CTA at PIC

Site report

Eli Carrasco

CTA Team @ PIC

PIC as a data center

- Spanish WLCG Tier-1 centre → Provides ~4% of Tier-1 data processing

- 23 employees, 9 at the Operations Team

- Support to many astronomy and cosmology projects

CPU: +8k cores under HTCondor v.9.0.17
Disk: 18.3 PB - running on dCache v8.2
Tape: 69.5 PB - running on Enstore v6.3.4-14

Our tape infrastructure

IBM TS4500

SL8500

2 Tape Libraries→SL8500 decommissioned after 2yr data migration

IBM TS4500 with:

→ 5 frames (L55+D55+3xS55)

→ 10 L8 drives + 11 L9 drives - x4 on each tapeserver

→ 5000 tape volumes, mostly LTO8

→ another S55 and a second accessor to be purchased soon

Enstore is our tape storage system since 2007 + **dCache** for disk

Soon **Enstore won't be supported** anymore by FNAL

→ started exploring CTA as a replacement

Why CTA?

- We need a system that integrates well with dCache

  - We knew the DESY team were working on integrating both systems

- Compatibility with CPIO tape format

- Possibility of having similar (although not identical) logical structure to Enstore

- Opensource

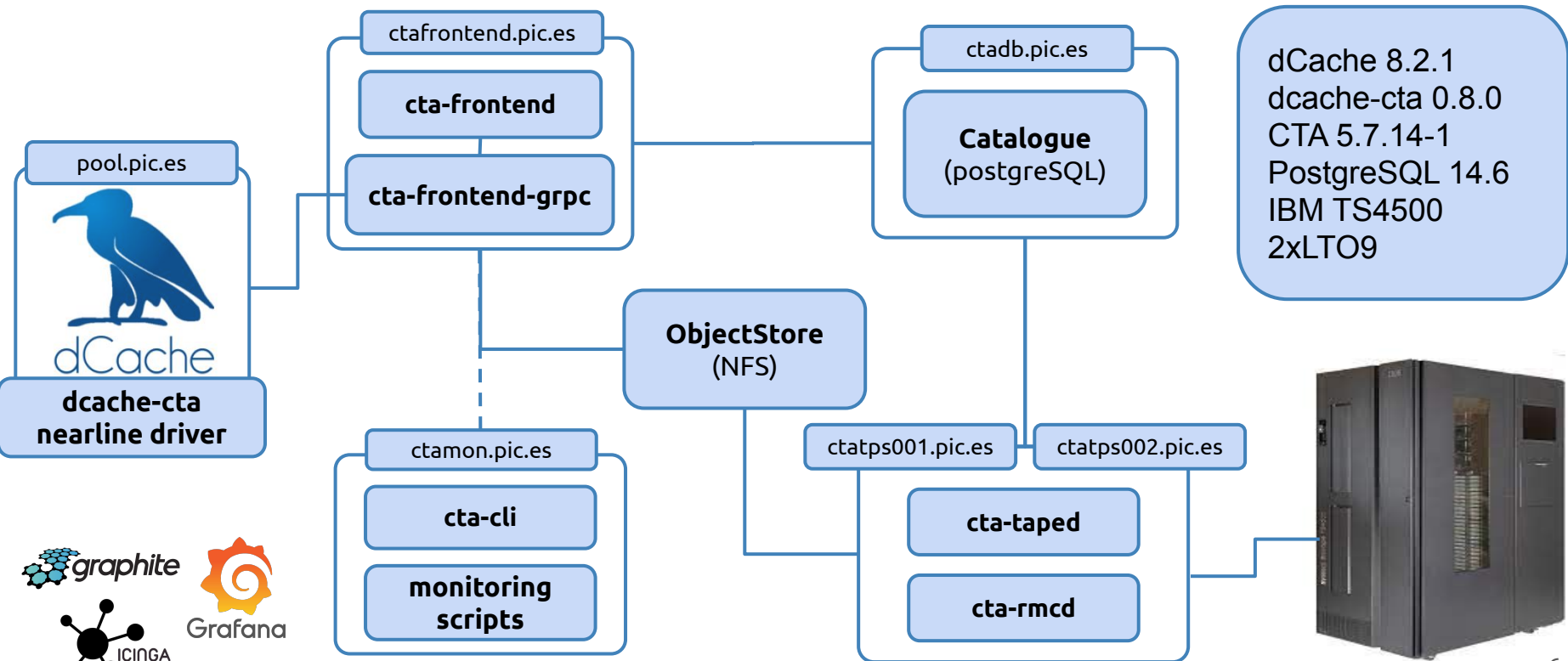- Wide adoption by the WLCG community

  - CERN, RAL, FNAL, IHEP, DESY…

CTA at PIC

Our team:

➔   1 full-time dedicated team member

➔   1 "part-time" team member who assists and helps me regularly

➔   Another team member that helps me with the dCache side as needed

Our approach → let's build a small test instance

➔   Package-based installation

  -   Limited experience with Kubernetes

  -   Allows us to fully understand how CTA works

➔   Started on a single server + MHVTL and progressively deployed our instance

Our dCache+CTA test instance

PIC port d'informació científica

ctafrontend.pic.es
- **cta-frontend**
- **cta-frontend-grpc**

pool.pic.es
- dCache
- **dcache-cta nearline driver**

ctadb.pic.es
- **Catalogue** (postgreSQL)

ObjectStore (NFS)

ctamon.pic.es
- **cta-cli**
- **monitoring scripts**

ctatps001.pic.es   ctatps002.pic.es
- **cta-taped**
- **cta-rmcd**

dCache 8.2.1
dcache-cta 0.8.0
CTA 5.7.14-1
PostgreSQL 14.6
IBM TS4500
2xLTO9

Steps done so far

- Built a functional logical structure while keeping the concept of File Families

- R/W operations

  - Understanding how does the scheduler work

  - Understanding how does CTA interact with dCache

- General testing: mount policies, repack operations, drive compression, …

- Version updates & upgrade to CTA v5

- Started building our system and monitoring tools:

  - Nagios check for alerts (services, drive status, tp occupation, failed requests…)

  - Metrics (collectd + custom scripts sent to graphite) →grafana dashboard

- Automated all our configuration with **puppet**

Some examples of our monitoring system

PIC
port d'informació
científica

LTO8 - Mounts
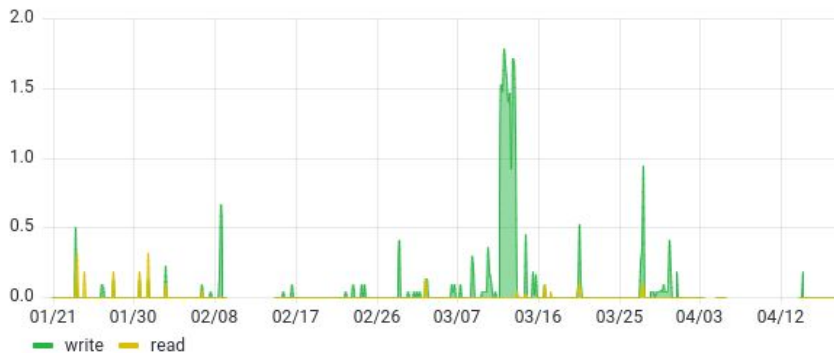
Tape Bandwidth by VO and FF
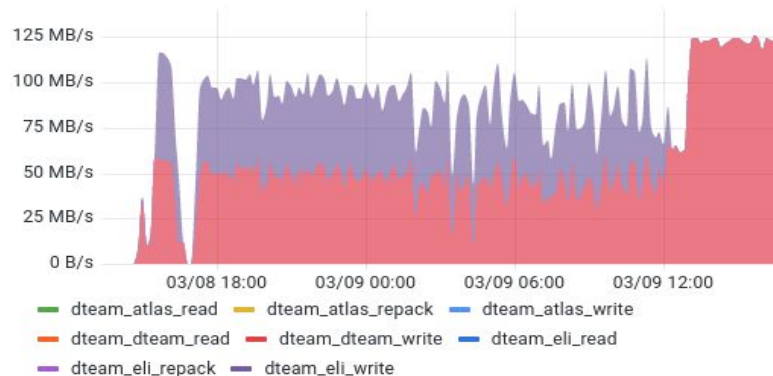
Host	Service	Status	Last Check	Duration	Attempt	Status Information
ctadb01-test.pic.es	CTA DB	OK	15:40:08	1d 21h 46m 29s	1/3	[OK] Services postgresql-14 running
ctafrontend01.pic.es	CTA Frontend	OK	15:40:08	1d 21h 46m 29s	1/3	[OK] Services cta-frontend running
	CTA Frontend GRPC	OK	15:31:12	1d 4h 25m 25s	1/3	[OK] Services cta-frontend-grpc running
ctamon01.pic.es	CTA Empty Tapes	OK	15:40:08	1d 21h 54m 59s	1/3	[OK] 4 tapes left in none tapepool
	CTA Failed Requests	WARNING	15:39:18	1d 2h 17m 19s	3/3 #1	[WARNING] There are 10 failed requests. Please ch
ctatps001.pic.es	CTA Frontend	OK	15:40:08	1d 21h 46m 29s	1/3	[OK] Services cta-frontend running
	CTA RMCD	OK	15:40:08	1d 21h 46m 29s	1/3	[OK] Services cta-rmcd running
	CTA Taped	OK	15:40:08	1d 21h 46m 29s	1/3	[OK] Services cta-taped running
ctatps002.pic.es	CTA Frontend	OK	15:40:08	1d 21h 46m 29s	1/3	[OK] Services cta-frontend running
	CTA RMCD	OK	15:40:08	1d 21h 46m 29s	1/3	[OK] Services cta-rmcd running
	CTA Taped	CRITICAL	15:42:08	0d 4h 46m 29s	3/3	[CRITICAL] Services cta-taped not running

Thoughts and experiences

- Confusion about dependencies, versioning + a few deployment issues until:
  - switched from MHVTL to our real tape library (IBM TS4500 partitioned)
  - public rpms were finally released after several months
- Documentation was a bit scattered and not very beginner friendly / assumed previous tape knowledge
  - surfing the **community forum** was of great help!
- My worst nightmare: `EOSReporter::AsyncQueryHandler::HandleResponse(): failed to XrdCl::FileSystem::Query() [ERROR] Operation expired code:206 errNo:0 status:1`
  - Good news it's over!
  - We found some kafka test machines guilty but can't figure out why?
- We still want to avoid Ceph and K8S

CTA's future at PIC

- Test tapeservers efficiency with 4 drives

- Test with more realistic loads to test performance & bandwidth

- Keep refining our monitoring system, centralise logs...

- Start using CTA for internal backups

- Discuss the use of compression - never enabled with Enstore

- Define how the data migration from Enstore will be done

- No date to switch to production yet!

The CTA(+dCache) Book

https://ctadcache.gitbook.io/

Preface

What is this book about?

This book is designed to help you get started with the *Cern Tape Archive* (CTA) system. It provides an overview of the general concepts and covers the installation process in detail. Additionally, the book includes answers to frequently asked questions as well as guidance to other procedures.

Please note that **this book is not the official documentation** and it is specifically focused on CTA+dCache integration for now. That means that some EOS explanations and configuration parts will be omitted and others will be exclusively applicable for dCache users.

So while some parts of the book may not be directly applicable for EOS users, the information provided on some concepts, installation, config and common troubleshooting procedures may (hopefully) still be useful for someone.

We welcome and encourage any feedback: feel free to submit any suggestions, missing pages or improvements to the book!

11

Thank you! Questions?

We are open to share our monitoring scripts, puppet code and whatever would be of help. Feel free to ask!

Big thanks to everyone @ the CTA Community!

elisabet@pic.es