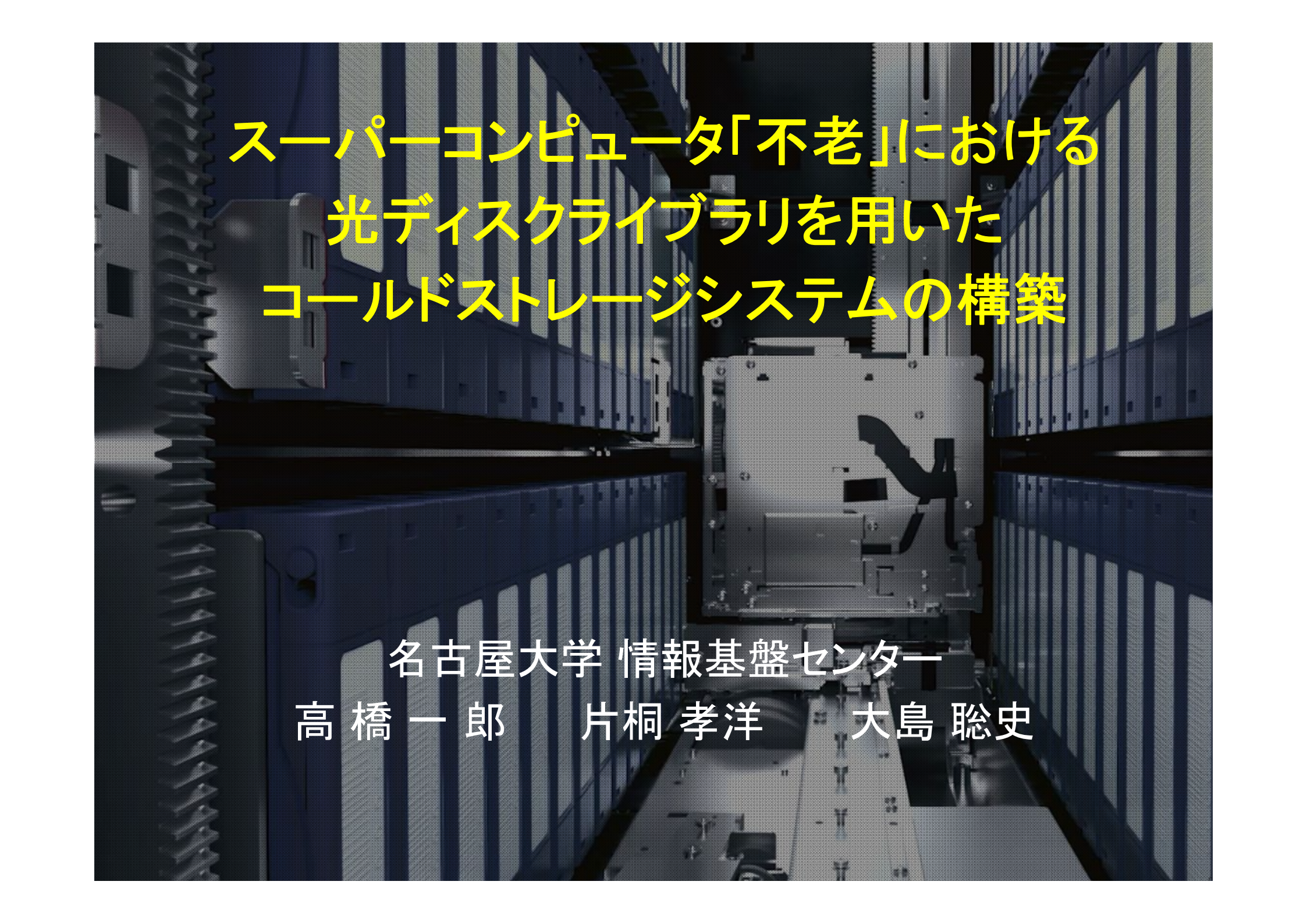




スーパーコンピュータ「不老」における 光ディスクライブラリを用いた コールドストレージシステムの構築

名古屋大学 情報基盤センター

高橋 一郎 片桐 孝洋 大島 聡史

アジェンダ

- ◆ 光ディスクストレージ システム Sony社製「ODA」とは
- ◆ 「ODA」を採用した理由
- ◆ コールドストレージシステム構成, 特徴等
- ◆ 運用(アプリケーション, 利用申請等)
- ◆ 利用方法
- ◆ 今後の課題



オプティカルディスク・アーカイブ (ODA)とは

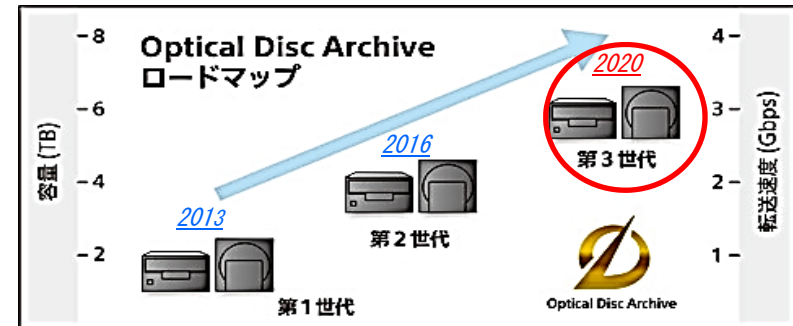
デジタルデータの長期保存(アーカイブ)を目的とした、大容量光ディスクストレージシステム

- ① Archival Disc (業務用次世代光ディスク)
記録メディアは、Write Onceです。



- ② ODAドライブユニット

- ③ PetaSite拡張型ライブラリー



		第1世代	第2世代	第3世代
転送速度	容量	1.5TB	3.3TB	~ 5.5TB
	読み出し	1.1Gbps	2Gbps	~ 3Gbps
転送速度	書き込み (w/Verify)	440Mbps	1Gbps	~ 1.5Gbps



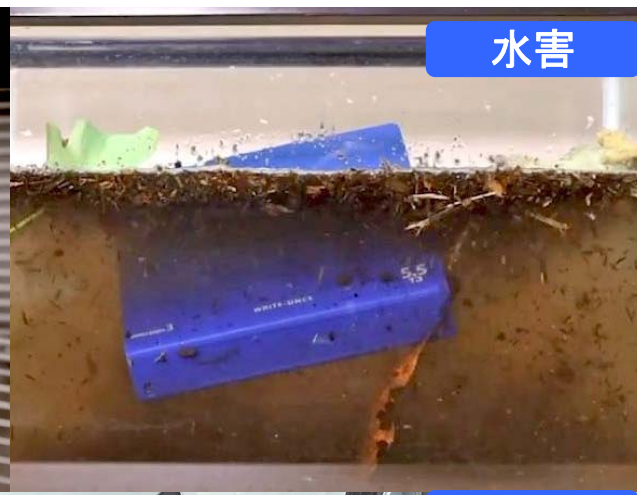
高温



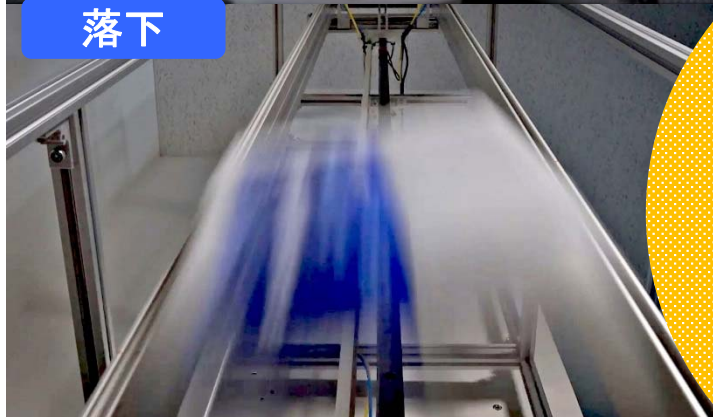
低温



水害



落下



加重



静電気

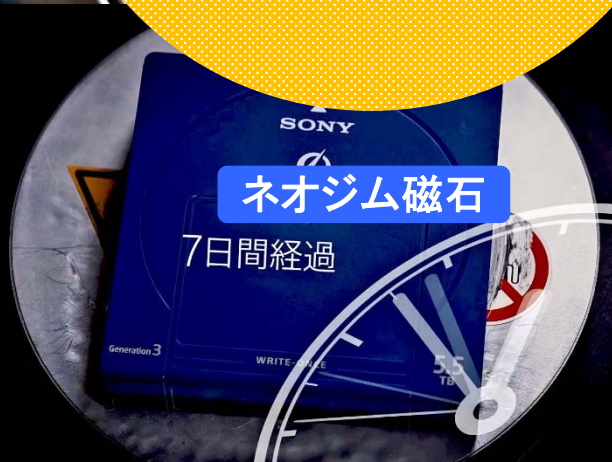


静電気

堅牢性が高い
安全性

ネオジム磁石

7日間経過



振動



Sony社提供

ODAライブラリーを採用した理由

- ① 堅牢性が高い
- ② 保存期間が長い(保存寿命100年以上:国際規格ISO/IEC 16963)
- ③ ライブラリーと単体ドライブが同じデータの記録方式
 - ・ 可搬型利用(ライブラリー⇔単体ドライブ)
 - ・ ボリューム単位で利用者に割当てて利用
 - ・ ライブラリー装置の更新時のファイル移行が不要
 - ・ 利用終了時にメディアを利用者に返却
- ④ Read/Writeアクセス性能

マイグレーション方式ではない

※ 単体ドライブユニットやライブラリー装置が必要(下位互換)

コールドストレージシステム構成

ODAライブラリ装置諸元(PetaSite拡張型ライブラリシステム)

	Phase1	Phase2
ODA制御サーバ数	1	4
ODAライブラリ数	1	4
機種名	PetaSite拡張型ライブラリシステム	
総物理容量	484TB	10.89PB (2.72PB×4ライブラリ)
総スロット数	88巻	1980巻 (495巻×4ライブラリ)
総ドライブ数	5	20 (5×4ライブラリ)
ロボット機構	1	4 (1×4ライブラリ)

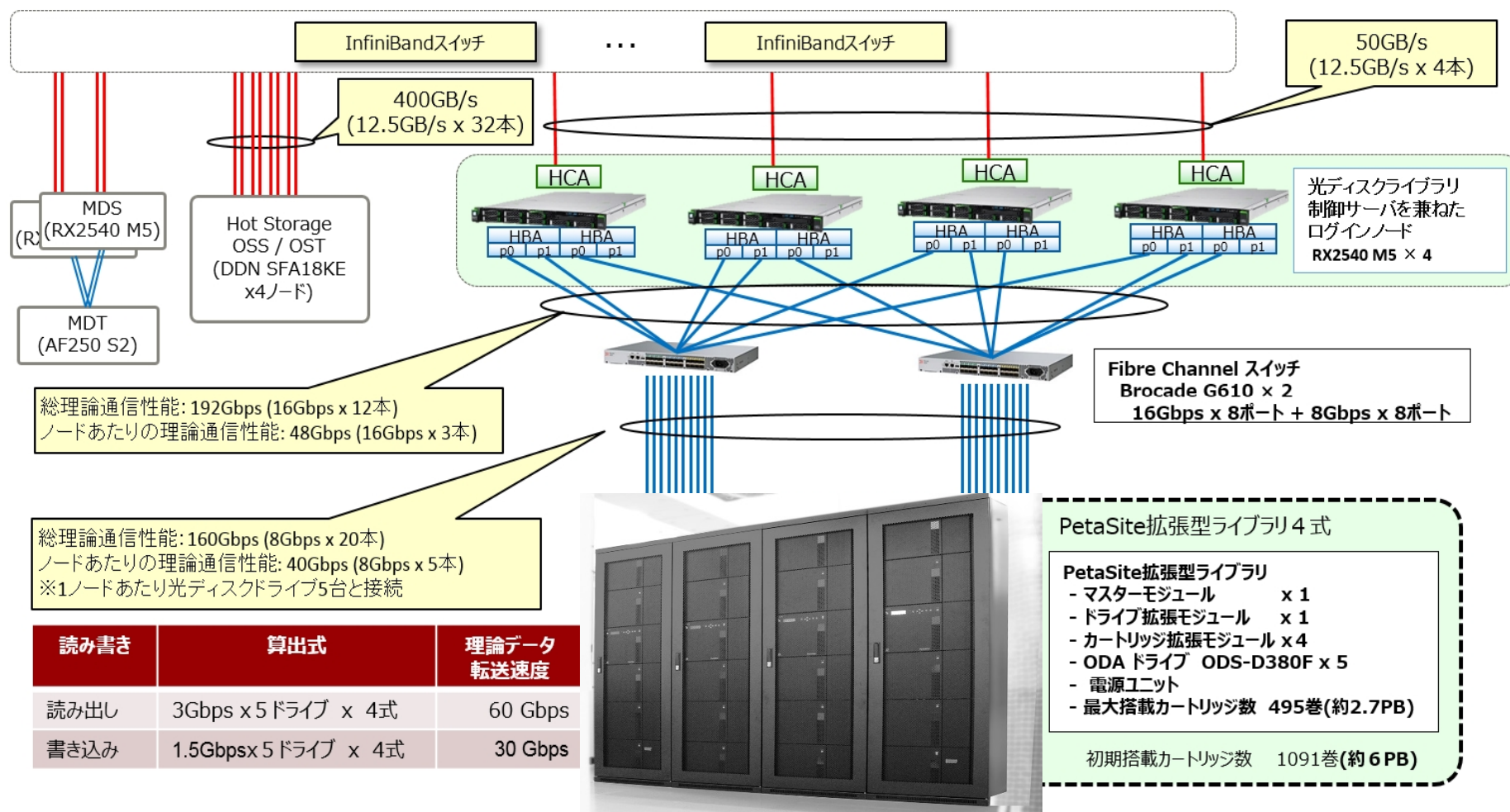
”4システム構成”
個別運用も可能



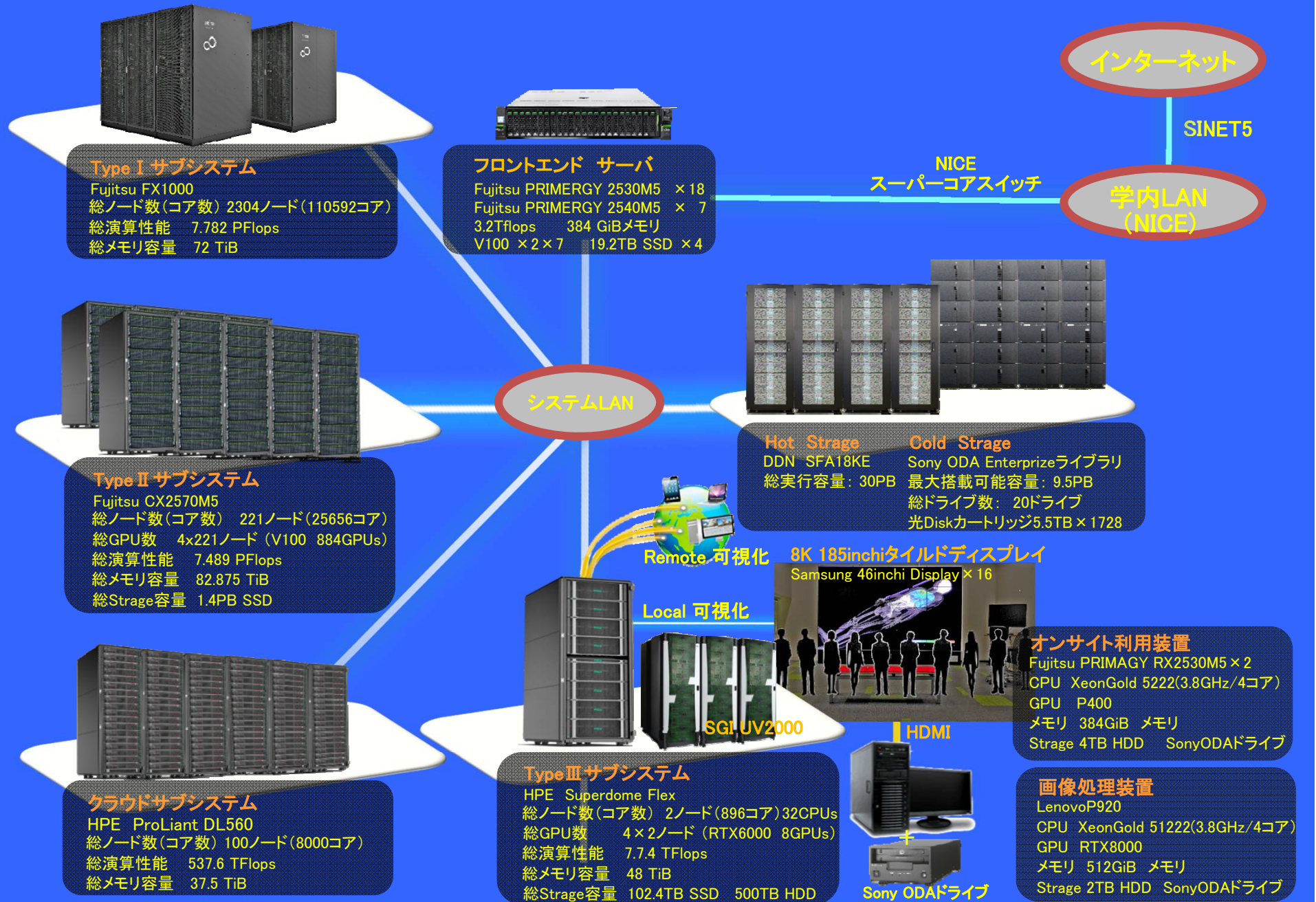
ODAライブラリ制御サーバ兼ログインノード諸元

項 目	仕 様
CPU	Intel Xeon Gold6248(2.5GHz/20コア)×2 理論演算性能 3.2TFLOPS
主記憶	384GiB (32GiB DDR4×12)
内蔵ストレージ	2TB SSD, 19.2TB SSD
インターコネク	InfiniBand EDR×1
SANインターフェース	FC16Gbps×4
ネットワークインターフェース	10GTwinax×2, 1000BASE-T×2

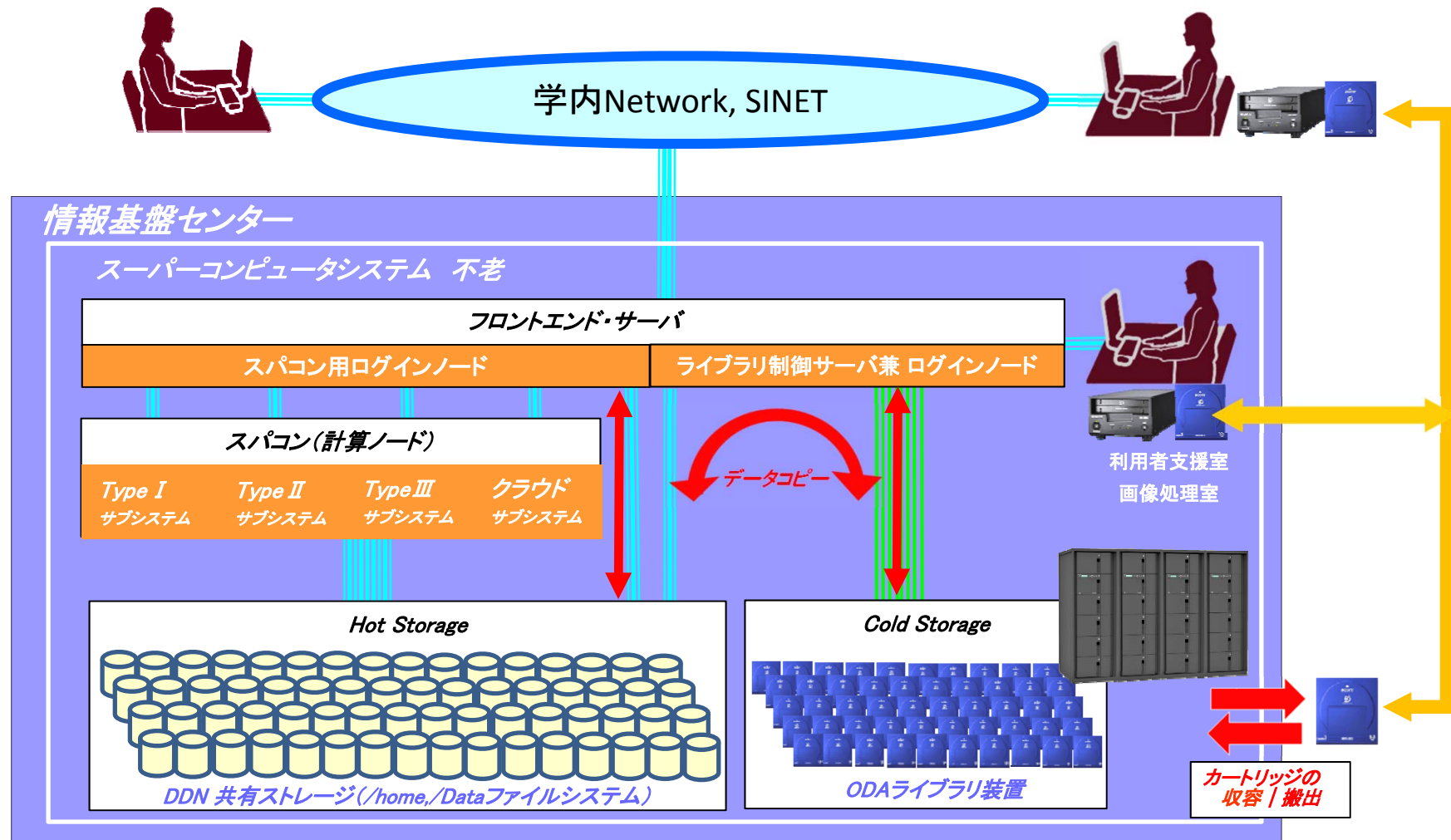
コールドストレージシステム構成(Phase2)



スーパーコンピュータ「不老」システム構成 (Phase2)



メディアとデータの流れ



コールドストレージシステムの特徴

- ◆ 1カートリッジを1ボリューム
- ◆ ライブラリ装置内のカートリッジはラベル(ボリューム通番)を付けて管理
ボリューム通番: ログイン名+3桁の通番
- ◆ ライブラリ装置と単体ドライブは ”同一記録形式”
カートリッジの持ち込み利用、搬出利用、可搬利用に対応
ライブラリ装置 ⇔ 単体ドライブ装置
- ◆ カートリッジに関する情報(所有者、記録情報等)を、DBを使って管理
※ DBを使った管理は、ライブラリ内のメディアに限る
- ◆ ベンダー提供のSDKを使って独自のストレージ管理ソフトウェアを開発
- ◆ 利用終了時に利用者にカートリッジを返却

運用アプリケーション(ベンダー提供SDK)

- ◆ カートリッジの管理情報を、DBを使って管理する機能
 - ボリューム情報, 利用者情報, 利用スペース情報, 記録ファイル情報等
- ◆ カートリッジに記録したファイル情報をDBを使って管理し、閲覧および検索する機能
- ◆ キャビネットとドライブ間のカートリッジを搬送する機能
 - ドライブへの取り付け、取り外し、搬送処理
- ◆ ドライブ状態を取得する機能
- ◆ カートリッジ単位で使用者を限定する機能。
- ◆ カートリッジをライブラリ装置外に搬出およびライブラリ装置内に収納する機能。

運用アプリケーション(センター開発)

- ◆ マウント操作を行う機能
所有者認証 → ドライブ状態確認 → 搬送 → マウントポイント作成 → マウント
- ◆ アンマウント操作を行う機能
所有者認証 → ドライブ状態確認 → アンマウント → ファイル情報をDBに反映
→ ドライブ開放 → キャビネットに返却 → マウントポイント消去
- ◆ ジョブ・スケジューラPBS proを使って、カートリッジにアクセスする
ジョブを投入・制御する機能
- ◆ ロボットアームを排他制御する機能
- ◆ ファイル記録情報を取得し検索する機能(Hotストレージ対応)
- ◆ カートリッジの管理(新規登録, 変更, 取消し, ライブラリ内に収納,
ライブラリ外に搬出)を管理する機能(開発中)

利用申請区分と経費

利用申請書 (区分)	ファイル利用負担経費 (年間)	内 容	備 考
新規	ファイル負担経費+管理費: 1口 (19万円 + 1万円)	カートリッジはセンター提供 1口:10カートリッジ(50TB相当)	ファイル負担経費は初年度 のみ 利用終了にカートリッジを利用 者に返却
継続	管理費 : 1口 (1万円)	1口:10カートリッジ	次年度以降
変更申請	な し	利用メンバーの変更 カートリッジの持出し利用 カートリッジの持込み利用 利用停止	
カートリッジ 持込み	管理費: 1口 (1万円)	カートリッジは利用者持込み 1口:10カートリッジ	

利用方法

- ◆ バッチジョブ
- ◆ インターラクティブ・バッチジョブ(コマンドライン)
- ◆ 会話型(コマンドライン)

【バッチジョブの利用例】

[oda01~]\$ @qsub ex1.bat

【会話型バッチジョブの利用例】

```
[oda01~]$ @isub
*****
* interactive job submiied.                *
*****
qsub: waiting for job 60.oda01 to start
qsub: job 60.oda01 ready
```

```
[oda01~]$ odaalloc x49999a001
[oda01~]$ cd /oda/x49999a/x49999a001
[oda01~]$ mkdir save
[oda01~]$ cp -r /home/x49999a/data01 save/.
[oda01~]$ ls -al save
[oda01~]$ odafree x49999a001
[oda01~]$ exit
```

ODAライブラリ操作コマンドの利用方法

コマンド名	機能
<code>odaalloc</code>	ライブラリ内のカートリッジのマウント処理を行う
<code>odafree</code>	ライブラリ内のカートリッジのアンマウント処理とDBのファイル情報を更新を行う
<code>@isub</code>	ライブラリ内のカートリッジにアクセスするインターラクティブ・ジョブを投入する
<code>@qsub</code>	ライブラリ内のカートリッジにアクセスするバッチジョブを投入する
<code>qstat</code>	ジョブの状態を表示する(ジョブスケジューラ PBSproのコマンド)
<code>qdel</code>	ジョブをキャンセルする(ジョブスケジューラ PBSproのコマンド)
<code>@vlist</code>	利用可能なライブラリ内のカートリッジの一覧情報を表示する
<code>@flist</code>	指定したカートリッジのDBに記録しているファイル情報を表示する ※ライブラリー内のカートリッジに限る
<code>@find</code>	カートリッジに記録しているファイル情報を、DBを参照してキーワードやワイルドカードで検索する
<code>@update</code>	指定したカートリッジの最新のファイル情報で、DBのファイル情報を更新を行う
<code>@tree</code>	@lsコマンドで格納したファイル情報をもとに、ディレクトリ構造のツリー表示を行う
<code>@ls</code>	指定したディレクトリ配下のディレクトリとファイル情報を、利用者が指定するファイルに格納する DBとは別に、利用者が独自にファイル情報を管理することができます。 Hotストレージでも利用できます
<code>grep</code>	Linuxのgrepコマンド。@lsコマンドで格納したファイル情報を、キーワードやワイルドカードを使って検索する

今後の予定

- ◆ Phase2のフル稼働に向けてシステムの整備を進める予定
- ◆ ODAライブラリの性能を引き出す利用方法を検討
- ◆ ODAライブラリの障害を回避する方法の調査(書き損等)
- ◆ 利用促進を図る予定

謝 辞

- ◆ ソニービジネスソリューション株式会社
- ◆ 富士通株式会社

他関係各位、この場を借りてお礼を申し上げます。

