

Centre de Calcul  
de l'Institut National de Physique Nucléaire  
et de Physique des Particules

# HPSS et robotique au CC-IN2P3

05/10/2020

# Nouvelle robotique

**Objectif : A terme, remplacer les robots Oracle/STK**

**Appel d'offre publié en septembre 2019**

- **Petite librairie pour valider les choix technologiques**
- **Capacité 20 Po**
- **12 Lecteurs Entreprise**
- **Marché à bon de commandes**
  - Extension de capacité
  - Lecteurs supplémentaire
- **Optionnel : Solution de supervision de la qualité des média**
  
- **5 réponses / 2 solution techniques**
  - TS4500 / Spectra Tfinity avec TS1160
- **ATOS titulaire du marché**
  - Offre basée sur la solution Spectra





## Installation de la nouvelle robotique

- 18-22 février 2020 par les équipes spectralogic
- Configuration
  - 5 frames (1 drive / 4 stockage)
    - ~ 3200 bandes
  - 2 teraporter (robot)
- 12 lecteurs IBM TS-1160
  - 400 Mo/s R/W
- 990 bandes 3592-60F JE
  - 20 To/cartouche
  - Terapack de 9 bandes



- Robotique en test dans l'infra HPSS Test (Février / Mars 2020)
- Configuration / Exploitation très différence des librairies Oracle
- Problème de performances en écriture :
  - Problème de performances sur en écriture
    - $\sim 200 \text{ Mo/s} \ll 400 \text{ Mo/s}$
  - HPSS ne supporte pas (encore) les « buffered tape marks »
  - Nécessite d'adapter les paramètres des classes de stockage.
    - Création d'agrégats plus gros
    - Performances satisfaisante (370 Mo/s)
- « Lenteur » des montages :
  - Problèmes dus à des paramètres de collecte de stats sur l'OS spectra
  - Désactivation de la fonctionnalité
  - Utilisation PVR amélioré (HPSS 7.5.3)
- Mise en production le 27/03 dans HPSS



- **Evolution de la librairie (aout/septembre 2020)**
  - Ajouts de 12 drives additionnels
  - Total : 24 drives TS1160
  - Ajout de 990 bandes supplémentaires (20 Po)
- **Configuration au 1<sup>er</sup> octobre 2020**
  - Total 1940 bandes et 23 lecteurs dans la partition HPSSPROD
  - Capacité : 39 Po
  - 40 bandes / 1 lecteur dans la partition HPSSTEST



# **Spectra Tfinity au quotidien**

## Retour d'expérience sur 6 mois d'utilisation



## Configuration uniquement via l'UI (web ou front panel)

- Simple, facile à utiliser et configurer (voir simpliste)
- Contrôle des Entrée/sorties de cartouches uniquement possible sur le Front Panel

## Messages et logs

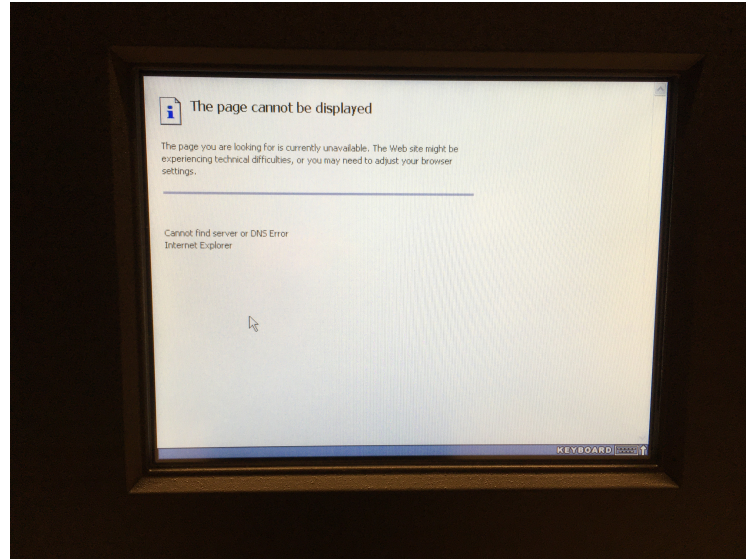
- Très limités
- Pas de logs texte à la mode ACSLS (`acsss_event.log` , `acsss_stats.log`)
  - Pas d'historique des montages !
- **Pas d'outil de ligne de commande en standard**
  - BlueScale repose sur Windows Embeeded (!)
  - SLAPI développé par LLNL  
<https://github.com/LLNL/slapi>

### SLAPI

SLAPI is a command-line tool that communicates with Spectra Logic tape libraries using their XML command reference. This provides a simple way to administer and monitor Spectra Logic tape libraries in large data centers.

Plantage régulier du panneau de commande de la librairie

→ Reboot automatique 1x/semaine (!)



- **Lenteurs de l'interface**

- 2 minutes pour que l'UI réponde
- Réponse du support :
- « Disable Drive performance monitoring and Power consumption monitoring »
- Améliorations mais toujours des lenteurs ponctuelles

## Routine Maintenance Configuration

This feature enables you to set up an ongoing maintenance routine for the LCM operating system. At the selected time, the LCM will automatically reboot, freeing up any unreleased memory and verifying the operating system integrity. This will not interfere with any backup operations; however, any active RLC sessions will need to be reestablished.

Enable:

Day of Week: Monday ▾

Hour: 5:00 AM ▾

» Save

« Previous



# Nouvelle robotique Spectra TFinity



- Incident du 11 septembre
  - Le démontage d'un lecteur échoue.
  - La bandothèque a été bloquée pendant 5 heures environ
    - Presque aucune requête de montage / démontage n'a été traité pendant ce temps
    - Le teroPorteur (bras robot) était positionné devant le lecteur en attendant le démontage.
- Bilan
  - Bande JE0552 cassée dans le lecteur
  - Récupération impossible
  - 20 To / 73 555 fichiers
  - Dizaine d'utilisateurs impactés (hess/antares/compass/dchooz/fazia ...)
  - Log en cours d'analyse chez Spectralogic



- **Robotique**
  - Environ 100 montages/ heure observé en pointe
    - À comparer aux ~ 450 montages heure sur les SL8500
    - Peu de relectures pour le moment
    - A voir en production dans le futur
  - Cycle d'accès au terapack plus long que sur une robotique classique
    - Extraction terapack de la chambre
    - Scan du terapack
    - Extraction de la cartouche
    - Chargement dans le lecteur
    - Reprise du terapack pour montage suivants.
- **Optimiser les relectures !**
  - Diminuer les montages

## Solution Oracle/Storagetek

- En production : 4 librairies / 56 lecteurs T10K / 24 LTO-7
- Ouverture des incidents automatisé via SDP2
  - Fonctionne bien pour les T10K et les composants robotique (bras, PTP, etc ...)
  - Toujours quelques problèmes sur les LTO
- **Traitement automatisé plutôt satisfaisant.**
  - Le technicien est autonome
  - Intervention souvent transparente.
- **Monitoring**
  - Storagetek Analytics (STA)
    - En production depuis 2014 : 4,5 M de montage en base
    - Génère des alertes par email (bandes / lecteurs / éléments robotique)






















## Media Lifecycle Management



- Donne une tendance sur l'état des bandes

### Media Lifecycle Management Report

Partition: HPSSPROD >> Go

Report: Media Health >> Go

Barcode	Media Health	Load Count
⚠ JE1019JE   	Average	6
⚠ JE0698JE   	Average	39
⚠ JE0790JE   	Average	10
⚠ JE1091JE   	Average	7
⚠ JE0344JE   	Average	110
⚠ JE0699JE   	Average	47
⚠ JE0142JE   	Average	155

 = Spectra Certified Media  = CarbideClean® Media


>> Save Report      Barcode Filter:  >> Filter

## Drive Lifecycle Management





















- Historique des 50 derniers montages

### Drive Lifecycle Management

>> Reset Health



Details for: FR3/DBA1/FTS11x0-DRV1  
Drive Health: 

Summary of tapes loaded in this drive

Barcode	Unmount Time	Tape Health at Unmount	Current Tape Health
JE0256JE	11/03/2020 13:53:41		
JE0517JE	11/03/2020 13:23:44		
JE0381JE	11/03/2020 12:39:21		
JE0515JE	11/03/2020 12:09:31		
JE0387JE	11/03/2020 11:14:15		
JE0352JE	11/03/2020 10:48:33		
JE1036JE	11/03/2020 09:59:44		
JE0520JE	11/03/2020 09:25:16		
JE0382JE	11/03/2020 06:28:42		
JE0912JE	11/03/2020 05:34:13		

History of tapes loaded in this drive (most recent first)

Mount 1

Time: 11/03/2020 13:53:41  
Barcode: JE0256JE  
Mfg SN: EYHV0TN450  
Tape Health:   
Drive Health:   
Tape Alerts: 00 00 00 00 00 00 00 00  
Display Character (SCD): U (No Data)



## Gestion automatisé des incidents :

- Moins avancée que chez Oracle
- Mécanisme d'ouverture des incidents non fonctionnel (pour le moment)
  
- Ex: Les drives n'ont jamais générés d'erreurs alors que 3 drives ont été remplacés

## Alerte par email synthétiques :

A Warning Message was generated on 10/22/2020 at 11:51:26.

Message Number - 33

Message Text The tape with barcode JE0529JE in drive FR3/DBA6/fTS11x0-DRV2 just encountered the following hard errors: 1 Hard Read Errors

Remedy Text Run the Drive Health Verification Test against drive FR3/DBA6/fTS11x0-DRV2. If the drive passes, investigate the media health of the tape with barcode JE0529JE.

- **Contrat de maintenance H+4 sur le robot et J+1 sur les lecteurs**
  - Maintenance sous-traité à « Stortrec »
  - 1 technicien sur Lyon est formé à ces interventions.
- **Ouverture des incidents directement chez Spectra Logic**
  - Au début les tickets étaient dupliqués et traités par stortrec
  - Redondant et compliqué (échange par mail)
  - Incident directement traités par Spectra
- **Support Spectra**
  - Ticket pris en charge rapidement
  - Personnel qualifié qui connaît le produit
    - Pas de workflow compliqué ni de call center en Inde
  - Souvent les mêmes interlocuteurs
- **Support Stortrec**
  - Pour le moment, Intervention uniquement pour les changement de lecteurs.
  - On réfléchi à faire la maintenance des lecteurs nous même
    - Remplacer un lecteur est un jeu d'enfant
    - Gain de temps dans la gestions des incidents

# Nouvelle robotique Spectra TFinity



- **Augmentation de la capacité déjà anticipée cet été 2020**
  - + 12 drives
  - + 990 Bandes (20 Po)
- **Evolution a venir (début 2021)**
  - Ajout d'une frame « drive »
  - Ajout de 3 frames médias
  - Configuration finale 9 frame
    - 36 drives
    - 6900 slots ( ~ 140 Po)
  - Bande
    - Encore + 20 Po cette année
- **Pour la suite (2022-2024)**
  - Ajout de lecteurs (+12)
  - Acquisition de librairies supplémentaires (10 frame)
    - En lieu et place d'un des complexe SL8500



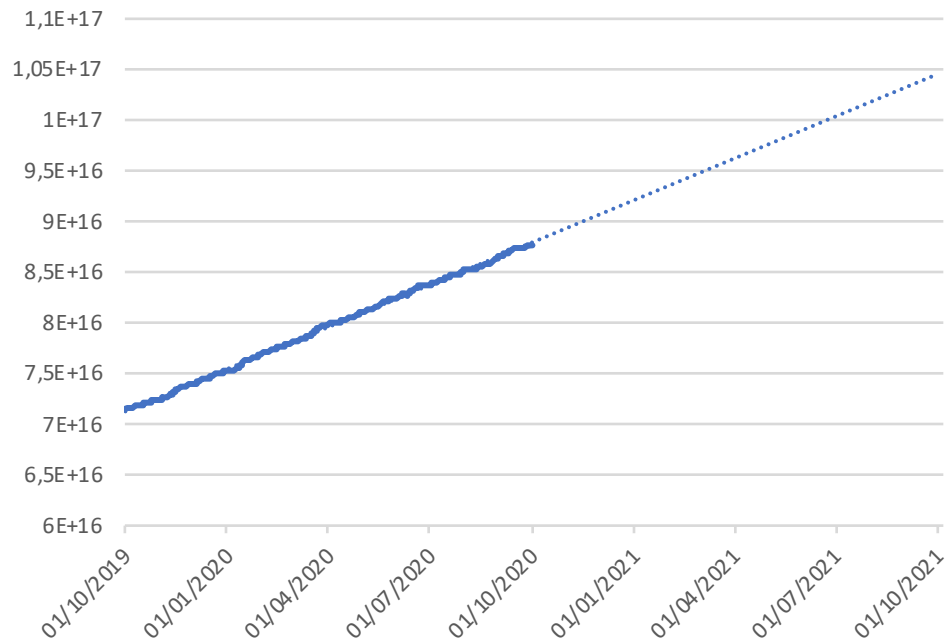


**HPSS**

- **Volumétrie stockée au 1<sup>er</sup> novembre**
  - 89 Po
- **Croissance stable (+16 Po/an) et très linéaire**
- **100 Po d'ici juillet 2021**

- **Utilisation Jaguar E au 1er novembre 2020**
  - 19,5 Po
    - ~ 10,7 Po de nouvelles données
    - ~ 8,8 Po de données repackés
  - Médias
    - 897 bandes EOM
    - 56 bandes partiellement remplis
    - 304 bandes allouées
    - 680 bandes non allouées

CROISSANCE HPSS DEPUIS 1 ANS



- **Migration HPSS (juin 2020)**

- HPSS v 7.5.3u12
- Nouveau PVR optimisé pour les libraries SCSI (Spectra/IBM)
- Logs basé sur syslog
- Gros changement dans la structure des logs
- Configuration syslog-ng via puppet

- **Evolution infrastructure**

- 4 nouveaux movers tapes
  - DELL R640
  - 2x 10 Gbits
  - 6 drives TS1160 / mover
- 23 drives en production

- **Reconfiguration des COS 11 / 12**

- Avant :
  - 64 Mo – 512 Mo : → COS 11
  - 512 Mo – 2 Go : → COS12
- Actuellement
  - 64 Mo – 2 Go : → COS12
  - 60 fichiers / agregats
  - 5 bandes écrites en //
- But :
  - Réduire le nombre de bande ouvertes en écriture
  - Consolider le cache disque



- **Robotique :**
  - Consolider les logs de la librairies Spectra
    - Historique de montage, etc.
    - Utilisation des l'API XML
- **HPSS :**
  - COS double copie (COS 50) toujours sur T10K-D
    - Problèmes lié à LBP
    - Actuellement en cours de migration (« recover »)
  - COS petits fichiers (COS 10) toujours sur T10K-D Sport
    - Des suggestions ?
  - Recréer les règles pour exploiter les logs dans Elasticsearch/Kibana
  - Adapter RFIO pour supporter les UID > 65535
- **Repack :**
  - 72 Po à recopier des T10K-D → Jaguar E
    - Repack COS12 et COS14
    - Changecons COS11 → COS12
  - Planning : 3 ans pour « vider » le robot Oracle.

## Migration HPSS 8.3 / 9.1

- 8.3 : Ordered Migration by Directory
- 9.1 : Buffered Tape Mark ! → Amélioration des performances en écriture

## Technologie bande

- Enterprise
  - TS1170 / Jag F
    - Lecteurs « Compatibility » : 40 To
    - Lecteurs « Capacity » : 60 To
- LTO
  - LTO9 : 18 TO
  - OPEN RAO !!!
  - Q1 2021

### Tape Drive Roadmap

LTO Generation	LTO-8	LTO-9	LTO-10 (Quantum-Safe)
Max Capacity (Native)	12 TB (L8)	18 TB (L9)	Up to 36 TB (LA)
Other Capacities (Native)	9 TB (M8) 6 TB (L7)	12 TB (L8)	18 TB (L9)
Native Data Rate	360 MB/s	Up to 400 MB/s	Up to 500 MB/s
Attachment	FC-8	FC-8, 12Gb SAS	FC-16, 12Gb SAS

4Q2017 → 4Q2018 [LTO-9 + 12 months] [LTO-10 + 12 months]

TS1100 Generation	TS1160	Gen-7 (Quantum-Safe)		Gen-8 (Quantum-Safe)	
Mode	Compatibility	Capacity	Compatibility	Capacity	Compatibility
Max Capacity (Native)	20 TB (JE)	Up to 60 TB (JF)	Up to 40 TB (JF)	Up to 120 TB (JG)	Up to 80 TB (JG)
Other Capacities (Native)	15 TB (JD) 10 TB (JD) 7 TB (JC) 4 TB (JC, R/O)	None	20 TB (JE) 15 TB (JD) 10 TB (JD)	None	Up to 40 TB (JF) 20 TB (JE)
Native Data Rate	400 MB/s	Up to 500 MB/s	Up to 500 MB/s	Up to 1000 MB/s	Up to 1000 MB/s
Attachment	FC-16, 10 GigE (RoCEv2), 12Gb SAS	FC-16, 12Gb SAS, 25 GigE (RoCEv2)	FC-16	FC-32, 12Gb SAS, 25 GigE (RoCEv2)	FC-32

Any statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Ed Childers/ Oct 2020 / © 2020 IBM Corporation

14