Enhancing User Access to Oceanographic Data Through Commercial Cloud Services

Roger Proctor, Integrated Marine Observing System (Australia), roger.proctor@utas.edu.au
Peter Blain, Integrated Marine Observing System (Australia), peter.blain@utas.edu.au
Dan Fruehauf, Integrated Marine Observing System (Australia), malkodan@gmail.com

The Integrated Marine Observing System (IMOS) is a national project funded by the Australian government established in 2007 to deliver ocean observations to the marine and climate science community. Its mission is to undertake systematic and sustained observations and to turn them into data, products and analyses that can be freely used and reused for broad societal benefits.

Historically IMOS information infrastructure has relied on national e-research infrastructure cloud services such as those supported by the University of Melbourne and the University of Tasmania. These services, being research grade, suffer from unplanned outages, unexpected loss of service and, at times, data corruption. As IMOS has matured as an observing system, now in its 10th year, and community exposure to IMOS data has increased, expectation on the system's availability and reliability has also increased. IMOS is now seen as delivering 'operational' information. In responding to this expectation, IMOS has relocated its services to the commercial cloud service Amazon Web Services (AWS). The relocation has significantly increased IMOS' measured availability.
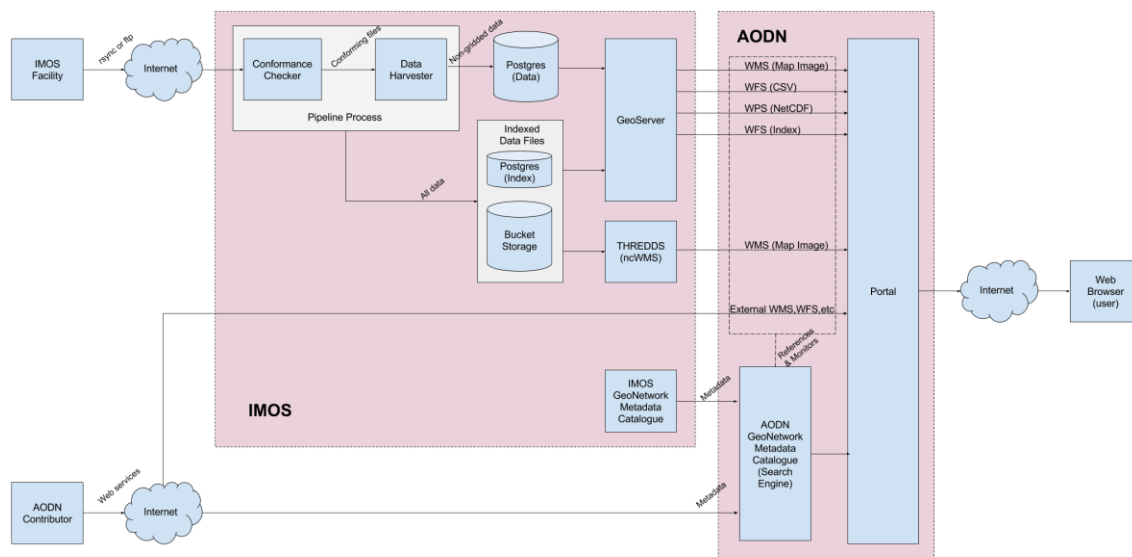

Figure 1 - IMOS infrastructure at AWS

Using AWS has enabled IMOS to improve the system architecture utilizing more advanced features like object storage (S3 - Simple Storage Service) as opposed to a traditional file system, Figure 1. S3 has virtually unlimited affordable storage and supports redundancy (stores copies in multiple facilities and devices) and versioning (keeps tracks of a single object's multiple versions). Redesigning the IMOS system architecture around S3 has improved data availability and resilience while protecting against human errors in data handling and providing a disaster recovery plan.