Designing New Data Replication Strategies Automatically

Syed Mohtashim Abbas Bokhari(✉) and Oliver Theel

Department of Computer Science, University of Oldenburg, Oldenburg, Germany
{syed.mohtashim.abbas.bokhari,oliver.theel}@uni-oldenburg.de

Abstract. A distributed system is a paradigm indispensable to the current world due to countless requests with every passing second. In distributed systems, reliability is of extreme importance. In this regard, data replication plays a vital role in making systems more reliable by increasing the availability of the access operations at a lower cost. However, availability and cost both cannot be achieved at the same time. Certainly, there are compromises between these objectives, thereby making different application-scenarios that may not be easily satisfied by contemporary strategies. This requires designing new strategies and the question still stands which strategy is the best for a given scenario or application class assuming a certain workload, its distribution across a network, availability of the individual replicas, and cost of the access operations. For this, the research exploits the heterogeneity between the strategies to generate new data replication strategies automatically through genetic programming. It uses and extends this genetic programming-based automatic mechanism to subsequently demonstrate its usefulness by reducing the cost significantly while not comprising too much on the availabilities of the access operations. It generates replication strategies there are innovative and such combinations have not been explored yet.

Keywords: Distributed systems · Fault tolerance · Data replication · Quorum protocols · Operation availability · Operation cost · Voting structures · Optimization · Machine learning · Evolutionary strategies · Genetic programming

1 Introduction

To provide highly available data access operations is a widely discussed prevalent problem in computer science. Relying on a single replica significantly confines the availability of the data. Therefore, the increase in the number of replicas to store the data objects is inevitable, which, when smartly applied, increases the availability of the data object and makes it more fault-tolerant. Because now, it can be accessed by approaching other replicas, too. But then the challenge comes up of managing those replicas and maintain consistency so that replicas always yield correct values [7]. The goal of the operations is also to behave in a replicated system the same as they would do in a non-replicated system. This is known as one-copy serializability (1SR) [1]. As for this, these replicas are managed by protocols known as data replication strategies (DRSs).

A. P. Rocha et al. (Eds.): ICAART 2020, LNAI 12613, pp. 308–331, 2021.
https://doi.org/10.1007/978-3-030-71158-0_15

These strategies impose a threshold of a minimal number of replicas known as read quorum (rq) and write quorum (wq) to be accessed to perform the preferred access operations. These access operations are either a read or a write operation. The decisions to choose suitable DRSs are trade-offs between choosing various quality metrics such as load, capacity, availability [2], scalability, and cost [3]. The availabilities of read and write operations are optimally point symmetrical to each other [4]. For instance, an increased availability for a write operation would compromise the availability of a read operation to a certain extent and vice versa. It is more like the same case with the cost of the read and write operations, too. The questions arise that what are those compromises, to what extent particular values can be compromised, and at the expense of what? These compromises could be highly application-specific and comprised of many scenarios, which will be discussed further in Sect. 2. This research intends to provide application-optimized DRSs to fulfill such specified scenarios. In this regard, this research is an extension of our research [5,7] particularly to demonstrate new instances by tweaking the algorithm slightly. It generates replication strategies there are innovative and such combinations have not been explored yet, which could significantly reduce the cost. This automatic mechanism evolves strategies as computer programs over many generations of evolution and also provides desirable trade-offs between the availability and cost of the access operations while restricting total nodes to the desirable limit. This can be visualized by a 3D representation, i.e., given in Fig. 17, where trade-offs are overt and relevant strategies can be easily picked at run-time. In this regard, details on a multi-objective optimization approach to data replication in distributed systems can be found here [6].

The paper is written as follows. Section 2 specifies and discusses the problem statement. Section 3 discusses the state-of-the-art DRSs and other contemporary approaches to address the problem and their limitations. Section 4 defines the fault model, describes the adopted methodology to approach the problem, and argues about the reason for picking this approach over others. Section 5 states the implementation aspects of the research. Section 6 presents the results and their comparisons, followed by a conclusion.

2 Problem Statement

The problem is illustrated by a triangle given in Fig. 1 where the consistency part is static because 1SR is maintained all the time. This leaves us room to fully operate around the availability and cost of the access operations (provided a threshold of the total number of replicas and the probability of individual replicas). It can be seen that there are many scenarios between the availability and cost of the access operations in a distributed paradigm. There exist many contemporary strategies to manage those distributed replicas, but the question still stands which strategy is best for a given scenario or application class. Considering the fact that not every strategy fulfills each scenario, leaves many scenarios unaddressed, for which no optimal strategy exists. Hence, there is no best solution (in terms of a global optima), but solutions that serve a particular purpose (i.e., local optima). Our research focuses on the automatic identification and design of such an optimized data replication strategy.

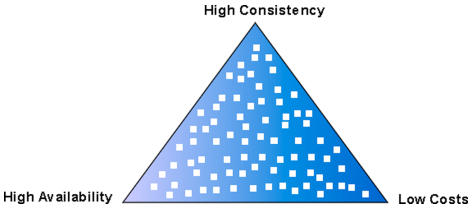

Fig. 1. Data replication scenarios [7].

3 Related Work

DRSs in general are categorized into two major classes: unstructured and structured DRSs. Unstructured DRSs, for instance, the Majority Consensus Strategy [9] use combinatorics and minimum quorum cardinalities to specify a quorum system. The Majority Consensus Strategy requires $\lceil n/2 \rceil$ replicas for the read and $\lceil (n+1)/2 \rceil$ for the write quorum to execute any operation in a system comprising n replicas. This threshold-based quorum system allows all the replicas an equal opportunity to be in a read or write quorum. However, it succumbs to high operational cost and scalability issues because of linearly increasing quorum cardinalities. This is not the case in structured replication strategies where structural properties and patterns are used to specify the quorum system. For instance, the Grid Protocol [13] imposes a logical rectangular i * j grid structure where i indicates column and j rows for a system comprised of i * j = n replicas. A read quorum consists of replicas from each column while a write quorum constitutes all the replicas from a column along with one replica from each column to satisfy the quorum system intersection property. There exist many other contemporary strategies such as Read-One-Write-All [8], the Tree Quorum Protocol (TQP) [10], the Weighted Voting Strategy [11], the Hierarchical Quorum Consensus [12], the Triangular Lattice Protocol (TLP) [14], etc., but the state-of-the-art has not much focused on a hybrid approach to explore new strategies.

There have been only a few limited efforts made towards hybrid strategies because of its cumbersome nature. So, there are some attempts, i.e., [5,17–19] on hybrid approaches, which manually design DRSs but lack automation. Moreover, there exist only a few papers, i.e., [15,16], etc., on hybrid approaches that primarily attempt to combine Tree Quorum Protocols with Grid Protocols, but they do not impose any unified structure on the nodes, which greatly limits the operability of the approach. Because of the diverse nature of topologies, there is less room for a hybrid approach to work effectively as it cannot incorporate the varied strategies freely [5]. As a consequence, many scenarios could be left unaddressed. Whereas, to address this issue, if a hybrid approach is applied to such a diverse nature of topologies, the problem easily goes out of hand. For this, this paper is an extension of our current [7] overall line of research. As it is ongoing research, it uses, extends, and optimizes the same foundational mechanism and follows the same research methodology to derive new results.

4 Methodology

Figure 2 shows the adopted methodology in a simplified manner. It starts with replication strategies being injected into a database repository and a scenario. Both, the nature of the repository and the scenario will be explained in detail in this section. The analysis and simulations (shown later in this paper) are performed on the repository until the desired solution is met, which then is inserted back to the repository for future use. Here, the question also arises of selecting the appropriate machine learning and simulation techniques for the identification and design of optimized data replication strategies. Let us dissect all these components one by one in the following.

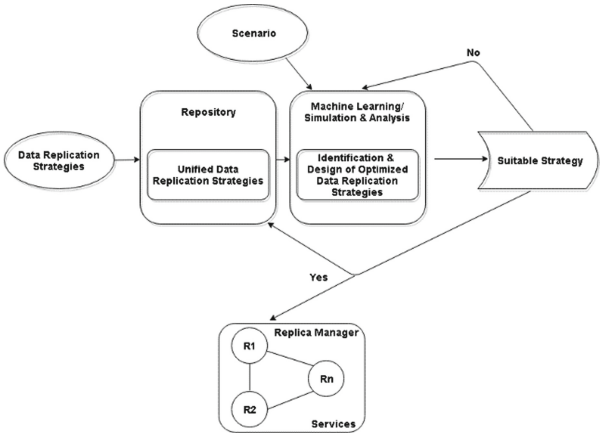

Fig. 2. Methodology [7].

4.1 Fault Model

Prior to discussing the components, we state the fault model and other assumptions first. The access operations are either read or write and are performed only when the proper quorum is acquired. The replicas are supposed to manifest a fail-silent behavior. All failures are assumed to be independent of each other. The network is supposed to be fully connected without communication failures. Only nodes (machines) with replicas can fail and the probability that a node has failed at any particular point in time is $(1 - p)$. p gives the probability that a node is available at an arbitrary point in time. The strategies are supposed to be version-based to avoid additional time synchronization issues, i.e., a replica does not only consist of some "payload" data but also a version number. A replica with the highest version number has the up-to-date payload.

4.2 Voting Structures

To address the mentioned topological and diversity issues between DRSs, a unified representation of these strategies by a concept like General Structured Voting [20] is

required for the simulation and machine learning approaches to be applied over it. Expert-based manual designs of optimized DRSs using the concept of voting structures have been presented in [18,19]. Figure 3 represents a quorum system by a directed acyclic graph (DAG) named a voting structure.

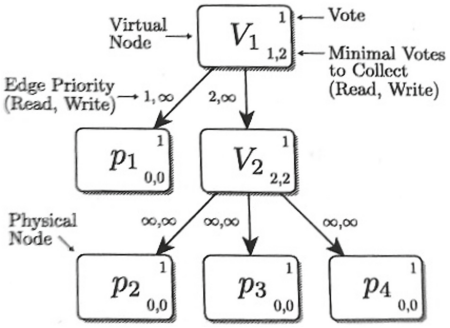

Fig. 3. Example of a voting structure [21].

A voting structure is traversed recursively by an algorithm to derive the quorums for respective access operations at run time independent of the varied topologies of the strategies. The nodes of a voting structure are either physical nodes representing actual replicas or virtual nodes that constitute the groupings of physical and virtual nodes. The virtual nodes are labeled V_i where i = 1, 2, ... while the physical nodes are labeled p_j where $1 \leq j \leq n$ and n represents the total number of replicas of a system. Irrespective of being a physical or virtual node, every node is endowed with votes comprised of a natural number (top right corner), which could also be comprehended as the weightage of that node. Furthermore, each node is equipped with a pair of minimal quorums (called "minimal votes" in the figure) to collect from its child nodes to build the read and write quorums. The minimal quorums for each node to gather per operation has to be less than or equal to the sum of the votes of its children. Some replication strategies, i.e., the Tree Quorum Protocol imposes a partial order on the quorums by which to use quorums for operation execution. The specification of such an ordering allows certain quorums to be used prior to others. In such cases, the directed edges of voting structures can be marked with operation-specific priorities imposing such orderings. An edge priority of 1 annotates the highest while the symbol ∞ represents the lowest priority. This voting structure is traversed by the recursive algorithm to derive respective quorums. It starts from the root node and queries as many of its child nodes as specified in the minimal quorums to orchestrate the quorums of physical replicas for the respective access operations. On each level, if the voting structure has a total number of votes V, then the quorums for intersection abide by the following rules in general:

$$rq + wq > V \ (to\ avoid\ read - write\ conflict) \tag{1}$$

$$wq > V/2 \ \ (to\ avoid\ write - write\ conflict) \tag{2}$$

Here, rq (wq) is a number representing the minimal read (write) quorum. For instance, the voting structure shown in Fig. 3 produces the following read (RQ) and write quorum sets (WQ):

$$RQ = \{\{p1\}, \{p2, p3\}, \{p2, p4\}, \{p3, p4)\}$$
$$WQ = \{\{p1, p2, p3\}, \{p1, p2, p4\}, \{p1, p3, p4\}\}$$

4.3 Scenario Parameters

A scenario for DRSs consists of constraints that determine the fitness of a strategy holistically to judge the goodness of a solution. These constraints may vary among different applications depending upon their nature, requirements, and resources. It comprises the following parameters [7].

Consistency of Operations. There exists a variety of data consistency models for DRSs ranging from strict data consistency to relatively weaker notions. As already stated, the consistency model opted for our approach is static and strictly meets the 1SR property. The 1SR property is maintained in a DRS when 1) every read quorum intersects every write quorum, 2) all write quorums intersect with each other, 3) replicas can be locked exclusively for write operations and locked shared for read operations.

Number of Replicas. There is a threshold imposed on the total number of replicas n that for any strategy, n cannot exceed the specified threshold value ϵ. This is because it certainly costs to create new nodes to host replicas.

$$\begin{aligned} & n, \epsilon \in \mathbb{N}^{+} \\ & \wedge n \leq \epsilon \end{aligned} \tag{3}$$

Availability of Access Operations. The probability that the data access operations are available for a DRS depends on the characteristics of the strategy, the probability of individual replicas p, and the number of replicas n. It is defined by A_r(p, n) and A_w(p, n) respectively, where A_r(p, n), A_w(p, n) $\in$ [0,1]. For some DRSs, there exist closed formulas to calculate the availability as well as the costs. However, generally, the equations given below are used to analyze the data access operations' availability of a DRS. All the RQs and WQs are derived from a DRS to calculate A_r(p, n) and A_w(p, n) for given p and n values. Equations 4 and 5 calculate the read and write operation availabilities respectively. For this, they rely on a so-called set of all possible read (write) quorums RQS (WQS). In the scope of the example given in Fig. 3, RQS equals (RQ $\cup$ {{p1, p2, p3, p4}}) and WQS equals (WQ $\cup$ {{p1, p2, p3, p4}}). The equations take the sum of the probability of all elements of RQS or WQS being available for a given probability p of individual replicas.

$$A_r\,(p, n) = \sum_{\forall q \in RQS} p^{|q|}\,(1-p)^{n-|q|} \tag{4}$$

$$A_w\,(p,n) = \sum_{\forall q \in WQS} p^{|q|}\,(1-p)^{n-|q|} \tag{5}$$

These availabilities are probabilities and constraints restrict them to be within the specified thresholds α, β.

$$\begin{aligned} A_r, A_w, \alpha, \beta &\in [0,1] \\ \wedge A_r &\geq \alpha \\ \wedge A_w &\geq \beta \end{aligned} \tag{6}$$

Cost of Access Operations. The average minimal costs for the data access operations are represented by C_r(p, n) and C_w(p, n) respectively. The read C_r(p, n) and write C_w(p, n) costs reckon the average minimal number of replicas out of the total number of replicas n, which are mandatory to perform an operation for a given probability of individual replicas p. This cost is calculated by taking the sum of the minimum number of replicas minRQ (minWQ) obligatory to form a read (write) quorum for each replica set in RQS (WQS) with the probability of the replica set appearing. Furthermore, the resulted sum has to be divided by A_r(p, n) or A_w(p, n) depending upon the particular access operation.

$$C_r\,(p,n) = \frac{\sum_{\forall q \in RQS}\ p^{|q|}\,(1-p)^{n-|q|} * minRQ(q)}{A_r\,(p,n)} \tag{7}$$

$$C_w\,(p,n) = \frac{\sum_{\forall q \in WQS}\ p^{|q|}\,(1-p)^{n-|q|} * minWQ(q)}{A_w\,(p,n)} \tag{8}$$

These costs are real positive numbers and constraints restrict them to be within the specified thresholds γ, δ.

$$\begin{aligned} C_r, C_w, \gamma, \delta &\in \mathbb{R}^+ \\ \wedge C_r &\leq \gamma \\ \wedge C_w &\leq \delta \end{aligned} \tag{9}$$

Fitness Weightage. We use a so-called fitness weightage (fw) that suggests a scenario to be biased towards either cost or availability (or even being neutral), to be able to convert a multi-objective into a single objective problem. This makes the optimization problem somewhat easier to solve.

$$fw \in [0,1] \tag{10}$$

Probability of Individual Replicas. There is a subtle difference between the availability of access operations and the availability of individual replicas p. p refers to the probability by which the replicas are available, which means the probability that a replica has failed at any particular point in time is $(1 - \mathrm{p})$ while the user performs the operations with access operations' probability. In a scenario, we restrict p to be in the interval between $\mathrm{p}_{min} \leq \mathrm{p} \leq \mathrm{p}_{max}$.

$$\begin{aligned} p_{min}, p, p_{max} &\in [0,1] \\ \wedge\ p_{min} &\leq p \leq p_{max} \end{aligned} \tag{11}$$

4.4 Database Repository

Figure 4 shows the data replication strategies Grid Protocol (left) and Triangular Lattice Protocol (right) being represented in a unified representation of a voting structure each. These voting strategies are stored in a scalable database repository in the form of JSON documents and can be queried upon any desirable criteria.

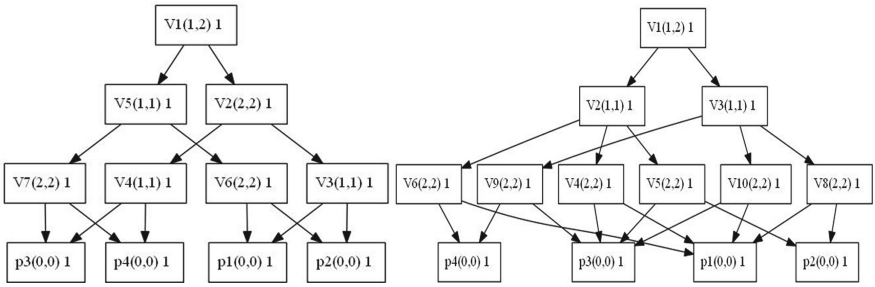

Fig. 4. Voting structures as DAGs representing DRSs [7].

4.5 Genetic Programming

The research proposes genetic programming (GP) [22,23] as a subset of machine learning to automatically identify or design application-optimized DRSs. The major difference between GP and other genetic variants of machine learning is the representation. GP is used to evolve computer programs. It consists of an encoding scheme, random crossover, mutation, a fitness function, and multiple generations of evolution to solve the specified task on its termination condition. The encoding scheme consists of a genotype (coding space) carrying an underlying set of traits and a phenotype (solution space), which is the behavioral expression of this genotype in a specific environment. Hence, the question arises which encoding scheme should be used since poor representations may lead to poor results. The crossover [24] operator mixes up the genetic material of parents in anticipation of forming a better offspring. It splits up the genome of two existing solutions at an arbitrary point and swaps them to create the offspring solutions inheriting properties from both of the parent solutions. The mutation operator changes the solution randomly but slightly, i.e., by flipping one or more bits from the previous offspring to generate a new altered child solution. In the pursuit of a solution, the questions of crossover and mutation types as well as points are also thought-provoking to address. Moreover, the population size also matters because a very small size implies a few possibilities of executing the crossovers. Therefore, only a fraction of the search space can be explored. Alternatively, a very large size may slow down the genetic approach. Although, it is highly problem-specific but very large populations do not solve the problem faster than moderate-sized populations. Figure 5 illustrates the problem in the context of genetic programming where we start from a scenario and an accordingly initial population. The initial population is analyzed based on its fitness to

the scenario in order to choose better strategies to perform crossover and sometimes also mutation in anticipation of a constant evolutionary trajectory until a solution is found.

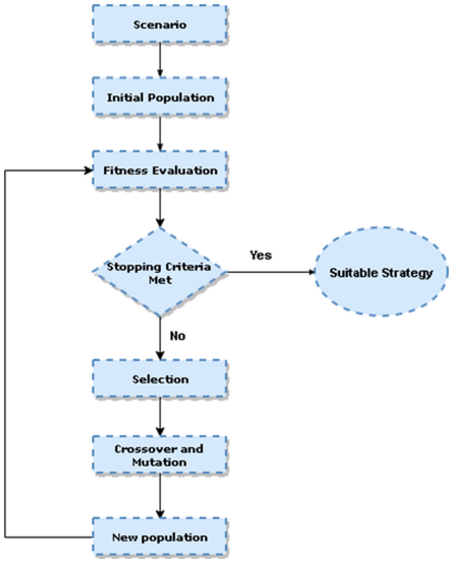

Fig. 5. Genetic programming [7].

5 Implementation

Having discussed the methodology, terminologies, and semantics, next we examine the implementation aspect comprising the parameters, functions, and the algorithm itself in detail. Once the scenario is specified to find a suitable DRS to fulfill it, the system parameters are set for the algorithm to run.

5.1 Mu, μ, and Lambda, λ

Having provided the repository to select the respective DRSs, the μ and λ values are also set as system parameters for the algorithm to start. μ is the restriction on the number of parents that are used to form the next generation and λ is the restraint on the number of offspring strategies generated using μ number of parent DRSs.

5.2 Crossover

There are various ways in which the DRSs can be combined and the resulting hybrid strategy will certainly exhibit different properties than the parents. The virtual nodes of two parent strategies are swapped to form new offspring strategies. Additionally, there are crossover points in every DRS represented by a Boolean variable which allows the crossover to be performed only on those points and in such a way that it maintains the DRSs' 1SR property throughout the process. While performing crossovers, the algorithm also restricts the number of replicas not to grow beyond a certain threshold ϵ specified in the scenario.

5.3 Mutation

The algorithm also performs mutation on the DRSs with the probability specified in the system parameters. This mutation modifies the votes of the strategies allowing some replicas to be more important in the weightage than other replicas. Once the votes are changed, the quorum also needs to be updated accordingly under the conditions (1) and (2) to uphold the 1SR property. In addition, the algorithm identifies the mutation points by a Boolean variable to avoid the DRS to be inconsistent and thereby, again, maintaining 1SR all the time.

5.4 Algorithm

Having specified a scenario and given it to the program, scenarioFitness is calculated. μ and λ are defined along with mutation probability. The list μList contains parent DRSs, the list λList comprises offspring DRSs, whereas the list initPopList consists of an initial population of DRSs. The Boolean variable isFit determines whether a strategy has achieved the expected level of fitness. The genetic program loops through all the passed on DRSs, calculates the fitness of every individual strategy, and selects the μ best strategies to the μList, in case, there is no satisfactory solution found in the initial population. This λList is then sent to the while loop to select the DRSs randomly from it (or from the initial population) and perform the crossovers and mutations to create λ offspring strategies. The use of the initial population here is for not letting the existing solutions vanish away in the next generations as the algorithm proceeds. The λList constitutes newly created strategies that are evaluated again to check if they satisfy the standard criteria. If the criteria are met, then the relevant newly generated optimized strategy is stored in the repository, the while loop terminates and so does the program. If not, it selects the μ best DRSs to the μList from (μList + λList) for the next generation. This process continues until a suitable strategy is found.

6 Experiments and Results

Figure 6 gives a relatively simple example of a hybrid DRS generated by the algorithm, which consists of 11 replicas. It can be seen that although the DRS is not very complex and maintains a tree-like structure rather than an acyclic one, yet it is so powerful and

Algorithm 1.

```
1  Specify a scenario;
2  Specify μ and λ;
3  Specify mutationProb;
4  Specify initPopListProb;
5  Define rand;
6  Initialize initPopList;
7  Initialize μList;
8  Initialize λList;
9  Double scenarioFitness = 0.0;
10 Boolean isFit = false;
11 Generate initial population of DRSs to the repository;
12 Retrieve, parse & store the generated DRSs to initPopList;
13 geneticProgrammingFunc () {
14    scenarioFitness = calculateFitness(scenario);
15    Loop through initPopList
16       Calculate fitness;
17       if (fitness ≥ scenarioFitness) {
18          isFit = true;
19          return;
20       }
21    END
22    Choose μ best DRSs to the μList;
23    Do{
24       Empty λList;
25       Loop to λ
26          Select randomly DRS1 from μList;
27          Select randomly DRS2 from (μList || initPopList);
28          Perform crossover1 of DRS1, DRS2;
29          Generate offspring DRSs;
30          if (rand(0,1) ≤ mutationProb) {
31             Perform mutation on the offspring;
32          }
33          Calculate fitness;
34          if (fitness ≥ scenarioFitness){
35             isFit = true;
36             Store offspring DRS into the repository;
37          }
38          Add offspring DRSs to the λList;
39       END
40       Select μ best DRSs to the μList from (μList + λList) for next generation;
41    }
42    While (! isFit);
43 }
```

optimized in terms of its availability and cost that it is competing with the Majority Consensus Strategy (MCS), which is believed to be the best in terms of its availability of write access operations. When compared, the hybrid DRS in terms of its availability is so close to MCS. It is almost the same for higher values of p, however, it is far better when it comes to the cost comparison.

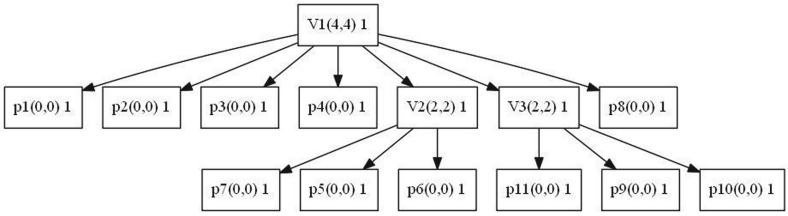

Fig. 6. Hybrid strategy 1 [7].

The availability and cost graphs on the discretized values of p are shown in Fig. 7 and Fig. 8, respectively, where Strategy 1 indicates the MCS while the Strategy 2 represents a hybrid DRSs. Both strategies consist of 11 replicas each. It can be seen that in terms of operational availability the hybrid strategy is converging on to the same values as MCS for higher values of p. This is a quite good availability but more importantly, it outclasses the MCS in terms of its cost in all the cases. Hence, it covers a scenario that could have been left unaddressed otherwise.

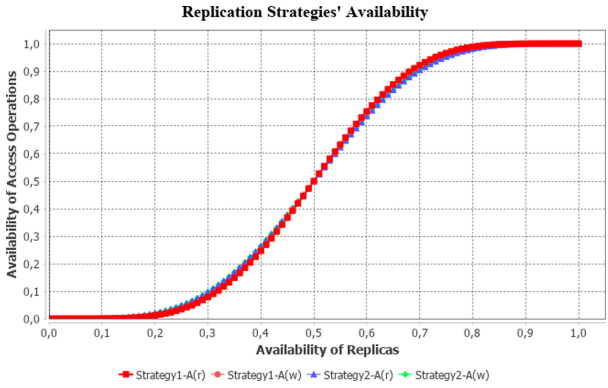

Fig. 7. Hybrid DRS 1, availability of the access operations [7]. (Color figure online)

In the best case, out of 11, it only takes four replicas each to perform a read and a write operation while the total cost for MCS is 12 for all the cases. This is a good example of a relatively less complicated DRS where we have not compromised the availability and yet reduced the cost significantly by using the hybrid approach via genetic programming.

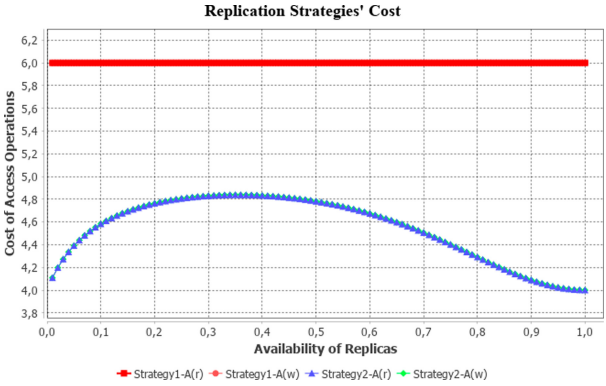

Fig. 8. Hybrid DRS 1, cost of the access operations [7].

Figure 9 shows a relatively complex but more economical example of an up-to-now unknown hybrid replication strategy designed via GP, exploiting the voting structures. It is comprised of both the Grid Protocol and the Triangular Lattice Protocol (TLP) where it unprecedentedly combines Grid Protocol comprising four replicas with TLP of six replicas, resulting in a total of ten replicas. It demonstrates an instance of a horizontal crossover, which has lowered the cost by a great value while maintaining a very good availability of the access operations.

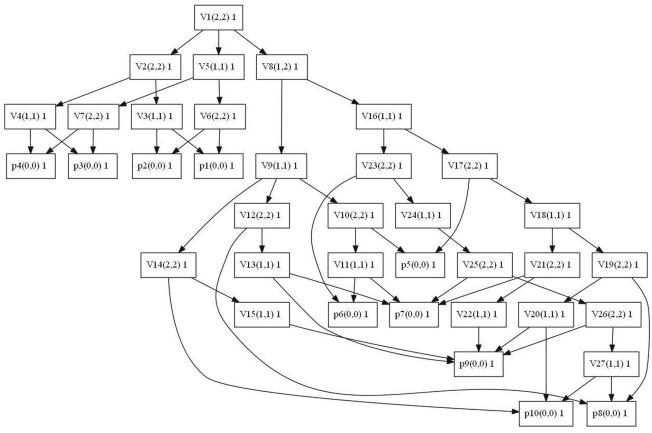

Fig. 9. Hybrid strategy 2.

Figure 10 presents and compares the availability of MCS with our hybrid approach of the same number of replicas. Red and pink lines represent the availabilities of the read and write operations, respectively, for the MCS. Whereas blue and green lines show availabilities of read and write operations, respectively, for the hybrid strategy.

The latter (hybrid) is competing fairly with the former (MCS), considering the fact that MCS is known to be the best for its availability, particularly for the critical write availability. In comparison, it can be noticed that availabilities are almost the same for the later values of p.

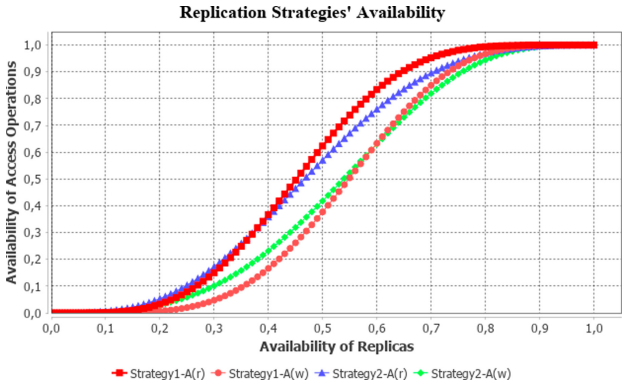

Fig. 10. Hybrid DRS 2, availability of the access operations. (Color figure online)

Figure 11 enables us a closer view where it can be observed that for hybrid strategy, the respective availabilities of the access operations converge onto almost the same values for later values of p, which is a very good operation availability considering the strong hardware nowadays.

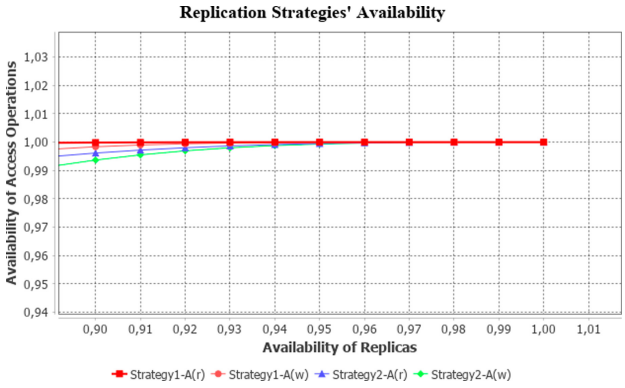

Fig. 11. Hybrid DRS 2, zoom-in availability graph.

As for the cost, as shown in Fig. 12, hybrid DRS is much cheaper as compared to the MCS. It costs almost half of the MCS, where in best cases, it only takes three replicas to perform an access operation, whereas the cost of the access operations for

MCS remains a constant of 11 replicas in total. Here, again, it is evident that we have significantly reduced the cost while not sacrificing on availability too much, covering another prospective scenario where a further reduced cost could be required.

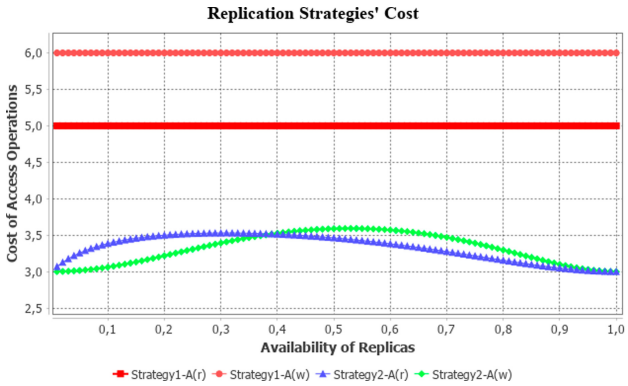

Fig. 12. Hybrid DRS 2, cost of the access operations.

We have demonstrated powerful examples of newly generated unknown voting structures via genetic programming, now we move on to specifying scenarios, explained earlier.

6.1 Scenario 1

Let us specify a sample scenario and apply our approach to find out whether a suitable replication strategy can be found. The scenario consists of the desired read and write availabilities and their respective costs, which must be achieved within the threshold of maximum 16 replicas and some availability p of individual replicas. However, cost is not important in this case, therefore, full weightage is given to availability.

6.2 Scenario Parameters 1

The desired read availability and write availability thresholds are 0.80 and 0.72, respectively, using a node availability of 0.6 inside a 16 replicas limit. Cost is specified being less than seven for each operation, but the fitness weightage determines the availability to be fully important.

$$\mathrm{p} = 0.6, \quad \epsilon = 16, \quad \alpha = 0.80, \quad \beta = 0.72,$$
$$\gamma = 7.0, \quad \delta = 7.0, \quad \mathrm{fw} = 1.0$$

6.3 System Parameters 1

Having defined the scenario, now the system parameters are set to run the algorithm accordingly. Here, the number of parent and offspring strategies are set to six and 15, respectively. The initial population is only used once, namely in the crossover process in the very first generation. The crossovers are performed all the time while the mutation is performed with a probability of 0.2.

$$\mu = 6, \lambda = 15,$$
$$\text{mutationProb} = 0.2$$

6.4 Results 1

This section shows the graphical visualization of the results generated by the algorithm on the provided parameters. It analyzes the fitness of every individual, every generation, and designs new strategies in the course of fulfilling the specified criteria when it is not found in the repository.

Fitness Analysis. Figure 13 depicts the fitness of every individual DRS and the way it evolves. The x-axis represents the number of DRSs and the y-axis denotes the fitness value of every individual strategy. The red line indicates the fitness of the DRSs while the pink and blue lines represent the availabilities of read and write operations, respectively. It can be noticed that it starts with only a few strategies of low fitness, which implies that the repository does not have a satisfactory solution to the problem. Then, the fitness improves and begins to evolve gradually through crossover and mutation operators of genetic programming until the loop stops over the desired termination condition.

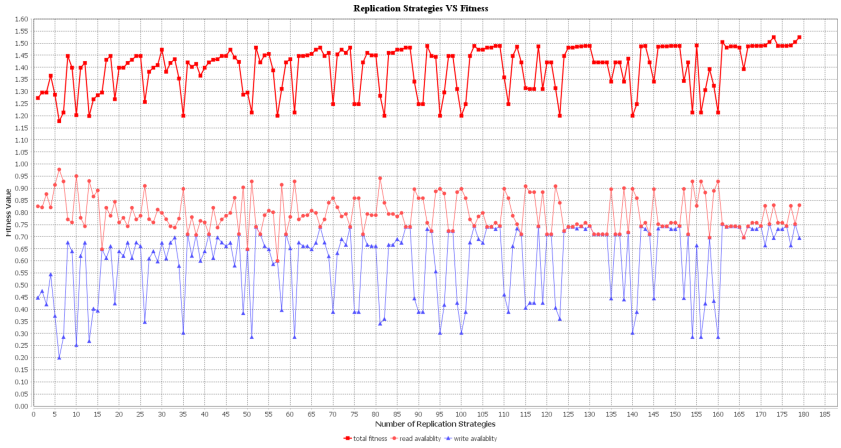

Fig. 13. Scenario 1, fitness of the DRSs [7]. (Color figure online)

Population Analysis. Figure 14 illustrates how the fitness of DRSs grows by every generation. The graph shows the fitness of the best DRSs among every generation. The x-axis represents the number of generations while the y-axis indicates the fitness value of the best replication strategy of a respective generation. It took 10 generations for the system to find a suitable DRS that satisfies the given scenario. It starts from a fitness of 1.365 and gradually but consistently continues to climb up until the desired fitness of 1.525 is achieved.

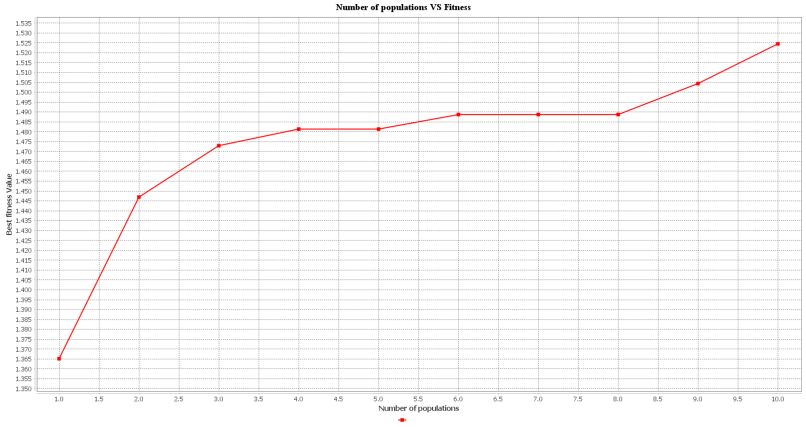

Fig. 14. Scenario 1, populations' evolution [7].

Hybrid Data Replication Strategy. Figure 15 shows the identified suitable strategy optimized for the mentioned scenario of Sect. 6.2. This strategy is comprised of 16 replicas that meet our threshold criterion ϵ. Moreover, the variable votes, quorums, and the structure itself reflect its hybrid nature that works together to serve the purpose and provide an up-to-now unknown replication strategy.

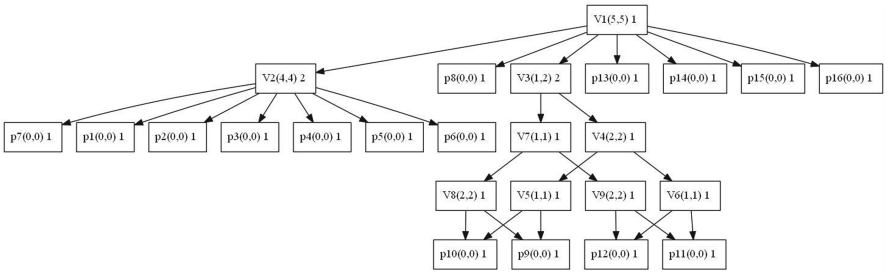

Fig. 15. Optimized hybrid DRS for scenario 1 [7].

Availability Analysis. Figure 16 shows the availability graph for the access operations of the identified DRS on discretized values of p. The newly designed DRS fulfills the specified scenario of thresholds. The x-axis represents the node availability while the y-axis indicates the availability of the access operations. The point symmetry of the graph overtly displays an extremely high availability for the access operations.

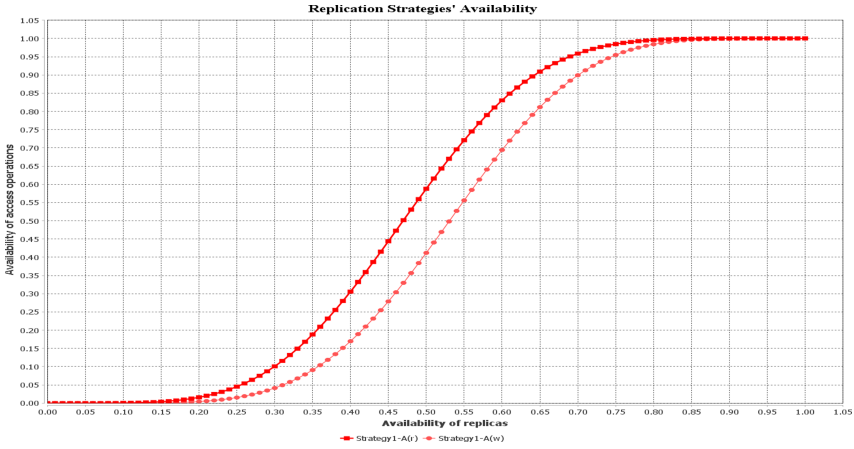

Fig. 16. Availability graph of read and write operations [7].

This availability is, again, very close to MCS (particularly for higher p values), which is considered the best in terms of the critical write operations' availability, and at the same time, our hybrid approach is reasonably economical in terms of cost. In best cases, it takes only five replicas each for the access operations out of 16 unlike MCS with a cost of 17 replicas in total for read and write, which is very expensive. A detailed comparison of this strategy is performed here [6].

6.5 Scenario 2

Having found a suitable optimized strategy for the specified scenario, now we specify a different, but relatively challenging scenario, as compared to the former one to show the effectiveness of our approach. Here, cost is also being taken into consideration, previously (Scenario 1) full weightage was given to availability, and cost was ignored in the evolutionary process.

6.6 Scenario Parameters 2

In this scenario, the probability of individual replicas is set to 0.7 and the total number of replicas is confined to 12, to achieve a read and a write availability of at least 0.85, whereas the expected cost for each operation is set to be no more than three replicas.

For this, 70% is being given to the availability of the access operations and rest to the cost.

$$\mathrm{p} = 0.7, \quad \epsilon = 12, \quad \alpha = 0.85, \quad \beta = 0.85,$$
$$\gamma = 3.0, \quad \delta = 3.0, \quad \mathrm{fw} = 0.7$$

6.7 System Parameters 2

As for system parameters, the number of parents and the number of children DRSs are set to 13 and 30, respectively, to explore the search space more. The use of the initial population in every generation is set to 30%. The mutation is performed with a probability of 0.2.

$$\mu = 13, \ \lambda = 30,$$
$$\mathrm{probInitialParentList} = 0.3,$$
$$\mathrm{mutationProb} = 0.2$$

6.8 Results 2

Figure 17 presents a 3D representation of the replication strategies generated through the process of genetic programming. The view represents the trade-offs between different objectives of the DRSs, making several possible trade-off scenarios. Here, the x-axis represents the availability while the y-axis cost of the access operations. Each strategy in this representation is given a unique color and clicking on it gives the relevant information about the DRS. The view can be divided into four equal quadrants. Quadrant 1 (top right) presents strategies with higher availabilities at higher costs. Quadrant 2 (top left) represents the strategies with lower availabilities at higher costs. Quadrant 3 (bottom left) presents strategies with lower availabilities at lower costs. Quadrant 4 (bottom right) presents strategies with higher availabilities and lower costs. The strategies are

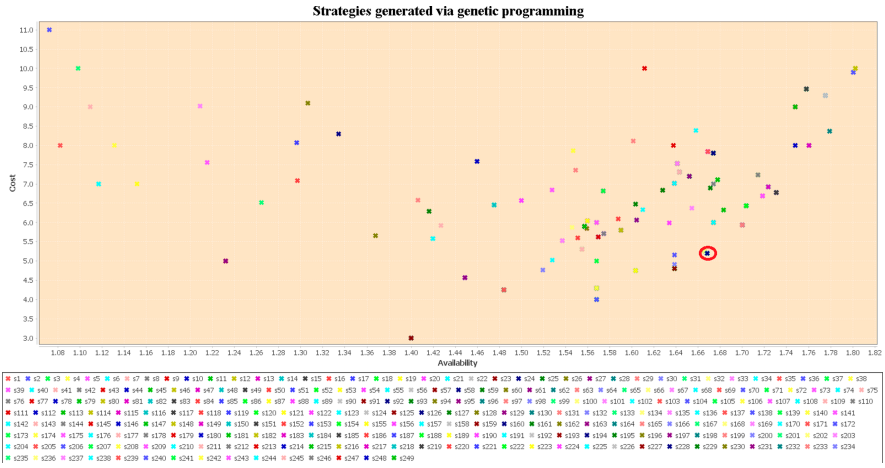

Fig. 17. Trade-off replications strategies.

getting closer to the total availability of 1.82 while the cheapest of the strategy takes three replicas in total for both the access operations to be executed. For the specified case, considering the desirable trade-offs, a DRS circled in red satisfying the criteria is chosen.

The graphs shown in Fig. 18, represent the fitness, read availabilities, and write availabilities of every individual strategy involved in the whole GP process. It is overt that it starts from a lower fitness and slowly with every passing generation it evolves. The parent strategies for the next generation are becoming better and better in terms of getting closer to the desired criteria until it hits the success.

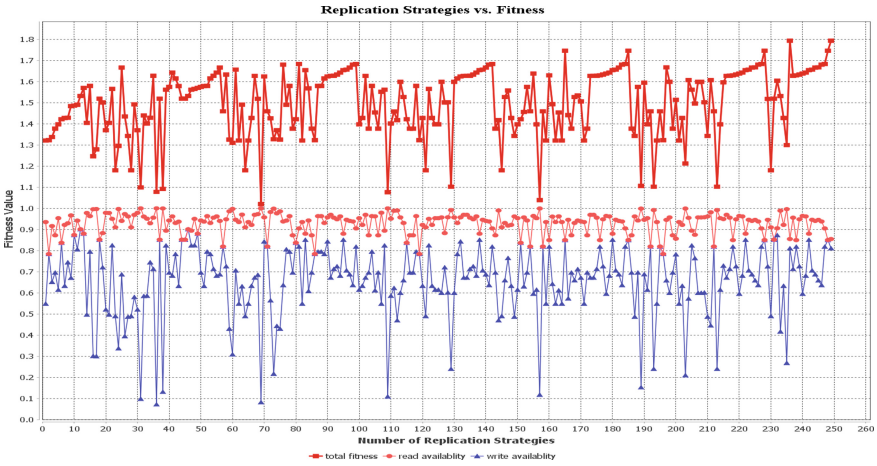

Fig. 18. Scenario 2, fitness of the DRSs.

Figure 19 shows selected strategies of best fitness among each generation. It shows an explicitly evolving trend among every generation starting from a lower fitness of 1.57 to the desired level 1.79. In this case, it takes six generations to find the respective solution.

Figure 20 displays the total number of replicas (y-axis, pink line) and the average minimal costs of read and write operations (y-axis, red line) for every DRS involved in the genetic process. The x-axis represents the strategy numbers to uniquely identify them. The graph starts with the strategies where there is not much difference between the total cost and the total number of replicas, which gradually fades away for the later strategies proving that the system is optimizing the DRSs in terms of their cost. Because now it takes less replicas to execute the access operations out of total replicas. The algorithm stops over a desired optimized strategy comprised of 11 replicas (better than the specified threshold of 12) with a total cost (sum of read and write costs) of almost five on given p.

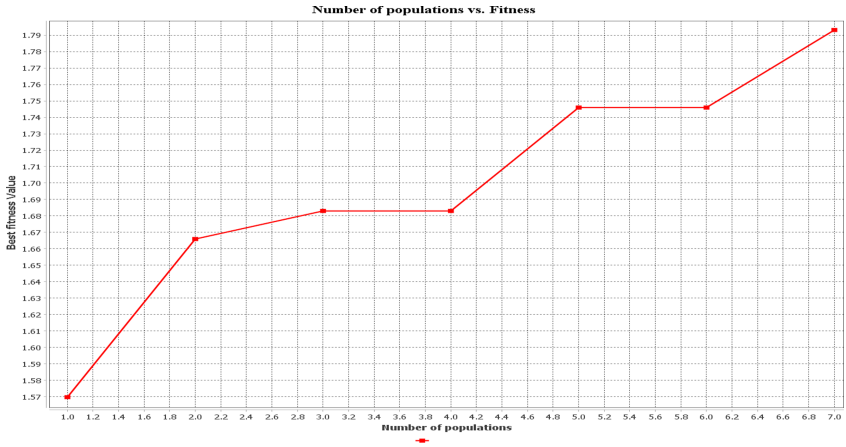

Fig. 19. Scenario 2, evolutionary trajectory.

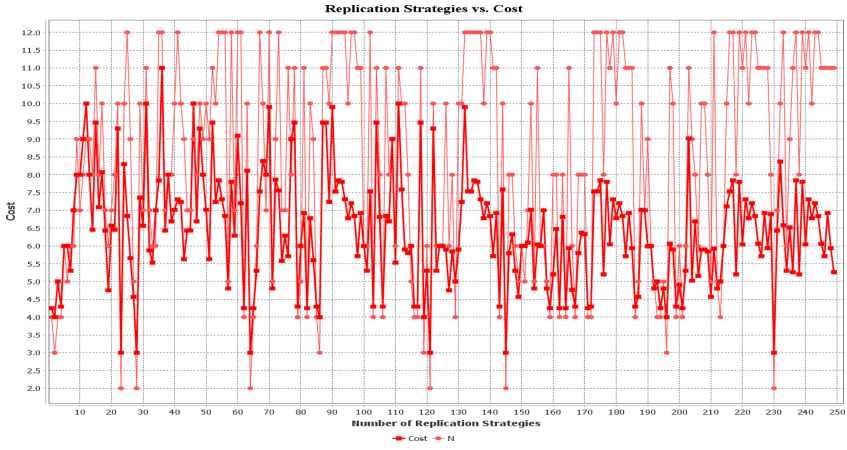

Fig. 20. Scenario 2, cost of the DRSs.

This specified scenario boils down to a fitness of 1.79 to be achieved. The algorithm finds a solution DRS achieving a total availability closer to 1.7 and total cost not more than six replicas (at best four replicas). The system takes six generations to evolve the strategies from a lower fitness 1.57 to the desired fitness of 1.793. Figure 21 shows the generated hybrid replication strategy, comprising MCS, TLP, and TQP (top to bottom, respectively). The system mixes up those strategies automatically in an intelligent way through vertical crossovers in order to satisfy the termination condition.

This instance demonstrates a very economical combination of these different strategies, which merely takes two replicas each to perform an operation in the best case, which is even cheaper than the famous TLP for both the access operations. This automatic mechanism to glue strategies together, in a certain fashion, on certain locations,

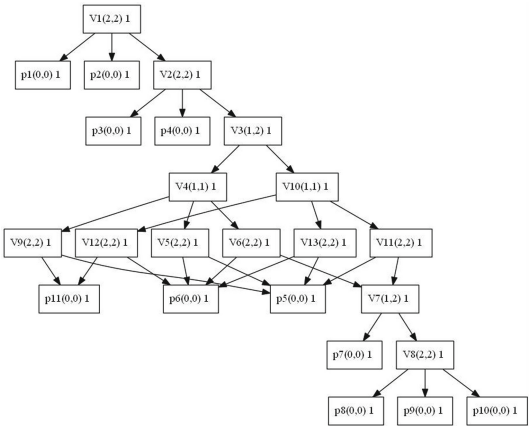

Fig. 21. Optimized hybrid DRS for scenario 2.

to make them optimized, opens up new possibilities of designing new replication strategies, until now unknown. In this manner, the proposed machine learning framework provides a strong opportunity to explore and design new unknown DRSs for any specified scenario and optimize them over several generations of evolution to meet the specified scenario-specific criteria.

7 Conclusion

The research demonstrates the use of machine learning, particularly, genetic programming for the data replication and fault tolerance. It uses this genetic programming-based automated mechanism for designing new hybrid optimized DRSs for specified application-specific scenarios for which no optimal strategy may exist. It designs new DRSs efficiently (without brute-forcing all the possible combinations) and evolve their population as computer programs to make them optimized. The novel approach does not only consider the availability aspect, but also the cost aspect and successfully models a scenario into a replication strategy. This paper demonstrates its usefulness by specifying different scenarios and accordingly designing innovative DRSs to date unknown, satisfying the criteria. This unification of the concepts of fault tolerance with the genetic programming has the potential to open whole new doors of exploring unknown replication strategies. As for future work, this research shall act as a building block to introduce new multi-type crossover as well as mutation operators to gain some more fine-grained control over the algorithm in anticipation of designing appropriate solutions, accordingly. More complex scenarios may also be taken into consideration, thereby indicating the sheer effectiveness of our approach as well as comparing with the state-of-the-art.

References

1. Bernstein, P., Hadzilacos, V., Bernstein, P.: Concurrency Control and Recovery in Database Systems. Addison Wesley, Boston (1987)

2. Naor, M., Wool, A.: The load, capacity, and availability of quorum systems. SIAM J. Comput. **27**(2), 423–447 (1998)
3. Ricardo, J., Marta, P., Bettina, K., Gustavo, A.: How to select a replication protocol according to scalability, availability, and communication overhead. In: Proceedings of the 20th IEEE Symposium on Reliable Distributed Systems (SRDS), New Orleans, LA, USA (2001)
4. Theel, O., Pagina, H.: Optimal replica control protocols exhibit symmetric operation availabilities. In: Proceedings of the 28th International Symposium on Fault-Tolerant Computing (FTCS-28), pp. 252–261. IEEE, Munich (1998)
5. Bokhari, S.M.A., Theel, O.: A flexible hybrid approach to data replication in distributed systems. In: Arai, K., Kapoor, S., Bhatia, R. (eds.) SAI 2020. AISC, vol. 1228, pp. 196–207. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-52249-0_13
6. Bokhari, S.M.A., Theel, O.: A genetic programming-based multi-objective optimization approach to data replication strategies for distributed systems. In: Proceedings of the IEEE Congress on Evolutionary Computation (CEC, WCCI), Glasgow, Scotland (2020)
7. Bokhari, S.M.A., Theel, O.: Design of scenario-based application-optimized data replication strategies through genetic programming. In: Proceedings of the 12th International Conference on Agents and Artificial Intelligence (ICAART), pp. 120–129. SCITEPRESS, Valletta (2020)
8. Bernstein, P., Goodman, N.: An algorithm for concurrency control and recovery in replicated distributed databases. ACM Trans. Database Syst. (TODS) **9**(4), 596–615 (1984)
9. Thomas, R.H.: A majority consensus approach to concurrency control for multiple copy databases. ACM Trans. Database Sys. (TODS) **4**(2), 180–207 (1979)
10. Agrawal, D., Abbadi, A.: The tree quorum protocol: an efficient approach for managing replicated data. In: Proceedings of the 16th International Conference on Very Large Data Bases (VLDB), pp. 243–254 (1990)
11. Gifford, D.: Weighted voting for replicated data. In: Proceedings of the Seventh ACM Symposium on Operating Systems Principles (SOSP), pp. 150–162. ACM (1979)
12. Kumar, A.: Hierarchical quorum consensus: a new algorithm for managing replicated data. IEEE Trans. Comput. **40**(9), 996–1004 (1991)
13. Cheung, S.Y., Ammar, M.H., Ahamad, M.: The grid protocol: a high performance scheme for maintaining replicated data. IEEE Trans. Knowl. Data Eng. **4**(6), 582–592 (1992)
14. Wu, C., Belford, G.G.: The triangular lattice protocol: a highly fault tolerant and highly efficient protocol for replicated data. In: Proceedings of the 11th Symposium on Reliable Distributed Systems (SRDS), pp. 66–73. IEEE Computer Society Press, Houston (1992)
15. Arai, M., et al.: Analysis of read and write availability for generalized hybrid data replication protocol. In: Proceedings of the 10th IEEE Pacific Rim International Symposium on Dependable Computing (PRDC), pp. 143–150. IEEE, Papeete (2004)
16. Choi, S.C., Youn, H.Y.: Dynamic hybrid replication effectively combining tree and grid topology. J. Supercomput. **59**(3), 1289–1311 (2012)
17. Theel, O.: Meeting the application's needs: a design study of a highly customized replication scheme. In: Proceedings of the Pacific Rim International Symposium on Fault Tolerant Computing, Australia, pp. 111–117 (1993)
18. Theel, O.: Rapid replication scheme design using general structured voting. In: Proceedings of the 17th Annual Computer Science Conference, New Zealand, pp. 669–677 (1994)
19. Pagnia, H., Theel, O.: Priority-based quorum protocols for replicated objects. In: Proceedings of the 2nd International Conference on Parallel and Distributed Computing and Networks (PDCN), Brisbane, Australia, pp. 530–535 (1998)
20. Theel, O.: General structured voting: a flexible framework for modelling cooperations. In: Proceedings of the 13th International Conference on Distributed Computing Systems, pp. 227–236. IEEE, Pittsburgh (1993)

21. Theel, O.: General Structured Voting: A Flexible Framework for Modelling Cooperations. Springer Vieweg, pp. 1–350 (2012)
22. Koza, J.: Genetic Programming: On the Programming of Computers by Means of Natural Selection. MIT Press, Cambridge (1992)
23. Banzhaf, W., Frank, D.F., Keller, R.E., Nordin, P.: Genetic Programming: An Introduction: On the Automatic Evolution of Computer Programs and its Applications, pp. 387–398. Morgan Kaufmann Publishers Inc., San Francisco (1998)
24. Syswerda, G.: Simulated crossover in genetic algorithms. In: Foundations of Genetic Algorithms (FOGA), pp. 239–255. Elsevier (1993)