



Ian Randall
Systems Analyst
DHS Team

ECMWF Site Presentation

European co-operation at its best: **pooling resources**



A world map with a light gray background. The continent of Europe is highlighted in a dark blue color. The rest of the world's continents are shown in a light gray tone.

ECMWF's role is to address the critical and most difficult research problems in medium-range NWP that no one country could tackle on its own

PLAYING A UNIQUE ROLE

- 
- A world map with a light gray background. The continent of Europe is highlighted in a dark blue color. The rest of the world's continents are shown in a light gray color with faint outlines of countries.
- 34 Member and Co-operating States
 - 350 staff +
 - Coming from 30 countries

EUROPEAN AT HEART



GLOBAL COLLABORATION

ECMWF The Data Handling System (DHS)

- Based on HPSS (7.5.3.0.u10) with bespoke MARS and ECFS front-end client.
- Installed in 2002 with most recent hardware upgrade being two 15 frame TS4500 Tape Libraries and currently 200 TS1160 tape drives.
- Archive size 451PB (312 million files) of data (December 2019). Stored on a mixture of 3592 JE (Primary data), T10K (legacy primary copy) and LTO (2nd copy backup)
- **Currently migrating T10K → 3592 JE at around 400-500 cartridges per week. Migration is expected to complete by July 2021.**
- Currently archiving ~260TB per day and restoring ~200TB per day.

ECMWF Current DHS Hardware

- 1 production Core server plus a standby core
- Dell R930 HPSS CORE (RHEL 7.4). 1TB of RAM FS900 Flash Array
- 15 disk movers (CentOS 7.4) and 60 tape movers (CentOS 7.4)
- 12 PB disk cache (IBM V7000, IBM V5000, DDN SFA7700 & DDN SFA14k)
- 16Gb switched FC, 20GbE external network (MARS & ECFS)
- 4 Oracle SL8500 Tape Libraries in one complex
- 187 Oracle T10KD tape drives
- ~27,000 T2 media, 43,000 LTO media (LTO6 LTO7 LTO8m), 7,000 3592 JE media

ECMWF TS4500 deployment (storlib1 & storlib2)

- 2 X 15 frame TS4500 tape library. Each with 11,620 licenced slots
- 192 production TS1160 tape drives.
- Delivery in late January 2019 and available for testing March 2019
- Production deployment delayed until June 2019 due to pending HPSS upgrade and AIX to RHEL conversion.
- 8 drives per host (Dell R440) direct connection. Two quad port HBA's
- Hosts could sustain maximum bandwidth 400MB/s read/write across all 8 drives in benchmarking.
- Several revisions of drive and library firmware later we are in a fairly stable environment 😊
- In production each physical library is split into two logical libraries:

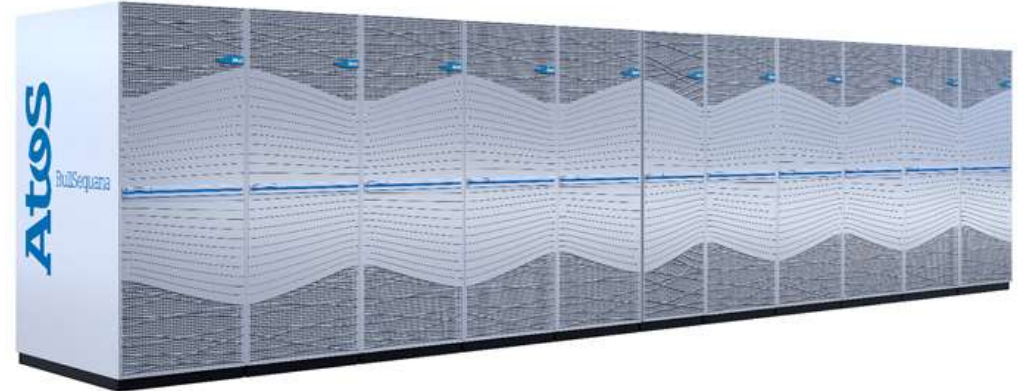
Drive Frames

STORLIBx_A = F3 → F6

STORLIBx_B = F11 → F13

ECMWF Our next Supercomputer

- Increasing the ensemble forecast horizontal resolution from a grid spacing of 18 km to about 10 km, which is expected to significantly improve forecasts of near-surface temperatures and winds
- Increasing the ensemble forecast vertical resolution from 91 layers to 137 layers, in line with the current high-resolution forecast
- Issuing extended-range forecasts daily rather than twice-weekly.
- To be hosted in our new data centre located in Bologna, Italy.



- BullSequana XH2000 Supercomputer
- AMD EPYC™ processors
- Mellanox DHR InfiniBand (200Gb/s)
- 91 PB of DDN EXAScaler Storage
- Water Cooling

ECMWF The BOND Project



- Located in Bologna, Italy
- The site will house all ECMWF compute (HPC, DHS, Servers)
- 25 permanent staff will be based there
- Building handover expected Q2 2020
- Users will still be based in Reading, UK and collaborators around the World will all connect remotely.

- DHS Team planning to move archive Q1 2021.
- Pre-built tape libraries will already exist in Bologna.
- Media will be transported via road. 2nd copies first then primary data.
- All HPSS disk cache will be migrated before disk systems are transported
- HPSS will be shutdown at UK site and spun up in Bologna.

ECMWF DHS Current Areas of Concern

- Concerns over ongoing supplier support (Oracle). We still have 27,000 cartridges to migrate to new media.
- Exponential data growth is not sustainable.
- Currently growing at 45% each year!
- Cost of physical media (LTO and 3592) isn't decreasing.



Any Questions?