Improving OSPF Dynamics on a Broadcast LAN

M. Goyal
W. Xie
S. H. Hosseini
K. Vairavan
D. Rohm
EECS Department
University of Wisconsin–Milwaukee
3200 N Cramer Street, Milwaukee, WI 53201
mukul@uwm.edu

In this article, the authors propose modifications to OSPF's interface state machine to reduce the time/processing requirements of the leader election process in a broadcast local-area network (LAN) environment. The proposed modifications are based on *dynamic* adjustment of wait time duration rather than using a static value. They compare the original and modified state machines both analytically as well as via testbed experiments.

Keywords: OSPF, interface state machine, broadcast LAN, leader election

1. Introduction

OSPF [1] is a link state routing protocol used for *intra-AS* routing. OSPF has been a popular choice for intra-AS routing and enjoys a large deployed base as it has a robust design that provides a good trade-off between the protocol's processing load/control traffic and the speed with which it converges to topology changes. However, over the past several years, the convergence speed requirements as well as the nature and size of the networks that OSPF serves have changed significantly. Hence, in recent times, there has been a significant effort to reevaluate the OSPF design. The primary goal of this effort has been to suggest protocol changes that speed up its operation/convergence speed without increasing the processing load on the routers. It is critical that the proposed modifications do not exacerbate the processing load on the routers since CPU overload in some routers may snowball into network-wide routing instability. In this article, we propose modifications to OSPF's interface state machine to speed up the leader election process in a broadcast local-area network (LAN) environment while reducing the processing requirements.

In OSPF, two neighbor routers are said to be *fully adjacent* to each other if their link state databases are synchronized. The need for database synchronization makes adjacency establishment and maintenance a costly process in terms of processing and time required. In a broadcast LAN environment, where every other router can be considered to be a neighbor, establishment/maintenance of full adjacency with every neighbor may put significant burden on a router. Hence, to reduce the number of full adjacencies required, the OSPF protocol requires the routers on a LAN to elect a leader among themselves to function as the *designated router* (DR). The DR originates a network-LSA (Link State Advertisement) on behalf of the LAN and establishes full adjacency with every other router on the LAN. The routers on the LAN also elect a *backup-designated router* (BDR) among themselves, which also establishes full adjacency with every other router on the LAN. In the event of the DR's failure, the BDR can quickly take over the responsibilities of the DR since it is already fully adjacent to all the other routers on the LAN. The other routers that are neither DR nor BDR establish full adjacency only with DR and BDR and consider themselves to be *DR Other*. The transition of a broadcast LAN interface from its initial *Down* state to one of the *DR/BDR/DROther* states is determined by OSPF's interface state machine (Fig. 1). Depending on the order in which different routers on the LAN come up, the routers may take a long time before settling on the identity of the DR and BDR for the LAN, performing several DR elections in the process.

Note that the leader election problem has been extensively studied in the context of the distributed systems [2-5]. In this article, we are concerned with the behavior of a particular leader election algorithm implemented by OSPF's interface state machine [1]. Furthermore, we are concerned with the behavior of OSPF's interface state machine in a *broadcast* LAN environment (e.g., Ethernet). The broadcast networks differ from nonbroadcast networks

SIMULATION, Vol. 82, Issue 2, February 2006 107-129

DOI: 10.1177/0037549706065924

in their ability to "broadcast" single physical messages to all of the routers on the network. In such networks, a router discovers other routers on the LAN dynamically by periodically broadcasting a *Hello* message [1]. On the other hand, in nonbroadcast networks (e.g., X.25), dynamic discovery may not be possible, and neighbors may need to be configured.

Network Model and Performance Metrics. The network model considered in this article consists of a number of routers connected over a broadcast LAN. Initially, all routers are down. The routers come up in a random order at random times. *Once up, the routers stay up*. Each router broadcasts its first Hello immediately upon coming up and subsequently after every hello interval H (10 seconds by default). Every Hello message carries the sending router's *Router ID* and *Router Priority* (discussed later), the list of known neighbors on the LAN, and the sending router's idea of DR and BDR for the LAN. We assume that no Hello packet is lost and the time required for a Hello to reach other routers as well as the time required for a router to process a Hello and perform a DR election is negligible. Let the *DR/BDR settling time* for a router be the time required by the router to settle on the final identity of the DR and BDR for the LAN (i.e., the time after which the router does not alter its view of the identity of DR and BDR routers). The DR/BDR settling time for the LAN as a whole can be defined as the latest DR/BDR settling time for a router on the LAN. In other words, the DR/BDR settling time for the LAN is the time required for all the routers on the LAN to settle on the identity of the DR and BDR for the LAN. In this process, each router may perform several DR elections. Some of these elections may be performed after the settlement of DR/BDR identity. We use the DR/BDR settling time for the LAN and the number of DR elections each router performs in the process as the metrics in our analysis.

Terminology. In the following discussion, we use the term *router* to mean its interface on the broadcast LAN where it is more convenient to do so. Also, the term *arrival* refers to the event of a router coming up. A received Hello is characterized as *two-way* or *one-way* based on whether it lists the receiving router as a neighbor. A router associates a *two-way* or better state with a neighbor when it receives a two-way Hello from the neighbor[1] [1]. A router is said to have established bidirectional communication with a neighbor when it associates a two-way or better state with the neighbor. A router elects a DR and a BDR from the set of routers consisting of itself and its bidirectional neighbors. Henceforth, we refer to this set as the set of *eligible* routers. Finally, we define an eligible router's *claim to DR/BDRship* as follows. If the eligible router is the router itself, we say that it claims DRship if the router considers itself to be the DR for the LAN. If the eligible router is a bidirectional neighbor, we say that it claims DRship if it has declared itself to be DR in its most recent Hello. The claim to BDRship is defined similarly.

The rest of the article is organized as follows. Section 2 explains the current interface state machine used by the OSPF protocol, and section 3 analyzes its behavior in different broadcast LAN scenarios. Section 4 contains a stochastic analysis for the expected number of DR elections performed by a router under the current interface state machine assuming exponentially distributed arrival times for the routers. Section 5 introduces the proposed modifications to the interface state machine and analyzes the behavior of the modified state machine in different scenarios. Section 6 obtains the expected number of DR elections performed by a router under the modified interface state machine, again assuming exponentially distributed arrival times for the routers. Section 7 compares the performance of the two state machines using testbed experiments. Finally, section 8 concludes the article. Note that this article is a significant extension of an earlier study [6], which did not contain stochastic analysis of the interface state machine before and after the proposed modifications (sections 4 and 6). The earlier study did not contain the experimental verification results either (section 7 and the appendix).

2. OSPF's Interface State Machine

Figure 1 illustrates the state machine followed by a router interface on a broadcast LAN. The interface is considered to be in the *down* state if it is not operational. On coming up, the interface starts the wait timer and transitions to the *waiting* state. The purpose of the waiting state is to allow the interface to establish bidirectional communication with other routers on the LAN and determine the identity of existing DR and BDR if any. The interface does not perform a DR election while it is in the waiting state. The interface emerges from the waiting state when one of the following events takes place:

- **Wait timer fires**. This happens when the wait time duration, W, is over since the coming up of the interface. The wait time duration W is set to be same as the *RouterDeadInterval*, which is four times the hello interval H (i.e., $W = 4H$) [1]. The wait time duration is set to a large value so as to allow the new interface to establish bidirectional communication with all the other routers on the LAN before it does its first DR election.

- **BackupSeen event is generated**. This happens when the interface has detected the presence or absence of a BDR either on receiving a two-way Hello claiming the BDRship for the sending router or receiving a two-way Hello claiming DRship for the sending router with no router listed as BDR.

The interface performs a DR election on emerging out of the waiting state and transitions to the *DR/BDR/DROther*

1. The neighbor state can also reach a two-way or better status when the router receives a *Database Description* packet from the neighbor [1].

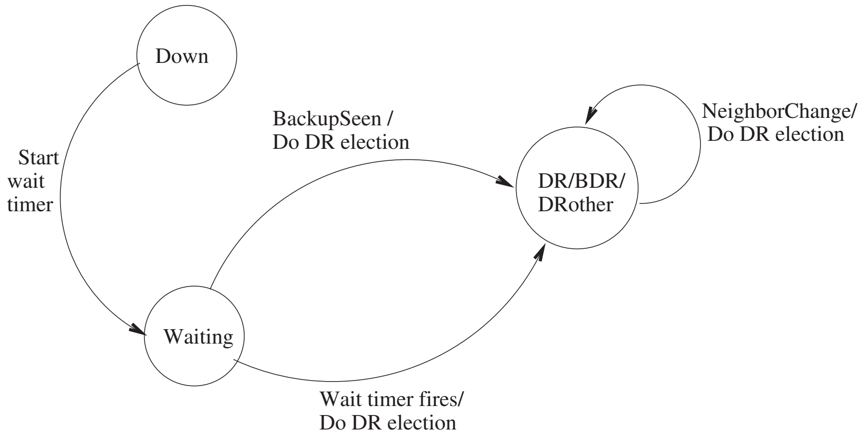

Figure 1. The current interface state machine in OSPF

state. Once out of the waiting state, the interface will perform a DR election whenever a *NeighborChange* event is generated. This event indicates a change in the set of bidirectional neighbors associated with the interface and is generated when

- the bidirectional communication is established (or breaks down) with a neighbor on the LAN,
- a bidirectional neighbor newly declares (or is no longer declaring) itself as DR/BDR for the LAN,
- the *Router Priority* changes for a bidirectional router. The Router Priority is used in the DR election, as described next.

The DR Election Algorithm. The DR election takes place in one or two passes. The second pass is performed if the router performing the DR election newly elects itself as DR/BDR in the first pass or had elected itself as DR/BDR in the previous election but does not do so in the first pass. The BDR election takes place first, as discussed next. If an eligible router claims DRship, it is not considered for election as BDR. If one or more eligible routers have claimed BDRship, the router with highest Router Priority is elected as BDR. In case of a tie, the router with highest Router ID is chosen as BDR. If no eligible router has claimed BDRship, the eligible router with highest Router Priority is elected as BDR. The ties are again broken using the Router ID. Once a BDR has been elected, the router elects a DR. If more than one router has claimed DRship, the router with highest Router Priority is chosen as DR. The ties are broken using Router ID. If no router has yet claimed DRship, the router elected as BDR is elected to be DR as well.

In the following analysis, we assume that all the routers have the same Router Priority.

3. Analyzing the Interface State Machine

In this section, we analyze the behavior of OSPF's interface state machine over a broadcast LAN. As discussed earlier, the network model consists of n routers coming up on a broadcast LAN in a random fashion. Based on the manner in which different routers arrive, we consider three different scenarios.

3.1 Scenario 1: A New Router Comes Up on a LAN with Settled DR/BDR

Consider the scenario where the first $n-1$ routers in the LAN already agree on the DR/BDR identity by the time the nth router, router z, comes up (Fig. 2). Let us assume that routers x and y are the existing DR and BDR, respectively, for the LAN. The router z sends out its first Hello immediately after coming up, which is received by all the other routers on the LAN and ensures that these routers mention router z as a neighbor in their next Hello. Within the next hello interval, router z will receive two-way Hellos from the other routers on the LAN (including routers x and y). On receiving the Hello from router y, router z will generate a BackupSeen event, come out of the waiting state, and do a DR election. If by this time, router z has already received a Hello from router x, it will elect router x as DR and router y as BDR. Otherwise, at this time, router y will be elected as both DR and BDR, and when the Hello from router x arrives, router x will be elected as DR. Thus, the new router elects the existing DR/BDR for the LAN as its own DR/BDR within one hello interval of coming up. Within this interval, the router may perform several DR elections. The first DR election is performed when the router receives the first two-way Hello from BDR. Every subsequent establishment of bidirectional communication with a new neighbor causes router z to perform a new DR

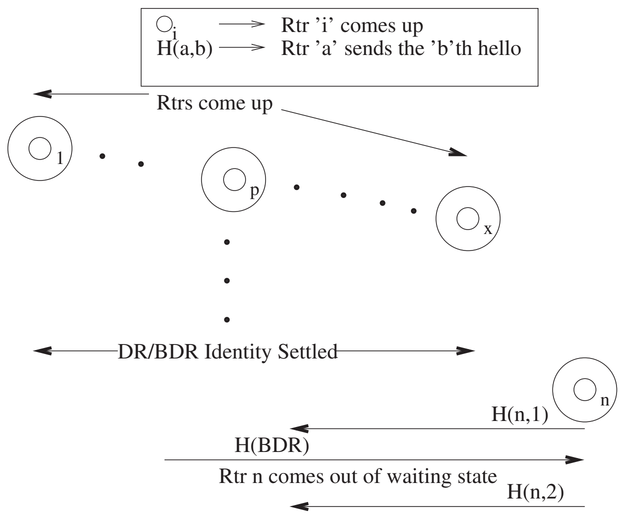

Figure 2. Scenario 1: A new router comes up on a local-area network (LAN) with settled designated router/backup designated router (DR/BDR)

election. Thus, router z may end up doing anywhere between 1 and $n-1$ DR elections. Assuming that router z may receive two-way Hellos from other routers on the LAN in any order with equal probability, it can be shown that the expected number of DR elections performed by router z is $n/2$.

3.2 Scenario 2: All Routers Bidirectional before the First Router's First DR Election

Now consider the scenario, shown in Figure 3, where all the n routers on the LAN are bidirectional with each other by the time the first router does a DR election on the firing of its wait timer. Let router x be the router with the highest Router ID among all the routers on the LAN, and let router y be the router with next highest Router ID. As the wait timer fires on a router, the router (except router x) will perform a DR election and elect router x as both DR and BDR. Router x will elect itself as DR and router y as BDR. In its next Hello, router x will declare itself to be the DR for the LAN. This Hello will be sent within one hello interval of router x electing itself as DR. On receiving this Hello, every router not in waiting state will perform a DR election electing router x as the DR and router y as the BDR. A router coming out of the waiting state subsequent to receiving this Hello will also elect router x as DR and router y as BDR. When router y sends out its next Hello, it will declare itself to be the BDR. On receiving this Hello, all the routers still in the waiting state will generate a BackupSeen event and perform a DR election. Other routers not in waiting state will also perform a DR election as a NeighborChange event is generated on receiving this Hello. In this DR election,

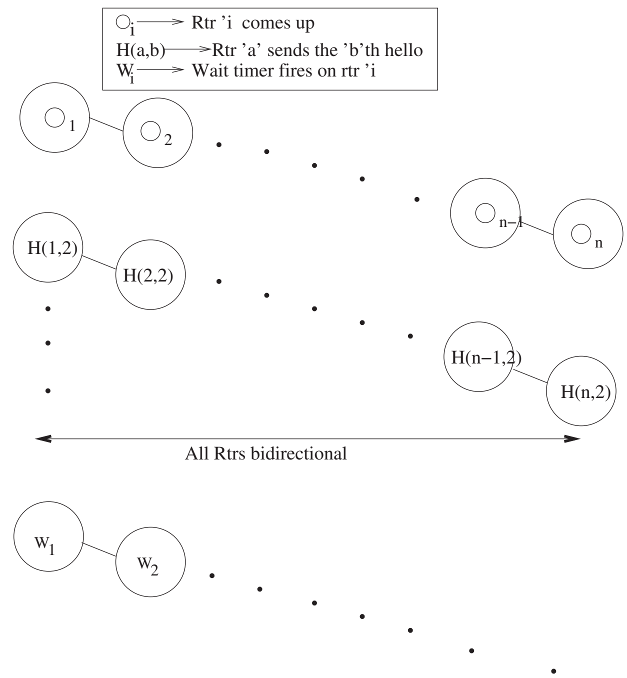

Figure 3. Scenario 2: All routers bidirectional before the first router's first designated router (DR) election

each router will (re)elect routers x and y as the DR and BDR, respectively.

3.3 Scenario 3: All Routers Not Necessarily Bidirectional before the First Router's First DR Election

Now we consider the most general scenario, shown in Figure 4, where all the routers on the LAN may not have established bidirectional communication with each other by the time the first router comes out of the waiting state. In fact, some of the routers may not yet be up by this time. In the following discussion, we make certain claims about the behavior of OSPF's interface state machine in this scenario.

Claim 1. If a router, say router x, is the only router to elect itself as DR by the time its Hello claiming the DRship is received by other routers, then the other routers will not elect themselves as DR and will eventually elect router x as DR.

Proof. Suppose router x is the only router to elect itself as DR by the time its first Hello claiming the DRship is received by other routers on the LAN. On receiving this Hello, all the routers not in waiting state will do a DR election and elect router x as DR since router x is the only router with a claim to DRship. Consider the routers that are in the waiting state when they receive this Hello. As

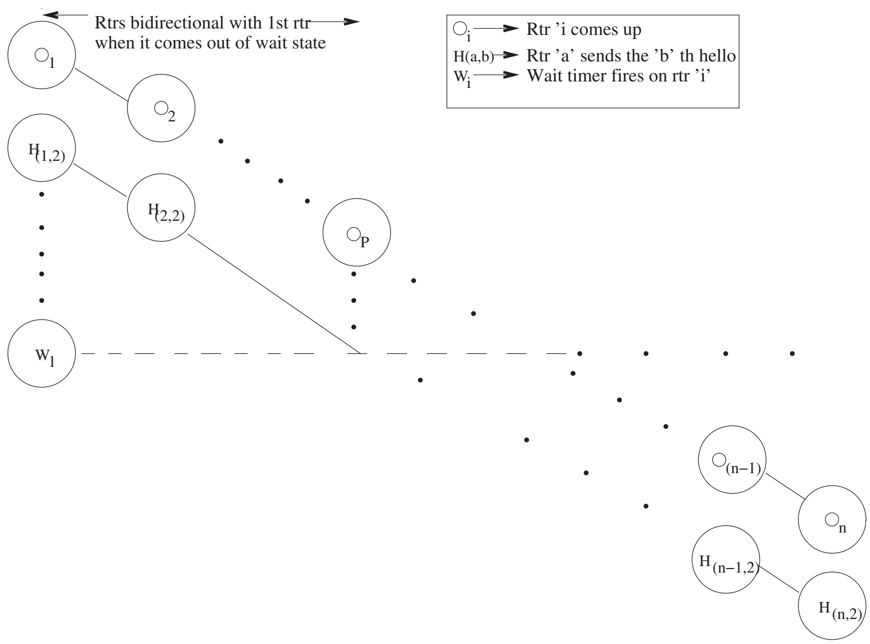

Figure 4. Scenario 3: All routers not bidirectional before the first router's first designated router (DR) election

each such router comes out of the waiting state, it will elect router x as DR since router x will still be the only router to claim DRship. Now consider the routers that come up after the time when router x sends out the first Hello claiming DRship. Suppose router c is the first such router to come out of the waiting state. If router c comes out of the waiting state as the result of its wait timer's firing, it would have received a Hello from router x claiming the DRship and would not have received any other claim to DRship. Hence, router c will elect router x as DR. Suppose router c comes out of the waiting state as the result of the BackupSeen event. If the BackupSeen event was caused by receiving the Hello from router x (the only DR) or router c had received a Hello from router x claiming DRship before the BackupSeen event, router c will elect router x as DR. If the BackupSeen event was caused by receiving the Hello from a BDR and router c had not received a Hello from router x claiming DRship by this time, router c will elect the BDR as both DR and BDR. When it finally receives the Hello from router x, it will elect router x as DR. Similarly, any other router that comes up after the time when router x sends out the first Hello claiming DRship will also elect router x as DR.

Claim 2. A router must have declared itself as the DR in its Hello, before another router elects itself as BDR.

Proof. If router y considers itself to be DR before a DR election and has not yet received any other DRship claim, it will not elect itself as BDR. Suppose router y does not consider itself to be DR before a DR election and elects itself as BDR in the first pass of the election. If no router has declared itself to be DR in its Hello so far, router y will elect itself as DR as well and perform the second pass of the DR election in which it will elect some other router as BDR. Thus, router y can elect itself as BDR in a DR election only if some other router has declared itself to be DR in its Hello.

Claim 3. If a router elects itself as DR for the LAN, it must have done so in its first DR election after coming out of the waiting state.

Proof. Consider a router, a, as it does a DR election on coming out of the waiting state. If, by this time, router a has received Hellos from one or more routers claiming the DRship, router a will always elect one of these routers as DR and will not claim DRship itself now or in future (as long as one or more of these routers continue to claim themselves as DR). If router a has not yet received a Hello claiming the DRship for the sending router but has received Hellos from one or more routers claiming the BDRship, router a will elect one of these routers as BDR as well as DR. Since, in this case, one or more routers are claiming BDRship, as per claim 2, one or more routers must have already claimed the DRship. When router a finally receives these claims, it will elect one of the claimant routers as DR and not claim DRship itself now or in future. If router a has not yet received any Hello claiming either DRship or BDRship for the sending router, it will elect the router with the highest Router ID among its set of eligible routers as DR. If, at this time, router a does not elect itself as

DR, there must be another bidirectional neighbor with a higher Router ID than router a, and in any subsequent DR election, router a will not elect itself as DR ahead of this router. Hence the proof.

Claim 4. Only one router elects itself as the DR for the LAN.

Proof. Suppose router x elects itself as DR at time W_x. According to claim 3, router x must have just come out of the waiting state. Also, by this time, router x must not have received a Hello from another router claiming the DR/BDRship (otherwise it would have elected this other router as DR and not itself). So, router x came out of the waiting state as the result of its wait timer firing and hence must have been up since time $0_x = W_x - W = W_x - 4H$. Router x will send out a Hello declaring itself as DR within one hello interval of electing itself as DR (i.e., within time $W_x + H$). According to claim 1, if router x is the only router to have elected itself as DR by the time router x's Hello claiming DRship is received by other routers on the LAN, then the other routers will not elect themselves as DR. Suppose the precondition for claim 1 does not hold; that is, another router z elects itself as DR at time W_z before receiving the Hello from router x claiming DRship ($W_z < W_x + H$). Using the same reasoning as above, we can claim that router z must have just come out of the waiting state at time W_z and must have been up since time $0_z = W_z - W = W_z - 4H$. Also, router z sends out a Hello within time $W_z + H$ claiming Drship, and router x must have come out of the waiting state before receiving this Hello ($W_x < W_z + H$). Clearly, $|W_x - W_z| < H$ and hence $|0_x - 0_z| < H$; that is, routers x and z must have come up within one hello interval of each other and hence must have been up for at least three hello intervals together before the first one among them comes out of the waiting state. Hence, routers x and z must have established bidirectional communication with each other by the time they come out of the waiting state. Therefore, both routers x and z cannot elect themselves as DR. Assuming that router x does elect itself as DR, router z cannot elect itself as DR, and hence the precondition for claim 1 holds, and no router other than router x elects itself as DR.

Claim 5. Each non-DR router may potentially elect itself as BDR for the LAN.

Proof. Consider the scenario where router x does not have any bidirectional neighbor when it comes out of the waiting state at time W. Router x does a DR election and elects itself as DR for the LAN with no BDR. All the other routers come up simultaneously a little while before router x sends out a Hello claiming DRship (say at time $W + H - 1$). These routers will send out their first Hellos immediately on coming up. Suppose router x receives these Hellos and immediately afterwards sends out its own Hello listing all other routers as neighbors and declaring itself as DR with no BDR. On receiving this Hello, every other router will generate a BackupSeen event (Hello from DR with no BDR indicated) and perform a DR election. Since at this time, router x is the only bidirectional neighbor for each of the other routers, these routers will elect themselves as BDR in the DR election.

Claim 6. If router y is the only router to elect itself as BDR by the time its first Hello claiming the BDRship is received by other routers on the LAN, then the other routers will not elect themselves as BDR and will elect router y as BDR.

Proof. When a router, a, receives the Hello from router y newly claiming the BDRship, it will perform a DR election irrespective of whether it was in the waiting state before receiving the Hello. If router y is the only router with a claim to BDRship at this time, it will be elected as BDR in this DR election. Thus, all the routers "up" at the time of receiving the Hello from router y claiming BDRship will elect router y as BDR. Consider the first router, b, to come up after router y sends its first Hello claiming BDRship. Router b will receive a two-way Hello from router y within a hello interval of coming up, thereby generating a BackupSeen event, and will elect router y as BDR in the subsequent DR election. Thus, all the newly up routers will also elect router y as BDR.

Claim 7. If a router elects itself as BDR for the LAN, it must have done so in its first DR election after receiving a Hello claiming DRship for the sending router.

Proof. Consider a router, a, that is not in the waiting state. Router a will perform a DR election on receiving the first Hello claiming DRship for the sending router. According to claim 2, router a could not have elected itself as BDR before this time. If, by this time, router a has received other claims to BDRship, it will always elect one of the claimant routers as BDR and will not claim BDRship itself now or in future (as long as one or more of these routers continue to claim BDRship). If router a has not yet received a claim to BDRship, it will elect the router with the highest Router ID among its set of eligible routers (excluding the DR) as BDR. If, at this time, router a does not elect itself as BDR, there must be another bidirectional non-DR neighbor with a higher Router ID than router a, and in any subsequent DR election, router a will not elect itself as BDR ahead of this router. Similarly, it can be proved that a router a, if it is in waiting state when it receives the first Hello claiming DRship for the sending router, will either elect itself as BDR in its first DR election on coming out of the waiting state or will not elect itself as BDR then or in future.

Claim 8. If the first Hello from a router claiming the DRship also mentions a BDR, only one router will elect itself as BDR.

Proof. Suppose the first Hello from router x claiming the DRship also mentions another router as the BDR. Receiving this Hello will not cause other routers in the waiting state to generate a BackupSeen event and perform a DR election. Suppose, as per claim 7, router y elects itself as BDR in its first DR election after receiving this Hello. Router y must not have received another claim to BDRship by this time (otherwise it would not have elected itself as BDR). Hence, router y could not have come out of the waiting state because of the BackupSeen event and must have been up for at least W seconds before electing itself as BDR. Router y will send out a Hello claiming BDRship within one hello interval of electing itself as BDR. As per claim 6, if router y is the only router to elect itself as BDR by the time its first Hello claiming the BDRship is received by other routers on the LAN, then the other routers will not elect themselves as BDR. Using the same methodology as in the proof for claim 4, it can be shown that no other router elects itself as BDR before receiving router y's Hello claiming BDRship. Hence, as per claim 6, only router y elects itself as BDR.

3.4 Upper Bounds on the Number of DR Elections Performed by a Router

Consider the scenario where multiple routers elect themselves as BDR. It is possible that each one of these routers sends out a Hello claiming BDRship. On receiving the BDRship claims from other routers, all these routers except one will renounce their claims. A router performs a DR election both on receiving a Hello newly claiming BDRship for the sending router and on receiving a Hello renouncing the sending router's claim to BDRship. Since, in a LAN of n routers, up to $n-1$ routers may elect themselves as BDR, a router may perform up to $n-1$ DR elections on receiving new claims to BDRship and $n-2$ elections as all the claims except one are renounced. Thus, a router may perform up to $2n-3$ DR elections in this manner. In addition, up to $n-1$ DR elections may be performed as a router establishes bidirectional communication with its neighbors after coming out of the waiting state. Two more DR elections may be performed as the result of the wait timer's firing and receiving a Hello from DR newly claiming the DRship. Thus, a router may perform up to $3n-2$ DR elections in its DR/BDR settling process. Note that if only one router elects itself as BDR for the LAN, the upper bound on the number of DR elections per router decreases to $n+2$. If all the routers establish bidirectional communication with each other before coming out of the waiting state (as in scenario 2), each router may need to perform at most 3 DR elections.

3.5 DR/BDR Settling Time

In the following discussion, we start counting time from the instant the first router comes up. Suppose router z is the last router to come up at time 0_n. Clearly, the overall DR/BDR settling time for the LAN will be at least 0_n. Also, the overall DR/BDR settling time for the LAN should at least be W, before which all the routers are either in waiting state or not up yet. Thus, the DR/BDR settling time for the LAN must at least be $max(W, 0_n)$. To determine the upper range for the DR/BDR settling time for the LAN, we consider all the possible scenarios:

- Consider the case where the other $n-1$ routers have already elected routers x and y as the DR and BDR by the time router z comes up at time 0_n. In this scenario, router z will come out the waiting state on receiving its first two-way Hello from router y (the BDR) and elect routers x/y as DR/BDR within one hello interval of coming up (i.e., within time 0_n+H). Thus, the overall DR/BDR settling time for the LAN in this case will be within 0_n+H.
- Consider the case where one router (router x) has elected itself as DR, and multiple routers have elected themselves as BDR by the time router z comes up. Thus, router z will receive one or more claims to BDRship within one hello interval of coming up and come out of the waiting state. Router z will also receive one claim to DRship (from router x) within one hello interval of coming up. After processing all these claims, router z (and other routers on the LAN) would have elected router x as DR and the router with the highest Router ID among all the claimants to BDRship as the BDR. Thus, in this case too, the overall DR/BDR settling time for the LAN will be within 0_n+H.
- Consider the case where one router (router x) has elected itself as DR but no router has elected itself as BDR by the time router z comes up. Clearly, the router(s) to elect itself (themselves) as BDR in future is (are) still in the waiting state. Router z is the last router to come up and will come out of the waiting state within time 0_n+W $(=0_n+4H)$. Therefore, all the future claimants to BDRship will also come out of the waiting state and elect themselves as BDR by time 0_n+W. These routers will send their claims to BDRship within time 0_n+W+H. By this time, the other routers must have also received router x's claim to DRship. On processing all these claims, all the routers will (re)elect router x as DR and the router with the highest Router ID among all the claimants to BDRship as the BDR. Thus, in this case, the overall DR/BDR settling time for the LAN will be within time 0_n+W+H $(=0_n+5H)$.
- Consider the case where no router has elected itself as DR or BDR by the time router z comes up. Clearly, the router to elect itself as DR in future is still in the waiting state. Router z is the last router to come up and will come out of the waiting state within time 0_n+W $(=0_n+4H)$. Therefore, the future claimant to DRship, say router x, will also come out of the waiting state and elect itself as DR by time 0_n+W and send its first claim to DRship by time 0_n+W+H. On processing this claim, the other routers will (re)elect router x as DR, settling the DR

identity. Now we consider the time taken to settle the BDR identity. Since the last router comes up at time 0_n, all the routers on the LAN would have established bidirectional communication with each other by time $0_n + H$. If router x sends out its first claim to DRship after this time, it would have elected (and will mention in its Hello carrying its DRship claim) the router with the highest Router ID among all the routers, excluding itself (say router y) as BDR. The other routers on the LAN, when they receive the DRship claim from router x, will also elect router y as BDR. Thus, in this case, the BDR identity will also settle by time $0_n + W + H$. If router x sends its first claim to DRship before time $0_n + H$, it is possible that multiple routers may elect themselves as BDR. These routers will send their claims to BDRship within time $0_n + 2H$. After processing these claims, each router would have elected the router with the highest Router ID among all the claimants to BDRship as BDR and router x as DR. Thus, in this case, the DR/BDR settling time for the LAN will be within time $0_n + 2H$.

Thus, we can claim that the DR/BDR settling time for the LAN ranges between $max(W, 0_n)$ and $0_n + W + H$.

4. A Stochastic Analysis of DR Elections Performed by a Router

As discussed above, a router may perform a large number of DR elections in its DR/BDR settling process. Most of these elections will be caused by multiple routers initially electing themselves as BDR and a router establishing bidirectional communication with its neighbors after coming out of the waiting state. In the following, we analyze these two scenarios assuming that the router arrival times are exponentially distributed with parameter λ. Thus, the mean arrival time for a router is $1/\lambda$ with a variance $1/\lambda^2$.

4.1 Expected Number of Routers Initially Electing Themselves as BDR

As per claims 5 and 8 in section 3.3, multiple routers may elect themselves as BDR initially if the DR is their first bidirectional neighbor and the first two-way Hello they receive from the DR does not list any router as BDR. The Hello from DR will not list any router as BDR as long as the DR is not bidirectional with any other router. This will happen if the DR is the first router to come up on the LAN and the second router comes up at least $3H$ seconds after the DR. Suppose the DR (first router) comes up at time t_1 and the second router comes up at time t_2; then, multiple routers may elect themselves as BDR initially if $t_2 > t_1 + 3H$. Otherwise, only one router will elect itself as BDR.

Suppose the second router arrives at least $3H$ seconds after the first one (i.e., $t_2 > t_1 + 3H$). This event happens with probability $e^{-3(n-1)\lambda H}$. Let $t_2 + t'$ be the time when the DR sends its first Hello after the coming up of the second router. Since the DR will become bidirectional with the second router at time $t_2 + H$, the Hello it sends at time $t_2 + t' + H$ will list some router as BDR. Hence, only the routers coming up between times t_2 and $t_2 + t'$ (including the second router) will elect themselves as BDR initially. Assuming that t', the time interval between the second router's arrival and the DR sending its next Hello, is uniformly distributed between 0 and H, the conditional probability that m routers elect themselves as BDR initially given that the second router arrives at least $3H$ seconds after the first one is given by $P(\text{m BDRs}|t_2 > t_1 + 3H) = \frac{1}{H}\int_0^H \binom{n-2}{m-1}(1-e^{-\lambda t'})^{m-1}(e^{-\lambda t'})^{n-m-1}dt'$. Hence, the conditional expectation of the number of routers electing themselves as BDR initially given that the second router arrives at least $3H$ seconds after the first one is given by $E[\#\,\text{BDRs}|t_2 > t_1 + 3H] = \sum_{m=1}^{n-1} m \times P[\text{m BDRs}|t_2 > t_1 + 3H]$. Finally, the expected number of routers electing themselves as BDR can be expressed as

$$\begin{aligned} E[\ \ \#\,\text{BDRs}\ \ ] &= E[\#\,\text{BDRs}|t_2 \le t_1 + 3H]P[t_2 \le t_1 + 3H] + \\ &\quad E[\#\,\text{BDRs}|t_2 > t_1 + 3H]P[t_2 > t_1 + 3H] \\ &= 1 \times (1 - e^{-3(n-1)\lambda H}) + \\ &\quad E[\#\,\text{BDRs}|t_2 > t_1 + 3H] \times e^{-3(n-1)\lambda H} \\ &= 1 + (E[\#\,\text{BDRs}|t_2 > t_1 + 3H] - 1)e^{-3(n-1)\lambda H}. \end{aligned}$$

Figure 5 shows the expected number of routers that initially elect themselves as BDR, using the expression above for $\lambda = 0.02$ and $\lambda = 0.05$ corresponding to mean router arrival times 50 seconds and 20 seconds, respectively. For reasonable mean router arrival times, the probability that the second router arrives at least $3H$ seconds after the first one becomes very small as the number of routers on the LAN increases. Consequently, the expected number of routers that initially become BDR is close to 1.

4.2 Expected Number of BDRship Claims if Multiple Routers Elect Themselves as BDR Initially

Suppose multiple routers elect themselves as BDR initially. On receiving a BDRship claim from another router, the remaining BDR claimants will perform a DR election and may no longer consider themselves as BDR as a result. Thus, very few of the initial claimants will be able to send out BDRship claims forcing other routers to perform a DR election. In the following, we show that if m routers elect themselves as BDR initially, the expected number of routers that will be able to send out BDRship claims is just H_m, the mth *harmonic number*.

Suppose m routers elect themselves as BDR initially ($m \ge 1$). Based on their router priority/Router ID combination, these routers can be ranked in a certain order such that on receiving a BDRship claim from a higher ranked

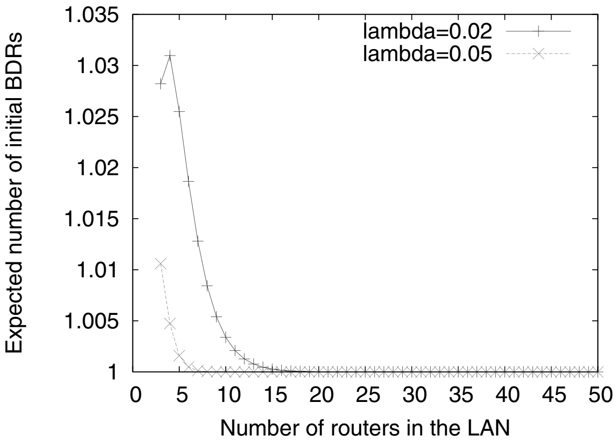

Figure 5. Expected number of routers that initially elect themselves as the backup designated router (BDR). H = 10 seconds.

router, all the lower ranked routers will give up their claim to BDRship. Thus, if the ith ranked router is the first router to send a BDRship claim, only i claimants will remain afterwards (all the routers with a higher ranking than i and the ith ranked router itself), and only top $i-1$ claimants may send any further claims. Since each router has equal probability of being the first router to send a BDRship claim, the expected number of total BDRship claims sent by m initial claimants can be written as $X(m) = 1 + 1/m \sum_{i=1}^{m} Y(i)$, where $Y(i)$ is the expected number of claims sent if top i claimants remain.

Suppose there are i claimants to BDRship at some stage. This means that top $i-1$ claimants are yet to send their BDRship claims. Since each one of these claimants is equally likely to send its claim first, using the same logic as above, the recurrence equation for $Y(i)$ is given by

$$Y(i) = \begin{cases} 1 + \frac{1}{i-1}\sum_{j=1}^{i-1} Y(j) & : \quad i \geq 2 \\ 0 & : \quad i = 1 \end{cases}$$

Using the equations above, we get

$$\begin{aligned}
&X(m) \\
&= 1 + \frac{1}{m}\sum_{i=1}^{m} Y(i) \\
&= 1 + \frac{1}{m}Y(m) + \frac{1}{m}\sum_{i=1}^{m-1} Y(i) \\
&= 1 + \frac{1}{m}(1 + \frac{1}{m-1}\sum_{i=1}^{m-1} Y(i)) + \frac{1}{m}\sum_{i=1}^{m-1} Y(i) \\
&= 1 + \frac{1}{m} + \frac{1}{m-1}\sum_{i=1}^{m-1} Y(i) \\
&= 1 + \frac{1}{m} + \frac{1}{m-1} + \frac{1}{m-2}\sum_{i=1}^{m-2} Y(i) \\
&= 1 + \frac{1}{m} + \frac{1}{m-1} + \frac{1}{m-2} + \cdots + \frac{1}{2} + Y(1) \\
&= 1 + \frac{1}{m} + \frac{1}{m-1} + \frac{1}{m-2} + \cdots + \frac{1}{2} \\
&= 1 + \frac{1}{2} + \cdots + \frac{1}{m-1} + \frac{1}{m} \\
&= H_m.
\end{aligned}$$

Hence, if m routers elect themselves as BDR initially, the expected number of routers that will be able to send out BDRship claims is just H_m, the mth *harmonic number*, which increases very slowly with increasing m. The results of this section and the previous one imply that the possibility of multiple routers electing themselves as BDR initially may not cause a significant number of DR elections.

4.3 Expected Number of DR Elections as a Router Establishes Bidirectional Communication with Other Routers

If a router is not in the waiting state, the establishment of bidirectional communication with a new neighbor causes a new DR election to take place. In this section, we determine the expected number of DR elections performed by a non-BDR router (say router z) as it establishes bidirectional communication with its neighbors. In this analysis, we assume that router z may come out of waiting state as a result of its wait timer firing or on receiving a BDRship claim from a neighbor. Furthermore, we assume that router z does not receive a DRship claim that does not list any router as BDR, and it does not receive BDRship claims from multiple routers. As discussed in sections 4.1 and 4.2, the probability of occurrence of these events is fairly low.

Suppose router z arrives at time t and the router that ultimately becomes BDR arrives at time t'. In the following discussion, we divide the time line in three parts based on t,t' values and calculate the expected number of DR elections performed by router z in each case. By multiplying the probability of occurrence of each case with the expected number of DR elections in that case and summing up the results, we can obtain the overall expected number of DR elections performed by router z.

$t \leq t'$. If router z arrives before the BDR, it will come out of waiting state as a result of its wait timer firing at time $t + 4H$. Hence, router z will be bidirectional with all the routers that arrive before time $t + 3H$. Assuming that the total number of routers on the LAN is n, the routers may arrive in any order with equal probability, and the probability that m routers arrive before router z $(0 \leq m \leq n-1)$ is $1/n$. Given that the remaining $n-m-1$ routers arrive after time t, the probability that p routers arrive between times t and $t + 3H$ is

$\binom{n-m-1}{p}(1-e^{-3\lambda H})^p(e^{-3\lambda H})^{n-m-1-p}$. Hence, the probability that q routers (excluding router z) arrive before time $t+3H$ is given by $\sum_{m=0}^{q}\frac{1}{n}\binom{n-m-1}{q-m}(1-e^{-3\lambda H})^{q-m}(e^{-3\lambda H})^{n-q-1}$. If q routers (excluding router z itself) arrive before time $t+3H$, router z would have established bidirectional communication with all these routers before coming out of waiting state at time $t+4H$. Thus, router z will perform just one DR election for these q routers. For each one of the remaining $n-1-q$ routers, router z will establish bidirectional communication with the router after coming out of waiting state and consequently perform one DR election. Thus, if q routers (excluding router z) arrive before time $t+3H$, router z will perform a total of $n-q$ DR elections as it establishes bidirectional communication with each router. Therefore, if $t \leq t'$, the expected number of DR elections performed by router z is given by $E[\#\text{ DR elections}|t \leq t'] = \sum_{q=0}^{n-1}(n-q)\sum_{m=0}^{q}\frac{1}{n}\binom{n-m-1}{q-m}(1-e^{-3\lambda H})^{q-m}(e^{-3\lambda H})^{n-q-1}$. Note that the probability that $t \leq t'$ is simply 0.5.

$t' < t \leq t'+3H$. In this case, the BDR arrives at time t' and will come out of the waiting state at time $t'+4H$ and elect itself as BDR. We assume that immediately afterwards, the BDR sends out a Hello claiming BDRship. Thus, router z, arriving at time t, will receive this Hello and come out of the waiting state at time $t'+4H$. Thus, router z will be bidirectional with all the routers coming before time $t'+3H$, including the BDR. The probability that m routers arrive before the BDR ($0 \leq m \leq n-2$) is $1/n$. Given that the BDR arrives at time t' and the remaining $n-m-2$ routers (excluding router z) arrive after time t', the probability that p routers arrive between times t' and $t'+3H$ is $\binom{n-m-2}{p}(1-e^{-3\lambda H})^p(e^{-3\lambda H})^{n-m-2-p}$. Hence, the probability that q routers (excluding router z but including the BDR) arrive before time $t'+3H$ is given by $\sum_{m=0}^{q-1}\frac{1}{n}\binom{n-m-2}{q-m-1}(1-e^{-3\lambda H})^{q-m-1}(e^{-3\lambda H})^{n-q-1}$. If q routers (excluding router z but including the BDR) arrive before time $t'+3H$, router z would have established bidirectional communication with all these routers before coming out of waiting state at time $t'+4H$. Thus, router z will perform just one DR election for these q routers. For each one of the remaining $n-1-q$ routers, router z will establish bidirectional communication with the router after coming out of waiting state and consequently perform one DR election. Thus, if q routers (excluding router z but including the BDR) arrive before time $t'+3H$, router z will perform a total of $n-q$ DR elections as it establishes bidirectional communication with each router. Therefore, if $t' < t \leq t'+3H$, the expected number of DR elections performed by router z is given by $E[\#\text{ DR elections}|t' < t \leq t'+3H] = \sum_{q=0}^{n-1}(n-q)\sum_{m=0}^{q-1}\frac{1}{n}\binom{n-m-2}{q-m-1}(1-e^{-3\lambda H})^{q-m-1}(e^{-3\lambda H})^{n-q-1}$. Note that the probability that $t > t'$ is simply 0.5, and the conditional probability that $t \leq t'+3H$, given that $t > t'$, is $(1-e^{-3\lambda H})$. Hence, the probability that $t' < t \leq t'+3H$ is $0.5(1-e^{-3\lambda H})$.

$t' < t, t'+3H < t$. In this case, the BDR will come out of waiting state at time $t'+4H$, and router z will come out of waiting state some time between t and $t+H$ on receiving a BDRship claim Hello from the BDR. We assume that the time, $t+t''$, when router z comes out of waiting state is uniformly distributed between t and $t+H$. The probability that m routers come up before router z is $1/n$. One of these routers is the BDR. Among the remaining $m-1$ routers, the expected number of routers with whom router z establishes bidirectional communication by the time it comes out of waiting state at time $t+t''$ is $\frac{(m-1)t''}{H}$. Router z establishes bidirectional communication with the BDR at time $t+t''$. Thus, router z establishes bidirectional communication with the remaining $(n-m-1)+(m-1)(1-\frac{t''}{H})$ routers after coming out of waiting state and thus does a total of $1+(n-m-1)+(m-1)(1-\frac{t''}{H}) = n-1-(m-1)\frac{t''}{H}$ DR elections. Thus, the expected number of DR elections performed by router z can be expressed as $E[\#\text{ DR elections}|t' < t, t'+3H < t] = \sum_{m=0}^{n-1}\frac{1}{n}\int_0^H(n-1-(m-1)\frac{t''}{H})\frac{1}{H}dt'' = \frac{3n-1}{4}$.

Note that the probability that $t > t'$ is 0.5, and the conditional probability that $t > t'+3H$, given that $t > t'$, is $(e^{-3\lambda H})$. Hence, the probability that $t' < t, t'+3H < t$ is $0.5(e^{-3\lambda H})$.

Figure 6 shows the expected number of DR elections performed by router z as it establishes bidirectional communication with other routers in the LAN, obtained by multiplying the probability of occurrence of each scenario mentioned above with the expected number of DR elections performed in that scenario and summing up the results. Clearly, the expected number of DR elections performed increases almost linearly with the number of routers in the LAN. Also, as the mean router arrival time increases, the number of bidirectional communications established before a router comes out of waiting state decreases, and hence the expected number of DR elections performed increases.

5. The Proposed Modifications to OSPF's Interface State Machine

In the previous section, we saw that a router, on a broadcast LAN of routers, may perform a large number of DR elections in the DR/BDR settling process. Most of these elections take place as the result of the router coming out of waiting state before establishing bidirectional communication with all the other routers on the LAN. Furthermore, it is possible that multiple routers elect themselves as BDR and force more DR elections as they send their BDRship claims and later retract them. A router may come out of the waiting state before establishing bidirectional communication with all the other routers on the LAN because of a BackupSeen event or because of the static/fixed wait time duration. Multiple routers may elect themselves as BDR because of the BackupSeen event generated on receiving a Hello from DR listing no router as BDR.

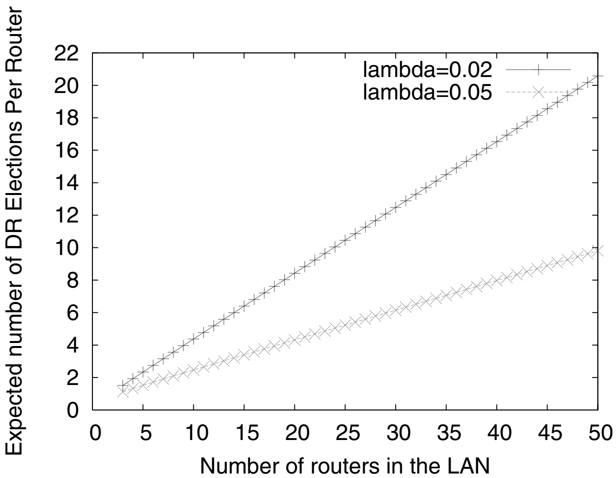

Figure 6. Expected number of designated router (DR) elections performed by a non–backup designated router (BDR) based on how many neighbors achieve bidirectional status after it comes out of waiting state

We also derived the DR/BDR settling time for the LAN to range between $max(W, 0_n)$ and 0_n+W+H $(= 0_n+5H)$. The DR/BDR settling time can be improved if the routers on the LAN could come out of waiting state soon after establishing bidirectional communication with all the other routers on the LAN rather than waiting for a fixed duration W. Further improvement can be achieved if a router sends out a Hello immediately after electing itself as DR/BDR.

Note that there is a trade-off involved in setting the value for the wait time. A small wait time duration may hasten the DR/BDR settling process but may cause a router to come out of the waiting state before establishing bidirectional communication with all the neighbors (and hence end up performing several DR elections). A large wait time duration may minimize the DR elections performed but cause a large delay in the DR/BDR settling process. It is possible to speed up the DR/BDR settling process and at the same time avoid too many DR elections if the wait time duration is adjusted dynamically as described next. The idea is to start the wait timer to fire after a small duration W (such that $H < W < 2H$) and restart the wait timer whenever a one-way Hello is received. In the following, we propose several modifications to the OSPF's interface state machine to improve its performance in terms of DR elections performed and the time required for the DR/BDR settling process.

Dynamic Wait Time Adjustment. We propose that a router must (re)start the wait timer whenever it receives a one-way Hello indicating the coming up of a new neighbor on the LAN. The wait time duration W is set to be sufficiently larger than one hello interval, thereby ensuring that the router would have established bidirectional communication with the new neighbor before the wait timer fires.

$H < W < 2H$. With the proposed restarting of the wait timer on receiving a one-way Hello, it is no longer required to set the wait time duration to be as large as four hello intervals. The wait time should just be large enough to allow a router to establish bidirectional communication with the neighbor whose one-way Hello triggered the (re)start of the wait timer. Assuming that the Hello packet losses are rare, it is sufficient that the wait time duration W is comfortably larger than one hello interval H (to take care of any randomization routers perform in determining the next time to send a Hello) but less than $2H$.

No BackupSeen Event. The BackupSeen event causes a router to detect the presence or absence of a BDR and come out of the waiting state immediately, thereby potentially speeding up the DR/BDR settling process. However, the BackupSeen event is also responsible for multiple routers electing themselves as BDR and a router coming out of the waiting state before establishing bidirectional communication with all the other "up" routers on the LAN. Thus, the BackupSeen event is responsible for most of the DR elections making up the $3n - 2$ upper bound on the number of DR elections performed by a router in the DR/BDR settling process. We propose eliminating the BackupSeen event from the interface state machine. As discussed later, the absence of the BackupSeen event does not cause any significant increase in the DR/BDR settling times because of speedup provided by dynamic wait time adjustment.

A New *Waiting2* State. If the LAN interface is in the waiting state (i.e., has not yet elected a DR or a BDR), restarting the wait timer on receiving a 1-way Hello ensures that a router does not come out of the waiting state while it has not yet established bidirectional communication with other "up" routers on the LAN. If the LAN interface is no longer in the waiting state, we would still like the router to avoid doing one DR election for every bidirectional communication it establishes with a new neighbor. We can achieve this by introducing a new *waiting2* state in the interface state machine. The new interface state machine is shown in Figure 7. Under the new state machine, the *DR/BDR/DRother* state is called the *nonwaiting* state. If the LAN interface is in the nonwaiting state, the receipt of a one-way Hello causes a transition to the *waiting2* state, and the wait timer is started. In the waiting2 state, establishing bidirectional communication with a new neighbor does not cause a DR election. The other events leading to the generation of a NeighborChange event (e.g., a router newly declaring itself as DR/BDR or taking its claim back) still cause a DR election to take place. This definition of NeighborChange event applicable during the waiting2 state can be called *NeighborChangeLite*. Receiving another one-way Hello while in the waiting2 state will cause the wait timer to be restarted. When the wait timer ultimately fires, the router does a DR election and transitions to the nonwaiting state.

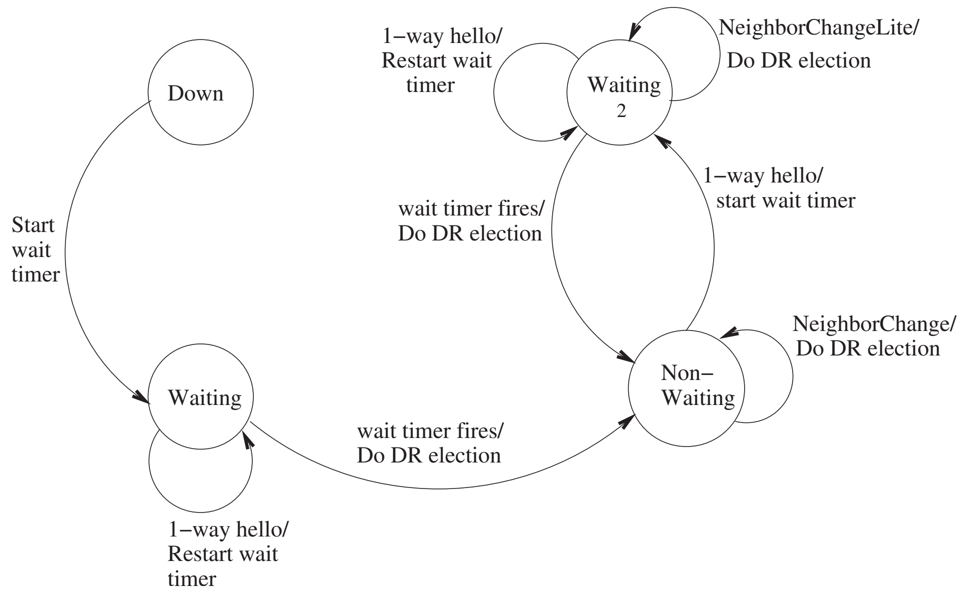

Figure 7. The proposed Interface State Machine for OSPF

Thus, under the new state machine, after coming out of the waiting state, a router does not perform one DR election for each newly bidirectional neighbor. Rather, the router performs one DR election after it has established bidirectional communication with all the neighbors from whom it has received a one-way Hello. Note that, under the new state machine, an interface's status as DR/BDR/DROther/none is an attribute of the state, called *DRAttribute*. In the waiting state, the interface has not performed a DR election, and its DRAttribute is *none*. In the nonwaiting and waiting2 states, the DRAttribute of the interface is either DR or BDR or DROther.

A Router Immediately Sends Out a Hello on Newly Electing Itself as DR/BDR. With this requirement, a router does not have to wait for up to one hello interval before sending out its claim to DR/BDRship. This provision speeds up the DR/BDR settling time for the LAN.

In the following, we analyze the performance of the new interface state machine in different scenarios.

Scenario 1: A New Router Comes Up on a LAN with Settled DR/BDR. Immediately after coming up, the new router will send out a one-way Hello to other routers on the LAN. On receiving this Hello, the other routers will start their wait timers and enter the waiting2 state. Thus, all the routers on the LAN (including the new router) will perform a DR election at time W after the arrival of the new router. By this time, the new router must have received the two-way Hellos from all the other routers on the LAN, including the current DR/BDR. Thus, in the DR election at time W, all the routers (including the new router) will (re)elect the existing DR/BDR as the DR/BDR. Thus, the settling time will be W, and only one DR election will be performed in the process. Under the old state machine, the router would have elected the existing DR/BDR as its DR/BDR within one hello interval of coming up but would have performed potentially several DR elections in the process. Since W has a value between H and $2H$, the delay in the settling time in the new scheme over the old scheme is not significant.

Scenario 2: Each Router Comes Up within Time *W* of Coming Up of the Previous Router. In this scenario, as a new router comes up and sends out a one-way Hello, all the other "up" routers will restart their wait timers. If the last router comes up at time 0_n, all the routers will come out of the waiting state at time $0_n + W$ and do a DR election. Let routers x and y have the highest and next highest ID among all the routers. In the DR election at time $0_n + W$, each router (except router x) will elect router x as both DR and BDR. Router x will select itself as DR and router y as BDR. Immediately afterwards, router x will send out a Hello claiming DRship. When other routers (including router y) receive this Hello, they will perform another DR election and elect router x as DR and router y as BDR. On electing itself as the BDR, router y will also immediately send out a Hello claiming BDRship, which will cause every other router to perform another DR election reelecting routers x and y as DR and BDR, respectively. Thus, the DR/BDR settling time for the LAN will be slightly more than $0_n + W$. Routers x and y perform two DR elections each while other routers perform three DR elections each.

Scenario 3: Each Router Does Not Come Up within Time W of the Previous Router. Consider a LAN of n routers, where the routers come up in two batches of p and $n-p$ routers, respectively. In each batch, each router (except the first one) comes up within time W of coming up of the previous router. The first router in the second batch comes up after a delay more than W of coming up of the last router in the first batch. Suppose 0_p is the time when the last router in the first batch comes up. The p routers, up by time 0_p, will come out of the waiting state at time $0_p + W$ and soon after elect two routers with the highest Router ID among themselves (say routers x and y) as DR/BDR performing two or three DR elections in the process, as discussed under scenario 2. When the first router in the second batch comes up and sends a one-way Hello, all the routers in the first batch will enter the waiting2 state. As the subsequent routers in the second batch come up and send their one- way Hellos, all the previously up routers will restart their wait timers. Suppose the last router in the second batch comes up at time 0_n. All the routers will emerge out of waiting or waiting2 state at time $0_n + W$ and perform a DR election. By this time, all the routers in the second batch must have received two-way Hellos from routers x and y claiming DRship and BDRship, respectively. Hence, in the DR election at time $0_n + W$, all the routers will (re)elect routers x and y as DR and BDR, respectively. Thus, the overall DR/BDR settling time for the LAN is $0_n + W$. Note that the routers in the first batch need to perform just one DR election for all the routers coming up in the second batch. The routers in the second batch perform just one DR election. Generalizing this scenario, suppose the routers on the LAN come up in m batches. The DR/BDR will be selected from the first batch of routers. The overall DR/BDR settling time for the LAN will still be $0_n + W$, where 0_n is the time when the last router comes up. The routers in a particular batch will need to perform just one DR election for each batch of subsequently coming up routers. Specifically, the routers in the first batch will perform $m+1$ or $m+2$ DR elections, and the routers in the tth batch ($t > 1$) will perform $m - t + 1$ DR elections.

In the following, we make certain claims about the behavior of the new interface state machine.

Claim 1. When an interface comes out of the waiting or waiting2 state, it has established bidirectional communication with all its known neighbors.

Proof. When router a receives a one-way Hello from a new neighbor router b, it restarts the wait timer. In addition, if the interface is not already in waiting or waiting2 state, it will transition to the waiting2 state. In its next Hello, router a will list router b as a neighbor. Router b receives this Hello and in its next Hello lists router a as a neighbor. Thus, router a will establish bidirectional communication with router b within one hello interval of router b's coming up. Since the wait time W is more than the hello interval H, router a would have established bidirectional communication with all its known neighbors by the time the wait timer fires, which also causes a transition out of the waiting/waiting2 state.

Claim 2. If a router, say router x, is the only router to elect itself as DR by the time its Hello claiming the DRship is received by other routers, then the other routers will not elect themselves as DR and will elect router x as DR.

Proof. Suppose router x is the only router to elect itself as DR by the time its Hello claiming the DRship is received by other routers on the LAN. On receiving this Hello, all the routers not in waiting state will do a DR election and elect router x as DR since router x is the only router with a claim to DRship. Consider the routers that are in the waiting state when they receive this Hello from router x. As each such router comes out of the waiting state, it will elect router x as DR since router x will still be the only router to claim DRship. Now consider the routers that come up after the time when router x sends out the first Hello claiming DRship. When each such router comes out of the waiting state, it would have received a two-way Hello from router x claiming DRship and would not have received any other claim to DRship. Hence, it will also elect router x as DR.

Claim 3. A router must have declared itself as the DR in its Hello, before another router elects itself as BDR.

Proof. The proof is similar to the proof for claim 2 in section 3.

Claim 4. If a router elects itself as DR for the LAN, it must have done so in its first DR election after coming out of the waiting state.

Proof. The proof is similar to the proof for claim 3 in section 3.

Claim 5. Only one router elects itself as the DR for the LAN.

Proof. Suppose router x elects itself as DR at time W_x. According to claim 4, router x must have just come out of the waiting state. So, router x must have arrived by time $0_x = W_x - W$. Router x must not have received a Hello from another router claiming DR/BDRship by time W_x (otherwise it would have elected this other router as DR and not itself). Router x will send out a Hello claiming DRship immediately after electing itself as DR (say at time $W_x + \epsilon$, where ϵ is a very small time duration). According to claim 2, if router x is the only router to have elected itself as DR by the time its Hello claiming DRship is received by other routers, then the other routers will not elect themselves as DR. Suppose the precondition for claim 2 does not hold; that is, another router z elects itself as DR at time W_z before receiving the Hello from router x claiming DRship ($W_z < W_x + \epsilon$). Using the same reasoning as above, we can claim that router z must have just come out of the waiting state at time W_z and must have arrived by time $0_z = W_z - W$. Also, router z sends out a Hello claiming DRship at time $W_z + \epsilon$, and router x must have elected itself as DR before receiving this Hello ($W_x < W_z + \epsilon$). Clearly, $|W_x - W_z| < \epsilon$ and hence $|0_x - 0_z| < \epsilon$ (i.e., routers x and z must have been "up" together for at least W seconds.) Since W is more than hello interval H, routers x and z must have established bidirectional communication with each other by the time they come out of the waiting state. Therefore, both routers x and z cannot elect themselves as DR. Assuming that router x does elect itself as DR, router z cannot elect itself as DR. Hence, the precondition for claim 2 holds, and no router other than router x elects itself as DR.

Claim 6. If a router, say router y, is the only router to elect itself as BDR by the time its Hello claiming the BDRship is received by other routers, then the other routers will not elect themselves as BDR and will elect router y as BDR.

Proof. The proof is similar to the proof for claim 2.

Claim 7. If a router elects itself as BDR for the LAN, it must have done so in its first DR election after receiving the Hello from DR claiming the DRship.

Proof. Consider a router, a, that is not in the waiting state. Router a will perform a DR election on receiving the first Hello from DR claiming the DRship. According to claim 3, router a could not have elected itself as BDR before this time. If, by this time, router a has received Hellos from one or more routers claiming the BDRship, router a will always elect one of these routers as BDR and will not claim BDRship itself now or in future (as long as one or more of these routers continue to claim themselves as BDR). If router a has not yet received a Hello declaring the sending router as BDR, it will elect the router with highest Router ID among its set of eligible routers (excluding the DR) as BDR. If at this time, router a does not elect itself as BDR, there must be another bidirectional non-DR neighbor with a higher Router ID than router a, and in any subsequent DR election, router a will not elect itself as BDR ahead of this neighbor. Similarly, it can be proved that if router a is in waiting state when it receives the first Hello from a router claiming the DRship, it will either elect itself as BDR in its first DR election on coming out of waiting state or will not elect itself as BDR then or in future.

Claim 8. Only one router, say router y, elects itself as the BDR for the LAN. Other routers will eventually elect router y as the BDR.

Proof. Suppose router x elects itself as DR and sends out the first Hello claiming DRship at time W_x. According to claim 7, no router can elect itself as BDR before this time. Suppose router y elects itself as BDR at time E_y in its first DR election after receiving a Hello from router x claiming DRship ($E_y > W_x$). Since router y does a DR election at time E_y, it must have arrived by time $0_y = E_y - W$. Router y must not have received a Hello from another router claiming BDRship by time E_y (otherwise it would not have elected itself as BDR). Router y will send out a Hello claiming BDRship immediately after electing itself as BDR (say at time $E_y + \epsilon$, where ϵ is a very small time duration). According to claim 6, if router y is the only router to elect itself as BDR by the time its Hello claiming BDRship is received by other routers, then the other routers will not elect themselves as BDR. Suppose the precondition for claim 6 does not hold; that is, another router z elects itself as BDR at time E_z before receiving the Hello from router y claiming BDRship ($E_z < E_y + \epsilon$). Using the same reasoning as above, we can claim that router z elected itself as BDR after receiving the Hello from router x claiming DRship ($E_z > W_x$), router z must have arrived by time $0_z = E_z - W$, and router z sends out a Hello claiming BDRship at time $E_z + \epsilon$ immediately after electing itself as BDR. Since router y elects itself as BDR before receiving any other claim to BDRship, $E_y < E_z + \epsilon$. Clearly, $|E_y - E_z| < \epsilon$, and hence $|0_y - 0_z| < \epsilon$ (i.e., routers y and z must have been "up" together for at least W seconds). Since W is more than hello interval H, routers y and z must have established bidirectional communication with each other by the time they elect themselves as BDR. Therefore, both routers y and z cannot elect themselves as BDR. Assuming that router y does elect itself as BDR, router z cannot elect itself as BDR. Hence the precondition for claim 6 holds, and no router other than router y elects itself as BDR.

Bound on Number of DR Elections Performed. On the basis of three scenarios discussed earlier and the claims made above, we can claim that, in a LAN of n routers, each router will perform at most $n + 2$ DR elections in the DR/BDR settling process. The first DR election will be performed as a result of the wait timer's firing. Up to $n - 1$ DR elections may be performed as the other $n - 1$ routers on the LAN come up in up to $n - 1$ "batches." If the router is not in waiting state, it may perform a DR election each on receiving claims to DR/BDRship from other routers. As per claims 5 and 8, only two such DR elections may be performed.

Bounds on DR/BDR Settling Time. With the assumptions that the time taken by a Hello packet to reach all the routers on the LAN, the time required to process/generate a Hello packet, and the time required to perform a DR election are negligible, we can claim that $0_n + W$ and $0_n + W + H$ constitute the lower and upper bounds on the DR/BDR settling time for the LAN, where 0_n is the time when the last router comes up. $0_n + W$ is the lower bound on the DR/BDR settling time for the LAN since the last router to come up at time 0_n will not come out of the waiting state before time $0_n + W$. The other routers will also come out of the waiting/waiting2 state at time $0_n + W$ and settle on DR/BDR identity soon after. Thus, in most cases, the DR/BDR settling time for the LAN is $0_n + W$. However, in some cases, it is possible that a router does not come out of the waiting/waiting2 state until time $0_n + W + H$. For example, this will happen if router z starts its wait timer on receiving the first one-way Hello from the last router at time 0_n and does not send its own Hello until after receiving the second one-way Hello from the last router at time $0_n + H$. The second one-way Hello will cause router z to restart its wait timer at time $0_n + H$. Thus, it is possible that router z does not settle on the DR/BDR identity until time $0_n + H + W$.

6. A Stochastic Analysis of DR Elections Performed by a Router under the Modified Interface State Machine

In the following, we determine the expected number of DR elections performed by a router under the modified interface state machine as it establishes bidirectional communication with its neighbors. As before, we assume that the router arrival times are exponentially distributed with parameter λ. Consider a router z in a LAN of n routers. Suppose router z is the $(n-m)$th router to arrive (i.e., m routers arrive after router z). Since the routers are equally likely to arrive in any order, the probability that m routers arrive after router z is $1/n$.

Let $N(m)$ denote the expected number of DR elections performed by router z if m routers will arrive after it. Since each one of these m routers arrives after an exponentially distributed time with parameter λ, the time interval, X, between the arrivals of router z and the next router will be exponentially distributed with parameter $m\lambda$. Clearly, if router z is the last router to arrive, it will perform just one DR election, that is, $N(0) = 1$. We now derive a closed-form expression for $N(m)$, $m \geq 1$. If the next router arrives within W seconds of router z, router z will restart its wait timer, and the expected number of DR elections it does will be $N(m-1)$. Otherwise, the expected number of DR elections can be expressed as $1 + N(m-1)$. Conditioning on X, we write

$$N(m|X = x) = \begin{cases} N(m-1) & : \quad x \leq W \\ 1 + N(m-1) & : \quad x > W \end{cases}$$

Applying the law of total probability, we obtain $N(m) = N(m-1)(1 - e^{-m\lambda W}) + (1 + N(m-1))e^{-m\lambda W} = e^{-m\lambda W} + N(m-1) = \sum_{i=0}^{m} e^{-i\lambda W} = \frac{1-e^{-(m+1)\lambda W}}{1-e^{-\lambda W}}$. Hence, the expected number of DR elections performed by router z is given by $E[\#\text{ DR elections}] = \sum_{m=0}^{n-1} \frac{1}{n} \frac{1-e^{-(m+1)\lambda W}}{1-e^{-\lambda W}} = \frac{1}{1-e^{-\lambda W}} - \frac{e^{-\lambda W}(1-e^{-n\lambda W})}{n(1-e^{-\lambda W})^2}$.

Figure 8 shows the expected number of DR elections performed by a router for $\lambda = 0.05$ (mean router arrival time 20 seconds) for different values of W as the number of routers in the LAN increases. As expected, a smaller value of W results in more DR elections. Figure 9 shows the expected number of DR elections for different λ values and fixed W (15 seconds) as the number of routers in the LAN increases. A decrease in λ means an increase in mean router arrival time as well as in the variance of router arrival times. Consequently, the probability that the next router will arrive within W seconds of the previous one decreases, which causes an increase in the expected number of DR elections. Finally, Figure 10 compares the expected number of DR elections performed by a non-BDR router as it establishes bidirectional communication with other routers before and after the proposed modifications. Clearly, the proposed modifications result in a significant decrease in the expected number of DR elections.

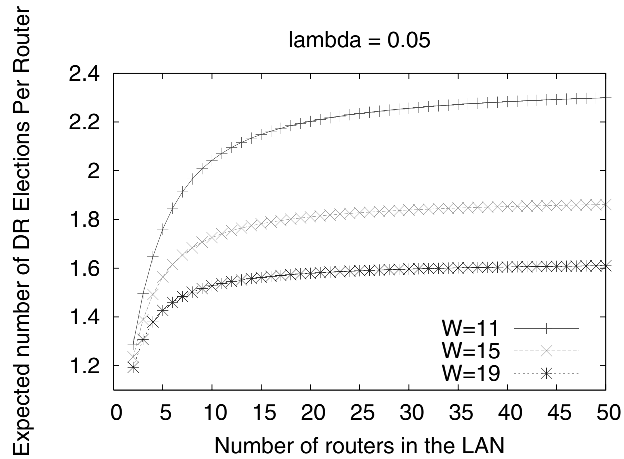

Figure 8. Expected number of designated router (DR) elections performed by a non–backup designated router (BDR) under the modified interface state machine based on when the router establishes bidirectional communication with its neighbors

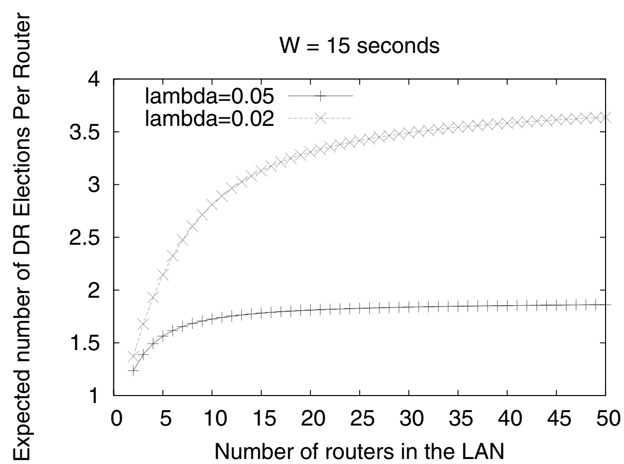

Figure 9. Expected number of designated router (DR) elections performed by a non–backup designated router (BDR) under the modified interface state machine for different mean router arrival times

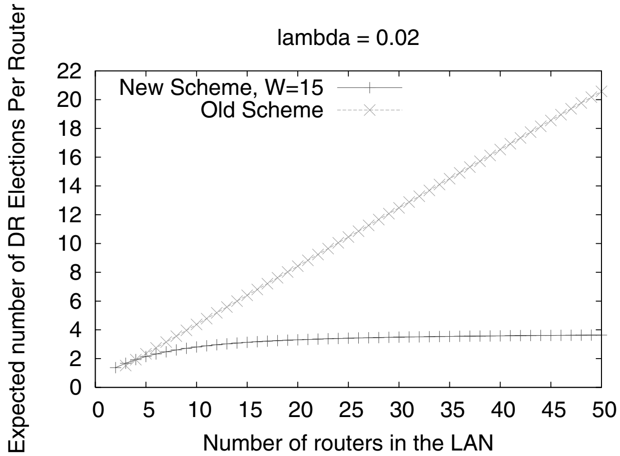

Figure 10. Expected number of designated router (DR) elections performed by a non–backup designated router (BDR) under the current (old) and modified (new) interface state machines

7. Experimental Verification

We compared the performance of the current and modified interface state machines via experiments as well. The experiments were conducted on an Ethernet LAN of eight Linux PCs running J Moy's OSPFD routing software [7]. This software already contains an implementation of OSPF's current interface state machine. We implemented the modified interface state machine in the software. By carefully selecting the times when the routers come up, we were able to create scenarios that highlight key differences in the performance of the two state machines. We used the *multixterm* script [8] running on a "control" PC to issue a command to all the router PCs simultaneously and thereby precisely control (within a few milliseconds) the router "up" times. The experiments for the modified state machine used a wait time of 15 seconds ($W = 15$). The router "up" times used in the experiments are shown in Table 1. The number of DR elections performed by each router and the DR/BDR settling times for the LAN in these experiments under the current and modified state machines are shown in Tables 2 and 3, respectively. In the following, we briefly discuss the experiments and the results. A detailed explanation of the results for each experiment can be seen in the appendix.

In experiment 1, all the routers were scheduled to come up at time 0. As per our analysis for the current state machine, routers 1 through 6 should perform three DR elections each, and routers 7 and 8 (the BDR and DR, respectively) should perform two DR elections each. The DR/BDR settling time for the LAN should be either 40 seconds or 50 seconds. A slight variation of a few milliseconds between the actual and scheduled "up" times resulted in routers 5, 7, and 8 starting a bit earlier than other routers, resulting in the observed results. Similarly, for the modified state machine, randomization in determining the next time to send Hello resulted in a few routers restarting their wait timers at time 10 on receiving a second one-way Hello from their neighbors. Thus, some routers emerge from waiting state at time 25.

Experiments 2 through 5 were designed to show the impact of the number of "batches" on the DR elections performed under the modified state machine. In experiment 2, each router comes up 12 seconds after the previous router, but clearly, all the routers still come up in one batch. Notice the significantly better performance of the modified state machine over the current state machine in this case. Also, note that router 2 performs just one DR election, while other non-DR, non-BDR routers perform three DR elections. This is because routers 8 and 7 (the DR and BDR, respectively) sent their Hellos claiming DR/BDRship immediately after electing themselves as DR/BDR (as required under the modified state machine), and router 2 was still in the waiting state when it received these Hellos. We observed the same phenomenon in several other experiments as well.

In experiment 3, each router comes up 20 seconds after the previous router. Thus, in this experiment, each router constitutes a separate batch. Hence, the modified state machine does not perform much better than the current state machine in terms of DR elections. However, the DR/BDR settling time is still better under the modified state machine. In experiments 4 and 5, the routers come up in two and three batches, respectively. The performance under the modified state machine is much better in these experiments when compared to performance under the current state machine.

Experiment 6 was designed to show the impact of multiple routers electing themselves as BDR under the current state machine. Here, each one of routers 1 through 7 initially elects itself as BDR under the current state machine. Notice the significantly reduced number of DR elections performed under the modified state machine. However, the current state machine performs better than the modified state machine in terms of DR/BDR settling time.

Experiment 7 was designed to show the impact, under the current state machine, of the routers establishing bidirectional communication with other routers after coming out of the waiting state. In this experiment, under the current state machine, routers 1 through 6 have not established bidirectional communication with any router by the time they come out of the waiting state. Hence, each one of these routers ends up performing seven DR elections as it establishes bidirectional communication with its seven neighbors on the LAN. Routers 7 and 8 are also out of waiting state when routers 1 through 6 come up and hence perform six additional DR elections as they establish bidirectional communication with routers 1 through 6. Under the modified state machine, routers 1 through 6 come out of waiting state only after establishing bidirectional communication with all the other routers on the LAN and hence perform only one DR election each. Routers 7 and 8

Table 1. Router "up" times

Exp ID	Router 1	Router 2	Router 3	Router 4	Router 5	Router 6	Router 7	Router 8
1	0	0	0	0	0	0	0	0
2	0	12	24	36	48	60	72	84
3	0	20	40	60	80	100	120	140
4	0	1	2	3	20	21	22	23
5	0	1	2	34	35	36	57	58
6	58	58	58	58	58	58	58	0
7	58	58	58	58	58	58	9	0
8	0	4	8	12	16	20	24	28
9	28	24	20	16	12	8	4	0
10	5	5	5	5	5	10	5	102
11	1	1	1	1	1	10	1	102

Table 2. Designated router (DR) elections performed and DR/backup designated router (BDR) settling time under current interface state machine

Exp ID	Router 1	Router 2	Router 3	Router 4	Router 5	Router 6	Router 7	Router 8	DR/BDR Settle Time
1	1	1	1	1	3	1	2	2	40
2	8	7	6	5	4	3	2	2	134
3	9	8	7	6	5	4	2	2	190
4	3	3	3	3	3	3	2	2	73
5	8	8	7	2	2	2	2	1	98
6	10	10	10	11	10	9	10	15	68
7	7	7	7	7	7	7	8	8	60
8	3	3	3	3	3	3	2	2	78
9	1	1	1	1	2	3	2	2	54
10	4	4	4	4	4	3	3	1	110
11	4	4	4	4	4	3	3	7	111

Table 3. Designated router (DR) elections performed and DR/backup designated router (BDR) settling time under modified interface state machine

Exp ID	Router 1	Router 2	Router 3	Router 4	Router 5	Router 6	Router 7	Router 8	DR/BDR Settle Time
1	2	2	2	2	2	2	1	2	25
2	3	1	3	3	3	3	2	2	99
3	9	7	6	5	4	3	2	1	155
4	2	3	2	3	1	1	1	1	38
5	5	4	4	2	2	2	1	1	73
6	2	2	2	2	2	2	1	3	83
7	1	1	1	1	1	1	3	3	83
8	1	3	1	1	3	1	1	2	43
9	1	1	1	1	3	1	2	2	43
10	2	4	2	4	2	3	3	1	117
11	2	2	4	4	4	3	3	1	117

perform just one additional DR election for routers 1 through 6 coming up in one "batch." However, the DR/BDR settling time is better under the current state machine as routers 1 through 6 accept routers 8 and 7 as DR and BDR, respectively, within 2 seconds of coming up. Under the modified state machine, routers 1 through 6 need to wait for up to 25 seconds before electing routers 8 and 7 as DR/BDR.

Experiments 8 and 9 cover the scenario where, under the current state machine, each router is bidirectional with other routers before coming out of the waiting state. The two experiments differ in the time when router 8 (the DR) comes out of waiting state. Under the modified state machine, in both experiments, the DR/BDR identity is settled 15 seconds after the coming up of the last router. However, under the current state machine, the time when router 8 comes out of waiting state determines the overall DR/BDR settling time for the LAN. In these experiments, the modified state machine performs better than the current state machine in terms of DR/BDR settling time. Both state machines complete the DR/BDR settling process with a small number of DR elections.

Experiments 10 and 11 cover the scenario where the first seven routers have settled on the DR/BDR identity by the time router 8 comes up. The two experiments differ in how many DR elections are performed by router 8 in its DR/BDR settling process under the current state machine. In experiment 11, under the current state machine, router 8 comes out of waiting state without establishing bidirectional communication with any router and ends up performing seven DR elections in total. Under the modified state machine, router 8 performs just one DR election in both experiments. However, the DR/BDR settling time for router 8 is shorter under the current state machine.

8. Conclusion

In this article, we analyzed the behavior of OSPF's interface state machine in a broadcast LAN environment. We showed that, under the current state machine, a router may need to perform several DR elections as it settles on the identity of DR and BDR for the LAN. Most of these elections are caused by the fact that the BackupSeen event or the fixed wait time may cause the router to come out of the waiting state before it has established bidirectional communication with all the other routers on the LAN. Furthermore, it is possible that multiple routers elect themselves initially as BDR and cause a large number of DR elections to take place as multiple BDRs stake their claims to BDRship and later retract them. Also, a fixed wait time of 40 seconds may cause the DR/BDR settling process to get unnecessarily delayed. Based on these observations, we suggested the following modifications to OSPF's interface state machine:

1. Wait timer should be restarted whenever a new one-way Hello is received.
2. Wait time should just be large enough to allow the router to establish bidirectional communication with the neighbor whose one-way Hello triggered the restart of the wait timer.
3. BackupSeen event should be eliminated.
4. A new *waiting2* state should be introduced in the interface state machine to allow a router that is no longer in the waiting state to avoid doing DR elections on establishing a new bidirectional communication while some of the neighbor states are not yet bidirectional.
5. A router should immediately send a Hello on newly electing itself as DR or BDR.

We analyzed the performance of the interface state machine before and after the modifications. We obtained upper bounds on the number of DR elections performed and the DR/BDR settling time in both cases. We performed a stochastic analysis to obtain an expected number of DR elections performed by a router in both cases, assuming that router arrival times are exponentially distributed. We also compared the performance in both cases via carefully controlled test-bed experiments. Based on this comprehensive comparison, we can state that the suggested modifications result in fewer DR elections and, in most cases, faster conclusion of the DR/BDR settling process.

9. Appendix: Detailed Explanation for Each Experiment

Experiment 1. In this experiment, all the routers were supposed to come up at time 0. Rather than coming up simultaneously, the routers come up within a few milliseconds of each other.

Current state machine: In the experiment performed using the current state machine, routers 5, 7, and 8 come up a little earlier than other routers. All the routers establish bidirectional communication with each other by the time they come out of the waiting state. Since routers 5, 7, and 8 start a bit earlier than other routers, the other routers are still in the waiting state when the following events take place: (1) routers 5, 7, and 8 come out of the waiting state; (2) router 8 elects itself as DR and immediately afterwards sends out a Hello claiming DRship; and (3) router 7 performs a DR election on receiving this Hello, elects itself as BDR, and sends out a Hello claiming BDRship. The other routers come out of the waiting state on receiving the Hello from router 7 claiming BDRship and perform their only DR election. Router 5 comes out of the waiting state at approximately the same time as routers 7 and 8 and hence does its first DR election as the result of its wait timer's firing. It does two more DR elections as it receives the first Hellos from routers 7 and 8 claiming BDRship and Drship, respectively. The DR/BDR settling time for the LAN in this experiment was 40 seconds.

Modified state machine: At time 10, router 2 receives a second one-way Hello from router 8. This is because router 8 came up a little while after router 2 and did not receive a Hello from router 2 until after sending its Hello at time 10. The one-way Hello from router 8 causes router 2 to restart its wait timer at time 10. The same thing happens to router 4 and router 7, with the difference that router 4 receives the one-way Hello from router 6 and router 7 receives from router 1. So routers 2, 4, and 7 come out of the waiting state at time 25, while routers 1, 3, 5, 6, and 8 come out of the waiting state at time 15. All routers establish bidirectional communication with each other by the time they come out of the waiting state. Router 8 comes out of the waiting state first, elects itself as DR, and immediately sends out a Hello claiming DRship. At this time, the other routers are still in the waiting state. Then routers 1, 3, 5, and 6 come out of the waiting state and perform their first DR election. They elect router 8 as DR and router 7 as BDR. At time 25, routers 2, 4, and 7 come out of the waiting state. Router 7 elects itself as BDR and sends out a Hello claiming BDRship instantly. On receiving this Hello, all other routers perform their second DR election and reelect router 8 as DR and router 7 as BDR. So under the modified interface state machine, the DR/BDR settling time for the LAN is 25 seconds, and each router performs one or two DR elections in the process.

Experiment 2. *Current state machine:* Under the current state machine, router 1 has established bidirectional communication only with routers 2 and 3 by the time it comes out of waiting state at time 40 seconds. It will perform five more DR elections as it establishes bidirectional communication with routers 4 through 8. It will perform two more DR elections when it receives the Hellos from routers 8 and 7 claiming DRship and BDRship, respectively. Thus, router 1 will perform eight DR elections in total. Similarly, it can be seen that routers 2, 3, 4, 5, and 6 will perform 7, 6, 5, 4, and 3 DR elections, respectively. Routers 7 and 8 perform two DR elections each, one on coming out of waiting state and the other on receiving the first DR/BDR claim from each other. Router 8 comes out of the waiting state at time 124, electing itself as DR and router 7 as BDR. It sends out the first Hello claiming DRship at time 134 seconds. On processing this Hello, every other router (re)elects routers 8 and 7 as DR and BDR, respectively. Thus, the DR/BDR settling time for the LAN is 134 seconds.

Modified state machine: In this experiment, all the routers come up in one "batch," with each router coming up 12 seconds after the previous router. The last router comes up at time 84 seconds. Hence, under the modified interface state machine, the DR/BDR settling time for the LAN is 99 seconds, and each router performs one, two, or three DR elections in the process. Router 1 comes up at time 0 and restarts its wait timer after every 12 seconds due to the one-way Hello from a newly coming up router. When it comes out of waiting state at time 99 seconds, router 1 performs its first DR election and elects router 8 both as DR and BDR. Routers 3, 4, 5, 6, and 7 also behave similarly. Router 8 comes up at time 84 seconds and comes out of the waiting state at time 99 seconds. Router 8 performs the DR election and elects itself as DR and router 7 as BDR. It sends out a Hello claiming DRship immediately. On receiving this Hello, routers 1, 3, 4, 5, 6, and 7 perform their second DR election and elect router 7 as BDR. On electing itself as BDR, router 7 instantly sends out a Hello claiming BDRship. Routers 1, 3, 4, 5, 6, and 8 receive this Hello and perform another DR election. This DR election does not change DR and BDR identity. Router 2 is still in the waiting state when it receives the Hellos from routers 8 and 7 claiming DRship and BDRship, respectively. Router 2 performs a DR election on firing of its wait timer and elects router 8 as DR and router 7 as BDR. In summary, routers 1, 3, 4, 5, and 6 perform three DR elections; routers 7 and 8 perform two DR elections; and router 2 performs one DR election.

Experiment 3. *Current state machine:* Under the current state machine, router 1 has established bidirectional communication only with router 2 by the time it comes out of waiting state at time 40 seconds. It will perform six more DR elections as it establishes bidirectional communication with routers 3 through 8. It will perform two more DR elections when it receives the Hellos from routers 8 and 7 claiming DRship and BDRship, respectively. Thus, router 1 will perform nine DR elections in total. Similarly, it can be seen that routers 2, 3, 4, 5, and 6 will perform 8, 7, 6, 5, and 4 DR elections, respectively. Routers 7 and 8 perform two DR elections each, one on coming out of waiting state and the other on receiving the first DR/BDR claim from each other. Router 8 comes out of the waiting state at time 180 seconds, electing itself as DR and router 7 as BDR. It sends out the first Hello claiming DRship at time 190 seconds. On processing this Hello, every other router elects routers 8 and 7 as DR and BDR, respectively. Thus, the DR/BDR settling time for the LAN is 190 seconds.

Modified state machine: In this experiment, routers come up in eight batches, with each router coming up 20 seconds after the previous router. The last router comes up at time 140 seconds. Hence, under the modified interface state machine, the DR/BDR settling time for the LAN is 155 seconds. Router 1 performs nine DR elections, router 2 performs eight DR elections, and so on. Router 8 performs one DR election. Router 1 comes up at time 0 and comes out of the waiting state at time 15 seconds. Router 1 performs the first DR election and elects itself as DR with no BDR. At time 20 seconds, router 2 comes up and sends out a Hello. On receiving this one-way Hello from router 2, router 1 goes into waiting2 state and comes out of it at time 35 seconds. In the second DR election, router 1 elects itself as DR and router 2 as BDR. Router 1 will go into waiting2 state and perform DR election six more times when routers 3 through 8 come up. All these DR elections will not change DR and BDR. Another DR election is performed by router 1 when it receives the first BDRship claim from router 2 at time 35 seconds. Thus, router 1 performs nine DR elections in total. Router 2 performs its first DR election on coming out of the waiting state at time 35 seconds and elects router 1 as DR, itself as BDR. Then, similar to router 1's behavior, router 2 will perform six more DR elections as routers 3 through 8 come up. We can explain the behavior of other routers similarly. Routers 3, 4, 5, 6, and 7 will perform 6, 5, 4, 3, and 2 DR elections, respectively. Router 8 performs one DR election on coming out of waiting state.

Experiment 4. *Current state machine:* Under the current state machine, all the routers have established bidirectional communication with other routers when they come out of the waiting state. On coming out of the waiting state, routers 1, 2, 3, 4, 5, 6, and 7 elect router 8 as both DR and BDR, while router 8 elects itself as DR and router 7 as BDR. Routers 1, 2, 3, 4, 5, 6, and 7 perform another DR election when they receive the first Hello claiming DRship from router 8. They reelect router 8 as DR and router 7 as BDR. Routers 1, 2, 3, 4, 5, 6, and 8 will perform one more DR election on receiving the first Hello claiming BDRship from router 7. Since the other routers perform the DR election on receiving the DRship claim from router 8 at time 73 seconds, the DR/BDR settling time for the LAN is 73 seconds.

Modified state machine: In experiment 4, different routers come up in two distinct batches, one after the other. Among the routers in the first batch (routers 1, 2, 3, and 4), router 4 is the first one to come out of the waiting state. Router 4 elects itself as DR and router 3 as BDR. Router 4 sends out a Hello claiming DRship immediately afterwards. Router 2 and router 3 receive this Hello just before they come out of the waiting state. So when they come out of waiting state, router 2 and router 3 elect router 4 as DR and router 3 as BDR. Router 3 sends out a Hello claiming BDRship immediately afterwards. On receiving this Hello from router 3, router 2 and router 4 perform another DR election. At this time, router 1 is still in the waiting state. When it comes out of the waiting state, router 1 performs one DR election and elects routers 4 and 3 as DR and BDR, respectively. Thus, before the coming up of the second batch, routers 1, 2, 3, and 4 perform one, two, one, and two DR elections, respectively. With the coming up of the second batch routers, routers in the first batch transition to waiting2 state. At time 38 seconds, all the routers come out of the waiting/waiting2 state and perform one DR election. Each router (re)elects router 4 as DR and router 3 as BDR. Hence, the DR/BDR settling time for the LAN is 38 seconds.

Experiment 5. *Current state machine:* Under the current state machine, routers 1, 2, and 3 have not established bidirectional communication with routers 4, 5, and 6 by the time they come out of the waiting state. On coming out of the waiting state (at time 40 seconds), routers 1 and 2 elect router 3 as both DR and BDR, while router 3 elects itself as DR and router 2 as BDR. Soon afterwards, routers 1, 2, and 3 establish bidirectional communication with routers 4, 5, and 6 and do three more DR elections in the process. After establishing bidirectional communication with these routers, routers 1, 2, and 3 consider router 6 to be the BDR. Some time afterwards, router 3 sends out its first Hello claiming Drship, causing routers 1 and 2 to perform one more DR election. Routers 4, 5, and 6 do not perform any DR election on receiving this Hello since they are still in the waiting state at that time. Routers 4, 5, and 6 are still in the waiting state when routers 7 and 8 come up. Routers 1, 2, and 3 perform two more DR elections each as they establish bidirectional communication with routers 7 and 8. In these DR elections, they also elect router 8 to be the new BDR. When routers 4, 5, and 6 come out of the waiting state, they have established bidirectional communication with routers 7 and 8 and hence elect router 3 as DR and router 8 as BDR. Finally, routers 7 and 8 come out of the waiting state and perform one DR election, electing router 3 as DR and router 8 as BDR. Router 8 comes out of waiting state at time 98 seconds. Thus, the DR/BDR settling time for the LAN is 98 seconds. Every other router performs one more DR election when router 8 sends out the first Hello claiming BDRship. This election does not change the DR and BDR identity.

Modified state machine: In experiment 5, different routers come up in three distinct batches, one after the other. When routers in the first batch (routers 1, 2, and 3) come out of the waiting state, the routers in the second batch (routers 4, 5, and 6) and the third batch (routers 7 and 8) are yet to come up. On coming out of the waiting state, router 1 and router 2 elect router 3 as both DR and BDR, while router 3 elects itself as DR and router 2 as BDR. Router 3 sends out a Hello claiming DRship immediately. On receiving this Hello, routers 1 and 2 perform another DR election and reelect router 2 as BDR. Router 2 sends out a Hello claiming BDRship right after the election. Routers 1 and 3 do one more DR election on receiving this Hello. Thus, before the coming up of the second batch, routers 1, 2, and 3 perform three, two, and two DR elections, respectively. At time 34 seconds, router 4 comes up. On receiving the one-way Hello from router 4, routers in the first batch transition into waiting2 state. These routers restart their wait timers on receiving the one-way Hello from router 5 and router 6. Routers 1, 2, 3, 4, 5, and 6 come out of the waiting2/waiting state at time 51 seconds and perform a DR election. All the routers elect router 3 as DR and router 2 as BDR. The same thing happens when routers in the third batch (router 7 and router 8) come up. Routers in the first and second batches transition into the waiting2 state at time 57 seconds and come out at time 73 seconds. Router 7 and router 8 also come out of the waiting state at time 73 seconds and perform one DR election, electing router 3 as DR and router 2 as BDR. So the DR/BDR settling time of the LAN is 73 seconds.

Experiment 6. *Current state machine:* Under the current state machine, router 8 is the only router on the LAN when it comes out of the waiting state at time 40 seconds and elects itself as DR with no BDR. The other routers come up at time 58 seconds and send out their first Hellos. Router 8 receives these Hellos and lists the other routers as neighbors in its next Hello at time 60 seconds. On receiving this Hello, the other routers generate a BackupSeen event and elect themselves as BDR since they are not yet bidirectional with any other router on the LAN except router 8, the DR. As these routers receive the other claims to BDRship, they will elect the router with the highest Router ID among themselves (router 7) as BDR. It turns out that routers 2, 4, 6, and 7 still consider themselves as BDR when they send their next Hello. Receiving each one of these Hellos will cause the other routers to perform a DR election. In their next Hellos, routers 2, 4, and 6 will take back their claims to BDRship and cause the routers receiving these Hellos to perform one more DR election per Hello. The routers will perform more DR elections as they establish bidirectional communication with other routers on the LAN. In many cases, a router, say router *a*, sends a two-way Hello to another router, say router *b*, which performs a DR election as a result, (re)elects itself or router *a* as DR/BDR, and initiates adjacency establishment with router *a*. On receiving a *Database Description* packet from router *b*, router *a* elevates the neighbor state associated with router *b* to bidirectional status and performs a DR election as a result. Later on, router *a* does another DR election as it receives a Hello from router *b* claiming DR/BDRship. Let us examine the DR elections performed by routers 2 and 8 in more detail. Router 2 does one DR election on coming out of waiting state on receiving the Hello from router 8 claiming DRship with no router mentioned as BDR. Three DR elections are performed as it establishes bidirectional communication with routers 3, 4, and 5 on receiving *Database Description* (DD) packets from them. Three more DR elections are performed as router 2 receives Hellos from routers 4, 6, and 7 claiming BDRship. One DR election is performed as router 1 sends its first two-way Hello. Finally, two more DR elections are performed as routers 4 and 6 take back their claims to BDRship. The DR elections performed by routers 1 and 3 through 7 can be explained similarly. Router 8 does 15 DR elections, as described next. First DR election is performed as its wait timer fires. The other routers, when they receive their first Hello from router 8 claiming DRship, elect themselves as BDR and hence initiate adjacency establishment with router 8 by sending it their DD packets. On receiving these packets, router 8 associates bidirectional status with these seven routers and performs seven more DR elections. Four DR elections are performed as routers 2, 4, 6, and 7 claim BDRship, and three DR elections are performed as routers 2, 4, 6 take their claims back. Let's consider the DR/BDR settling time. Routers 1 through 6 and 8 settle on the DR and BDR identity (router 8 as DR and router 7 as BDR) at time 68 seconds after they receive the Hello claiming BDRship from router 7. The DR/BDR settling time for router 7 is 60 seconds (i.e., when it comes out of the waiting state). So the DR/BDR settling time for the LAN is 68 seconds.

Modified state machine: In this experiment, only router 8 comes up at time 0, while all other routers come up at time 58 seconds. At time 15 seconds, router 8 comes out of the waiting state and perform the first DR election, electing itself as DR with no BDR. At time 58 seconds, all other routers come up. On receiving one-way Hellos from these routers, router 8 transitions into the waiting2 state. Due to the differences between scheduled and actual router up times, router 1 comes up a little earlier than router 7. So, router 7 does not receive its first Hello from router 1 until time 68 seconds. Receiving a one- way Hello from router 1 at time 68 seconds causes router 7 to restart its wait timer. Thus, routers 1, 2, 3, 4, 5, 6, and 8 come out of the waiting/waiting2 state at time 73 seconds, but router 7 comes out of the waiting state at time 83 seconds. On coming out of the waiting state at time 73 seconds, routers 1, 2, 3, 4, 5, 6, and 8 perform one DR election, electing router 8 as DR and router 7 as BDR. At time 83 seconds, router 7 comes out of the waiting state and elects itself as BDR. Router 7 sends out a Hello claiming BDRship immediately. On receiving this Hello, routers 1, 2, 3, 4, 5, 6, and 8 perform one more DR election. Thus, routers 1, 2, 3, 4, 5, and 6 perform two DR elections in total, router 7 does one DR election, and router 8 does three DR elections. The DR/BDR settling time for the LAN is 83 seconds.

Experiment 7. *Current state machine:* Under the current state machine, router 8 comes out of the waiting state at time 40 seconds and elects itself as DR and router 7 (its only bidirectional neighbor) as BDR. Router 7 comes out of the waiting state at time 49 seconds and elects router 8 as both DR and BDR. At time 50 seconds, router 8 will send its first claim to Drship, causing router 7 to perform a DR election, electing itself as BDR. The other routers, routers 1 through 6, come up at time 58 seconds and come out of the waiting state at time 59 seconds on receiving a Hello from router 7 claiming BDRship. Router 8 also performs a DR election on receiving this Hello. Soon afterwards, routers 1 through 6 will receive a Hello from router 8 claiming DRship at time 60 seconds. Thus, routers 1 through 6 elect routers 8 and 7 as DR and BDR, respectively. They also initiate adjacency establishment with routers 8 and 7, causing routers 8 and 7 to establish bidirectional communication with them and perform six DR elections, one for each router. Thus, routers 8 and 7 end up doing eight DR elections each. Routers 1 through 6 will perform five more DR elections as they establish bidirectional communication with their remaining five neighbors. Thus, routers 1 through 6 end up doing seven DR elections each. The DR/BDR settling time for the LAN is 60 seconds.

Modified state machine: In this experiment, router 8 comes up at time 0, router 7 comes up at time 9 seconds, and all the other routers come up at time 58 seconds. Router 8 restarts its wait timer when it receives the one-way Hello from router 7 at time 9 seconds. Router 8 comes out of the waiting state at time 24 seconds and elects itself as DR and router 7 as BDR. Router 8 sends out a Hello claiming DRship immediately afterwards. Router 7 also comes out of the waiting state at time 24 seconds and elects router 8 as both DR and BDR. On receiving the Hello from router 8, router 7 does its second DR election and elects router 8 as DR and itself as BDR. Router 7 sends out a Hello claiming BDRship immediately afterwards. On receiving this Hello, router 8 performs a DR election. At time 58 seconds, other routers come up. Routers 7 and 8 go into the waiting2 state when they receive the one-way Hellos from other routers. All routers come out of the waiting/waiting2 state and perform a DR election at time 73 seconds, except router 2. As it happened in experiment 6, rather than coming up simultaneously, router 1 comes up a little earlier than router 2. At time 68 seconds, router 2 receives its first one-way Hello from router 1, which restarts router 2's wait timer. Router 2 comes out of the waiting state at time 83 seconds and does a DR election. Thus, the DR/BDR settling time for the LAN is 83 seconds.

Experiment 8. *Current state machine:* Under the current state machine, routers 1 through 7 have established bidirectional communication with all other routers by the time they come out of waiting state at times 40 to 64 seconds. They elect router 8 as both DR and BDR. Router 8 comes out of the waiting state at time 68 seconds and elects itself as DR with router 7 as BDR. At time 78 seconds, router 8 sends out its first Hello claiming DRship. On receiving this Hello, routers 1 through 7 perform one DR election and reelect router 8 as DR and router 7 as BDR. At time 84 seconds, router 7 sends out its first Hello claiming BDRship. On receiving this Hello, routers 1 through 6 and 8 perform another DR election. But this election does not change the DR and BDR identity for the routers. Thus, the DR/BDR settling time for the LAN is 78 seconds.

Modified state machine: In this experiment, all the routers come up in one "batch," with each router coming up 4 seconds after the previous router. The last router comes up at time 28 seconds. Hence, under the modified interface state machine, the DR/BDR settling time for the LAN is 43 seconds, and each router performs one, two, or three DR elections in the process. Router 1 comes up at time 0 and restarts its wait timer every 4 seconds due to the one-way Hello from a newly up router. Routers 2, 3, 4, 5, 6, and 7 behave similarly. Router 8 comes up at time 28 seconds and comes out of the waiting state at time 43 seconds. Router 8 performs a DR election and elects itself as DR and router 7 as BDR. It sends out a Hello claiming DRship immediately. When router 7 receives this Hello, it is still in the waiting state. Soon after, router 7 comes out of the waiting state and elects router 8 as DR and itself as BDR. Router 7 also sends out a Hello claiming BDRship immediately. On receiving this Hello, router 8 performs one more DR election. Routers 1, 3, 4, and 6 also receive the Hellos claiming DRship/BDRship from router 8/router 7 before they come out of the waiting state. So when they come out, they also elect router 8 as DR and router 7 as BDR. Routers 2 and 5 come out of the waiting state before they receive the Hellos from router 8 and router 7. Thus, they perform three DR elections.

Experiment 9. *Current state machine:* Under the current state machine, the DR/BDR settling time of the LAN is 54 seconds, as discussed next. Router 8 comes up at time 0. When it comes out of the waiting state at time 40 seconds, it elects itself as DR and router 7 as BDR. Then at time 44 seconds and 48 seconds, router 7 and router 6 come out of the waiting state, respectively, and both elect router 8 as DR and BDR. At time 50 seconds, routers 7 and 6 receive the first Hello claiming DRship from router 8. This triggers a DR election. They reelect router 8 as DR and router 7 as BDR. At time 52 seconds, router 5 comes out of the waiting state and performs a DR election. At time 54 seconds, routers 1 through 6 and 8 all receive the first Hello claiming BDRship from router 7. On receiving this Hello, routers 5, 6, and 8 perform another DR election, and routers 1, 2, 3, and 4 come out of the waiting state due to the BackupSeen event. Thus, these routers perform one more DR election, and the DR/BDR settling time for the LAN is 54 seconds.

Modified state machine: In this experiment, each router comes up 4 seconds after the previous router. Thus, all the routers come up in one batch. But this time, the coming-up sequence is 8 and 7 through 1. The DR/BDR settling times and the reasons for DR elections are same as in experiment 8. But this time, routers 1, 2, 3, 4, and 6 receive the Hellos claiming DRship/BDRship from router 8/router 7 just before they come out of the waiting state. Thus, these routers perform just one DR election. Router 5 receives the Hellos after coming out of the waiting state, so it performs three DR elections. Router 7 also comes out of the waiting state before receiving the Hello from router 8. Hence, router 7 performs two DR elections.

Experiment 10. *Current state machine:* Under the current state machine, routers 1 through 5 and 7 come out of the waiting state at time 45 seconds. Routers 1 through 5 elect router 7 as both DR and BDR, while router 7 elects itself as DR and router 6 as BDR. Router 6 comes out of the waiting state at time 50 seconds and elects router 7 as both DR and BDR. At time 55 seconds, routers 1 through 6 receive the first Hello claiming DRship from router 7. They perform a DR election and reelect router 7 as DR and router 6 as BDR. At time 60 seconds, routers 1 through 5 and 7 receive the first Hello claiming BDRship from router 6 and perform one more DR election. Thus, before the coming up of router 8, routers 1 through 5 perform three DR elections and routers 6 and 7 perform two DR elections. Then router 8 comes up at time 102 seconds and immediately sends out a Hello. The other routers in the LAN receive this Hello and hence will list router 8 as a neighbor in their next Hello. Routers 1 through 5 and 7 will send their next Hello around time 105 seconds. On receiving these Hellos, router 8 will establish bidirectional communication with these routers but will not perform any DR election since it is still in the waiting state. At time 110 seconds, router 6 will send out a Hello, causing router 8 to generate a BackupSeen event and perform a DR election. In this DR election, router 8 will select routers 7 and 6 as DR and BDR, respectively. No more DR elections will be performed by router 8 since it has already established bidirectional communication with the other routers on the LAN. Thus, the DR/BDR settling time of the LAN is 110 seconds.

Modified state machine: In this experiment, router 8 is the last router to come up at time 102 seconds. By this time, the other routers in the LAN have already elected routers 7 and 6 as DR and BDR, respectively. Let's first consider the DR elections done by routers 1 through 7. Routers 1 through 5 and 7 come up at time 5 seconds, while router 6 comes up at time 10 seconds. Routers 2, 4, 6, and 7 come out of the waiting state at time 25 seconds. Routers 2, 4, and 6 elect router 7 both as DR and BDR, while router 7 elects itself as DR and router 6 as BDR. Router 7 sends out a Hello claiming DRship immediately. On receiving this Hello, routers 2, 4, and 6 do another DR election and reelect router 7 as DR and router 6 as BDR. Router 6 also sends out a Hello claiming BDRship immediately. This Hello triggers another DR election in routers 2, 4, and 7. Routers 1, 3, and 5 have received the DRship Hello from router 7 and the BDRship Hello from router 6 before they come out of the waiting state. Thus, they perform just one DR election and elect router 7 as DR and router 6 as BDR. At time 102 seconds, routers 1 through 7 receive a one-way Hello from router 8, and all of them transition into the waiting2 state. All the routers, including router 8, come out of the waiting2/waiting state at time 117 seconds and perform a DR election. Hence, the DR/BDR settling time for the LAN is 117 seconds. In total, routers 1, 3, and 5 perform two DR elections; routers 2 and 4 perform four DR elections; routers 6 and 7 perform three DR elections; and router 8 performs one DR election.

Experiment 11. *Current state machine:* In this experiment, router 8 comes up at time 102 seconds. It does its first DR election at time 110 seconds on receiving a Hello from router 6 (the BDR) and elects router 6 as both DR and BDR. Router 8 does six more DR elections at time 111 seconds on establishing bidirectional communication with routers 1 through 5 and 7 (on receiving their two-way Hellos). In the DR election triggered by the Hello from router 7, router 8 elects router 7 as DR and router 6 as BDR. Thus, in this experiment, router 8 does seven DR elections, and the DR/BDR settling time is 111 seconds.

Modified state machine: In this experiment, router 8 is the last router to come up at time 102 seconds. Routers 1 through 5 and 7 come up at time 1 second, and router 6 comes up at time 10 seconds. The DR elections performed in this experiment can be explained in a similar manner as in experiment 10. Since router 1 and router 2 receive the Hellos claiming DRship/BDRship from routers 7 and 6 before they come out of the waiting state, the two routers perform just one DR election before the coming up of router 8. Routers 3, 4, and 5 come out of the waiting state before receiving the DR/BDRship claims from routers 7 and 6. Hence, they perform three DR elections before the coming up of router 8 at time 102 seconds. Routers 6 and 7 perform two DR elections before time 102 seconds, as was the case in experiment 10. On receiving a one-way Hello from router 8 at time 102 seconds, routers 1 through 7 enter the waiting2 state. All the routers come out of the waiting2/waiting state at time 117 seconds, performing one more DR election and (re)electing routers 7 and 6 as DR and BDR, respectively. Hence, the DR/BDR settling time in this experiment is 117 seconds.

10. References

[1] Moy, J. 1998. OSPF version 2. Request for Comments (Standards Track) RFC 2328, Internet Engineering Task Force, April.

[2] Garcia-Molina, H. 1982. 'Elections in a distributed computing system. *IEEE Transactions on Computers* 31:48-59.

[3] Fredrickson, N., and N. Lynch. 1987. Electing a leader in a synchronous ring. *Journal of the ACM* 34:98-115.

[4] Huang, Y., and P. McKinley. 1997. Group leader election under link-state routing. In *Proceedings of ICNP'1997*, pp. 95-104.

[5] Lynch, N. 1996. *Distributed algorithms*. San Francisco: Morgan Kaufmann.

[6] Goyal, M., W. Xie, S. H. Hosseini, and K. Vairavan. 2005. Improving OSPF dynamics over broadcast LAN. In *Proceedings of IEEE MASCOTS'2005*.

[7] Moy, J. Accessed 27th April 2006. OSPFD routing software resources. http://www.ospf.org

M. Goyal *obtained his PhD in Computer and Information Science from the Ohio State University in 2003. He is currently an Assistant Professor of Electrical Engineering and Computer Science at the University of Wisconsin-Milwaukee. His research interests include performance evaluation of computer networking protocols.*

W. Xie *is a PhD student at the University of Wisconsin-Milwaukee, Department of Electrical Engineering and Computer Science, Milwaukee, USA.*

S. H. Hosseini *obtained his PhD in Electrical Engineering from the University of University of Iowa. He is currently an associate Profssor of Electrical Engineering and Computer Science at the University of Wisconsin-Milwaukee. His research interests include Computer Networks, Computer Architecture, Fault-Tolerance, and Performance Evaluation.*

K. Vairavan *obtained his PhD in Electrical Engineering from the University of Notre Dame in 1968. He is currently a Professor of Electrical Engineering and Computer Science at the University of Wisconsin-Milwaukee. His research interests have spanned many areas including computer networks.*

D. Rohm *is a graduate student at the University of Wisconsin-Milwaukee studying Computer Science, and an Instructor / Unix Systems Administrator at St. Norbert College in De Pere, Wisconsin.*