

А. Т. Лисовский
(УО «ГрГУ им. Я. Купалы», Гродно)

ТЕОРЕМА CAP В РАСПРЕДЕЛЕННЫХ ХРАНИЛИЩАХ ДАННЫХ

Распределенные хранилища данных наиболее широко знакомы как распределенные базы данных. Большинство распределенных баз данных – это не реляционные базы данных NoSQL, ограниченные семантикой ключ-значение. Они обеспечивают невероятную производительность и масштабируемость за счет консистентности или доступности данных. Распределенное хранилище данных сложно описать без предварительного представления теоремы CAP.

Теорема CAP (известная также как **теорема Брюера**) утверждает, что в любой реализации распределённых баз данных возможно обеспечить не более двух из трёх следующих свойств:

согласованность данных (англ. *consistency*) – во всех вычислительных узлах в один момент времени данные не противоречат друг другу;

доступность (англ. *availability*) – любой запрос к распределённому хранилищу завершается корректным откликом, однако без гарантии, что ответы всех узлов системы совпадают;

устойчивость к разделению (англ. *partition tolerance*) – расщепление системы на несколько изолированных секций не приводит к некорректности отклика от каждой из секций [1].

Практика показывает, что большинство приложений больше ценят доступность. Приложениям не всегда нужна сильная согласованность данных, а также доступность может стать причиной несвоевременной синхронизации данных из-за задержек сети, что может привести к несогласованности данных. Примером распределенной базы данных является Cassandra. Она является распределенной NoSQL базой данных, которая предпочитает свойства AP из теоремы CAP.

Литература

1 Brewer, E. A. A Certain Freedom: Thoughts on the CAP Theorem / E. A. Brewer // Proceeding of the XXIX ACM SIGACT-SIGOPS symposium on Principles of distributed computing. – N. Y. : ACM, 2010. – Vol. 29, № 1. – P. 335-336.