Tenth ECMWF Workshop on
Meteorological Operational Systems

14-18 November 2005, Reading

The Logical Data Store

Bruce Wright, John Ward & Malcolm Field

Contents

The presentation covers the following sections

Current situation and the way forward

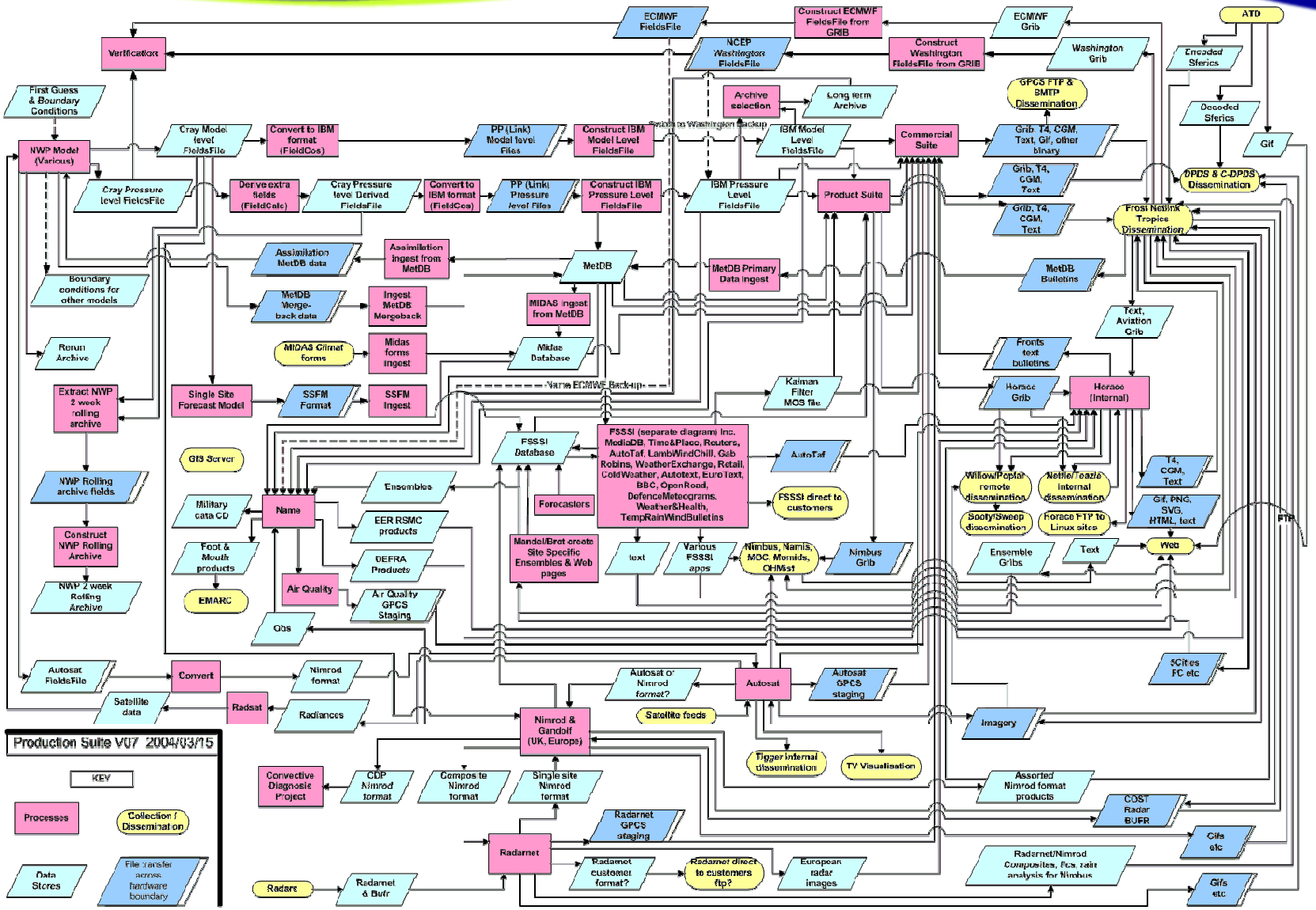
‘**Organic**’ un-governed growth has lead to:

- *complex, incoherent, undocumented IT systems*
- *striking resemblance to ‘spaghetti and meatballs’*

The ‘**information silo**’ results in:

- *inconsistent, locally processed data*
- *proliferation of data & data access mechanisms*
- *poor access to ‘enterprise’ information assets*

Current data flows



Current situation and the way forward

‘**Organic**’ un-governed growth has lead to:

- *complex, incoherent, undocumented IT systems*
- *striking resemblance to ‘spaghetti and meatballs’*

The ‘**information silo**’ results in:

- *inconsistent, locally processed data*
- *proliferation of data & data access mechanisms*
- *poor access to ‘enterprise’ information assets*

The **business drivers** for change are:

- *cost efficiency*
- *improved agility*
- *improved consistency*

The ‘**LDS**’ **(Logical Data Store)** is the means by which these issues will be addressed

Logical Data Store concept

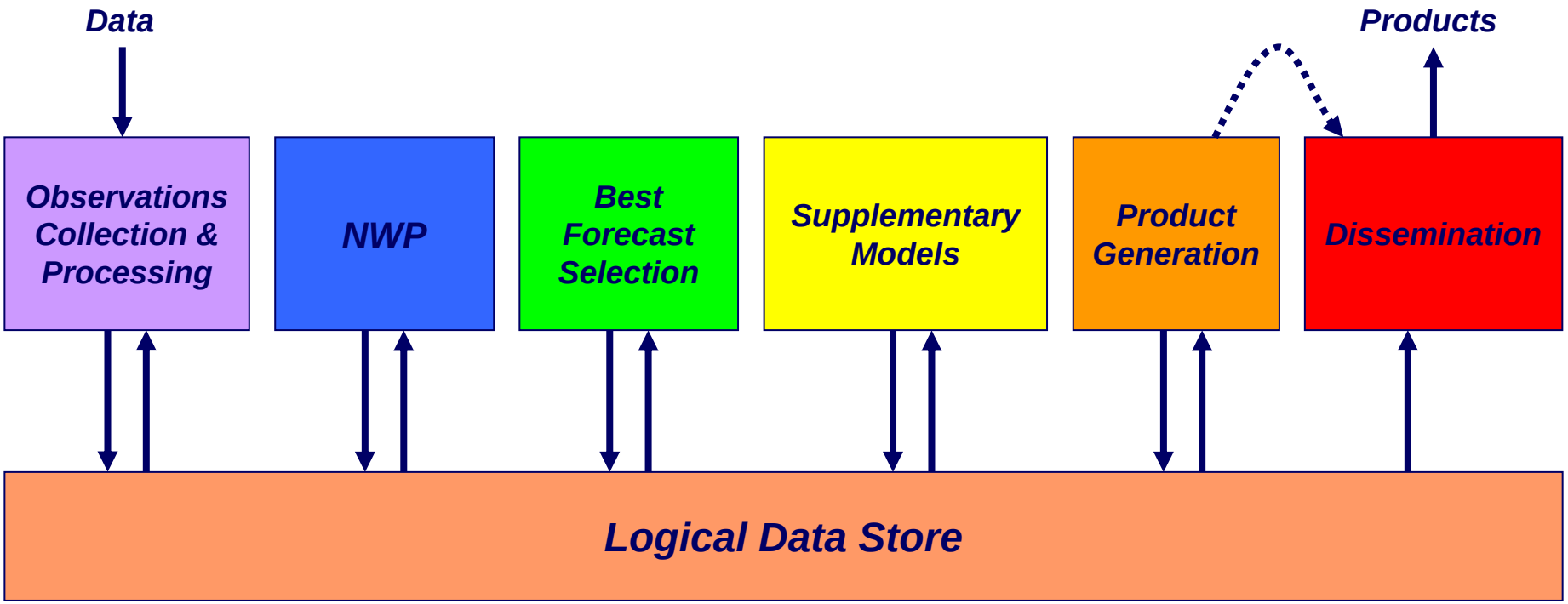
Vision

A single, logical repository for all core (shared) enterprise meteorological ‘datasets’ and products

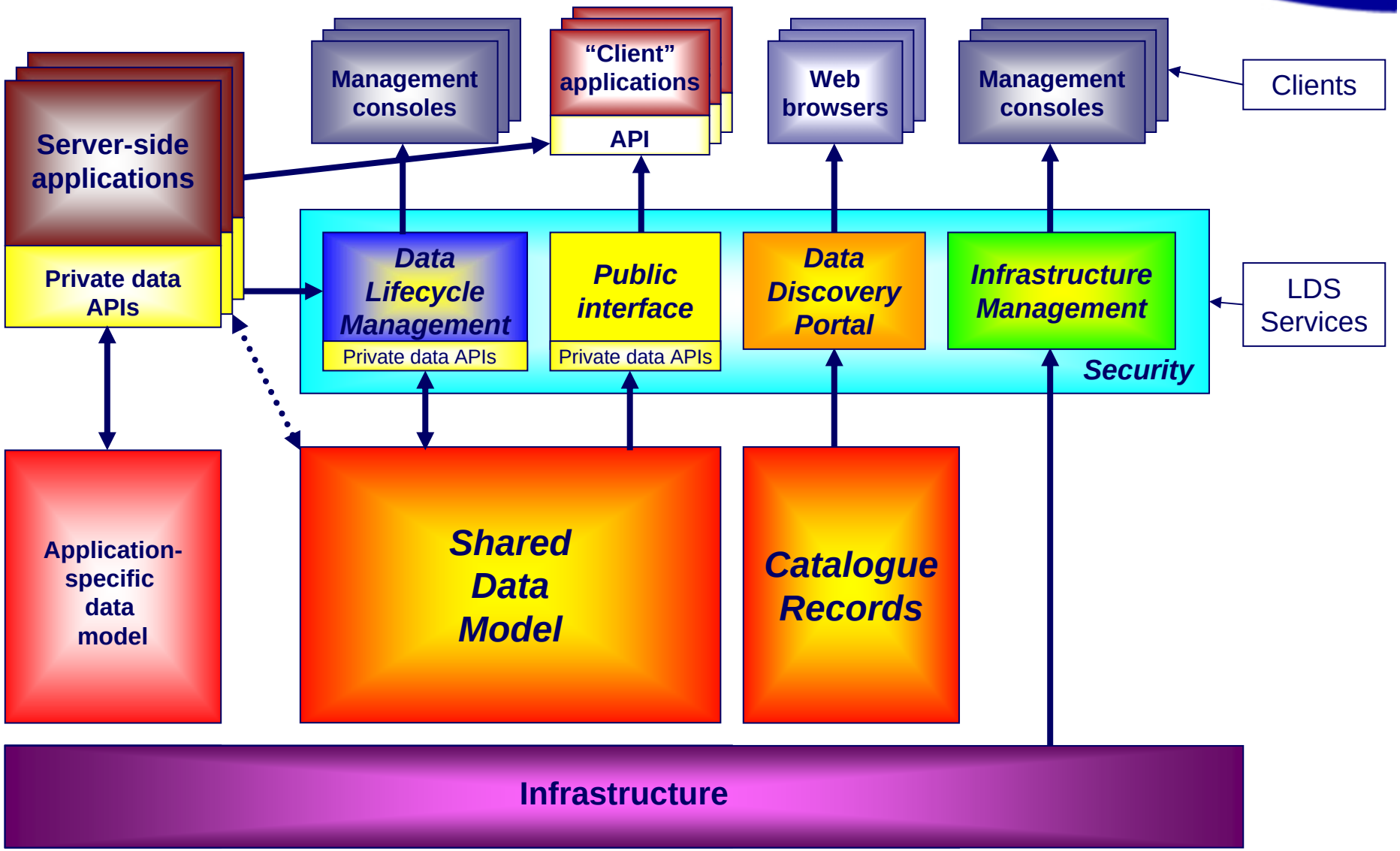
Aims

- Consistent meteorological data
- Uniquely identifiable
- ‘Spatially-enabled’ (facilitating spatial manipulation & querying)
- Accessible through a set of common interfaces
- Managed in a standard way

Information architecture



Logical Data Store



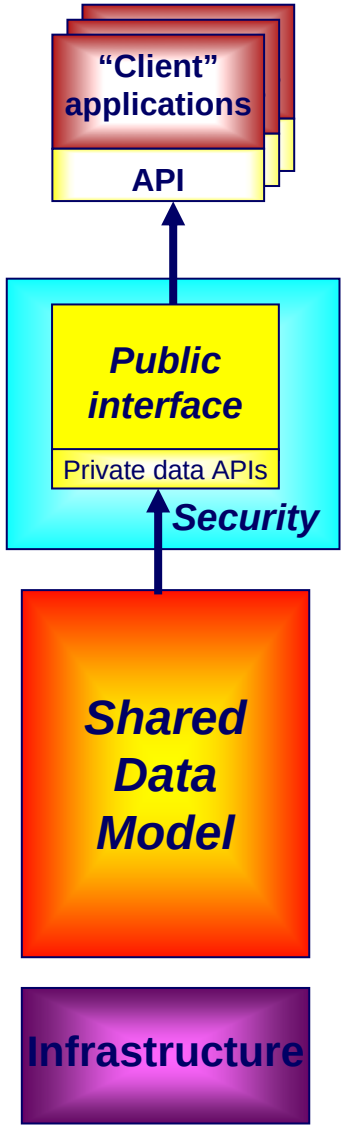
Key to Logical Data Store concept

The “Public Interface”

- Hides the complexity of:
  - Databases & archives
  - Formats & codes
  - Interfaces to different data types
- De-couples the client application from the data store
- Provides:
  - A single way to access all data in the LDS
  - Using a standard request (metadata)

Note: In the development of the LDS, we also intend to:
- Rationalise and consolidate data stores
- Take advantage of new data management technologies

LDS 'Public' Interface



…Provision of data and products

- Get resource
  - Dataset
  - Subset
  - Process
  - Format
  - Return

Standard form

Support for range of clients: Web browser, Java, C, Fortran, Open GIS client

This refers to a set of data and not a physical file (e.g. an NWP model run, all the UK climate observations)

Geospatial extent (e.g. lat-long box)
Temporal extent (e.g. validity time)
Parameter / attribute (e.g. Temp.)
Domain restriction (e.g. T > 28C)

Transform (e.g. units)
Derive (e.g. T_d from T, q, p)
Re-project (e.g. lat-long to National Grid)
Interpolate (re-grid)
Filter (sub-sample)

Change from 'Standard' to: PP/Fieldsfile, NetCDF, GRIB, BUFR, CSV

Proposed solution for LDS Public Interface

Web Services

- Use an HTTP Transport for messages (like web pages)
  - Highly interoperable
  - Clear and simple client-server interaction
- Use XML as a standard form for the request and response
  - Self-describing data
  - Implement metadata standard
  - Can use a standard schema (e.g. GML)
- Possible Issues:
  - Performance? (esp. for voluminous data)
  - Will it work? (new technology risk)

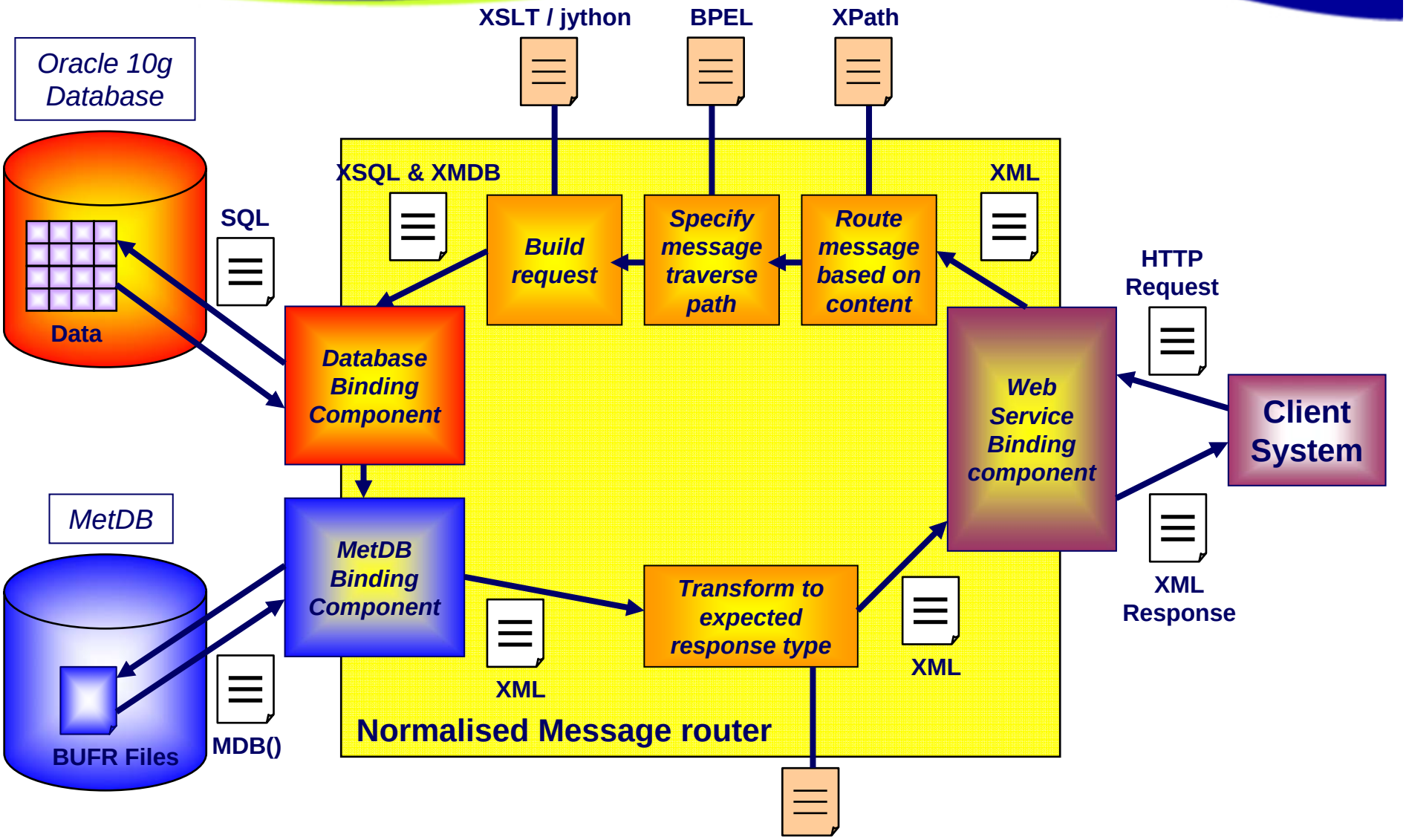
Development already completed

- Consistent use of Oracle RDBMS to hold a range of data types
- Standard Java interfaces
- Web Services with XML for data exchange
- Draft Met Office Metadata Standard building on the ISO191xx, WMO standards and CF Convention
- Standard components for deriving best climatological observation values from our archive

Investigations underway

- Lightning location database operational demonstrator
  - Store direct to Oracle Database (data and products)
  - Provide Web Services interface (probably as Web Feature Service)

Lightning location Web Service demonstrator



Investigations underway

- Lightning location database operational demonstrator
  - Store direct to Oracle Database (data and products)
  - Provide Web Services interface (probably as Web Feature Service)

- NWP use of Oracle RDBMS proof of concept
  - Pull observations directly to supercomputer from database
  - Store forecast output direct from supercomputer into database
  - Provide data access using current (Fortran-based) APIs

Investigations underway

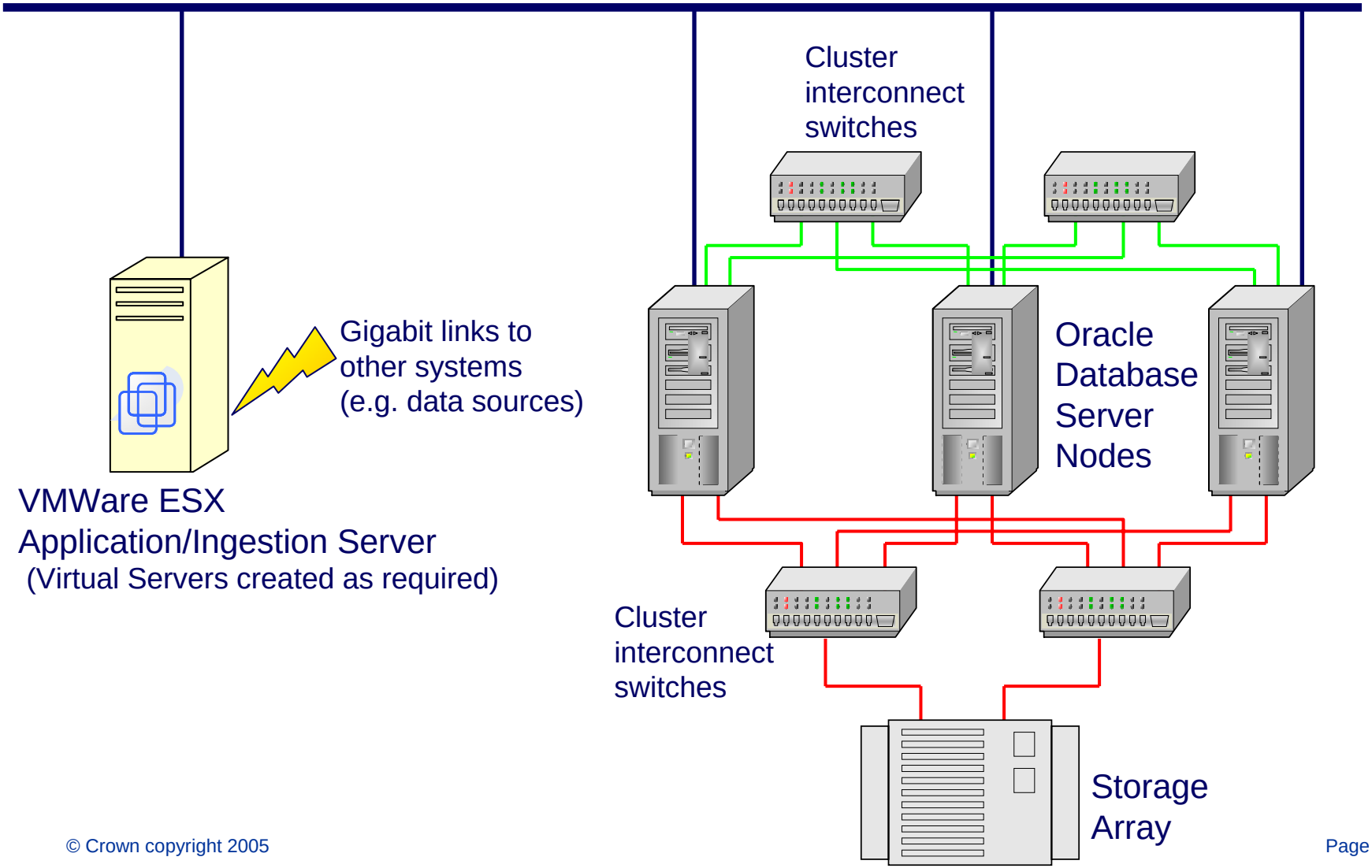
- Lightning location database operational demonstrator
  - Store direct to Oracle Database (data and products)
  - Provide Web Services interface (probably as Web Feature Service)

- NWP use of Oracle RDBMS proof of concept
  - Pull observations directly to supercomputer from database
  - Store forecast output direct from supercomputer into database
  - Provide data access using current (Fortran-based) APIs

- Oracle 10g Database cluster functionality investigations
  - 3-node Dell cluster using Oracle RAC (Real Application Cluster)

Proposed hardware architecture

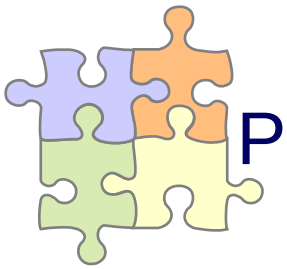
Corporate Ethernet LAN backbone (CDN)



Approaches to be adopted

NOT Big Bang

- Very high risk
- Huge amount of effort
- Long time before getting any benefit

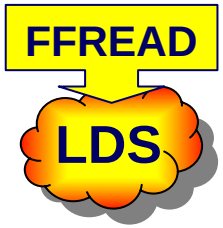


Piecemeal

- Focus of specific data types, to:
  - Prove the approach (technical solution)
  - Demonstrate end-to-end capability
  - Address existing problems
  - Provide ‘quick wins’
- But, as part of a long term Roadmap

Wrapping

- Interface to existing data stores (for the present):
  - Where migration costs are high (e.g. Archive)
  - To provide simple migration paths to use LDS



Proxy

- Provide temporary proxy interfaces
  - For those widely used
  - To allow partial/gradual data migration
  - To allow gradual application migration

Other parallel activities

SIMDAT

- EU co-funded project to promote the use of GRID technology
- Developing catalogue for managing distributed data
- Collaboration with ECMWF, DWD, MeteoFrance & EUMETSAT to deliver a ‘meteorological scenario’ for the Future Weather Information System

DEWS – Developing Environmental Web Services

- DTI co-funded collaborative project
- Using leading edge technology in real scenarios: health & marine
- Academic (BADC, ESSC) & commercial (Lost Wax, BMT, IBM) input

Questions & Answers

Email: bruce.wright@metoffice.gov.uk