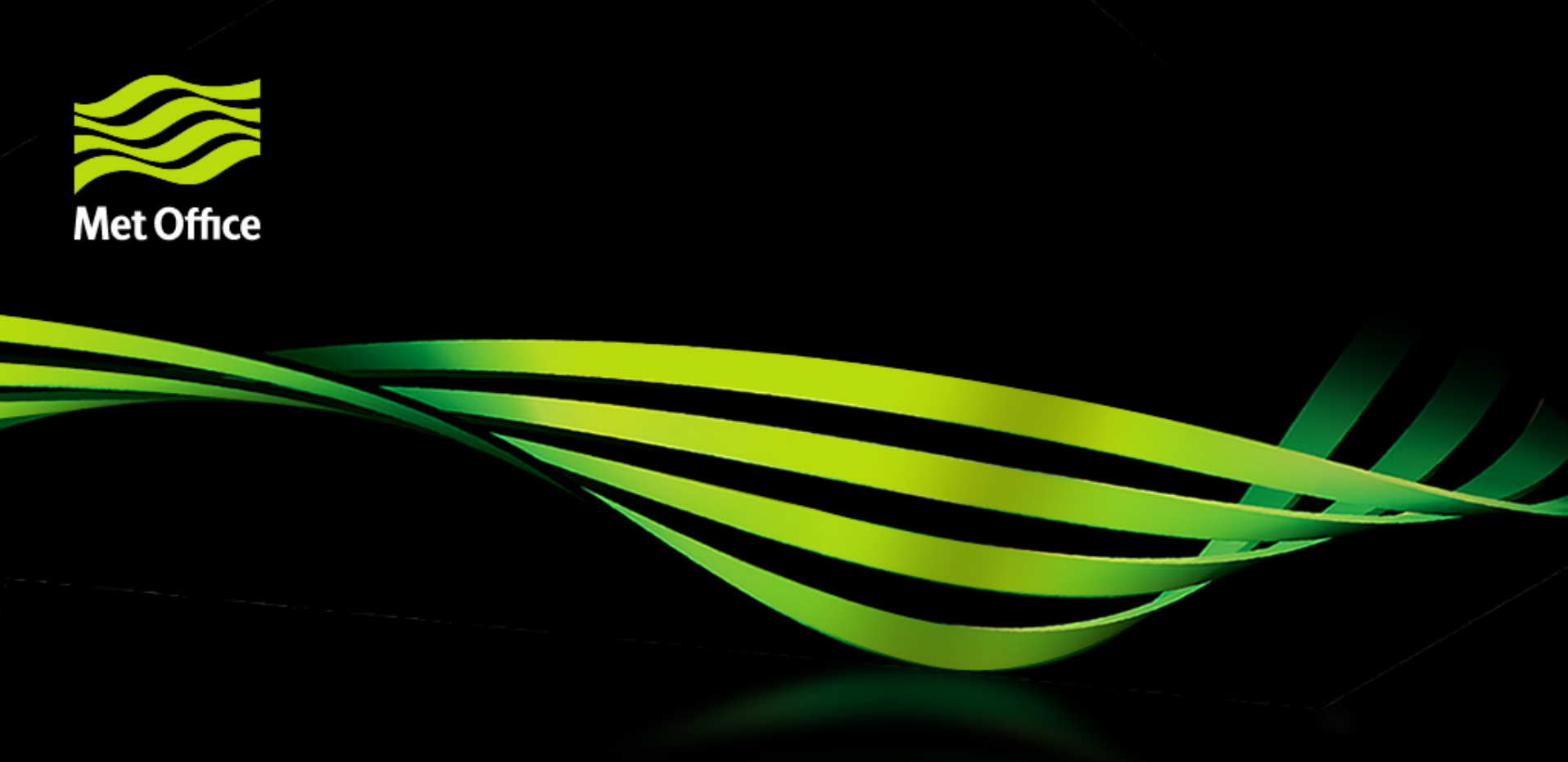


Results from the UK Met Office HPC Procurement

Paul Selwood

© Crown copyright Met Office

Introduction

© Crown copyright Met Office

The Met Office

- National Weather Service for the UK
- Climate Prediction (Met Office Hadley Centre)
- Operational and Research activities
- About 1,700 staff in 60 locations around the world

© Crown copyright Met Office

Met Office Supercomputers

- 21 node NEC SX-8 cluster

- 2 NEC SX-6 clusters
  - 15 node for NWP
  - 19 node for Climate

- 2 x 16 CPU Itanium front-ends
- 4 x 8 CPU Itanium file servers

- Post processing on IBM System Z mainframe

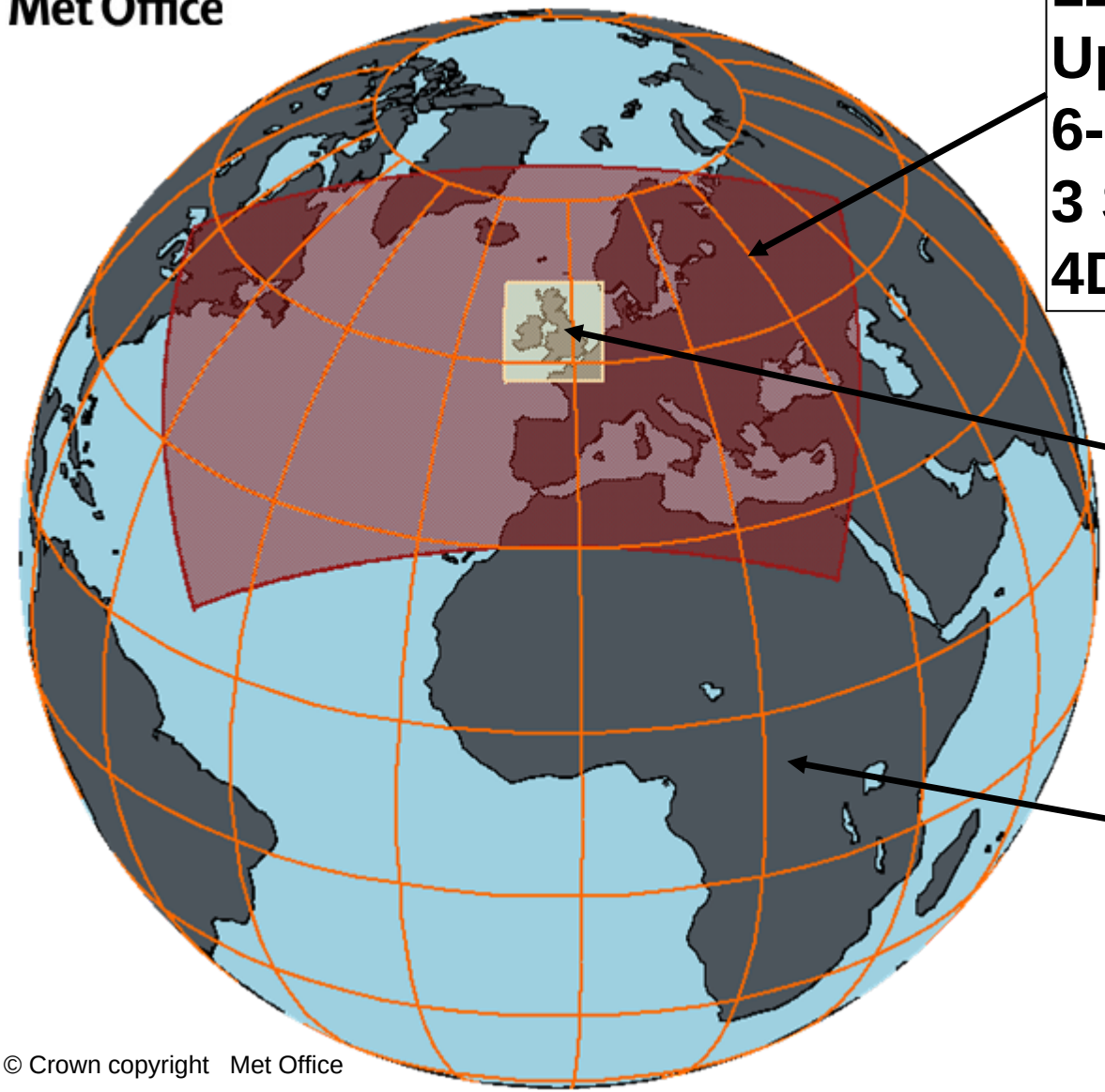
© Crown copyright Met Office

The Met Office Unified Model

- Climate and Forecast model
- Atmosphere, Ocean and Coupled (also sea-ice, atmospheric chemistry, aerosols, river transport, …)
- Atmosphere
  - Non-hydrostatic, semi-Lagrangian, semi-implicit, Arakawa C grid, Charney-Phillips vertical coordinate
- ~ 800K LOC
- MPI parallelisation

© Crown copyright Met Office

Deterministic Forecasts



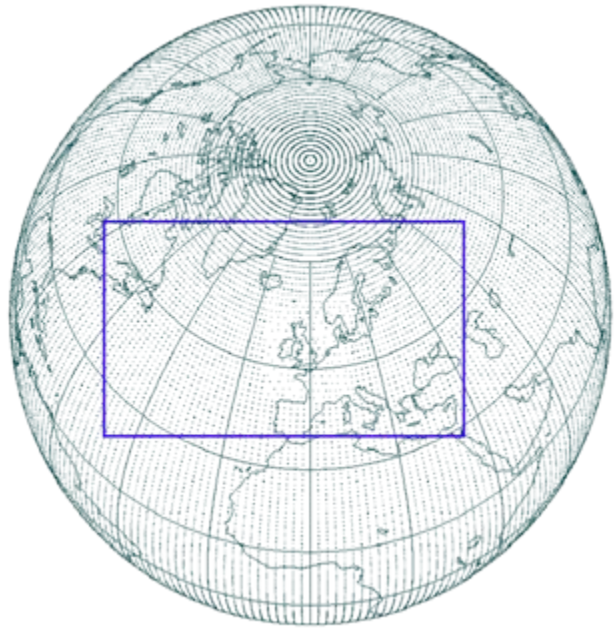
12km – 38 levels
Up to 48hr f/c
6-hourly update
3 SX-8 nodes
4D-Var

4km
Up to 36hr f/c
1 SX-8 node
3D-Var

40km - 50 levels
Up to 7 day f/c
6-hourly update
3 SX-8 nodes
4D-Var

© Crown copyright Met Office

MOGREPS



- 24 members – 3 SX-6 nodes

Global

- Run to T+72
- N144 (~ 90 km)
- Uses Ensemble Transform Kalman Filter (ETKF) for generating initial perturbations
- Stochastic physics – random perturbation of parameterisation schemes
- Also MOGREPS-15, medium range run at ECMWF

LAM

- Run to T+36
- 24 km
- North-Atlantic Europe
- Takes initial and boundary conditions from global model
- Stochastic physics

© Crown copyright Met Office

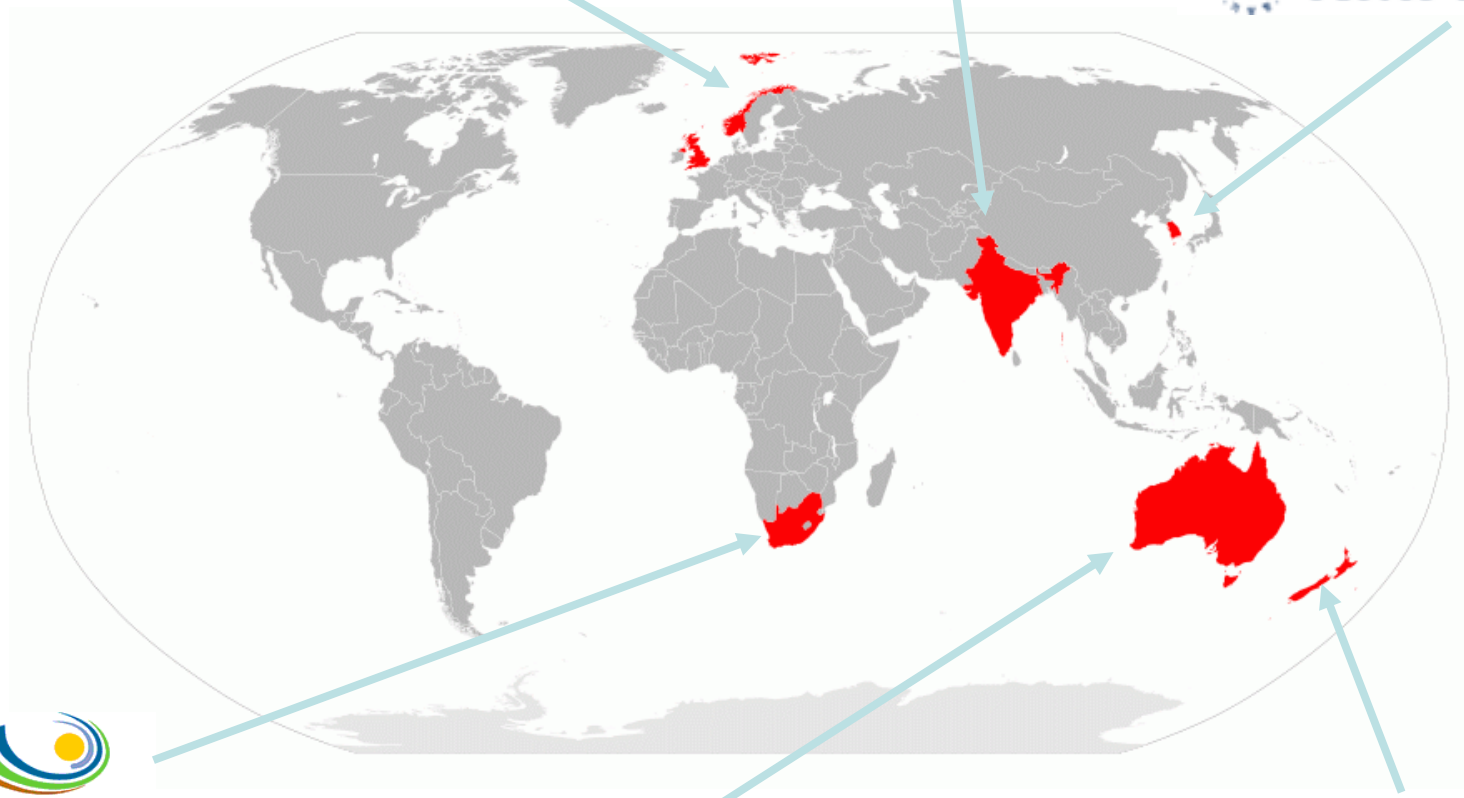
Collaboration

Met Office

Norwegian Meteorological Institute
met.no

सत्यमेव जयते

KMA KOREA METEOROLOGICAL ADMINISTRATION



South African Weather Service

Australian Government
Bureau of Meteorology

CSIRO

NIWA
Taihoro Nukurangi

© Crown copyright Met Office

Storage

© Crown copyright Met Office

MASS to MASS-R

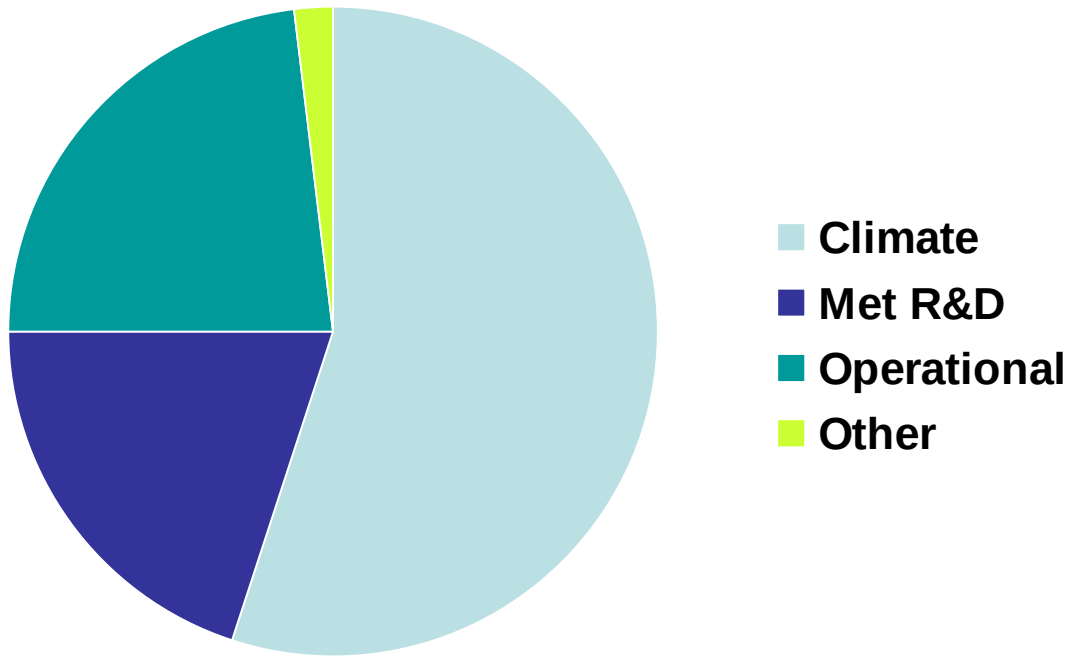
- Managed Archive Storage System
- MASS is a FileTek system using StorageTek, SUN and COPAN hardware

- MASS-R will replace
- IBM HPSS
- Contract awarded March 2008
- First Phase accepted September 2008

© Crown copyright Met Office

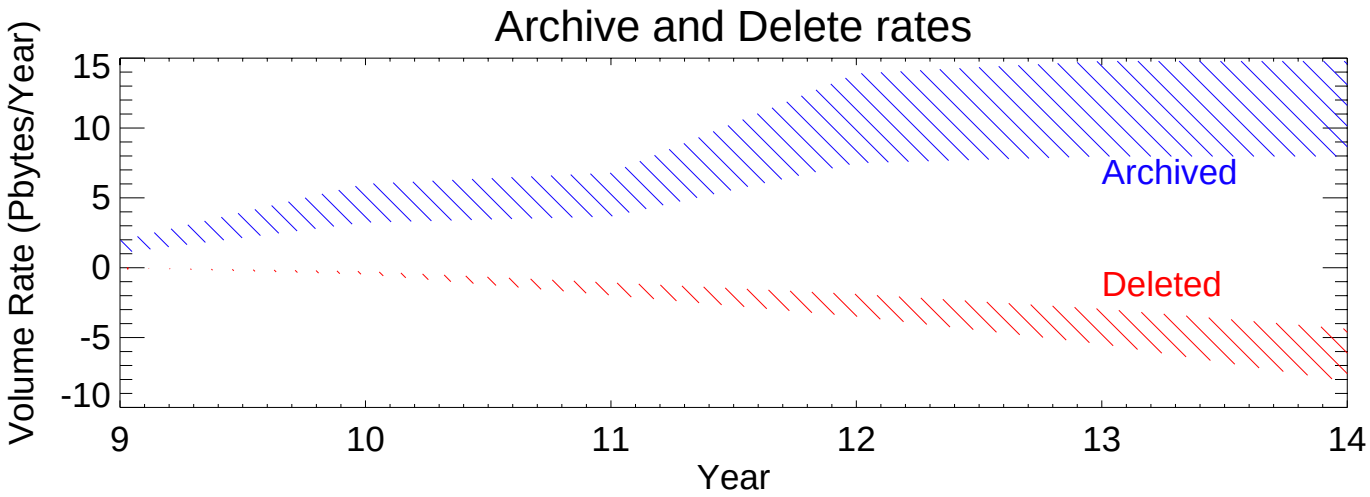
Current storage profile

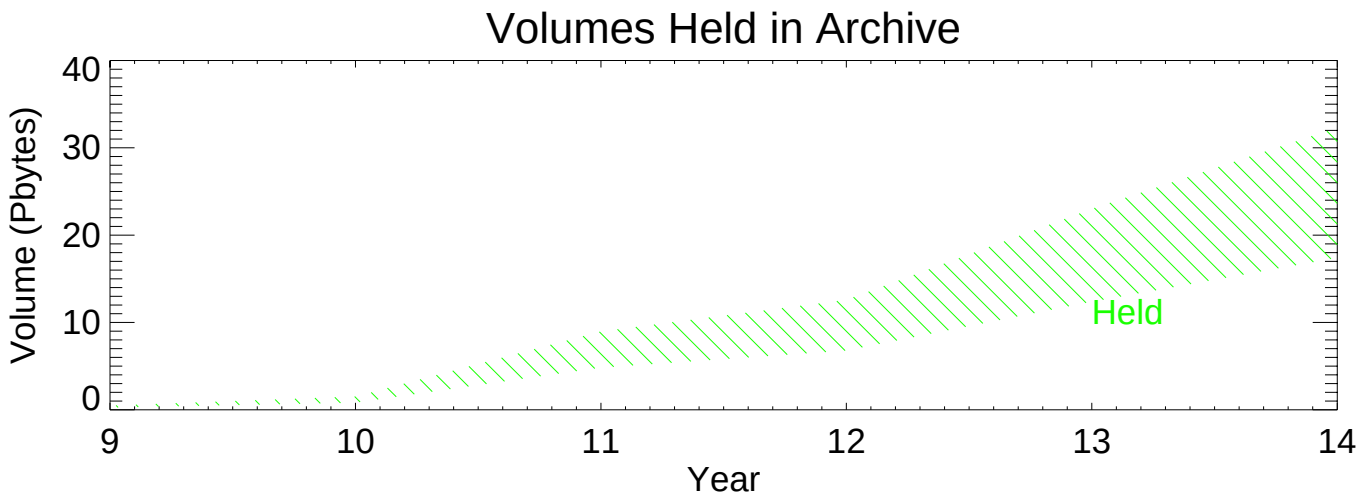
- Archive 1.5 - 2Tbytes/day
- Retrieve ~500Gbytes/day
- Stored volume ~1.5Petabytes



© Crown copyright Met Office

Projected storage





© Crown copyright Met Office

HPSS installation

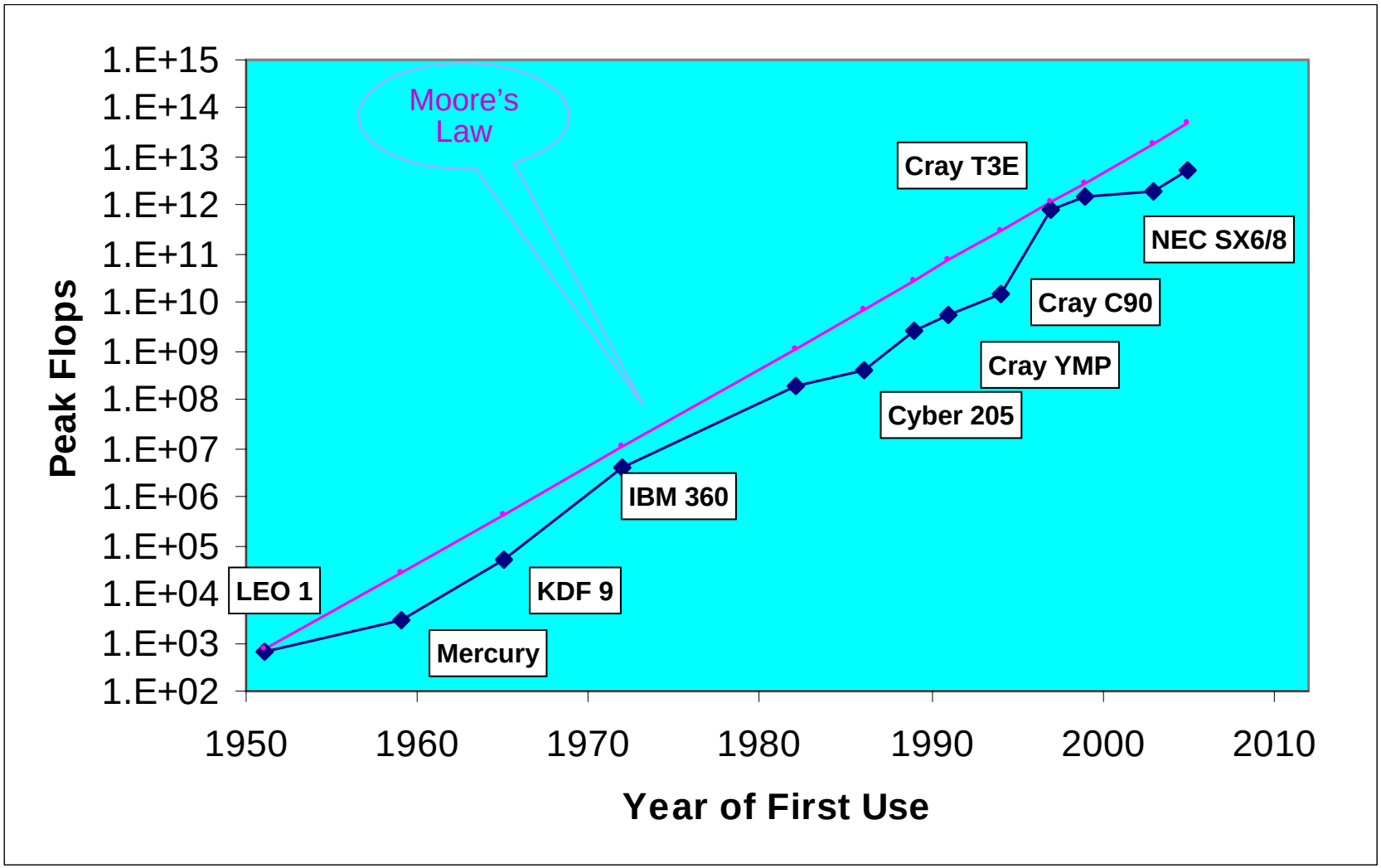
	Disk (>40 days)	**Tape (primary copy)**
Sep 2008	80 Tbytes	400 Tbytes
Jan 2009	~350 Tbytes	5.8 Pbytes
Mar 2014	~1.5 Pbytes	~30 Pbytes

© Crown copyright Met Office

HPC

© Crown copyright Met Office

60 Years of Met Office Computers



© Crown copyright Met Office

What Benchmarks?

- UK 1.5km - L76
- Global N512L76 (24km)
- HadGEM2-A
- NEMO – ½ degree climate configuration
- HadOCC ocean
- 4DVAR – N144L70
- OPS - RTTOV

© Crown copyright Met Office

Benchmark weights

Benchmark	Weight	G Copies	H Copies
HadGEM2-A	0.2	128	256
N512L76	0.2	8	16
UK 1.5	0.25	4	8
NEMO	0.05	64	128
HadOCC	0.1	128	256
4DVAR	0.15	8	16
OPS	0.05	16	32

Speedups combined via harmonic mean to get G and H

© Crown copyright Met Office

Procurement Process

- OJEU notice issued July 2007
  - ~13 responses
  - All provided with initial benchmarks and full Operational Requirement
- 6 completed Pre-Qualification Questionnaire
  - All passed
- Full response to OR in November 2007
  - Evaluation just on technical grounds
  - 4 shortlisted

© Crown copyright Met Office

Process cont

- Updated OR response in February 2008
  - Meetings to clarify/strengthen bids
  - Site reference visits
  - Evaluation 50% technical, 50% financial
  - 2 Shortlisted
- Best and Final Offers in June 2008
  - Technical, commercial and contracts negotiations
  - Final evaluation 100% financial on whole life costs and benefits

© Crown copyright Met Office

The Winner

IBM

G1 = 6.5

H1 = 8.5

© Crown copyright Met Office

Dave Kay (IBM) and John Hirst (Met Office) sign the contract on 01/08/08

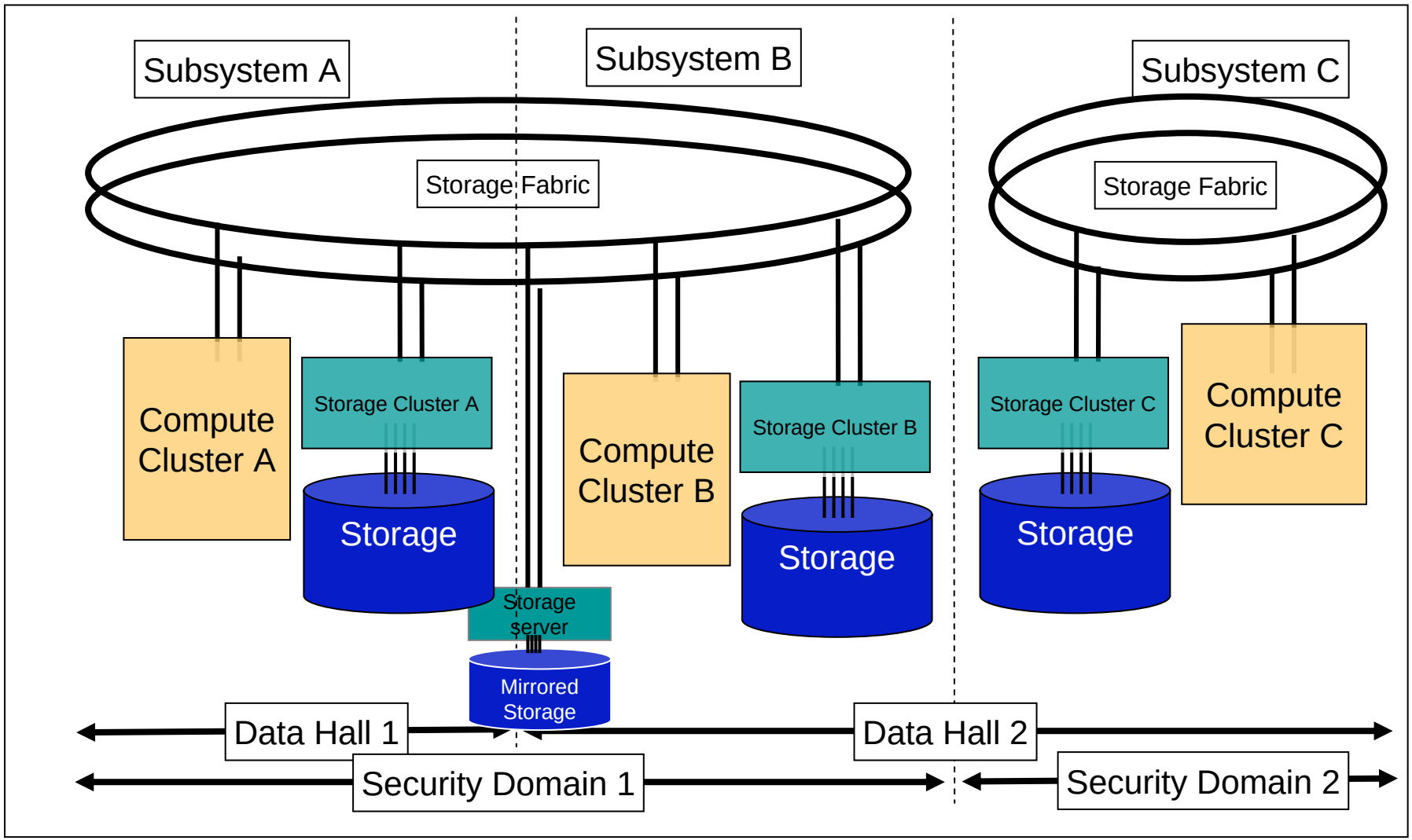
© Crown copyright Met Office

What are we getting?

- 2 main clusters
  - Production / Research
  - 90 compute nodes per cluster
  - 5760 GB memory
  - Shared GPFS
  - 48 GB/s aggregate disk bandwidth
  - 690 TB disk
- Collaboration cluster
  - 28 nodes
  - Memory and Storage proportional to main clusters
  - Non-shared GPFS
  - Shared with NERC (50:50 funding)
- Test and Dev cluster
  - 2 compute nodes
  - 6TB storage

© Crown copyright Met Office

System Overview



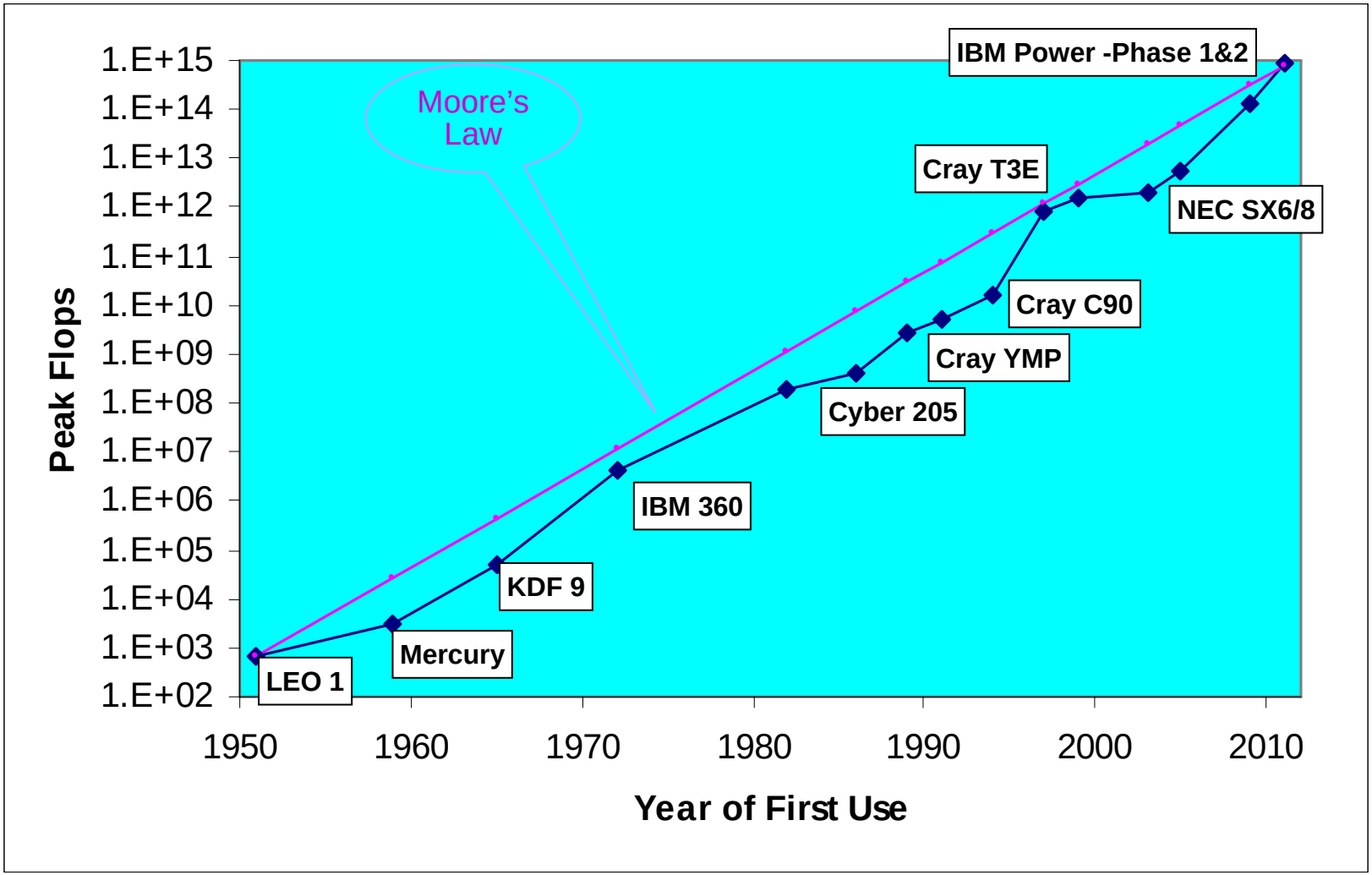
© Crown copyright Met Office

Summary compared with NEC

	NEC SX6/8 3 systems	IBM Power6 3 systems	Factor
CPUs or Cores per Node	8	32	4
Peak Performance per node (GFLOPS)	128 **(for SX8)**	600	4.6
Number of Nodes	59 (SX6/SX8 mix)	208	3.5
Total Peak Performance (TFLOPS)	5.4	125	23.1
Number of CPUs / Cores	472	6656	14.1
Total Memory (TBytes)	2.7	13.3	4.9
Total Disk (TBytes)	36	776	21.5
Disk Performance (GB/s)	~0.15	>1 (24 total per cluster)	~7

© Crown copyright Met Office

60 Years of Met Office Computers redux



© Crown copyright Met Office

Implementation plans

- Test System for code porting - acceptance mid-October 2008
- Collaboration cluster (Met Office use) – acceptance 31/12/08
- Remove NEC SX-6 – January 2009
- First/second main clusters - acceptance early April/end May 2009

  (user access ~ 6 weeks before acceptance)
- Production models operational end June 2009
- Service for NERC users – go-live 01/04/09 tbd
- Phase 2 hardware upgrade – October 2010 to June 2011
- Implementation complex owing to other related projects eg. MASS, HPC for NERC, new security architecture

© Crown copyright Met Office

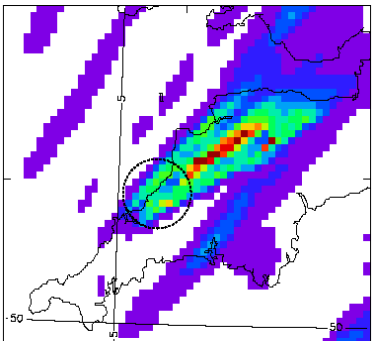
Provisional Model Upgrade Timetable 2009

- May 2009
  - System Acceptance
  - UK 1.5km Model trial
- September 2009
  - Global and NAE to 70 levels
  - Global EPS to 60km / 70levels
- December 2009
  - Global Model to 25km
  - NAE EPS to 16km / 70 levels

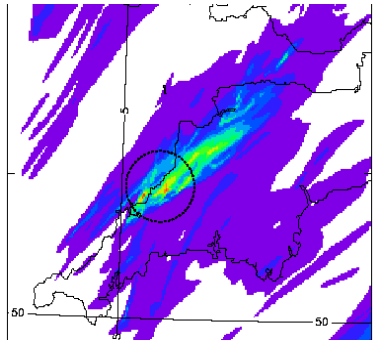
© Crown copyright Met Office

Our longer term Forecasting Strategy…

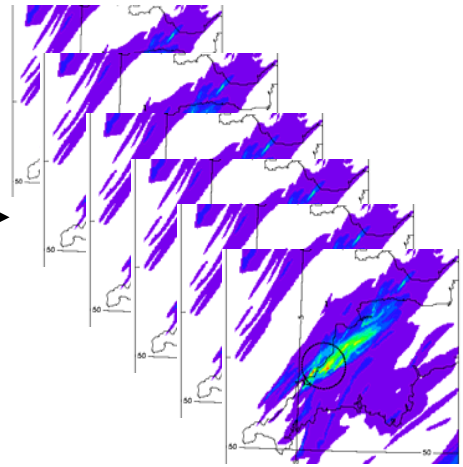
Now | 2009 | 2011/12 | 2014



4km UK



1.5km UK



1.5km UK Ensemble

- 40km Global
- 12km NAE
- 24km NAE Ensemble

- 25km Global
- 12km NAE
- 16km NAE Ensemble

- 16km Global
- (12km NAE)
- 12km NAE Ensemble

© Crown copyright Met Office

Addressing the Scalability Challenge

- Our mid-life upgrade gives a big scalability challenge
- By 2013 may need to run UM on 50-100k cores
- EPSRC funded project “Towards Generic Scalability of the UM”
- Small team focussed on performance and scalability for all Met Office applications
- IBM application consultancy
- Collaboration

© Crown copyright Met Office

Questions and answers

© Crown copyright Met Office