J Braz Comput Soc (2012) 18:61–80
DOI 10.1007/s13173-012-0062-x

LADC 2011

On the reliability and availability of replicated and rejuvenating systems under stealth attacks and intrusions

Luís Teixeira d'Aguiar Norton Brandão ·
Alysson Neves Bessani

Received: 19 September 2011 / Accepted: 28 January 2012 / Published online: 1 March 2012

Abstract This paper considers the estimation of reliability and availability of intrusion-tolerant systems subject to non-detectable intrusions caused by stealth attacks. We observe that typical intrusion tolerance techniques may in certain circumstances worsen the dependability properties they were meant to improve. We model intrusions as a probabilistic effect of adversarial efforts and analyze different strategies of attack and rejuvenation. We compare several configurations of intrusion-tolerant replication and proactive rejuvenation, and varying mission times and expected times to node-intrusion. In doing so, we identify thresholds that distinguish between improvement and degradation of dependability, with a focus on security. We highlight the complementarity of replication and rejuvenation, showing improvements of resilience not attainable with any of the techniques alone, but possible when they are combined. We advocate the need for thorougher system models, by showing vulnerabilities arising from incomplete specifications.

Keywords Reliability · Availability · Resilence · Security · Dependability · Intrusion tolerance · Stealthiness · Replication · Rejuvenation · Models

1 Introduction

The design of *dependable and secure distributed systems* usually considers fault-tolerant or intrusion-tolerant architectures as a way to cope with faults and intrusions. In particular, techniques of redundancy in space (e.g., replication [19]) and time (e.g., rejuvenation [12]) allow systems to behave correctly even though some of its components may err or be intruded: *replication* enables a system to withstand the failure of some *nodes* (also known as *replicas* or *components*) up to a certain fault tolerance threshold, e.g., f out-of n; *rejuvenation* (also known as *repair* or *recovery*) allows malfunctioning or intruded nodes to be restored to a healthy state.

From a *reliability theory* [2] standpoint, fault tolerance has been extensively studied as a broad approach to deal with fail-prone components. In the context of malicious attacks, *intrusion tolerance* [10, 27] goes beyond traditional fault tolerance. Besides enabling dependable systems to cope with crashes and (typically random) abnormal behaviors, it also allows them to tolerate *undetected intrusions*, where parts of the system become under the control of a stealth adversary. Intrusion tolerance explicitly aims to preclude such intrusions from causing global security failures, e.g., loss of confidentiality. Common techniques used to improve the dependability of systems in traditional fault tolerance contexts sometimes imply different qualitative effects

A previous version of this paper [5] appeared at LADC 2011, the Latin-American Symposium on Dependable Computing (© 2011 IEEE).

L. Teixeira d'Aguiar Norton Brandão (✉)
Electrical & Computer Engineering Department,
Carnegie Mellon University, 4720 Forbes Ave, CyLab,
Collaborative Innovation Center, Pittsburgh, PA 15213, USA
e-mail: luis.papers@gmail.com

L. Teixeira d'Aguiar Norton Brandão
e-mail: lbrandao@cmu.edu

L. Teixeira d'Aguiar Norton Brandão · A. Neves Bessani
LaSIGE, Faculdade de Ciências, Departamento de Informática,
Universidade de Lisboa, Edifício C6, Piso 3, Campo Grande,
1749-016, Lisboa, Portugal

L. Teixeira d'Aguiar Norton Brandão
e-mail: lbrandao@di.fc.ul.pt

A. Neves Bessani
e-mail: bessani@di.fc.ul.pt

in intrusion tolerance contexts. The different requirements of each context usually imply different levels of sophistication, thus resulting in systems with distinct properties. Still, a common first-sight intuition (though sometimes wrong) considers that fault tolerance and intrusion tolerance, obtained by architectural augmentation of an initial system (e.g., replicating several components and requiring a majority vote for each decision), are aligned with dependability just because they allow some components to fail or be intruded. In particular, one could naively believe that the increase of the threshold of faulty components that a system can withstand always leads to the improvement of dependability of the overall system. Contrarily, in this paper, we highlight attack scenarios for which the dependability of systems tolerating intrusions is lower than that of the respective non-augmented systems. We show how over-simplified system models, with incomplete specifications, leave room for vulnerabilities. Our arguments are based on high-level aspects of the redundancy architectures and ignore the cost of their implementation and operation.

The choice of terminology related with *dependability* and *security* is a matter of interesting discussion [1]. We do not intend to enter such discussion in this paper, but we shall make a quantifiable analysis of dependability based on well-defined metrics, while considering different attack models and intrusion-tolerant configurations. By *dependability improvement,* we mean higher *reliability* ($\mathcal{R}$) or *availability* ($\mathcal{A}$), both dependability attributes, with $\mathcal{R}$ measuring the probability of never failing to maintain a certain property during a certain *mission time*, and with $\mathcal{A}$ measuring the probability of correctness at a random instant within an intended *mission time*. We envision (but do not discuss) a security perspective, where these metrics might be used to establish (or compare) the ability of systems in accomplishing or maintaining certain security goals.

An example Consider a file-storage server (a node) controlling the access of clients (external users that interact with the node) to some data. Consider that, due to dependability concerns, this storage system is augmented to a new system made up of n *replicated* nodes, such that the correct access to data requires the interaction and combination of *votes* of a large enough subset of correct nodes. For example, if the single main security concern is confidentiality, then this replication could be achieved in the way of a secret sharing scheme [20]. If more security properties are involved, such as integrity and availability, then a more sophisticated system could be proposed [3]. The motivation for such *augmentation by replication* is typically supported by an implicit (or explicit, but not justified) assumption: that a replicated system should be more dependable than a single node, namely that it should have less *likelihood* of failing (or be failed in) its mission. In this paper, we challenge the coverage of such assumption throughout several examples that illustrate the opposite scenario. In fact, we show that, within some models and domains of configurations, there is room for both upgrade and downgrade of the properties that one would typically expect to improve. In other words: *techniques that augment the dependability or security of systems in one environment might decrease them in another related environment*. We thus propose that assumptions about dependability improvement, namely those brought upon by techniques of replication and rejuvenation, should be justified, rather than implicitly assumed.

Still in the example of a replicated system with n nodes, consider a protocol that is guaranteed to perform correctly if and only if at most f nodes are in erroneous state. In our context of attacks, we call *intrusion of a node* to the process of transitioning it from a correct (*healthy*) state to an erroneous (*intruded*) state, and denote f as the *threshold of tolerable intrusions*. The functional relation between the threshold f of tolerable intrusions and the total number n of nodes usually depends on the type of protocol and the nature of intrusions. For example, it is common to have systems allowing crash fault tolerance with $n \geq f + 1$, while Byzantine fault tolerance usually requires $n \geq 3f + 1$ [6, 18, 26].

Besides issues that may be specific to a particular protocol or system, there is a quantifiable effect on the dependability of a system, which arises from a direct relation between some high-level aspects of its intrusion-tolerant configuration (e.g., the $\langle n, f \rangle$ relation), the dynamics of intrusion of each component (e.g., the way in which an attack promotes an intrusion) and the intended mission time of the system. For example, it is well known that a *Triple Modular Redundant* architecture (i.e., $n = 3$ and $f = 1$) under accidental random faults (e.g., crash of components that ware out with time) is less reliable than its non-redundant counterpart (i.e., $n = 1$ and $f = 0$), if the mission time of the system is long enough compared with the *expected time to failure* (ETTF) of each component [13]. In this paper, we revisit this result while considering a security perspective, where intrusions happen as a result of stealth attacks. We compare the dependability of different families of intrusion-tolerant configuration (characterized by certain f/n ratios), including proactive rejuvenation of nodes [6, 12, 21], for a range of mission times. In doing so, we identify thresholds that make the difference between improvement and degradation of *dependability*. It is our goal to emphasize that the dependability/security enhancement being sought with intrusion tolerance may sometimes be jeopardized, if the estimation of reliability or availability is neglected.

Goal and contributions With this paper, we aim to highlight the importance of system model specifications that allow a quantitative (or at least comparative) evaluation of the dependability properties being sought. We pursue this goal by exemplifying: a model of relationship between attack and intrusion, allowing such quantification; and variations of *reliability* and *availability* brought upon by different *attack models*, *replication configurations* and *rejuvenation strategies*. We present the following technical contributions:

1. we formalize an *intrusion model* directly dependent on the *adversarial effort* for intruding nodes and compare results for different instantiations of attack;
2. we identify scenarios where *intrusion-tolerant replication* decreases *reliability* and *availability* of a system under attack;
3. we find configurations toward *reliability* and *availability improvement goals*, for finite, unbounded and infinite mission times;
4. we highlight the possible complementarity between *replication* and *rejuvenation*.

Organization The remainder of the paper is organized as follows: Section 2 introduces a preliminary system model, modeling attacks and intrusions, and defines several dependability attributes; Section 3 illustrates analytic and quantitative results, focusing on *reliability* and formalizing a notion of *relative-resilience*; Section 4 extends the system model to consider *rejuvenations*, shifting the focus to *availability* and obtaining respective results; Section 5 describes some related work; Section 6 concludes the paper with some final remarks; the Appendix collects the mathematical formulas that sustain most of the results presented throughout the paper.

2 Preliminary system model

In this section, we define a preliminary[1] system model and the metrics that we shall use to characterize it. On purpose, we define a system model that is able to span a family of configurations, so that we can study the variation of characteristics across different instantiations.

Definition 1 An *intrusion-tolerant replicated* system, $\langle n, f\rangle$ (with $0 < f < n$), is a system composed of n nodes, correct while the simultaneous number of *intruded* nodes does not exceed f. $\langle 1, 0\rangle$ is called the *reference system*—one that fails when its single node is intruded.

With "intruded", we intend a meaning more general than is usually denoted by "faulty" or "erroneous". In particular, an *intruded* node might continue to execute correctly, from some operational point of view, despite being already under the control of a *malicious* adversary. Such control may be as subtle as the ability, at any time decided by the adversary, to interfere with the service running on the node.

We are interested in comparing characteristics of $\langle n, f\rangle$ with those of $\langle 1, 0\rangle$, when the former is built as an architectural augmentation of the later, using intrusion-tolerant replication. Many implementations fit this model. For example: $f = n - 1$ for some synchronous crash fault-tolerant (Crash FT) protocols (e.g., [19]); $f = \lfloor (n-1)/2 \rfloor$ for some *Byzantine fault-tolerant* (BFT) systems with synchrony (e.g., [19]) or using trusted components (e.g., [7]); $f = \lfloor (n-1)/3 \rfloor$ for general BFT systems (e.g., [6, 26]).

Definition 2 The *mission time* (MT) of a system is the uninterrupted interval of time during which the system is intended to be correct. MT may be finite and known, finite but unknown, or (assumed to be) infinite.

We do not consider MT to be a deadline for a mission to be accomplished, but instead the duration of time during which some property should hold valid (e.g., be available to perform an operation, or ensure the confidentiality of some information).

Definition 3 The *reliability* ($\mathcal{R}$) of $\langle n, f\rangle$ is the probability that the system will never fail during its MT.

Definition 4 The *availability* ($\mathcal{A}$) of $\langle n, f\rangle$ is the probability that the system is not failed at an instant of time randomly and uniformly chosen from the MT period.

Equivalently, we say that $\mathcal{A}$ is the expected proportion of MT during which the system is correct.

Definition 5 A dependability property (e.g., $\mathcal{R}$ or $\mathcal{A}$) of a $\langle n, f\rangle$ system is said to be *desirable* if it is *better* than that of $\langle 1, 0\rangle$.

For example, if $\mathcal{R}_{n,f} > \mathcal{R}_{1,0}$, then $\langle n, f\rangle$ is said to have desirable $\mathcal{R}$.

Assumption 1 (Intrusion model) The system has a $\langle n, f\rangle$ architecture, with state represented by vector $\vec{\phi}(t)$, of length n, at each instant t in time. The state of each node j, with $j \in \{1, \dots, n\}$, is given by $\phi_j(t) \in \{0, 1\}$, with 0 denoting a *healthy* state (H) and 1 denoting an *intruded* state (I). Each node starts in state H, at $t = 0$, and transitions probabilistically to state I according to an *intrusion rate* (IR) $\lambda_j(t)$ (a probability density) that is directly proportional to an *intrusion adversarial effort* (IAE) exerted on the node at instant t. The proportionality ratio IR/IAE is the same for all nodes and shall henceforth be assumed to be 1.

[1] We call it "preliminary" because we shall extend its properties, later in the text (see Section 4.1).

Assumption 1, distinguishing IR and IAE, defines the dynamics of intrusion but does not assert anything about the intensity with which the nodes might be attacked. Instead, it simply states how the process of intrusion occurs, in this model, as a probabilistic result of an attack (i.e., of an adversarial effort). The proportionality relation implies that all nodes have the same probability of being intruded when subjected to the same IAE for the same amount of time, even though an attacker could still choose to attack different nodes with different variations of effort. Since we assume a proportionality constant of 1, henceforth we shall use $\lambda_j(t)$ to specify both IR and IAE.

We have just introduced an *intrusion model*. We can also look at Definition 1 as a *failure model*, once n and f are fixed: *the system is failed whenever more than f nodes are simultaneously in intruded state*. In many real systems, some deviations from correctness do not necessarily imply failure (or at least immediate failure). Nonetheless, since we focus on environments with malicious attacks, we opt for a conservative estimation of dependability properties. If considering reliability, this means that a system fails *as soon as* more than f nodes are in intruded state. If considering availability, namely for systems with rejuvenation (see Section 4) where the number of intruded nodes is not a monotonic function of time, the system is failed *only during the periods in which* the number of intruded nodes is higher than f. This contrasting perspective of "fails as soon as" versus "failed only during the periods in which" is a good way to differentiate the concepts of reliability and availability.

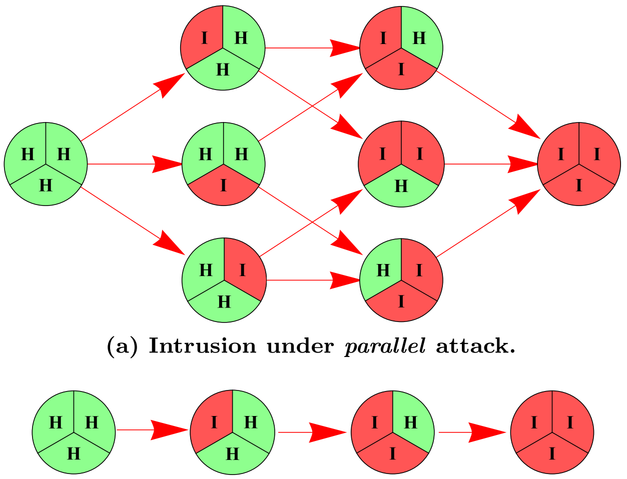
We still need to model the process of attack that leads to intrusions. We proceed with two *alternative* attack models, both of practical interest (see (1) and (2) in the Appendix, for mathematical details).

Assumption 2 (Attack models) The system may be attacked in one of the following manners:

- *Parallel Attack* (‖)—The IAE is equal on all healthy nodes and has constant intensity (λ);
- *Sequential Attack* (.·.)—The IAE targets one healthy node at a time, with constant intensity (λ).

Our analysis will be based on mathematical abstractions, but from a practical point of view we envision cases where the IAE is a pressure impressed directly by an *attacker*. We do not consider cases where intruded nodes could become a helper of the attacker, contributing to the IAE over the remaining healthy nodes.

The diagram in Fig. 1a illustrates the possible states and state-transitions of a system with $n = 3$. Note that the $(i + 1)^{\text{th}}$ leftmost column of circles contains all global states with exactly i intruded nodes. Thus, a $\langle 3, f\rangle$ system, for any $f < 3$, is failed whenever its global state corresponds to any circle to the right of the $(f + 1)^{\text{th}}$ leftmost column. The diagram naturally suits the parallel attack model, if each arrow represents a constant IR λ (resulting from a constant IAE). However, it could also fit a sequential attack model if for each circle only one outbound arrow (does not matter which) is allowed to have a positive IR, i.e., if only one transition is possible.



(a) Intrusion under *parallel* attack.

(b) Intrusion under a particular choice of *sequential* attack.

Fig. 1 State diagrams of intrusion for a system with $n = 3$ nodes. In each sub-figure, each *circle* represents a global state of the system, with each *inner triangle* representing the state of a node: healthy (H) or intruded (I). Each *arrow* represents a transition where a single node changes from H to I. Each *arrow* corresponds to a constant *intrusion rate* (λ)

The diagram in Fig. 1b, which is actually a sub-diagram of the one in Fig. 1a, naturally suits the representation of a particular choice of sequential attack, if one considers that: each position of inner triangle (inside a circle) represents a particular node; and each arrow has an associated constant IR (λ). Actually, while not considering rejuvenations, all paths of sequential attack leading to failure are equally efficient, i.e., the ordering in which nodes are attacked is irrelevant. With some flexibility (and this will be important when interpreting more complex diagrams in the remainder of the paper), this diagram can also be interpreted as representing a parallel attack model, if: the $(i + 1)^{\text{th}}$ leftmost circle stands for all possible states containing exactly i intruded nodes (see Fig. 1a); and the respective outbound arrow of each circle stands for $(3 - i)$ possible transitions from each imagined source state to the respective possible $(3 - i)$ destination states. In particular, each of the middle circles ((I, H, H) and (I, H, I)) in Fig. 1b would represent 3 possible states (the circles in the respective column in Fig. 1a), and each of the lateral circles ((H, H, H) and (I, I, I)) in Fig. 1b would represent a single state (equivalent in Fig. 1a). Later in the paper (see Section 4), we shall augment these types of diagrams to include also the effect of rejuvenations.

A note on diversity The assumptions made so far do not consider cases of exploitation of *common-mode vulnerabilities*, capable of leading all nodes to immediate simultaneous intrusion. If this were to be possible, the discovery of a vulnerability in a node could facilitate the intrusion of remaining healthy nodes (a dependence between intrusions that would favor the attacker). In practice, avoiding such vulnerabilities is a hard-to-solve problem. A common technique to mitigate the problem involves implementing intentional *diversity* in the replication and rejuvenating process of nodes, on the dimensions that sustain the vectors of attack [16] that are likely to be exploited. It is not our goal to discuss the feasibility or effectiveness of such techniques—we simply focus on constructed examples that fit well the model of independence of intrusions. However, *we strongly emphasize that we are not simplifying as in making a* wish *for added security, but rather to show that, even with probabilistic independence of intrusions across nodes, dependability properties might still be brought down by the intrusion-tolerant techniques* (*e.g., replication*) *whose application would typically intend otherwise.*

A note on dependence It is also worth mentioning a commonly overlooked fact: although, from a defensive point of view, *independence of intrusions* is better than a possibility of *simultaneous collective intrusion*, it is not an optimal situation. As noted in [24], "better than independence can actually be attained". Actually, there are two orthogonal axes of dependence: one refers to the probabilistic aspect of intrusions (our model is indeed of independence, because the ratio IR/IAE does not depend on the number of intruded nodes); another refers to architectural aspects of attack (e.g., in the parallel-attack model nodes are attacked independently of others, whereas in the sequential-attack model there is a *good* dependence in the (defensive) sense that each node under attack protects all remaining healthy nodes from being attacked).

Examples of potential attack scenarios We emphasize that it might not be within the reach of an attacker to decide freely about the characteristics of a possible attack to a system. For example, a goal of stealthiness may require a limitation of the IAE upon each node. Also, the architecture of the system might protect itself from exposure to a certain type of attack ($\|$ or $\therefore$). Of course, each system might lend itself to different *vectors of attack*, i.e., to several ways of having one or more vulnerabilities being exploited. The following informal examples are consistent with our assumptions and illustrate possible constraints on attacks:

– *IAE limited to ensure stealthiness.* Consider a set of online nodes, each protected with a random-one-time-password, under a $\|$-attack using random password attempts, with equal frequency in all nodes. If the system is prepared to sound an alarm if too many incorrect passwords are attempted in a given window of time, then the attacker must limit its IAE in order to remain undetected. In this case, a $\|$-attack on n nodes cannot be replaced by a $\therefore$-attack with a focused *effort* n times higher in a single node at a time.

– *Parallel type required by architecture, IAE limited by reactiveness.* Consider a server application with a certain buffer-overflow vulnerability, leading to immediate intrusion if exploited with a certain code-injection. If a $\langle n, f\rangle$ system were to be built with n online servers (nodes) with the same application, then an adversary could potentially intrude all of them simultaneously (i.e., with the same code injection). To prevent such dependency, consider that an *instruction set randomization* (ISR) mechanism [24] is used, where the server-application of each node corresponds to a randomized version, indexed by an independent small key. The ISR might not remove the vulnerability of each node, but simply obfuscate it, such that a different code injection, unknown in advance to the attacker, is necessary to provoke intrusion. The attacker might still intrude each node, by trial and error attempts until it guesses the respective randomization key, but the intrusion success is independent between nodes. The frequency of such attempts (and thus, proportionally, the IAE) might be limited if each unsuccessful buffer-overflow attempt makes the server crash and reboot. Additionally, let the communication between a client (the attacker) and a set of servers (the nodes) be mediated by a proxy which, for each client-request, establishes a connection with a random server. In this example, the attacker is limited to a $\|$-attack, because, from a coarse time-granularity point of view, each server experiences the same average of intrusion attempts per amount of time (i.e., the same IAE).

– *Sequential Attack due to attacker's limitations.* Consider a single-person (the attacker) that is well skilled in a type of social-engineering attack, requiring human physical presence for a continued amount of time. If the system being targeted is a set of geographically dispersed nodes, then the individuality of the attacker only allows him to perform a $\therefore$-attack. For a similar type of example, consider an attack that requires a distinct learning phase for each node (e.g., learning a language). If each learning task is more efficient when performed in a focused way, then a $\therefore$-attack type might be preferable. For compatibility with Assumptions 1 and 2, each intrusion should not provide any advantage to the next intrusion, or, more precisely, the proportionality ratio between IAE and IR remains constant and the IAE itself remains constant.

– *Sequential Attack with ordering defined by architectural properties of the system.* Consider a system that protects itself with a nested layering of defenses—for example, a vault inside a vault, inside a vault, and so forth. If the only known feasible attack requires breaking the outer layer and proceed sequentially through the inner layers, then only a ∴-attack type can be performed.

In the next section, we shall compare how reliability is affected by different types of attack, among other varying parameters. We argue that it is pertinent to compare different models, because in practice the same system might be subject to different adversarial environments.

3 Time, reliability, resilience

In this section, we consider the *reliability* ($\mathcal{R}$) of $\langle n, f\rangle$ systems, under each model of attack and in several perspectives:

1. Which $\langle n, f\rangle$ systems have a *desirable expected time to failure* (ETTF)?
2. For which *mission time* (MT) does a $\langle n, f\rangle$ system have a *desirable* $\mathcal{R}$?
3. Given a MT, a goal of $\mathcal{R}$ and a functional relation (e.g., a ratio) between *replication degree* n and *intrusion tolerance threshold* f, how to adjust f or n?
4. How to define goals of $\mathcal{R}$*-improvement* and how to achieve them?

To be practical, we shall group systems by functional relations $n(f)$ or $f(n)$, relating the degree of replication (n) with the intrusion tolerance threshold (f). We shall use suggestive labels, such as *Crash* and *Byzantine* (in *synchronous* or *asynchronous* environment), to identify such groups. For example, simple *Crash* fault-tolerant systems are often achieved with $n = f + 1$, i.e., $f = n - 1$. It is also common to see *Byzantine* fault-tolerant systems with $n = 2f + 1$ or $n = 3f + 1$, i.e., $f = \lfloor (n-1)/2\rfloor$ or $f = \lfloor (n-1)/3\rfloor$. However, we emphasize that, despite the labeling, the analysis ahead will not be based on the type of faults, but only on the relation between n and f.

3.1 Expected time to failure

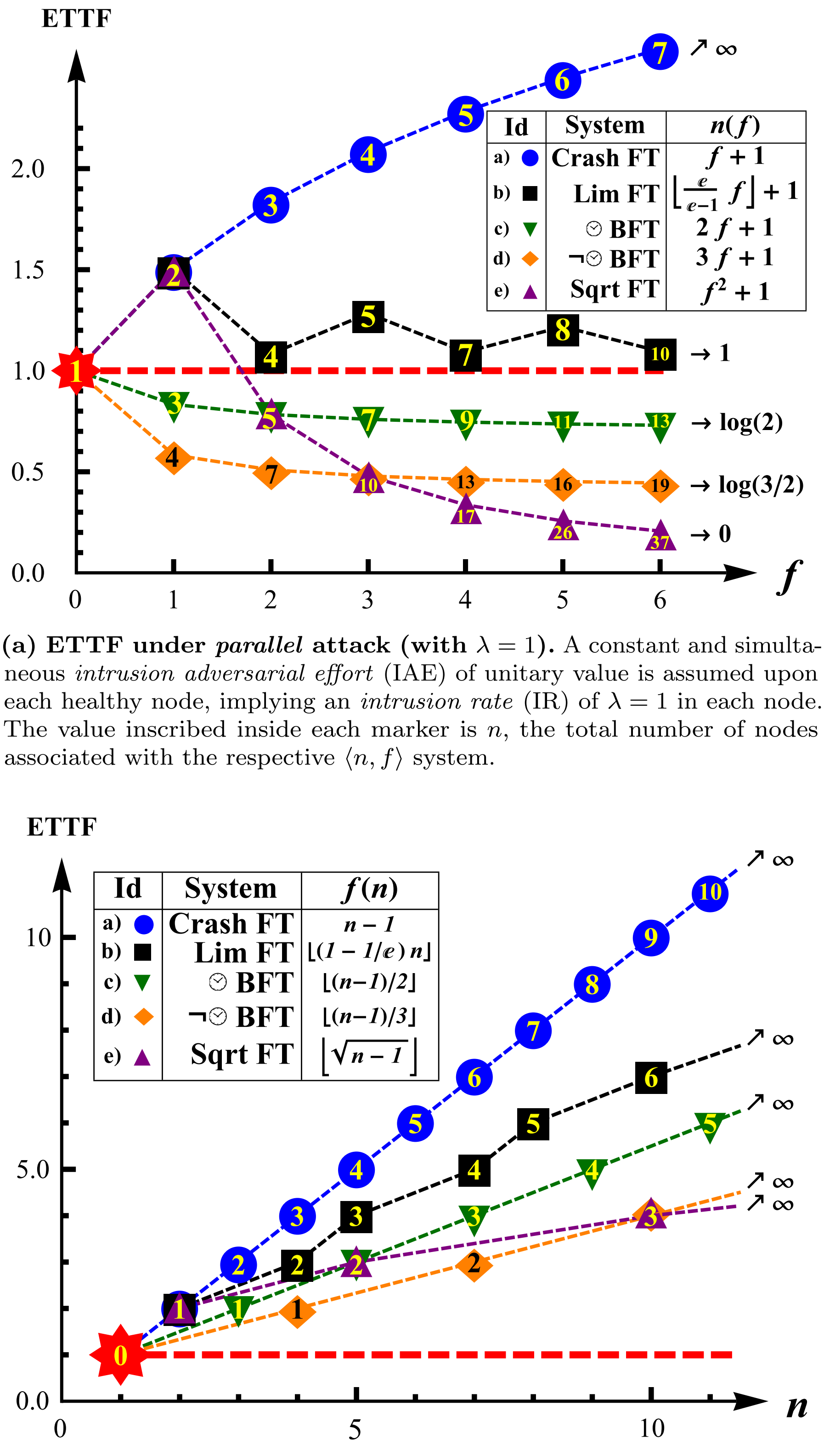
For the reference system, $\langle 1, 0\rangle$, the parallel (∥) and sequential (∴) models of attack are equivalent. The probability of the single node becoming intruded follows an exponential distribution and the respective ETTF ($\mu_{1,0} = 1/\lambda$) is the inverse of the node's *intrusion rate* (IR) (λ) (see (3), (4), and (5) in the Appendix).

The ETTF is a metric often used to obtain a quick intuition about the reliance of a system in terms of time, e.g., about the duration of time for which the system should be trusted to hold some security property. Also, the MT of a system is often defined as a function of its ETTF. Thus, we now determine the circumstances in which the ETTF increases or decreases with the number of nodes (n). Let $\mu_{n,f}$ stand for the ETTF of a $\langle n, f\rangle$ system. By Definition 5, a system has a desirable ETTF if $\mu_{n,f} > \mu_{1,0}$ or, equivalently, when the ratio $\mu_{n,f}/\mu_{1,0}$ is higher than 1. We shall now analyze this ratio for different families of $\langle n, f\rangle$ configurations.

ETTF under parallel attack In this model, the ratio is $\mu_{n,f}/\mu_{1,0} = \sum_{i=n-f}^{n} 1/i$, as deduced in [25]. Intuitively, the $f + 1$ terms in the sum correspond to the $f + 1$ intrusions that would lead to a failure. Figure 2a shows curves for several cases, assuming a simultaneous unitary *intrusion adversarial effort* (IAE) upon each node, i.e., $\lambda_j(t) = 1 - \phi_j(t)$. In the extreme of higher ETTF is the type of system (●) that works correctly while at least one node is healthy ($f = n - 1$), having a ratio of $\mu_{n,n-1}/\mu_{1,0} = \sum_{i=1}^{n} 1/i$ (a sum with n terms). When the *intrusion tolerance threshold* ratio f/n decreases below a certain limit, the system eventually transitions to an undesirable ETTF. The *Lim FT* curve (■) illustrates, for several values of f, the limit case of desirable ETTF. Asymptotically (in the limit $n \to \infty$), the transition occurs for $f/n = (1 - 1/e) \approx 0.63$, with $e \approx 2.718$ being Euler's number. For lower f/n ratios, the global ETTF decreases while the threshold f increases, as seen in curves with $f = (n-1)/2$ (▼) and $f = (n-1)/3$ (♦), typically used in *Byzantine fault-tolerant* (BFT) systems. Though decreasing, for these cases the ETTF still converges to a positive value. For example, with $f = (n-1)/3$, the ETTF tends to $\log(3/2) \approx 40.5\%$ of $\mu_{1,0}$. In a further extreme, when the ratio f/n itself converges to 0, while increasing f, the ETTF also converges to 0, as shown with the *Sqrt FT* curve (▲), with $n = f^2 + 1$. The lowest ETTF happens without intrusion tolerance, i.e., $f = 0$ (only illustrated for $n = 1$), for which the ETTF decreases inversely proportional to n, i.e., $\mu_{n,0}/\mu_{1,0} = 1/n$.

ETTF under sequential attack In this model, the ETTF is much higher, with $\mu_{n,f}/\mu_{1,0} = f + 1$ (also deduced in [25]), if λ is fixed when varying n. Each node has an *expected time to intrusion* of $\mu_{1,0}$, out only when it starts being attacked. The higher increase of ETTF with f is now the result of a (*good*) dependence between the IAE on different nodes. Intuitively, a node being attacked draws all the attention from the attacker, and thus, while *healthy*, it protects the other nodes from being attacked. Figure 2b highlights the ETTF in function of n, for different $\langle f, n\rangle$ systems. Note that, if this graphic was plotted in function of f, all curves would superpose, as $\mu_{n,f}$ is now a pure function of f. The set of systems labeled as *Lim FT* (*b*, ■) is printed just as a curiosity, as for a sequential attack they do not correspond to any interesting threshold. The *stranger* form of this curve is due to the nonmonotonicity of the ratio f/n for the sequence of plotted points (enabling f from 0 to 6)—note in $f(n)$ the division by Euler's number (a non-integer).

Fig. 2 Expected time to failure (ETTF) under attack. In each sub-figure, each point (a marker along a dashed line) indicates the ETTF (the position in the vertical axis) of a specific intrusion-tolerant system $\langle n, f\rangle$, with n being the total number of nodes and f being the threshold of tolerated intrusions. The marker ✷ represents the reference system $\langle 1, 0\rangle$. Each other type of marker (●, ■, ▼, ♦, ▲) represents a specific functional relation between n and f (as detailed in the auxiliary box in the upper area of each sub-figure). The *vertical axis* (labeled ETTF) actually measures the ratio $\mu_{n,f}/\mu_{1,0}$, between the ETTF of the respective $\langle n, f\rangle$ system and the ETTF of $\langle 1, 0\rangle$. For $\lambda = 1$, it follows that $\mu_{1,0} = 1/\lambda = 1$, so the ratio is indeed the ETTF of $\langle n, f\rangle$. The *horizontal dashed line*, starting to the right of marker ✷, highlights the threshold between *desirable* and *undesirable* ETTF. The value to the right of each curve, and prefixed with a small arrow, indicates the limit ETTF as $f \to \infty$



(a) ETTF under *parallel* attack (with $\lambda = 1$). A constant and simultaneous *intrusion adversarial effort* (IAE) of unitary value is assumed upon each healthy node, implying an *intrusion rate* (IR) of $\lambda = 1$ in each node. The value inscribed inside each marker is n, the total number of nodes associated with the respective $\langle n, f\rangle$ system.

(b) ETTF under *sequential* attack (with $\lambda = 1$). A constant *intrusion adversarial effort* (IAE) of unitary value is assumed upon one healthy node at a time, implying a respective *intrusion rate* (IR) of $\lambda = 1$ in the node being attacked. The value inscribed inside each marker is f, the maximum number intrusions tolerated by the respective $\langle n, f\rangle$ system.

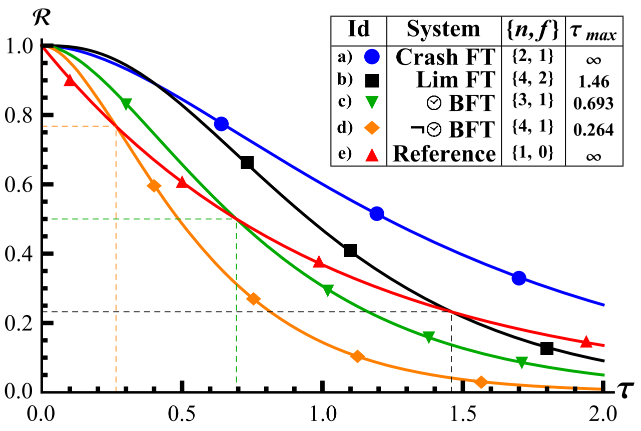


Fig. 3 Reliability ($\mathcal{R}$) under parallel attack. The horizontal axis measures the *mission time* (MT) τ in a scale normalized to the ETTF of the reference system $\langle 1, 0\rangle$, i.e., normalized to $\mu_{1,0} = 1/\lambda$. For each curve, associated with a specific $\langle n, f\rangle$ configuration, the vertical axis measures $\mathcal{R}^{\parallel}_{n,f}(\tau)$, and the respective $\tau_{\max}$ (in the rightmost column of the auxiliary box on the upper right) is the value satisfying $\tau \in [0, \tau_{\max}] \Leftrightarrow \mathcal{R}^{\parallel}_{h,f}(\tau) \ge \mathcal{R}^{\parallel}_{1,0}(\tau)$

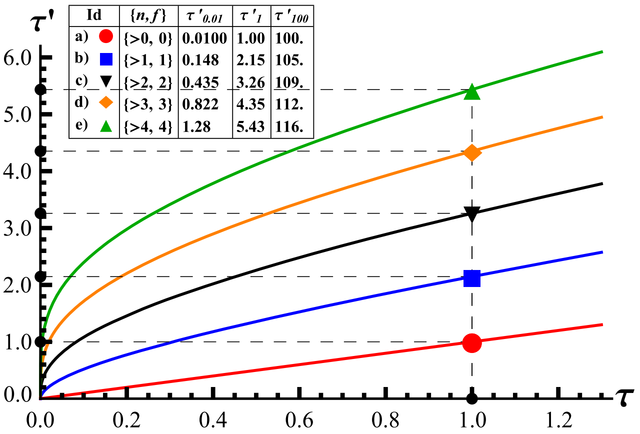


Fig. 4 Mission time (MT) for the same reliability as that of $\langle 1, 0\rangle$, under sequential attack. The *horizontal axis* measures the MT (τ) of the reference system $\langle 1, 0\rangle$. Each *curve* is associated with an *intrusion tolerance threshold* f, but is independent of the *replication degree* n. The *vertical axis* measures another MT (τ'), such that, if the curve associated with $\langle n, f\rangle$ includes point $\langle \tau, \tau'\rangle$, then $\mathcal{R}^{\therefore}_{n,f}(\tau') = \mathcal{R}^{\therefore}_{1,0}(\tau)$. Curve *a* (●), with $f = 0$, is the identity $\tau = \tau'$

In conclusion, the differences in types of attack ($\parallel$ versus $\therefore$), may make the difference between improving or worsening the ETTF of a system, when *augmenting* its configuration from $\langle 1, 0\rangle$ to $\langle n, f\rangle$. This should bring to attention the importance of considering architectural aspects that may limit the types of attack, when deciding on how to achieve *intrusion tolerance*.

3.2 Reliability per mission time

The ETTF is a useful metric, but there is no fundamental reason for it to be the desired MT. Thus, we now consider a more dynamic perspective and analyze the *reliability* ($\mathcal{R}$) for different MT values. We are interested in knowing *what are the mission times for which intrusion-tolerant replication does not worsen the reliability of a system, when compared to that of* $\langle 1, 0\rangle$. This information is important when one wants to define an adequate MT given a $\langle n, f\rangle$ system, or, vice-versa, select the best $\langle n, f\rangle$ system given a predetermined MT.

Henceforth, symbol τ shall be used to express time normalized to $\mu_{1,0} = 1/\lambda$, the *expected time to intrusion* (ETTI) of a node under attack. When considering this unit, one can assume $\mu_{1,0} = 1$ (and consequently $\lambda = 1/\mu_{1,0} = 1$). Equivalently, whatever λ, one can assume $\tau = t/\mu_{1,0} = \lambda t$, where t is (the wall-clock) time used to measure $1/\lambda$ (recall that λ is a rate).

The analytic formulas for $\mathcal{R}$ in the $\parallel$-attack model and in the $\therefore$-attack model are given in the Appendix (see (9) and (13), respectively).

Reliability under parallel attack Figure 3 and Table 1 show the variation of $\mathcal{R}^{\parallel}_{n,f}(\tau)$ for several pairs $\langle n, f\rangle$. When a *small* amount of time has passed, an intrusion-tolerant system with $f > 0$ has desirable $\mathcal{R}^{\parallel}$, because it is not yet likely that *many* nodes have been intruded. As time passes, more nodes are likely to have been intruded, and thus a low ratio f/n may imply lower $\mathcal{R}^{\parallel}$. In Fig. 3, we show solutions ($\tau_{\max}$) of the MT for which $\mathcal{R}^{\parallel}$ transitions from *desirable* to *undesirable*. In other words, $[0, \tau_{\max}]$ is the interval for which $\mathcal{R}^{\parallel}_{n,f}(\tau) \ge \mathcal{R}^{\parallel}_{1,0}(\tau)$.

For example, consider a context that requires $n = 3f + 1$ and for which each node under attack has an estimated *expected time to intrusion* of 1 year. In Table 1, we see that, when compared to $\langle 1, 0\rangle$, a system $\langle 4, 1\rangle$ has *desirable* $\mathcal{R}^{\parallel}$ for $\tau = 0.2$, i.e., a MT of 2.4 months, because $\mathcal{R}^{\parallel}_{4,1}(0.2) > \mathcal{R}^{\parallel}_{1,0}(0.2)$. However, for $\tau = 0.5$, i.e., a MT of 6 months, the respective $\mathcal{R}^{\parallel}$ is *undesirable*, because $\mathcal{R}^{\parallel}_{4,1}(0.5) < \mathcal{R}^{\parallel}_{1,0}(0.5)$. In Fig. 3, we see $\tau = 0.264$ as the transition value ($\tau_{\max}$) of $\langle 4, 1\rangle$.

This example clearly illustrates why *replication* is not on its own aligned with dependability—one must consider the intrusion tolerance threshold f and the mission time (or, more precisely, $MT/\mu_{1,0}$) before determining if a $\langle n, f\rangle$ intrusion-tolerant configuration brings an advantage or a disadvantage in terms of dependability (e.g., reliability). These mathematical results are already well established in the literature (e.g., see reliability of *Triple Modular Redundant* architectures for accidental faults in [13]). One of our contributions here is in highlighting the MT thresholds that make a difference between improvement and degradation of dependability, while having a security perspective in mind. Also, we call the attention to the impact of different adversarial characteristics of the environment in which a system might be placed (e.g., parallel versus sequential attack).

Table 1 Reliability ($\mathcal{R}$) under parallel attack. Each row corresponds to a different $\langle n, f\rangle$ system. Each column (below the top merged cell labeled $\mathcal{R}$) corresponds to a different MT τ, normalized to the ETTF of the reference system $\langle 1, 0\rangle$. The values of *desirable* reliability (i.e., those that are higher than the reliability of $\langle 1, 0\rangle$ for the same MT) are highlighted in slightly larger font size

System Type	*n*	*f*	$\mathcal{R}$				
			$\tau=0.2$	$\tau=0.5$	$\tau=1$	$\tau=2$	$\tau=5$
Reference	1	0	0.819	0.607	0.368	0.135	0.00674
No FT	2	0	0.670	0.368	0.135	0.0183	0.0000454
Crash FT	2	1	*0.967*	*0.845*	*0.600*	*0.252*	*0.0134*
⏲ BFT	3	1	*0.913*	*0.657*	0.306	0.0500	0.000136
Crash FT	3	2	*0.994*	*0.939*	*0.747*	*0.354*	*0.0201*
¬⏲ BFT	4	1	*0.847*	0.487	0.144	0.00891	1.22×10^{-6}
Lim FT	4	2	*0.979*	*0.828*	*0.469*	0.0911	0.000270
Crash FT	4	3	*0.999*	*0.976*	*0.840*	*0.441*	*0.0267*
⏲ BFT	5	2	*0.955*	*0.694*	0.264	0.0200	3.03×10^{-6}
¬⏲ BFT	7	2	*0.883*	0.434	0.0684	0.000751	2.88×10^{-10}
⏲ BFT	7	3	*0.976*	*0.723*	0.230	0.00834	7.10×10^{-8}
Lim FT	7	4	*0.997*	*0.910*	*0.509*	0.0568	0.0000105

On a more global look to Fig. 3 and Table 1, we note that different functional relations between n and f imply different MT-ranges of desirable $\mathcal{R}^{\parallel}$:

1. any MT—e.g., the *Crash FT* curve, in representation of any curve with $n = f + 1$, is higher than the *Reference* curve for any positive MT;
2. MT up to some $\tau_{\max} > 1$—e.g., the *Lim FT* curve, in representation of any curve with $n = \lfloor \frac{e}{e-1} f \rfloor + 1$ and $f \geq 2$, intersects the *Reference* curve for $\tau > 1$;
3. MT up to some $\tau_{\max} < 1$—e.g., the *BFT* curves, in representation of any curve with $n = 2f + 1$ or $n = 3f + 1$ (for $f > 0$), intersect the *Reference* curve for $\tau < 1$;
4. never—e.g., system $\langle 2, 0\rangle$ (see Table 1), in representation of any replicated but non-intrusion-tolerant system (i.e., $n > 1$ and $f = 0$), has lower $\mathcal{R}^{\parallel}$ than that of the *Reference*, for any positive MT.

Reliability under sequential attack In this model, the time required to intrude more than f nodes is independent of the total number of nodes (n). For any MT, the reliability always grows with the intrusion tolerance threshold f (see (14). Still, for any $\langle n, f\rangle$ system, reliability converges to 0 as time increases (see (14)).

For the sequential-attack model, a graphic equivalent to the one in Fig. 3 (i.e., $\mathcal{R}^{\iddots}$ versus τ) would have no curve intersections. Thus, we proceed directly to a new perspective, showing in Fig. 4 how an increase of f allows an increase of MT, from τ to τ', without changing the $\mathcal{R}^{\iddots}$ of the overall system. Note that the ratio τ'/τ is much smaller near $\tau = 1$ than it is for smaller values of τ. For example, a reference system $\langle 1, 0\rangle$ used for a MT of $\tau = 0.01$ has the same $\mathcal{R}^{\iddots}$ has an intrusion-tolerant system with $f = 4$ used for a MT of $\tau' = 1.28$, i.e., 128 times higher. However, if the reference of comparison is $\langle 1, 0\rangle$ for a MT of $\tau = 1$, then the replicated system with $f = 4$ has higher reliability only when used up to a MT of $\tau' = 5.43$, i.e., only 5.43 times higher. The solution of $\mathcal{R}^{\iddots}_{1,0}(\tau) = \mathcal{R}^{\iddots}_{n,f}(\tau')$ in order of τ' is presented in the Appendix (see (15)).

3.3 Time periods with *relative-resilience*

It is easy to understand what it means to increase the MT by a multiplicative factor. However, with *reliability* ($\mathcal{R}$), a probability, the scale is not linear and thus it may not be meaningful to ask for a linear improvement of $\mathcal{R}$ (e.g., to improve $\mathcal{R}$ by a factor of 2). Nonetheless, in the interest of intuition, we would like to be able to make comparisons in a linear scale, while still relating with the concept of reliability. To deal with this, we define a new metric, to which we suggestively call *resilience* (ρ), increasing linearly with the number of bits with which $\mathcal{R}$ is close to 1.[2] In other words, improving ρ by one unit means increasing the $\mathcal{R}$ by halving its distance to 1 (see (16) in the Appendix).

We can now make significant questions in a linear scale, such as: *what are the values of mission time (MT) for which the resilience (ρ) of $\langle n, f\rangle$ is at least c times higher than that of* $\langle 1, 0\rangle$ (see (17) and (18) in the Appendix). Note that we may talk about a relative-resilience improvement brought upon by a $\langle n, f\rangle$ configuration, if $c > 1$, even though the absolute resilience ($\rho_{n,f}(t)$) decreases with time (i.e., with the increase of MT) for any $\langle n, f\rangle$ configuration. We emphasize that, consistently with the enunciated goals of this paper, this is an objective way of measuring a *dependability improvement* brought upon by intrusion-tolerant replication in our system model.

Resilience under parallel attack Table 2 presents some numerical solutions for the periods of MT for which a $\langle n, f\rangle$ system, under parallel attack ($\parallel$), should be designed for when intending a certain relative-resilience factor (c). Some interesting facts:

- Every $\langle n, f\rangle$ system has a maximum relative-resilience factor that it can sustain. For $n = f + 1$, any factor (c) is valid either for any MT ($\tau_{\max} = \infty$) or for none at all ($\tau_{\max} = 0$). For the other illustrated systems, any $c > 0$ is valid only for a finite duration.

[2] This approach can be found in related areas. For example: the "nines of availability" counts the nines in the decimal expansion of the value of *availability* ($\mathcal{A}$); a cryptographic algorithm is sometimes said to have a security strength of k bits, if *breaking* an encryption requires an amount of work equivalent to what would take, for a certain reference symmetric encryption algorithm with key-size k, to find an encryption key by trial and error (i.e., an exhaustion attack in a space of size 2^k). We differ from the "nines of $\mathcal{A}$" example by using a base 2 (binary) instead of 10 (decimal), and differ from both examples by having a measure in the domain of reals, instead of just integers.

Table 2 Time Periods (τ) with relative-resilience (c) under parallel attack ($\parallel$). Each row corresponds to a different $\langle n, f\rangle$ configuration. Each column (below the top merged cell defining $\tau_{\max}$) corresponds to a specific relative-resilience factor c. τ is a measure of time normalized to the ETTF of the reference system $\langle 1, 0\rangle$, i.e., such that $\mu_{1,0} = 1$. Each cell, intersection of a column with value c and a row with configuration $\langle n, f\rangle$, contains the maximum mission time value ($\tau_{\max}$) for which the relative-resilience of the $\langle n, f\rangle$ configuration is at least c. Values $\tau_{\max}$ are highlighted in slightly larger font-size if the respective c is valid for τ up to at least 1

System Type	n	f	$\tau_{max} = \left[max\,(\tau) : \rho^{\parallel}_{n,f}(\tau) \geq c \times \rho^{\parallel}_{1,0}(\tau)\right]$						
			c=0.1	c=0.5	c=1	c=1.25	c=1.5	c=2	c=3
Reference	1	0	∞	∞	∞	0	0	0	0
No FT	2	0	*2.25*	0.481	0	0	0	0	0
Crash FT	2	1	∞	∞	∞	∞	∞	∞	0
⏲ BFT	3	1	*3.36*	*1.59*	0.693	0.382	0.144	0	0
Crash FT	3	2	∞	∞	∞	∞	∞	∞	∞
¬⏲ BFT	4	1	*1.73*	0.746	0.264	0.120	0.0306	0	0
Lim FT	4	2	*4.06*	*2.33*	*1.46*	*1.15*	0.871	0.405	0
Crash FT	4	3	∞	∞	∞	∞	∞	∞	∞
⏲ BFT	5	2	*2.19*	*1.20*	0.693	0.512	0.360	0.129	0
¬⏲ BFT	7	2	*1.14*	0.579	0.296	0.201	0.128	0.0319	0
⏲ BFT	7	3	*1.78*	*1.07*	0.693	0.559	0.445	0.259	0.0313
Lim FT	7	4	*2.82*	*1.86*	*1.36*	*1.18*	*1.03*	0.761	0.337

- With replication (i.e., $n > 1$) a lack of intrusion-tolerance (i.e., $f = 0$) always implies lower resilience, i.e., for any MT the relative-resilience is always lower than 1 (see (19) in the Appendix).
- For any $f \geq 1$, some ρ-improvements (i.e., $c > 1$) can be obtained for a small MT. However, only large ratios f/n allow ρ-improvements for large MT.

As an example, consider a MT $\tau_{\max} = 0.0319$, as obtained in Table 2 for $c = 2$ and $\langle 7, 2\rangle$ (a possible BFT system with configuration $n = 3f + 1$). Using (9) and (16) (in the Appendix), we calculate the *reliability* ($\mathcal{R}$) and respective *resilience* (ρ):

- for $\langle 1, 0\rangle$, $\mathcal{R}_{1,0}(0.0319) \approx 96.9\% \Rightarrow \rho \approx 5.0$;
- for $\langle 7, 2\rangle$, $\mathcal{R}_{7,2}(0.0319) \approx 99.9\% \Rightarrow \rho \approx 10.0$.

Thus, a system $\langle 7, 2\rangle$ is approximately 2 times ($c \approx 10.0/5.0 = 2$) more resilient than the reference (non-replicated) system $\langle 1, 0\rangle$, for a mission time of $t \approx 0.0319 \times \mu_{1,0}$. If $\mu_{1,0}$ (the ETTI of each node) is 1 year, then the (at least) double resilience is valid for a MT of about 11.4 days (0.0319 × 1 year).

Resilience under sequential attack Under sequential attack, the resilience increases with the intrusion tolerance threshold f, as a consequence of the reliability also increasing. In Table 3 we show some numerical solutions relating MT (τ) and relative-resilience factors (c), for several values of f. An interesting qualitative difference can be noted in comparison with the parallel attack model. In the sequential model, even though the absolute resilience still decreases with the increase of time, the relative-resilience factor actually increases with MT (thus Table 3 refers to $\tau_{\min}$, instead of $\tau_{\max}$).

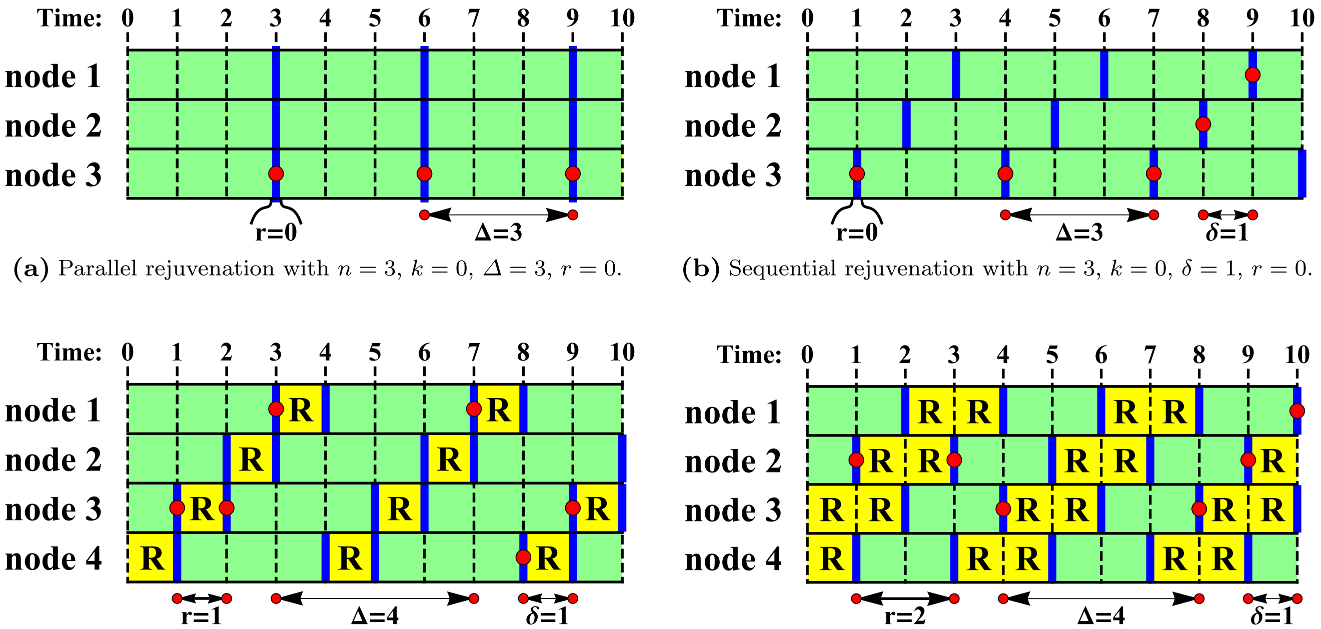
Table 3 Time periods (τ) with relative-resilience (c) under sequential attack ($\therefore$). Each row corresponds to a different intrusion tolerance threshold f, for which any $\langle n, f\rangle$ system with $n > f$ applies. Each column (below the top merged cell defining $\tau_{\min}$) corresponds to a specific relative-resilience factor c. τ is a measure of time normalized to the ETTF of the reference system $\langle 1, 0\rangle$, i.e., such that $\mu_{1,0} = 1$. Each cell, intersection of a column with value c and a row with value f, contains the minimum mission time ($\tau_{\min}$) for which the relative-resilience of the $\langle n, f\rangle$ configuration is at least c. 0^+ indicates that any positive value of mission time satisfies the condition of relative-resilience higher than c (note that the comparison operation is $>$ and not $\geq$, so that $\tau_{\min}$ is not trivially 0 for any c). Values $\tau_{\min}$ are highlighted in slightly larger font-size if the respective c starts before some τ smaller than 1. $\tau_{\min}$ is 0^+ whenever $c \leq f + 1$

System Type	n	f	$\tau_{min} = \left[min\,(\tau) : \rho^{\therefore}_{n,f}(\tau) > c \times \rho^{\therefore}_{1,0}(\tau)\right]$						
			c=1	c=2	c=2.5	c=3	c=4	c=5	c=10
f=0	>0	0	*0⁺*	∞	∞	∞	∞	∞	∞
f=1	>1	1	*0⁺*	*0⁺*	*0.382*	1.15	2.56	3.78	8.99
f=2	>2	2	*0⁺*	*0⁺*	*0⁺*	*0⁺*	*0.217*	*0.742*	2.77
f=3	>3	3	*0⁺*	*0⁺*	*0⁺*	*0⁺*	*0⁺*	*0.0450*	1.33
f=4	>4	4	*0⁺*	*0⁺*	*0⁺*	*0⁺*	*0⁺*	*0⁺*	*0.627*

4 Availability and the role of rejuvenations

In this section, we analyze the dependability enhancement brought upon by the use of *proactive rejuvenation* [6, 18, 22]. *Rejuvenation* is a process that restores the state of a node to *healthy*, regardless of its previous state. Consistently with our model of intrusions and attacks (Assumptions 1 and 2), we assume that the eventual intrusion of a node, at a given time, does not make easier the future intrusion of other nodes, not even of the same node after rejuvenation. This type of independence is usually achieved by the use of *diversity*, which might be effective for certain vectors of attack. Within our scope, we keep agnostic to the implementation of diversity, simply assuming that it might be effective in some cases of practical interest, and thus we measure dependability in a conservative way.

In the previous section, we omitted the analysis of *availability* ($\mathcal{A}$). When not considering *rejuvenations*, $\mathcal{A}$ can be deduced by integrating the *reliability* ($\mathcal{R}$) across time and normalizing the result to the MT (see (19) in the Appendix). Both $\mathcal{R}$ and $\mathcal{A}$ increase with rejuvenations, because it becomes more difficult for an attack to succeed in surpassing the *intrusion tolerance threshold* f. However, with rejuvenations $\mathcal{A}$ has the extra benefit of accounting also the moments of correctness obtained after a first global failure. Thus, $\mathcal{A}$ is positive even for an infinite *mission time* (MT). This is pertinent whenever global failure is not considered as a catastrophic event and the re-establishment of service is considered worthy. The focus of $\mathcal{A}$ is not on the first global failure (probability of never failing), but instead on the accumulated delivery of service (probability of not being failed at a random instant).



(a) Parallel rejuvenation with $n = 3$, $k = 0$, $\Delta = 3$, $r = 0$.

(b) Sequential rejuvenation with $n = 3$, $k = 0$, $\delta = 1$, $r = 0$.

(c) Sequential rejuvenation with $n = 4$, $k = 1$, $\delta = 1$, $r = 1$.

(d) Sequential rejuvenation with $n = 4$, $k = 2$, $\delta = 1$, $r = 2$.

Fig. 5 Timeline of different Rejuvenation Models. In each sub-figure, time flows in the horizontal axis from left to right, starting at $t = 0$. Each *row* represents the timeline of one of the n nodes in a system. The *empty slots* stand for online time; the *slots with letter R inside* stand for (offline) rejuvenation time. In each column, there are exactly $n - k$ nodes online and k nodes rejuvenating. The *thicker vertical segments* mark the beginning and ending instants of each rejuvenation, either as an instantaneous process ($r = 0$) when $k = 0$ (sub-figures 5a and 5b), or as a process taking some time ($r > 0$) when $k > 0$ (sub-figures 5c and 5d). The auxiliary *small circles*, on top of some thick vertical segments, examplify the horizontal extremities of the measure of some parameter of the rejuvenation scheme (as respectively exemplified below the lowest timeline row): r is the time a node takes in each rejuvenation; Δ is the period between the beginning of consecutive rejuvenations of the same node; δ is the minimum time between rejuvenations of different nodes

4.1 Extended system model

If we could detect attacks and/or intrusions, then a *reactive rejuvenation* scheme could be implemented [21]. For example: a detected attack could be mitigated by rejuvenating components more frequently; a detected intrusion could be amended by immediately rejuvenating the respective node. However, in our context of stealthiness, we can rely only on proactive rejuvenation schemes, of which we shall describe two models: *parallel* ($\parallel$) and *sequential* ($\therefore$).

Assumption 3 formalizes both types of rejuvenation and Fig. 5 illustrates the timeline of node rejuvenations for several specific rejuvenation schemes.

Assumption 3 (Periodic Rejuvenations) Let $n > 0$ be the total number of nodes in a system that initiates its operation at instant 0. At any instant of time $t > 0$, let $k \in \{0, \dots, n-1\}$ be the constant number of rejuvenating (offline) nodes and let $n' = n - k$ be the number of online nodes. Let $\Delta > 0$ be the (periodic) interval of time between the beginning of rejuvenations of the same node. Let δ (with $0 \leq \delta < \Delta$) be the smallest time between the beginning of rejuvenations of different nodes (δ is equal to 0 if different nodes rejuvenate simultaneously). Let $r \geq 0$ be the time duration of each rejuvenation of any node. Nodes 1 through n' become online for the first time simultaneously at instant 0. For $j \in \{n'+1, \dots, n\}$, node j becomes online for the first time at instant $(n - j + 1) \times \delta$; before that it is considered to be in its 0th rejuvenation. For $j \in \{1, \dots, n\}$, node j begins its ith *rejuvenation* (with $i \in \mathbb{N}_1$), at instant $(n' - j + 1) \times \delta + (i - [j \leq n']^?) \times \Delta$, where $[j \leq n']^?$ is 1 if $j \leq n'$ and 0 otherwise. Moreover, k is a constant integer satisfying $r = \delta \times k$ and $r = \Delta \times (k/n)$. Rejuvenation schemes are distinguished in two types: *parallel* ($\parallel$), if $\delta = 0$, or *sequential* ($\therefore$) otherwise. A system without rejuvenation is denoted as a $\parallel$-rejuvenating system with $\Delta = \infty$.

Figure 5 illustrates clearly some differences between the timelines of distinct types of rejuvenations ($\parallel$ and $\therefore$) and different number of simultaneous rejuvenating nodes (k). For parallel rejuvenations (Fig. 5a): nodes rejuvenate simultaneously ($\delta = 0$) after every interval of Δ time units; since (by assumption) the duration of rejuvenation of each node is proportional to δ, if follows that rejuvenations are instantaneous[3] ($r = \delta \times k = 0$) and thus nodes are never offline

[3] In Section 4.4, when making a practical comparison between different types of rejuvenation, we shall substantiate the possibility of instanta-

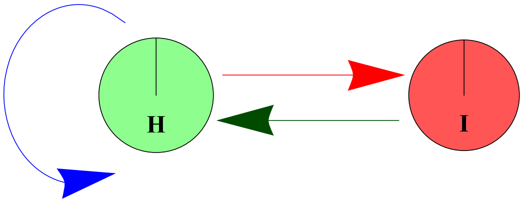


Fig. 6 State diagram of a rejuvenating system with $n = 1$ node, under attack. Each circle represents the state of the single node: *healthy* (H) or *intruded* (I). A rejuvenation *heals* ($H \rightarrow H$ or $I \rightarrow H$) the node. An intrusion *intrudes* ($H \rightarrow I$) the node

for a continuous amount of time (i.e., $k = 0$). For sequential rejuvenations, nodes can also rejuvenate instantaneously, if $k = 0$ (Fig. 5b), but each one does so at a different instant in time (i.e., $\delta > 0$); if $k = 1$ (Fig. 5c) then once a node becomes online another one immediately starts rejuvenating; finally, if $k > 1$ (Fig. 5d), then several nodes can be in rejuvenating state simultaneously, but starting at different instants of time. In any case, each node is *online* for durations of $\Delta - r$, interleaved with *offline* durations of r, and the number of *online* nodes is a constant ($n' = n - k$).

Extended model By combining the models of attack, intrusion, and rejuvenation, we get an extended model where healthy nodes can be intruded and then be reverted back to a healthy state. In this new system model, n accounts also with k offline nodes. In typical systems, the parameters n, f and k are related in a linear way, i.e., as $n = af + bk + c$, for non-negative integers a, b, and c. For simplicity, we shall restrict the remaining comparison examples to cases with $b = 1$ and $c = 1$. Thus, a triplet $\langle n, f, k \rangle$ will henceforth be used, to denote the full constitution of the system in terms of numbers of nodes (an exception is made to the reference system $\langle n, f \rangle = \langle 1, 0 \rangle$, which clearly implies $k = 0$). We assume that attacks can influence the rate of state transitions (as determined in Assumption 1), but cannot influence the schedule of rejuvenations.

For each $\langle n, f, k \rangle$ system, the rejuvenation duration (r) of a node is related with the periodicity of rejuvenations by $r = \Delta \times (k/n)$ or $r = \delta \times k$, respectively for rejuvenations of type $\|$ or $\therefore$. Thus, we may characterize a system with only two extra parameters in subscript:

- $\langle \|, \Delta \rangle$: *parallel* ($\|$) rejuvenations with period Δ, and assuming $\delta = 0$.

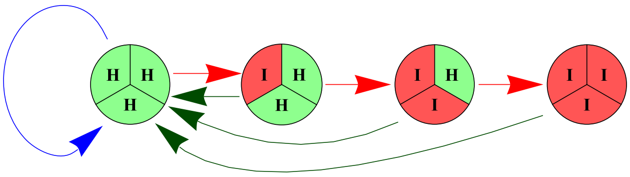


Fig. 7 State diagram for parallel rejuvenation of a system with $n = 3$ nodes, under a particular choice of sequential attack. Each *circle* represents the set of 3 nodes and their states. For parallel rejuvenations, the order in which nodes are intruded is irrelevant, and only the number of intruded nodes matter. Thus, a more general interpretation (suitable also for the case of parallel attack), considers that each circle with i triangles in state I is representative of all global states with exactly i nodes intruded

- $\langle \therefore, \delta \rangle$: *sequential* ($\therefore$) rejuvenations, with consecutive nodes being rejuvenated at instants separated by δ, and with $\Delta = n \times \delta$.

In terms of parameters characterizing the external environment, we shall continue to use $\|$ or $\therefore$ for the type of attack (parallel or sequential) and λ for the *intrusion adversarial effort* (IAE) upon each node.

Types of rejuvenation The choice of rejuvenation type might not be arbitrary. By assuming a scenario of stealth attacks and intrusions, proactive rejuvenations must be implemented with a protocol that is resilient to intruded nodes, even though they might be indistinguishable from healthy ones. For example, if the system implements non-stop operations, the rejuvenation process might require transfer of state from online nodes to rejuvenating nodes, thus making a *sequential* rejuvenation scheme more appropriate than a *parallel* one. In such cases, parameters k and r are relevant in terms of implementation. Actually, an eventual inability to enforce a fixed bounded limit on r may result in security vulnerabilities for some protocols, as noted in [22].

The 2 models of attack and 2 models of rejuvenation give 4 possible types of combinations. However, for a system made of a single node ($n = 1$) all combinations collapse into the same model—Fig. 6 shows the respective state diagram. Notably, for the reference system $\langle 1, 0 \rangle$ (or actually any other with $f = 0$), *rejuvenation* does not affect *reliability* ($\mathcal{R}$), because: (1) the *intrusion* of a node corresponds to the immediate failure of the system; and (2) the rejuvenation of a healthy node does not alter its *intrusion rate* (IR). Consequently, if there is no intrusion tolerance then a $\mathcal{R}$-improvement can only be obtained by using more reliable nodes. Nevertheless, availability ($\mathcal{A}$) is improved with rejuvenation even for the reference case with $n = 1$. For $n > 1$, we analyze the models separately.

4.2 Parallel rejuvenation

In each instantaneous *parallel* rejuvenation, a $\langle n, f \rangle$ system (necessarily with $k = 0$) is reset to a completely healthy

neous rejuvenations by considering the existence of virtual nodes (*virtual* in the sense of never being online, and not being accounted in parameter n), whose role is only to help preparing the future instantaneous rejuvenation of *real* nodes. In that case we shall still refer to instantaneous rejuvenations, even though r will be considered as a positive value given by $r = \Delta \times (k + vk)/n$, where vk is the number of virtual nodes.

state, i.e., with all nodes healthy (see (20) in the Appendix). As an example, Fig. 7 shows the state diagram for a system with $n = 3$ and $k = 0$, subject to parallel rejuvenations. In comparison with Fig. 1b, only the rejuvenation transitions were added.

In the parallel rejuvenation model, the overall *reliability* as a function of time ($\mathcal{R}_{n,f,\|,\Delta}(t)$) can be obtained as a product of *reliabilities* ($\mathcal{R}_{n,f}$) for time-windows of width Δ (i.e. $\mathcal{R}_{n,f}(\Delta)$) and less (i.e. $\mathcal{R}_{n,f}(m)$ for some $m < \Delta$) (see (21) in the Appendix). For $\langle n, f\rangle$ systems with $f > 0$, rejuvenation might heal *intruded* nodes before the number of simultaneous intrusions exceeds f. Recalling Fig. 3, we conclude that intrusion-tolerant replication and rejuvenation may have complementary roles in dependability:

- *intrusion-tolerant replication*, with $f > 0$, improves $\mathcal{R}$ for small MT, *but* for small ratios f/n it is prejudicial for large MT;
- *rejuvenation* cannot bring benefits before its first application, *but* it reduces the long-term degradation effects on dependability, by periodically bringing the system back to its initial overall state (i.e., with all nodes healthy—see (20) in the Appendix).

By applying both techniques together (rejuvenation and intrusion-tolerant replication), the $\mathcal{R}$ improvement might be valid even for an unbounded MT (finite but not known in advance). To achieve such overall improvement, a $\langle n, f\rangle_{\|,\Delta}$ system must have a low enough period Δ, namely less than the threshold value of time (in Fig. 3) for which $\langle n, f\rangle$ (without rejuvenation) transitions to *undesirable* $\mathcal{R}$. In this way, even configurations $\langle 3, 1\rangle$ and $\langle 4, 1\rangle$ under parallel-attack may have *desirable* $\mathcal{R}$. This amends the negative result (for dependability) that we had achieved with the preliminary system model. However, if $MT = \infty$ then $\mathcal{R}_{n,f,\|,\Delta}$ is simply 0, whereas $\mathcal{A}_{n,f,\|,\Delta}$ is still positive (see (22) in the Appendix).

4.3 Sequential rejuvenation

A more challenging analysis is that of sequential rejuvenations, for which there is no periodic interval for which the overall system state is reset, even though the instants of rejuvenation are periodic. This happens because nodes are rejuvenated one at a time, thus not guaranteeing that the number of intruded nodes goes back to 0. In particular, a strong-enough attacker may be able (probabilistically) to intrude nodes at a faster pace than their rejuvenation. As a consequence, the number of intruded nodes may potentially be maintained above the threshold f for durations much longer than $\delta \times n$. Moreover, for a sequential-attack there are paths of attack with different effectivenesses, because of their relation with the ordering of rejuvenations. In this respect, *we always assume an optimal IAE sequence*, from the point of view of the attacker, as stated in Assumption 4.

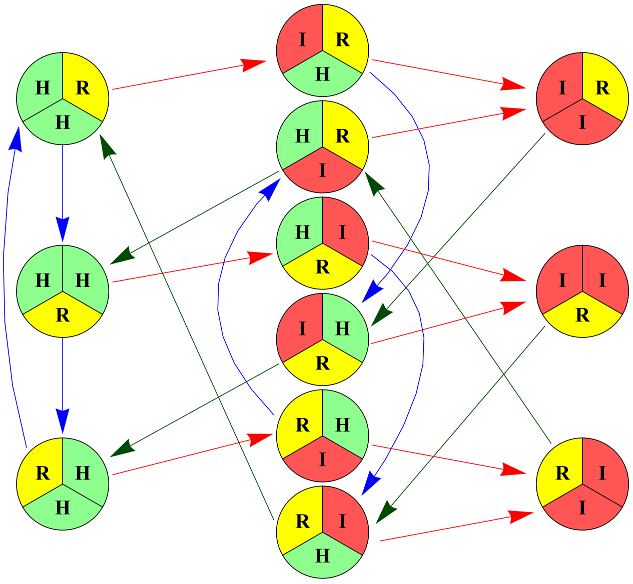
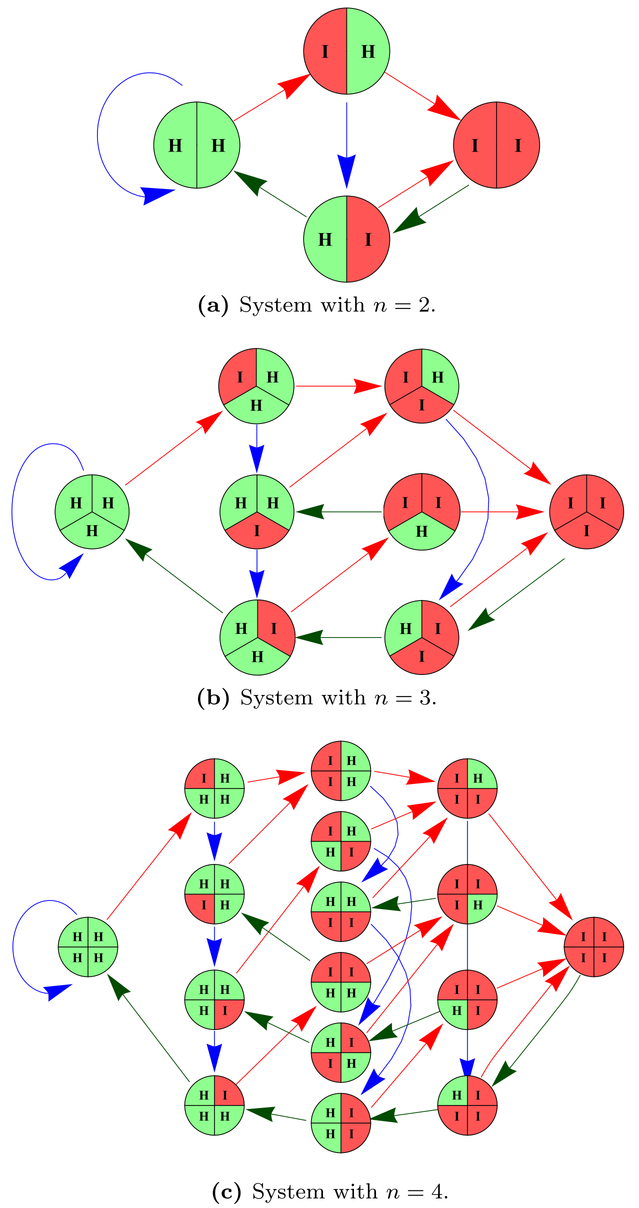


Fig. 8 State diagram of sequential rejuvenation of a system with $n = 3$ and $k = 1$, under optimal (i.e., most effective) sequential attack. Each *circle* represents a set of $n = 3$ nodes and their states: *healthy* (H), *intruded* (I) or *rejuvenating* (R). The types of transition are illustrated with arrows with different directions: *rightward*, when a previously healthy node transitions to intruded state; *leftward*, when a previously intruded node transitions to rejuvenating state; *vertical* (*downward* or *upward*), when a previously healthy node starts being rejuvenated

Assumption 4 (Optimal IAE sequence) Under sequential rejuvenation, a sequential attack always targets the (yet) healthy node which will remain un-rejuvenated for the longest time.

State diagrams showing rejuvenations Figure 8 shows a state diagram for a system with $n = 3$ and $k = 1$, where there are always 2 online nodes and 1 rejuvenating (offline) node. Some transitions triggered by rejuvenation occur between circles in the same column, as they correspond to the starting of rejuvenation of a previously healthy node, thus keeping constant the number of intruded nodes. Also, given our assumption of an optimal attack sequence, each circle of the leftmost column has only one outbound arrow corresponding to intrusion, leading to a circle in the middle column for which the next rejuvenation will not reduce the number of intrusions. In this diagram there are cycles that never go back to a completely healthy state, contrarily to what would happen in a diagram for parallel rejuvenations (e.g., Fig. 7).

State diagrams hiding rejuvenations In our model of rejuvenations, a node being rejuvenated is offline, and thus not available for interaction, namely not available to be attacked. Thus, from the point of view of an attacker, there are only $n' = n - k$ nodes available and each rejuvenation is done instantaneously ($r = 0$). If in Fig. 8 we remove (hide)



(a) System with $n = 2$.

(b) System with $n = 3$.

(c) System with $n = 4$.

Fig. 9 State diagrams of sequential rejuvenations with $k = 0$, under optimal sequential attack. Each sub-figure depicts the state diagram of a system with a different number of nodes ($n \in \{2, 3, 4\}$). In each sub-figure, each *circle* represents a set of n nodes and their states: *healthy* (H) or *intruded* (I). A rejuvenation *heals* ($H \rightarrow H$ or $I \rightarrow H$) the right-upper triangle and then rotates the circle counter-clock-wise by $2\pi/n$ (i.e., by $1/n$ of a full circle rotation). An intrusion *intrudes* ($H \rightarrow I$) the healthy triangle further away (in time) being healed

the rejuvenating node of $\langle 3, f, 1\rangle$, we are left with a system of only two nodes, i.e., $\langle 2, f, 0\rangle$, as shown in Fig. 9a. In other words, Fig. 8 can be mapped onto Fig. 9a, by mapping each 3 circles and 3 arrows onto an equivalent single circle and arrow, in the same column. Henceforth, we shall use these simplified diagrams that hide the complexity of rejuvenating nodes. Also, when making simulations to determine the availability of systems with sequential rejuvenation, we shall actually simulate $\langle n - k, f, 0\rangle$ instead of $\langle n, f, k\rangle$, which is equivalent, after doing the necessary adjustments to the values r and Δ.

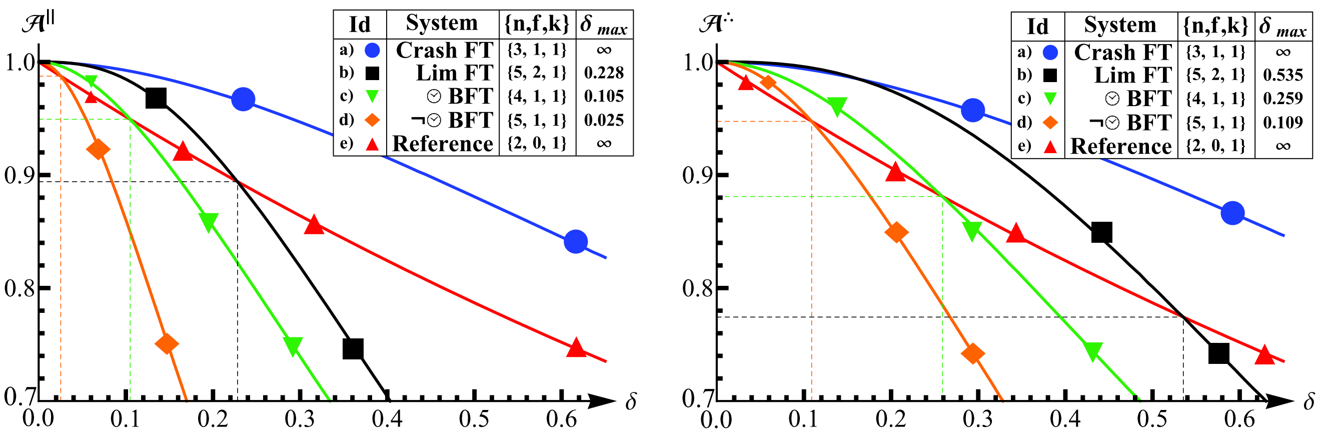
Consider in Fig. 9a the rejuvenating transition from (I|H) to (H|I), represented by the downward arrow (↓) in the middle. Note that, despite the flipping of positions of letters I and H, the same nodes remain intruded and healthy. What changes is the time-distance that the intruded node is away from its future rejuvenation. The transition corresponds to moving from a state with *a single intruded node being two rejuvenating-steps way from healing*, to a state with *the same intruded node being only one step away from healing*. In other words, the transition corresponds to the case where the rejuvenation is applied to a healthy node, thus not altering the number of intruded nodes of the system, but bringing the intruded node one step closer (in time) to being rejuvenated (i.e., it will be healed in the next rejuvenating step). Also, note that from the leftmost column to the middle column only one intrusion arrow exists—the arrow (↗) going from (H|H) to (I|H). This is consistent with Assumption 4, under which a sequential attack always targets the healthy node that is further away from rejuvenation.

We have seen how to go from Fig. 8 to Fig. 9a. Following the same logic, we can simplify the analysis of sequential-rejuvenating systems with non-instantaneous rejuvenations (i.e., with $k > 0$) for other values of n. To simplify, we shall look instead to the system with $n' = n - k$ nodes and $k' = 0$ offline nodes at any time (and consequently with instantaneous rejuvenations). In the remainder of this section, we shall compare different $\langle n, f, k\rangle$ systems, the biggest of which being $\langle 4, 1, 1\rangle$ (i.e., $n = 2f + k + 1$, with $f = 1$ and $k = 1$) and $\langle 5, 1, 1\rangle$ (i.e., $n = 3f + k + 1$, with $f = 1$ and $k = 1$). Their respective equivalent diagrams are depicted in Fig. 9b ($n = 3$ and $k = 0$) and Fig. 9c ($n = 4$ and $k = 0$).

The rules of probabilistic transition between states are easy to define and simulate. As an example, Fig. 10 shows results of availability ($\mathcal{A}$) for a parallel attack model (sub-figure 10a) and for sequential attack model (sub-figure 10b), when varying δ (the offset between sequential rejuvenations). We consider cases with $k = 1$, and thus $\delta = r$. Let $n' = n - k$. The curves shows that different $\langle n', f\rangle$ systems have *desirable* $\mathcal{A}$ (i.e., higher than that of $\langle n', f\rangle = \langle 1, 0\rangle$) for different offsets δ of rejuvenations: for any δ if $\langle n', f\rangle = \langle 2, 1\rangle$; only for $\delta \lesssim 0.10$ or $\delta \lesssim 0.26$, if $\langle n', f\rangle = \langle 3, 1\rangle$, under $\|$ or $\therefore$ attack, respectively; only for $\delta \lesssim 0.024$ or $\delta \lesssim 0.11$ if $\langle n', f\rangle = \langle 4, 1\rangle$, under $\|$ or $\therefore$ attack, respectively.

4.4 A practical comparison of configurations

So far we have compared a few $\langle n, f, k\rangle$ configurations, two models of attack, two models of rejuvenation, and a few perspectives of parameter selection. In real cases, further practical restrictions may condition the criteria for optimal configuration. As an illustrative comparison example, consider



(a) Under parallel attack (with $\lambda = 1$).

(b) Under sequential attack (with $\lambda = 1$).

Fig. 10 Availability ($\mathcal{A}$) with *sequential* rejuvenations. Each sub-figure corresponds to an environment under a particular type of attack: parallel ($\|$) in sub-figure 10a; sequential ($\therefore$) in sub-figure 10b. In each sub-figure: each curve corresponds to a $\langle n, f, k\rangle$ system, where n is the total number of nodes, f is the threshold of tolerable intrusions, and k is the number of rejuvenating (offline) nodes at any given instant; the horizontal axis measures δ, the time offset between rejuvenations of different nodes; the vertical axis measures $\mathcal{A}$, the expected proportion of time for which the number of intruded nodes is at most f; in the rightmost column of the auxiliary box in the upper right corner of each sub-figure, each value $\delta_{\max}$ is the maximum value of parameter δ for which the $\mathcal{A}$ of the respective $\langle n, f, k\rangle$ system is better (i.e., higher) than that of the reference system (curve e, ▲); the reference curve was obtained from the analytical expression $(1 - e^{r})/r$; all other curves (i.e., for systems with $f > 1$) were obtained by joining pairs $(\langle\delta, \mathcal{A}^{\|}\rangle)$ or $(\langle\delta, \mathcal{A}^{\therefore}\rangle)$, with δ spaced in intervals of at most 0.01, and with $\mathcal{A}^{\|}$ or $\mathcal{A}^{\therefore}$, respectively, being an average over the result of 100 probabilistic simulations with a *mission time* $\delta \times 10^5$

that an intrusion-tolerant system must be built, subject to the following constraints:

1. the underlying protocol requires $n = 2f + k + 1$, e.g., a typical synchronous or stateless BFT system with rejuvenation (e.g., [21]);
2. resources are limited to a maximum of 4 nodes;
3. considering two possibilities of implementation, the system may either be attacked sequentially with a focused IAE of $\lambda = 3$ per node, or in parallel with a dispersed IAE of $\lambda = 3/(n - k)$ per online node;
4. the rate at which nodes can rejuvenate is proportional to the number of available offline nodes, e.g., new (diversified) software replicas are generated using the computational resources of nodes that are not online.

With these restrictions, *what is the configuration that enables a higher $\mathcal{A}$, for an infinite MT?*

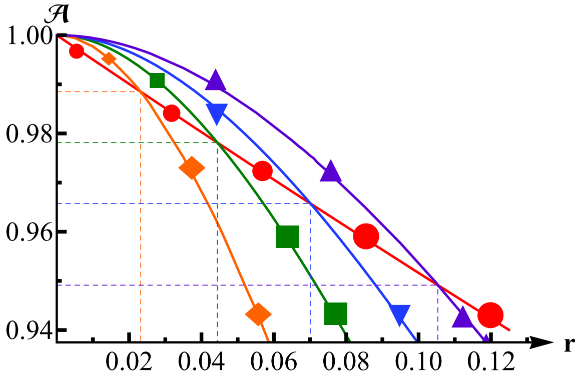
Making a fair comparison The instantaneous rejuvenation ($r = 0$) of nodes, in the case of parallel rejuvenations, still seems somewhat far-fetched. To substantiate it, we allow the existence of offline *virtual nodes* (vk), helping in the preparation of new replicas. We characterize them as *virtual* because they are not to be accounted in the value n (the total number of *real* nodes) as defined in Assumption 3. However, for the purpose of this example, we make the virtual nodes count toward the limit of 4 nodes, i.e., $n + vk = 4$. To compare different systems in an equal standing, we require that a virtual node must work for time r in order to prepare the instantaneous rejuvenation of a real node, where r is the exact same time that a (real) offline node takes to rejuvenate in a sequential rejuvenating scheme. Thus, we are now ready to consider the above question for different values of r.

Comparable scenarios Considering the restrictions and the guidelines for fair comparison just stated, we shall compare 5 different scenarios:

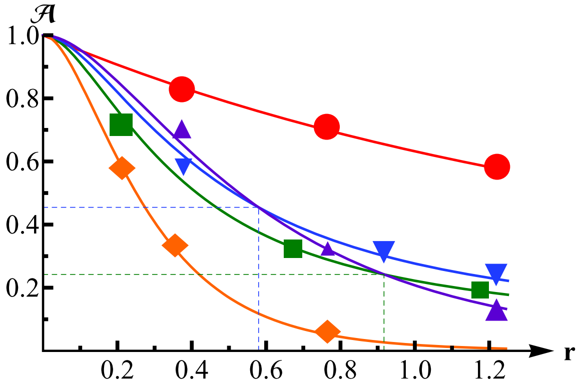
- *Single node case:* The reference system is characterized by a single node online at any time, i.e., $n - k = 1$. In this case, both parallel and sequential attacks are the same, and so we set $\lambda = 3$, in accordance to the above guidelines. Also in this case, both rejuvenating schemes are equivalent, and so we can choose arbitrarily between two notations. If seeing it as a $\therefore$ rejuvenating scheme, then $\langle n, f, k, vk\rangle = \langle 4, 0, 3, 0\rangle$ and $\langle rej, \delta, \Delta\rangle = \langle \therefore, r/3, (4/3)r\rangle$ and $\lambda = 3$. If seeing it as a $\|$ rejuvenating scheme, then $\langle n, f, k, vk\rangle = \langle 1, 0, 0, 3\rangle$, $\langle rej, \delta, \Delta\rangle = \langle \|, 0, r/3\rangle$ and $\lambda = 3$.
- *Two sequential-rejuvenation cases, with $f > 0$:* The configuration of nodes is limited to $\langle n, f, k, vk\rangle = \langle 4, 1, 1, 0\rangle$; defining the time parameters in terms of r we get $\langle rej, \delta, \Delta\rangle = \langle \therefore, r, 4r\rangle$. Finally, there are two distinct variants of this scenario: $\lambda = 1$ for $\|$-attack; $\lambda = 3$ for $\therefore$-attack.
- *Two parallel-rejuvenation cases, with $f > 0$:* The configuration of nodes is limited to $\langle n, f, k, vk\rangle = \langle 3, 1, 0, 1\rangle$. The virtual node has to prepare 3 rejuvenations per period Δ. The time parameters are again described in terms

Id	n	f	k	vk	rej	att	λ	Δ/r	δ/r	$r_{\max}$
a) ●	4	0	3	0	∴	∴	3	4/3	1/3	∞
b) ■	3	1	0	1	‖	∴	3	3	0	0.044
c) ▼	3	1	0	1	‖	‖	1	3	0	0.070
d) ◆	4	1	1	0	∴	∴	3	4	1	0.023
e) ▲	4	1	1	0	∴	‖	1	4	1	0.105

(a) Legend: n is the total number of *real* nodes; f is the threshold of tolerated intruded nodes; k is the number of *real* nodes (i.e., accounted in n) that are offline rejuvenating at any given instant; vk is the number of *virtual* nodes (i.e., not accounted in n) that are offline helping with the preparation of instantaneous rejuvenation of online nodes; *rej* denotes *rejuvenation type*; *att* denotes *attack type*; ‖ denotes *parallel*; ∴ denotes *sequential*; *any* denotes ∴ or ‖, arbitrarity, λ is the *intrusion adversarial effort* (IAE) placed in a node under attack, resulting in a similar *intrusion rate* (IR); Δ is the shortest time difference between rejuvenations of the same node; δ is the shortest time difference between rejuvenation of different nodes; r is the time that a node takes to rejuvenate if a single node (itself or a virtual node) applies its computational power to that effect; in each row, $r_{\max}$ is the maximum r for which the respective system has $\mathcal{A}$ not less than that of the reference system (curve, ●).



(b) Zoom In ($r \in [0, 0.125]$).



(c) Zoom Out ($r \in [0, 1.25]$).

Fig. 11 Availability ($\mathcal{A}$) in function of rejuvenation time (r) per node. Curves a) (●), b) (■) and c) (▼) were obtained from the respective analytic expressions of availability; the curves of sequential rejuvenation cases, were obtained by joining pairs $\langle r, \mathcal{A}\rangle$, with r spaced in intervals of at most 0.01, and with $\mathcal{A}$ being an average over the result of 100 probabilistic simulations with a mission time $\delta \times 10^5$

of r. $\langle rej, \delta, \Delta\rangle = \langle \|, 0, 3r\rangle$. This case also has two different variants, depending on the attack type: $\lambda = 1$ for ‖-attack; $\lambda = 3$ for ∴-attack.

For any of the 5 cases under comparison (see Fig. 11a):

- at any given time, there are: $n - k$ real nodes online; k real nodes rejuvenating; vk extra virtual nodes helping with the preparation of rejuvenations;
- the global rejuvenation period (i.e., the time between two rejuvenations of the same node) is $\Delta = r \times n/(k + vk)$;
- the minimum time between rejuvenations of different nodes is $\delta = r/k$ for ∴-rejuvenations, and $\delta = 0$ for ‖-rejuvenations;
- the sum of IAE across all healthy nodes is at most 3, i.e., $\sum_{j=1}^{n} \lambda_j(t) \le 3$—in particular, for a ‖-attack it is proportional to the number of healthy nodes, and for a ∴ it is constant while there is at least one healthy node.

In Fig. 11, we plot the availability of such systems, in function of parameter r (time required to recover each node). This figure shows interesting results:

1. For each rejuvenation type, a focused ∴-attack ($\lambda = 3$) is more effective than a dispersed ‖-attack ($\lambda = 1$). This was expected given that for the ‖-attack the sum of IAE (across all nodes) decreases with the number of healthy nodes. Moreover, when in ∴-rejuvenations, the ∴-attack is more effective by pursuing an optimal IAE sequence.

2. For each attack type, and with $f = 1$, as r grows, the availability of ∴-rejuvenations eventually becomes lower than that of ‖-rejuvenations—see Fig. 11b or Fig. 11c for curve b (■) versus curve d (◆); see Fig. 11c for curve c (▼) versus curve e (▲). This was expected, as ∴-rejuvenations cannot guarantee a periodic complete recovery. Thus, a fast enough intrusion of nodes (or, equivalently, a slow enough rejuvenation of nodes) may keep the system failed for a long time. Yet, it is interesting to see that, in the sole case of ‖-attack, the ∴-rejuvenation is more effective than the ‖-rejuvenation if $r < 0.58$. Our intuition is that this happens because, for a strong ‖-attack (or equivalently, for a low enough r), a ∴-rejuvenation allows a higher frequency of healing of eventually intruded nodes. This qualitative comparison between rejuvenation types does not hold for ∴-attacks, because the *optimal IAE sequence* targets first the nodes that are further always from rejuvenation.

3. As r grows, the system with lowest intrusion tolerance threshold ($f = 0$) but higher rejuvenation rate (i.e., lower Δ/r) eventually becomes more available than the alternatives. This means that, if single nodes cannot be rejuvenated quickly enough, then it is better to increase k than f.

5 Related work

Intrusion tolerance Much research has been done on intrusion-tolerant protocols (e.g., [6, 7, 26, 27]). We do not focus on protocols, but instead on high level properties, such as the functional relation between replication degree n and intrusion tolerance threshold f. One of our main motivations was to show that intrusion tolerance is not necessarily aligned with reliability or availability. Such alignment depends on a set of parameters that must be combined together in a way that gives rise to desirable dependability properties.

Reliability Reliability has been widely studied [2], both in theory and practice. Many works consider detailed estimations of reliability. The work in [14] (one out of many possible examples) studies a particular type of system and analyzes the probability that simultaneous faults actually lead to failure, thus distinguishing fatal from nonfatal faults. We instead follow a high level approach and, focusing on a context of malicious attack, base our estimates on simple and conservative modeling decisions: intrusions cannot be detected and any number of intrusions above the threshold implies immediate failure. Our analysis used some analytic results from [25].

Rejuvenations The effect of rejuvenation schemes is the topic of previous research works. For example, tradeoffs between proactive and reactive recoveries are evaluated in [8, 12, 21]. In a similar way, we compare different models of rejuvenation, but avoid reactive schemes, given our scope of stealth intrusions. The work in [23] mentions the infeasibility of enforcing a threshold of intrusions and considers proactive recovery as a possible mitigation. It also points out caveats in asynchronous systems that depend on synchronous rejuvenations. In this paper, we are not concerned with proving the feasibility of rejuvenations or rejuvenations—we just assume their possibility and then study the configurations that provide an enchancement to reliability and availability.

Diversity Much research has been done on the need of *diversity* in systems with rejuvenation (e.g., [4, 9, 15, 16, 18]) and on how to avoid common-mode failure (e.g., [11]). We do not address the problem of node vulnerabilities, but are instead just concerned with the specification of intrusions as the result of direct attack efforts. We are interested in finding configurations that allow the best dependability properties. Nevertheless, we show that degradation is possible even when intrusion independence exists.

6 Conclusions

In this paper, we showed how some (often neglected) parameters play an important role in determining the *reliability* and *availability* of intrusion-tolerant systems. We focused on the impact of mission time, rejuvenation strategy, and attack model. Based on our analytical and simulation-based study we found four main insights that should be taken into account when designing dependable systems based on intrusion-tolerant replication and rejuvenation:

(1) In order to assess the concrete benefits of replication and rejuvenation, it is important to specify the mission time, or, more precisely, its relation with the *expected time to intrusion* of individual nodes. Its non-specification allows opportunity for undesired levels of *reliability* and/or *availability*. For example, intrusion-tolerant replication may be counter-productive in the long term if parallel attacks are in place and malicious stealth intrusions are expected. Even a simple distinction between finite, unbounded or infinite mission time might help distinguishing configurations in terms of the dependability enhancement they provide.

(2) The choice of rejuvenation type—*sequential* or *parallel*—is important for the overall reliability and availability of the system. For example, *sequential* rejuvenations, incapable of guaranteeing that the overall system is reset to a completely healthy state, allow a subtle time-window of attack not present in truly periodic *parallel* rejuvenations.

(3) Rejuvenation and (some configurations of) intrusion-tolerant replication have complementary roles by improving the dependability (e.g., *reliability* and *availability*) of systems for two opposite extremes of a mission timeline: the short term and the long term. The two techniques can complement each other to provide an improvement of reliability that is valid for any finite and possibly unknown (unbounded but not infinite) mission time. This benefit can be expressed quantitatively, for example using the defined measure of *resilience* that formalizes goals-of-improvement in a linear way.

(4) The impossibility of predicting the power and behavior of an adversary should not stop intrusion-tolerant systems from being objectively measured in terms of some of its dependability properties. For example, by specifying a relation between an effort of attack and the respective intrusion rate of nodes, it is possible to analyze how a system behaves within a range of adversarial intrusion effort and compare it objectively against other systems with different configurations. (In our examples, we considered that an "effort" exerts a proportional probabilistic rate of intrusion.)

The study presented in this paper is a step toward understanding how to use *intrusion tolerance* techniques to provide *tolerance to uncertainty of assumptions*, as a way to improve the design of dependable systems that better withstand a variety of adversarial environments with some hidden and unspecified parameters.

Acknowledgements The research by the first author was partially supported by FCT (Fundação para a Ciência e a Tecnologia—Portuguese Foundation for Science and Technology) through the Carnegie Mellon Portugal Program under Grant SFRH/BD/33770/2009, while a student in the Dual PhD program in ECE at CMU-ECE and

FCUL-DI-LaSIGE. This research was also partially supported by FCT through project PTDC/EIA-IA/100581/2008 (REGENESYS) and the Multiannual (LaSIGE) program. We thank the anonymous reviewers of LADC 2011 and JBCS for their helpful comments about earlier versions of this paper.

Appendix

A.1 Acronyms and symbols

Acronyms:

- BFT (*Byzantine fault-tolerant*)
- CDF (*cumulative distribution function*)
- ETTF (*expected time to failure*)
- ETTI (*expected time to intrusion*)
- IAE (*intrusion adversarial effort*)
- IR (*intrusion rate*)
- MT (*mission time*)
- PDF (*probability density function*)

Symbols:

- $\mathcal{A}$ (*availability*)
- f (threshold of tolerable intrusions)
- ϕ_j (state of node j, taking value 0 to denote *healthy* state and 1 to denote *intruded* state; j is an identifier of the node—an integer between 1 and the total number of nodes)
- k (number of rejuvenating nodes)
- λ (IAE, or IR)
- λ_j (IAE upon node j, or IR of node j)
- n (total number of real nodes)
- vk (number of virtual nodes, not accounted in n but helping real nodes to prepare their instantaneous rejuvenations)
- $\mu_{n,f}$ (ETTF of a system with configuration $\langle n, f\rangle$)
- $\mathbb{N}_1$ ($\{1, 2, \ldots\}$)
- p (PDF of global failure)
- P (CDF of global failure)
- $\mathcal{R}$ (*reliability*)
- ρ (*resilience*)
- t (wall-clock time, independent of $\mu_{1,0}$ or λ)
- τ (time normalized to $\mu_{1,0}$, i.e., to $1/\lambda$)
- $\|$ (parallel)
- $\therefore$ (sequential)
- $\gg$ (much greater than)

A.2 Formulas

In this Appendix, we make explicit the mathematical formulas that sustain the graphics and calculations of the paper. As a rule of notation, we reserve subscripts for internal configuration parameters (replication degree, intrusion threshold, rejuvenation parameters) and superscripts for external parameters (attack intensity and type).

Attack models Assumption 2 defined two models of attack: parallel ($\|$) and sequential ($\therefore$). Their respective formal conditions are expressed in (1) and (2), with $\{1, \ldots, n\}$ being the set of node identifiers, with $\phi_j(t)$ being the state (0 or 1, respectively, for *healthy* or *intruded*) of node j at instant t, and with $\lambda_j(t)$ being the IAE upon node j at instant t.

$$(\exists\lambda > 0)\big(\forall j \in \{1, \ldots, n\}\big)\big(\lambda_j(t) = \lambda \times \big(1 - \phi_j(t)\big)\big) \tag{1}$$

$$\begin{aligned}&(\exists\lambda > 0)\big[\big(\exists j \in \{1, \ldots, n\} : \phi_j(t) = 0\big)\big] \Leftrightarrow \\ &(\exists j \in \{1, \ldots, n\})\big[\big(\lambda_j(t) = \lambda\big) \wedge \big(\forall j' \neq j\big)\big(\lambda_{j'}(t) = 0\big)\big]\end{aligned} \tag{2}$$

Intrusion process From Assumptions 1 and 2, we consider the case of a constant IAE and of a proportionality between IAE and IR. Thus, IR is a constant λ and the intrusion of a node is modeled probabilistically with an associated PDF (p) of intrusion, a CDF (P) of intrusion and an associated ETTI ($\mu_{1,0}$) (per node under attack), as defined in (3), (4), and (5), respectively.

$$p_{1,0}^{(\lambda)}(t) = \lambda \times e^{-\lambda t} \tag{3}$$

$$P_{1,0}^{(\lambda)}(t) = 1 - e^{-\lambda t} \tag{4}$$

$$\mu_{1,0}^{(\lambda)} = 1/\lambda \tag{5}$$

The subscripts 1 and 0 stand for the parameters n and f of the reference system $\langle 1, 0\rangle$, composed of a single node.

Reliability ($\mathcal{R}$) Let $P_{n,f}(t)$ stand for the probability of a $\langle n, f\rangle$ system ever failing up to instant t. The overall *reliability* of $\langle n, f\rangle$ is, by definition, the probability of the system never failing up to instant t, and can thus be given by (6).

$$\mathcal{R}_{n,f}(t) = 1 - P_{n,f}(t) \tag{6}$$

Reliability always converges to zero as time grows (see (7))

$$(\forall\lambda > 0)(\forall n > f > 0)\Big(\lim_{t\to\infty} R_{n,f}^{\lambda}(t) = 0\Big) \tag{7}$$

When necessary, we shall use superscripts to inform the parameters of attack, namely the type of attack (parallel or sequential) using the respective symbols ($\|$ or $\therefore$), and the *intrusion adversarial effort* (IAE) λ. Whenever clear in the context, we may omit these superscripts.

When considering rejuvenations (parallel or sequential), we shall use the respective subscripts and symbols ($\|$ and $\varDelta$ or $\therefore$ and δ). Note that symbols $\|$ and $\therefore$ are used both for attack types and rejuvenation types.

Reliability ($\mathcal{R}$) under parallel *($\|$) attack* The CDF of failure is in (8). Calculating the sum, and subtracting it from 1, *reliability* becomes as in (9), with ${}_2F_1$ being the *Hypergeometric2F1* function.

$$P^{\|,\lambda}_{n,f}(t) = \sum_{i=f+1}^{n} \binom{n}{i} P_{1,0}(t)^i \times \big(1 - P_{1,0}(t)\big)^{(n-i)} \tag{8}$$

$$\begin{aligned}\mathcal{R}^{\|,\lambda}_{n,f}(t) = 1 - \big(e^{-\lambda t}\big)^{n-(f+1)}\big(1 - e^{-\lambda t}\big)^{f+1} \\ \times \binom{n}{f+1} \times {}_2F_1\big(1, f+1-n; f+2; 1-e^{\lambda t}\big)\end{aligned} \tag{9}$$

When under parallel attack, the qualitative effect produced on reliability, by varying either the replication degree (n) or the intrusion tolerance threshold (f) alone, is described in (10). In particular, reliability decreases by increasing n while fixing f, or by decreasing f while fixing n.

$$\begin{aligned}&(\forall t, \lambda > 0)(\forall n' > n > f' > f) \\ &\big(\mathcal{R}_{n',f'}(t), \mathcal{R}_{n,f}(t)\big) \vee \big(\mathcal{R}_{n,f}(t) < \mathcal{R}_{n,f'}(t)\big).\end{aligned} \tag{10}$$

Reliability ($\mathcal{R}$) under sequential *($\therefore$) attack* The probability density $p^{\therefore}_{n,f}(t)$ that the $(f+1)$-th node is intruded exactly at instant t, is defined recursively in (11), with $p^{\therefore,\lambda}_{0,1}(t) \equiv p^{(\lambda)}_{1,0}(t)$.

$$p^{\therefore,\lambda}_{n,f}(t) = \int_{t'=0}^{t} p^{\therefore,\lambda}_{n,f-1}\big(t'\big) p^{(\lambda)}_{1,0}\big(t - t'\big)\, dt' = \frac{(\lambda t)^f}{f!} \lambda e^{-\lambda t} \tag{11}$$

The global probability of failure $P^{\therefore,\lambda}_{n,f}(t)$ is defined in (12).

$$P^{\therefore,\lambda}_{n,f}(t) = \int_{t'=0}^{t} p^{\therefore,\lambda}_{n,f}\big(t'\big)\, dt' \tag{12}$$

Let $Q(a, z_0, z_1) = \Gamma(a, z_0, z_1)/\Gamma(a)$ stand for the *generalized incomplete regularized gamma function* [17], where $\Gamma(a, z_0, z_1) = \int_{t=z_0}^{z_1} t^{a-1} e^{-t}\, dt$ and $\Gamma(a) = \Gamma(a, 0, \infty)$. The *reliability* of a $\langle n, f\rangle$ system is defined in (13).

$$\mathcal{R}^{\therefore,\lambda}_{n,f}(t) = 1 - \mathcal{P}^{\therefore,\lambda}_{n,f}(t) = Q(f+1, \lambda t, \infty) \tag{13}$$

Here, reliability always grows with the intrusion tolerance threshold f (see (14)).

$$\begin{aligned}&(\forall t, \lambda > 0)(\forall n' > f' > f)(\forall n > f > 0) \\ &\big(\mathcal{R}^{\therefore,\lambda}_{n',f'}(t) > \mathcal{R}^{\therefore,\lambda}_{n,f}(t)\big)\end{aligned} \tag{14}$$

Mission time (MT) for the same reliability ($\mathcal{R}$) For *sequential* attacks, solving $\mathcal{R}^{\therefore,\lambda}_{1,0}(t) = \mathcal{R}^{\therefore,\lambda}_{n,f}(t')$ in order of t' gives (15), with $Q^{\langle 0,0,-1\rangle}$ being the inverse of $Q(a, z_0, z_1)$ in the 3rd argument.

$$t' = Q^{\langle 0,0,-1\rangle}\big(f+1, \infty, e^{-\lambda t}\big)/\lambda \tag{15}$$

Resilience (ρ) In Section 3.3, we define *resilience* as a measure that grows linearly with the number of bits to which reliability is close to 1. The formal definition is given in (16). Finding the *mission times* for which a $\langle n, f\rangle$ system is at least c times more *resilient* than the reference system $\langle 1, 0\rangle$ is accomplished by solving (17) in order of τ. The respective translation to *reliability* terms is given in (18).

$$\rho_{n,f}(t) = -\log_2\big(1 - \mathcal{R}_{n,f}(t)\big) \tag{16}$$

$$\rho_{n,f}(t) \geq c \times \rho_{1,0}(t) \tag{17}$$

$$\mathcal{R}_{n,f}(t) \geq 1 - \big(1 - \mathcal{R}_{1,0}(t)\big)^c \tag{18}$$

The effect of varying either the replication degree (n) or intrusion tolerance threshold (f) alone has the same qualitative effect (increase versus decrease) in resilience as in reliability. In other words, (7), (10) and (14) are valid upon replacing symbol $\mathcal{R}$ with ρ.

Availability ($\mathcal{A}$) Availability is the probability that the system is healthy at a random (uniformly selected) instant of time within the *mission time* (MT) interval. It can be obtained by computing the expression in (19). Note that $\mathcal{A}(t)$ does not mean the availability at instant t, but the availability of a system with a MT t.

$$\mathcal{A}_{n,f}(t) = \frac{1}{t} \int_{t'=0}^{t} \mathcal{R}_{n,f}\big(t'\big)\, dt' \tag{19}$$

Parallel ($\|$) rejuvenations Let Δ be the time period of $\|$-rejuvenation, such that the system is restored to a completely healthy state at every instant $i \times \Delta$ of time. The property of periodic global health resetting of the system, with all nodes simultaneously and instantaneously becoming (or just remaining) healthy, is described in (20).

$$\sum_{j=1}^{n} \phi_j(\Delta \times i) = 0, \quad \text{for } i \in \mathbb{N}_1 \tag{20}$$

Let $M = \lfloor t/\Delta \rfloor$ and $m = \text{mod}_\Delta t$ be auxiliary variables related to Δ. The *reliability* (see (21)) and *availability* (see (22)) can be obtained in function of the formulas without rejuvenations, by partitioning the MT into windows of size Δ.

$$\mathcal{R}_{n,f,\|,\Delta}(t) = \mathcal{R}_{n,f}(\Delta)^M \mathcal{R}_{n,f}(m) \overset{t \gg \Delta}{\approx} \mathcal{R}_{n,f}(\Delta)^{(t/\Delta)} \tag{21}$$

$$\begin{aligned}\mathcal{A}_{n,f,\|,\Delta}(t) &= (1 - m/t) \times \mathcal{A}_{n,f}(\Delta) \\ &\quad + (m/t) \times \mathcal{A}_{n,f}(m) \overset{t \gg \Delta}{\approx} \mathcal{A}_{n,f}(\Delta)\end{aligned} \tag{22}$$

Sequential rejuvenations For the sequential rejuvenation model, the limit availability, $\mathcal{A}_{n,f,\therefore,\delta}(\infty)$ (e.g., required to plot curves in Figs. 10 and 11), was computed approximately as an average across several simulations (e.g., 100), each performed for a large mission time (e.g., MT $= \delta \times 10^5$). In each simulation: the instants of intrusion of a node under attack were computed probabilistically, consistently with (3) and (4); then, $\mathcal{A}_{n,f,\cdot,\delta}(\infty)$ was approximated to the amount of time during which the system had at most f nodes intruded and divided by the total amount of time.

Note Software *Mathematica for Students* [28] was used to perform the simulations needed for Figs. 10 and 11, plot all the figures and tables and help deducing (9), (11), (13), and (15).

References

1. Avizienis A, Laprie JC, Randell B, Landwehr C (2004) Basic concepts and taxonomy of dependable and secure computing. IEEE Trans Dependable Secure Comput 1:11–33
2. Barlow RE (2002) Mathematical and statistical methods in reliability. In: Mathematical reliability theory: from the beginning to the present time, vol 7. World Scientific, Singapore, pp 159–175
3. Bessani A, Correia M, Quaresma B, André F, Sousa P (2011) DepSky: dependable and secure storage in a cloud-of-clouds. In: Proceedings of the 6th conference on computer systems (EuroSys'11). ACM, New York, pp 31–46
4. Bessani A, Daidone A, Gashi I, Obelheiro R, Sousa P, Stankovic V (2009) Enhancing fault/intrusion tolerance through design and configuration diversity. In: Proceedings of the 3rd workshop on recent advances on intrusion-tolerant systems (WRAITS'09)
5. Brandão LTAN, Bessani A (2011) On the reliability and availability of systems tolerant to stealth intrusion. In: Proceedings of the 5th Latin-American symposium on dependable computing (LADC 2011). IEEE Computer Society, Los Alamitos, pp 35–44
6. Castro M, Liskov B (2002) Practical Byzantine fault tolerance and proactive recovery. ACM Trans Comput Syst 20:398–461
7. Correia M, Neves NF, Veríssimo P (2004) How to tolerate half less one byzantine nodes in practical distributed systems. In: Proceedings of the 23rd IEEE international symposium on reliable distributed systems (SRDS'04). IEEE Computer Society, Washington, pp 174–183
8. Daidone A, Chiaradonna S, Bondavalli A, Veríssimo P (2008) Analysis of a redundant architecture for critical infrastructure protection. In: de Lemos R, Di Giandomenico F, Gacek C, Muccini H, Vieira M (eds) Architecting dependable systems V. Lecture notes in computer science, vol 5135. Springer, Berlin, pp 78–100
9. Forrest S, Somayaji A, Ackley DH (1997) Building diverse computer systems. In: Proceedings of the 6th workshop on hot topics in operating systems (HotOS-VI). IEEE Computer Society, Washington, p 67
10. Fraga JS, Powell D (1985) A fault- and intrusion-tolerant file system. In: Proceedings of the 3rd international conference on computer security, pp 203–218
11. Garcia M, Bessani A, Gashi I, Neves N, Obelheiro R (2011) OS diversity for intrusion tolerance: myth or reality. In: Proceedings of the international conference on dependable systems and networks (DSN'11), Hong Kong
12. Huang Y, Kintala CMR, Kolettis N, Fulton ND (1995) Software rejuvenation: analysis, module and applications. In: Proceedings of 25th international symposium on fault tolerant computing (FTCS'95). IEEE Computer Society, Washington, pp 381–390
13. Koren I, Krishna CM (2007) Fault tolerant systems. Morgan Kaufmann Publishers Inc, San Francisco
14. Koren I, Shalev E (1984) Reliability analysis of hybrid redundancy systems. IEE Proc E, Comput Digit Tech 131:31–36
15. Littlewood B, Strigini L (2004) Redundancy and diversity in security. In: Samarati P, Ryan P, Gollmann D, Molva R (eds) Computer security—ESORICS 2004. Lecture notes in computer science, vol 3193. Springer, Berlin, pp 423–438
16. Obelheiro RR, Bessani AN, Lung LC, Correia M (2006) How practical are intrusion-tolerant distributed systems? DI-FCUL TR 06–15, Dep. of Informatics, Univ of Lisbon
17. Olver FW, Lozier DW, Boisvert RF, Clark CW (2010) NIST handbook of mathematical functions. Cambridge University Press, New York
18. Roeder T, Schneider FB (2010) Proactive obfuscation. ACM Trans Comput Syst 28:4:1–4:54
19. Schneider FB (1990) Implementing fault-tolerant service using the state machine approach: a tutorial. ACM Comput Surv 22:299–319
20. Shamir A (1979) How to share a secret. Commun ACM 22:612–613
21. Sousa P, Bessani AN, Correia M, Neves NF, Verissimo P (2010) Highly available intrusion-tolerant services with proactive-reactive recovery. IEEE Trans Parallel Distrib Syst 21(4):452–465
22. Sousa P, Neves NF, Veríssimo P (2005) How resilient are distributed f fault/intrusion-tolerant systems. In: Proceedings of the 2005 international conference on dependable systems and networks (DSN'2005). IEEE Computer Society, Washington, pp 98–107
23. Sousa P, Neves NF, Verissimo P (2007) Hidden problems of asynchronous proactive recovery. In: Proceedings of the 3rd workshop on hot topics in system dependability (HotDep'07). USENIX Association, Berkeley
24. Sovarel AN, Evans D, Paul N (2005) Where's the FEEB? The effectiveness of instruction set randomization. In: Proceedings of the 14th conference on USENIX security symposium (SSYM'05). USENIX Association, Berkeley, p 10
25. Trivedi KS (2001) Probability and statistics with reliability, queuing and computer science applications, 2nd edn. Wiley, Chichester
26. Veronese GS, Correia M, Bessani AN, Lung LC (2009) Spin one's wheels? Byzantine fault tolerance with a spinning primary. In: Proceedings of the 28th IEEE international symposium on reliable distributed systems (SRDS'09). IEEE Computer Society, Washington, pp 135–144
27. Veríssimo P, Neves N, Correia M (2003) Intrusion-tolerant architectures: concepts and design. In: de Lemos R, Gacek C, Romanovsky A (eds) Architecting dependable systems. Lecture notes in computer science, vol 2677. Springer, Berlin, pp 3–36
28. Wolfram Research I (2011) Mathematica 8.0 for students