CLOUD COMPUTING CONCEPTS

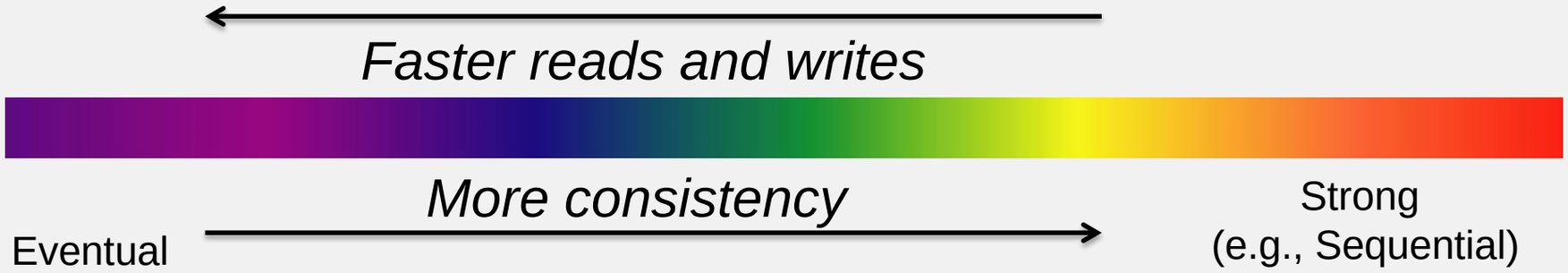
with Indranil Gupta (Indy)

KEY-VALUE STORES NoSQL

Lecture D

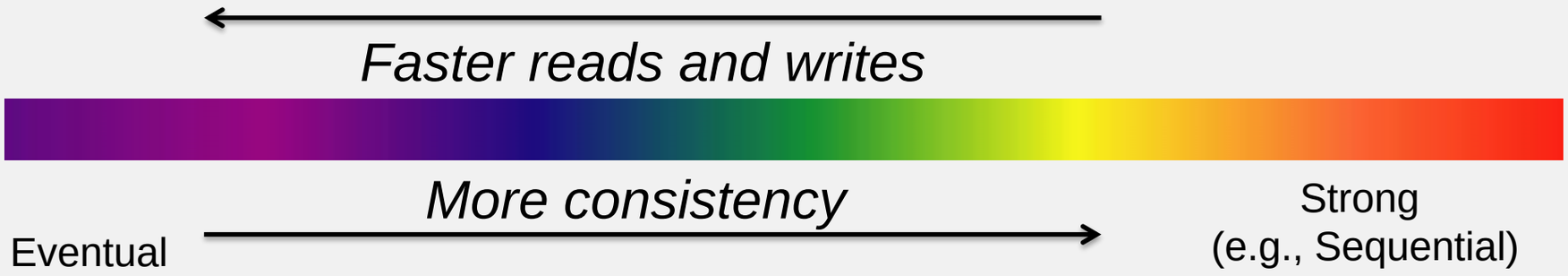
THE CONSISTENCY SPECTRUM

Consistency Spectrum



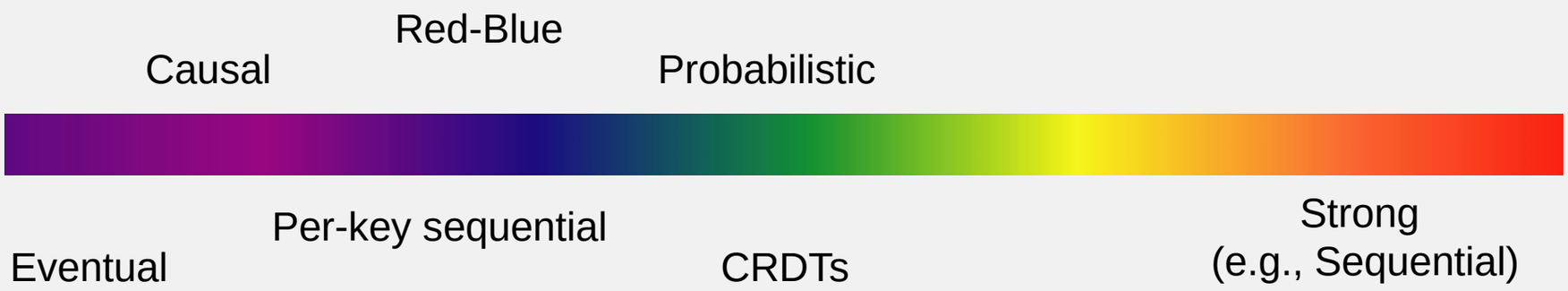
Consistency Spectrum

- Cassandra offers eventual consistency
  - If writes to a key stop, all replicas of key will converge
  - Originally from Amazon's Dynamo and LinkedIn's Voldemort systems



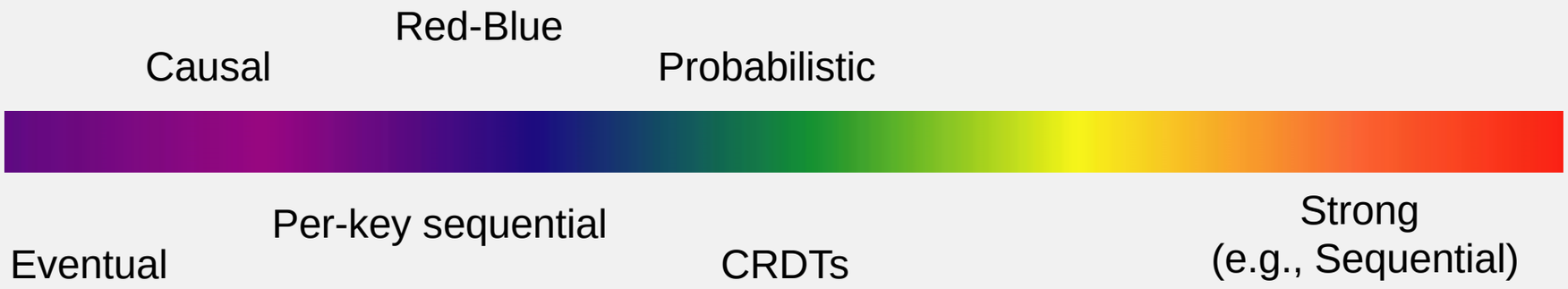
Newer Consistency Models

- Striving towards strong consistency
- While still trying to maintain high availability and partition-tolerance



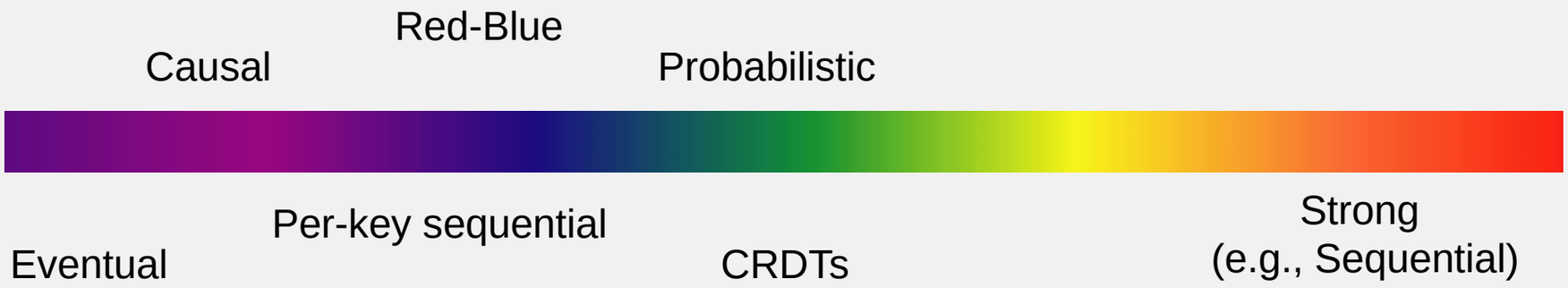
Newer Consistency Models (Contd.)

- **Per-key sequential**: Per key, all operations have a global order
- **CRDTs** (Commutative Replicated Data Types): Data structures for which commutated writes give same result [INRIA, France]
  - E.g., value == int, and only op allowed is +1
  - Effectively, servers don't need to worry about consistency



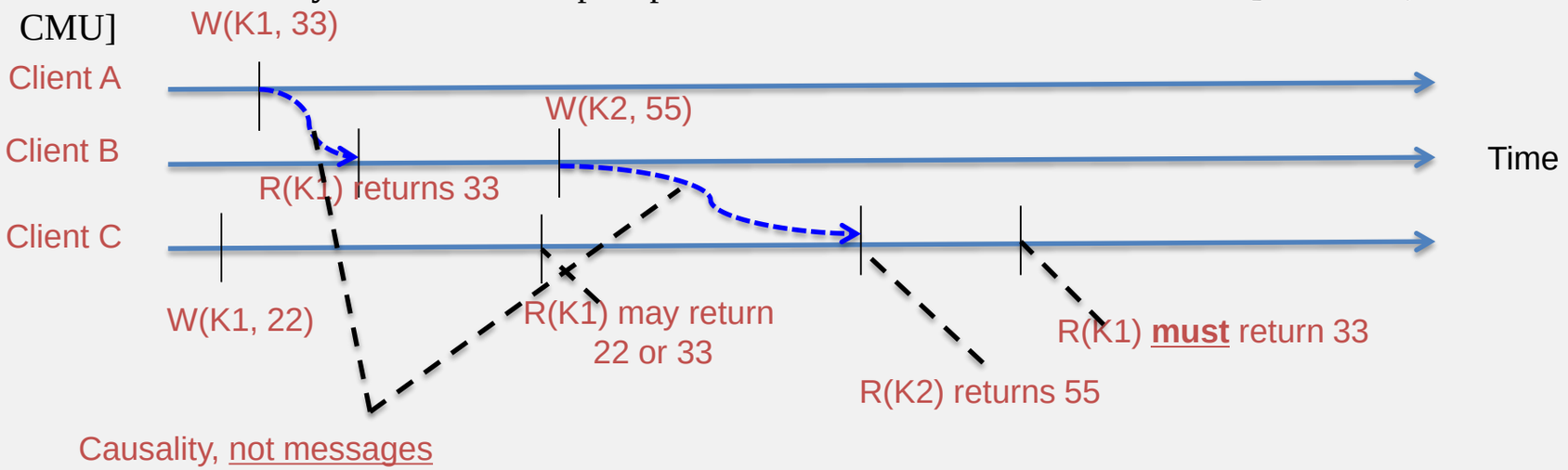
Newer Consistency Models (Contd.)

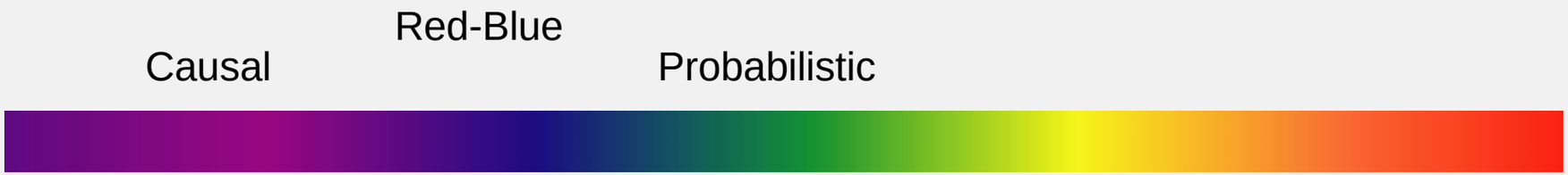
- **Red-blue consistency**: Rewrite client transactions to separate ops into red ops vs. blue ops [MPI-SWS Germany]
  - Blue ops can be executed (commutated) in any order across DCs
  - Red ops need to be executed in the same order at each DC



Newer Consistency Models (Contd.)

Causal Consistency: Reads must respect partial order based on information flow [Princeton, CMU]





Strong Consistency Models

- **Linearizability**: Each operation by a client is visible (or available) instantaneously to all other clients
  - Instantaneously in real time
- **Sequential Consistency** [Lamport]:
  - *"... the result of any execution is the same as if the operations of all the processors were executed in some sequential order, and the operations of each individual processor appear in this sequence in the order specified by its program.*
  - After the fact, find a "reasonable" ordering of the operations (can re-order operations) that obeys sanity (consistency) at all clients, and across clients.
- Transaction ACID properties, e.g., newer key-value/NoSQL stores (sometimes called "NewSQL")
  - Hyperdex [Cornell]
  - Spanner [Google]
  - Transaction chains [Microsoft Research]