PostgreSQL の災害対策

災害への備えは万全に! ログの順序性保証で確実なレプリケーション

－富士通の技術者に聞く！PostgreSQL の技術－

2018 年は都市部の局所的な集中豪雨や台風被害、大地震の発生数増加などが大きな問題となり、災害に備える必要性が再認識されました。企業活動における災害対策の中でも特に重要なのが、生命線とも言える「データ」の保全です。データは一度失われると二度と復旧できません。このためデータベースの災害対策が急務と考えるシステム管理者も多いと思います。
ここでは PostgreSQL における災害対策について富士通のポリシーや考え方、実際の取り組みについて開発担当である谷口 和博に聞きました。

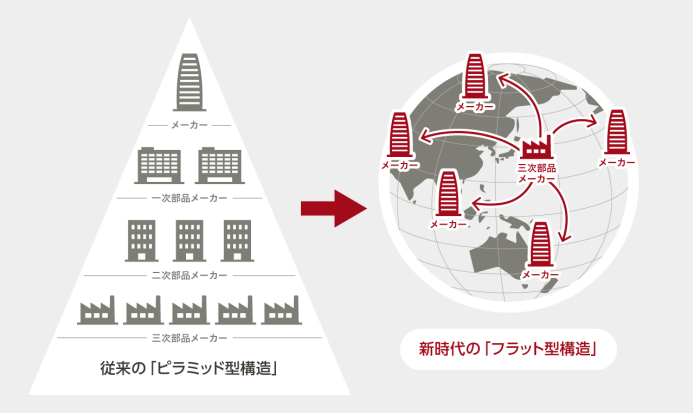
谷口 和博 Kazuhiro Taniguchi
富士通株式会社　ミドルウェア事業本部 データマネジメント・ミドルウェア事業部
専門分野：データベース
入社以来、データベース管理システムの開発・保守に携わり、主に高性能・高信頼なミッションクリティカル機能を中心に担当。4 年ほど前から PostgreSQL をベースとする FUJITSU Software Enterprise Postgres の開発に参画。

災害対策の重要性

近年の災害を見ると昔より被害規模の大きいケースが増えていると感じられます。このため、データベースへの災害対策はますます重要になるのではないでしょうか。

谷口

そうですね。以前は法令に従って災害対策をシステムとして備えていることが重要でしたが、実際に発動しなければならないような状況が増えていると感じます。これは近年、災害が多い国内に限った話だけでなく、海外でも同様に地球規模での気候変動やテロなどのリスクが背景にあるかと思います。また、企業コンプライアンスの観点でも必要性は高まりつつあります。万が一、データセンターが被災して大切なデータを失ってしまった場合、自社の信用を低下させるだけでなく、ステークホルダーにも損害を与えてしまうことが考えられます。
例えば産業界では、グローバル化や新興国の台頭、インターネットの普及などによって産業構造そのものが変化しています。具体的には、従来のメーカー、一次部品メーカー、二次部品メーカー、三次部品メーカーと積み重なるような「ピラミッド型構造」から、技術力のある三次部品メーカーなどが自社の製品や部品を国内外のメーカー各社へ直接かつタイムリーに納品するような「フラット型構造」に変化しています。しかし、このような構造は特定企業の業務停止が他の企業の業務停止を連鎖的に拡大させるといったリスクを抱えています。このため企業の基幹システムには 24 時間 365 日の安定稼働が求められるため、今後ますます災害対策のニーズが増加するものと推測しています。



商用データベースでは災害対策に向けたさまざまな機能が実装されていますが、オープンソースデータベースにおける取り組みはいかがでしょうか？

谷口

一口に災害対策と言っても、お客様の BCP（Business continuity planning：事業継続計画）によってシステムに求められる要件が異なります。具体的には、RPO（Recovery Point Objective）や RTO（Recovery Time Objective）に対して、どの仕組みが最適なのか？を選択することになります。代表的なオープンソースデータベースである PostgreSQL では、定期的なバックアップとレプリケーションによる災害対策の仕組みが提供されています。

用語解説

BCP	企業が災害やテロ攻撃などの緊急事態に遭遇した場合に、速やかに事業を復旧させつつ損害を最小限にするために、平常時の活動や緊急時の活動および事業復旧手段などの段取りを決めておく計画
RPO	過去のどの時点までのデータを保障して復旧させるかという目標値
RTO	被災時点からどれだけの時間で業務を復旧させるかという目標値

なるほど。しかし最近では、オープンソースデータベースを基幹系システムに採用する事例が増えていますが、災害対策の仕組みは十分なのか？と不安を感じているお客様もいらっしゃるのではないでしょうか。

谷口

確かに、オープンソースの世界では、国内で求められるような災害対策に着目した取り組みはあまり見受けられません。これは PostgreSQL がご存じのとおりグローバルなコミュニティーの中で成長しており、国内と海外で災害対策に対するニーズが異なるという一面もあります。一部、ヨーロッパでは洪水のリスク対策としてのニーズはあるようですが、国内ほどではないとの認識です。 このような観点から FUJITSU Software Enterprise Postgres（以降、Enterprise Postgres と略します）では、PostgreSQL の持つ災害対策に必要な機能の強化に取り組んでいます。加えて、自然災害大国である日本において、Enterprise Postgres で実績を積んだ結果をコミュニティーにフィードバックすることが私たちの役割と考えています。

Enterprise Postgres における強化ポイント

グローバルなコミュニティーにとって重要な役割ですね。Enterprise Postgres における具体的な取り組みについて教えてください。

谷口

先ほど説明したとおり PostgreSQL でも、定期的なバックアップとレプリケーションによる災害対策の仕組みが提供されています。しかし、基幹系システムに適用するには機能が足りないと考え、Enterprise Postgres の最新版であるバージョン 10 では、レプリケーションに付加価値を提供しています。

レプリケーションへの付加価値ですか？もう少し詳しく教えてください。

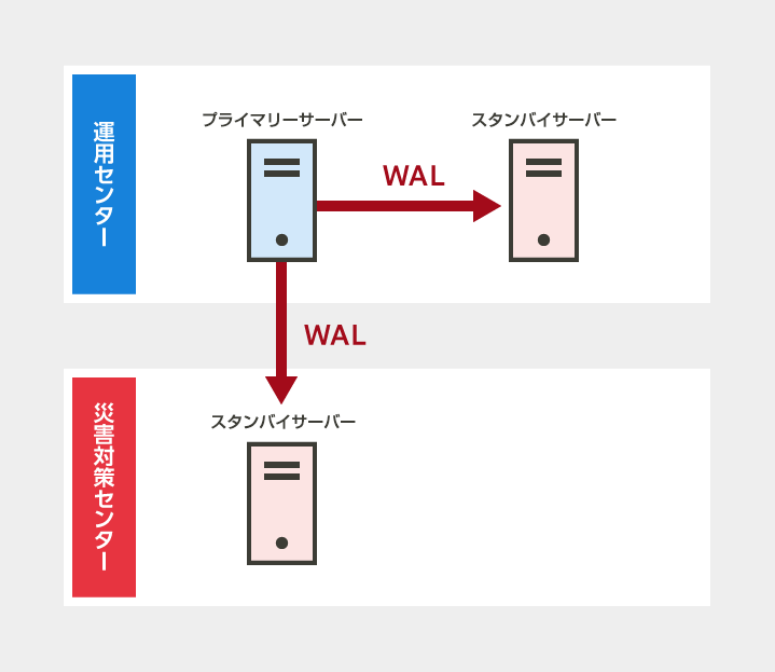
谷口

データベースに災害対策を必要とするようなシステムにおいては、必ずと言ってよいほど運用センター内でのシステム冗長化が求められます。この場合、運用センター内でデータベースを多重化して可用性を高めたうえで、レプリケーションを使用して遠隔地にバックアップを行うのが一般的な運用スタイルになります。しかし、このような運用スタイルでは、運用センター内で障害が発生してフェイルオーバーが発生すると、影響が伝播して災害対策センターの復旧が必要になるケースがあります。この復旧中に運用センターが被災すると、せっかく設定した RTO や RPO の目標を達成できなくなってしまいます。また、災害対策センターを活用する目的で参照業務を稼働させている場合には、災害対策センター復旧のために業務が停止してしまうというリスクがあります。

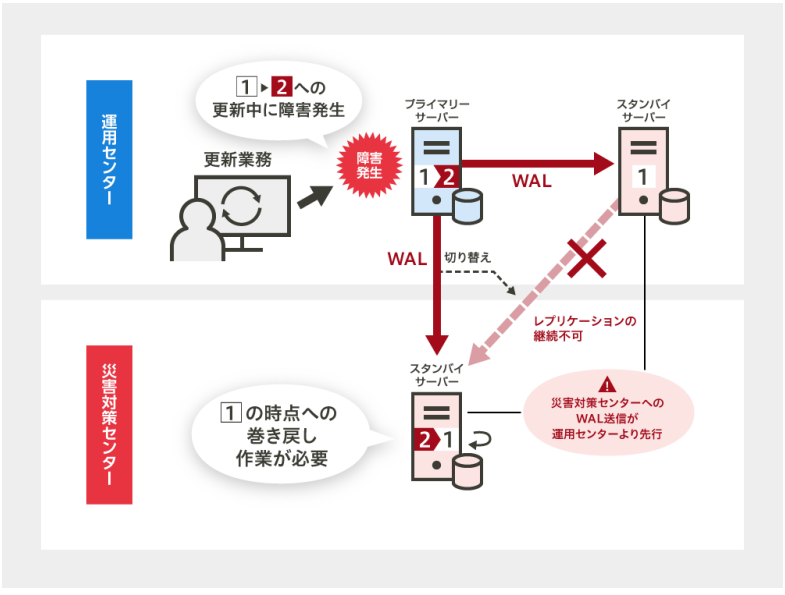
運用センターで発生した障害が災害対策センターに影響を及ぼす可能性があるということですね。なぜそのような事象が発生するのでしょうか？

谷口

PostgreSQL に搭載されているストリーミングレプリケーションで災害対策を行う場合、運用センター内のサーバーを冗長化したうえで災害対策センターのサーバーとも冗長化することで実現します。要は、運用センターのプライマリーサーバーから運用センターのスタンバイサーバーと災害対策センターのスタンバイサーバーに対して同時にトランザクション更新ログ（WAL：Write Ahead Logging）を送信するアーキテクチャーです。



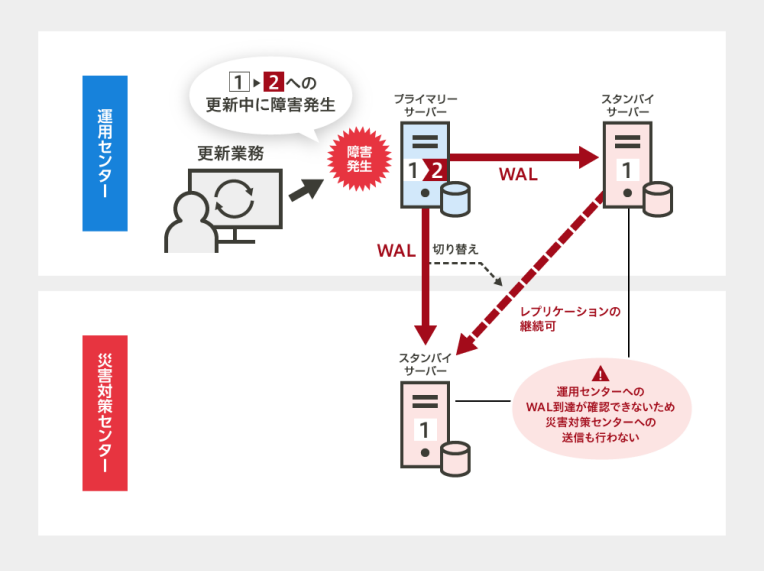
このアーキテクチャーでは、2 つのスタンバイサーバーに対して WAL の送信が順序性も無く行われることから、運用センターのスタンバイサーバーよりも災害対策センターのスタンバイサーバーへの送信が先行する可能性があります。このとき、運用センターのプライマリーサーバーで障害が発生すると、待機していたスタンバイサーバーが災害対策センターとのレプリケーションを継続しようと試みますが、WAL の連続性が損なわれていることが原因で、レプリケーションを継続することができなくなります。その結果、災害対策センターでデータベースの巻き戻しが必要となります。



タイミングによって発生する可能性があるということですね。どのようにして解決したのでしょうか？

谷口

Enterprise Postgres で提供しているレプリケーションは、PostgreSQL に搭載されているストリーミングレプリケーションがベースになっています。ここに対し、富士通の独自技術によって、運用センター内でのレプリケーションと、運用センターから災害対策センターへのレプリケーションを互いに制御することによって解決しています。具体的には、運用センターのスタンバイサーバーに対する WAL の送信と書き込みが完了したことを確認してから、災害対策センターのスタンバイサーバーへ WAL を送信して書き込みを行うことで同期制御します。これにより、災害対策センターのスタンバイサーバーを停止してデータベースの巻き戻しを行う必要がなくなります。



富士通ならではの“こだわり”とは

ここまでのお話を伺っている限りこのような事象が発生することは稀のような気がしますが、何故ここまで考慮するのでしょうか？

谷口

確かに、一般的に考えて運用センター内同士における近距離通信と、遠隔地に設置されている災害対策センターへの遠距離通信についての回線性能を考慮すると、災害対策センターへの WAL 送信が先行することは極めて稀だと考えています。しかし、サーバーや通信回線が高負荷になることで起きるラグや運用センター内における回線切断など、本事象が発生する可能性が理論上有り得ないと言い切ることは絶対できません。
そのため、当社では例え稀なケースであっても理論上発生しないようにする仕組みを実装していることこそ、システムの高信頼化やお客様への安心感に繋がり、ゆるぎない「あんしん」を提供できると考えています。

お客様が感じる「あんしん」を追求しているのですね。開発にあたって苦労した点はなんでしょうか？

谷口

この仕組みは、拡張モジュールなどの後付方式で実装することが不可能であり、コミュニティーのコアモジュールに実装しなければなりませんでした。
富士通としては PostgreSQL をエンタープライズ市場向けに強化することが役割と考えており、これらの高信頼化機能もオープン化する必要があると考えています。そのため、当社に在職している PostgreSQL コントリビューターと協調しながらコアモジュールへの適用性を持たせたプログラムというものを意識した開発を心掛けました。

常にオープン化を意識した開発を心掛けているのですね。最後に今後の取り組みについて教えてください。

谷口

今後は、災害対策センター側の信頼性向上にも取り組んでいきます。
現状、定常運用中でも災害対策センター内においてストリーミングレプリケーションによる冗長化はできますが、災害対策センター内で障害が発生すると自動フェイルオーバーすることが難しいため、災害対策センターの可用性という点においては運用センターより劣っている状態です。これを、定常運用中に災害対策センター内で障害が起きても自動フェイルオーバーできる仕組みを実装することで、更に信頼性が向上できるようになります。

基幹系に求められる信頼性がますます強化されそうですね。PostgreSQL の今後が楽しみです。ありがとうございました。

2019 年 2 月 4 日