Chapter 6

Conclusion

With the exponential growth of the Internet, there is a critical need to design efficient, scalable and robust protocols to support the network infrastructure. In this thesis, we examined a new class of protocols that accomplish these goals by using multipoint, announcement-style communication.

We observed that many multipoint protocols rely on a small set of very powerful techniques to accomplish scalability. We studied a number of the scalability techniques that repeatedly appear in both proposed and deployed Internet protocols: Suppression, Announce-Listen and Leader Election. We primarily examined scenarios where group members arrived into the system simultaneously and where steady state was maintained. We also considered the effect of process arrivals and departures on these techniques.

We analyzed and simulated each algorithm using a probabilistic failure model. We derived performance metrics to evaluate each technique, focusing on metrics for delay, bandwidth usage, and memory overhead, where relevant. We also determined metrics for the level of consistency that can be achieved by these algorithms given their loosely-coupled messaging model and for the speed with which convergence can occur. The model exposed several key parameters that allowed us to study parts of the operating space that had not been fully investigated before. In particular, we presented a more sophisticated loss model, examining the impact of both correlated and uncorrelated loss on each algorithm, and allowing multiple versus single losses to occur in the system.

We studied the Suppression algorithm within both lossless and lossy networks. The first allowed us to scrutinize the importance of the timer distribution function on the tradeoff between metrics for response time and messaging overhead. We found that the benefits of one distribution over the other is very context dependent, which conflicts with the the standard practice of employing the uniform distribution by default. We also discovered that highly-tuned distribution functions ("designer functions") can be created that optimize the performance of a given Suppression metric, provided that certain network parameters can be bounded. In addition, examining Suppression in a lossless environment provided an opportunity to evaluate the algorithm in simplified terms, separate from other algorithms (with which it is often combined) and separate from any particular operating context (such as reliable multicast).

We then re-examined the Suppression algorithm under lossy conditions, refining the known metrics to more accurately reflect their performance in a lossy network and proposing several new metrics, including the completion time of the algorithm, as well as the expected number of required versus extra messages.

For Announce-Listen, we proposed an alternate model to existing models. The model not only allowed closer examination of parameters for caching and transmission delay, but also placed AL within a multipoint context where we could study the impact of group size more directly. We focused on the derivation of a metric for consistency, which was then used to frame the discussion of other metrics. Most importantly, we validated our model through simulation.

Unlike many other approaches to Leader Election, we did not tie the LE algorithm to any particular network topology. Instead we characterized its performance in terms of the bounds set by its transmission delay. In addition, we contrasted a more traditional version of the algorithm with one that employs a round of Suppression at the beginning of the algorithm. We found that Suppression offers superior scalability properties for large groups in lossy networks. We showed that a leader selection criterion based on greatest process identifier is limited, and that a leader selection policy based on the Suppression wake-up time is superior.

We also explored a theory of composition, that the metrics for SUP and AL were suitable metrics for an LE algorithm built from them. We discovered that they bound the performance of LE under certain ideal conditions, but were more aptly described as indicators of performance trends. The divergence between the performance of SUP and AL as separable components and the performance when combined into LE was due to changes in the functionality of SUP when it was used as a building block. An open issue is to explore the results when there exists a direct correspondence between components and composed algorithm.

Much of our analysis was validated through extensive simulation. Using our analysis, designers will be able to parameterize, as well as to predict the performance of these algorithms. Hence, designers will be better informed about how to optimize the metrics of choice. More broadly, through the analysis of these algorithms we established a methodology for examining other scalability techniques.

In the future, we would like to revise our network assumptions. We would like to examine different delay models, studying the variance in transmission delays. We would like to investigate other timer distribution functions in the Suppression algorithm, especially for differentiation between classes of group members. A future plan is to create a parameterized version of the basic Suppression technique; the decision to suppress would be based on network loss conditions and on the number of messages a process received while waiting for the Suppression timer to expire. In the future, we would like to model and to analyze the consistency of alternate versions of Announce-Listen, in which certain processes only listen sporadically to the group address or rely on proxies or hierarchies of proxies to distribute announcements on their behalf. For Leader Election, an open issue is to compare the traditional usage of process id with other election schemes, which are geared toward optimizing certain system metrics, such as response time, consistency, load, or topological placement. Finally, we would also like to identify and evaluate other scalability techniques upon which multicast algorithms are built.