

# Data Handling System at ECMWF

F. Dequenne

May 2004

[francis.dequenne@ecmwf.int](mailto:francis.dequenne@ecmwf.int)

# What is ECMWF?

## *European Centre for Medium-range Weather Forecast*

- **European meteorological organisation supported by 25 European States.**
- **Develops medium-range and seasonal forecasting through numerical methods.**
- **Provides on daily basis to its Member States 10-days worldwide-weather forecast and other weather-related products.**
- **Provides to its Member States an organised on-line access to 27 years of weather forecasts and observations.**
- **Provides to the world meteorological and research community retrieval facilities for specific datasets.**
- **Provides extensive facilities for weather modelling research.**

# Member States

---

**Belgium**

**Denmark**

**Germany**

**Spain**

**France**

**Greece**

**Ireland**

**Italy**

**Luxembourg**

**The Netherlands**

**Norway**

**Austria**

**Portugal**

**Switzerland**

**Finland**

**Sweden**

**Turkey**

**United Kingdom**

## *Co-operation agreements or working arrangements with:*

**Czech Republic**

**Croatia**

**Iceland**

**Hungary**

**Romania**

**Serbia & Montenegro**

**Slovenia**

**ACMAD**

**EUMETSAT**

**WMO**

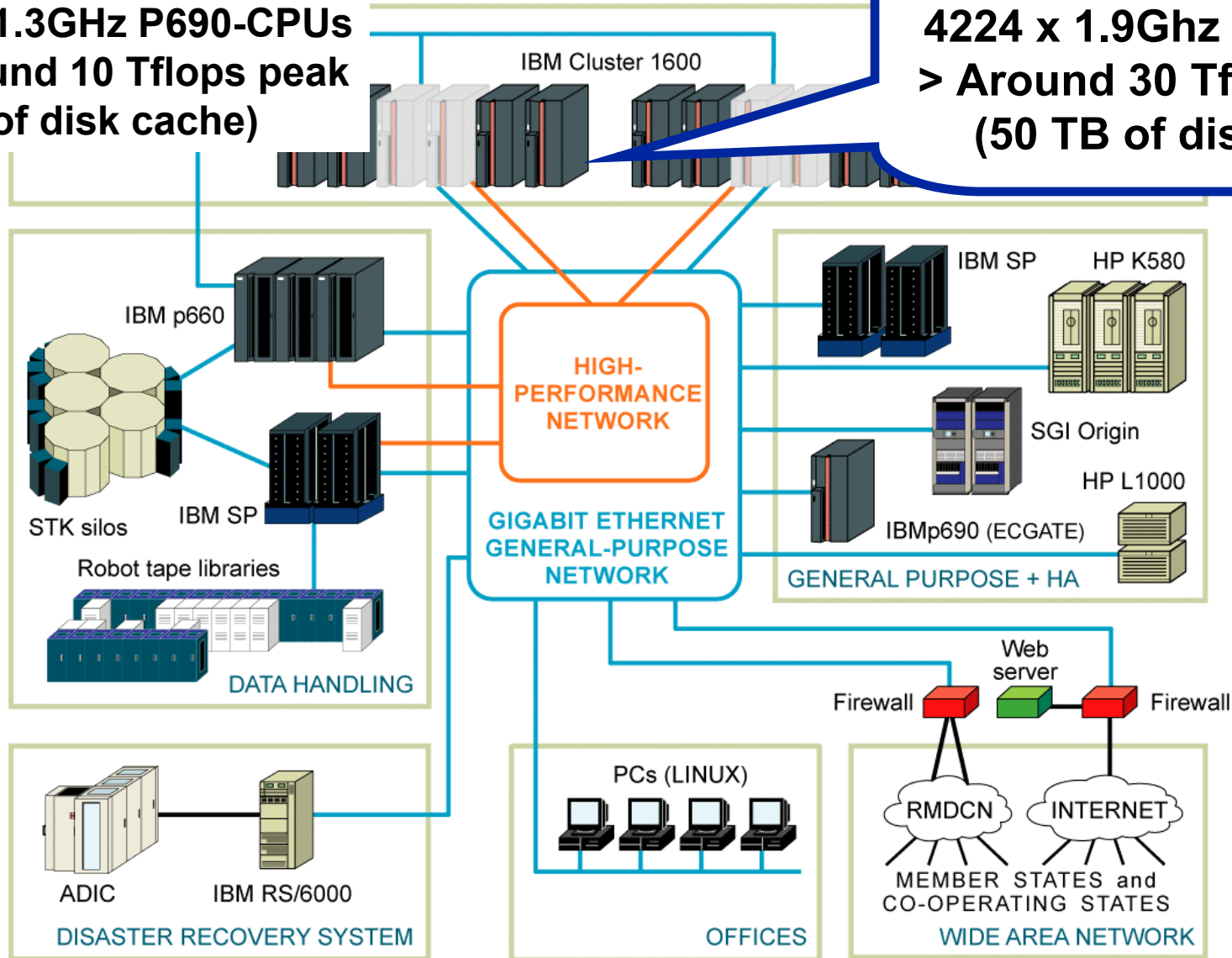
**JRC**

**CTBTO**

# ECMWF Computer Environment

920 x 1.3GHz P690-CPU  
> Around 10 Tflops peak  
(13 TB of disk cache)

To be replaced by  
4224 x 1.9GHz P690-CPU  
> Around 30 Tflops peak  
(50 TB of disk cache)



# Data Handling Applications

- **MARS**

- **Meteorological Archive and Retrieval System.**

- **Bulk of the data, few files**

- **Interfaced through an ECMWF application.**

- **Depends heavily on tape get-partials.**

- **ECFS**

- **HSM-like service for “ad-hoc” files.**

- **Millions of files, many very small.**

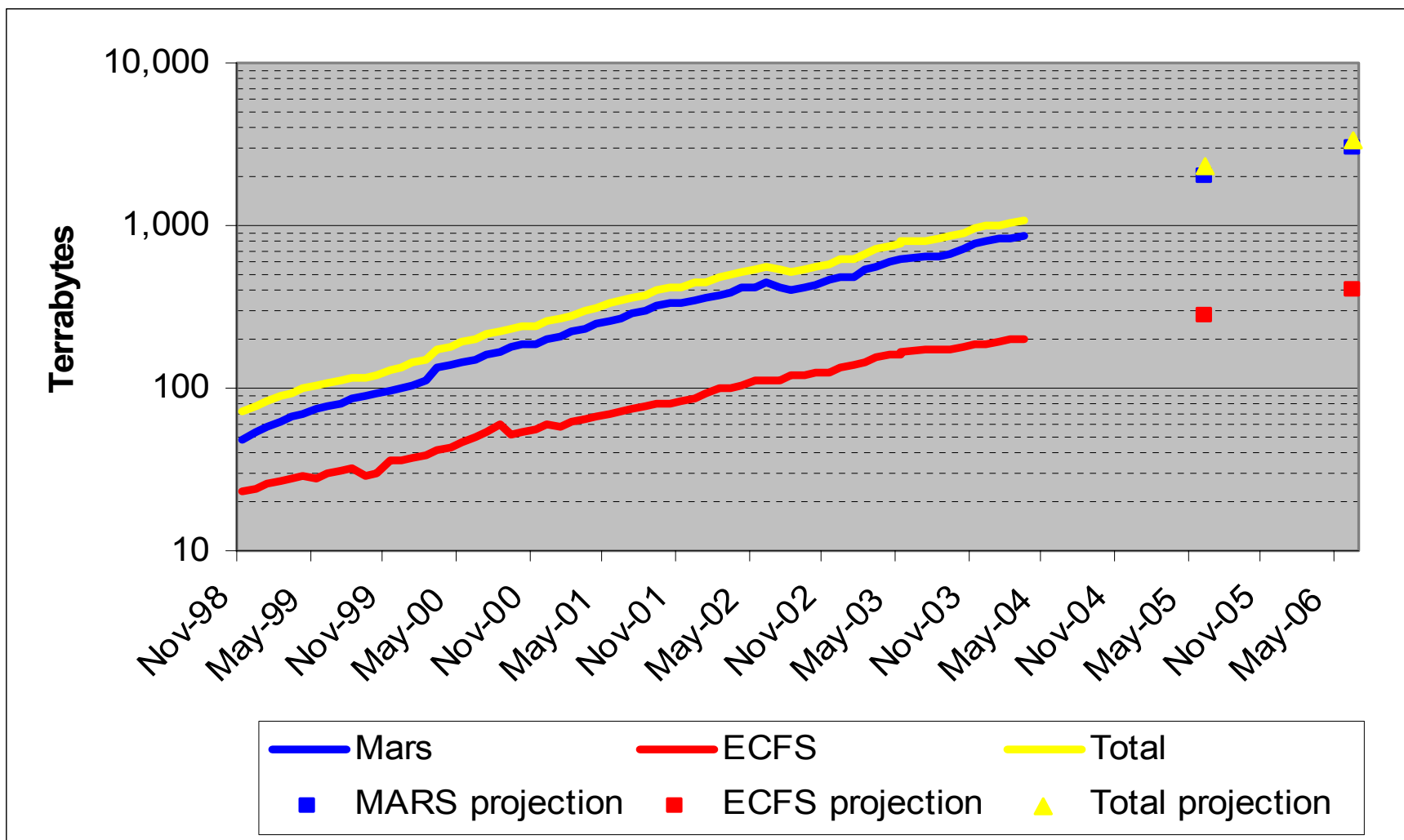
# DHS Services: MARS

- Close to 900TB of data, all stored in HPSS.
- 1 million files.
- 1 TB new data stored every day.
- Data indexed by in-house application, providing a powerful virtualisation engine.
- Requires many tape drives able to load and position tapes quickly.
- Medium to long term archive.
- Comprises
  - MARS Operational (40% of the data, WORM, backed-up)
  - MARS Research (60%, WORS, no backup)

# DHS Services: ECFS

- 200 TB of data (+ backup copy)
- 10 Million files
- 500 GB of data stored daily (but up to 50GB/hour peaks)
- Volatile data
- A lot of small files
  - 10% of files < 10KB
  - 35% of files < 512 KB
- The migration to HPSS has started, and should complete by end 2004
  - Most new data is now saved in HPSS
  - The old data is progressively copied to HPSS

# Volume of data stored



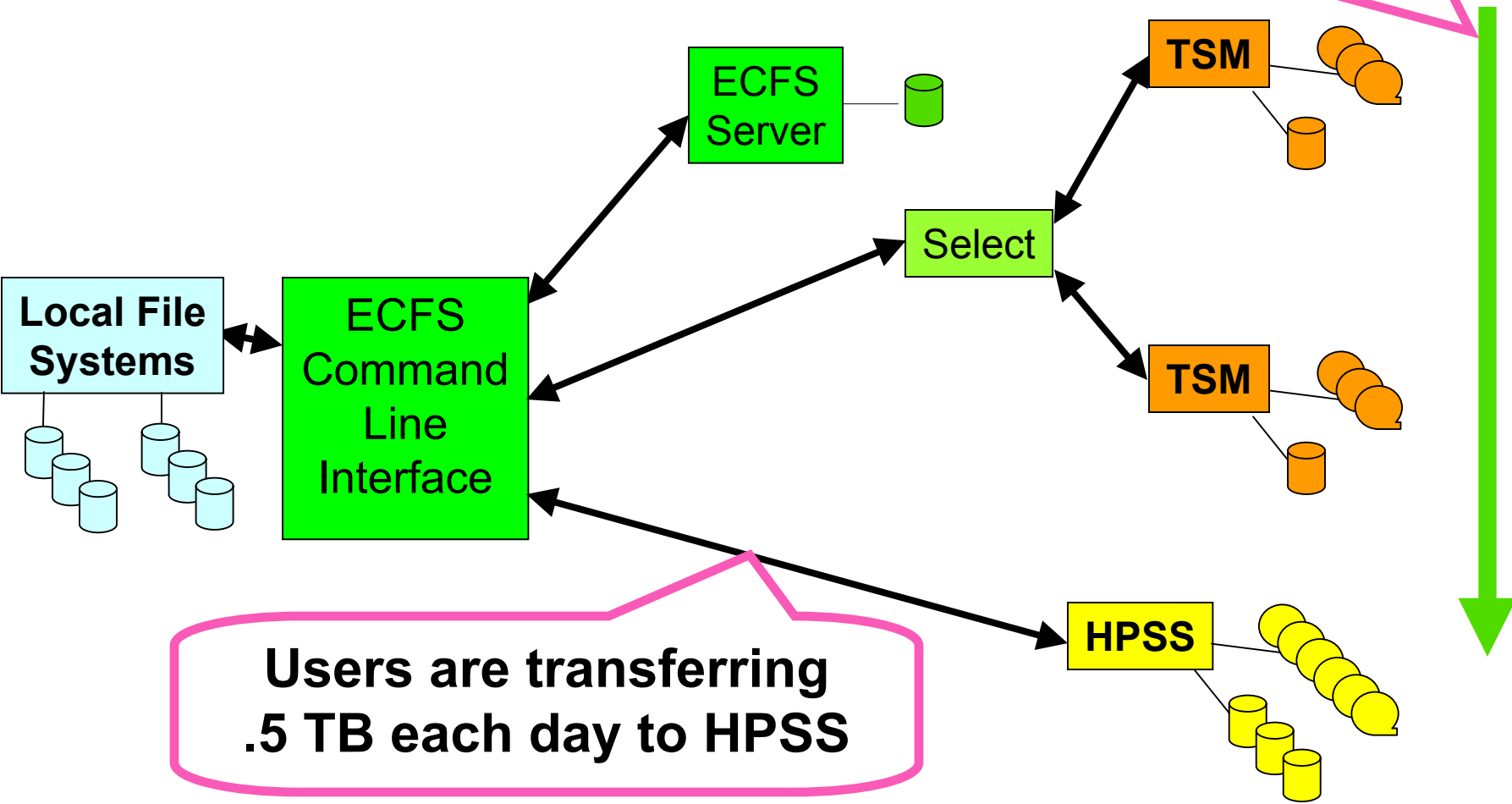
**NB: These values do not include the second backup copy of our most critical data.**

# What happened since the last HUF?

- **MARS fully ported to HPSS. (September)**
- **ECFS is being migrated to HPSS. (Started in March)**
- **We use disk/tapes hierarchies... (In ECFS context, March)**
- **IBM 3592 drives for MARS (part) and ECFS (all) data. (December)**
- **AML robots with LTO/1 and LTO/2 drives. (July?)**
- **We reached 1 PetaByte of primary data stored in HPSS. (April)**
- **We beta-tested HPSS 5.1 and put it in production. (November)**
  - ➔ **Thanks to the support and development teams for their help!**

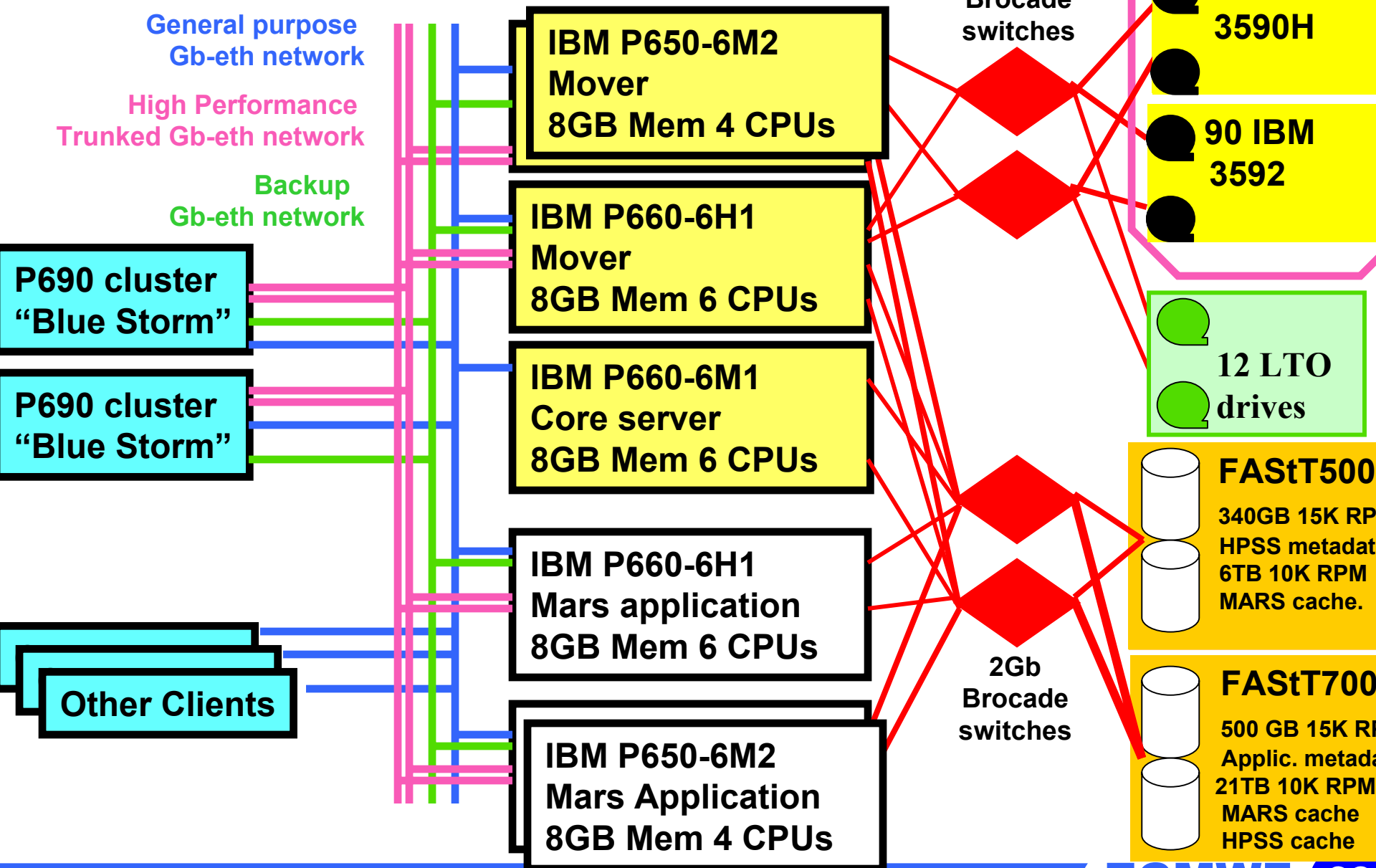
# ECFS implementation

.5-1.5 TB is copied each day from TSM to HPSS

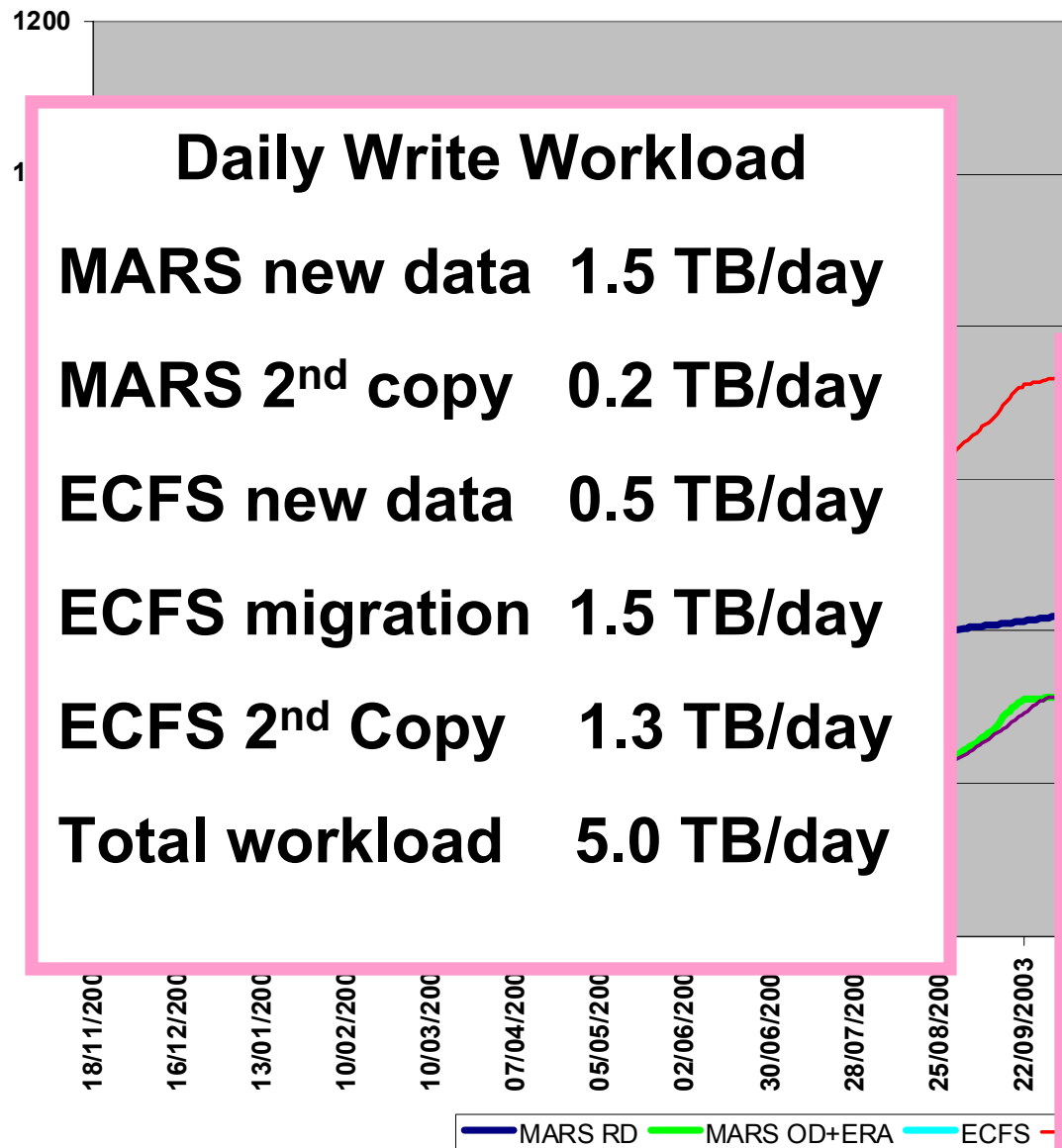


Users are transferring .5 TB each day to HPSS

# HPSS configuration by End 2004



# HPSS data growth



# Issues.

- **Poor performances of writing small files to tapes. We are impatient to see user-transparent file-aggregation mechanisms.**
- **Need to stop PVL to add devices.**
- **No automatic attempt to access secondary tape copies if first copy is down.**
- **CLI for all administrative functions.**
- **Disk fragmentation.**
- **Slow Migration rates.**

# Future developments.

- **Conclude ECFS migration.**
- **Deploy a more resilient Disaster Recovery environment.**
- **Support the increased load generated by new super computers.**
- **Get rid of DCE.**
- **Deployment of small-files aggregation tools.**
- **Consider deploying SAN3P.**
- **Start considering new robotics.**