

# 高速化だけじゃない! 最適化による性能改善

ご購入はこちら

松尾 玲央馬

## 3-1 最適化の基礎

### Q1 最適化の定義

Q 性能を最適化する目的は？

A レイテンシの短縮/スループットの向上/リソース効率の改善を目指す

性能最適化の目的は、レイテンシ、スループット、リソース効率の3つの軸での改善を目指すことです。アプリケーションの用途に応じて優先順位を明確にすることが、効果的な最適化への第一歩です。

性能最適化というとプログラムを速くすることと考えがちですが、実際にはもっと多面的な取り組みです。現代のコンピュータ・システムでは、単純な実行速度だけでなく、次の3つの観点から性能を評価し、改善していく必要があります。

### ● 3つの最適化の軸

プログラムの性能最適化には大きく3つの方向性があります。それぞれの特徴と、どのような場面で重要になるかを説明します。

#### ▶レイテンシの短縮

レイテンシとは、図1に示すように処理を開始してから結果が得られるまでの時間を指します。例えば、オンライン・ゲームでボタンを押してからキャラクタが動くまでの時間や、ウェブ・アプリケーションでクリックしてから画面が切り替わるまでの時間がこれに当たります。

#### ▶スループットの向上

スループットとは、図2に示すように単位時間当た

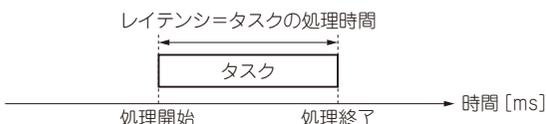


図1 レイテンシは処理を開始してから結果が得られるまでの時間

りの処理量を指します。ウェブ・サーバであれば1秒間に処理できるリクエスト数、動画変換ソフトウェアであれば1分間に処理できるフレーム数がスループットの指標となります。バッチ処理システムやデータ分析システムでは、個々の処理の速さよりも、全体として多くのデータを処理できることが重要です。

#### ▶リソース効率の改善

リソース効率とは、同じ処理をより少ないリソースで実行すること、または同じリソースでより多くの処理を実行することを指します。CPU使用率やメモリ使用量はもちろん、ディスクI/O、ネットワーク帯域といった資源を効率的に使うことも含まれます。リソース効率を高めることで、サーバ・コストの削減や、同じハードウェアでより多くの処理を並行実行できるといったメリットにつながります。

### ● 3つはトレードオフの関係

これら3つの軸は、しばしばトレードオフの関係にあります。例えば、ウェブ・サーバでキャッシュを多用してレイテンシを短縮すると、メモリ使用量は増加しますが、CPU使用率は削減され、全体的なリソース効率は用途によって向上する場合があります。

このように、1つの指標を改善すると、別の指標が悪化することがよくあります。性能最適化を成功させるには、アプリケーションの用途とユーザの期待を明確にし、3つの指標の優先順位を決定することが重要です。

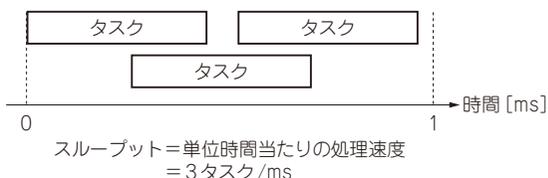


図2 スループットは単位時間あたりの処理量