

性能情報を利用したストレージディスク障害予兆検知

鈴木 克典† 増田 峰義† 江丸 裕教†
 (株)日立製作所 研究開発グループ†

1 はじめに

ストレージシステムでは、ディスク故障に伴う性能や可用性の低下が問題となる。これに対し、個々のディスクから取得できる SMART 等の情報で障害予測を行う既存研究がある [1]。しかし、これらディスク単位の情報は粒度が細かく、大量ディスクで構成されるストレージシステムから必ずしも取得できる情報ではない。

一方、性能情報は運用で一般的に監視される情報であり、ストレージシステムから取得できる可能性も高い。また、ディスク障害に伴い性能の異常も見られる場合があることから、本稿では性能情報を用いたディスク障害検知方法を検討した。

2 提案手法

2.1 項で本稿で対象とするストレージシステムを、2.2 項でディスク障害に伴う想定性能変化を示す。提案手法は 2.2 項で示す性能変化を検知するため、通常時のデータを学習し、測定値とモデル予測値との乖離を検知する。本稿では、ディスク障害前後で顕著な性能情報間の関係性の変化が見られた Write 性能について報告する。最後に 2.3 項でモデリング方法を示す。

2.1 想定ストレージシステム

図 1 に想定ストレージシステムの動作を示す。ホストが書き込みを行うと、想定ストレージシステムはキャッシュにデータを書き込み、即座にホストに応答を返す。キャッシュ上のデータは、ホストへの応答と非同期にディスクに書き込むことによりレスポンス低下を防いでいる。また、以下の性能情報を取得することができることを想定する。

Write Response 書き込み時レスポンスタイム

Dirty Write Occupation Rate キャッシュ容量に対する、ディスク書き込み待ちデータの容量

Disk Utilization Rate 単位時間あたりのディスクが I/O を処理している時間の割合

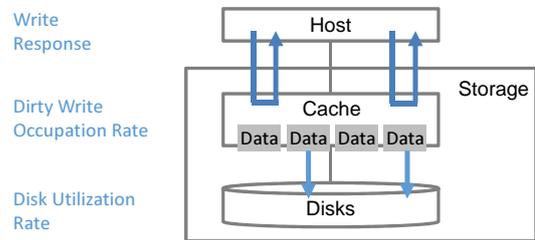


図 1: 想定ストレージシステム

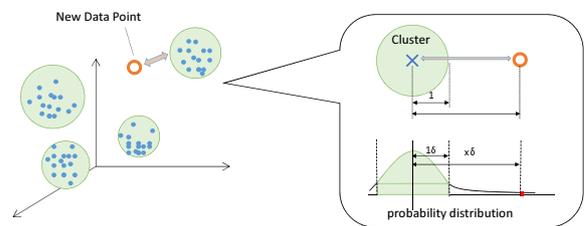


図 2: クラスタリングによる異常値検知

2.2 Write 性能変化のモデリング

ホストの書き込みが少ない場合、Write Response はキャッシュ書き込み時間で律速される。一方、ホストの書き込みが多い場合、キャッシュ上ディスク未書込のデータが増え始める。キャッシュに空き領域が少なくなると、ディスク未書込データをディスクに書き込み、空き領域ができるまで新規の書き込みが待たされる。

ここで、ディスク障害の予兆として、書き込みエラー等が発生し始めると、同じ Write 量であっても書き込みに要する時間が増え、Disk Utilization Rate が高まると考えられる。また、それによって Disk Write Occupation Rate が低下しづらくなることで、新規書き込みの待ち時間が増加し、Write Response が遅くなると考えられる。そこで、本研究では Write Response, Dirty Write Occupation Rate, Disk Utilization Rate の 3 性能情報について、通常時の状態を学習させた。

2.3 クラスタリングによる異常値検知

本稿ではクラスタリングをベースにした方法を用いて通常状態を学習させた(図 2)。モデル生成時には、2.2 項で選択した性能情報を軸とする空間に、各時刻のデータをプロットし、クラスタリングを行う。予測時には、

An Anomaly Detection Method for Disk Failures Based on Storage System Performance Data

†Katsunori SUZUKI †Mineyoshi MASUDA †Hironori EMARU
 †Center for Technology Innovation, Research & Development Group, Hitachi, Ltd.

データが正規分布に従うとみなし，測定データと最近傍クラスタ中心点間の距離と，最近傍クラスタ内各点とクラスタ中心点間の距離の標準偏差の比（以降，異常度と記載）を計算する [2]．本稿では異常度が2 以下の場合に，測定データが既存クラスタに属するとした．異常度が2 以上のデータとは，測定データとクラスタ中心点間の距離が，平均に比べて標準偏差の2 倍以上離れているデータである．統計的にはそのようなデータがクラスタに属する確率は5%以下であり，稀なデータであると考えられる．

3 評価と考察

3.1 評価方法

4 件の実ストレージのディスク故障時性能データ（データセット 1-4）を用いて，2.3 項で示した性能変化が検知されるか，評価した．用いたデータは，ディスク故障前後 2 週間の性能データ（1 時間平均）である．ディスクが故障すると RAID の復旧機能により，スペアディスクが用いられるようになる．そこで，故障ディスク交換後 10 日程のデータを通常時とみなして学習し，ディスク故障前 2 週間のデータについて検査した．

3.2 異常値検知結果

図 3 に提案手法の適用結果を示す．4 件の内，1 件（データセット 1）でディスク故障 5 日前に異常度 2.5 以上のデータを検出した．同ディスクの1ヶ月間の測定データは，すべて異常度 0.5 以内に収まることから，検出したデータは通常からの外れ具合が大きいことが分かる．

次に，図 4 にデータセット 1 の Write 性能データを示す．まず，ホストの秒間書き込み回数を示す Write IOPS は1ヶ月を通して変化せず，ワークロードの変化はない．しかしながら，Write Response はディスク障害後はほぼ 1ms 以下だが，ディスク障害前の 1 週間前から 2ms を超える場合が頻発していた．また，想定ではディスク障害前に Disk Utilization Rate も増加するが，Write Response 以外には変化は見られなかった．これは，障害発生前には瞬間的な状態変化が発生しており，これに対して特に敏感な Write Response のみが反応出来ているためと考えられる．そのため，ディスク障害前後で性能情報間の関係性が変化しているものの，変化量が小さく，提案手法では特に関係性の変化が大きい一時点のみを検出したと考えられる．

4 おわりに

本研究では，ストレージシステムから取得できる可能性の高い性能情報を用いた，ディスク障害の予兆検知手

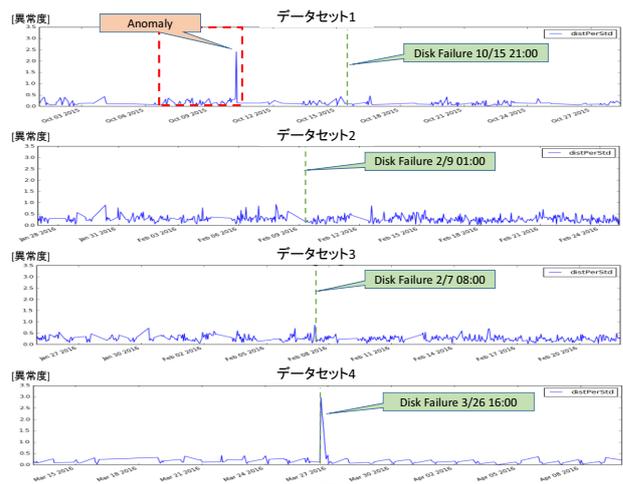


図 3: データセット 1-4 の外れ値検査結果

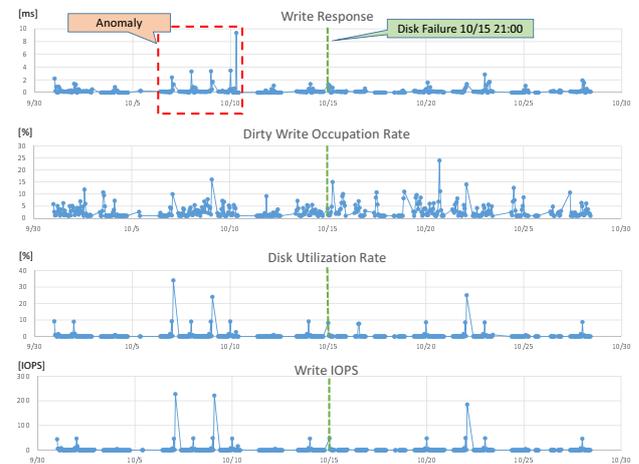


図 4: データセット 1 の Write 性能情報

法を検討した．提案手法では，ディスク障害に伴う性能情報間の関係性の变化を想定し，通常時データを学習する．そして予測値と測定値間の乖離を見る．本稿ではクラスタリングでモデリングし，実データを用いて提案手法の検証を行った．その結果，4 件中 1 件のデータにおいて，ディスク障害前に想定する Write Response の低下を検知できた．今後はより多くのデータによる検証と，より予測精度を上げるためのモデリング方法検討を行っていく．

参考文献

- [1] B. Zhu and et al. Proactive drive failure prediction for large scale storage systems. *MSST 2013*, pp. 1–5.
- [2] 鈴木忠志, et al. 高度保守サービスに貢献する予兆診断システム. *日立評論*, Vol. 95 No. 12, 2013.