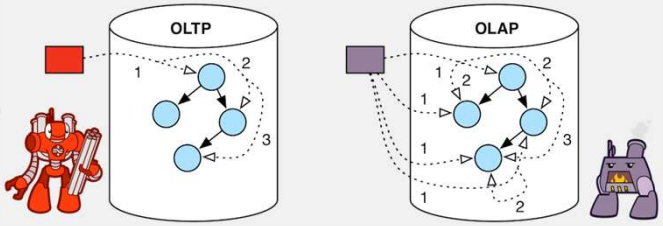
CLOUD COMPUTING APPLICATIONS

Graphs: Databases - Categories

Prof. Reza Farivar

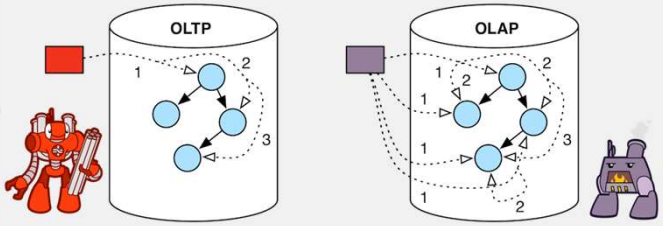
Graph Databases For Transaction Processing

- OLTP-based graph systems allow the user to query the graph in real-time
  - e.g. What soda does Sean buy?
- Typically, real-time performance only possible when a local traversal is enacted
  - A local traversal is one that starts at a particular vertex (or small set of vertices) and touches a small set of connected vertices (by any arbitrary path of arbitrary length).
  - OLTP queries interact with a limited set of data and respond on the order of milliseconds or seconds.
- Neo4J
  - Cypher language
  - graph pattern-match query language Cypher
- TinkerPop and Gremlin
- Amazon Neptune
- Azure CosmosDB



Graph Analytical Processing Systems

- With OLAP graph processing, the entire graph is processed
  - e.g. What is the average price for a soda paid by people like Sean?
  - every vertex and edge is analyzed
    - Possibly more than once for iterative, recursive algorithms
    - Standard or custom algorithms
  - Results are typically not real-time
    - Can take on the order of minutes or hours for massive graphs .
- Bulk Synchronous Parallel (BSP) programming model
  - Pregel -> Giraph -> Spark GraphX -> GraphFrames
  - Gremlin VertexProgram()



Four Graph Communities

- Four famous, yet somewhat disconnected, graph communities:
  - Semantic Web
    - RDF data model
    - SPARQL Query language
      - Declarative
  - Graph database (OLTP)
    - Labeled Property Graph data model
      - where data is organized as nodes, relationships, and properties (data stored on the nodes or relationships).
    - Pattern Matching Graph Traversal / Motif finding
      - Imperative: Gremlin
      - Declarative: Cypher
  - Big Data Graph Processing (OLAP)