Software Module Clustering

Exhauro Project

Problem Definition

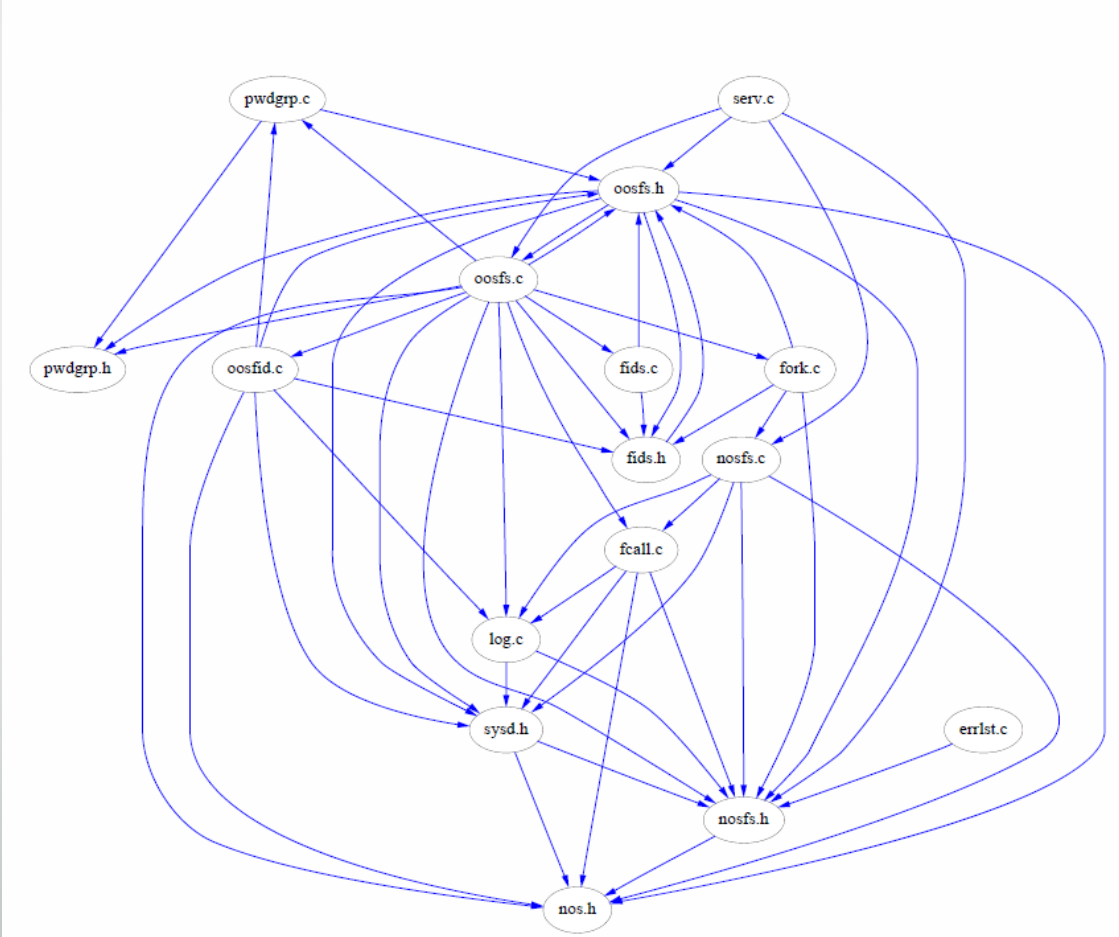


Figure 2. Module Dependency Graph of the File System

Problem Definition

$$MF_k = \begin{cases} 0 & si\ i = 0 \\ \frac{i}{i + \frac{1}{2}j} & si\ i > 0 \end{cases}$$

i: intra-connectivity
j: inter-connectivity

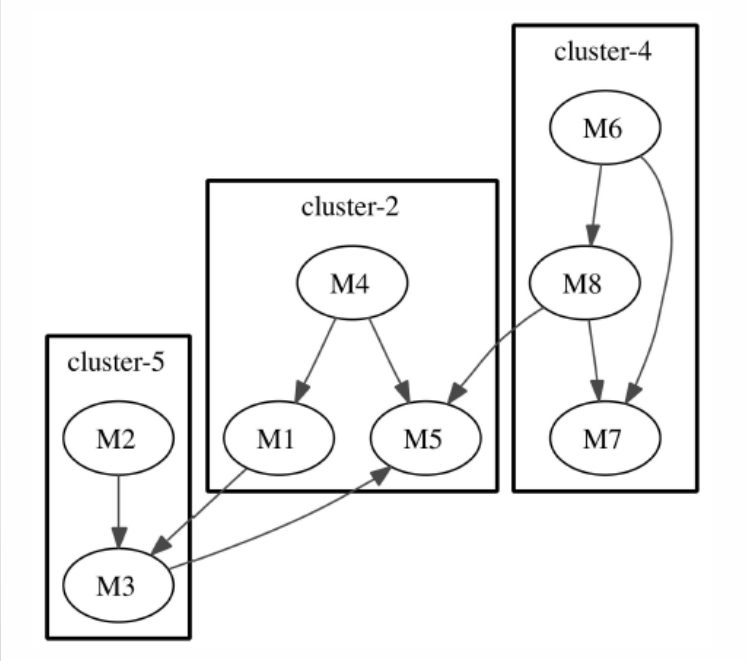
Modularization factor

$$MQ = \sum_{k=1}^{n} MF_k$$

Modularization quality

Praditwong et al. (2011)

Example



$$MF_1 = MF_3 = MF_6 = MF_7 = MF_8 = 0$$

$$MF_5 = \frac{1}{1 + \frac{1}{2} \cdot 2} = \frac{1}{2}$$

$$MF_2 = \frac{2}{2 + \frac{1}{2} \cdot 3} = \frac{4}{7}$$

$$MF_4 = \frac{3}{3 + \frac{1}{2} \cdot 1} = \frac{6}{7}$$

$$MQ = \frac{1}{2} + \frac{4}{7} + \frac{6}{7} = \frac{27}{14} = 1.928571\ldots$$

Praditwong et al. (2011)

Multi-objective Formulation

Maximizing Cluster Approach

- the sum of intra-edges of all clusters (maximizing),
- the sum of inter-edges of all clusters (minimizing),
- the number of clusters (maximizing),
- MQ (maximizing),
- the number of isolated clusters (minimizing).

Equal-size Cluster Approach

- the sum of intra-edges of all clusters (maximizing),
- the sum of inter-edges of all clusters (minimizing),
- the number of clusters (maximizing),
- MQ (maximizing),
- the difference between the maximum and minimum number of modules in a cluster (minimizing).

Praditwong et al. (2011)

State of The Art