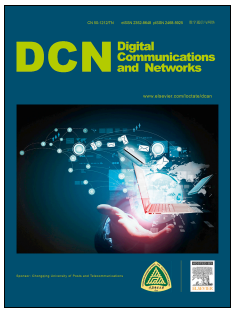
Accepted Manuscript



PRECESION: Progressive recovery and restoration planning of interdependent services in enterprise data centers

Ibrahim El-Shekeil, Amitangshu Pal, Krishna Kant

PII: S2352-8648(17)30084-6

DOI: 10.1016/j.dcan.2017.08.001

Reference: DCAN 95

To appear in: *Digital Communications and Networks*

Received Date: 26 February 2017

Revised Date: 2352-8648 2352-8648

Accepted Date: 4 August 2017

Please cite this article as: I. El-Shekeil, A. Pal, K. Kant, PRECESION: Progressive recovery and restoration planning of interdependent services in enterprise data centers, *Digital Communications and Networks* (2017), doi: 10.1016/j.dcan.2017.08.001.

ACCEPTED MANUSCRIPT
Digital Communications and Networks(DCN)

HOSTED BY
ELSEVIER

Available online at www.sciencedirect.com
ScienceDirect

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

PRECESION: Progressive Recovery and Restoration Planning of Interdependent Services in Enterprise Data Centers

Ibrahim El-Shekeil, Amitangshu Pal, and Krishna Kant

Computer and Information Sciences, Temple University, Philadelphia, PA 19122
E-mail:{ielshekeil,amitangshu.pal,kkant}@temple.edu

Abstract

The primary focus of this paper is to design a progressive restoration planning in an enterprise data center environment after a partial or full disruption. Repairing and restoring disrupted components in an enterprise data center takes a significant amount of time and human involvement. After a major disruption, the recovery process involves multiple stages, and during each stage, the partially recovered infrastructures can provide limited services to the users at some degraded service level. However, how fast and efficiently an enterprise infrastructure can be recovered depends on how the recovery mechanism will restore the disrupted components considering the inter-dependencies in between the services, along with the limitations of expert human operators. The entire problem turns out to be NP-hard and rather complex, and we devise efficient meta-heuristics to solve the problem. By taking some real-world examples, we show that the proposed meta-heuristics gives very accurate results and still runs ~600-2800 times faster compared to the optimal solution obtained from a general purpose mathematical solver [1].

KEYWORDS:
progressive restoration planning, enterprise data center, genetic algorithm, integer linear program, multi-layer networks

1. Introduction

Disruptions in enterprise data centers may occur as a result of a hardware failure, operating system or software failure, intrusion, virus outbreak or natural disaster. For example, the 2011 Japan earthquake and tsunami that damaged major data centers in Tokyo [2], and the 2012 Hurricane Sandy in the USA affected some data centers in New York due to flooding [3]. Other examples include storms and lightning that took down Google's St. Ghislain data center operations for five days in 2015 [4] and technical hiccups that affected the services of Bank of America [5] and Amazon [6] centers for 4-6 days. Such scenarios or disruptions may also arise as a result of relocation and upgrade of a data center in a different site, which needs proper pre-move planning and expertise. Depending on the scale, such disruptions can be partial that impact some applications or full that impact the entire data center. When such disruptions occur, they cause significant downtime which may lead to a substantial financial and legal impact. According to a recent study by Ponemon Institute, a data center outage can cost an average of $5,600 per minute [7].

In light of the above, there is a growing need to optimize the post-disruption recovery and restoration process for the enterprise data centers. A complete post-disruption restoration process for a large data center requires multiple stages as backup resources are brought to the field and installed, which sometimes requires a few weeks to several months [9]. Within this entire recovery stage, the partially restored infrastructures will still have to operate in a degraded manner and provide some partial level of service for clients.

A key design challenge of the restoration plan is to support partial business continuity that allows applications to *progressively* come back online after failures or disruptions. Fig. 1 shows such a data center recovery process where the critical applications or services are gradually recovered over a duration of 30 hours. Notice that the recovery process of all the services cannot be started simultaneously because the services in an enterprise data center are typically inter-dependent. The sequencing of data center services that are grad-