

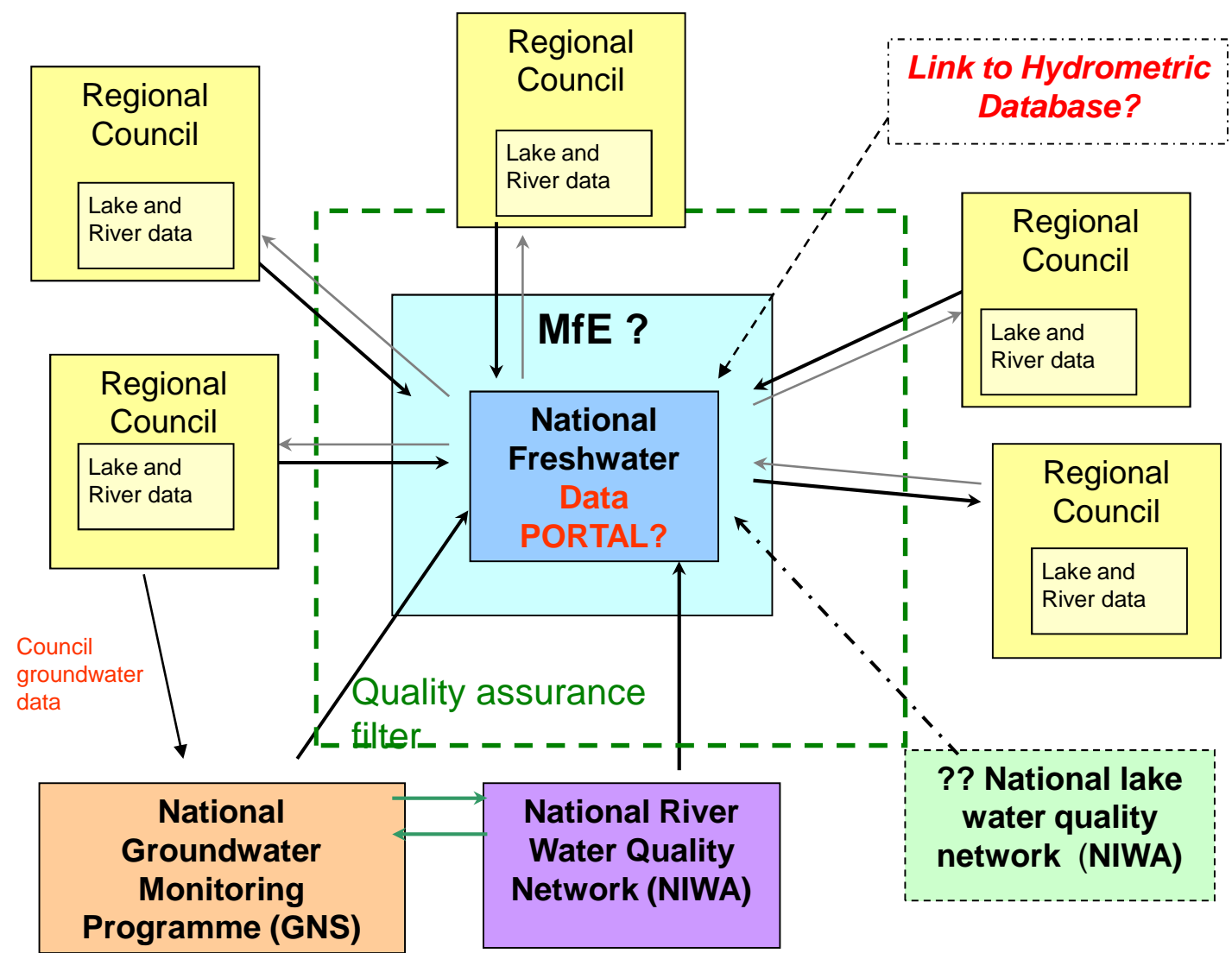
# Thoughts on a National Freshwater Quality Data Base (NFWQDB)

Clive Howard-Williams

National Institute of Water and Atmospheric Research  
Ltd.



# Concept: **Federated database system** (Modified from last meeting)



**Those who generate the data, "own" the data**

# Benefits of a NFWQDB

- Faster, easier more timely development and validation of WQ guidelines and standards (nutrient caps)
- Inter-regional comparisons, cross regional analyses
- Transference of knowledge and understanding among regions
- Reference sites in adjacent regions
- Benefits to all data providers in QA/QC, data reduction, technical support and advice (ie optimisation etc)
- Faster, easier more timely national scale analyses and reporting
- Faster, easier more timely international reporting and comparisons

## Experience shows the following have to be accounted for in federating data:

- Concurrent flow data with water quality sampling sites
- Sampling frequency and duration for trend analyses
- Representation of environmental drivers (ie REC classes, land-use classes etc) for state analyses
- Lack of “reference” sites
- Variation amongst data providers in measured analytes
- Differences in field methods (and competencies and training) of field staff
- Differences in laboratory methods, detection limits, lab inter-comparisons for QA/QC
- Multiple (nested) sites in catchments (needed for catchment studies but potential problem in national reporting)

**These can all be addressed**

# Where to from here?

## Four ways to generate a national database:

1. Opportunistic: Collate existing data, groom and make available on the web (relatively simple and cost effective); intermittent; Collation and grooming step **has been done before.**
2. Distributed data base (as per diagram) gives firm sense of ownership but needs DB architecture and could be IT intensive, needs standards, hosting issues?.
3. Conduct 1,2 above in parallel
4. Develop true national database ('single data container') regularly updated - expensive to develop and maintain and an overhead on Providers. **Need to consider this only in the longer term.**

# Federated data base system

A “federated data base” model has data accessible through a single and multiple portals (ie through each Data Provider) recognising that: 1. Not all current Providers will want to contribute; 2. National data bases are complex and will be expensive. **Provider buy-in required**

- Recognise that Data Providers have their own front-ends for data delivery and then identify the data that will be available for a national database. **Individual Providers report back**
- Work through the issues (see earlier slide) related to consistency. **Needs to be done by inter-Provider agreement**
- Identify criteria for inclusion of sites and data (time series, commitment etc). **Needs to be done by inter-Provider agreement**
- Decide on a format (architecture) into which the data can be collated and portal host arrangement. **Provider agreement and agreement on funding**

**Look forward to progress – MfE should continue to facilitate**