

大規模な汎用実験環境における資源管理

宮地利幸[§]

知念賢一[§]

篠田陽一[†]

北陸先端科学技術大学院大学 情報科学研究科[§] / 情報科学センター[†]

1 はじめに

大規模なネットワーク実験を行う際には、ネットワーク実験に利用できるノードの管理が重要である。実験に利用可能なノードは、基本的に、実験に利用されていない状態と利用されている状態という2つの状態を遷移する。しかし、ノードは故障する可能性があり、多くのノードを取り扱うほど、存在する故障ノードの数は多くなる。

故障ノードを実験トポロジの一部に利用した場合、正しい実験結果が得られない可能性がある。さらに、故障ノードが存在すること自体の認識、および認識した後の故障ノードの特定に長い時間がかかり、実験の遂行自体に大きな影響を及ぼす可能性がある。したがって、故障検出手法と、発見された故障ノードを実験から排除する手法の確立が必要である。

2 資源管理モデル

資源管理システムは、基本的に利用者から要求された条件を満たすノードを検索し利用者へ貸し出す。またそれに加え、ノードの故障を検知し、故障ノードを利用不可の状態へ移動した後、故障を復旧するための処理を行う。あるノードが複数の利用者の実験に属することを防ぐ、ノードの排他制御を実現するために、資源管理システムは、環境で1つだけ動作する。

2.1 ノード状態の遷移

ノードの状態には、待機、貸出し、利用禁止、利用停止の4つを設ける。ノードの状態遷移を図1に示す。各ノードが最初に資源管理システムの管理下に置かれる際には、待機状態となる。資源管理システムは、利用者からノードの貸し出し要求を受けると、利用者へ貸出し可能で待機状態にあるノードから、要求された条件を満たすノードを必要数利用者へ貸し出す(貸出し状態)。利用者により実験利用に好ましくないと判断されたノードは、故障の可能性がある場合は利用停止状態に移され、実験の性格に合わない場合は待機状態へ戻される。利用停止状態のノードに対しては、管理者による検証が行われ、その結果、故障が発見された場合は利用禁止状態に移行し、修理が完了後、待機状態へ遷移する。また、健全であることが確認されたノードは再び待機状態へ遷移する。

2.2 ノードの故障検知とその復旧

ノードの故障検知は、貸し出された利用者が利用停止状態に遷移させる場合と、何らかの健全性確認システムと協調し、その結果を反映させる場合がある。健全性確認

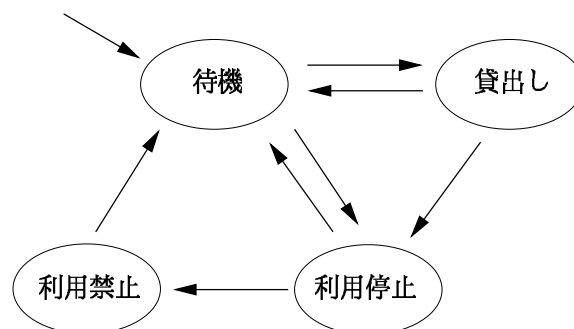


図1: ノードの状態遷移図

システムにも様々な種類があり、現在考えている確認手法には、ping を利用しての疎通確認や各ノードの vmstat や netstat の結果、特定ノードのトラフィック監視による状態確認、SNMP を利用した状態管理などがある。

このようなチェックを定期的に行うことで、ある程度の故障を検知できる。また、実験遂行システムと協調することで、利用者の実験中に不審な動作をするノードを発見した場合は、利用者へ通知するとともに、実験によってはノードの置き換えなどを行う。システムが故障を検知した際に行う処理は、利用者が指定できる必要がある。

基本的に利用者による実験中の検証は、実験による影響を受ける可能性があるため、利用者が不審な動作をするノードを発見した場合は、ノードを利用禁止状態に遷移させるのではなく、利用停止状態へ遷移させるにとどめ、実際に故障しているかどうかの検証は管理者が行う。また、夜間や週末に大規模な検証を行うことも可能である。この管理者による一斉検証は、検証対象のノードを利用停止状態へ遷移させた後に行い、その結果、検証対象のノードは待機状態または利用禁止状態へ遷移する。これらの検証の結果、故障が発見された場合は、ノード群の管理者に通知し修理を促す。

3 今後の方針

今後は、本稿で提案した設計をもとに、StarBED¹ における資源管理システムを実装し運用する。これにより定義した状態と健全性の確認方法の妥当性や、健全性確認システムや実験遂行システムとの協調方法について検討する。さらにこの結果を新たな設計実装に反映することで、より良いシステムを構築する。

¹ StarBED Project (<http://www.starbed.org/>)

北陸 IT 開発支援センター (<http://www.hokuriku-it.nict.go.jp/>)