Reliability Engineering and System Safety 169 (2018) 127–136

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Heterogeneous 1-out-of-*N* warm standby systems with online checkpointing

Gregory Levitin[a,b], Liudong Xing[c], Yuanshun Dai[a,*]

[a] Collaborative Autonomic Computing Laboratory, School of Computer Science, University of Electronic Science and Technology of China, China
[b] The Israel Electric Corporation, P. O. Box 10, Haifa 31000, Israel
[c] University of Massachusetts, Dartmouth, MA 02747, USA

A R T I C L E I N F O

Keywords:
Online checkpoint
Warm standby
Mission cost
Mission reliability
Mission time
Optimization
Real-time
Sequencing

A B S T R A C T

As a common practice in computing-related applications, checkpointing is used to facilitate an effective system recovery in the case of the occurrence of failures. Checkpoints are performed to save data associated with completed portion of a mission task. In the case of a failure, through rollback and data retrieval the system can resume the mission task from the last successful checkpoint instead of from the very beginning of the mission, saving time and cost. This paper models and optimizes 1-out-of-*N*: G warm standby systems subject to uneven online checkpointing, where checkpoints can be performed in parallel with execution of the primary mission task for improving efficiency of computing elements. Both data checkpoint and retrieval take dynamic time, depending on the amount of work completed. System elements can be heterogeneous in the time-to-failure distribution, performance, and level of readiness to take over the mission task during the warm standby mode. A numerical method is first suggested to evaluate mission performance indices including mission success probability, expected mission completion time, and expected mission operation cost. Examples are provided to demonstrate influence of mission deadline and element resource sharing parameter (i.e., CPU time distribution between the checkpointing procedure and the primary mission task) on the mission performance metrics. The optimal checkpoint distribution and optimal element activation sequencing problems are considered for different combinations of optimization objectives and constraints. A co-optimization problem is further addressed, which aims to find the optimal combination of checkpoint distribution and element activation sequence. Example optimization solutions illustrate the tradeoff among the three mission requirements (reliability, completion time, operation cost) for warm standby systems with online checkpoints.

1. Introduction

In computing or information technology related applications, checkpointing has been commonly used to facilitate effective system recovery in the case of failures occurring due to factors such as hardware malfunctions, software errors, or user mistakes [1–4]. Specifically, even checkpoints [5–8] or uneven checkpoints [9–12] are performed to save data associated with completed part of the task to a reliable storage during the mission. When a failure takes place, the system is able to resume the mission task from the last successfully performed checkpoint through rollback and data retrieval. Without checkpointing, the system has to restart the entire mission task from scratch, which is costly in time and capital especially for long-running applications.

The checkpointing technique has been implemented in different ways. Particularly, an incremental checkpoint only saves data generated since the previous successful checkpoint while a full checkpoint saves all data generated from the beginning of the mission [13]. The incremental checkpoint can be quick to perform involving low overhead or time for saving data. Also it requires small capacity on storage. However, the system recovery tends to be slow because all incremental checkpoints performed before the failure are needed for system restoration. On the other hand, the full checkpoint involves higher overhead (longer time for data saving and retrieval) and needs more significant capacity on storage, but it usually can restore the system function quicker than the incremental checkpointing [14]. Hybrid checkpointing techniques are also implemented to tradeoff the checkpoint overhead and system restoration time, which combine occasional full checkpoints with more frequent incremental checkpoints during the mission [13,15,16]. The restoration of systems implementing the hybrid checkpoint technique requires retrieving data saved by the latest full checkpoint as well as data saved by all incremental checkpoints following the latest full checkpoint. This work focuses on uneven incremental or full checkpoints. Even checkpoints appear as a special case of uneven checkpoints.

* Corresponding author at: Collaborative Autonomic Computing Laboratory, School of Computer Science, University of Electronic Science and Technology of China, China.
E-mail addresses: levitin@iec.co.il, levitin_g@yahoo.com (G. Levitin), lxing@umassd.edu (L. Xing), 1125105129@qq.com (Y. Dai).

http://dx.doi.org/10.1016/j.ress.2017.08.011
Received 14 February 2017; Received in revised form 6 August 2017; Accepted 10 August 2017
Available online 15 August 2017
0951-8320/

Acronyms and Abbreviations

cdf	cumulative distribution function
CPU	central processing unit
GA	genetic algorithm
MTO	mission task operation
pdf	probability density function
pmf	probability mass function
WS	warm standby

Nomenclature

N	number of elements in the WS system
R	mission success probability or reliability
D	expected mission completion time
C	expected total mission operation cost
$T_{\max}$	maximum allowed mission time or deadline for completing the mission
$Y_{\max}$	maximum allowed number of time intervals in the mission
τ	minimal recognized time interval
H	number of checkpoint procedures to be completed during the entire mission
M	total number of operations to be performed during the mission (excluding checkpoints)
w_j, z_j	per unit time cost of element j being in the operation mode, and WS mode, respectively
V_j, λ_j	replacement cost, time of WS element j
G_j	performance (number of operations per unit time) of element j
d_j	life-time deceleration factor for element j
ω	element resource sharing parameter
π_j	fraction of the entire mission task that should be performed between the (j-1)-th and j-th checkpoints
$\boldsymbol{\pi}$	checkpoint distribution vector: $\boldsymbol{\pi}=(\pi_1,\dots,\pi_H)$
B_h	number of operations required for the h-th checkpoint procedure
U_h	number of operations required for retrieving data after the h-th checkpoint
$A_{j,h}(m)$	number of time intervals needed by element j (activated after completion of h checkpoints) to complete m additional checkpoints
$b(x)$	function returning number of operations required for saving data generated after performing fraction x of the entire mission task
$u(x)$	function returning number of operations required for retrieving data saved after performing fraction x of the entire mission task
$s(k)$	index of the element (given it is still working), which should be activated after elements with indices $s(1),\dots s(k\text{-}1)$ have failed
S_k	a sequence of elements $s(1)$, $s(2)$, …, $s(k)$
$Q_k(h,Y)$	probability that the number of the last checkpoint completed by the sequence of elements S_k is h and the number of the time interval when the last element from this sequence failed is Y
$\varphi_{j,h}(i)$	number of checkpoints completed by element j activated after completion of h checkpoints and operated during i intervals
$\Theta_j(Y)$	expected cost of using element j that fails in the WS mode not later than in time interval Y
$\overline{\Theta}_j(Y)$	expected cost of using element $s(k)$ given it does not leave WS mode and does not fail until switching off in time interval Y
$E_j(h,Y)$	expected cost of using element j given that it should be activated in interval Y after the h-th checkpoint
$E_{WS}(j)$	expected cost of using element j given that it remains in the WS mode during the whole mission
$F_j(t)$, $f_j(t)$	time-to-failure *cdf, pdf* for element j

The benefits of checkpoints in assisting system recovery do not come without a cost; they can have negative effects on the system performance due to the additional overhead incurred by performing the checkpointing procedures. To reduce or minimize the negative effects, diverse checkpoint placement policies such as fixed even [13] or uneven [16], dynamic [17,18], adaptive [19], age or provenance-dependent [20,21] checkpoints have been proposed. However, the existing works on modeling and optimizing checkpoint policies have mostly focused on single or distributed systems [22], but not on warm standby systems that abound in mission critical or safety critical applications such as space missions, flight controls and power systems [23–26].

A warm standby system has one or multiple operating and online element, and extra standby elements experiencing less or no (cold standby) operational stresses. When an on-line element malfunctions, it is replaced with an available standby element which is fully activated first and then takes over the task from the failed element [27,28]. The failure behavior of a warm standby element is dynamic; it is different before and after the activation, which makes modeling and analysis of warm standby systems difficult [29,30]. Recently an even checkpoint model with fixed time of data saving and negligible time of data retrieval has been suggested for cold standby systems (a special case of warm standby systems) [5]. An extension to considering uneven checkpoints with dynamic data saving and retrieval time was performed for warm standby systems [31]. These existing works on standby systems with checkpoints have assumed that the system is dedicated to the checkpointing procedure whenever it occurs. In practice the system performing the checkpoint procedures often has high performance and its valuable processing time can be wasted when interacting with the slower storage device for saving checkpoint data [32]. Thus to improve efficiency and further save the overall mission completion time, checkpointing can be performed in parallel with the main mission task. Such a checkpointing mechanism is referred to as online checkpoint hereinafter.

This work makes advancement to the state-of-the-art by modeling a 1-out-of-N: G warm standby system with non-identical elements performing uneven online checkpoints. The system mission is real-time, which is successful only if a specified amount of task can be accomplished by a pre-specified deadline. We model and optimize mission indices of success probability, expected cost and expected completion time of the considered warm standby system.

The rest of the paper is arranged as follows. Section 2 presents the warm standby system model studied in this work. Section 3 derives the proposed algorithm for analyzing mission success probability, expected mission completion time and expected mission cost for the considered warm standby system. Section 4 illustrates the proposed evaluation algorithm using examples. Effects of system parameters on the system performance are investigated. As applications of the suggested evaluation algorithm, Section 5 gives formulation and example solutions of relevant optimization problems. Section 6 concludes the paper and indicates a few directions of the future research.

2. Warm standby system model

The system has N non-identical elements, which are activated to operate on the system mission according to a pre-specified sequence $s(1)$, …, $s(N)$. Initially, $s(1)$ is online and operating and the remaining elements wait in the warm standby (WS) mode. When the online element malfunctions, it is replaced by the next available WS element in the