Computer Developers of “cloud computing” systems face a conundrum: To achieve elasticity, today's applications must abandon consistency. In effect, the developer is told to assume that scalability entails running on stale data, tolerating incorrect answers, and offering few if any guarantees. For example, cloud database systems often weaken serializability and durability. The underlying principle has been standardized as a design maxim (“CAP”), which posits an inevitable weakening of consistency in applications that desire availability and partition tolerance), and a theoretical argument that supports it.

If CAP is correct, how will we ever exploit cloud computing in settings that have sensitive data and where correct actions are important, for example in domains like health care, finance, power grid control, etc?

But what if CAP is simply wrong? This talk will look at the technical underpinnings of CAP, focusing on the basic assumptions made in cloud computing environments. We’ll see that CAP may actually reflect a simple but faulty assumption. This, in turn, leads to the realization that some styles of strongly consistent applications may actually be more scalable and faster than their weakly consistent counterparts. The trick is to leverage an idea from web data mining to work around what otherwise looks like a hard (NP-complete) problem.