ABSTRACT

Dual Heap Overlay Resource Discovery Protocol for Mobile Grid proposed in this work locates mobile nodes in a dynamic mobile grid. The protocol resolves exact / partial match and range queries specifying multi-attribute-based search for mobile nodes. Organizing dynamic nodes with frequently changing attributes to facilitate query resolution is not trivial. Issues specific to centralized approaches and P2P approaches employing single or multiple Distributed Hash Tables (DHTs) to discovery are overcome by non-DHT overlays used by the protocol.

The protocol uses two non-DHT min-heap overlays namely, the sum heap and the attribute heap. A mobile grid built on a cellular network is investigated. A central node closer to the Base Station (BS) receives periodically the state information on mobile nodes and maintains the centralized sum heap. Each node in the sum heap serves as an index to all nodes having same attribute sum. Attribute heap is distributed among the peers based on one of the attributes.

Upon receipt of the query the central node computes the sum of the attributes in the query. Nodes are disqualified and not considered for the query request whose attribute sum exceeds the sum as per the sum heap. Nodes if any, whose sum matches, are examined as per the order in the attribute heap. If no examined node satisfies all attributes individually, nodes with higher sum are
compared resulting in ‘slack’. The query gets propagated until a node is located or the search fails after exhaustively checking all eligible nodes. Thus the protocol has one of three outputs: QWS - success without slack (exact match), QS- success with slack and QF - failure.

The number of disqualified nodes and hence the resulting rejection rate determines the efficiency of the protocol. This can be estimated by knowing how many nodes share same attribute sum. Excluding replication, the theoretical maximum of number of resources with k attributes contributing same attribute sum has been estimated recursively. This helps in determining the efficiency of the protocol.

The extended protocol resolves range queries where an approach to reduce the slack value is proposed resulting in an optimal solution. A linear programming problem is formulated to achieve the same.

The protocol is designed to handle node mobility and disconnections prevalent in mobile grid. When nodes voluntarily disconnect, the protocol identifies suitable destination node for the migrated process using a max-heap. Max-heap organizes nodes based on their spare computational capacity. The node with highest spare capacity is trivially extracted and the search for destination thus is performed with constant time complexity.

Mobile nodes being resource-constrained, a factor of safety is introduced to guarantee that there is no degradation in the performance below accepted levels when such migration occurs. During migration dynamic load balancing is done. An algorithm is devised to give option to a mobile node for accepting or rejecting the
migrated process. A dedicated bandwidth solution is proposed to export / import the process-state in Long-Term-Evolution-Network (LTE) based mobile grid.

To handle multiple attributes the protocol is a novel approach; it does not have multiple indices and hence reduced maintenance overheads; it follows a simple and efficient structure for the overlay; in literature no other work attempts to reduce the search set as attempted in the protocol. The protocol is scalable. Its search time is stable as the number of attributes increase. The routing table being very small is easy to maintain. Features peculiar to a mobile environment like mobility, dynamism and disconnections, bandwidth changes and device constraints inherent to a mobile grid are taken care by the protocol efficiently.

There is a scope for further research in the domain of resource discovery in mobile grid. The local scheduling in a cell performed by the protocol may be extended to perform high-level scheduling involving a cluster of cells. It can also consider other aspects of resource management system like fault tolerance, co-allocation, interoperability among different VOs, etc. Further, to host a migrated process the protocol tests only the ability of a node to support computation. A combination of computing and storage abilities can be considered while selecting a destination. Moreover, a task can be migrated to a node belonging to a different cellular region by extending the overlay to cover the entire grid. This can bring in favorable load balancing conditions enhancing the response time of the grid application.