

規子と哲樹の、おしえて！ミッションクリティカル パーティショニング技術—上手に分割していただきたく(1/2)

前回は「うどん店を繁盛させるため、いかに厨房を効率よく回せばいいか」について、うどん店の経営に詳しい日立の長江さんにテデーしてもらいました…じゃなくて、ミッションクリティカルなデータベースシステムにおける並列処理「パラレルサーバ」について、うどん店のたとえを挙げながら紹介いたしました。「うどん屋のたとえ、思った以上に使える…！」との手ごたえを得た我々は、今回も前回に引き続き、ミッションクリティカルなうどん店におけるデータベースのゆで方やパーティショニングなどについてテデーしてもらったのであります。もはやデータベースの話をしているのか、うどん店の話をしているのか、あいまいになってきて……。



* この記事は、翔泳社が運営するEnterpriseZine に、2015年2月10日 に掲載された記事より転載しています。

ミッションクリティカルに欠かせないデータベース技術「パーティショニング」

吉村 さて、今月も始まりました、「規子のうどんでGO！」

長江 そんな連載ではありません。

吉村 えっ、じゃあ、今まで一体何の話をしてきたんですか……。

長江 今までもこれから、ミッションクリティカルなデータベースについて話しています。

吉村 拝承。

長江 ちなみに今回のテーマは「パーティショニング」です。

吉村 うどんのたとえで説明していただきたく。あと、そろそろ新しい日立用語をテデーしていただきたく。

長江 それはもうちょっと待っていただきたく。パーティショニング、何となく分かりますよね？

吉村 テーブルを分割するってやつですよ。何となくは知ってるつもりなんですけど、実は詳しくは知らなかったりします。

長江 大規模システムやミッションクリティカルなシステムの設計・開発に携わったことがない方だと、パーティショニングという言葉自体は聞いたことがあっても、実際に触る機会はありません。実際、一般的なリレーショナルデータベース製品では、パーティショニングの機能はオプション扱いになっていることが多いですね。

吉村 確かに、パーティショニングの機能を上手に扱えるデータベース技術者は「上級者」というイメージがあります。

長江 そうですよ。でもHiRDBのように、そもそもミッションクリティカル用途を前提に作られているデータベース製品では、パーティショニングって割と普通に使われているんですよ。

吉村 そもそも、なぜパーティショニングはミッションクリティカルなシステムで多く使われているんでしょう？

長江 第1回でも説明したように、ミッションクリティカルなシステムで使われるデータベースは、高負荷トランザクションに耐えられなくてははいけません。パーティショニングもそのための機能の1つなんです。具体的には、表を複数に分割することで、ある特定のエリアのデータに対するアクセスが集中してボトルネックになるような事態を未然に防いだり、ある特定のエリアの障害やメンテナンスがシステム全体に及ぼす影響を最小化するための技術です。

吉村 ちなみにパーティションって、パーティ・ションっていう人が発明したって知ってました？ 嘘ですけど。

長江 ……前回、データの処理を複数のサーバに分割して並列処理させることで、高負荷トランザクションに対応できることを説明しましたよね。でも、サーバをいくらたくさん並べて並列処理しても、それらが皆同じ巨大なテーブルにアクセスするのは、アクセス競合が頻発してボトルネックになってしまいます。これを回避するには、テーブルを複数のパーティションに分割して、それぞれでアクセスを均等化してやればいいわけです。

吉村 なるほど。うどん店のたとえで言うなら、うどんの長さをぴったり同じにそろえれば商売繁盛間違いなしということですね！

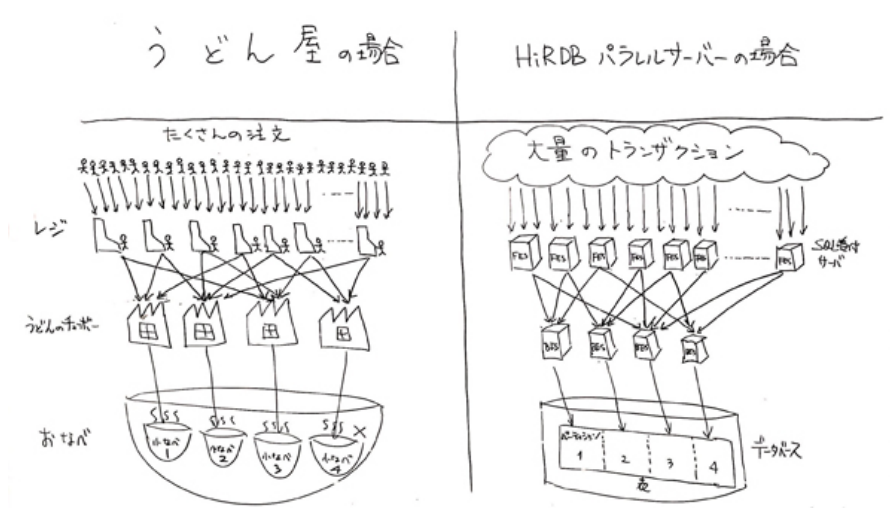
長江 全然違います。強いて言えば、うどんを茹でる鍋を複数用意するというイメージでしょうか。厨房でうどんを茹でる料理人をいくらたくさんそろえても、鍋が大鍋1個だけだと明らかに非効率ですよ。

吉村 確かに。注文はばらばらに入ってきますし、茹でる人も複数いるので、大鍋1個だけだと茹でる量の調整が難しそうですし、作業の待ち時間も発生して無駄が多そうです。

長江 その点、小さめの鍋を複数並べて同時並行で茹でた場合は、複数いる調理人が常に無駄なく稼働できますし、ランダムに入ってくる注文にも柔軟に対応できます。それに、もし何らかの事情で鍋を交換しなくてはなくなった場合でも、大鍋を丸ごと交換となるとその間オーダーを完全にストップしなくてははいけません。小鍋に分けていればほかの鍋で茹でる間にさっと交換できます。空の鍋を常に用意しておけば、鍋に問題が起きたときや、茹汁を交換したいときなどにすぐ交換できますし、何なら予備の鍋を常に火にかけて準備しておけば、ほぼ時間のロスなく交換できます。

吉村 それって、まさに「コールドスタンバイ」と「ホットスタンバイ」ですね！

長江 その通りです。パーティショニングは、大規模データベースの運用性や可用性を担保する上でとても有効な技術なんです。



規子と哲樹の、おしえて！ミッションクリティカル パーティショニング技術—上手に分割していただきたく(2/2)

パーティショニングの効果は「データの分け方」によって大きく左右される

吉村 ああ、何となく分かってきましたよ、パーティショニング！ じゃあミッションクリティカルなシステムでは、皆パーティショニングをがまん使ってハッピーになろうぜ的な？

長江 そう簡単にはいかないところが、ミッションクリティカルの難しいところです。実は、パーティショニングの設計には結構コツがあるんです。中でも一番重要なのが、「データをどう分割すればいいか」ですね。

吉村 逆に言うと、いろんな分割の仕方があるということですか？

長江 その通りです。しかし、データをどう分割すればいいか分からないので、「とりあえず均等配置しておけばいいんじゃないか？」と「ハッシュ分割」を選んでしまうケースが多いです。複数のパーティション間でアクセスがなるべく均等化されるよう、データ格納先のパーティションをシステムが自動的に決定する方式です。

吉村 ー、とりあえずハッシュ分割しておけばいいんじゃないか？

長江 いくないんです。特にミッションクリティカルでは、何も考えずにハッシュ分割で均等配置するだけでは、場合によっては逆効果になりかねないんです。

吉村 (笑)

長江 そこ、笑うところじゃないです。

吉村 すみません。そのあたり、もう少し詳しく教えていただけますか。

長江 具体的には、アプリケーションが扱う業務の特性や、データベースで管理するデータの特性、さらにはシステム構成なども加味した上で、ハッシュ分割以外の分割法もうまく活用する必要があります。例えば履歴データなどは「年度ごと」「月ごと」といったように日付の範囲ごとに分割しておくとか何かと便利です。あるいはアプリケーションの特性によっては、取引データなどを地域別のパーティションに分けておいた方が処理がスムーズに運ぶケースもあるでしょう。この場合は「キーレンジ分割」（地域名など不連続な値をグループ化する場合は「リスト分割」と区別して呼ぶデータベース製品もあります）という手法をとります。いずれにせよ、データそのものの特性や、そのデータをどう使うかによっては、ハッシュ分割ではなく、あらかじめデータベース設計者が明示的にデータの分割方法を指定した方がいいケースも多いんです。

吉村 ーん、いまいちイメージが……。

長江 さっきのうどん店のたとえで言えば、「鍋をどう分割するか？」ということですね。例えば、讃岐うどんだけを出すお店があったとしましょう。厨房では同じ麺しか茹でませんから、同じ種類の鍋を複数用意しておいて、それらがすべて均等に稼働するように厨房で仕事を機械的に割り振ればいわけです。

吉村 機械的に。

長江 一方、太麺もあれば細麺もある、さらにはきし麺や蕎麦まで出しているお店ではどうでしょう？ 同様に、同じ鍋を並べて仕事を機械的に割り振ってもいいですけど、うどんや蕎麦の種類ごとに茹でる時間や手順は異なりますから、「太麺専用の鍋」「細麺専用の鍋」「きし麺専用の鍋」とそれぞれ分けた方が効率がいいですよ。

吉村 確かに、その方が仕事ははかどりそうです。

長江 それに、もし誤って鍋を引っくり返しちゃって交換するのに時間がかかりそうな場合でも、それが太麺の鍋なら「ちょっとの間、太麺だけオーダーをストップして！」とオーダー係に指示できますよね。

吉村 「太麺ストップして」。

長江 これが、どの鍋も同じ仕事をしているとなると、どのメニューの調理にどれだけの影響が出るか分からないので、厨房もホールも混乱してしまいます。データベースも同じことで、ハッシュ分割で機械的にデータをパーティションに割り振っている場合には、ある特定のパーティションに障害が発生したり、メンテナンスのために一時停止させる際、どのサービスにどれだけの範囲の影響が及ぶか、明示しにくいんです。これでは、ミッションクリティカルシステムにふさわしいサービスレベルは維持できませんよね。

吉村 そのあたり、「人間が極限まで知恵を絞って」とか「あらゆる英知を終結させて」のあたりをもうちょっとkwskお聞かせいただきたく！確かに、「太麺にどれだけかかるか分からない」と言われるより、「いまキシ麺はストップしてるけど、何分後にはできるようになります」と明確に説明してくれた方が、お客さんの納得度は高いかも。太麺だけに。

長江 ええ。高いレベルの接客が求められる高級店では、そこは重要ですよ。高いサービスレベルを担保する必要があるミッションクリティカルなシステムでも、同様なんです。何も考えずにハッシュ分割でパーティショニングするのではなく、データやアプリケーションの特性を十分に理解し、さらに可用性や運用性も加味した上で適切なデータ分割を行う必要があるというわけです。

吉村 なかなか奥が深いんですね……ちなみに、具体的にどういうデータ特性の場合には、どんなデータ分割を行えばいいんですか？

長江 これは技術的にかなり深いテーマなので、この連載ではちょっとカバーしきれないですね。なので、興味のある方は、日立がホームページ上で公開している技術文書をぜひ参照していただきたく。

吉村 いただきたく！



特記事項

この記事は、翔泳社が運営するEnterpriseZine に、2015年2月10日 に掲載された記事より転載しています。

記載の会社名、製品名は、それぞれの会社の商標または登録商標です。

記載の仕様は、製品の改良などのため予告なく変更することがあります。