

ロバストなビッグデータ利活用基盤技術の研究開発

R&D for fundamental technology for use of robust big data

研究代表者

北林巧巳, 株式会社 Skeed

Takumi Kitabayashi, Skeed Co. Ltd.

研究分担者

飯澤徹平†
Teppei Iizawa†

神武直彦††
Naohiko Kohtake††

嶋田大輔†
Daisuke Shimada†

†株式会社 Skeed ††慶應義塾大学大学院システムデザイン・マネジメント研究科
†Skeed Co. Ltd. ††Graduate School of System Design and Management, Keio University

1. ビッグデータを取り扱うシステムの課題

2. 自律分散型P2P(Pear to Pear)通信技術

ビッグデータのデータ流通/処理基盤として自律分散P2P技術の適用を提案

→「ロバストなビッグデータ利活用基盤」

3. 「ロバストなビッグデータ利活用基盤」のアーキテクチャ

4. 検証実験結果

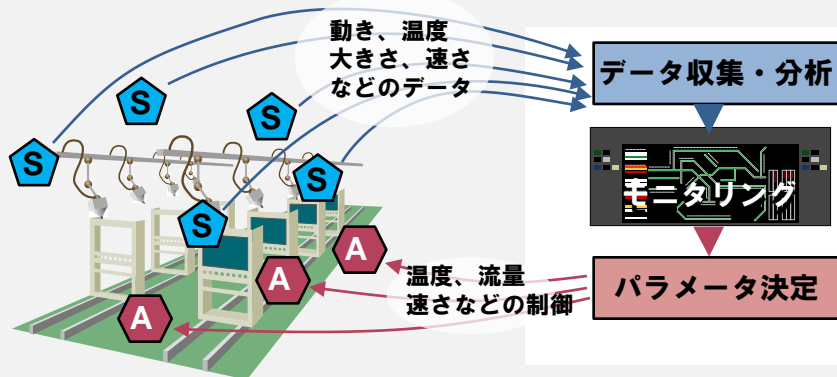
5. 今後の事業化構想

1. ビッグデータを取り扱うシステムの課題

多数のセンサーとクラウド上のコンピュータが連携して産業や社会インフラを支援

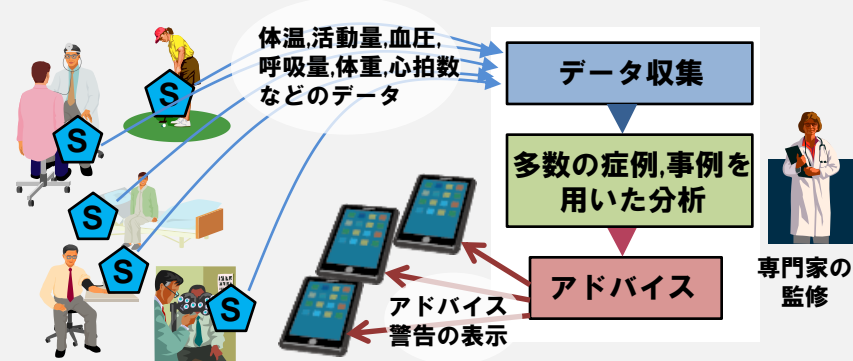
◆ 産業応用

工場の生産設備等に設置したセンサー類から送信されるデータに基づいて、設備全体を制御



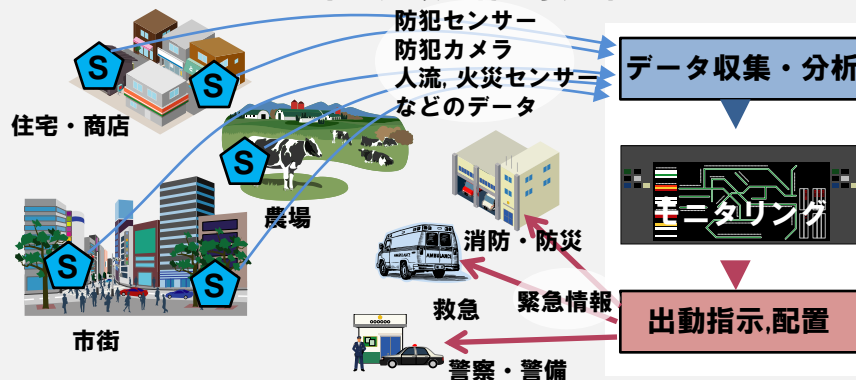
◆ 健康/ヘルスケア

人が持つセンサーデバイスから活動量や心拍数などを送信し、データを分析して健康に役立つ提案をする



◆ セキュリティ

施設に設置されたカメラや動物体センサー赤外線センサーなどからのデータを集め施設全体の安全管理をする



◆ スマート社会インフラ

街の施設や住居に設置されたエネルギーや水などの資源消費量センサーの情報を集め、街全体の省エネ化を進める



1-2. ネット接続される機器の激増

ネット接続機器の拡大傾向について

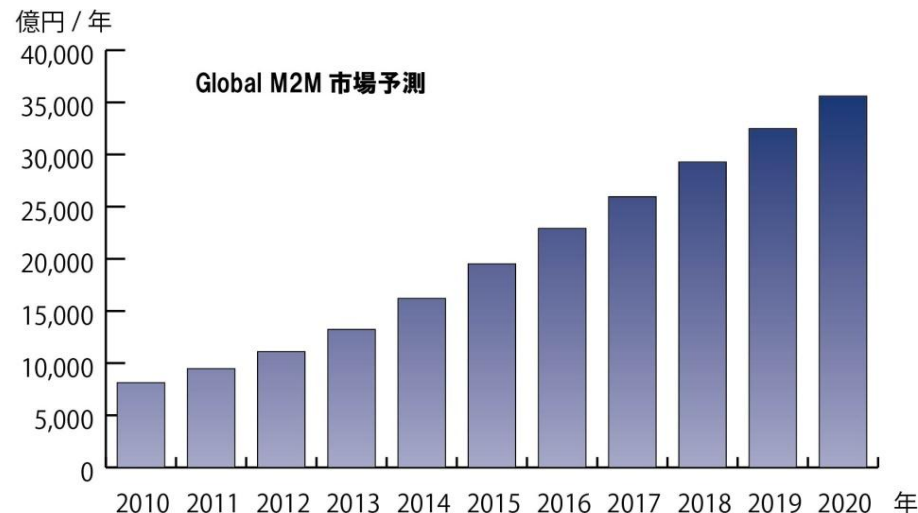
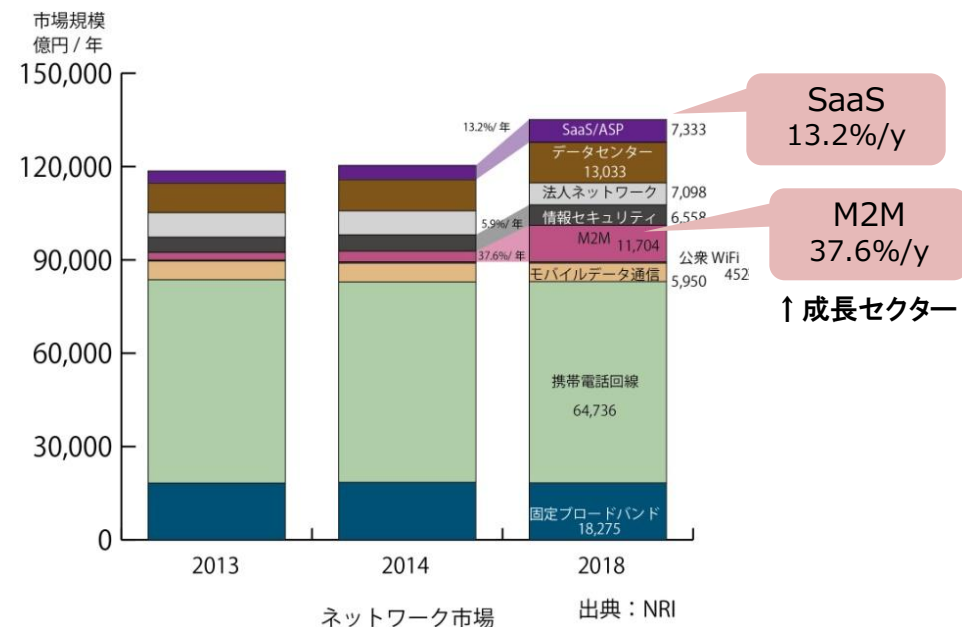
- 2020年までにネットに接続されるエンドポイントは300億台と予測

機器同士が直接ネットで通信する形態

→M2M(Machine to Machine)通信

M2M市場規模予測

- IoT市場の売り上げは8兆9000億
- 経済価値は1兆9000億ドル



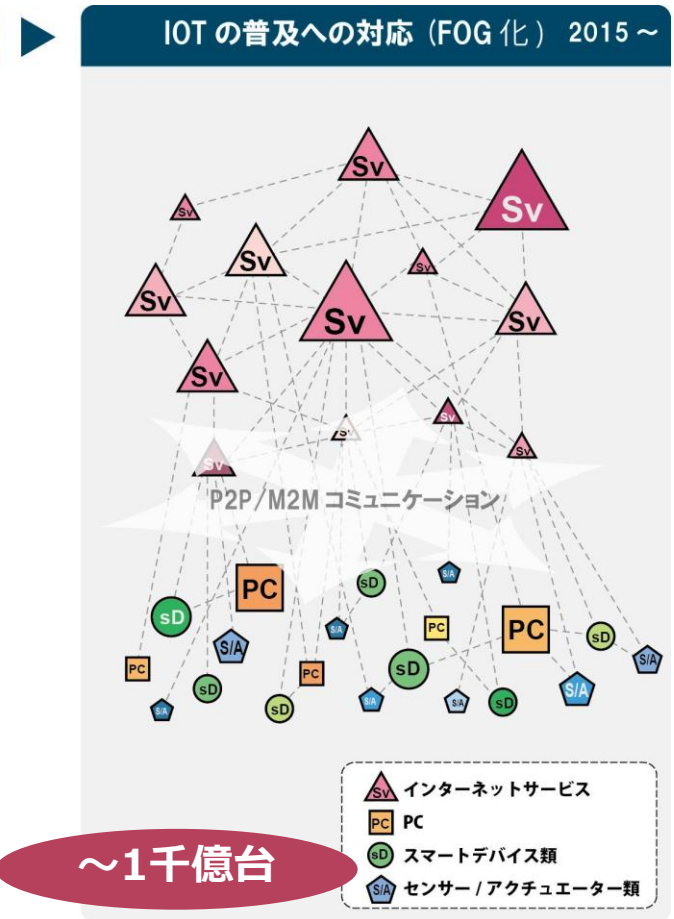
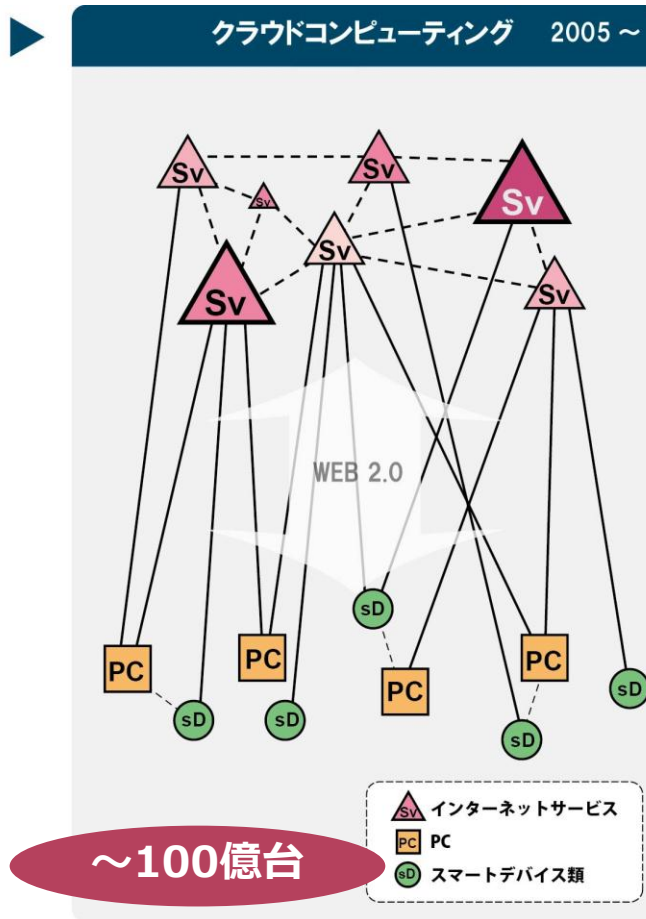
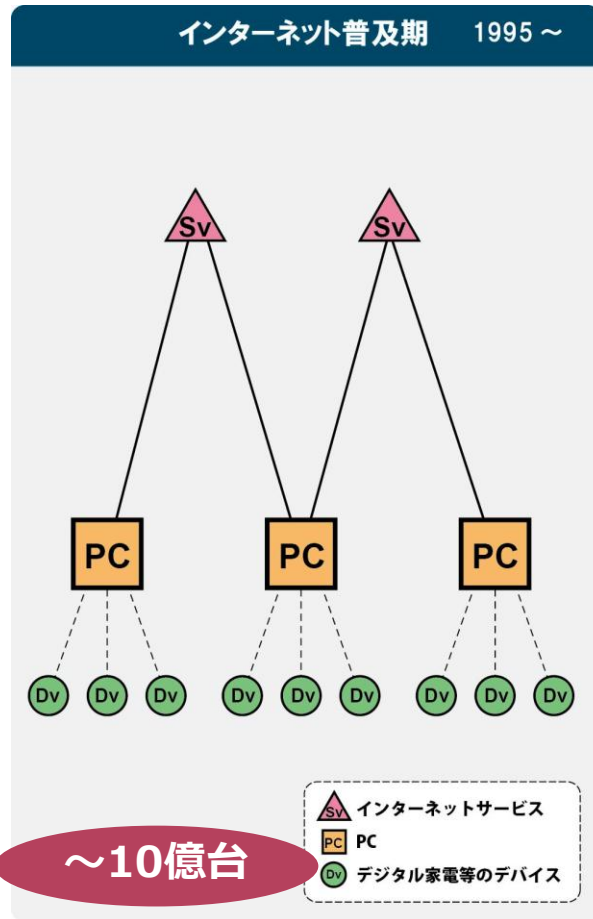
市場規模内訳: IoT デバイス, 電子部品, サービス用プラットフォーム (ハード/ソフト), サービス事業

※2014/3/28 矢野経済研究所発表資料より引用

1-3. ビッグデータ利用拡大によるネットワーク形態の変化

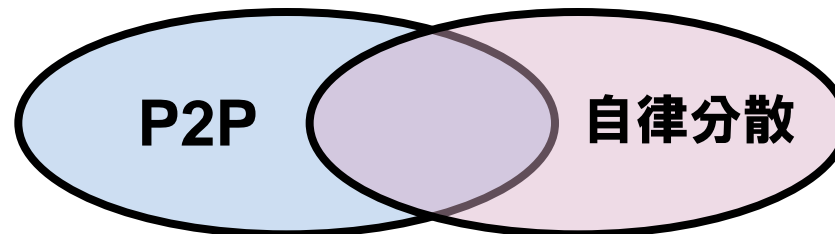
❖ 端末数の拡大と、各端末の専用端末化に伴い通信トラフィックの流れが変化

- M2M端末の数の増大、各端末の機能/性能の制限により、クライアント-サーバ型の集中型通信から装置同士や装置-サーバー間の分散型直接通信が増大する

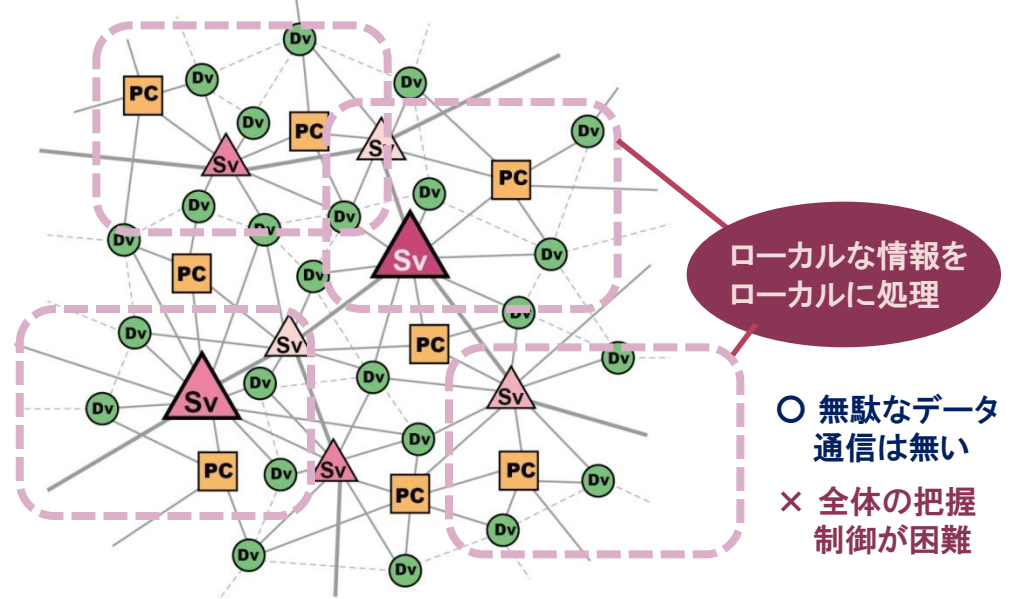
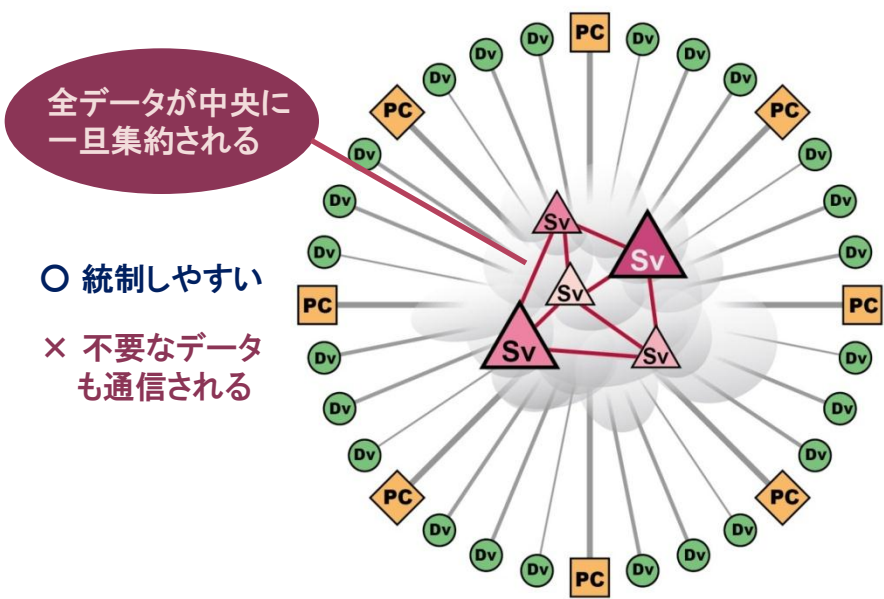


2. 自律分散型P2P(Pear to Pear)通信技術

ビッグデータのデータ流通/処理基盤として自律分散P2P技術の適用を提案
→「ロバストなビッグデータ利活用基盤」

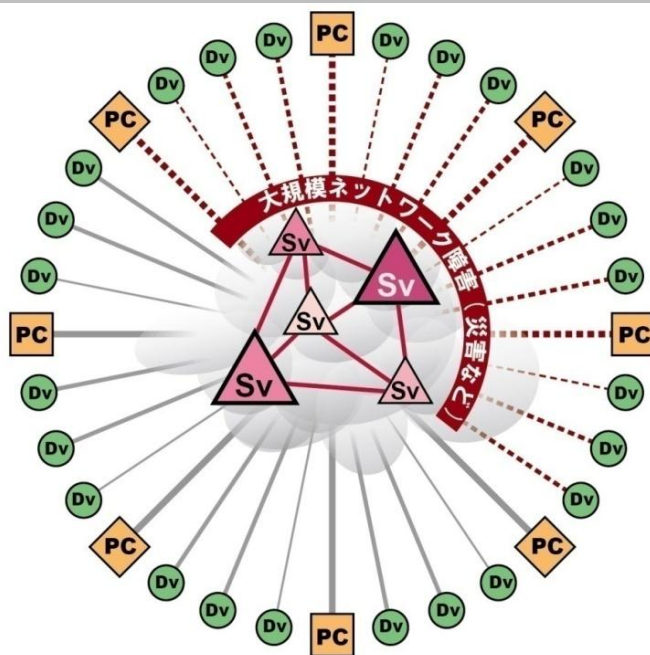


- ❖ P2Pではすべてのデータを中央のサーバに集約して処理するのではなく、様々な機能をネットワークエッジで分散して処理する。
- ❖ 情報の地産地消により、トラフィックの偏在による課題(コスト,信頼性)を解決



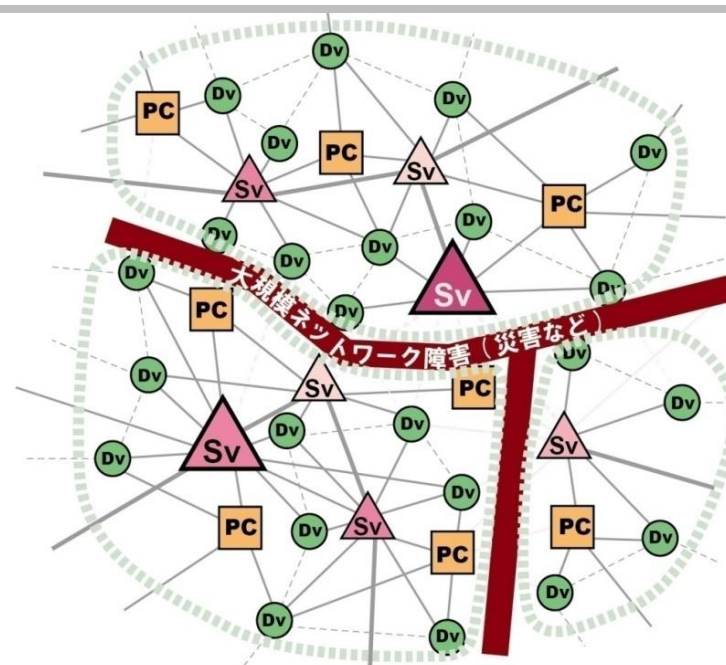
	クライアント - サーバ型 集中処理	P2P通信活用分散処理
○利点	管理/統制が容易 これまでの実績, 技術の蓄積	無駄な通信が少ない (情報の地産地消) <i>耐障害性 (→ 次ページ)</i>
×課題	不要なデータも通信される <i>耐障害性 (→ 次ページ)</i>	全体の把握と制御が煩雑 (→ 自律分散技術)

- ❖ サービス(サーバ)やネットワークの一部に障害が発生しても、残る部分が自律的に機能提供を続行する
 - 障害のみでなく、メンテナンスや更新、サーバの追加、廃止などの時も同様



集中型におけるサーバやネットワークの障害

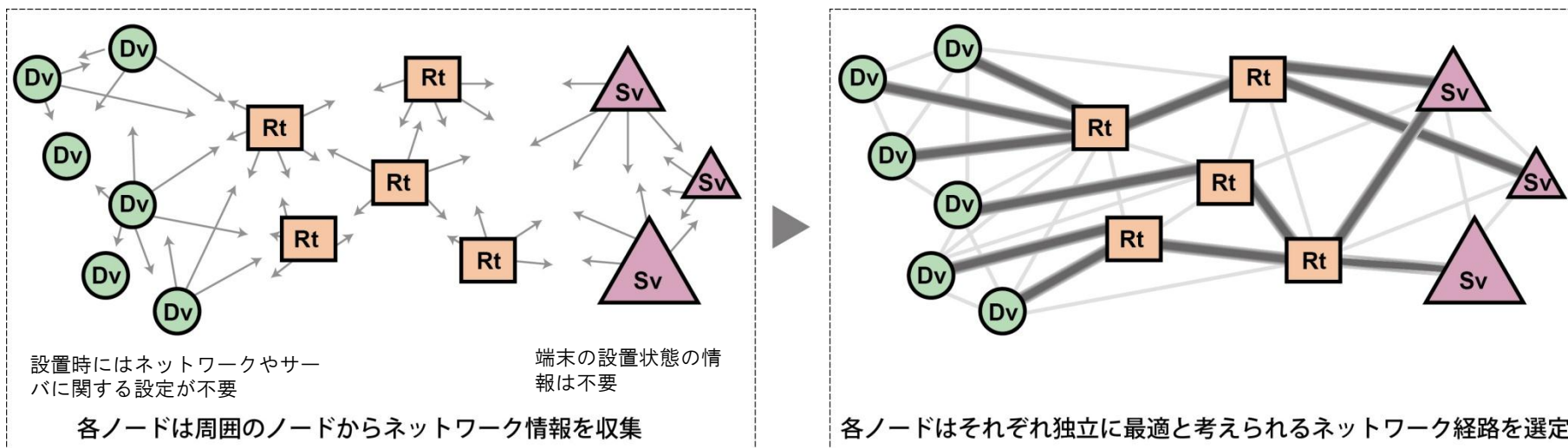
障害が発生したサーバに括り付けられた端末や障害が発生したネットワークに接続された多数の端末が機能できなくなる。



自律分散システムにおける障害

障害時も残ったリソースやネットワーク経路を使ってサービスを続行する
ネットワークが分断されてもその“島の中”で可能な限りの処理を継続する

- ❖ (前述): P2Pネットワークは、そのままでは全体の状況の把握が困難で制御も煩雑
 - 全体を制御/管理する装置がネットワークに存在しない
- ❖ そこで各端末が自律的に**自端末の周囲のネットワークを構築する機能を搭載**することが望ましい。
 - 人の手や、上位のサーバなどの指示による設定が不要
 - 動的に変化するネットワークに対して、常に最適なリソース配分になるよう自動調整



設置時にはネットワークやサーバに関する設定が不要

● Dv データ発生端末 □ Rt データ中継装置 ▲ Sv データ処理装置

初期状態

ネットワーク構築成功状態

3. 「ロバストなビッグデータ利活用基盤」のアーキテクチャ

自律分散P2Pを実現するための三要素

自律分散ネットワーク

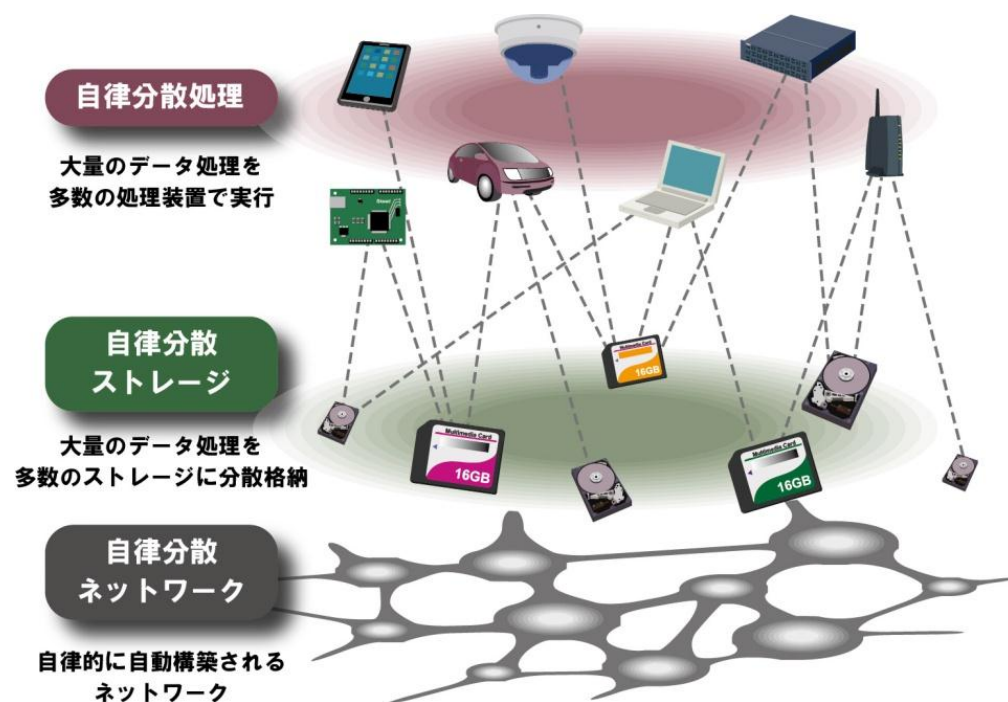
- それぞれの端末が自律的に経路を構築するネットワーク

自律分散ストレージ

- 複数の物理的に冗長的に分散蓄積するデータストレージ

自律分散処理

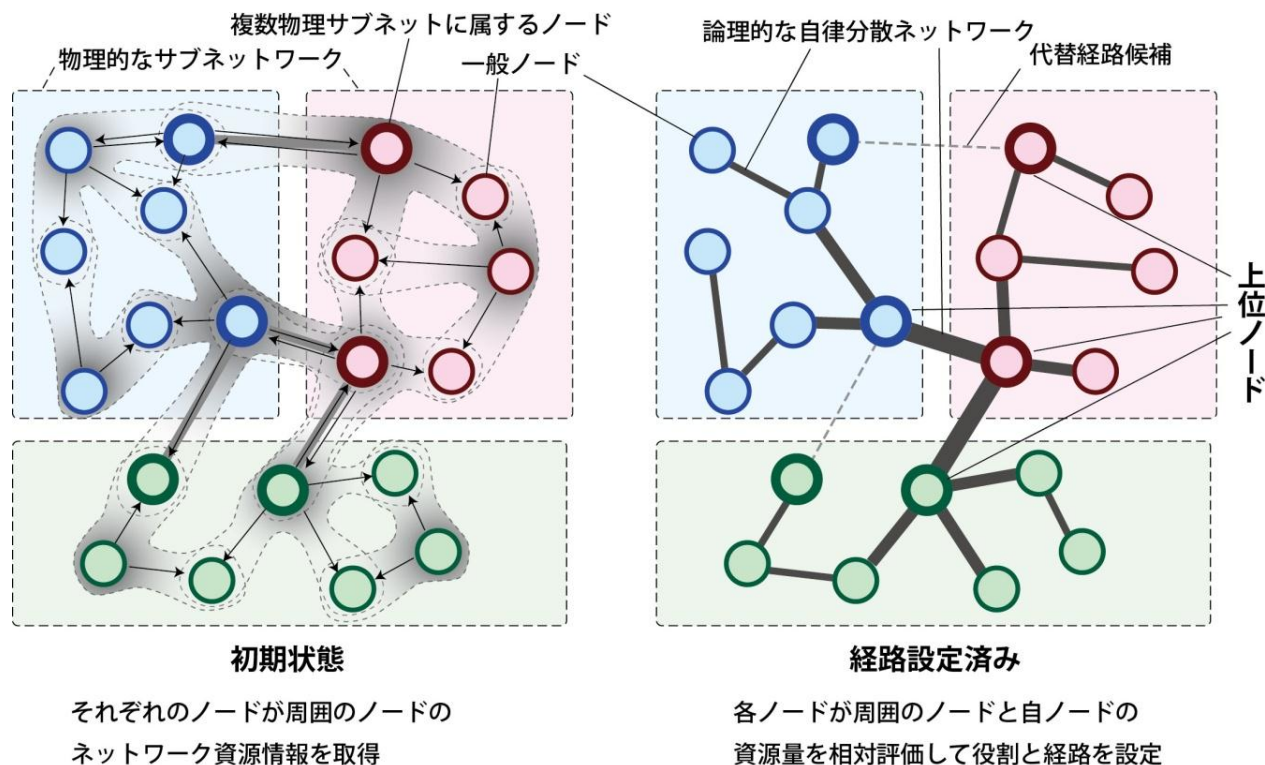
- 多数の演算リソースを使った並列分散処理機構



三要素の検証実験概要

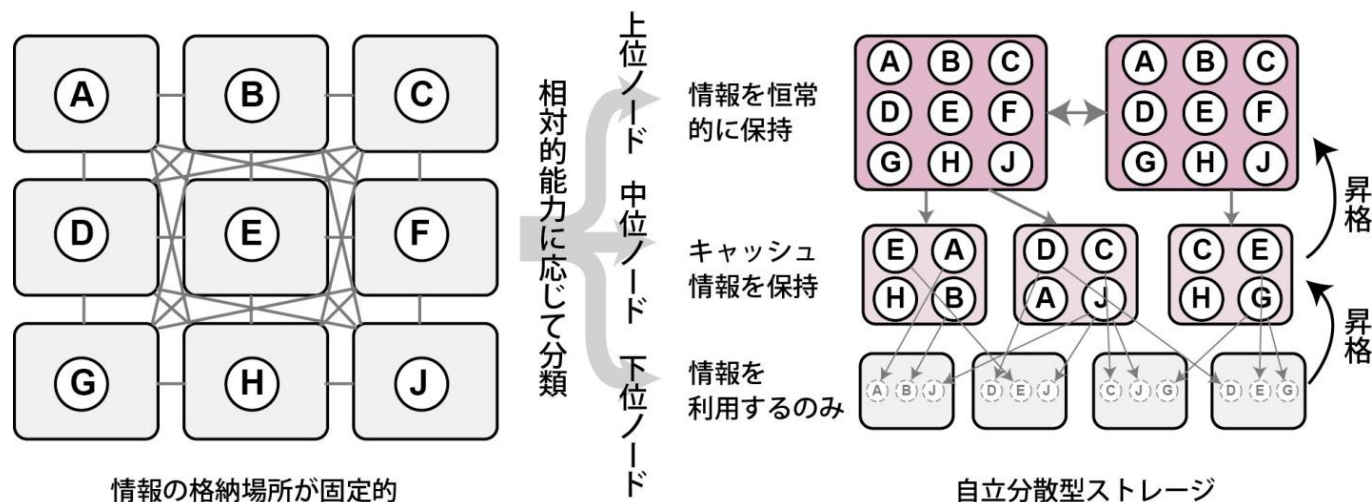
	自律分散型ネットワーク	自律分散ストレージ	自律分散処理
検証した技術	自動的経路探索機能	データの最適な分散配置方式	ネットポロジ-変化に適応する処理の分散配置
	経路情報の管理手法	障害時におけるデータの復元	データの格納場所に適応する処理の分散配置
	P2P向け通信プロトコル	重複データの排除	データ分割による機密性向上

- ❖ 近傍のノードと相互通信して自律的にグループと通信経路を構成
 - 複数の経路候補を保持し障害時には迂回経路を利用する 【通信到達性, 耐障害性】
- ❖ 自ノードや周囲のノードの特性を相対評価しノードの役割(階層)を自動設定
 - グループの最大数を制限することにより経路解決を高速化 【経路探索コスト低減】
 - グループをまたがるノードを能力評価に基づいて自動選出して昇格 【通信効率向上】



自律的な経路構築の概念図

- ❖ 複数ストレージ装置に分散格納するKVS(Key Value Storage)型ストレージシステム
 - 分散と冗長化により、一部に障害が発生してもデータを再現することが可能
- ❖ 全体を統括する装置を置かず、近傍での相互評価により各ノードの役割を自動調整
 - 連続稼働時間(電源状態)とストレージ容量を主とした総体評価により上中下の3階層に分化
 - 上位層:構造化オーバーレイによる分散KVSを構築
 - 多重化されたKVSで検索の確実性を保証 データの保存と検索に使用 【冗長化】
 - 中間層:非構造化オーバーレイによるキャッシュを構築
 - 検索内容と検索結果のキャッシュを層全体で共有 【低遅延化】
 - 下位層:検索結果と取得したデータのキャッシュとして利用
 - 検索内容と検索結果のキャッシュを近傍のみで共有 【低遅延化】



ノードの役割の自動調整例

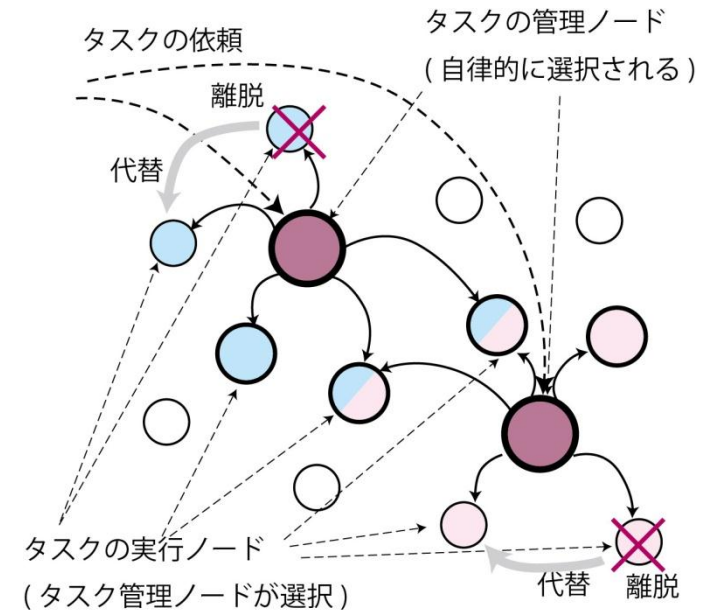
❖ 全体を統括する装置を持たない自律的な分散並列処理

❖ 演算性能と連続接続時間を元に自律的に3階層に分類

- 上位層を管理用エンドノードを選出
- 中位層を作業者として演算に使用
- 下位層は自律分散処理には利用しない

❖ 構成変化への対応／耐障害性の実現

- ホットスワップにより管理機能を継続
 - データ処理中のノードや、データを蓄積しているノードが動的に離脱・参加を繰り返しても、データ処理機能、蓄積データを保持
 - ノード離脱を検知した段階で他の有効なノードに蓄積・処理が引き継がれ、参加したノードがネットワークに認識されると、蓄積・処理が自律的に割り当てられる
- 新たなエンドノードの選出と演算の再割当
 - 処理リクエスト毎に、そのリクエストに対する管理ノードが選出される。同時に副管理ノードも決定することで、管理ノードが離脱することがあっても処理を継続【動的に最適なものが選ばれる】



ノードの役割の自動調整

4. 検証結果

❖ 検証の目的

- **大規模なデータ収集・蓄積・処理能力を実現するために**
 - スケールアウトが可能
 - ロバスト性(耐障害性)
 - セキュリティの確保

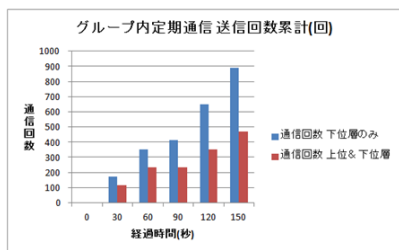
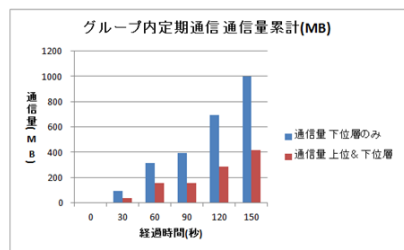
など兼ね備えるシステムの一つの方法として、「自律分散型P2Pネットワーク」を基盤とするシステムがこれらの要件を満たしていることを検証し、その特性を明らかにする。

❖ 検証内容:前記3要素それぞれについて目標を設定し、検証ソフトウェアを主にJava言語で開発、NICTのStarBED³などのx86サーバ上の最大1万ノード超を使用して機能性能の検証を実施した。

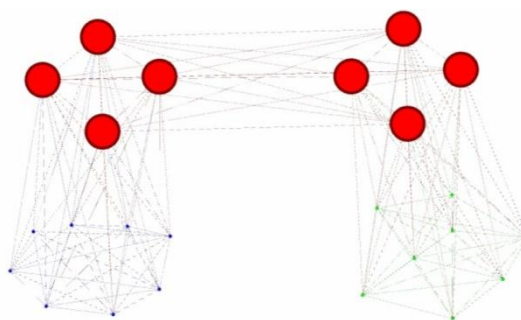
- 1. 自律分散ネットワーク
 - 経路の自律的構築 経路構築時間 障害時の代替経路構築時間
- 2. 自律分散ストレージ
 - レスポンスタイム 一部ノード障害時のデータ復元性
- 3. 自律分散処理
 - 一部ノード障害時も処理を完結すること 処理能力が利用可能ノード数に応じて拡大

🔗 目標と検証結果

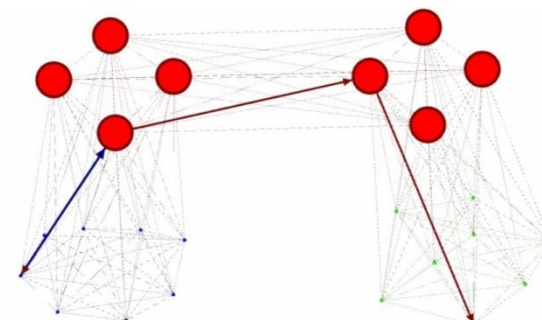
到達目標	評価項目	目標	達成	備考
数百から数百万ノードまでのネットワーク規模で、自律分散型ネットワークの構築に要する時間を5秒以内とする。	新規ネットワークの構築	5秒以内	◎	400ノードにて0.69秒を達成 数百万ノード予測値も同等
ノード数の50%の増減が発生した場合や、ネットワーク経路の半数程度の切断などの変化が発生した場合でも、自律分散型ネットワークの再構築に要する時間を3秒以内とする。	既存ネットワークへの参加	3秒以内	◎	0.15秒を達成
	既存ネットワークの大規模な経路変更	3秒以内	◎	0.39秒を達成
	通信の暗号化		○	



グループ化による通信量削減



ネットワーク構築完了状態



Pear2Pear通信中

階層的グループ間での通信

❖ 目標と検証結果

到達目標	評価項目	目標	達成	備考
・データの機密性を維持する。	適切な権利者による検索	可能であること	○	
	権利を持たないデータ	検索結果に反映されないこと	○	
・データを蓄積するノード数が20%減少した場合でも蓄積されたデータの完全性を損なわない。	ノードの20%離脱	全てのデータにアクセスできる	○	
	ノードの20%離脱時の継続性	データの再配置が行われている最中でも、すべてのデータにアクセス可能なこと	○	
・自律分散型ネットワーク全体のトラフィック偏在を生じさせないようデータの蓄積先を自律的に変更する。	ノードの20%離脱によるデータの偏在	自律的に解消されること	○	
・十分な安全性・信頼性が評価された暗号を用いる	蓄積データの暗号化	蓄積データが判読できないように暗号化されていること	○	
・暗号化時の分散蓄積するための所要時間は3倍以内とする。	データ保存の所要時間	ローカル保存時の3倍以内	○	1.45ms (0.5msに対し)
・暗号化した蓄積データ百万件に対して検索を行った場合の処理応答時間を1秒以内とする。	検索速度	百万件の暗号化蓄積済みデータへの検索に1秒以内で応答すること	○	平均0.2秒 (百万ノード)

❖ 目標と検証結果

到達目標	評価項目	目標	達成	備考
・データを処理するノード数が増加した場合、その増加に応じて処理能力を向上させながら処理を継続する。	処理ノードの増加	増加しない場合に比べて処理完了までが短くなること	○	
・処理中にノード数が頻繁に変化する場合でも処理を完全に完了することが可能。	処理ノードの頻繁な変化	処理が正常に完了すること	○	
・検索処理に分散処理技術を適用し、分散処理にかかるノード間のメッセージとノード内で処理中のデータを暗号化したまま、このデータの所有者が検索を完了する。	暗号化	処理に関わる通信及びノードで処理中のデータが暗号化されていること	○	
	百万件に対する検索処理の実行	1秒以内。	○	平均0.2秒

✦ 検証された特長

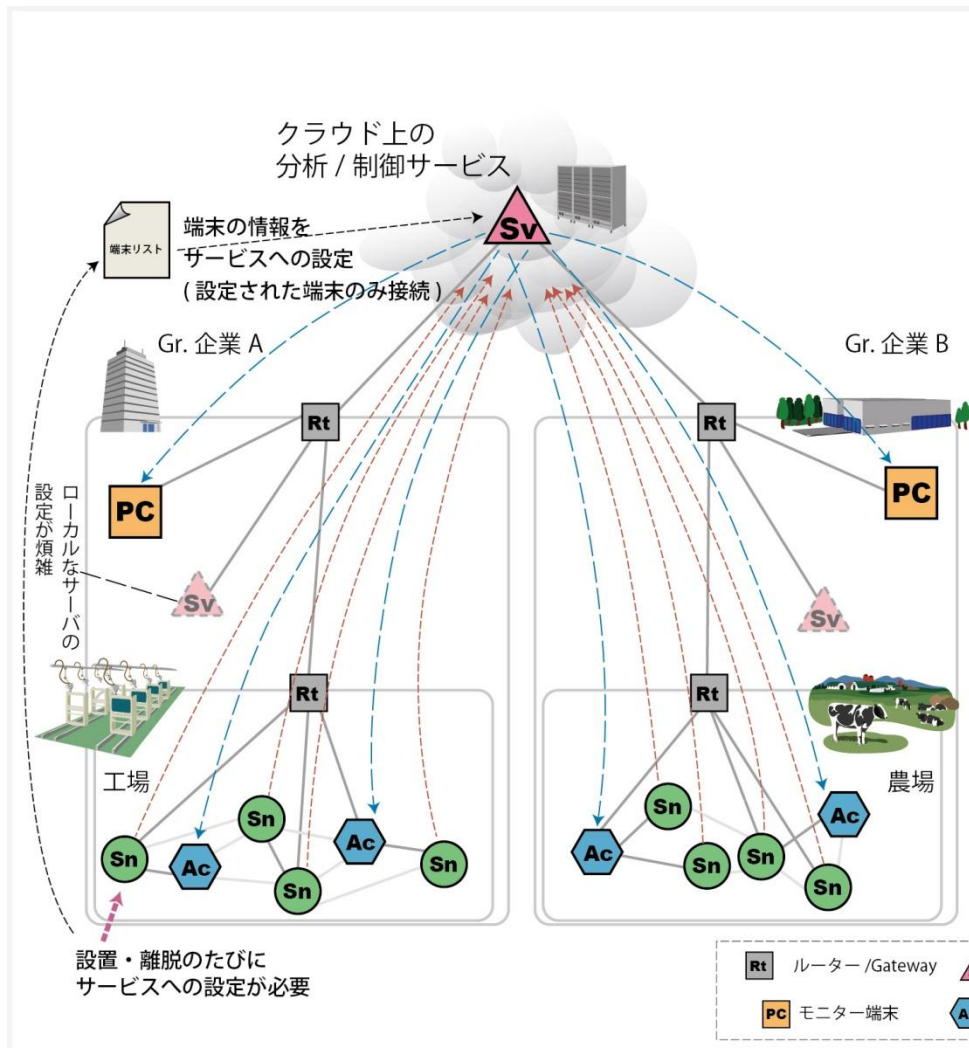
	構造上の特徴	得られるメリット・特長
共通	固定的な管理ノードを持たない	環境変化への 動的な適応 理論上無限のスケールアウト性 耐障害性
自律分散ネットワーク	自律的, 動的な経路構築	通信到達性確保 耐障害性
	相対的な端末特性に応じた自動的な役割アサインとグループ化	経路構築の高速化 伝送効率の向上
自律分散ストレージ	分散格納と冗長化	ストレージコスト削減 データ保持の確実性向上
	相対的な端末特性に応じた自律的な役割アサイン	アクセス時レスポンス向上 耐障害性
自律分散処理	分散処理	処理の高速化 データの近傍で処理(通信コスト)
	相対的な端末特性に応じた動的な役割アサイン	環境変化への 動的な適応 耐障害性

5. 今後の事業化構想

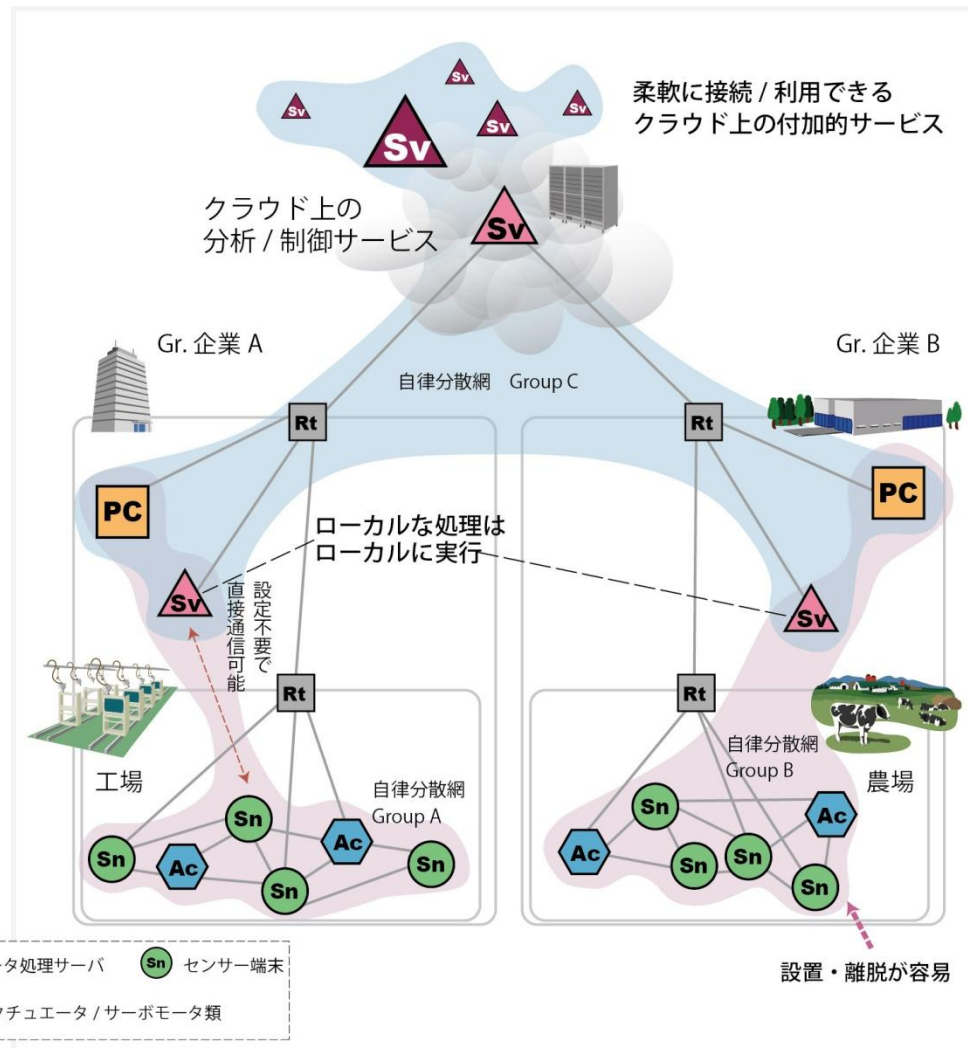
4-5. 今後のロバストなビッグデータ利活用基盤の適用先例

環境が過酷で変化が激しい場所に設置されるセンサーデバイスの通信にP2Pを活用

- 設置, 設定, 代替が容易 情報の地産地消によりトラフィックの集中を回避



クライアント-サーバ型接続による M2M



P2P 型接続を活用した M2M

4-6. 適用先の具体的例 - 1 パーソナル

分野	対象製品 サービス	適用ケース	効果・メリット例
ヘルスケア/ ウェルネス	ヘルスケア	活動量計を搭載したウェアラブルデバイスによる健康維持・管理および医療サービスとの連携による未病	高齢者のQOL向上 医療費の削減
	ペットケア	活動量計を搭載したウェアラブルデバイスによるペットの健康維持・管理および飼い主とのコミュニケーション促進	高齢者のQOL向上 医療費の削減 犬猫殺処分の低減
	介護・見守り	活動量計を搭載したウェアラブルデバイスによる高齢者、乳幼児の介護・見守り支援および介護施設や保育施設における安心・安全の向上	女性の社会進出の促進 介護サービスの高付加価値化
	スマートハウス	ネットワーク家電や仮定内センサを活用した防犯、エコ・省エネ、火事支援。ヘルスケア、ペットケア、介護・見守りサービスとの連携	女性の社会進出の促進 省エネ・省電力化の促進 高齢核家族世帯におけるQOL向上

4-7. 適用先の具体的例 - 2 社会インフラ

分野	対象製品 サービス	適用ケース	効果・メリット例
社会インフラ	災害・減災	災害時におけるネットワーク通信手段の確保。被災状況、避難誘導情報の伝達。センサーを活用した災害予測の実現	災害時における損傷被害の軽減およびライフラインの強靭化
	ITS	車車間、車路間におけるセンサーレスプロブの実現による交通情報および有用情報の共有	交通渋滞の緩和 交通事故の低減
	FEMS/BEMS	プラント・工場・事業施設におけるデバイスやセンサを活用したインフラ管理、遠隔操作による自動化・省力化の実現、事前察知による自己防止	企業の設備・インフラの効率的運用促進、安全対策および人件費抑制
	スマートシティ	コミュニティ内およびコミュニティ間における総合的なサービスの高度化およびインフラ（道路・建物・ライフライン）の強靭化	地域社会の道路や建物をはじめとするインフラの再構築・維持コストを低減

END

株式会社 Skeed
慶應義塾大学大学院システムデザイン・マネジメント研究科