

課題番号 jh160052-ISH

## 耐災害性・耐障害性の自己検証機能を具備した広域分散仮想化基盤に関する研究とその実践的運用

柏崎礼生（大阪大学）

概要 広域に分散した研究組織が計算機資源を提供し合うことにより広域分散ストレージシステムを構築する。現在、国立情報学研究所、金沢大学、奈良先端科学技術大学院大学、京都大学、大阪大学、高知工科大学、広島大学、九州大学からなる拠点により仮想化環境の構築が完了しており、北海道大学、東北大学、琉球大学が設定中、また新たな拠点として北陸先端科学技術大学院大学を拠点として追加する予定である。このシステムを利用した仮想化環境を構築し、この上で動作する様々なアプリケーション、特に Software Defined Storage の検証と評価を行っている。上記システム、仮想化環境、およびアプリケーションに対して Software Defined Network (SDN)を用いて意図的な障害を発生させ、耐災害性・耐障害性の検証を行う基盤を構築した。

### 1. 共同研究に関する情報

#### (1) 共同研究を実施した拠点名

北海道大学、東北大学、国立情報学研究所、金沢大学、奈良先端科学技術大学院大学、京都大学、大阪大学、高知工科大学、広島大学、九州大学、琉球大学の 11 拠点。

ティ)

- 九州大学：岡村耕二（Content Centric Network）
- 琉球大学：長田智和（インターネット運用技術）

#### (2) 共同研究分野

##### □ 超大規模情報システム関連研究分野

#### (3) 参加研究者の役割分担

- 北海道大学：棟朝雅晴（分散システム技術）、杉木章義（インタークラウド技術）
- 東北大学：菅沼拓夫（サイバーリアルコンピューティング）、中村隆喜（高可用性ストレージ技術）
- 国立情報学研究所：横山重俊（アカデミッククラウド）
- 奈良先端科学技術大学院大学：市川晃平（仮想ネットワーク技術）
- 京都大学：岡部寿男（クラウドプライバシ技術）、小谷大祐（スケーラブル SDN 技術）
- 大阪大学：中川郁夫（分散オブジェクト技術）、柏崎礼生（研究総括）
- 高知工科大学：菊池豊（超広域分散環境）
- 広島大学：近堂徹（インタークラウドマイグレーション）、渡邊英伸（クラウドセキュリ

### 2. 研究の目的と意義

一定年数の研究データの保存義務や、大災害に対する障害回復の観点から研究組織における遠隔拠点へのバックアップの配置への需要が存在している。また日本国内に留まらず、国際的な研究において透過的かつ効率的にデータを共有することでより円滑な研究進捗が期待することができる。透過的ファイルシステムを提供する広域分散ストレージは、OS からそのストレージ領域をマウントすることが容易であるため、仮想計算機としての利用やこれを用いた広域ライブマイグレーションによる災害回避など、前述の需要や期待に対して高いパフォーマンスで応じることが可能である。また本基盤自身が SDN を用いて耐障害性・耐災害性を検証することにより、パフォーマンスだけでなく災害発生時のパフォーマンス劣化やサービス不能状態の規模を事前に把握することが可能となり、正常時だけでなく異常時も含めた統合的なサービス品質を定量的に評価することが可能となる。

### 3. 当拠点公募型共同研究として実施した意義

日本をはじめ、環太平洋地域の島嶼国においては特に、自然災害による情報インフラストラクチャの破壊が他地域と比較して高い頻度で発生している。センサー端末やモバイル端末から収集された時系列データを用いた防災・減災のための取り組みがこれらの地域では重要視されているが、広域分散システムは大規模自然災害による影響を受けやすい。そのため広域分散システムは単に構築するだけではなく、その耐災害性・耐障害性を検証して実際の大規模災害時に情報収集と解析が困難とならないよう十分な災害訓練を事前に行うことが求められる。

学際大規模情報基盤共同利用・共同研究拠点 (JHPCN) にはスーパーコンピュータと HPCI システムが具備されており、ビッグデータと称される高い流速・多様性・大容量を有するデータの発生源となっている。そのためデータを保持・共有する需要とマッチングしやすい。また広域分散システムの開発や検証では常時広帯域のネットワークトラフィック要求を発生させるため、十分な帯域で拠点間が接続されていなければ高い精度で問題解決を行う事が難しい。学際大規模情報基盤共同利用・共同研究拠点は超広帯域ネットワークにより接続されているため、このような研究開発と検証のための環境として極めて有益である。

また、インタークラウドは新しい研究テーマとして位置づけられている。複数の独立なクラウド間でのデータ共有は難しいとされるが、本研究が提唱するアーキテクチャでは、複数の独立なクラウドをまたいで計算機リソースを透過的に共有・参照するアーキテクチャを実現することが期待できる。本研究の成果によって、研究データの遠隔保存や共有が実現できるだけでなく、複数の独立に運用されるクラウド間での透過的なファイル共有が可能になり、例えば、産業界で生成・蓄積されるデータを、学術クラウド上で分析・解析することや、学術クラウド上のデータを商用の解析機能

を用いて処理するなど、国際・業際での透過的なデータ共有を容易にすることが考えられる。

広域分散ストレージや、これを用いた広域分散環境を国際的に展開するにあたり、まず国内で十分に広域な環境を構築した上での検証が重要である。本研究課題は複数・多数の、地理的に離れた場所にある拠点間を相互に接続し、また相互に補完しあうことで、予期できない大規模な災害・震災時であっても、ネットワークシステム全体としての機能を維持、提供できるための仕組みを実現する。現在、JHPCN を構成する拠点のうち 5 拠点に distcloud のための機材が設置され、計算機資源を提供して頂いている。最終的にはこれを全 JHPCN 拠点に拡大することを目論んでおり、拠点数の拡大により分散ストレージのパフォーマンスが向上することもさることながら、各拠点での担当者と連携することにより、様々な研究組織、様々な地域における需要に応じて本基盤が提供できるサービスの新たな展開を模索することが可能となることを重視している。具体的には既に述べたような巨大データの複数拠点での共有や災害回復手法としての利用の他に、広域ライブマイグレーションを用いた高 SLA を実現する IaaS、PaaS 基盤や、センサー端末やモバイル端末により広域から収集された大量のデータを収集・解析して高速に応答する IoT サービスなどを想定している。

各拠点の担当者が、その人的接続性を十分に活用して新たな研究との連携を発掘することにより、より多様で複雑な学際的研究を展開することが期待できる。

#### 4. 前年度までに得られた研究成果の概要

(平成 28 年度新規採択課題である)

#### 5. 今年度の研究成果の詳細

平成 27 年度まで本研究提案の代表者、分担者らは国内外 11 の研究組織（大阪大学、広島大学、金沢大学、国立情報学研究所、奈良先端科学技術大学院大学、京都大学、高知工科大学、北海道総合通信網データセンター、株式会社オキッドデータセンター、東北大学、カリフォルニア大学サンディエゴ校）からなる広域分散仮想化環境「distcloud」を構築し、太平洋横断ライブマイグレーションをはじめとして様々な実証実験を行ってきた。その後、SDN を用いて広域分散システムの耐災害性・耐障害性を検証する「DESTCloud」を開発し、災害シナリオを用いて広域分散システムを構成する実ネットワークに対して意図的な障害を発生させることで耐災害性を検証するプラットフォームを構築した。これらの基盤は情報通信研究機構が提供する JGN-X を用いた拠点構成で実現されていたが、JGN-X の終了と SINET5 との部分的統合、本研究提案の参加拠点の多くで 100Gbps 接続が実現されたことから、この基盤を SINET5 へ移行することとした。この作業は国立情報学研究所、金沢大学、奈良先端科学技術大学院大学、京都大学、大阪大学、高知工科大学、広島大学、九州大学、琉球大学においては終了しており、現在、東北大学のネットワーク構成を整備中である。この整備が終わり次第、これら二拠点も distcloud、DESTCloud の拠点として利用可能となる。

本計画では distcloud の拠点数をさらに拡大させ、その拠点間で DESTCloud を動作させることで広域分散仮想化基盤と、その耐災害性・耐障害性の検証を同時に実現できるプラットフォームを構築し、その実践的な運用の知見を得ることを目的とした。また、既に参加しているカリフォルニア大学サンディエゴ校のほか、ニューヨーク、ロサンゼルス、シンガポール拠点の構築を進捗させ、国際的な学術研究環境を構築することを予定していた。これ

らの予定については情報通信研究機構との共同研究体制の下で推進する予定であったが、この体制の確立に失敗したため、現在は PRAGMA との連携体制のもと、環太平洋地域の拠点の拡大に努めることとしている。この環境の上で透過的かつ効率的にデータを共有するための透過的分散オブジェクト技術や、広域分散アプリケーションの実証実験を行っている。具体的には各拠点に設置された物理計算機を仮想化ホストとして動作させ、その上で複数の仮想計算機 (Virtual Machines: VM) を稼働させる。VM には分散アプリケーションのサービスを動作させ、拠点間は SINET5 の L2VPN および L3VPN で接続されている。拠点間の接続は SINET5 のトポロジに殉じるが、より現実に即した経路での通信を模擬できるよう、QinQ 技術を用いて拠点間の仮想的なパスを複数構成している。これらの拠点を構築するためのドキュメントも整備されており、本研究に参加を希望する拠点はより少ないコストで広域分散環境と耐災害性・耐障害性検証実験に参加することができる。DESTCloud のコントローラは現在広島拠点に設置された計算機上で VM として動作している。

また、北海道大学のデータサイエンスクラウドが提供する物理サーバ 2 台と仮想サーバ 8 台を用いて distcloud 環境の模擬環境を構築している。物理サーバ 1 台は、distcloud 環境を利用するユーザが NFS プロトコルを用いてファイルシステムをマウントするためのアクセスサーバとして利用し、もう 1 台は Linux container を用いた動的なコアサーバ（後述）のデプロイ環境として検証を行う。仮想サーバ 8 台は distcloud のバックエンドを支えるコアサーバとして動作し、実際のデータの保存と分散 Key-Value Store のノードとして動作する。データを保存するため、この仮想サーバにはそれぞれ 1TB のストレージを接続する。また、現在東北大学に設置した distcloud 環境の整備を行い、北海道大学と共にこれらの環境を distcloud 環境に SINET を用いて接続し、広域分散ストレージと、その上で動作するアプリケーションの評価

実験を行う。また distcloud 環境の新たな取り組みとして、distcloud を支えるストレージエンジンを高速かつプログラマブルに変更可能な仕組みとする手法について実装を行いその検証を行った。この仕組みを導入することにより、開発版と安定版のストレージエンジンを並行稼働させることができ、シームレスな環境のアップグレードを実現することが期待できる。

## 6. 今年度の進捗状況と今後の展望

国内拠点数の増大については当初計画を遂行できているほか、北陸先端科学技術大学院大学のように計画当初にはなかった拠点の参加を得ることもできた。また北海道大学とは本研究提案とは別にデータのレプリケーションサイトを相互に所有し合う共同研究を推進させているが、本研究提案で取り組んでいる distcloud のノードとして追加することも想定することが可能であり、当初予定よりも本研究提案の活動が拡大している。

その一方で海外拠点の拡大については当初の目標を完遂することが出来なかったことは事実として認めなければならない。前章で述べたように PRAGMA との関係性を発展させ、当初予定の拠点とは異なる海外拠点を設置する方向性で拡大させることができている。また、本研究提案の分担者である横山教授が APAN (Asia Pacific Advanced Network) の Cloud working group の co-chair を退任されたことに伴い、本研究提案の代表者である柏崎が co-chair として就任することとなった。これにより PRAGMA のみならず、APAN の参加組織に対しても本取り組みについて広報する機会を得たこととなる。APAN ではオセアニアから西アジア、南アジア、東南アジア、東アジアに渡る広範な学術組織が参加している。2016 年度は APAN42、APAN43 でそれぞれ発表を行い、distcloud および DESTCloud の活動に関する取り組みを紹介した。

これらの普及活動と並行し研究協力の申し出を受けて distcloud 用の 120 コア、640GB の計算機資

源を確保することに成功した。この計算機資源は新規拠点や計算機資源が不足している拠点に対して重点的に配分していく予定である。DESTCloud については現在、4 つのサブテーマを確立して研究開発を推進している。

### (1) システムの進化プロセスの形成

堅牢性を高め、より安定したシステムを構築・運用していくためには、PDCA サイクルにおいて定量的な「Check (検証・評価)」をもとに、最適化されたプロセスへと漸近させる「Action (改善のための行動)」へ繋げるための「Reflection (反映)」が重要となる。対象とする広域分散システムを本プラットフォームで検証・評価するなかで抽出した問題点に対して、改善フェーズの自動化・最適化やシステムへの適切なフィードバックを実現するための手法について検討する。

### (2) 人的プロセス・人工災害を含めた応用ドメインの拡張

現状は人的な要素がほとんど勘案されておらず、人的要素への対応で応用範囲を拡張し、より汎用性を高める。まず、ICT 基盤の運用プロセスにはソフトウェアにより自動化された制御構造が含まれるが、人間(オペレータ)による判断が含まれる場合もある。さらに、災害を自然災害だけでなくネットワークアタック等の人工災害に拡張することで、より適用範囲が広がる。そこで収集コストの低減と、行動ログ解析手法の確立を目指し、運用プロセスに関わる人間の行動や外部/内部からの人工的な攻撃も含めた耐災害性・耐障害性の検証を行う。

### (3) レジリエンス評価基準の確立

災害時・障害時におけるサービスを品質を評価することにより、ICT 基盤における耐災害性・耐障害性の定性的、汎用的な評価基準を確立する。耐障害性を高めるプロトコルの分類と、災害・障害に関する体系的な整理を行うことにより、災害・障害に対して対応可能であるかどうかを判定し、実際に防災訓練を

実施することで確認を行う。また、災害・障害の規模の格付けを行うことにより、それに対応可能な耐障害性・耐災害性の格付けを確立する。このような取り組みを通して最終的には定量的・汎用的な指標の定義を目指す。

#### (4) 持続的な研究を実現するためのモデル確立

耐災害性・耐障害性を検証する活動を今後、公的資金による補助なしで持続させるための方法論について検討し、ビジネスモデルの策定とその実証を通して継続的な研究開発を実現する。より学際的な連携、より実学的な研究開発を持続し、啓発活動、普及活動、ロビー活動を行うことと並行して、これらの活動を評価し、改善に結びつける。

この活動を推進するため平成 29 年度からの科学研究費基盤研究(B)特設分野研究に応募を行ったほか、平成 29 年度戦略的創造研究推進事業による RISTEX、さきがけの提案を行う予定である。

JHPCN による共同研究体制のみならず、北海道大学情報基盤センター、東北大学電気通信研究所、国立情報学研究所とそれぞれ平成 29 年度の共同研究を行う体制を確立している（平成 29 年 5 月 15 日現在）。distcloud 拠点としては東京工業大学、北陸先端科学技術大学院大学のほか、PRAGMA との協働によりフロリダ大学への設置も検討中である。このように本研究提案は 2015 年度における申請当時より拡大した展開と、具体的な進捗を達成することができている。

## 7. 研究成果リスト

### 7.1 課題期間前半

#### (1) 学術論文

- “広域分散システムの耐障害性を評価する検証プラットフォームの実装と評価”，北口善明，柏崎礼生，近堂徹，市川晃平，西内一馬，中川郁夫，菊池豊，情報処理学会論文誌，Vol. 57, No. 3, pp. 958--966 (2016)
- “自動構成機能を有する大規模キャンパスネ

ットワーク管理システムの実装と評価”，近堂徹，田島浩一，岸場清悟，岩田則和，相原玲二，情報処理学会論文誌，Vol. 57, No. 3, pp. 998-1007 (2016)

#### (2) 国際会議プロシーディングス

- “A design and implementation of global distributed POSIX file system on the top of multiple independent cloud services”，Ikuo Nakagawa，Hiroki Kashiwazaki，Shinji Shimojo，Kohei Ichikawa，Tohru Kondo，Yoshiaki Kitaguchi，Yutaka Kikuchi，Shigetoshi Yokoyama and Shunji Abe，In Proc. of the 5th International Congress on Advanced Applied Informatics (IIAI AAI 2016)，pp. 867-872 (2016)
- “Design and Implementation of Global Reference and Indirect Method Invocation Mechanisms in the Dripcast”，Ikuo Nakagawa，Masahiro Hiji，Hiroshi Esak，Proc. of IEEE COMPSAC，pp. 338-343 (2016)
- “PRAGMA-ENT: An International SDN Testbed for a Cyberinfrastructure in the Pacific Rim”，Kohei Ichikawa，Pongsakorn U-chupala，Che Huang，(全 21 名，1 番目)，Concurrency And Computation: Practice And Experience，(in press) (2016)
- A Simple Multipath OpenFlow Controller using topology-based algorithm for Multipath TCP’，C. Nakasan，Kohei Ichikawa，H. Iida，P. Uthayopas，Concurrency And Computation: Practice And Experience (in press) (2016)
- SAGE-based Tiled Display Wall Enhanced with Dynamic Routing Functionality Triggered by User Interaction’，Yoshiyuki Kido，Kohei Ichikawa，Susumu Date，Yasuhiro Watashiba，Hirotake Abe，Hiroaki Yamanaka，Eiji Kawai，and Shinji Shimjo，Future Generation Computer Systems (FGCS)，

Vol. 56, pp. 303-314 (2016)

### (3) 国際会議発表

- “DESTCloud: now and future”, Hiroki Kashiwazaki, Asia Pacific Advanced Network 42 meeting (APAN42), Disaster Mitigation Workshop.

### (4) 国内会議発表

- 柏崎礼生: 地域間インターネットクラウドと減災の 5 年間, 第 10 回地域間インターネットクラウドワークショップ, ITRC Technical Report [to appear] (2017).
- 中川郁夫: インタークラウド環境での秘匿分散統計解析手法を応用したビジネスモデルの検討, 第 10 回地域間インターネットクラウドワークショップ, ITRC Technical Report [to appear] (2017).

### (5) その他 (特許, プレス発表, 著書等)

特になし

## 7.2 課題期間後半

### (1) 学術論文

特になし

### (2) 国際会議プロシーディングス

特になし

### (3) 国際会議発表

- Hiroki Kashiwazaki: Brief discussion of container leveraged clouds, APAN43: The 43rd Asia-Pacific Advanced Network Meeting, Cloud working group (2017).
- Hiroki Kashiwazaki: distcloud: distributed virtualization platform, APAN43: The 43rd Asia-Pacific Advanced Network Meeting, Cloud working group (2017).

### (4) 国内会議発表

- 柏崎礼生: Short reports from NZNOG2017/NANOG69/APAN43, 第 11 回地域間インターネットクラウドワークショップ, ITRC Technical Report [to appear] (2017).
- 北口善明: クライアント OS の IPv6 実装調査, 第 11 回地域間インターネットクラウドワークショップ, ITRC Technical Report [to appear] (2017).
- 小谷大祐: アジャイル型共創による高齢者補助ロボット用ネットワークプラットフォーム技術の研究開発, 第 11 回地域間インターネットクラウドワークショップ, ITRC Technical Report [to appear] (2017).
- 柏崎礼生: post DESTCloud: RICC の次の取り組み, 第 40 回インターネット技術第 163 委員会研究会, ITRC Technical Report [to appear] (2017).
- 北口善明: DESTCloud プロジェクトのこれまでとこれから, 第 40 回インターネット技術第 163 委員会研究会, ITRC Technical Report [to appear] (2017).
- 柏崎礼生: 地域間インターネットクラウドと減災の 5 年間, 第 10 回地域間インターネットクラウドワークショップ, ITRC Technical Report [to appear] (2017).
- 柏崎礼生: SCOPE と DESTCloud の 2 年間とこれからの RICC/D4Cloud, 第 39 回インターネット技術第 163 委員会研究会, ITRC Technical Report [to appear] (2017).

### (5) その他 (特許, プレス発表, 著書等)

特になし