広域分散ファイルシステムGfarmのSLA評価手法

渡邉 英伸[1]　亀澤 祐一[2]　高杉 英利[2]　平野 一樹[3]　今井 潔[3]　村田 健史[1]　建部 修見[4]

概要：クラウドサービスにおいて，広域分散型のストレージシステムは重要なインフラの一つであり，高い可用性が要求される．一般的にクラウドサービスの可用性を示す指標として SLA(Service Level Agreement) が用いられているが，広域分散型ストレージシステムの SLA を評価するための標準的な方法は存在しない．本論文では，広域分散ファイルシステム Gfarm を対象に，広域分散型ストレージシステムの SLA の評価手法を提案する．SLA の評価実験より，Gfarm の機能を活用することで 99.99%以上の可用性を保証できる可能性があることを確認した．

キーワード：広域分散型ストレージシステム, SLA, Gfarm

SLA evaluation method for a distributed file system with Gfarm

Abstract: A distributed storage system is one of important infrastructures in cloud services and requires high availability. Many cloud service providers often use SLA (Service Level Agreement) to indicate high availability, but there is no standard method to evaluate SLA of a distributed storage system. We propose the SLA evaluation method for a distributed storage system with Gfarm and confirmed that Gfarm has possibility to guarantee more than 99.99% availability by utilizing Gfarm functions well.

Keywords: distributed storage system, SLA, Gfarm

1. はじめに

近年，デジタルデータは飛躍的な成長を遂げており，科学研究分野においても，いわゆるビックデータの波は押し寄せている．科学データのサイズや種類は，大規模化・多様化するまでに至っており，今後ますます科学データは増え続け，解析できない程の量と種類の科学データが埋もれる懸念がある．このような背景の中，データ指向型研究手法 [1] が提唱されている．

本国においても，文部科学省よりデータ指向型研究手法に関する検討会 [2] が実施されており，大学や研究所など学術機関での利用を想定した学術系クラウド環境 (アカデミッククラウド) が提供されてきている [3][4]．情報通信研究機構 (NICT) においても，国内外の観測拠点で生成される観測データやスパコンで生成されるシミュレーションデータなど，あらゆる科学データを収集・蓄積すると同時に解析環境も提供する科学研究向けのクラウドシステム (NICT サイエンスクラウド [5]) を構築している．データ指向型科学研究手法として学術系クラウドを考える際，最も重要なインフラの一つがストレージである．多種多様な観測データや研究データを預かる学術系クラウドでは，ファイルサイズやファイル数がエクサオーダーになることが想定され，広域に分散可能なストレージシステムが重要視される．現在，広域分散型のストレージシステムを確立する広域分散ファイルシステムが数多く開発されており，オープンソースとして公開されるまでに至っている [6][7][8][9]．NICT サイエンスクラウドでも，オープンソースの広域分散ファイルシステムである Gfarm[10] を用いて国内 5 地区 (東京, 名古屋，京都，大阪，沖縄) にあるデータセンターに分散配置した計算機を 10Gbps の L2 高速バックボーンネットワーク網である JGN-X[*1] で接続し，約 3PB の広域

1 情報通信研究機構
NICT, 4-2-1, Nukui-Kitamachi, Koganei, Tokyo 184–8795, Japan
2 株式会社エヌ・ティ・ティネオメイト
NTT NEOMEIT CORPORATION
3 株式会社 SRA
Software Research Associates, Inc.
4 筑波大学計算科学研究センター
Center for Computational Sciences University of Tsukuba

[*1] http://www.jgn.nict.go.jp/

分散型ストレージシステムを構築している．2013 年 8 月現在で管理ファイル数は約 1.4 億という広域分散型のストレージシステムとして国内有数の規模を有する．

広域分散ファイルシステムは，ストレージのスケーラビリティー，管理データファイルの冗長性，バックアップレス，BCM(Business Continuity Management) モデルなど，多様な可能性を秘めている．特に，スケールアウトと地域分散を実現することによって高い可用性を保証する点は，広域分散ファイルシステムの最大の特徴となる．民間のクラウドサービスでは，広域分散ファイルシステムを用いたクラウドストレージサービスにおける可用性の水準を数値で示すことによって，SLA(Service Level Agreement) を保証するケースも数多く存在する．学術系クラウドにおいても，広域分散型ストレージシステムの高可用性を保証することは例外ではないと言える．

一方で，広域分散ファイルシステムを用いたストレージシステムの SLA を評価するための標準的な方法は存在しない．そのため，クラウドストレージサービスを提供する民間のクラウド業者では，過去の経験や社内サービス規定に基づいて定義し評価を行っているのが現状であり，一般的に SLA の評価方法は不透明かつ非公開である [11][12]．一定の指標になり得る SLA 評価手法が明確になる事は，オープンソースを主体にクラウド間の連携も見据えている学術系クラウドにおいて自組織のクラウドをアピールする，あるいは多組織のクラウドを判定する意味でも重要な事であると考える．

本論文では，広域分散ファイルシステム Gfarm を対象に，広域分散型ストレージシステムの SLA の評価手法を提案する．そして，提案手法をもとに Gfarm の最小構成における広域分散型ストレージシステムの可用性に対する SLA の評価結果を報告する．

以下の論文構成について説明する．2 章では，提案する SLA 評価手法について述べ，3 章で，SLA の評価結果を考察する．最後に，4 章で本稿のまとめについて述べる．

2. SLA 評価手法

2.1 SLA の現状

クラウドサービスの品質や成果を定量的に把握し評価することは，クラウドの利用者および提供者の双方にとって大きなメリットとなる．特に，利用者においてはシステム構成やサービス提供体制がブラックボックス化されていることが多いため，クラウドサービスの見える化は重要な課題である．既に民間のクラウドサービスでは，SLA の導入によりサービスの目標値を明確化しサービス品質の向上につなげている事例は数多く存在する．

SLA には，一般的に広義と狭義の使い方がある．図 1 に SLA の階層構造を示す．広義の使い方としては，サービスに関する契約行為とその証である契約書（併せて契約）

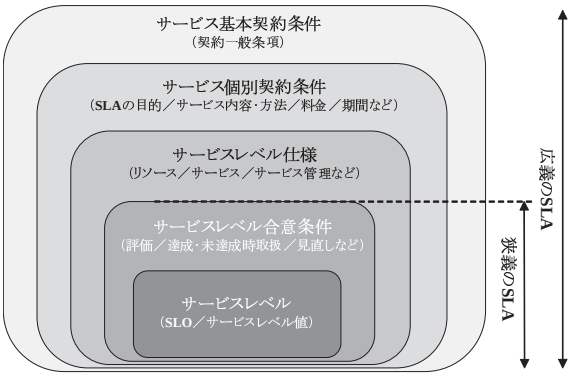

図 1 SLA 階層構造
Fig. 1 SLA hierarchal architecture.

である．サービスの範囲・内容およびサービスレベルに関する取り決めを規定し，契約当事者間で合意したものとなる．狭義の使い方としては，サービスレベルの規定とその取扱いを記述した文書である．SLA により合意したサービスの範囲や水準を数値化し，サービスレベルの項目とサービスレベルの値の組み合わせで表現するものとなる．民間のクラウドサービスにおいても，品質を 99.9%(スリー・ナイン) や 99.99%(フォー・ナイン) で示しているように，これらが狭義の SLA に該当する．

また，SLA はサービス利用者とサービス提供者の双方の協議の上で決定するため，プライベートクラウドのような個々と契約する形態の場合，利用者の個別目的に応じてサービスレベルを設定することから，その内容を外部に開示することはほとんどない．パブリッククラウドのような不特定多数の利用者と契約する形態の場合では，公開することが一般的であるため内容をある程度把握することが可能となるが，公平性の観点から原則サービスレベルを変更することは許されない．その他にも，利用するサービスの種類や利用目的によって，サービスレベルを保証値と位置づける場合と目標値と位置づける場合がある．個々と契約する形態の場合，サポートや管理などサービス管理品質を注力する傾向にあり，サービスレベルは目標値とすることが多い．不特定多数の利用者と契約する形態の場合では，サービスの機能や性能に重きを置く傾向にあり，サービスレベルは保証値とするケースが多い．このように，サービス提供形態によって SLA の特徴が異なることを認識しておく必要がある．

国内外ともに SLA の標準化に向けた取り組みも進められている [13][14][15][16]．しかしながら，これらの多くはガイドラインである．SLA の概念ならびに一般的な対策が記述されているのみであり，具体的な SLA の評価手法が明記されているわけではない．加えて，クラウドのような新たな IT サービスを対象とした SLA に関しては，標準化の取り組みの方が追い付いておらず，クラウドで提供される広域分散型のストレージサービスに対する SLA 評価方法

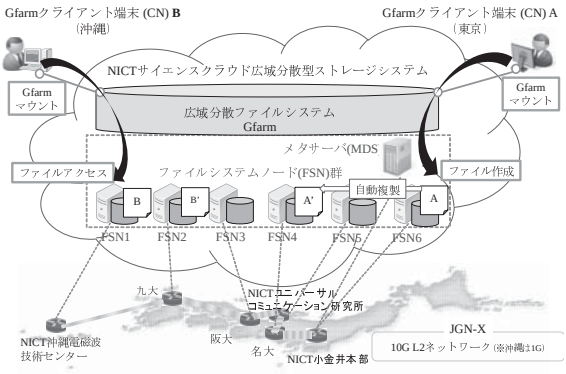

図 2 NICT サイエンスクラウド広域分散型ストレージシステム構成
Fig. 2 A distributed storage system architecture with Gfarm in NICT Science Cloud.

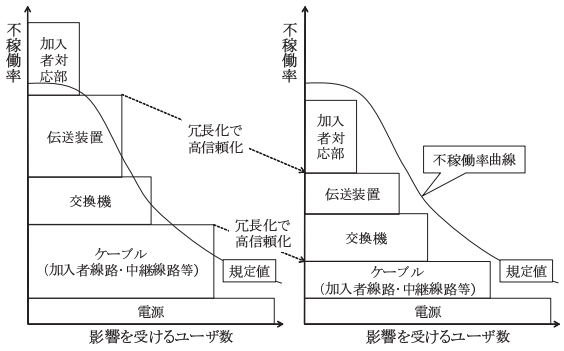

図 3 通信網における不稼動率曲線活用方法
Fig. 3 The method with a non-utilization rate of a curve for a network communication.

は存在しないのが現状である．本論文では，Gfarm を用いた広域分散型ストレージシステムに対する SLA の評価モデルならびに定式化方法を提案する．これにより，広域分散型ストレージシステムにおける SLA の見える化の促進を目指す．なお，本論文で扱う SLA は狭義の SLA とし，サービスレベルを保証値として扱うことを前提とする．加えて，可用性を示す際に用いられる稼働率を SLA の対象として議論する．

2.2 SLA 評価モデル

本節では，最初に Gfarm について概説した後，SLA 評価モデルの構築方法について述べる．Gfarm は，複数組織からなるグリッド環境における高信頼のファイル共有および高性能分散並列処理を実現するために開発された広域分散ファイルシステムである．ファイルの保存場所などメタ情報を集中管理するメタサーバ (MDS) とファイルの実体を保持するファイルシステムノード (FSN) で構成される．図 2 に NICT サイエンスクラウド上で Gfarm を用いた広域分散型ストレージシステムの構成を示す．Gfarm は，分散している FSN の物理保存領域をあたかも 1 つの論理保存領域に見せることができる．ユーザのクライアント端末 (CN) は，Gfarm が用意する専用のクライアントライブラリを導入することで，Gfarm が提供する論理領域をマウントして利用することが可能となる．MDS は冗長化構成をとることが可能で，基本的にファイルの情報を全てメモリ上で管理する．これにより，CN からの問い合わせなどに対して高い応答性を確保している．また，CN がファイルを保存するあるいはファイルにアクセスする際には，必ず MDS を介して CN から最小の RTT(Round Trip Time) に該当する FSN にアクセスさせる設計となっている．加えて，動的に複製を作成するレプリケーション機能も備えおり，高い可用性を確保している．

Gfarm を用いた広域分散型ストレージシステムの SLA 評価モデルを構築するにあたり，通信網の規定で用いられる不稼働率曲線を用いる手法を参考にした．通信サービスでは，稼働率を永続的に維持してくための評価モデルとして，いわゆる社会的迷惑量への期待値を等しくする不稼働率曲線が用いられている．本節では，このモデルの考え方を広域分散型ストレージシステムに当てはめることにより，SLA の評価モデルを構築する．通信網における不稼動率曲線は，通信サービスを提供するための各システムの不稼働率をそのシステムが故障した時に影響を受けるユーザ数に応じて積み上げたものとなる．この不稼働率曲線をどのようにすべきかという一般的な解は無く，サービスを提供する側で過去の経験や社内サービス規定に基づいて作成されているのが現状である．図 3 に通信網における不稼動率曲線活用方法を示し，図 4 に Gfarm を用いた広域分散型ストレージシステムの不稼動率曲線活用モデルを示す．例えば，既にサービス提供しているシステムを更改する場合，まずそれぞれの機能要件に併せて不稼働率を積み上げることを行う．この結果，図 3(左図) のように不稼動率曲線 (規定値) 内に収まらない場合は，冗長化構成を採用するなど各装置の信頼性要求条件を定めていくことで図 3(右図) のようにサービスレベルを維持するようにしていく．

これに併せて，Gfarm を用いた広域分散型ストレージシステムの不稼動率曲線活用モデルを構築すると，図 4(左図) のようになる．広域分散型ストレージシステムにおける可用性は，情報にいつでもアクセスできることであることから，提案する広域分散型ストレージシステムの不稼動率曲線活用モデルでは，x 軸を影響を受けるファイル数としている．積み上げる項目として，まず最も影響を与えやすい MDS に関する項目を下位層に配置している．そして，全ての FSN が停止する場合，エリア単位で FSN が停止する場合，ノード単位で FSN や CN が停止する場合の順で積み上げることとした．Gfarm 機能を十分に活用しない場合，ハードウェア故障等の影響を考慮しなければならずシステムの運用上，故障監視やオンサイト対応，ビルの停電対応など非常に多くの要因を包含しなければならない．一方，Gfarm 機能を十分に活用すれば，停電対応やハードウェア

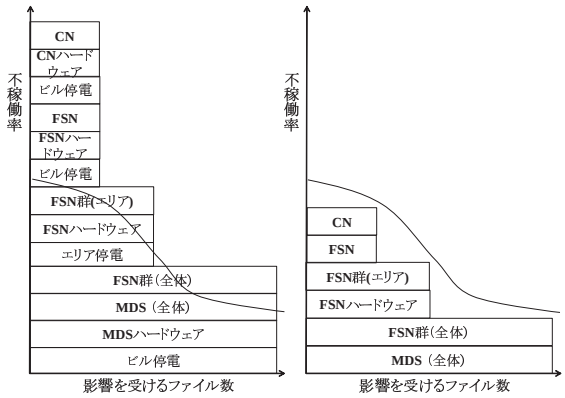

図 4 Gfarm を用いた広域分散型ストレージシステムの不稼動率曲線活用モデル

Fig. 4 The model with a non-utilization rate of a curve for a distributed storage system with Gfarm.

故障などの影響を無視することが可能となり，図 4(右図) とすることができる．ここで，本論文における Gfarm を用いた広域分散型ストレージシステムの不稼動率曲線活用モデルを構築するための前提条件を以下に示しておく．

- ネットワークに関しては，SLA99.99 以上であることとし，冗長化構成を図っているものとする．
- MDS は，冗長化構成かつ地理的分散を図っていることとする．
- FSN も MDS と同様に，地理的分散を図っていることとする．
- レプリケーション機能によって必ずファイルが 2 つ以上複製され，地理的に分散して保存する設定であることとする．

2.3 SLA の定式化

一般的に稼働率における SLA の定式化には，システム構成からサービスの稼動率を算出する方法と実際の運用データからサービスの稼働率を算出する方法の 2 つがある．

2.3.1 システム構成から算出する稼働率

システム構成から算出する稼働率は，いわゆるシステムハードウェアの稼働率であり，平均故障間隔 MTBF(Mean Time Between Failure) と平均復旧時間 MTTR(Mean Time To Repair) で算出することができる．サーバの冗長化構成をとらない場合，稼働率 As は，

$$As = MTBF/(MTBF + MTTR), \tag{1}$$

で算出することができる．一方，サーバの冗長化構成をとる場合の稼働率 Ap は，

$$Ap = 1 - (1 - As) * (1 - As), \tag{2}$$

で算出することが可能である．

Gfarm を使った広域分散型ストレージシステムを単純化した 3 パターンの SLA 評価モデルを図 5 に示す．形態 1 のように冗長化をしない構成では，MDS または FSN は直列接続となり稼働率 $A(system)$ は，

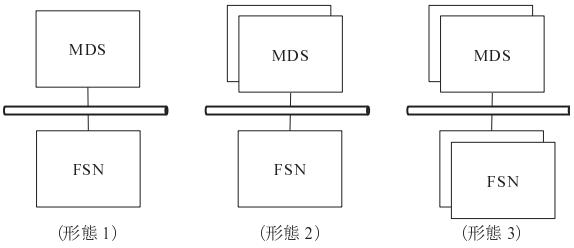

図 5 Gfarm を用いた広域分散型ストレージシステムを単純化した SLA 評価モデル

Fig. 5 The simplified SLA evaluation model of a distributed storage system with Gfarm.

$$A(system) = As(MDS) * As(FSN), \tag{3}$$

のように定式化される．形態 2 や形態 3 のように冗長化の構成を取る場合には，MDS と FSN は並列接続となり，形態 2 の場合の稼働率 $A(system)$ は，

$$A(system) = Ap(MDS) * As(FSN), \tag{4}$$

となり，形態 3 の場合の稼働率 $A(system)$ は，

$$A(system) = Ap(MDS) * Ap(FSN), \tag{5}$$

のように定式化ができる．

2.3.2 運用データから算出する稼働率

今回，SLA の対象が広域分散型ストレージシステムであることから，運用データから稼働率を算出する方法は，ファイルアクセスの成功率で代替することが可能と考える．すなわち，これはソフトウェアを意識した稼働率ということになる．本論文では，ユーザのファイル操作に対するアクセシビリティの平均値として，ある期間における Gfarm でアクセスされた全てのファイル数からアクセスエラーとなったファイル数を調査することで，稼働率 (ファイルアクセス成功率) を算出することとした．

ファイル数 F とした場合の稼働率 $F(success)$ は，

$$F(success) = (F(all) - F(error)/F(all)), \tag{6}$$

のように定式化できる．

2.3.3 広域分散型ストレージシステムの稼働率

システム構成から算出する稼働率と運用データから算出する稼働率は，相互に稼働率の低下を招くような関連性が無いことから，我々は最終的に双方の稼働率を掛け合わせることで，広域分散型ストレージシステムの稼働率を算出できると考えている．したがって，本論文で提案する Gfarm を用いた広域分散型ストレージシステムの稼働率 $A(Gfarm)$ は，

$$A(Gfarm) = A(system) * F(success), \tag{7}$$

のように定式化ができる．

3. SLA 評価実験

本章では，NICT サイエンスクラウドで運用中の Gfarm 環境を用いてシステム構成から算出する稼働率と運用データから算出する稼働率を測定し，最終的な広域分散型ストレージシステムの稼働率を算出する．その結果をもとに Gfarm が可用性に対してどれだけのポテンシャルを有しているかを評価する．Gfarm を使った広域分散型ストレージシステムの SLA(稼働率) を低下させる要因を以下に示す．

人的要因　オペレーションにおいてシステム停止を招くようなケース

性能要因　トラフィックなどの状況に応じた設備増強などが適切に行われずファイルの読み書きがエラーとなるようなケース

ソフトウェア要因　システム環境の変更やファイル更新などによりバグが顕在化するようなケース

システム構成要因　サーバ，ルーター，L2 スイッチ，ネットワークケーブルなどの故障に起因するようなケース

本来であれば，全ての要因を考慮した SLA の評価が必要であるが，今回の SLA 評価実験においては，Gfarm の可用性におけるポテンシャルを純粋に評価したいため，人的故障要因ならびに性能要因を考慮しないものとする．なお，NICT サイエンスクラウドが提供しているストレージシステムとしては，これらの要因に対してオペレーションを含めた対応を適切に実施していることを補足しておく．また，NICT サイエンスクラウドにおける広域分散型ストレージシステムの SLA について評価するにあたり，以下の環境を前提として評価を進める．

- ネットワークは冗長化されているものとし，SLA99.99 以上であるとみなす．
- NICT の所内 (小金井ならびにけいはんな) は停電対策済みとする．
- RAID 構成により，単一サーバ内のデータの冗長性は確保済みとする．
- Gfarm で使用する認証等の周辺システムは考慮しない．
- データはレプリケーションされ，地域分散保存されているものとする．
- MDS および FSN は冗長化構成が図られ，地域分散しているものとする．

なお，JGN-X は公式にネットワーク SLA を評価していないものの，2011 年・2012 年の障害情報を参考にする限り，国内におけるネットワーク障害は発生していないことから，SLA99.99%以上のネットワークであるとみなしても現状問題は無いと考えられる．

図 6 に評価実験環境を示し，表 1 にマシンの仕様を示す．評価実験環境として，4 拠点（小金井，けいはんな，阪大，名大）に分散配置され，JGN-X の L2 ネットワークで接続された計 17 台の FSN を用いた．MDS は，小金井の

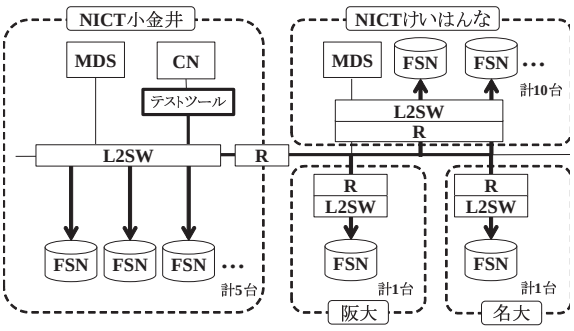

図 6　評価実験環境
Fig. 6　The experiment environment on NICT Science Cloud.

MDS を主系，けいはんなの MDS を従系とした地域分散型の冗長化構成となっている．ネットワーク遅延 (RTT) は，小金井とけいはんな間/阪大間は 10 ミリ秒，小金井と名大間は 15 ミリ秒である．なお，実際のネットワーク構成はもっと複雑であるが，今回の評価実験では最低限の構成とした．ソフトウェア構成は，Gfarm 2.5.8rc1，FUSE 2.74，PostgreSQL 8.3.12 となっている．

3.1 評価結果

3.1.1 システム構成要因としての SLA

システム構成に起因する SLA の低下としては，設備故障が代表的なものである．ここでは，前述の前提条件と図 6 に示した評価実験環境に基づきシステムを構成する機器の平均故障間隔 MTBF と平均復旧時間 MTTR により SLA を評価する．MTBF 値および MTTR 値については，本来実際の運用を想定した評価により算出するべきではあるが，今回は時間の関係上製造・販売サイドからヒアリングなどで得た情報を利用することとした．本論文で利用する MTBF 値を以下に示す．なお，MTTR 値は 24H としている．

ルータ：　MTBF(R) = 200000H
L2 スイッチ：　MTBF(SW) = 120000H
MDS：　MTBF(MDS) = 50000H
FSN：　MTBF(FSN) = 50000H

初めに，MDS，FSN，L2 スイッチ，ルータの単体稼働率を算出する．MDS 単体の稼働率 $As(MDS)$ は，

$$As(MDS) = MTBF/(MTBF + MTTR) = 50000/(50000 + 24) = 0.9995202, \quad (8)$$

となる．同様に FSN 単体の稼働率 $As(FSN)$ も

$$As(FSN) = MTBF/(MTBF + MTTR) = 50000/(50000 + 24) = 0.9995202, \quad (9)$$

となる．また，L2 スイッチ単体の稼働率 $As(SW)$ およびルータ単体の稼働率 $As(R)$ は，

$$As(SW) = MTBF/(MTBF + MTTR)$$

表 1　マシンの仕様.
Table 1　Specification of machines.

	MDS		FSN / CN			
場所	小金井	けいはんな	小金井	けいはんな	阪大	名大
CPU(Xeon)	E5506 2.13GHz	X5675 3.07GHz	E5645 2.40GHz	X5675 3.07GHz	E5645 2.40GH	
Memory	192GB		96GB			
OS	CentOS 5.7		CentOS 6.4	CentOS 5.7	openSUSE 12.1	
HDD	SAS 6 本 700GB (RAID5)	SAS 3 本 1.2TB (RAID5)	SATA3 24 本 60TB (RAID6)	SATA3 60 本 112TB (RAID6)	SATA3 24 本 55TB (RAID6)	
NIC	1GbE	10GbE	1GbE	10GbE	1GbE	
備考	主系	従系	CN も同仕様	-	-	-

$$= 120000/(120000 + 24) = 0.9998000, \quad (10)$$

$$As(R) = MTBF/(MTBF + MTTR)$$
$$= 200000/(200000 + 24) = 0.99988001, \quad (11)$$

になる．

次に，MDS の冗長化構成および FSN の冗長化構成における各々の稼働率を算出する．図 6 に示した評価実験環境の MDS は，L2 スイッチとルータに直列に接続されており，その環境が小金井とけいはんなに冗長化されている．よって，冗長化構成における MDS の稼働率 $Ap(MDS)$ は，

$$Ap(MDS) = 1 - (1 - As(MDS) * As(SW) * As(R)) * (1 - As(MDS) * As(SW) * As(R)) = 0.9999866, \quad (12)$$

となる．FSN が最小構成の場合，L2 スイッチとルータに直列に接続された環境が 2 箇所 (例えば，小金井とけいはんな) に冗長化されていることとなり，冗長化構成における FSN の稼働率 $Ap(FSN)$ は，

$$Ap(FSN) = 1 - (1 - (As(FSN) * As(SW) * As(R)) * (1 - As(FSN) * As(SW) * As(R)) = 0.9999866, \quad (13)$$

となる．

これらの結果より，システム構成要因としての Gfarm を用いた広域分散型ストレージシステムの稼働率 $A(system)$ は，

$$A(system) = Ap(MDS) * Ap(FSN) = 0.9999732, \quad (14)$$

となる．

3.1.2　ソフトウェア要因としての SLA

ソフトウェアに起因する SLA の低下としては，運用前の検証不足やシステム環境の変更に伴い顕在化する不具合が代表的なものである．ここでは，NICT サイエンスクラウドにおける過去のファイルアクセス量を参考にファイルアクセスの負荷を掛けることで SLA を評価する．なお，NICT サイエンスクラウドでは，Gfarm の導入試験 (正常系・準異常系・異常系) で Gfarm が正常に動作することは確認済みである．また，この評価については性能評価ではないことから，過負荷によってアクセスエラーが発生しないような状況で実施することとした．表 2 にファイルアクセスの負荷印加条件を示す．

負荷印加のテストツールとして，Gfarm と同時にインストールされる gfruntest r7733 を用いた．テストツールは，SSH のセッションをテスト実行する端末 (今回は小金井の CN 自身) に張り，その端末を介して該当する全ての FSN に並列で負荷を印加する．具体的には，表 2 のような各ファイルサイズのファイルを指定した間隔毎に一連の処理 (ファイル作成，数回のファイル読み書き，ファイル削除) を一定期間繰り返すだけである．

今回，評価実験を実施した期間は 22 日間 (2013/2/8〜2013/3/1) である．図 7 に負荷変動の結果を示す．なお，負荷印加期間の前半 (2013/2/8〜2013/2/18) では動作確認を兼ねて 16Mbps 程度の負荷（表 2 における 5MB の列）を印加し，その後負荷を増大させている．

22 日間のアクセスログを整理した結果，総読み書き回数は，2,424,615 回であった．その内 84 件のエラーが発生した．内訳を以下に示す．

80 件：　テストツールによるローカルホストへの SSH ログイン失敗

4 件：　ファイルアクセスエラー

これらの結果より，Gfarm を用いた広域分散型ストレージにおけるソフトウェア要因としての稼働率 $F(success)$ は，

$$F(success) = (F(all) - F(error)/F(all)) = (2424615 - 4)/2424615 = 0.9999983, \quad (15)$$

となる．なお，SSH ログイン失敗のエラーの 80 件については，Gfarm の機能とは関係のない要素のため，稼働率の算出には反映させていない．

表 2　ファイルアクセス負荷印加条件
Table 2　The load test conditions to file access.

ファイルサイズ	1KB	50KB	500KB	5MB	50MB	500MB	5GB
ファイル作成奸悪 [秒]	3.0	1.0	2.0	2.5	30.0	40.0	600.0
1 分あたりのファイル作成数	20.0	60.0	30.0	24.0	2.0	1.5	0.1
負荷 [Mbps]	0.003	0.31	1.93	16.00	13.33	100.0	68.67

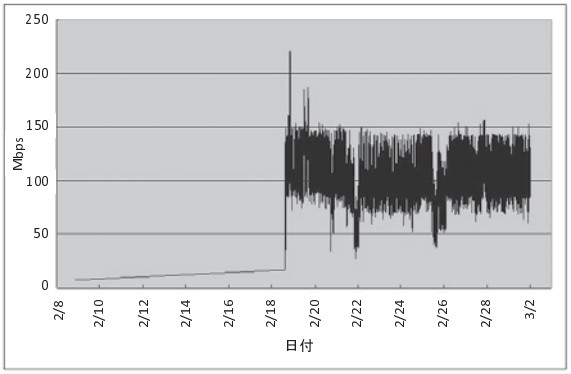

図 7　負荷変動結果
Fig. 7　The result of load variation.

3.1.3　Gfarm の SLA

システム構成要因としての稼働率とソフトウェア要因としての稼働率の算出結果はこれまでに示した通りである．これらの結果より，NICT サイエンスクラウドにおける Gfarm を用いた広域分散型ストレージシステムの稼働率 $A(Gfarm)$ は，

$$\begin{aligned} A(Gfarm) &= A(system) * F(success) \\ &= 0.9999732 * 0.9999983 \\ &= 0.9999715, \end{aligned} \tag{16}$$

となり，最小構成の Gfarm 環境を想定したとしても Gfarm の機能を活用することで 99.99%(フォー・ナイン) 以上の可用性を達成できる可能性を秘めていることが分かった．

3.2　考察

負荷印加評価実験のエラーに対して考察する．今回の評価では，80 件の SSH ログイン失敗エラーと 4 件のファイルアクセスエラーが発生した．前者は，テストツール側のエラーではあるが，LDAP のエラーログが同時に出力されていた．後者においても，頻繁に LDAP サーバにアクセスしていたことがログより確認できた．NICT サイエンスクラウドでは，LDAP でユーザ情報を管理している関係で，CN にも LDAP クライアントが設定されている．このことも含め，原因は LDAP に起因したものと考えられる．Gfarm は，MDS から得たファイル情報に対して紐付くユーザ情報の整合性を確認する仕組みとなっている．そのため，LDAP クライアントが有効だった場合，大量のファイルに対してアクセスがあると，MDS からレスポンスがある度に LDAP サーバへの問い合わせが集中することになる．その結果，CN から遠隔地の FSN へファイルアクセスを行おうとしている際に，あるタイミングでタイムアウトを引き起こしたことでこれらのエラーが発生したと推察される．この問題は，LDPA 情報を一定期間キャッシュすることが可能な nscd(name server cache daemon) を利用することで解決する予定である．nsdc を CN で起動させておけば，LDAP プロトコル送信を抑制することなる．よって，ファイルアクセスのエラー確率を減らすことが可能と考えられる．

4.　おわりに

クラウドサービスにおいて SLA は，信頼性の指標の一つとして重要な位置づけとなっている一方で，クラウドに代表される広域分散型ストレージシステムの標準的な SLA 評価手法が存在しない現状がある．このような背景の中，本論文では Gfarm を用いた広域分散型ストレージシステムの可用性に対する SLA 評価手法を提案した．具体的には，広域分散型ストレージシステムの SLA を評価するためのモデル構築手法として，通信網における不稼動率曲線活用方法を応用した手法を示した．また，その評価モデルをもとに SLA を定式化する手法として，システム構成から算出する稼働率と運用データから算出する稼働率を掛け合わせる手法を提案した．また，提案手法を用いて NICT サイエンスクラウドの環境上で Gfarm を使用した広域分散型ストレージシステムの可用性を検証した結果，Gfarm の最小構成を想定したとしても，Gfarm の機能を十分に活用することで 99.99%以上の可用性を達成できる見込みを得ることができた．

今後の課題としては，提案手法の汎用性ならびに実用性の検証を行う．また，標準化への取り組みも並行して実施していく予定である．

謝辞

本研究は，NICT サイエンスクラウド上の計算機リソースを用いて実施しています．ここに記して謝意を表します．

参考文献

[1] Tony Hey, Stewart Tansley, and Kristin Tolle: *The Fourth Paradigm: Data-Intensive Scientific Discovery*, 入手先 ⟨http://research.microsoft.com/en-us/collaboration/fourthparadigm/⟩, (2009).
[2] 文部科学省: アカデミッククラウドに関する検討会, 入手先 ⟨http://www.mext.go.jp/b_menu/shingi/chousa/

shinkou/027/index.htm⟩, (2012).
[3] 北海道大学情報基盤センター: 北海道大学アカデミッククラウドのご紹介, 入手先 ⟨http://www.hucc.hokudai.ac.jp/ã10019/kosyu/pdf2/cloud_1⟩ (2011).
[4] 国立情報学研究所: アカデミック・クラウド, NII Today, No.56, (2012).
[5] Ken T. Murata, S. Watari, T. Nagatsuma, M. Kunitake, H. Watanabe, K. Yamamoto, Y. Kubota, H.Kato, T. Tsugawa, K. Ukawa, K. Muranaga, E. Kimura, O. Tatebe, K. Fukazawa and Y. Murayama: *A Science Cloud for Data Intensive Sciences*, Data Science Journal, Vol. 12, pp.139-146, (2013).
[6] Tom White, *Hadoop*, ISBN-10: 487311439X, ISBN-13: 978-4873114392, (2010).
[7] GrusterFS, 入手先 ⟨http://www.gluster.org/⟩
[8] F. Hupfeld, T. Cortes, B. Kolbeck, E. Focht, M. Hess, J. Malo, J. Marti, J. Stender, E. Cesario: *XtreemFS - a case for object-based storage in Grid data management*, VLDB Workshop on Data Management in Grids. In: Proceedings of 33rd International Conference on Very Large Data Bases (VLDB) Workshops, (2007).
[9] Sage A. Weil. Scott A. Brandt. Ethan L. Miller. Darrell D. E. Long. Carlos Maltzahn: *Ceph: A Scalable, High-Performance Distributed File System*, Proceedings of the 7th Conference on Operating System Design and Implementation (OSDI '06), (2006).
[10] O. Tatebe, K. Hiraga, N. Soda: *Gfarm Grid File System*, New Generation Computing,Ohmsha,Ltd.and Springer,Vol.28,No.3,pp.257-275, (2010).
[11] Amazon S3, 入手先 ⟨http://aws.amazon.com/jp/s3/⟩.
[12] NTT Communications Cloudn, 入手先 ⟨http://www.ntt.com/cloudn/data/sla.html⟩.
[13] OGC: *The Introduction to the ITIL Service Lifecycle (Official Introduction)*, ISBN-10: 0113311311, ISBN-13: 978-0113311316, (2010).
[14] 経済産業省: 情報システムに係る政府調達への *SLA* 導入ガイドライン, 入手先 ⟨http://www.meti.go.jp/policy/it_policy/tyoutatu/sla-guideline.pdf⟩, (2009).
[15] 総務省: 公共 *IT* におけるアウトソーシングに関するガイドライン, 入手先 ⟨http://www.soumu.go.jp/denshijiti/pdf/060213_03.pdf⟩, (2003).
[16] JEITA: 民間向け *IT* システムの *SLA* ガイドライン第四版, ISBN-10: 4822262642, ISBN-13: 978-4822262648, (2013).