A Markov Random Field Approach for Modeling Correlated Failures in Distributed Systems

Jorge E. Pezoa
Electrical Engineering Department and The Center for Optics and Photonics (CEFOP)
Universidad de Concepción, Concepción, Chile E-Mail: jpezoa@udec.cl

***Abstract*—In this paper, logically and spatially correlated failures affecting a distributed-computing system (DCS) have been modeled in a stochastic manner by means of a Markov random field (MRF) approach. The MRF is induced by the topology of the communication network, and is specified locally by the reliability of each node and the degree of interaction between a node and its nearest neighbors. Thus, the MRF introduces a global probability distribution function for the failure patterns of nodes in the DCS, which is parameterized using n values per node, where n is the number of nodes in the DCS. The statistical analysis conducted on test networks has shown that, compared to independent failures, correlated failures increase: (i) the average number of failed nodes due to failures propagate among the nodes; and (ii) the probability of observing a large fraction of failed computing nodes.**

***Keywords*-Distributed computing; Reliability; Markov Random Fields**

I. Introduction

Distributed computing (DC) is computing paradigm that allows to process computationally and data-intensive workloads in a parallel and cooperative fashion using a large number of computing nodes. Unlike other parallel computing environments, in a distributed-computing system (DCS) the memory is not shared, and furthermore, is geographically distributed. Consequently, computing nodes must exchanging data and control messages using an communication network that is usually bandwidth constrained [1].

A DCS is very complex system due to the heterogeneous processing capabilities of the nodes, the large number of elements in the system, the tight coupling of the nodes, the data dependencies in the workloads, and the concurrent dynamics of the nodes, workloads, and network, among other factors. In spite of this complexity, there exists a deep understanding on the computing performance and resource utilization of DCS [2]. The understanding, however, of DCSs's reliability and availability is not as deep as in the case of the aforementioned subjects.

Reliability and availability in DCSs are indeed extremely complicated to model and analyze [2]. To obtain results and achieve conclusions, researchers have simplified the problem assuming that the components in a DCS may fail independently. Under this (on occasions oversimplified) assumption, the reliability of DCSs has been vastly assessed, and studies have been conducted regarding either the computing nodes, the communication network, the application being processed by the DCS, or DCS's management software as the basic component in the analysis. Examples of such results are the works by Ravi *et al.* [3] as well as the work by Srinivasan and Jha [4] where the system reliability of complex DCSs was maximized using task reallocation, or the work by Vidyarthi and Tripathi [5] where both the safety and reliability of a DCS were jointly maximized.

The assumption on the independent node failure in a DCS is a popular because it simplifies the analysis. However, it is clear that such an assumption is not realistic for the type of failures occurring in DCS scenarios, because a DCS is a highly heterogeneous computing environment that imposes a significant communication latency [2]. Furthermore, a DCS becomes highly dynamic due to the communication network and the nodes are affected by a wide class of anomalies that change the topology of the system in a random fashion [6], [7]. These anomalies do exhibit some type of spatial and/or temporal correlation when they result, for instance, from wide-area power or network outages, communication network failures, or missed data dependencies. In addition, these correlated failures may also induce further failures in other nodes, as a result of the lack of reliable communication between the components of the DCS. For instance, by analyzing DCSs' logs stored at The Failure Trace Archive [8] we noticed that the first failure triggered by a power outage indeed produced a correlated failure in two nodes of the high-performance computing (HPC) system at Los Alamos National Laboratory. In fact, other authors have thoroughly analyzed failure logs in large-scale systems and concluded that: (i) correlated node failures are frequent events affecting such systems; and (ii) correlated failures reduce the reliability of a DCS [9], [10]. Finally, we add that it has been quantified that correlated failures may reduce the system unavailability by orders of magnitude [11], [12].

In this paper, we have tackled the problem of modeling the reliability of DCSs affected by correlated component failures. To model correlated failures, a Markov random field (MRF) [13] approach has been undertaken to derive a Gibbs probability distribution [13] for the patterns of correlated failures affecting the DCS. The Gibbs distribution is induced by the underlying network topology of the DCS, which has been abstracted using graph theory. In addition, the Gibbs

 ISBN: 978-1-61208-234-9

distribution has been parameterized by individual and group node parameters, such as the reliability of each node and the degree of interaction between a node and its nearest neighbors. Equipped with this global Gibbs distribution function, patterns of correlated-failure can be sampled using an algorithm whose inputs are: the graph modeling the topology of the DCS and n parameters per node, where n is the number of nodes in the DCS.

The rest of this paper is organized as follows. In Section III, we build the model for correlated failures. In Section IV, we present simulations results on the impact of correlated failures in the reliability of DCSs and, in Section V, the conclusions of this work are outlined.

II. Related Work

Under scenarios of application different from DC, correlated failures have been extensively modeled. For instance, in computer networks correlated failures have been modeled using naive yet effective methods, such as regarding clusters of nodes, whose joint probability of failure is "large enough," as prone to fail in a correlated manner [14]. Jiang and Cybenko used hidden Markov models to detect temporally and spatially correlated failures in a network security system [15]. Fu and Xu reported in [16] a proactive management system for node failures using a failure predictor based on spatial and temporal correlations. Other researchers have developed models for correlated failures triggered by massive natural and/or man-made events [17]–[19]. These events occur within a certain geographical region and physically damage several nodes. For example, correlated failures were modeled using geographical distances and failure probabilities in [17], while the concept of probabilistic shared-risk groups was used in [18], and a Strauss spatial point process was employed in [19] to capture the aforementioned failures.

In the context of storage systems, Bakkaloglu *et al.* [20] modeled the availability of a storage system in the presence of correlated failures introducing using two representations. The first representation used the so-called correlation level parameter, which was defined as the conditional likelihood of failure at a unit, given that another system unit has already failed. The second representation relies on the capability of the Beta-Binomial distribution to capture correlation among interconnected storage units. Later, Nath *et al.* modeled correlated failures in wide-area storage systems by fitting a bi-exponential distribution for the number nodes failing in a correlated manner [21]. In software development, Goseva *et al.* [22] and by Dai *et al.* [23] modeled correlation in software reliability using a Markov renewal process which incorporated the dependencies among successive software runs. In the context of system monitoring, Fiondella and Gokhale derived analytical expressions, based on pairwise component correlations, for the reliability in an on-demand system exhibiting correlated failures [24].

Some models for correlated failures in DCSs have been also proposed in the literature. To the best of our knowledge, the work by Tang and Iyer is the first paper on modeling correlated failures in multicomputer systems [6]. In this pioneering work, the authors tackled the modeling problem by analyzing traces from real systems and proposed a two-phased hyperexponential model for the time between failures. It is noteworthy to mention that correlation was modeled in the time domain assuming that failures propagate among nodes. Following the same ideas, Nath *et al.* studied the effects of failure patterns on the availability of DCSs using traces from real-world systems [25]. Dai *et al.* evaluated the reliability of a grid computing system by modeling the failure correlation appearing in the different subtasks executed by the grid. Chen *et al.* reported in [7] a model for temporally correlated failures in HPC systems which captures cyclic dependencies among the tasks executed by the nodes. Gallet *et al.* created a database of system logs and modeled correlated failures in a probabilistic fashion using parametric models with time-varying parameters [9].

III. Model for correlated failures

The key idea is to develop a model for correlated failures capturing the logical and spatial interaction among the nodes in a DCS. To do so, we first abstract the logical and geographical connections between the nodes in a DCS by means of the underlying topology of the network connecting the nodes. Next, the ability of MRFs to model correlated phenomena has been exploited by defining meaningful local interactions that are simple to specify. These interactions in turn, define a global Gibbs distribution of logically and spatially correlated failures. The technical details of the model are provided next.

A. Markov random fields approach for modeling spatially correlated failures

Suppose that the undirected graph $G = (V, E)$ represents the topology of a DCS, where $V = \{1, \ldots, n\}$ is the set of nodes and $E \subset V \times V$ represents the underlying topology of the communication network connecting the nodes. In order to capture both logical as well as spatial correlations in a MRF setting, the following neighborhood system is introduced:

$$\mathcal{N}_v \triangleq \{u : d_W(v,u) \leq D_{\max} \vee d_L(v,u) = 1,\ u,v \in V\}. \tag{1}$$

In words, two nodes are neighbors if their Euclidean (geographical) distance is within the range $D_{\max}$ or if they have a direct connection with each other. From this definition of neighborhood, the graph G induces the neighborhood system $\mathcal{N}$.

Suppose now that X_i is a binary random variable representing if a node has failed ("1") or not ("0"). The definition of neighborhood-system in conjunction with the collection of binary random variables $\mathbf{X} = \{X_i,\ i \in V\}$ taking values on the configuration space $\Omega = \{0,1\}^n$ is employed here to introduce a MRF. The definition of the MRF is complete when the Markovian condition is specified, that is, the MRF is completely determined when the likelihood of failure of a node, conditional on the failed or working state of its neighbor nodes is specified.

Requirements. It is of interest here analyzing the performance of DCSs in scenarios where the failure of a node induces failures in other functioning nodes, for instance, due to the inability of the working nodes to exchange data and information with a failed node. It is also of interest to this work to model situations where the geographical or logical proximity of a functioning node to a failed node increases the probability of failure on the functioning node and its neighbor nodes. To fulfill all these requirements, the following local specification for the probability of failure of the node, say, v, given the failed or working state of its neighbor nodes is proposed:

$$\begin{aligned}&\mathsf{P}\{X_v = x_v | \mathbf{X}(\mathcal{N}_v) = \mathbf{x}(\mathcal{N}_v)\}\\ &= \frac{\exp\left(-T^{-1}\, x_v(r_v - \sum_{u\in\mathcal{N}_v} s_{v,u}\ x_u)\right)}{1+\exp\left(-T^{-1}(r_v - \sum_{u\in\mathcal{N}_v} s_{v,u}\ x_u)\right)},\end{aligned} \tag{2}$$

where T is a constant, r_v is a non-negative parameter modeling the resilience of the vth node to failures and $s_{v,u}$ is a non-negative parameter modeling the strength of interaction between the nodes u and v.

Note that, due to the summation in (2), the likelihood of node v of being in a failed state, $x_v = 1$, effectively increases when one or more of its neighboring nodes are also in a failed state, $x_u = 1$. Moreover, consider the following definition for $s_{v,u}$, the strength of interaction parameter:

$$s_{v,u} = \begin{cases} \frac{D_{\max}}{d_W(v,u)} + s_L & \text{, if } u\in\mathcal{N}_v \wedge d_W(v,u) \le D_{\max} \\ s_L & \text{, if } u\in\mathcal{N}_v \wedge d_W(v,u) > D_{\max} \\ 0 & \text{, if } u\notin\mathcal{N}_v \end{cases}, \tag{3}$$

where s_L is a non-negative parameter modeling the logical strength of interaction between nodes v and u. This inhomogeneous definition for $s_{v,u}$ clearly increases the likelihood of failure of nodes when they are geographically or logically close to failed nodes.

The equivalence between MRFs and Gibbs fields can be exploited to determine the energy function. By invoking the law of total probability, the definition (2) and recalling the Markovian condition it is straightforward to obtain the energy function in terms of second-order Gibbs potentials:

$$\begin{aligned}\mathcal{E}(\mathbf{x}) &= \sum_{v\in V} r_v\ x_v - \sum_{v\in V}\sum_{u\in\mathcal{N}_v} s_{v,u}\ x_v x_u,\\ &= \mathbf{x}^T\mathbf{r} - \mathbf{x}^T\mathbf{A}\mathbf{x},\end{aligned} \tag{4}$$

where $\mathbf{x} = (x_1 \ldots x_n)^T$, $\mathbf{r} = (r_1 \ldots r_n)^T$, and $\mathbf{A} = (s_{v,u})_{n\times n}$. Thus, the Gibbs distribution associated with this energy function is

$$\pi_{\mathbf{X}}(\mathbf{x}) = \frac{1}{Z_T}\exp\left(-\frac{\mathbf{x}^T\mathbf{r} - \mathbf{x}^T\mathbf{A}\mathbf{x}}{T}\right). \tag{5}$$

Note that the local specification (2) is independent of the normalizing constant Z_T, while the Gibbs distribution depends on it. Note also that, when the strength of interaction parameters are equal to zero, the Gibbs distribution reduces to the case of independent failures.

B. A Gibbs sampler for generating correlated failures

Realizations of spatially correlated failures following a Gibbs distribution can be sampled, in theory, from (5). Unfortunately, the normalizing constant T is usually hard to compute since due to the large dimension of the configuration space. In order to circumvent this problem, sampling algorithms such as Gibbs or Metropolis samplers can be employed to generate realizations of (5). These sampling algorithms yield realizations of MRFs by constructing a field-valued, homogeneous Markov chain that has as its stationary distribution, the desired Gibbs distribution. The idea of the algorithm is to generate a realization of a Markov chain that, at a large number of iterations, will be close to (5). A key result in MRFs theory proves that a Markov chain having as a stationary distribution (5) can be constructed using the local specifications [13].

From (2), the local specifications for the vth node are:

$$p_0 = \pi\left(0|\mathbf{x}(\mathcal{N}_v)\right) = \frac{1}{1+\exp\left(-T^{-1}(r_v - \sum_{u\in\mathcal{N}_v} s_{v,u}\ x_u)\right)}, \tag{6}$$

$$p_1 = \pi\left(1|\mathbf{x}(\mathcal{N}_v)\right) = \frac{\exp\left(-T^{-1}\left(r_v - \sum_{u\in\mathcal{N}_v} s_{v,u}\ x_u\right)\right)}{1+\exp\left(-T^{-1}(r_v - \sum_{u\in\mathcal{N}_v} s_{v,u}\ x_u)\right)}. \tag{7}$$

Once these expressions are known, the Gibbs sampler can be implemented. Algorithm 1 shows the details of the sampling process, whose main idea is the following: Starting with an initial random configuration, at each iteration of the algorithm a node is randomly picked, say the vth node. The value of the realization x_v associated with the random variable X_v, is updated according to either p_0 or p_1. This process is repeated a large number of times, K, and as a result of these K iterations a sample from (5) is obtained.

IV. Simulation results

To demonstrate the ability of the MRF-based model for generating correlated failures, DCSs with representative network topologies have been considered. In the examples a nationwide DCS has been considered, where nodes are located at several cities in the USA, as shown in Fig. 1. The network topology of the first DCS considered is a realization from the class of the so-called random networks. In the second and third DCSs considered, the underlying communication networks correspond to modified versions

 ISBN: 978-1-61208-234-9

Algorithm 1 Gibbs sampler for the distribution (5). Algorithm taken from [26].

Require: $G = (V, E)$, T, r_v, s_L, $D_{\max}$, $\mathcal{N}_v$, and K
Ensure: $\mathbf{x}$
  Set $\mathbf{x}^0$ to any random value in Λ^V
  Set $k = 0$
  while $k \leq K$ **do**
    $\mathbf{x}^{k+1} \leftarrow \mathbf{x}^k$
    Randomly pick $v \in V$
    Compute p_0 using (6)
    Generate a random number $\alpha \sim U[0, 1]$
    if $\alpha < p_0$ **then**
      Set $x_v^{k+1} = 0$
    else
      Set $x_v^{k+1} = 1$
    end if
    $k \leftarrow k + 1$
  end while
  $\mathbf{x} \leftarrow \mathbf{x}^k$

of the AT&T IP backbone network 2Q2000 [27]. The DCSs studied in this section comprise 20, 38 and 17 nodes, and the Fiedler connectivity of the communication networks are 0.47, 0.45 and 0.23, respectively. (The Fiedler connectivity is defined as the second smallest eigenvalue associated with the Laplacian matrix of the graph G modeling the topology of the network [28].)

Samples of correlated failures have been drawn using the Gibbs sampler shown in Algorithm 1. To model the resilience of the nodes to failures, the r_i parameters were set, for simplicity, to be homogeneous for all the nodes. The coupling or strength of interaction parameters, $s_{i,j}$, can be defined as homogeneous ($s_{i,j} = s$ for all $i, j \in V$) or heterogeneous ($s_{i,j}$). Here, such parameters are mainly heterogeneous because they depend on the geographical distance of the nodes; however, for simplicity the parameter s_L modeling the logical strength of interaction between neighboring nodes was defined to be homogeneous. Thus, unless otherwise stated, the following parameters have been used to generate patterns of correlated failures on the DCS: $T = 1$, $r_i = 2$, and $s^L = 1$. Additional parameters employed are: (i) the Gibbs sampler iterates $K = 50000$ times before yielding a sample of the MRF; and (ii) covariance matrices were estimated using 2000 realizations of the MRF. For comparison, the case of independent failures has been also simulated by setting all the strength of interaction parameters to zero, due to when $s_{i,j} = 0$ for all i and j in (5), the Gibbs distribution reduces to a product of exponential distributions, which corresponds in fact to a probability distribution for independent failures.

Correlated failure patterns have been tested by generating a total of 2000 failure realizations, and sampled covariance matrices have been computed. Each off-diagonal element of such matrices was statistically tested for correlation using a t-test for the hypothesis of no correlation with a confidence of 99%. The results of these tests and a sample realization of correlated failures in the DCS were used to construct the

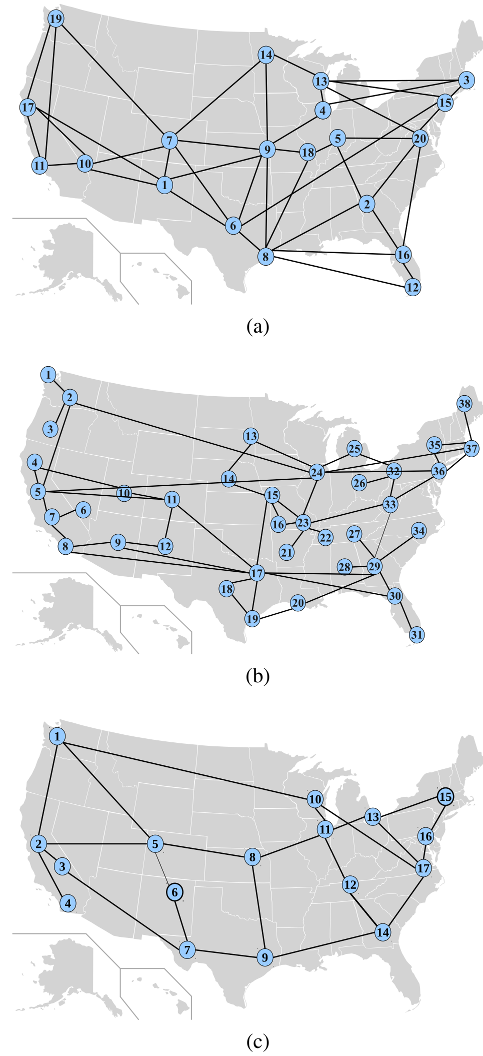

Figure 1: (a) Sample DCS composed of 20 nodes. (b) DCS interconnected by means of the AT&T IP backbone network 2Q2000 [27]. (c) DCS interconnected by means of a simplified version of the AT&T IP backbone network 2Q2000 [27].

images shown in Fig. 2. The elements in the diagonal of the matrices correspond to a sample realization from the Gibbs distributions obtained using Algorithm 1. A node, say, the ith node has become failed if the ith diagonal element is “1” (black rectangle) and is in a working state if the ith element is “0” (no rectangle). The off-diagonal elements of the matrices shown in Fig. 2 represent if there is correlation (value “1” depicted as a gray rectangle) or not (value “0” depicted using no rectangle) between the ith and jth nodes. Note that the pattern of correlated failures in Fig. 2(a) shows that four out of twenty nodes have failed. These four nodes are nodes 1, 7, 9, and 14, which clearly form a cluster of nodes in Fig. 1(a). Note also that the off-diagonal elements of the correlation matrix depicted in Fig. 2(a) confirm with a 99% confidence the correlation between the failures at the

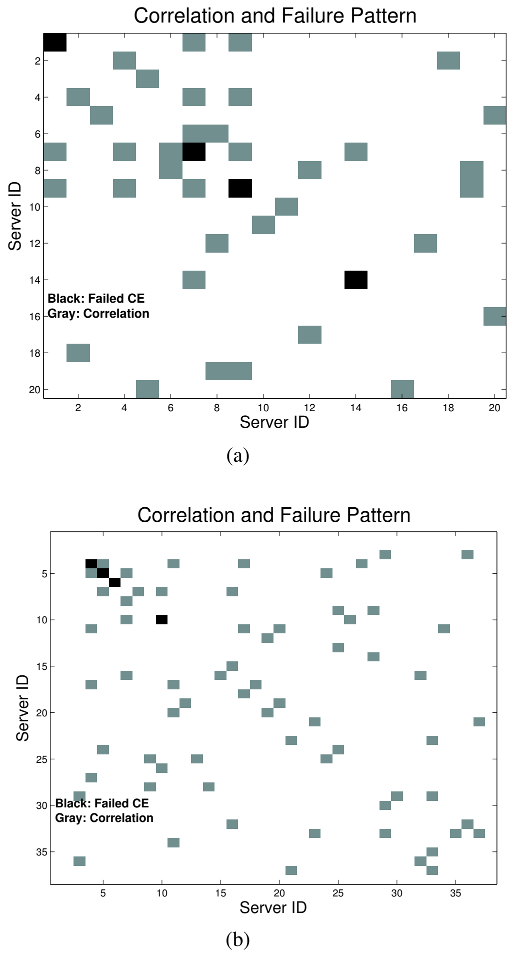

Figure 2: Matrices showing the spatial correlation in the case of the (a) sample network with 20 nodes, and (b) the AT&T IP backbone network 2Q2000. The elements in the diagonal of the matrices correspond to a sample realization from the Gibbs distributions. A red color means a failed node. The off-diagonal elements show if there is (blue color) or not (white color) spatial correlation between the failures at the nodes.

aforementioned nodes. Similarly, Fig. 2(b) shows a sample failure pattern where the directly connected nodes 4, 5, and 10, as well as the stub node 6, have failed in a correlated manner.

Figures 3 and 4 show the effect of both the resilience parameter and the strength of interaction parameter on the average number of failed nodes for the DCSs depicted in Fig. 1(a) and (b). As expected, it can be observed from the figures that the average number of failed nodes increases as the robustness parameter decreases. In addition, as either the homogeneous or the heterogeneous strength of interaction parameter increases so it does the average number of failed nodes. This behavior suggests that failures propagate more intensely as these parameters increase, thereby reflecting the fact that failures become more correlated as the coupling

Table I: FAILURE PATTERNS IN CORRELATED AND INDEPENDENT FAILURE SCENARIOS FOR THE DCS IN FIG. 1(c).

	Probability of failure patterns		
	Correlated	Independent	Failure pattern
	0.081	10^{-42}	All nodes
Clustering	0.063	10^{-35}	All except 2, 4, 10, 11
	0.030	10^{-41}	All except node 17
Effect	0.030	10^{-41}	All except node 7
	0.030	10^{-41}	All except node 2
Inhibition	10^{-5}	0.006	One node only
	4×10^{-5} to 10^{-7}	4×10^{-5}	Two nodes only
Effect	5×10^{-5} to 10^{-9}	4×10^{-7}	Three nodes only

between nodes increases. Also as expected, when the robustness parameter is fixed, the average number of failed nodes is larger in the case of correlated failures as compared to the case of a independent failures. Finally, note that the slopes of the plots in Figs. 3(a) to (c) are steeper than those shown in Figs. 4(a) to (c). This is attributed to the connectivity of the underlying networks associated with the DCSs. Note that the topology of the 20-node DCS is more connected than the 38-node DCS. Such a fact is clearly reflected in the Fiedler eigenvalue [28]. As a consequence of this greater connectivity, the logical coupling between nodes is naturally accentuated due to the larger number of relative connections in the 20-node DCS as compared to the connections in the 38-node DCS.

Table I compares the effect of correlation parameters s_L and $D_{\max}$ on some interesting failure patterns for the 17-node DCS shown in Fig. 1(c). The normalizing constant has been calculated by considering all the values in the configuration space for independent and correlated failures. With this, the probability of each specific failure pattern can be calculated from the Gibbs distribution (5). Results show that the probability of having a large fraction of the nodes failing is much higher in the correlated-failure case than in the independent-failure case. As expected from such a model, the correlation parameters $s_{v,u}$ can be used to control the degree of failed-nodes clustering or bunching. Similarly, the probability of failure patterns with very few failed nodes is much lower in the correlated-failure scenario than that corresponding to the independent case. Namely, there is a weaker "inhibition" effect in the correlated-failure scenarios compared to the independent-failure scenario.

V. Conclusion and Future Work

This paper presented a novel framework, based on MRFs and graph theory, for modeling correlated failures of computing nodes. The model abstracts the arbitrary topology of the underlying network connecting the nodes in a DCS. The developed failure model captures the spatial correlation between nodes with logical and geographical connections and captures also the percolation effect of node damage across the DCS. The model was developed by defining local conditional specifications of failure probabilities, which can

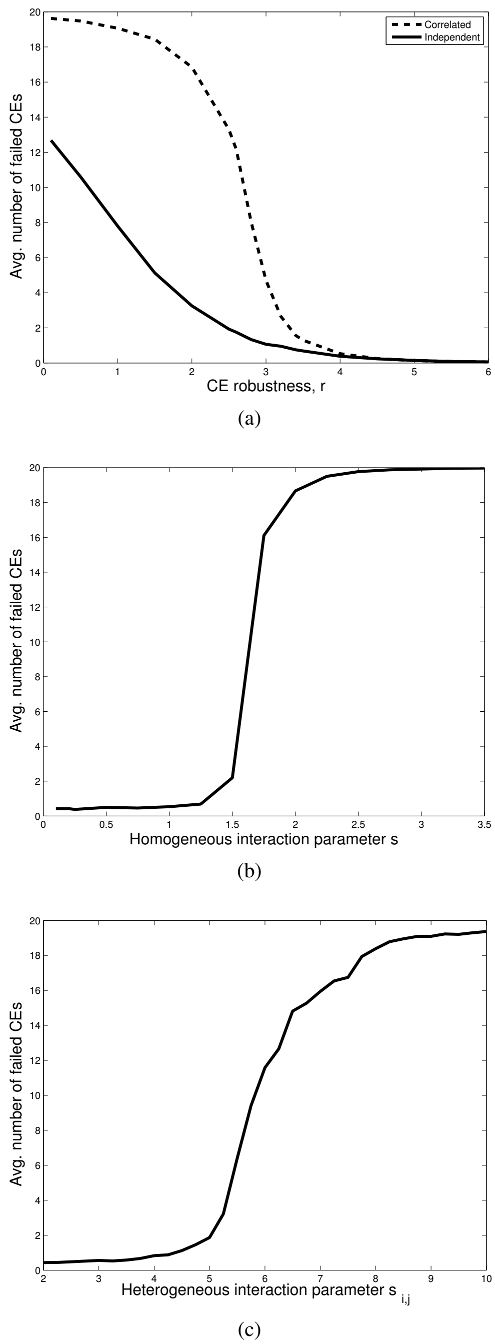

Figure 3: Average number of failed nodes versus (a) r_v parameter, (b) s_L parameter, and (c) $D_{\max}$ parameter for the DCS with 20 nodes shown in Fig. 1(a).

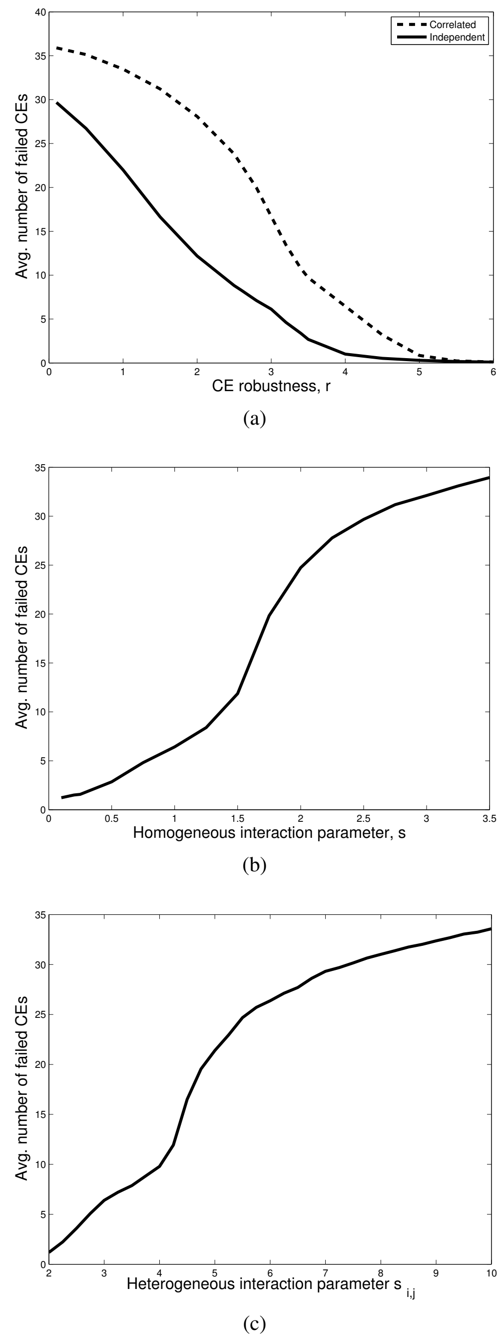

Figure 4: The average number of failed nodes versus (a) r_v parameter, (b) s_L parameter, and (c) $D_{\max}$ parameter for the DCS with 38 nodes shown in Fig. 1(b).

be easily specified in practice since they are related directly to both the geographical and logical relations imposed by the topology of the DCS. Key in the development of the model are the set of parameters termed as the strength of interaction between nodes, which quantifies the degree of interaction between nodes in terms of physical distances and also in terms of logical coupling.

The statistical analysis conducted on realizations obtained from the model for correlated failures has shown that the failure of a single node does propagate to other functioning nodes, and the degree of propagation depends on the intensity of the so-called inter-node strength of interaction parameter. The analysis confirms also that the average number of failures increases as the logical and geographical strength of interaction between nodes increases and, as expected, the average number of failed nodes also increases, as compared to the case of independent failures, when correlated failures affect the nodes of a DCS. Analytical results show also that the probability of having a failure pattern involving a large fraction of the nodes is considerably higher than in the case of independent failures, when correlated failures affect the system. Moreover, the strength of interaction parameters specified in the model can be used to limit the number of failed-nodes. This result is of practical interest in order to identify the vulnerabilities associated with coordinated attacks on the network infrastructure of the DCS.

As a future work, traces available on Internet from production systems will be studied and the proposed Gibbs distributions will be fitted in order to validate the proposed model.

ACKNOWLEDGMENT

The author acknowledges the support of both Fondecyt Project 11110078 and the Basal Project FB024.

REFERENCES

[1] R. Shah, B. Veeravalli, and M. Misra, “On the design of adaptive and decentralized load balancing algorithms with load estimation for computational grid environments,” *IEEE Trans. Parallel and Dist. Systems*, vol. 18, pp. 1675–1686, 2007.

[2] G. Bolch, S. Greiner, H. de Meer, and K. S. Trivedi, *Queueing Networks and Markov Chains*, 2nd ed. John Wiley & Sons, Inc., 2006.

[3] V. Ravi, B. Murty, and J. Reddy, “Nonequilibrium simulated-annealing algorithm applied to reliability optimization of complex systems,” *IEEE Trans. Reliability*, vol. 46, pp. 233–239, 1997.

[4] S. Srinivasan and N. Jha, “Safety and reliability driven task allocation in distributed systems,” *IEEE TPDS*, vol. 10, pp. 238–251, 1999.

[5] D. Vidyarthi and A. Tripathi, “Maximizing reliability of a distributed computing system with task allocation using simple genetic algorithm,” *J. Syst. Arch.*, vol. 47, pp. 549–554, 2001.

[6] D. Tang and R. K. Iyer, “Analysis and modeling of correlated failures in multicomputer systems,” *IEEE Trans. Comput.*, vol. 41, no. 5, pp. 567–577, 1992.

[7] X. Chen and X. He, “Tolerating temporal correlated failures from cyclic dependency in high performance computing systems,” in *Proc. 14th IEEE Int. Conf. Parallel and Distributed Systems*, Washington, DC, USA, 2008, pp. 509–516.

[8] INRIA, “The Failure Trace Archive,” http://fta.inria.fr, 2012, [Online; accessed May-2012].

[9] M. Gallet, N. Yigitbasi, B. Javadi, D. Kondo, A. Iosup, and D. Epema, “A model for space-correlated failures in large-scale distributed systems,” in *Proc. 16th Euro-Par Conf. on Parallel processing*, 2010, pp. 88–100.

[10] P. Joshi, H. S. Gunawi, and K. Sen, “Prefail: A programmable failure-injection framework,” EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2011-30, Apr 2011. [Online]. Available: http://www.eecs.berkeley.edu/Pubs/TechRpts/2011/EECS-2011-30.html

[11] R. B. Kiran, K. Tati, Y. chung Cheng, S. Savage, and G. M. Voelker, “Total recall: System support for automated availability management,” in *In NSDI*, 2004, pp. 337–350.

[12] P. Yalagandula, S. Nath, H. Yu, P. B. Gibbons, and S. Seshan, “Beyond availability: Towards a deeper understanding of machine failure characteristics in large distributed systems,” in *Proc. Of Usenix Workshop On Real, Large Distributed Systems*, 2004.

[13] P. Bremaud, *Markov chains, Gibbs fields, Monte Carlo simulation, and queues*, 2nd ed. Springer-Verlag, New York, 2001.

[14] H. Weatherspoon, T. Moscovitz, and J. Kubiatowicz, “Introspective failure analysis: Avoiding correlated failures in peer-to-peer systems,” *Reliable Distributed Systems, IEEE Symposium on*, 2002.

[15] G. Jiang and G. Cybenko, “Temporal and spatial distributed event correlation for network security,” in *Proc. of American Control Conference*, Boston, MA, 2004.

[16] S. Fu and C. Z. Xu, “Quantifying temporal and spatial correlation of failure events for proactive management,” in *Proc. 26th IEEE International Symp. on Reliab. Dist. Systems*, 2007.

[17] K. Kim and N. Venkatasubramanian, “Assessing the impact of geographically correlated failures on overlay-based data dissemination,” in *GLOBECOM*, 2010, pp. 1–5.

[18] S. Neumayer, G. Zussman, R. Cohen, and E. Modiano, “Assessing the Vulnerability of the Fiber Infrastructure to Disasters,” *Networking, IEEE/ACM Transactions on*, vol. PP, no. 99, p. 1, 2010.

[19] M. Rahnamay-Naeini, J. E. Pezoa, G. Azary, N. Ghani, and M. M. Hayat, “Modeling stochastic correlated network failures and assessing their effects on reliability,” in *Proc. IEEE Int. Conf. Computer Communications Networks*, 2011.

[20] M. Bakkaloglu, J. J. Wylie, C. Wang, and G. R. Ganger, “On correlated failures in survivable storage systems,” Parallel Data Laboratory, Carnegie Mellon University, Tech. Rep., 2002.

[21] S. Nath, H. Yu, P. B. Gibbons, and S. Seshan, “Subtleties in tolerating correlated failures in wide-area storage systems,” in *Proceedings of the 3rd conference on Networked Systems Design & Implementation - Volume 3*, 2006.

[22] K. Goseva-Popstojanova and K. S. Trivedi, “Failure correlation in software reliability model,” *IEEE Trans. Reliability*, vol. 49, pp. 37–48, 2000.

[23] Y. Dai, M. Xie, and K. Poh, “Modeling and analysis of correlated software failures of multiple types,” *IEEE Trans. Reliability*, vol. 54, pp. 100–106, 2005.

[24] L. Fiondella and S. S. Gokhale, “Estimating system reliability with correlated component failures,” *International Journal of Reliability and Safety*, vol. 4, no. 2–3, pp. 188–205, 2010.

[25] S. Nath, S. Nath, H. Yu, H. Yu, P. B. Gibbons, P. B. Gibbons, S. Seshan, and S. Seshan, “Tolerating correlated failures in wide-area monitoring services,” Intel Research, Tech. Rep., 2004.

[26] J. E. Pezoa, “Theory of resource allocation for robust distributed computing,” Ph.D. dissertation, The University of New Mexico, Albuquerque, NM, USA, 2010.

[27] M. Dodge and R. Kitchin, *The Atlas of Cyberspace*. Addison Wesley, 2008.

[28] C. Godsil and G. Doyle, *Algebraic Graph Theory*. Springer Science+Bussiness, New York, NY, USA, 2006.