Controlling the Coverage of Grid Information Dissemination Protocols

Bhavin Gandhi, Sameer Tilak, Michael J. Lewis and Nael B. Abu-Ghazaleh
Department of Computer Science
State University of New York, Binghamton NY 13902
{bgandhi, sameer, mlewis, nael}@cs.binghamton.edu

Abstract

Grid information dissemination protocols distribute information about the dynamic state of computational resources throughout interconnected wide area Grids. Performance metrics for these protocols include the overhead of information packets, and the accuracy of the information at the time it is used to schedule applications. Our previous work advocated non-uniform protocols to keep dissemination local to the information source, as a method of keeping overhead manageable while achieving adequate freshness and accuracy. This paper considers the problem of providing better control over the dissemination of information and influencing the "coverage footprint" that defines where the information reaches within the Grid. The paper describes work that investigates the coverage characteristics of existing protocols and refines and combines them into hybrid protocols that are more controllable. We consider this work to be a necessary step toward adaptive dissemination protocols that will be able to react to the state of Grid resources to change dynamically how and where information is disseminated. This in turn will increase the effectiveness of Grid schedulers under various load levels and distributions.[1]

1 Introduction

Computational Grids will increase by orders of magnitude the amount of interconnected computing resources readily available for use by scientists and programmers with computationally demanding applications. In order to effectively schedule such applications, Grid schedulers require fresh and accurate information about Grid resources. The problem of tracking dynamic resource states is extremely challenging because of the large number of resources, their dynamic state, and the many candidate scheduling points. In our model, resources proactively disseminate state information periodically.

The size of computational Grids makes many possible dissemination approaches impractical. Optimized multicast reduces overhead, but requires Grid nodes to maintain topology information to build and track the multicast tree [1]. Flooding provides a simple localized solution, but has an extremely high overhead. Gossiping protocols [9] and random walks [6] may result in uncontrollable loss of coverage.

Ideally, schedulers would have just enough information to schedule applications and to effectively balance the load. In previous work [4], we advocated non-uniform information dissemination, to capitalize on a general observation that information about Grid resources is of greater interest to the nearby client population. This results in lightweight information dissemination that maintains the resource state information with high accuracy where it will be most beneficial.

To meet diverse resource-specific requirements, dissemination protocols should be dynamic and configurable. To control dissemination, ideally Grid designers would simply manipulate forwarding probabilities. Unfortunately, detailed analysis reveals that the coverage patterns of our protocols do not follow intuition, rather they exhibit unpredictable coverage.

Our goal then, is to develop protocols with more controllable coverage to enable resource providers to determine the footprint of dissemination. Controlling the dissemination footprint is important for several reasons. First, different providers will have different requirements for dissemination. Second, we view controllability as a necessary prerequisite for adaptive dissemination protocols. Only adaptive protocols, which dynamically alter the aggressiveness with which information is disseminated, will be able to react to changing Grid load characteristics. Controllability in and of itself may not be sufficient, but we feel that it is certainly a necessary attribute for building adaptive protocols.

The first contribution of this paper is to show that the effect of the forwarding probabilities on the network coverage achieved by basic protocols [4] is complex and non-linear. We show that a small range exist where small changes of forwarding probabilities result in large changes in the coverage patterns. The second contribution of this paper is to propose

hybrid protocols that synthesize desired coverage patterns by changing the forwarding probabilities at different hops from the source to best match the desired coverage.

Our protocols are similar to gossiping (e.g., [5]) which attempts to cover an entire network with reduced overhead via probabilistic forwarding. However, our work adopts the idea of non-uniform information dissemination [10] [4] to cover local nodes more aggressively. Butt et. al. proposed "flocking" [2] of Condor pools, combined with the Pastry [8] peer-to-peer overlay for scalable discovery of resources. Pastry is locality-aware, so Condor ends up mapping resource requests to nearby Condor pools that can service them, keeping application deployment time down. This approach achieves non-uniformity in a different way than the randomized protocols that we characterize in Section 2; however, Condor shares our goal of servicing requests locally. Maheswaran et. al. [7] also share the approach of associating higher value with nearby information. They introduce the notion of "Grid potential," which weighs a Grid resource's capability with its distance from the application "launch point". The authors propose and study the tradeoffs between three different protocols, but do not consider controllability as a metric.

2 Coverage Analysis of Existing Protocols

In this section, we describe the basic Biased and Unbiased protocols [4] and analyze their coverage patterns as a function of the forwarding probability P. Coverage C_d is defined as the average percentage of nodes at distance d from a source that receive a disseminated packet.

Unbiased (Gossiping) Protocol: In this protocol, the forwarding probability is constant at all nodes, no matter their distance from the source. Coverage analysis indicates that for large values of P, the disseminated information reaches almost all of the nodes in the overlay (Figure 1 shows this effect for a random topology of 1000 nodes).[2] The coverage is not responsive to the forwarding probabilities in this range; this is consistent with the known bimodal effect of gossiping protocols [5]. For $P \geq 0.65$, the dissemination covers more than 80% of the nodes. Varying the size and topology of the network shows similar trends: the coverage is affected by the connectivity of the network, but in general, high forwarding probabilities do not result in non-uniform coverage.

In Figure 2 we observe that for lower values of P, the dissemination footprint does exhibit non-uniform coverage; in this range, the coverage is highly sensitive to P. For $P \leq 0.52$, there is high coverage ($> 90\%$) for the initial $k = 4$ hops, and thereafter the coverage drops gradually with each hop. For $P \geq 0.58$, the coverage is high for the initial $k = 7$ hops and then it drops from 85% (at the 8th hop) to 20% (at

---

the 26th hop). We argue that the "footprints" obtained with these values for P are static; the resource provider does not have the flexibility to define the coverage pattern. Therefore, if the resource dissemination model requires higher coverage for k' hops (where $k' > k$), low values of P cannot achieve this. And as described earlier, increasing P to try to increase k results in uniform (full) coverage.
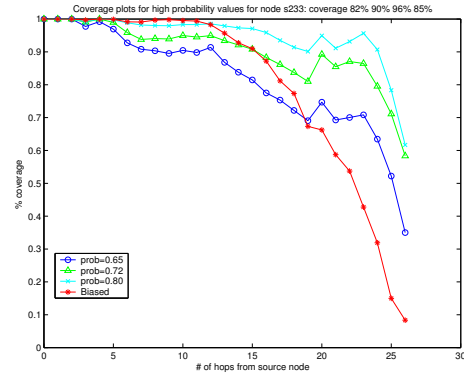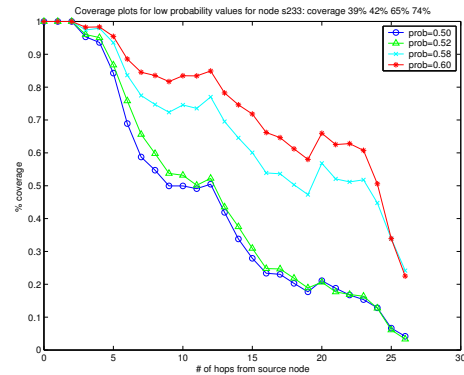

Figure 1. Coverage (High Prob.)

Figure 2. Coverage (Low Prob.)

Biased Protocol: This protocol uses the following function at each hop to determine the forwarding probability: $P_i = (n - d_i)/n$, where n is some constant (typically the network diameter) and $d_i <= n$ indicates node i's distance from the source along the path through which the packet was received. Because of the way the Biased function is set up, for relatively large values of n, the forwarding probabilities at the initial k hops are high, and therefore the coverage in these initial hops is uniform. After these k hops, the source cannot control the coverage pattern.

The effect can be seen in Figure 1 where coverage is almost 100% for initial $k = 12$ hops. After that, the coverage drops quickly. Thus, the Biased protocol is also not source

---

[2]We use a custom message-level simulator for all studies in this paper; the simulator abstracts away lower level network details in favor of efficiency [3].

268

controllable, since given an estimate of high coverage in the initial k' hops (where $k' \neq k$), the source can not configure the protocol to obtain the desired high coverage in k' hops.

3 Addressing Protocol Limitations

The Unbiased and Biased protocols define static footprints; the resource provider cannot control dissemination coverage. In this section, we address that limitation by proposing a Hybrid protocol that uses different values of P at different distances from the source to best match the desired coverage.
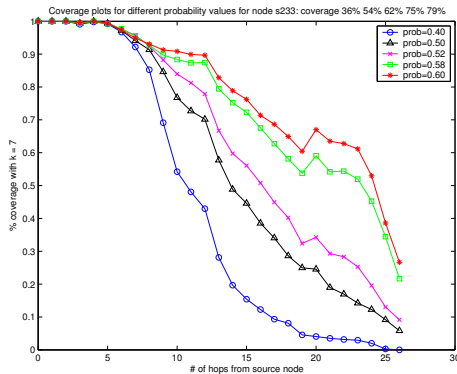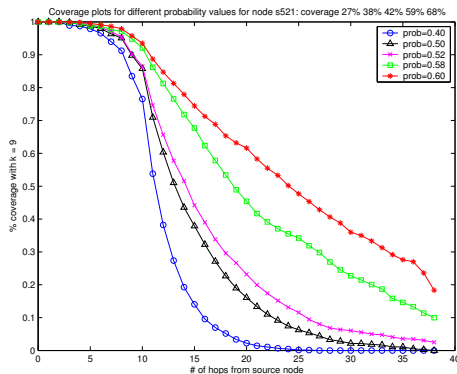

Figure 3. Random Toplogy: Coverage

Figure 4. Mesh Topology: Coverage

Figure 3 shows the coverage characteristics of a coarse Hybrid protocol on the random topology previously used in Section 2. The resource provider (node s233 in the overlay) configures the Hybrid protocol with $P = 0.8$ for the initial $k = 7$ hops. For $k > 7$ hops, we use different values of $P < 0.65$ to observe the coverage patterns. We note that our selection of $P < 0.65$ is guided by the fact that values of P more than 0.65 can result in uniform coverage as discussed in the previous section. A comparison of this graph with the

coverage graphs of pure non-uniform protocols presented in Figures 1 and 2 suggests that the Hybrid protocol offers a high degree of freedom to the resource provider in selecting the coverage footprint. This protocol avoids the two extremes of coverage, full coverage and uncontrolled local coverage. We argue that none of the pure non-uniform protocols could have provided the types of footprints that the Hybrid protocol offers. With lower values of P after k hops, the drop in coverage is steep as can be seen from the curve of 0.40. If the coverage requirements at intermediate nodes are high, higher values of P could be used. We show the coverage of the Hybrid protocol on a 1000 node 20-by-50 mesh topology in Figure 4. In this case, the resource provider (node s521 in the overlay) configures the protocol with $P = 0.8$ for $k = 9$ hops. The results here show a similar trend to the random topology. These graphs provide insight into the possible mechanisms for adaptive dissemination.

In a realistic application, the information distribution requirements might be more sophisticated then the two phase configuration studied above.

To motivate the problem, consider the following three different scenarios. (i) A supercomputing center located on the east coast might decide that it needs to serve all users on the east coast, 60% of users in the midwest region, 40% of users in the southern region, and 20% of users on the west coast (labelled "Supercomputer" in Figure 5(a)). (ii) A 32 node cluster owned by a university Computer Science department might wish to facilitate scientific research and encourage interdisciplinary collaboration by making the cluster available across the campus, and at all other universities localted within 200 miles (labelled "Cluster" in Figure 5(b)). (iii) Finally, a company providing a *computation on demand* service through its high priced cluster of 256 nodes may wish to maximize profit by attracting as many users as possible. After an indepth analysis of cost-benefit ratio, the company executives decide to advertise aggressively in four neighboring states, to 80% of users in the next six states, and to just 10% of users from the remaining states. (This scenario is labelled "Organization" in Figure 5(c)).

In each of the above figures, the dashed line represents the network-centric footprint that the application desires, and the solid line represents a curve obtained using a Hybrid protocol. Figure 5 demonstrates that the Hybrid protocol can be configured to closely match varied coverage requirements.

The flexibility offered by the Hybrid protocol makes it possible to configure the degree of localization. This leads us to consider two possible dissemination strategies. One is where we use aggressive forwarding for small number of initial k hops and use high probability after k hops. The other is to use aggressive forwarding for high number of initial k hops and use low probability after k hops. In general, we observe that keeping the value of k small and using high forwarding probabilities at later hops results in almost the same

(a) Supercomputer curve.   **(b)** Cluster curve.   **(c)** Organization curve.
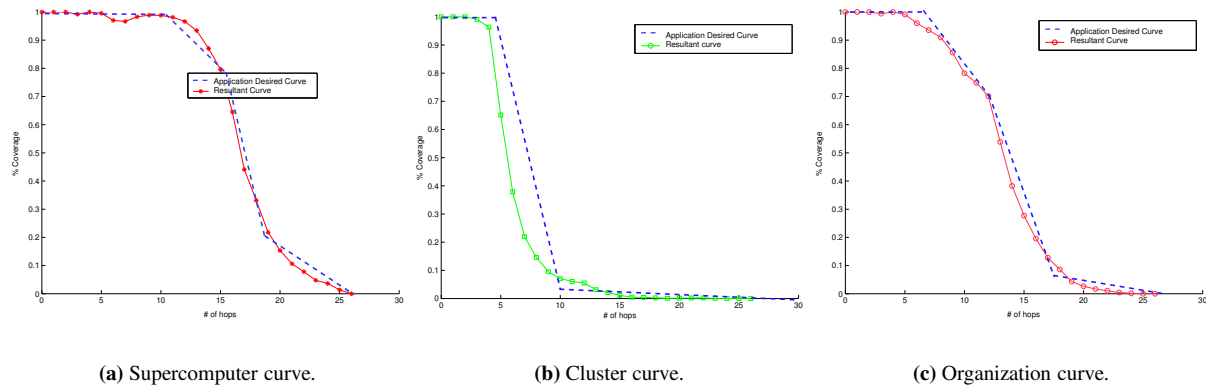
Figure 5. Coverage characteristics for three sample scenarios

coverage at a considerably lower overhead than by choosing a high value of k and then using small forwarding probabilities at later hops. In our experiments, we observed that when k is set to 7 and the forwarding probability is set to 0.52, the coverage at 25 hops is 10% at an overhead of approximately 2300 packets. On the other hand, to obtain 10% coverage at nodes 25 hops away from the source, when k is set to 19 (with forwarding probability 0.40 afterwords), an overhead of approximately 4700 packets is incurred. Due to space constraints, we have presented only one representative scenario over here. We refer readers to [3] for detailed results on this study as well as more detailed analysis of the proposed protocols in this paper.

4 Summary and Future Work

Our previously proposed randomized dissemination protocols assumed that coverage can be directly controlled by setting forwarding probabilities to fixed or linearly decreasing levels. However, we show that the relationship between forwarding probabilities and coverage is complex. Specifically, coverage is not sensitive to forwarding probabilities at high probability ranges, nor at low probability ranges. It is only sensitive to them in a relatively small range where small changes of forwarding probabilities can lead to large changes in achieved coverage. As a result, we propose to control the dissemination footprint by setting the forwarding probabilities to values that best fit the desired coverage. We show using examples how different application-desired footprints can be achieved at low overhead. This controllability is an important capability that is necessary (but not necessarily sufficient) for adaptive control of dissemination to optimize scheduling in the face of changing Grid resource states and query generation patterns. We plan to use this work as a building block in our future work on adaptive dissemination

protocols.

References

[1] S. Banerjee, B. Bhattacharjee, and C. Kommareddy. Scalable application layer multicast. Technical report, University of Maryland, College Park, 2002.

[2] A. R. Butt, R. Zhang, and Y. C. Hu. A self-organizing flock of condors. *SC '03*, November 15-21, 2003, Phoenix, AZ.

[3] B. Gandhi, S. Tilak, M. Lewis, and N. Abu-Ghazaleh. Coverage control for grid resource information dissemination protocols, 2005. Available on the web at `http://grid.cs.binghamton.edu`.

[4] V. Iyengar, S. Tilak, N. B. Abu-Ghazaleh, and M. J. Lewis. Nonuniform information dissemination for dynamic grid resource discovery. *Proceedings of IEEE NCA*, pages 97–106, 2004.

[5] L. Li, J. Halpern, and Z. Haas. Gossip-based ad hoc routing. In *IEEE Infocom*, 2002.

[6] Q. Lv, P. Cao, E. Cohen, K. Li, and S. Shenker. Search and replication in unstructured peer-to-peer networks. *SIGMETRICS Perform. Eval. Rev.*, 30(1):258–259, 2002.

[7] M. Maheswaran and K. Krauter. A parameter-based approach to resource discovery in grid computing system. In *GRID*, pages 181–190, 2000.

[8] A. Rowstron and P. Druschel. Pastry: Scalable, decentralized object location, and routing for large-scale peer-to-peer systems. *Lecture Notes in Computer Science*, 2218:329–350, 2001.

[9] Y. Sasson, D. Cavin, and A. Schiper. Probabilistic broadcast for flooding in wireless mobile ad hoc networks, 2002.

[10] S. Tilak, A. Murphy, and W. Heinzelman. Non-uniform information dissemination for sensor networks. In *11th IEEE International Conference on Network Protocols (ICNP'03)*, 2003.