式会社)



ニッケル水素電池
(ジーエス・ユアサコーポレーション)



低燃費タイヤの開発
(住友ゴム工業)



軽量気泡コンクリート
(旭化成／旭建材)



タングステン高効率リサイクル技
術(住友電気工業)

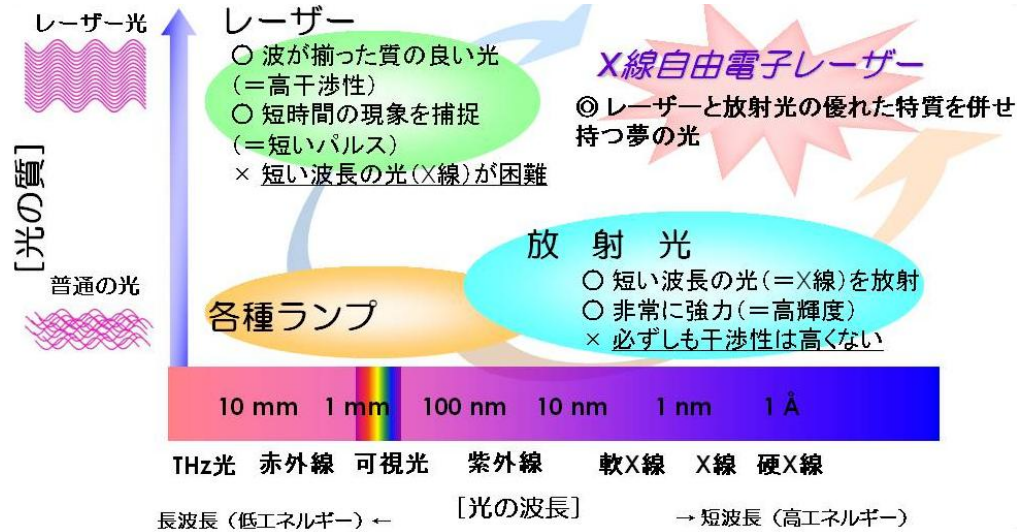


むし歯予防ガム(江崎グリコ)



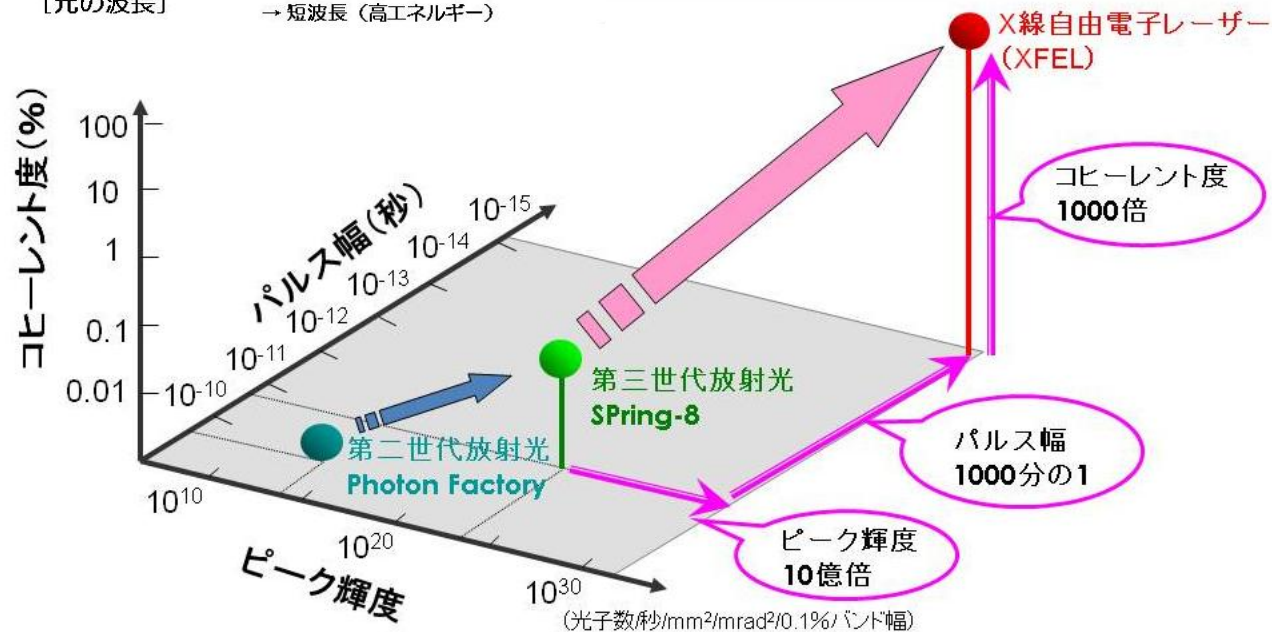
ヘアケア製品
(花王・P&G)

X線自由電子レーザー(XFEL)とは



To see Invisible X線自由電子レーザー が実現する 三つの「見る」

- ◎ 短い波長：原子レベルを見る。
- ◎ フェムト秒パルス：原子や分子の反応を見る。
- ◎ 高干渉性：分子や物質の形をありのままに見る。



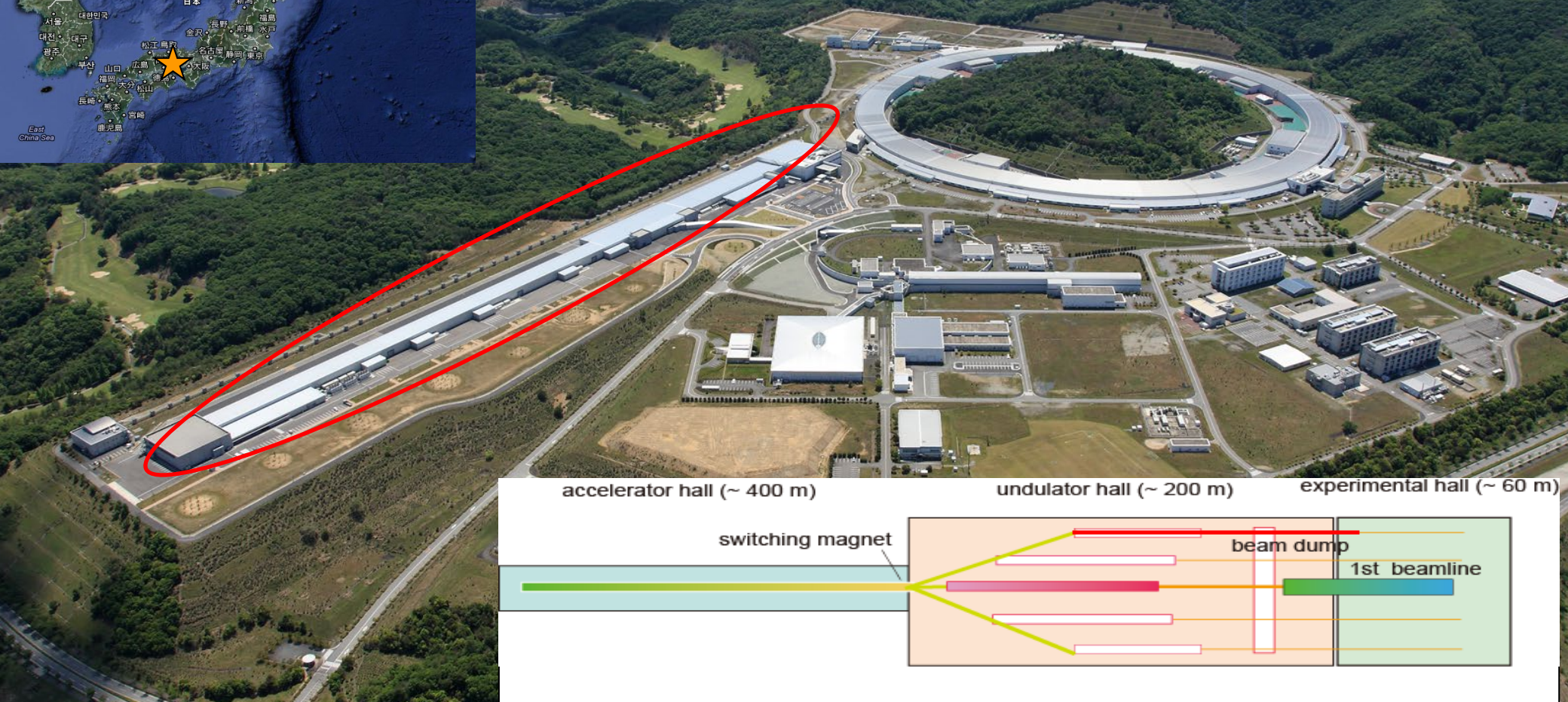
SACLA @SPring-8

First Compact XFEL

Construction: FY2006~2010

First lasing: June 7th 2011

User Operation: March 7th 2012 ~



accelerator hall (~ 400 m)

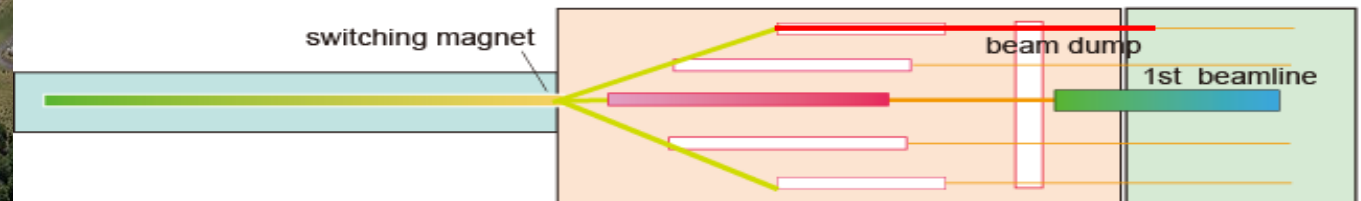
undulator hall (~ 200 m)

experimental hall (~ 60 m)

switching magnet

beam dump

1st beamline



XFEL光の特徴を活かした観測法

- 短時間パルス ($\sim \text{fs}$)
- 高い空間コヒーレンス
- 高輝度 (10^{11} photons/pulse)



- 原子や分子の運動に影響されない
- コヒーレンス利用による先端的ナノイメージング



- プローブ(X線)による損傷が起こる前の情報が得られる
- 化学反応、相転移などの高速現象を追跡可能
- 溶液など、実環境に近い状態下の試料が扱える

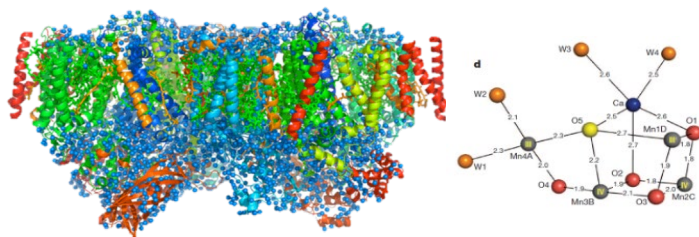
Key scientific results of SACLA in diverse fields

Damage-free protein crystallography



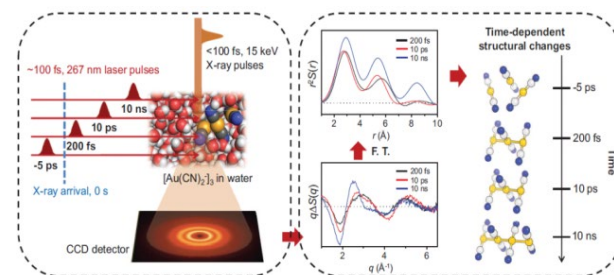
Prof. Shen
(Okayama U)

Photosystem II



M. Suga, et al., Nature 517, 99 (2015).

Ultrafast dynamics in chemical reaction



Prof. Adachi
(KEK)

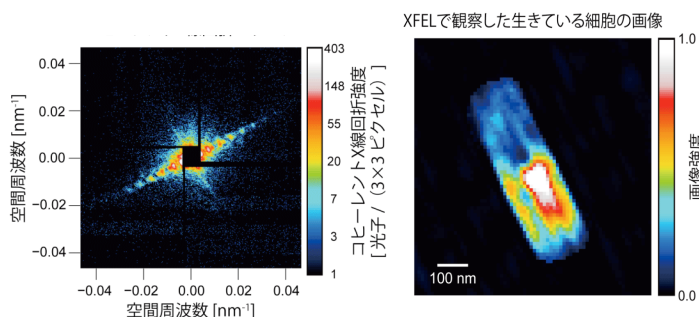
K. Kim, et al, Nature 518, 315 (2015).

Femtosecond snapshot of bioparticles



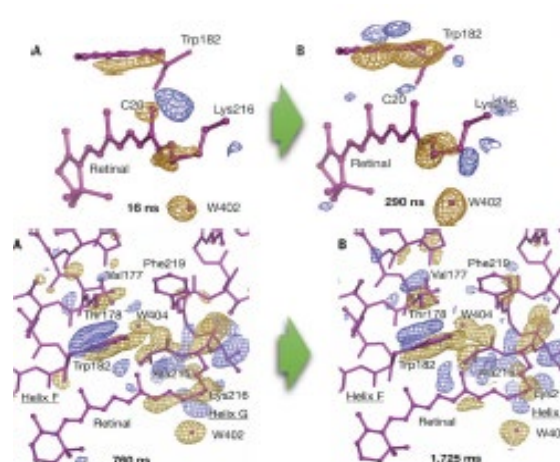
Prof. Nishino
(Hokkaido U)

Live cell



T. Kimura, Y. Joti, et al., Nature Comm. 5, 3052 (2014).

Protein structure dynamics



Prof. Iwada
(Kyoto U)

E. Nango, et al, Science 354, 1552 (2016).

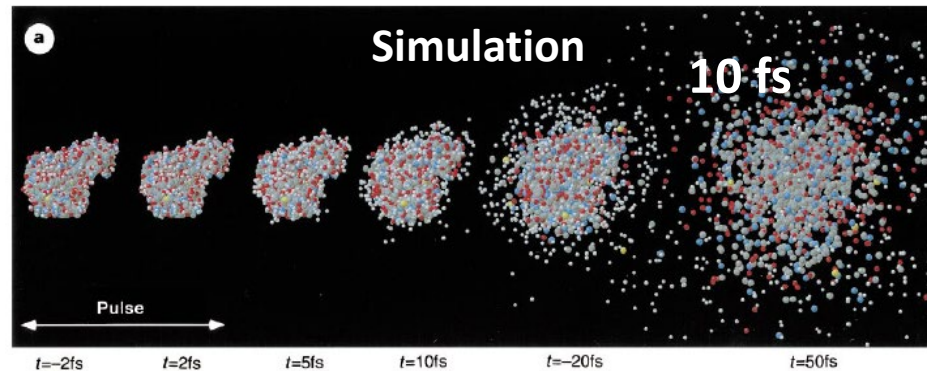
“Diffraction-before-Destruction”

Neutze et al., *Nature*
406, 752 (2000)



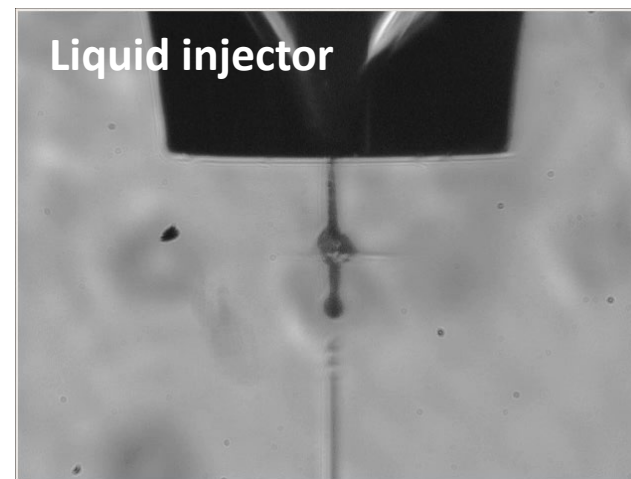
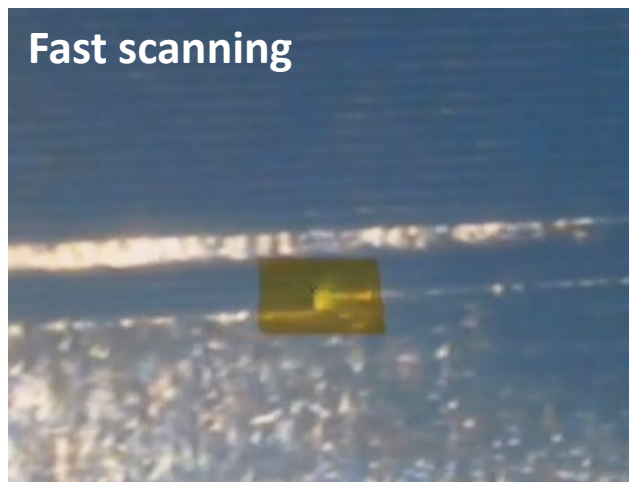
XFELs could outrun radiation damage process with ultrafast ($< 10\text{fs}$) pulses.

→ Diffraction signal from samples before destruction with XFEL.

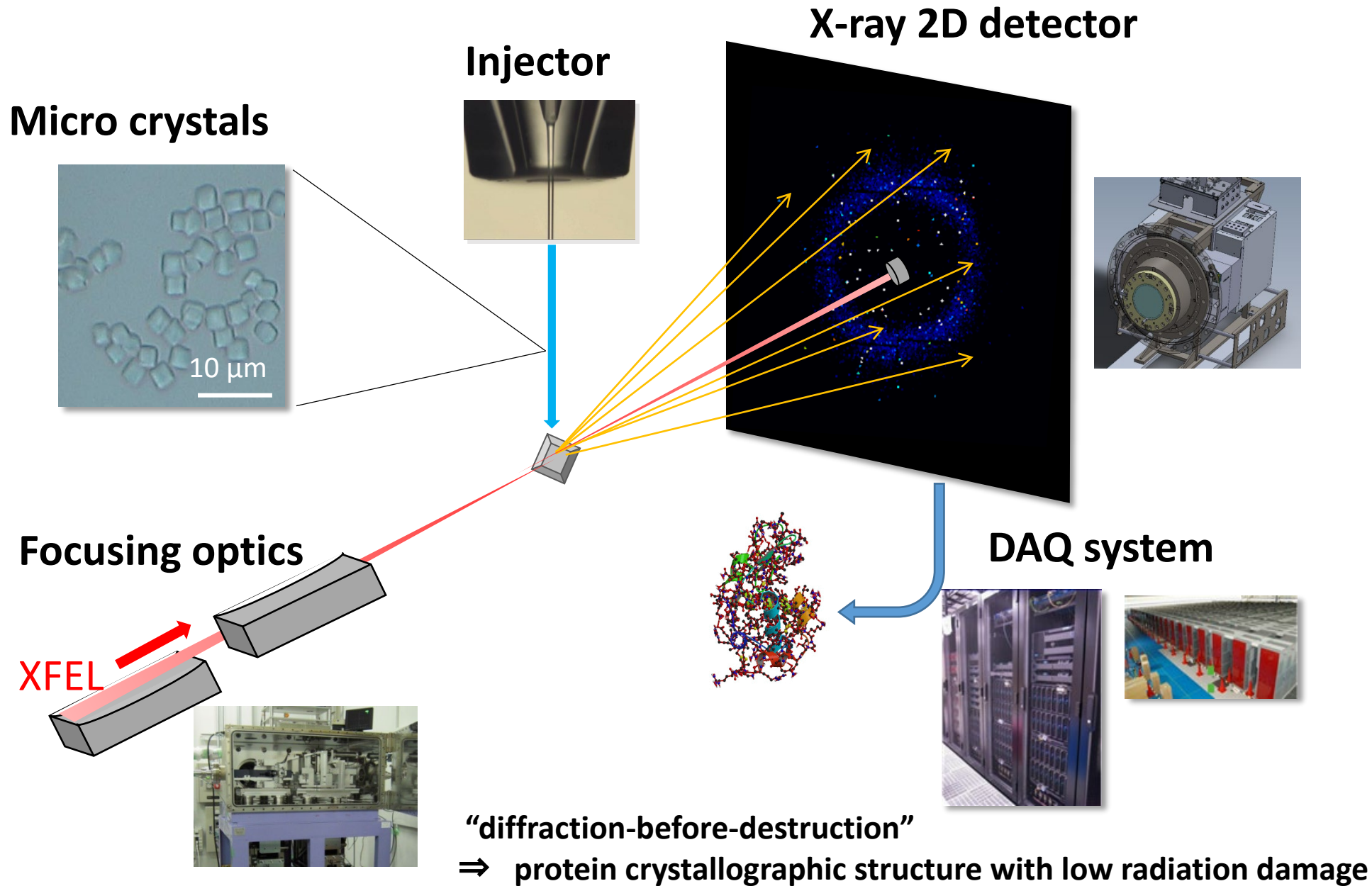


- femtosecond snapshots almost free from atomic and molecular motion
- fast phenomena such as chemical reaction in physiological condition.

This measurement scheme requires *single-pulse measurement* with pulse-by-pulse sample exchange.



Serial Femtosecond Crystallography (SFX)



SACLAのデータ基盤開発

SACLA DAQの要件

- XFELパルス毎に同期して実験計測を行う。最大スループットの設計値は6Gbps(MPCCDセンサー12個@60Hz)。

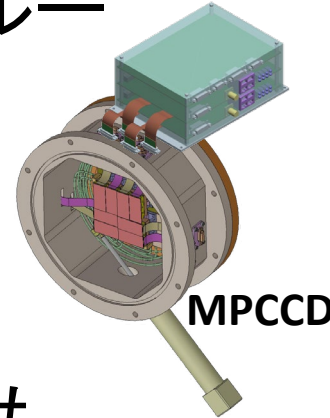
高速・確実なデータ保存

- SASEの特性によりパルス毎に揺らぐXFEL光の状態も記録し、サンプルからのシグナル解析の際に利用する。

相関データ解析

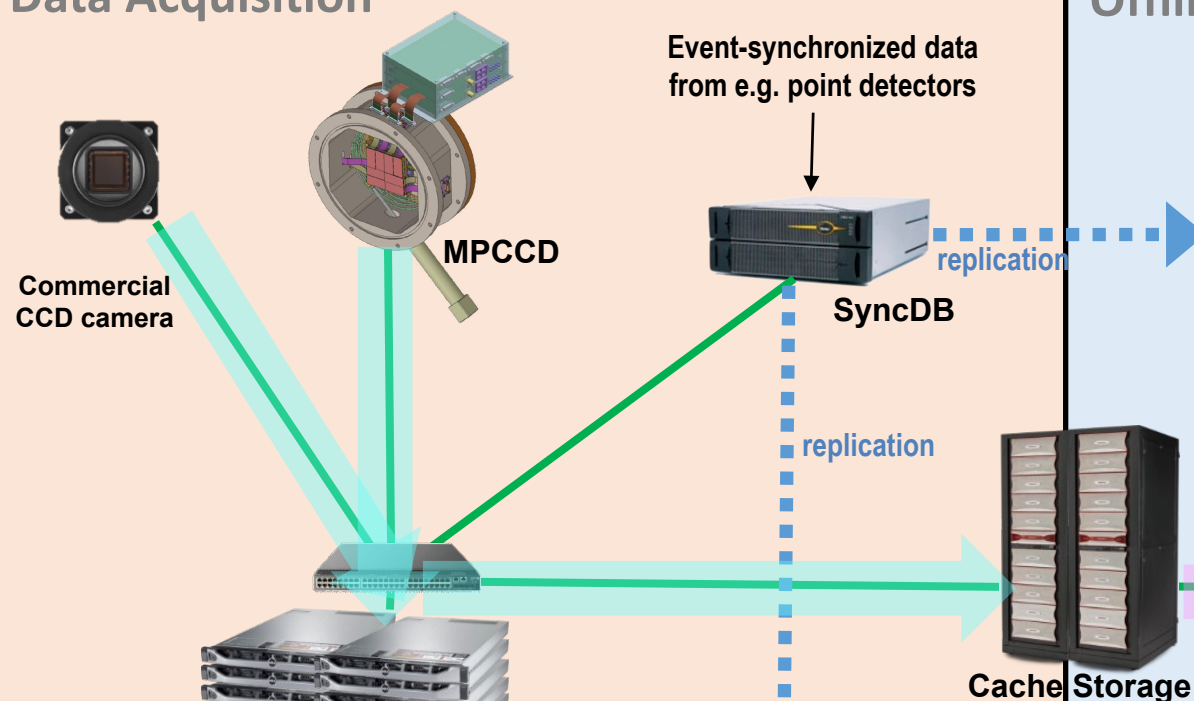
- 実験中にデータを確認しながら測定を行い、必要に応じて実験条件にフィードバックする。

良質データの取得のための効率的データ解析

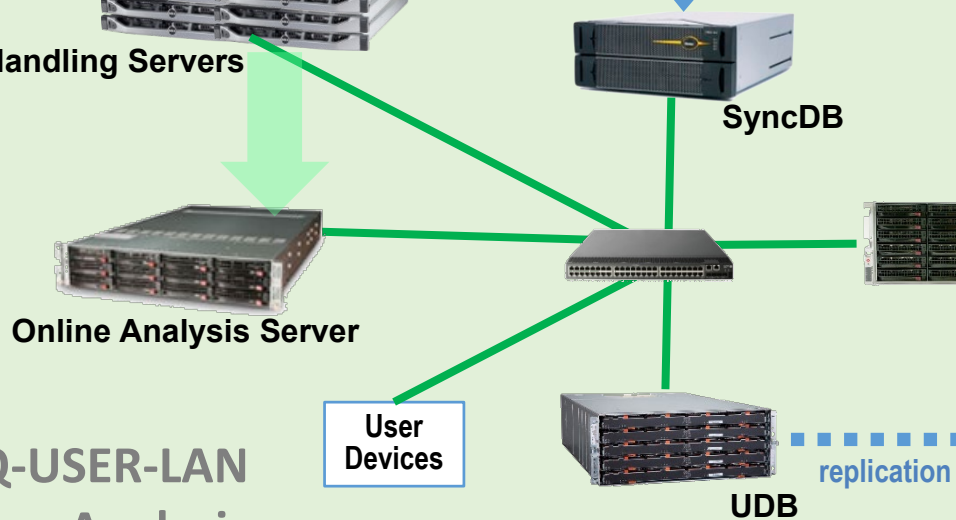


DAQ-LAN

Data Acquisition

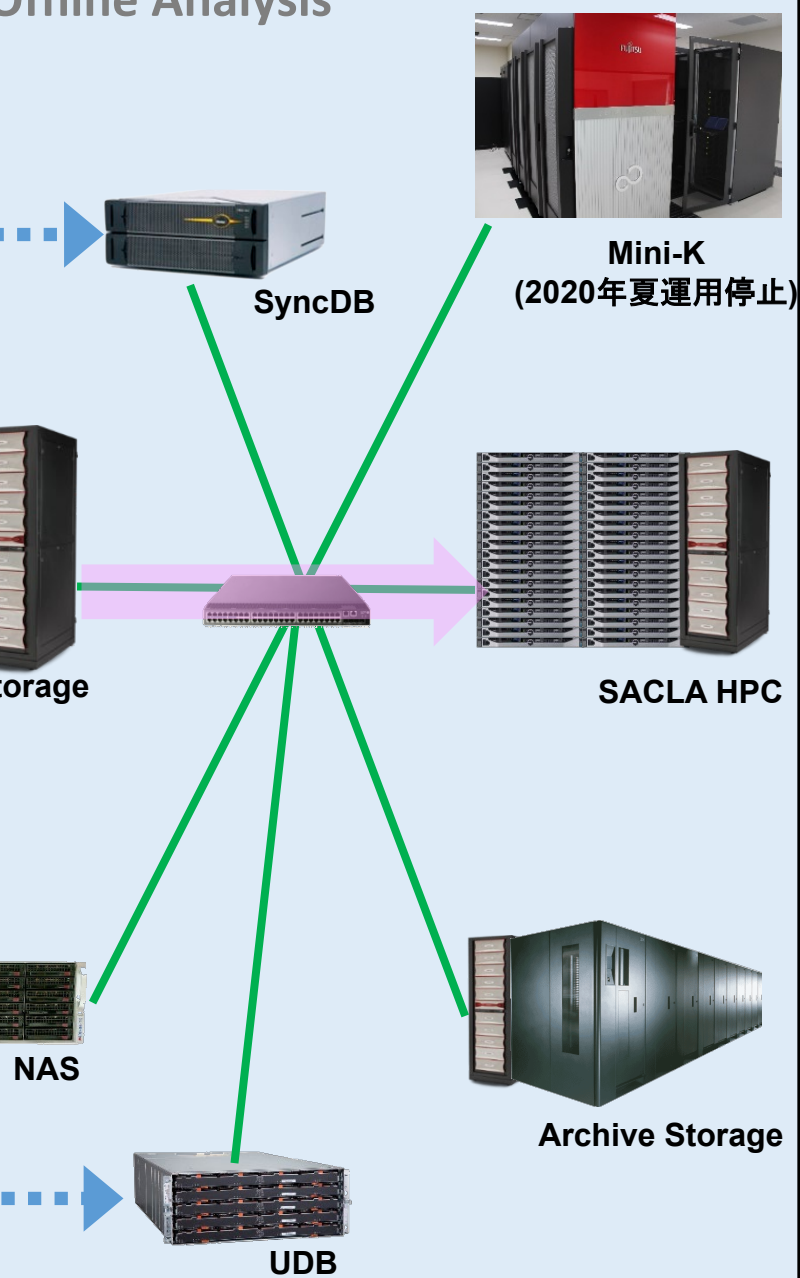


Data-Handling Servers



DAQ-USER-LAN Online Analysis

HPC-LAN Offline Analysis

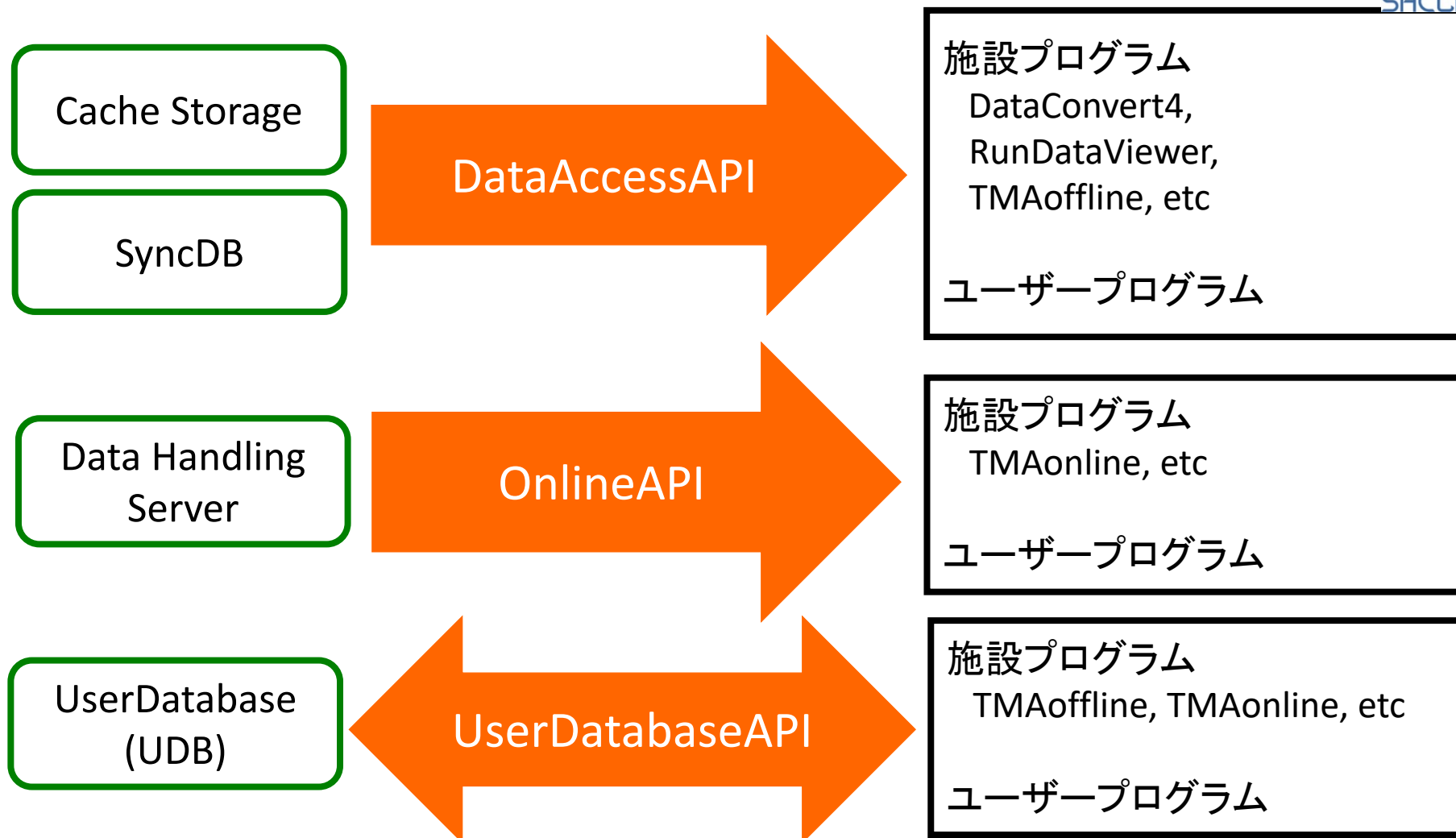


— Ethernet

各計算機システムの仕様

Cache storage system	[System for BL1 experiments] Hardware: Storage capacity: File system:	DDN SFA7700X ~100 TB GPFS (I/O ~2 GB/sec)
	[System for BL2 experiments] Hardware: Storage capacity: File system:	DDN SFA7700X ~380 TB GPFS (I/O ~7 GB/sec)
	[System for BL3 experiments] Hardware: Storage capacity: File system:	DDN SFA7700X ~510 TB GPFS (I/O ~7 GB/sec)
Archive storage system	Hardware: Storage capacity: File system of the disk part: Hierarchical management:	Disk: DDN SFA10K, Tape: IBM System Storage TS3500 & TS1140 (5 drives) Disk: 1 PB & Tape: 8 PB (can be increased up to 28 PB) GPFS (I/O > 4 GB/sec) Tivoli Storage Manager (I/O ~200 MB/sec)
HPC system	Hardware: Theoretical peak performance: Total memory capacity: Interconnect: Storage capacity: File system:	Compute nodes: DELL PowerEdge M630(52 nodes) Storage: DDN SFA12K ~40 TFLOPS ([intel E5-2660 v4 × 2] × 50 + [Intel E5-2699 v4 × 2] × 2) ~3.5 TB (64 GB × 38 + 512 GB × 2) Infiniband FDR ~2 PB Lustre (I/O ~10 GB/sec)

データ解析のためのAPI群 (C/C++, Python)



施設フォーマット, 各データが何処にあるか等をユーザーが意識する必要なし
ユーザーの自由度・利便性を向上

SACLAの計算機システム上のソフトウェア構成

データ解析
用カスタム
GUI

データ解
析CUI

市販データ
解析ツール
(Matlab etc.)

HDF5フォー
マットへの
コンバータ

Web Portal
経由の解析

短時間解析を実現

データ解析API
(Python, Fortran, C++ etc.)

多言語対応

データアクセス(user level)

C言語の実装
ユーザインターフェイスの標準化

データアクセス(low level)

C言語の実装
データの場所、構造、フォーマットの隠匿化

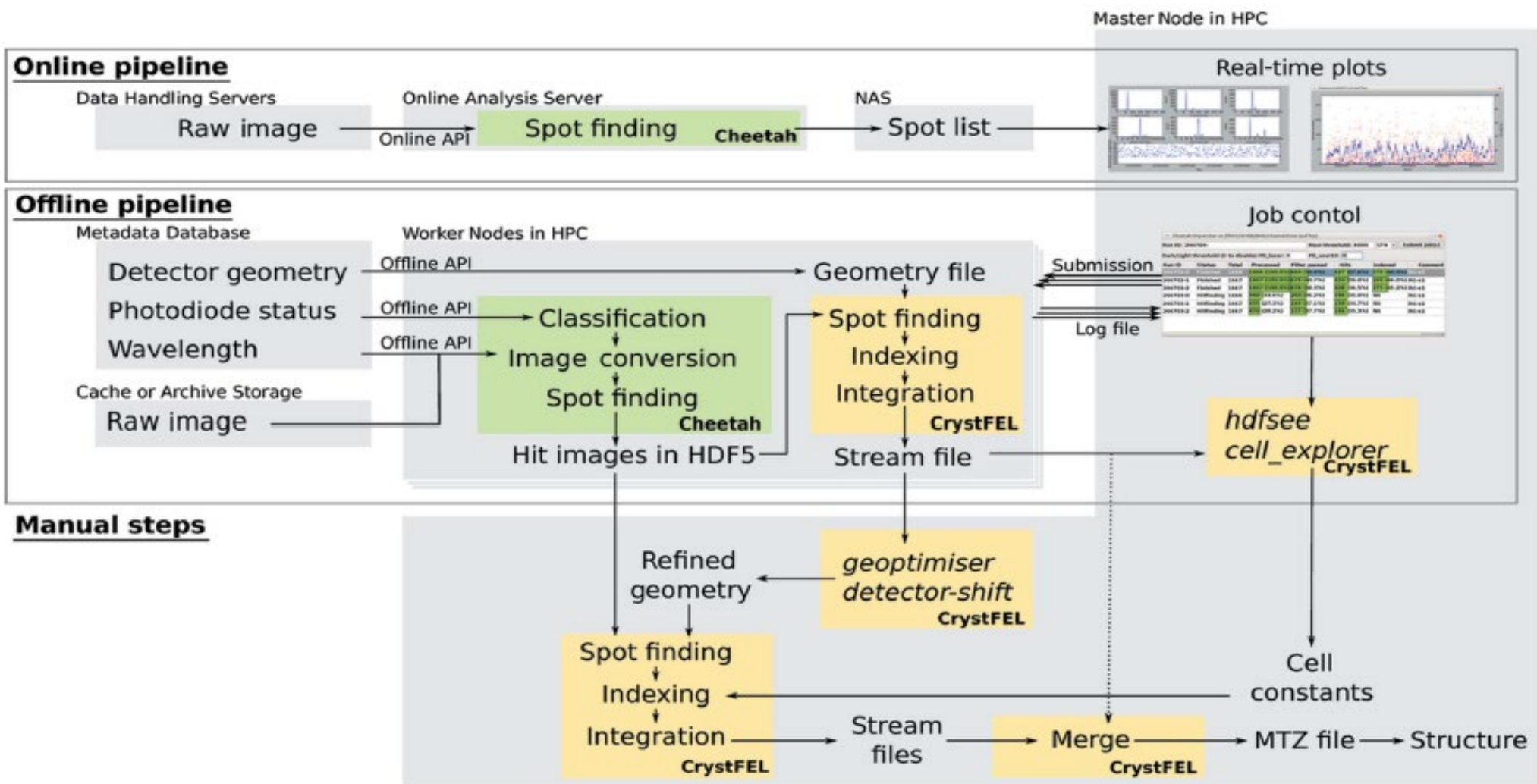
高速データアーカイブ
(画像データ格納)

高速データベース
(メタデータ、強度データ、一時分
析結果)

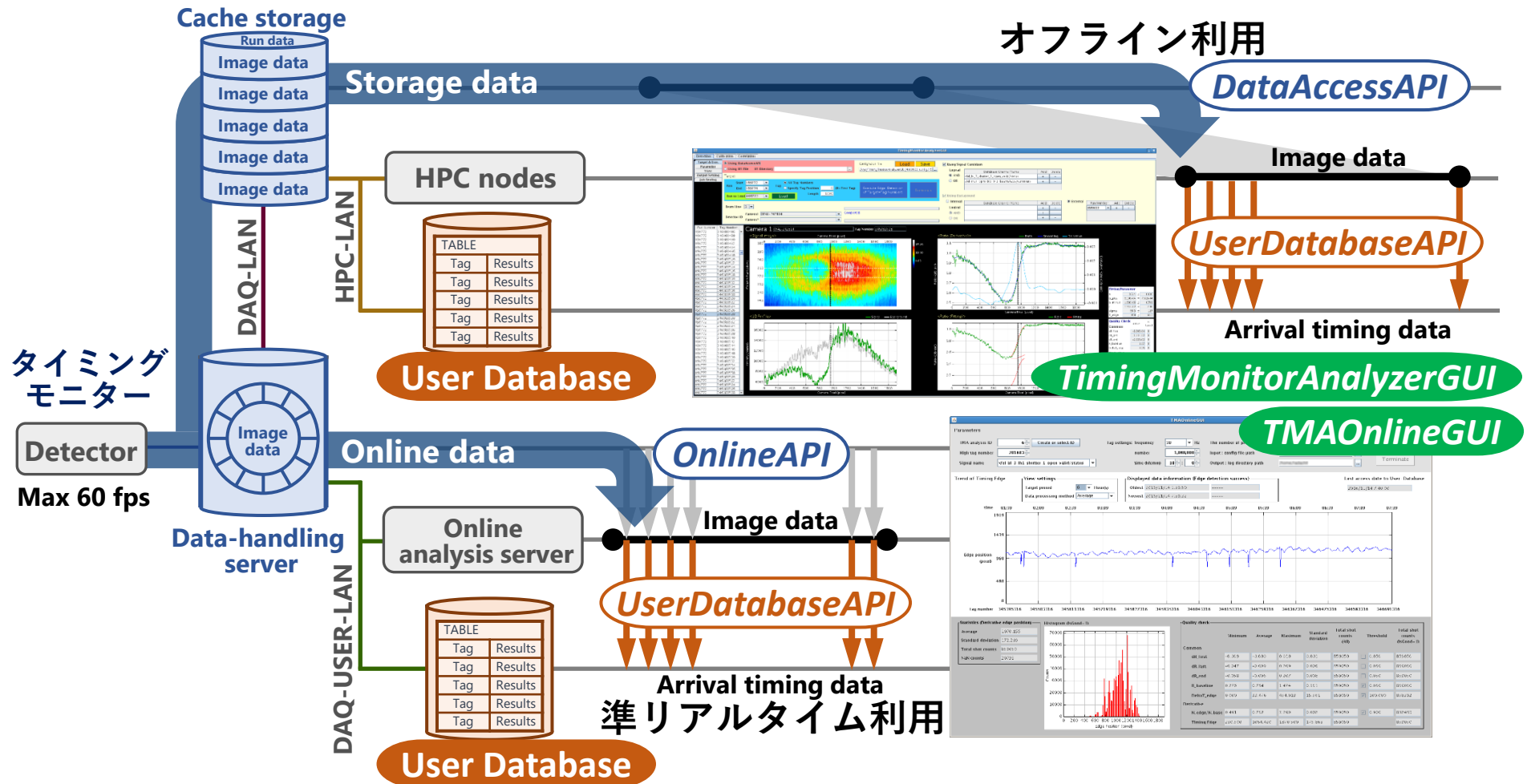
ハードウェア

解析例: SFXの迅速データ処理

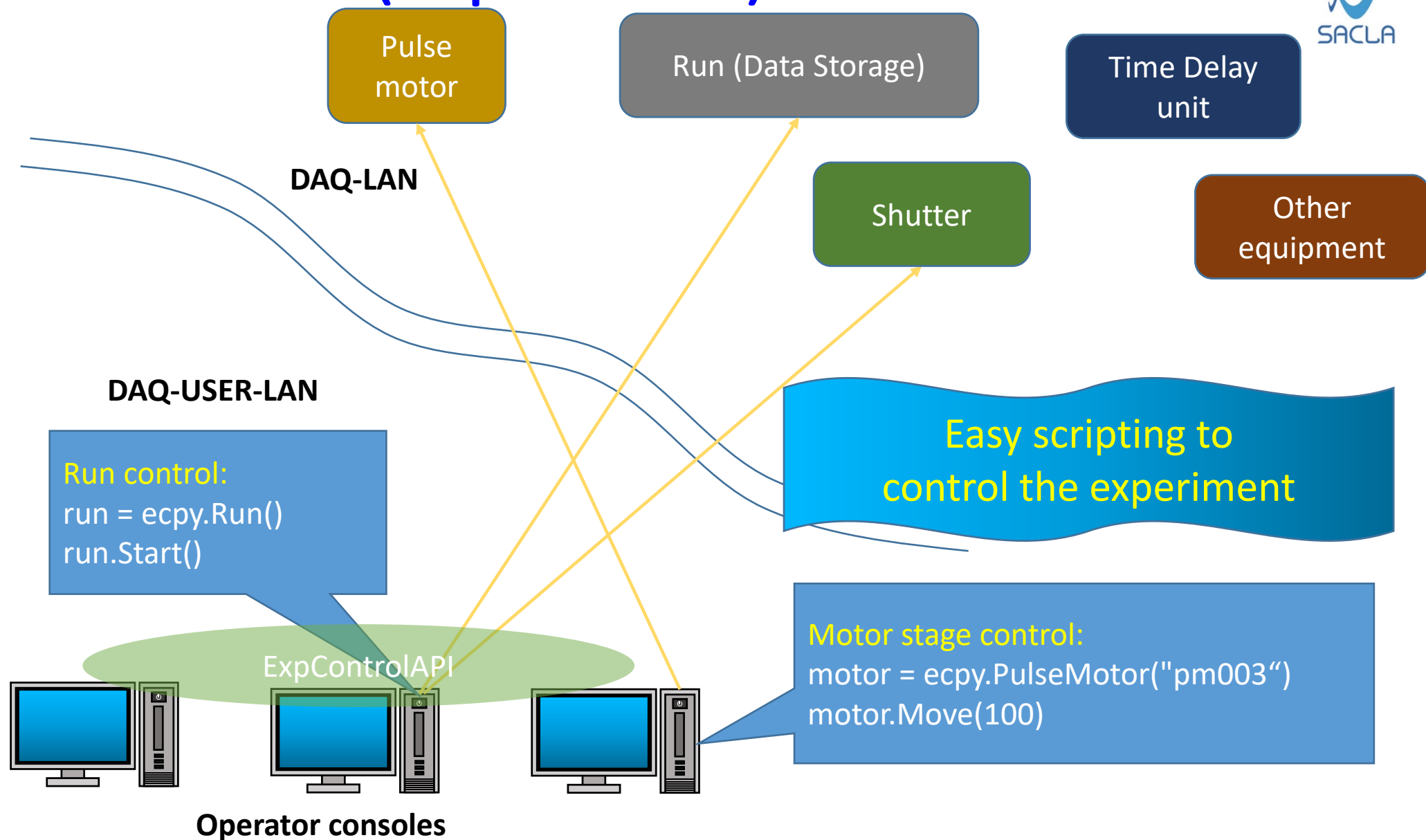
Data processing pipelines based on Cheetah(Barty et al.) and CrystFEL (White et al)



タイミングモニターアナライザー(TMA)



実験制御(ExpControl)API



個別実験の自由度を向上させ、効率的な実験制御プログラム作成を可能に.

→ 実験の自動制御を可能とする基盤が揃いつつある.

生データ生成量の最大値(データ収集システムの設計値)

6Gbps (MPCCD12センサー一分)

2TB/時

現状のデータ量(BL3本合計)

[生データ] ~ 4PB (2017- 1.6 PB/年)

[2次データ(UserData)] ~ 3PB

設立趣旨

SPRING-8/SACLAの大量データに関し、高度なデータ活用が期待されている。SPRING-8/SACLAの供用のための共用実験データ・ネットワークシステムの円滑な運用と利用実験の高度化に応じた継続的な拡充のため、データ・ネットワークシステムにかかる整備・運営の意志決定機関として設立

委員会メンバー

委員長、理研RSCセンター長、JASRI研究担当理事、SPRUC会長、SPRING-8利用推進協議会 会長、専用BLの各代表(FSBL連合体, 東大, 兵庫県, QST, JAEA, NSRRC, NIMS, サンビーム, 京大RISING, 阪大LEPS, 豊田, 阪大蛋白研) (計17名)

これまでに実験データ保持ポリシーを策定。今週末、データ・ネットワークシステム利用・運用における利用者-施設者間の基本的な方針を議論予定。

SACLAのデータ管理

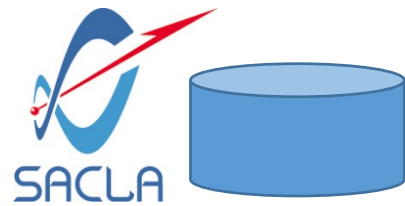
2020年2月に策定したデータ保持ポリシーに基づき2021年3月に運用変更予定

分類	所有者	Cache system	新Disk system	Tape system
同期収集データ	施設	240日	3年	+ 7年
/ExpData(生データ)	ユーザ	90日	3年	+ 5年
/UserData	ユーザ	—	3年	—

- ✓ 各システムで、保持期限、容量制限、利用期限を設定
- ✓ データ消去運用開始までに移行期間
- ✓ ネットワーク、LTO等データ持ち帰り手段の強化を検討中
- ✓ ホームページにて情報公開

http://xfel.riken.jp/users/pdf/20200716_SACLA_data_retention_e.pdf

R-CCSとの連携によるデータ転送環境の導入(HPCI)



外部転送ノード

新設予定
(2021年1月)



10G

gfpcopyで約900MB/sec

2022年のSINET6移行に合わせて
SACLA-RCCSの帯域増強を予定



レプリケーション

100G (SINET5)

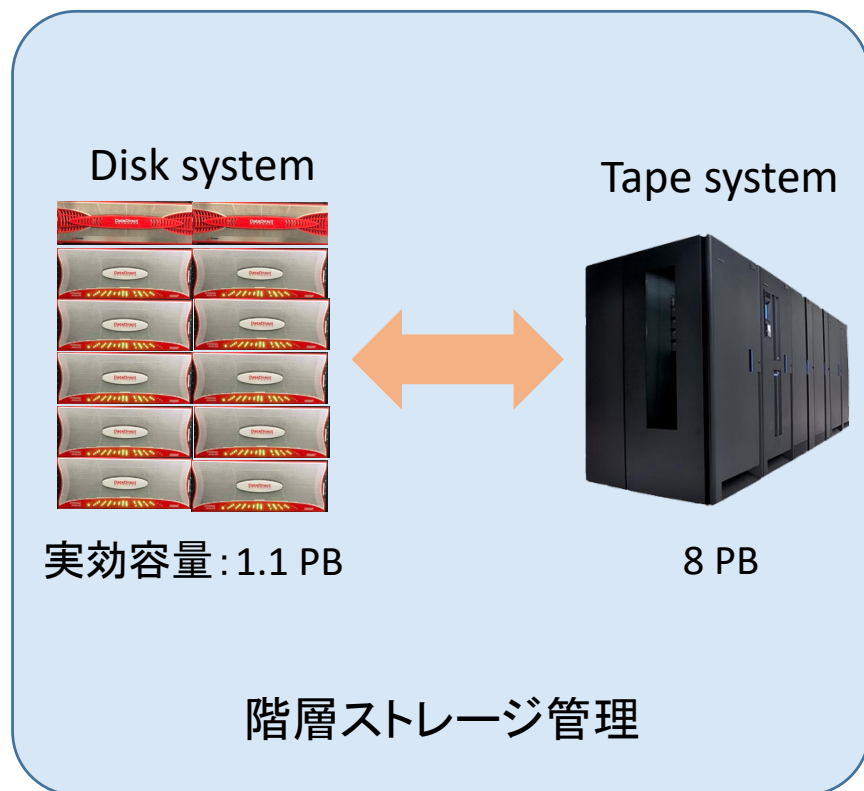


ユーザーが使いやすいデータ転送システムを
R-CCSと共同で開発中

- 大型実験施設とHPCIの連携利用課題に貢献
- 保持期限切れのデータの移動先の一つ

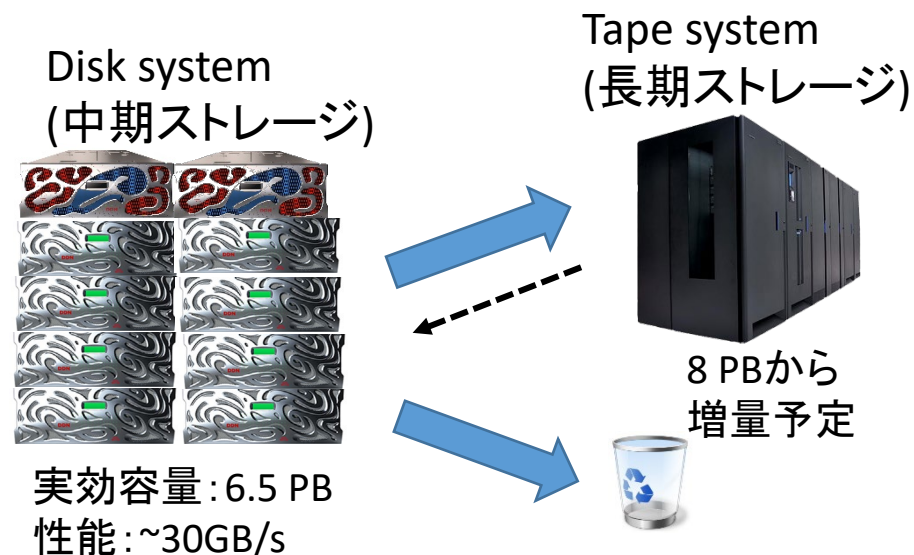
アーカイブストレージの更新

OLD(-2021.3)



- 一定期間が過ぎたデータはディスクからテープに自動的に移動
- テープにあるデータはアクセスすれば自動的にリコール

NEW(2021.3-)



中期ストレージと長期ストレージの分離

- 一定期間が過ぎた生データはディスクからテープに移動
- テープにあるデータは、ユーザーの申請に応じてリコール
- 保持期限が過ぎたユーザーデータはディスクから消去

大型計算機システムの今後の調達計画

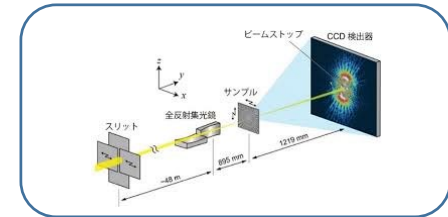


	FY2020	FY2021	FY2022	FY2023	FY2024	FY2025	FY2026
BL1Cache	買取		データ領域の有効活用・調達コスト低減のために統合予定 5年リース				
BL2Cache	7年リース						
BL3Cache	5年リース						
UserCache	買取						
			リプレイス				
HPC	5年リース	5年リース					5年リース
		リプレイス					リプレイス
FX10	買取	運用停止					
Archive Storage	買取	改造して運用継続					
		ディスク廃棄 テープを分離					
中期	5年リース					5年リース	
★データ・ネットワーク委員会設立			★CITIUS 20M運用開始		リプレイス		
★データ保持運用の変更							

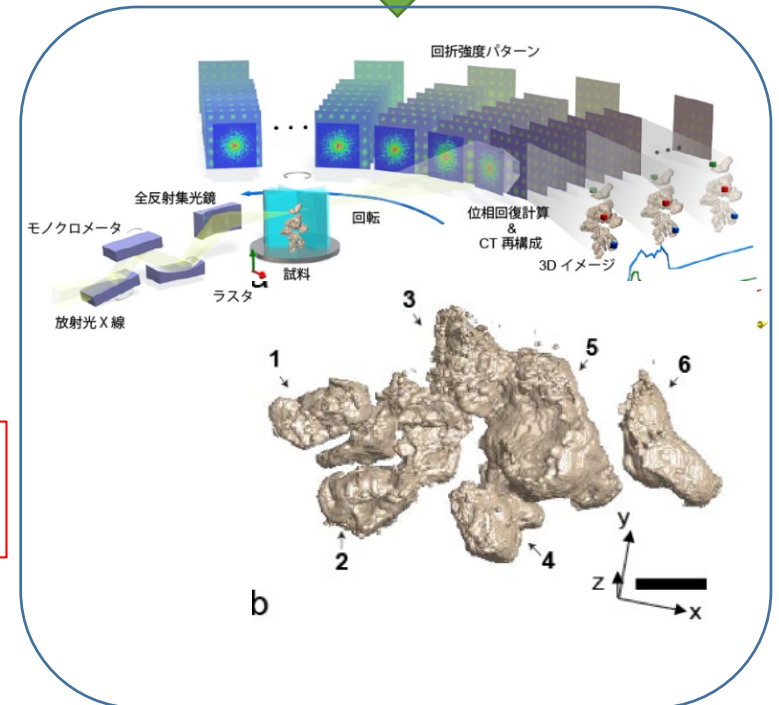
SPring-8のデータ基盤高度化の 展望と取組

SPring-8の高度化計画(SPring-8-II)

- 8GeV $\Rightarrow$ 6GeV
省電力化
- より明るく
明るさ100倍程度
- コヒーレント度強化

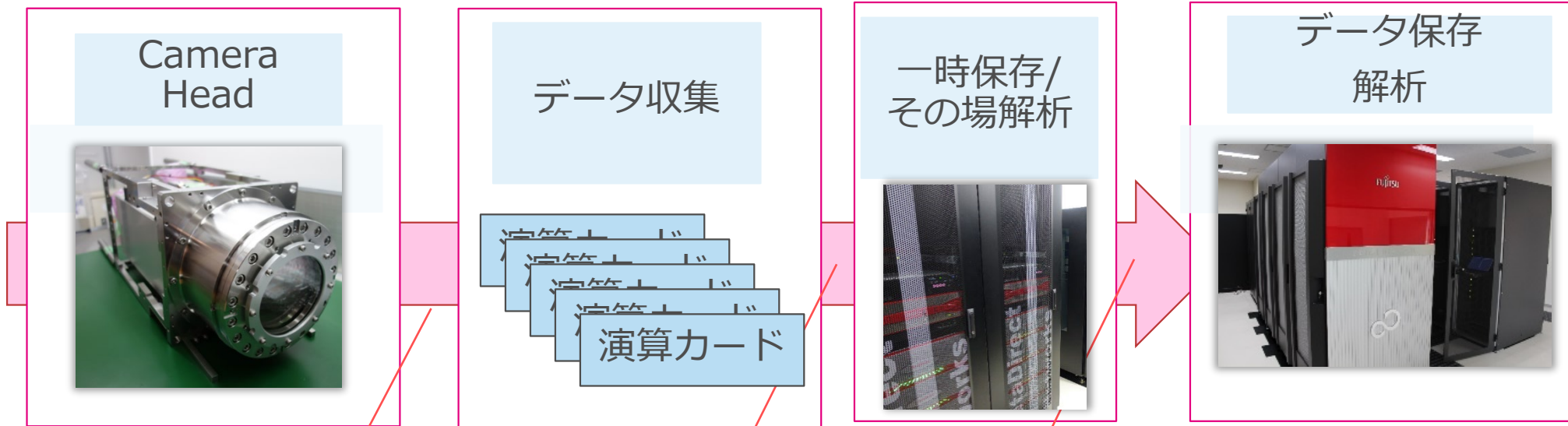


光源・光学系・検出器の性能向上
に伴うデータの大容量化

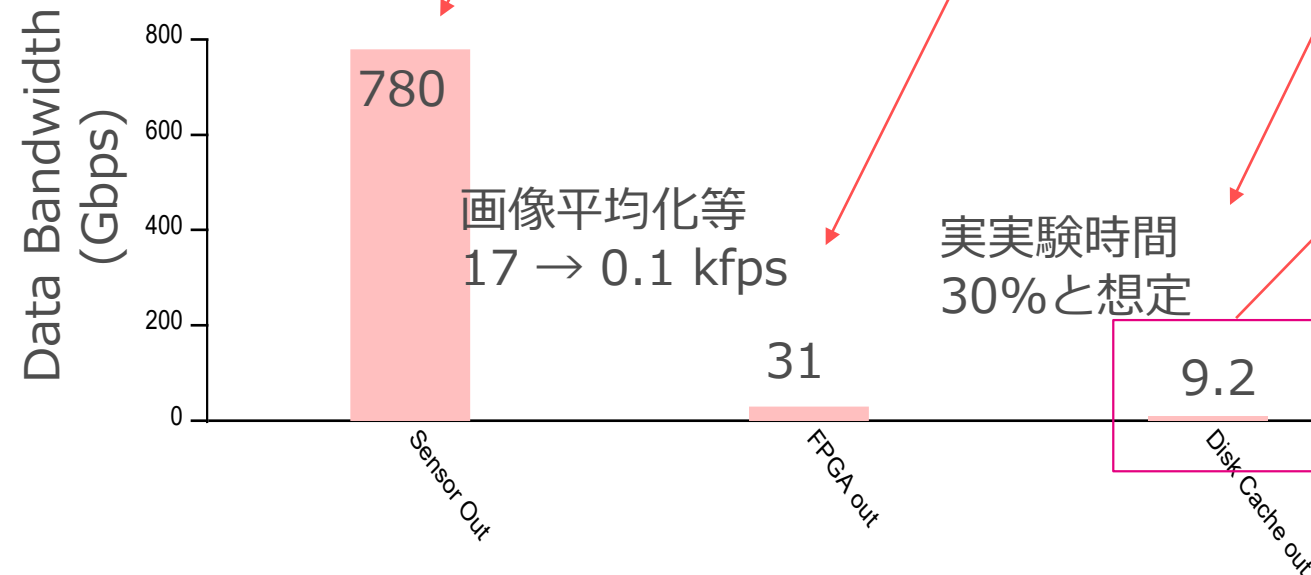


2014年9月 Conceptual Design ReportをWeb公開
<http://rsc.riken.jp/pdf/SPring-8-II.pdf>

CITIUS検出器: データ補正などのデータ処理系の開発 データ可視化・その場解析、保存技術の開発 (イメージ図)



9.6 Mpixel



16 PByte/年/system
(参考)

SPring-8共用ビームライン
(26本) 生データ合計
0.32 PB/年 @2017年

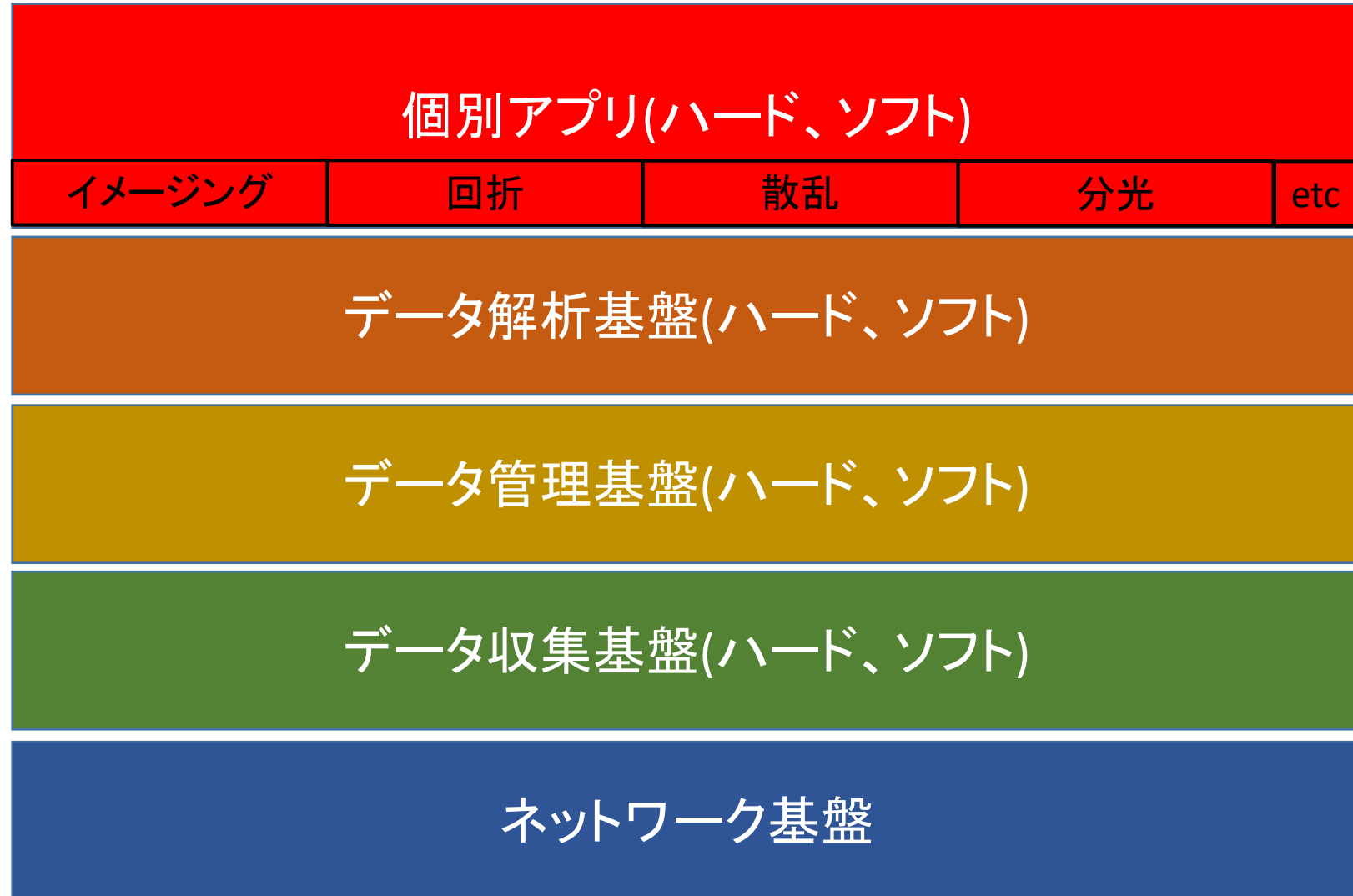
**SACLAの大規模データ
処理(1.6PB/年)の経験
を生かしたDAQ開発を検討**

SPring-8におけるデータ管理について(展望)

- 次世代光源により高コヒーレントのX線が提供され、検出器の高性能化および実験自動化が進むとデータ量は爆発的に増大。
- 個別のビームラインでデータを管理し、実験者が生データを持ち帰るやり方は時代遅れ。
- 施設として実験データを管理・利活用するためのインフラを整備していくことが必要。
- インフラ整備のためには、データの所有権、管理期限等のデータポリシーについて施設とユーザーで議論しながら整理していくことが重要。

← SPring-8データ・ネットワーク委員会

SPring-8高度化に向けたデータ基盤の構成案



SPring-8の共用実験ネットワーク高度化

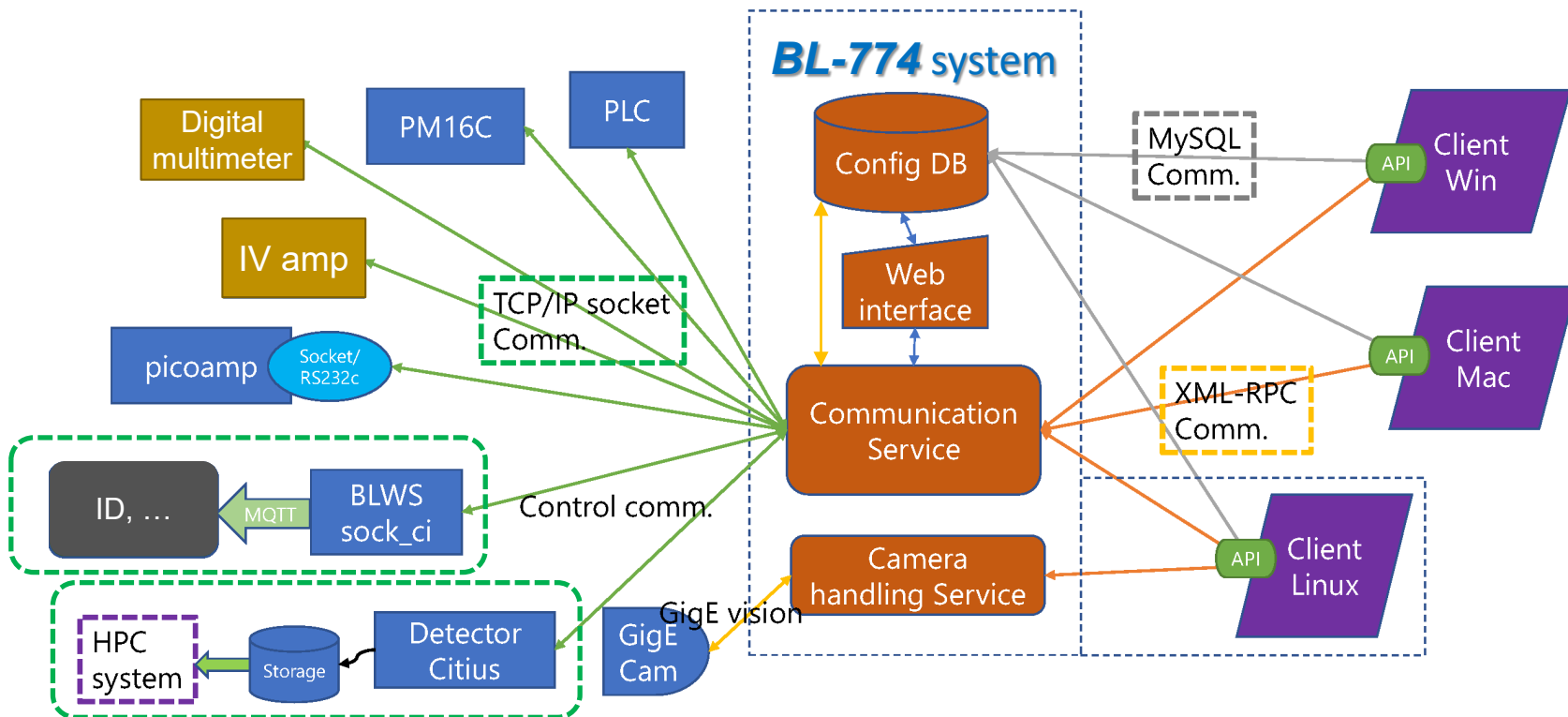
- SPring-8建設時から歴史的な理由で複雑化したネットワーク構成を根本的に改修
- 利用実験に関わるもの(**共用実験ネットワーク**)を大きく4つに分類
 - ユーザー支援ゾーン
 - ビームライン制御ゾーン
 - データ収集ゾーン
 - データ解析ゾーン
- 現在のOA-LANの実験利用を見直し
- 共用実験用のリモート接続環境の構築(VPN等)

FY2020	FY2021	FY2022	FY2023	FY2024	FY2025	FY2026
★共用ネットワーク関係規定の整備					★全系高度化完了	
		★SINET6への切り替え				
部分更新(+ R&D)			全系高度化		コンポーネント増強により 必要な区間から100G化	

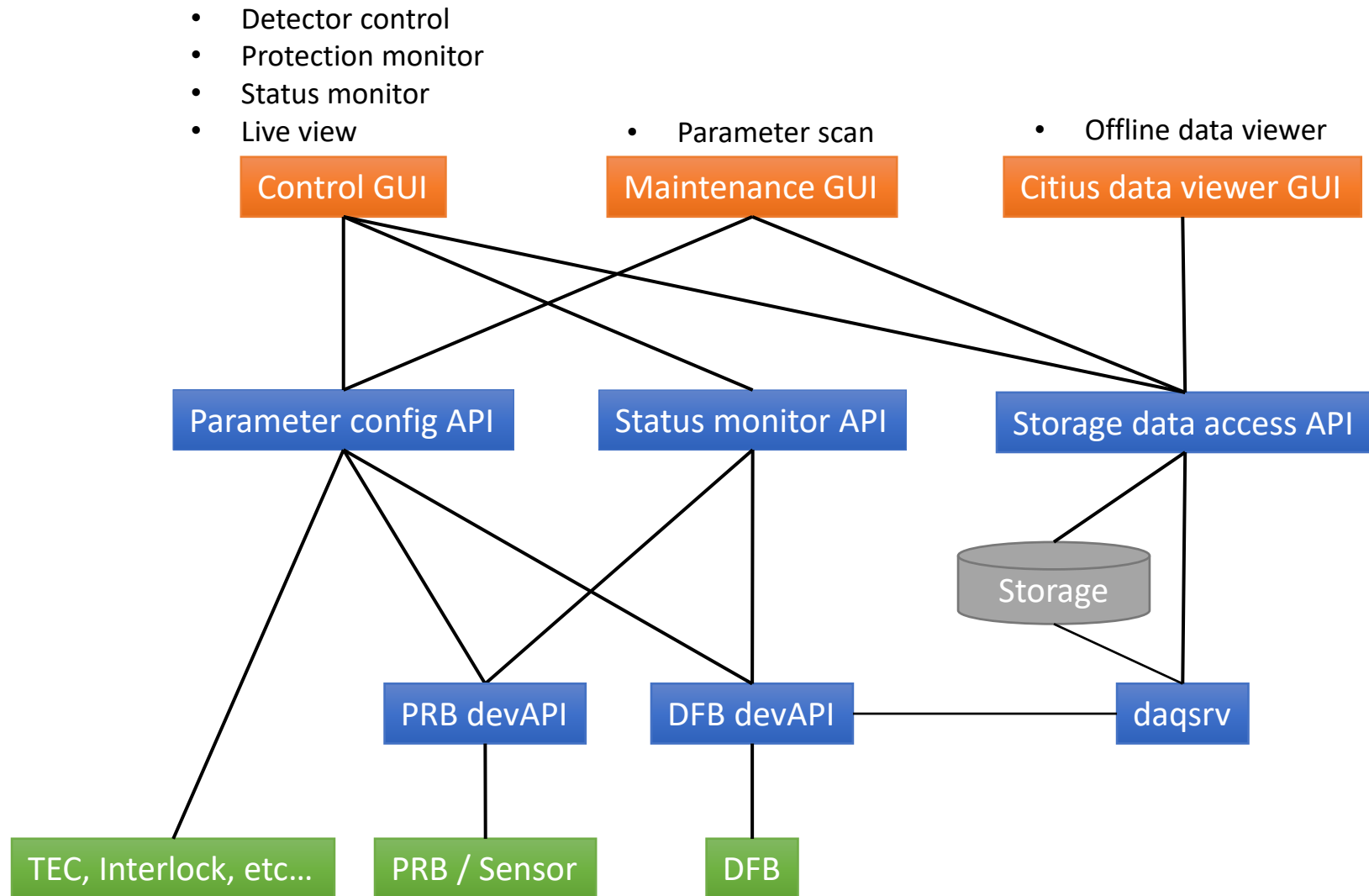
データ収集基盤: BL-774

実験コントロールプラットフォームの標準化

- Python API (LabVIEWサポートあり)
- XML-RPCサーバを介したTCP/IPソケット通信



データ収集基盤: CITIUS-APIの整備



2020年初期バージョンリリース

SPring-8のデータインフラの展望

- 光源・光学系・検出器の性能向上に伴い増大する実験データを活用するための**共用実験データシステム**の導入を検討中
- 2021年度にプロトタイプ計算機を導入し、R&Dにより本番仕様を検討予定
 - 効率良いデータ圧縮法の検討
 - 複数BLの大量データを活用するためのデータ管理法の検討
- 2023年度中の共用実験データ受け入れを予定
- 外部リソースの利用も検討中(HPCI等)

FY2019	FY2020	FY2021	FY2022	FY2023	FY2024	FY2025	FY2026
★共用データ関係規定の整備							
演算装置		Proto HPC					
		共用HPC					
		★データ受け入れ開始					
ストレージ		←		→			
		データ管理の仕組みを検討					
					中期(HDD)		
						長期(光disk)	

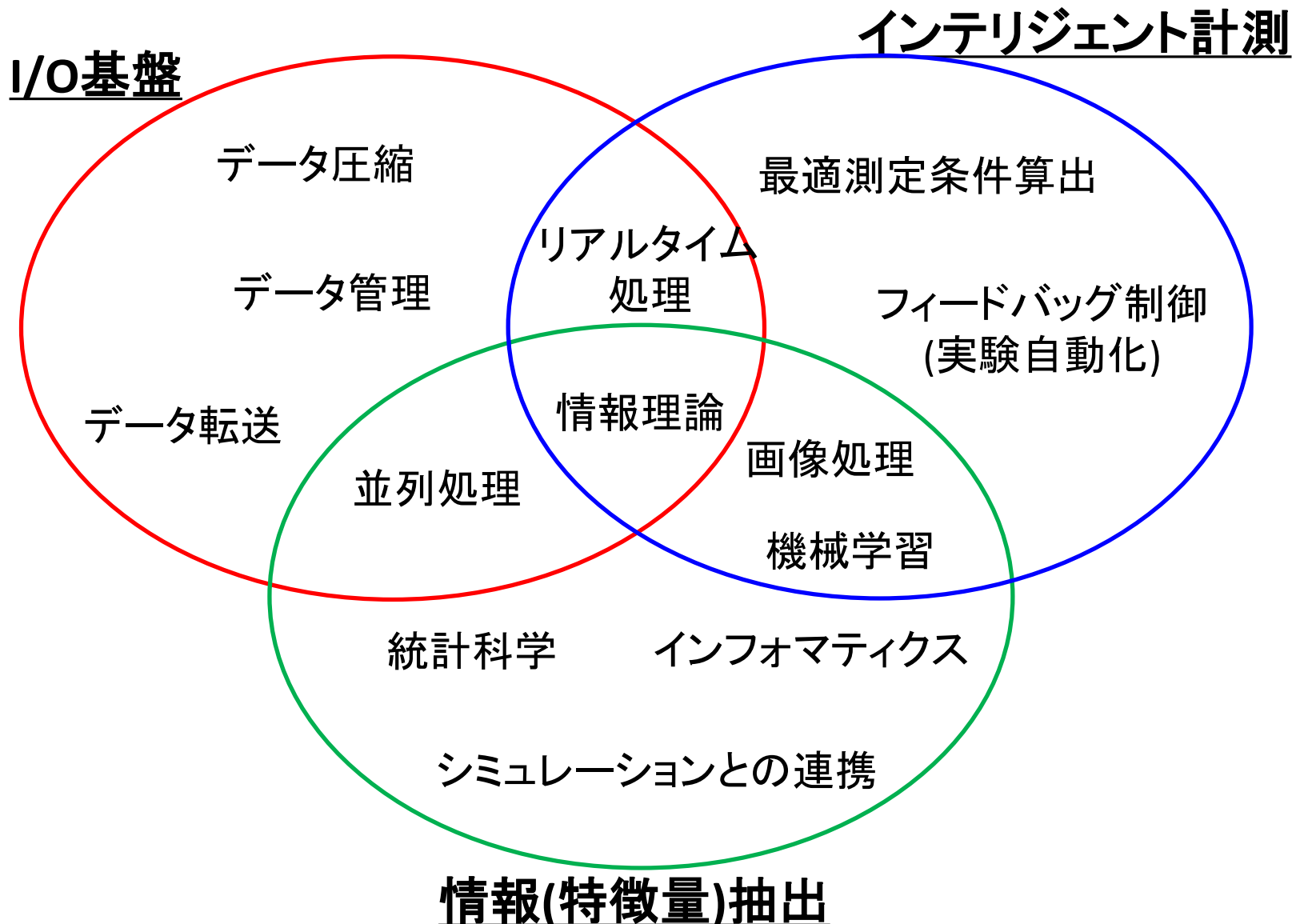
データ解析支援はどの程度必要か？

データインフラを整備しても、実験データが利活用されなければ意味がない。

- 実験ユーザーに期待する範囲
- データサイエンスの研究者との連携に期待する範囲
- 施設が面倒をみる範囲

を整理して議論していきたい。

大規模データの効率的ハンドリングに向けて



要素技術を整理し、インテグレーションを可能とする基盤を整備

まとめ

- 6GBpsで計測されるSACLAの大量データを確実に保存・効率よく解析するためのデータ基盤を開発
- SPring-8の高度化で爆発的な増大が見込まれるデータを効率的に取り扱うデータ基盤の標準化を検討中
 - <= SACLAの大規模データ取得・解析基盤の開発経験を生かす
- SPring-8/SACLAのデータ管理方法の見直し
 - ✓ データ・ネットワーク委員会の設立と方針決定

謝辞

[SACLAのデータ基盤開発]

亀島敬(JASRI), 本村幸治(理研), 中嶋享(理研), 山鹿光裕(JASRI), 阿部利徳(JASRI),
杉本崇(JASRI), 岡田謙介(JASRI), 古川行人(JASRI), 大端通(JASRI),
田中良太郎(JASRI), 初井宇記(理研), 矢橋牧名(理研),
RSCエンジニアリングチーム

[SPRING-8高度化に向けたデータ基盤の検討]

中嶋享(理研), 本村幸治(理研), 平木俊幸(理研), 杉本崇(JASRI),
山鹿光裕(JASRI), 阿部利徳(JASRI),
初井宇記(理研), 矢橋牧名(理研),
RSCエンジニアリングチーム