

DAME: Distributed Engine Health Monitoring – Data Architecture for the Grid

Tom Jackson
University of York





Project Partners



- EPSRC Funded, £3.2 Million, 3 years, commenced Jan 2002.
- UK pilot project for e-Science (£220 million programme)
- 4 Universities:
 - University of York, Dept of Computer Science
 - University of Sheffield, Dept of Automatic Control and Systems Engineering
 - University of Oxford, Dept of Engineering Science
 - University of Leeds, School of Computing and School of Mechanical Engineering
- Industrial Partners:
 - Rolls-Royce
 - Data Systems and Solutions
 - Cybula Ltd



DAME Objectives



- Building a demonstration system as proof of concept for Grid technology in the aerospace diagnostic domain.
- Three primary Grid challenges:
 - Management of large, distributed and heterogeneous data repositories;
 - Rapid data mining and analysis of fault data;
 - Information management and data fusion for diagnosis/prognosis applications;
- Other key (commercial) issues:
 - Remote, secure access to flight data and other operational data and resources;
 - Management of distributed users and resources;
 - Quality of Service issues (and Service Level Agreements)



Business & Technical Drivers

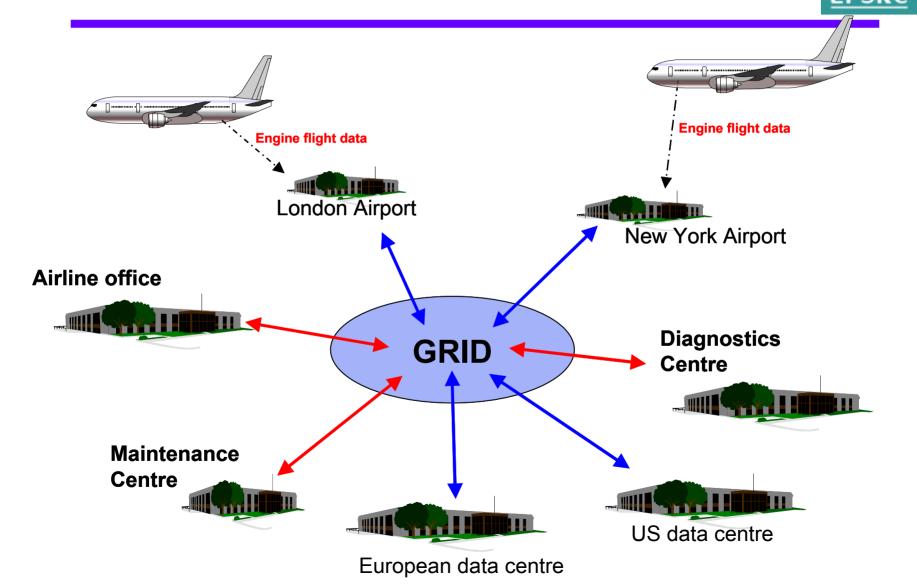


- As research partners, RR and DS&S have provided a potential business scenario that has been used to drive technical developments in several areas:
 - Virtual Organisations (VO)
 - Supporting multiple, distributed stakeholders within the business process
 - Large scale distributed data management and mining
 - How to maximise the commercial benefit of data derived from health monitoring systems
 - Secure, distributed data
 - How to protect assets and control access to data assets within a VO
 - Quality of Service
 - How to guarantee levels of service for remote services and systems.



Operational Scenario







How Big is the Data Problem?



An example scenario:

- Heathrow capped at 36 landings per hour.
- If half have 4 engines and half have 2, average aircraft carries 3 engines.
- Each engine generates around 1GB of data per flight.
- $36 \times 3 \times 1 = 108GB$ raw engine data per hour.
- Factor in the working day and the rest of the world...
- ...Terabytes and up!



Demonstrator Objectives



The DAME demonstration system provides a diagnosis workbench (portal) which brings together a suite of analysis services via Grid technology;

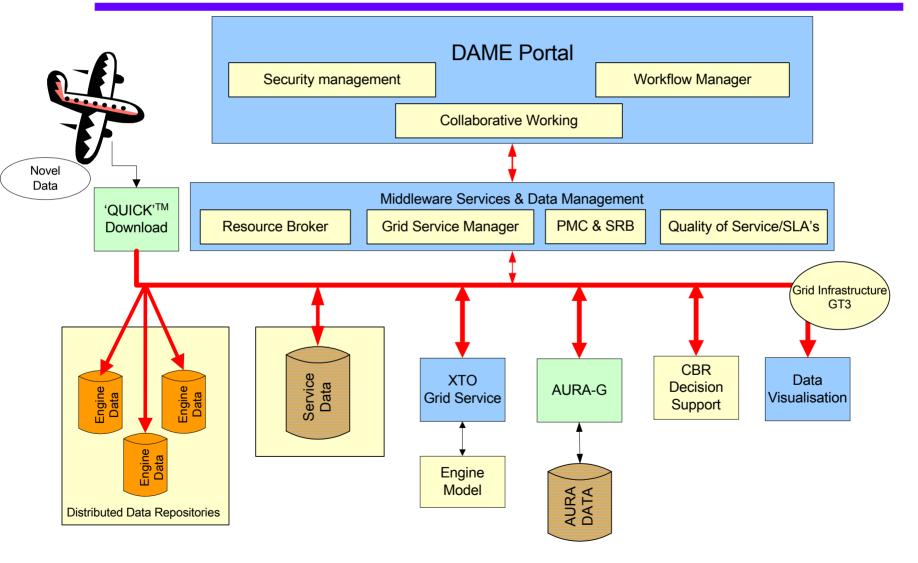
- Provides access to a range of analysis tools for the engine diagnosis process
- Will act as central control point for automated workflows
- Manages issues of distributed diagnosis team and virtual organisations
- Manages issues of security and user roles.



DAME Functional Overview





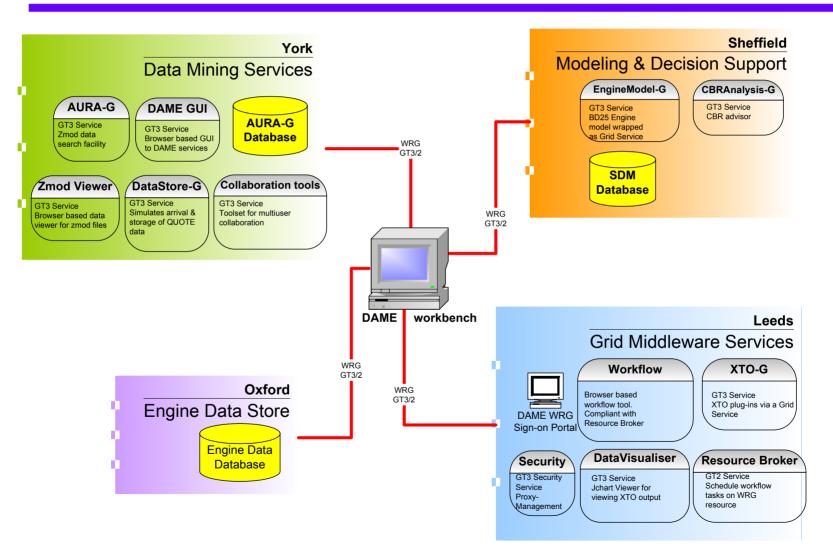




White Rose Grid Distribution









Data Mining Activities



- Objectives have been to develop a data mining service to search fleet archives of QUICK engine data within operational time constraints.
 - To support diagnosis and prognosis activities
 - To support long term fleet predictive maintenance
 - Business assumption is that data is archived remotely
- Two tools have been developed:
 - AURA-G: Grid enabled signal search engine;
 - Signal Data Explorer: Interactive search GUI for signal data
- These tools are generic to any pattern matching domain, but have been tuned to RR & DS&S requirements for DAME.



AURA-G Service & PMC



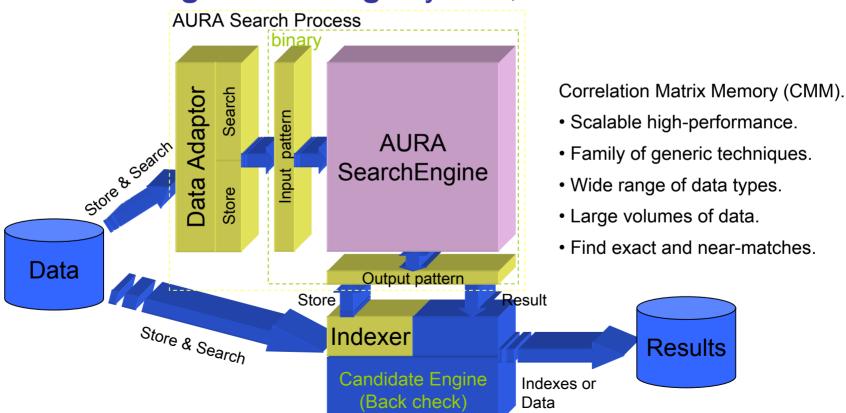
- AURA is a highly scalable, binary associate memory technology, developed for advanced pattern matching applications.
- DAME has developed a distributed Grid enabled version of this technology; matches commercial challenges of data growth and data distribution.
- Provides capability to mine huge volumes (Terabytes) of complex signal data with sub second response times.
- AURA-G is managed by a DAME middleware service called Pattern Match Controller:
 - Provides mechanisms for distributing and managing search queries
- A virtual data archiving system, Storage Request Broker(SRB) provides the middleware for data cataloguing and archiving.



AURA Integration & Deployment



 AURA is a collection of processes; data adaptors, search-engine and back-check. It wraps around an existing data storage system;

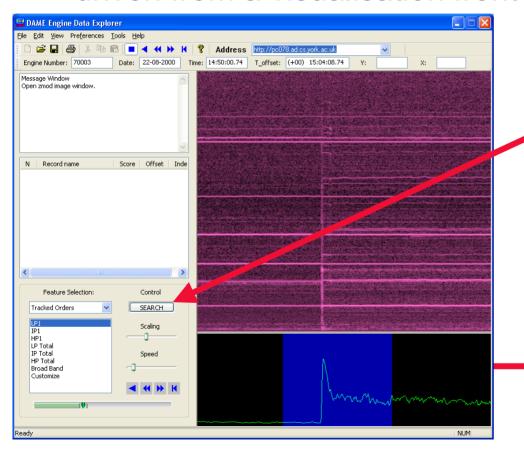




Distributed Pattern Search



 Complex time-series pattern matching process driven from a visualisation front end.



A search is launched across the fleet data archives using the Grid enabled AURA search engine

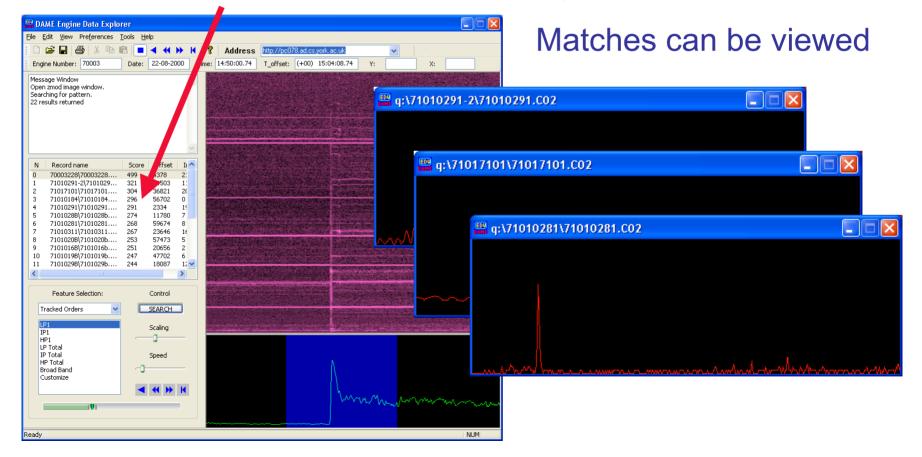
Engineer selects region of interest



Distributed Pattern Search, 2



 All matched pattern records are retrieved from the fleet archives and ranked according to similarity

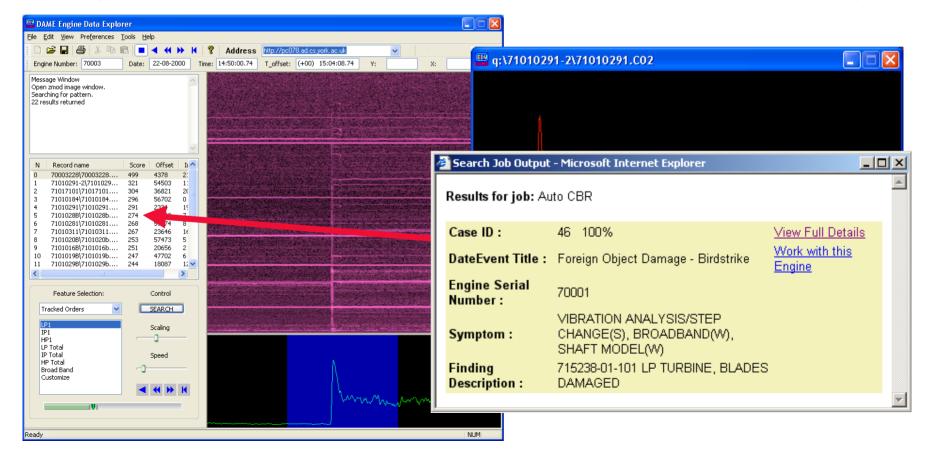




Distributed Pattern Search, 3



 The case history associated with any meaningful signature can be recalled from the fleet maintenance logs



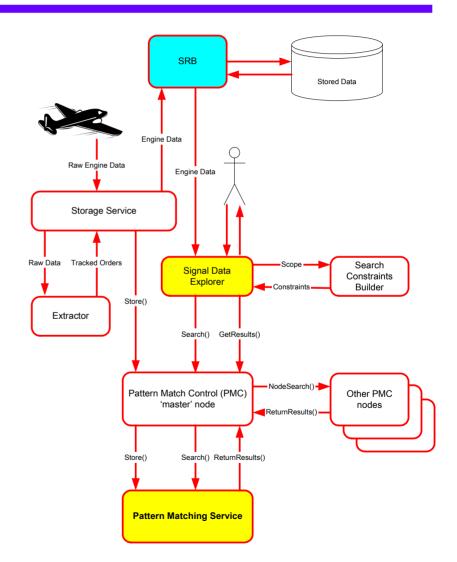


Pattern match Control





- The SDE interfaces to the PMC middleware
- GT3.0 enabled service components;
- PMC provides:
 - Distributed search
 - Interface to data archive system (SRB or other)
 - Scalablity
 - Robustness





PMC cont



 PMC architecture has been developed on business premise of remote data.

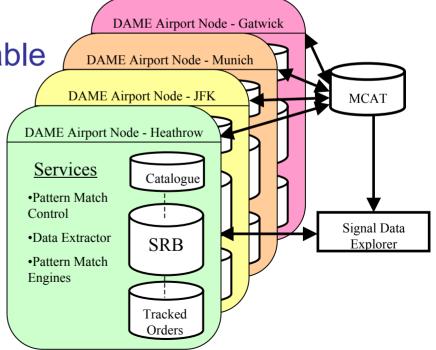
E.g. Airports act as data repositories for Engine

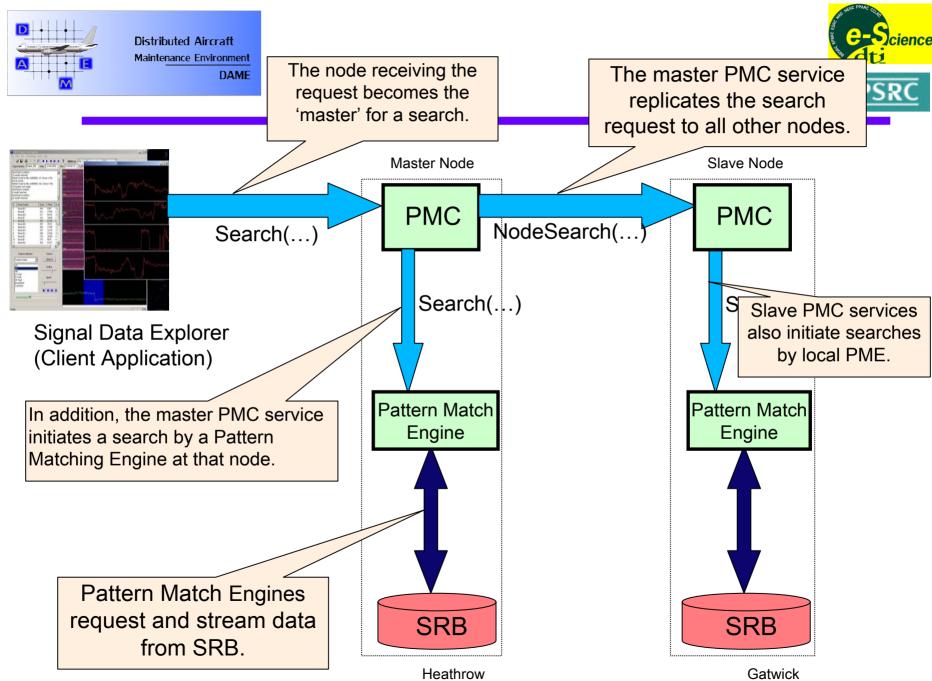
health data

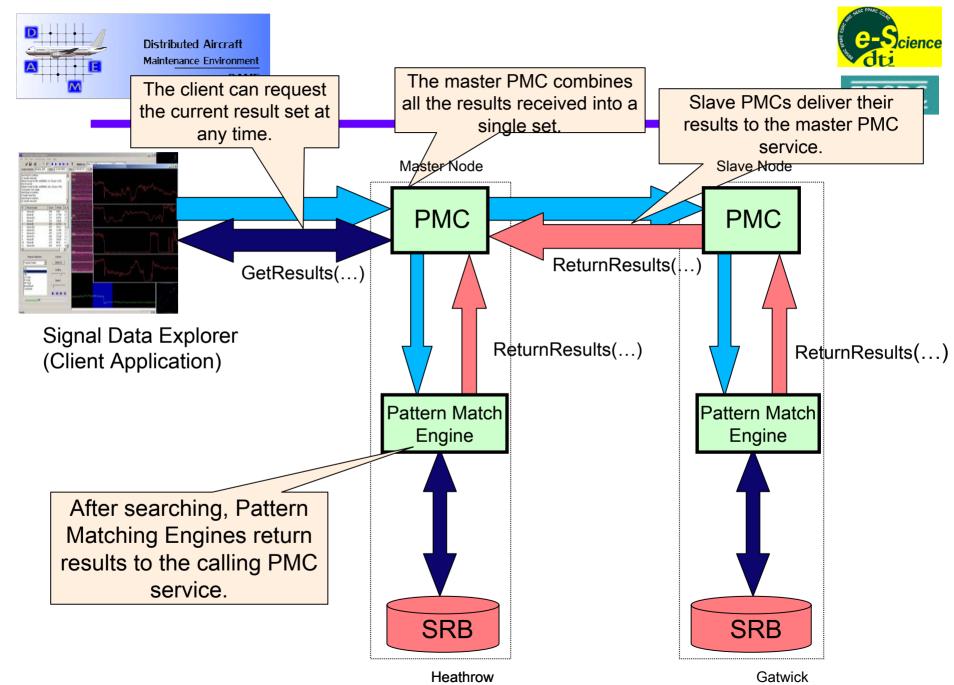
SRB provides hugely scalable

virtual catalogue

& index system









Decision Support tools



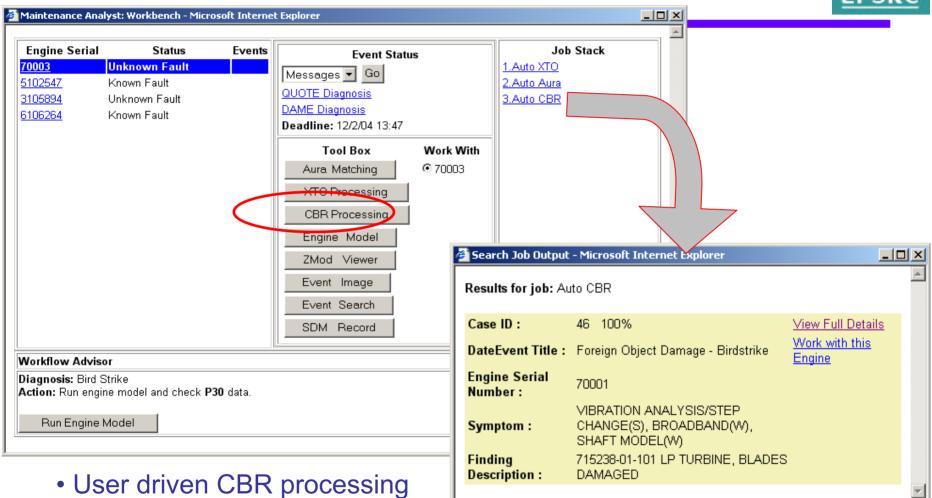
- Objectives are to maximise use of available data assets to support business processes:
 - Diagnosis & Prognosis
 - Fleet predictive maintenance
 - Workflow management
 - Domain knowledge capture
- Achieved through integration of diverse data assets through Grid enabled Case Based Reasoning tools.
- Grid provides standardised interface to legacy systems and data servers.



CBR Maintenance Analysis in DAME







 Automated CBR processing via Workflow Manager



Summary



- DAME is demonstrating the potential of Grid-based diagnostics for health-monitoring applications;
- Grid paradigm offers:
 - Scalable solutions for management of distributed data
 - Secure access to services and data
 - Role based access and virtual collaboration on fault data
- DAME has shown how data growth problems could potentially be addressed within scope of distributed data assets.