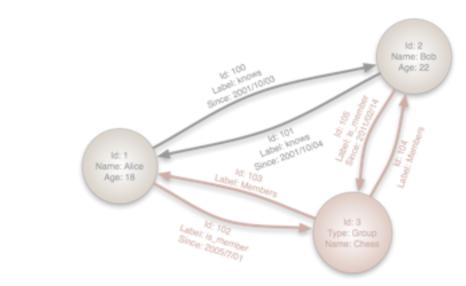
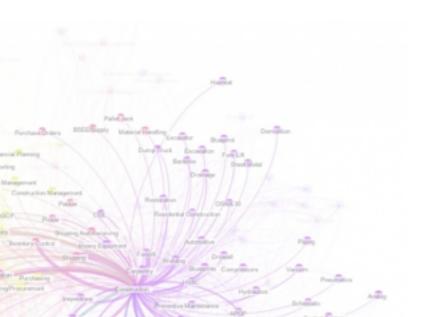


# Intro to Neo4j and Graph Databases

David Montag Neo Technology

david@neotechnology.com



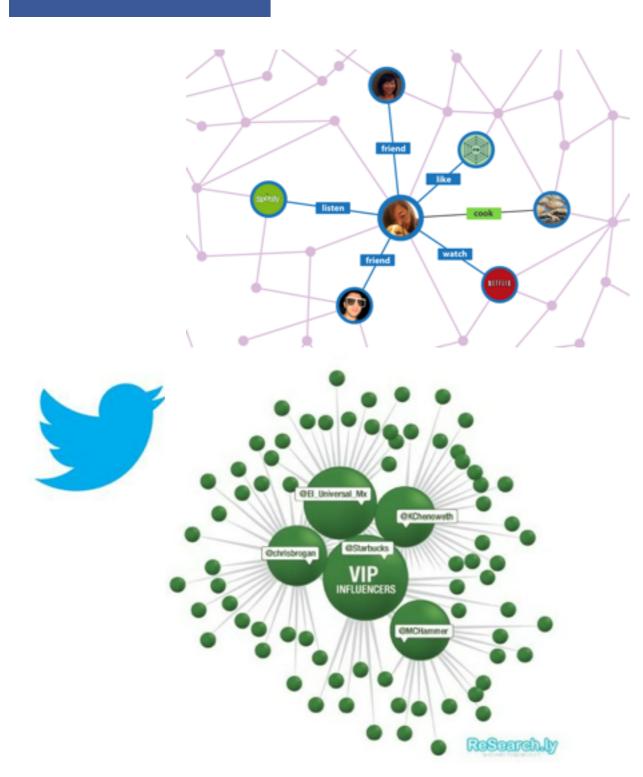






### Early Adopters of Graph Technology

facebook.





### The Knowledge Graph

Learn more about one of the key breakthroughs behind the future of search.

### Survival of the Fittest



#### Evolution of Web Search

#### Pre-1999 WWW Indexing

Index

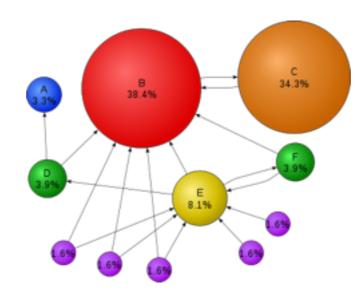
Page numbers in **bold face** refer to key term definitions Page numbers in *thalics* offse to images or diagrams Page numbers followed by a "t" indicate a table

absolute temperature scale, 350-331 sheekste zwo, 351 acceleration of gravity, A 23t accuracy, A.5 aretic acid (CHyCOOH) buffers, 575-576, 581-582 conjugate acid-base pairs, 540 ionization constant, 553, 554; manufacture of, 451 titrations, 590-592 as weak usid, 1440, 145, 551-552 arid bare pairs, conjugate, \$40-544 arid bare martinus, \$38 autoionization of water, 545-547 gus-freming exchange, 150–151 net-ionic equations for, 140–150 neutralization, 146–150, 561–566 of sults, 146–151, 561–566 acid-bare titrations, 527-594 acidic solutions, 546, 650-632, A 328-A 338 acid ionization constant expressions, 550-551acid ionization constants  $(K_{a})$ . Sor ionization co aride (Kg) arideeis, 576 eride, 143. See also arid-base martinese, ionizatio constants, acids (Kg), specific entries, e.g. curbosolic acids, lartie aride Branated Lowey concept, 532-544 buffer solutions, 575-536 ronjagate acid-base pairs, 540-344 equilibrium constants, 405t ionization ronstants, 350-341, A 201 Lawis, 566-562 organic, 544 pH scale, 547-550 properties, 143-145 solubility of salas, 397-598 solutiona, 546, 650-652, A.326-A.338 strengthe, 143-146, 353-556 titrationa, 517-394 water's role, 540 actinides, 55, 230-251 activated complex, 433 activation energies (Eq), 434, 430–640, 443, 647, 449–450 active sites, 449 activities, 466, 547, 703-706 activity point, metal, 159-160 actual yields, 120 addition, 43, A.6, A.9 ublition, significant figures in, A.6

ui: 342-340, 366-370, 380-301, 706 sklyl groups, 70-71 sirohols, 64, 503-507 aldebydes, 278-279 alkali metala, 55, 106 alkaline batteries, 670 ultuline eath metals, 55 ultuline fael cells, 674 alkalosis, 576 silveres, 68-71, 277-278, A.25-A.36 silveres, 280-283, A.26-A.27 sikyi gova, 70-71, A.25-A.26 sikyass, 281, A.27 silotoopes, 23-24, 201, 403-405 ájóha particles, 32–39, **693–**696, 697, 699–700 ájóha radiation, **693** alpha rays, 36-37 slaminum (AJ), 7, 21, 103, 634-635, 682 animes, 544-645 ammonia (NRA) anines, 545 Bestated-Lowey base, 738-539 complex iona, 567 ionization constant, 554, 561 standard molar esthalpy of formation, 2101 structure, 6-R structure, 6+8 synthesis, 107–108, 462, 494–495 V3EPR model, 312–313 ar weak bare, 145, 539 anmonium ions, 74, 77 anophous solids, 396 umpheres (A), 658 unphipentic species, 540 unphotesic metal hydrosides, 567, 602 unpätude, 223 unu (utomic mass unit), 46 analytical chemistry, 114 Anderson, Carl, 696 ungolar geometries, 314 unions, 72, 76, 78, 257-258 unodes, 37, 654. See also electroch unodic inhibition, 683-614 antihonding molecular orbitals, 298 untimetter/antielectrone, 696 aqueour equilitein. Ser also unid-bare titratione, buffers factors affecting solubility, 597-602 precipitation, 603-604 solubility product constant, 394-597 equeous solutions (eq), 100. See also buffer solutions, nobutions electrolysis, 675-678 electrolytes, 82-83, 136, 527-528, 654-655 iosir compounds, 136-139 molarity, 166-162 standard reduction potentials in, A.32t-A.34t stainhiometric relationships in, 167

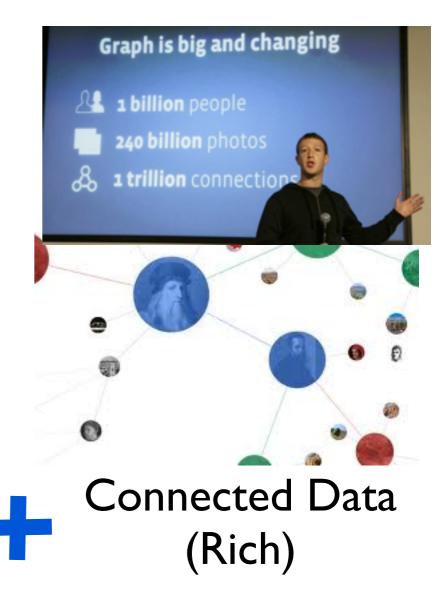
### <u> 1999 - 2012</u>

Google Invents PageRank



<u>2012-?</u>

Google Knowledge Graph, Facebook Graph Search



Discrete Data 🗕

Connected Data (Simple)

#### Survival of the Fittest **Evolution of Online Recruiting**





#### Home > Job Search

Search over 150,000 U.S. jobs. Perform your search below or get tips on searching.

#### Location Search:

----- Select a location ------Alaska-Anchorage Alaska-Fairbanks Alaska-Juneau Alaska-Valdez Alabama-Anniston

#### Category Search:

----- Select a category ------Accounting/Finance/Banking Administrative/Clerical Creative Arts/Media Education/Training Engineering/Architecture/Design

#### Keyword Search:



#### Discrete Data







Invite more friends — ask them to share their connections

#### Jobs with Connections

glassdoor

9

Sr. Statistical Analyst, Product Innovation Netflix - Los Gatos, CA From: Job.com - 1 days ago

Creative Director frog design - San Francisco, CA From: Experteer - 8 days ago

Java Server Software Engineer Electronic Arts - Redwood City, CA From: Experteer - 3 days ago Want better jobs? Tell us your current job title

EMERGENCY MEDICAL TECH U.S. NAVY - Fremont, CA From: Monster - 14 days ago

#### **Connected** Data

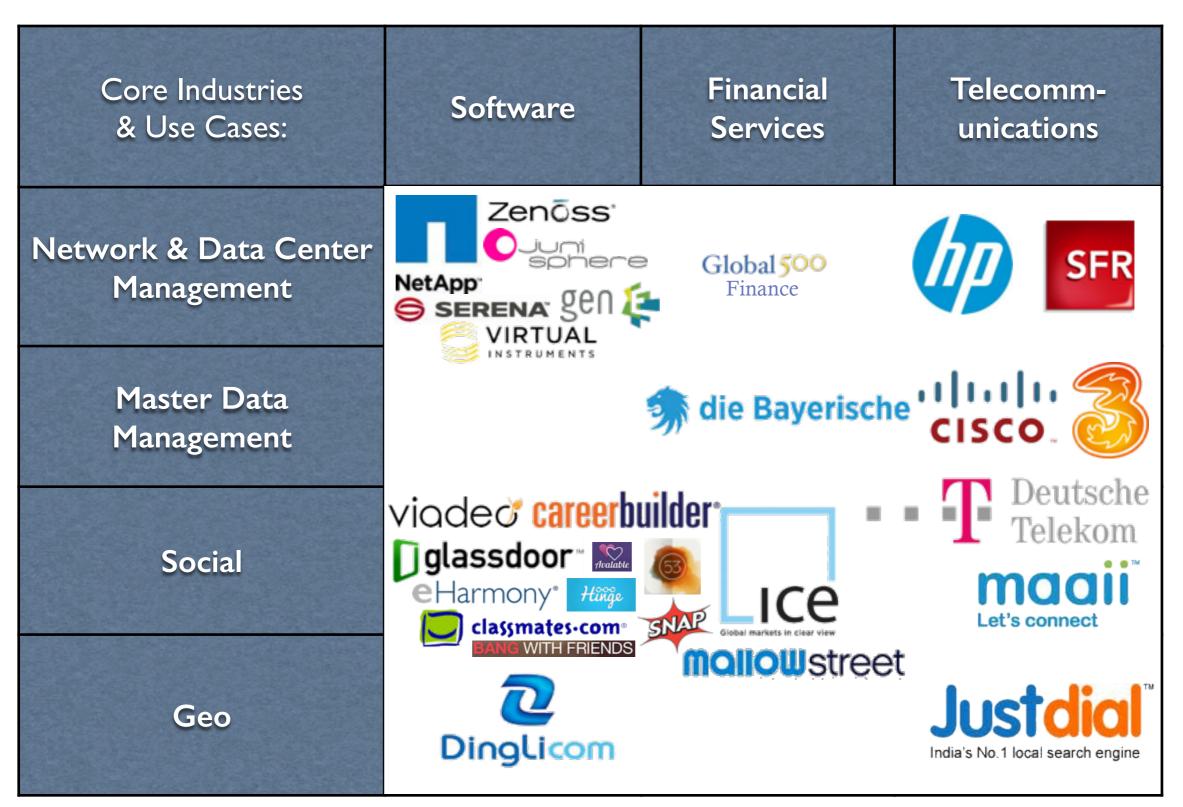
# Early Adopter Segments (What we expected to happen - view from several years ago)

Core Industries & Use Cases:	Software	Financial Services	Telecomm- unications
Network & Data Center Management			
Master Data Management			
Social			
Geo			

### Neo4j Adoption Snapshot

Select Commercial Customers\*





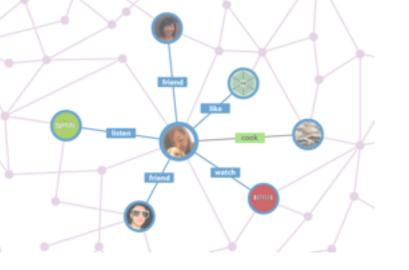
\*Community Users Not Included

Neo Technology, Inc Confidential

### **Neo4j Adoption Snapshot**

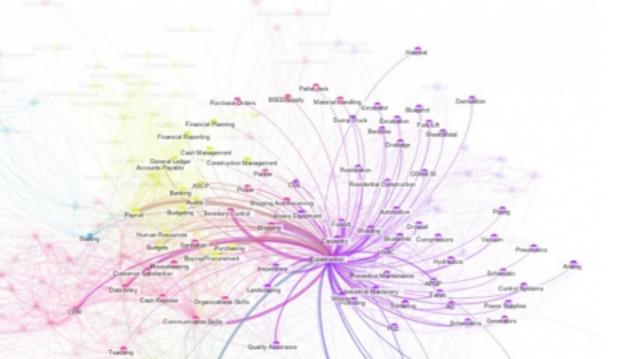


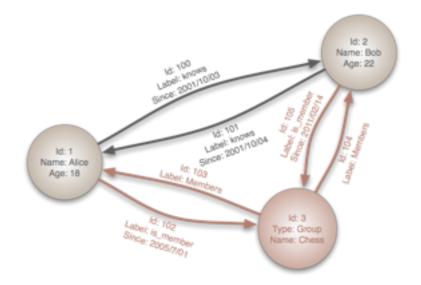
\*Community Users Not Included

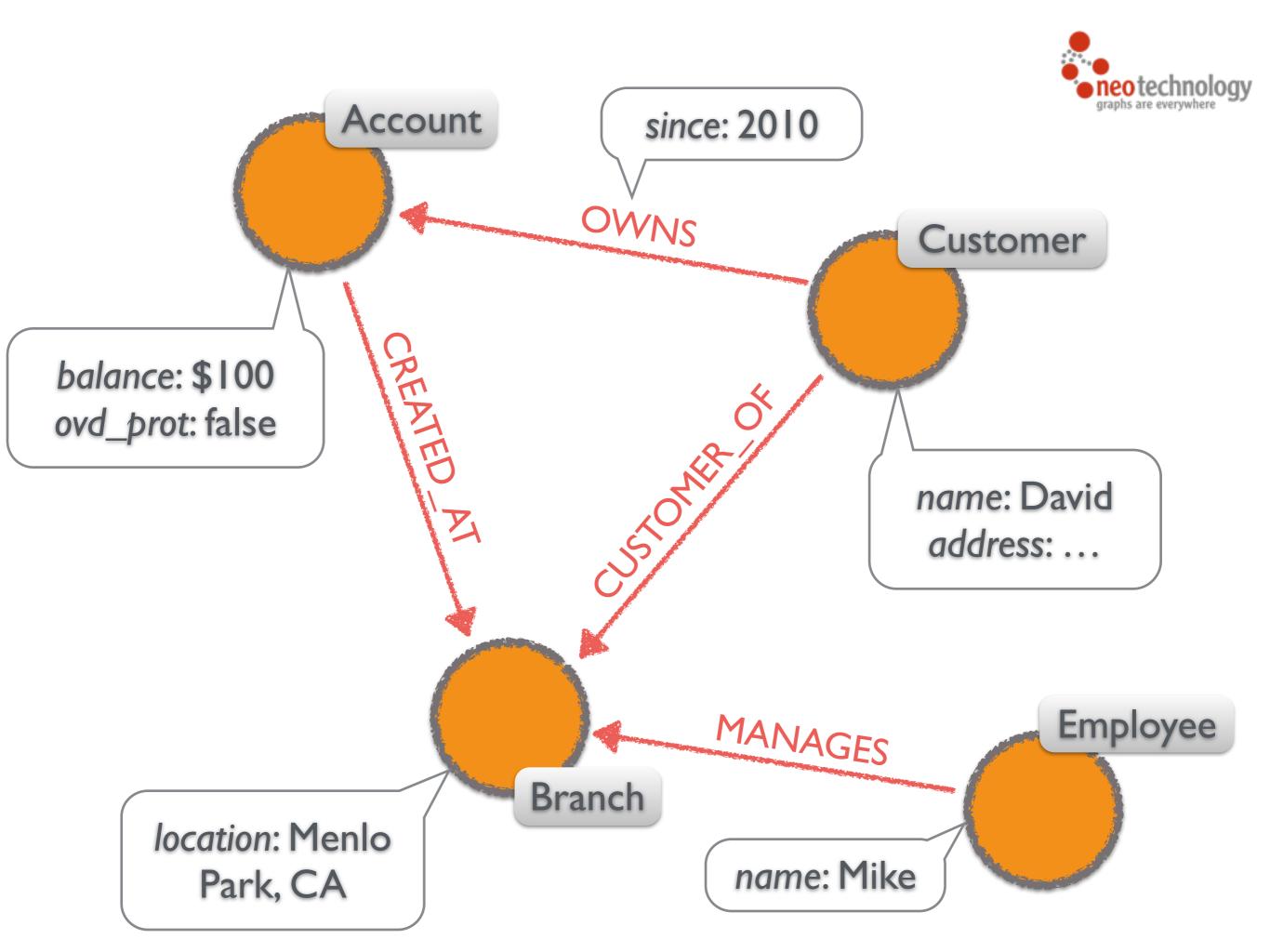




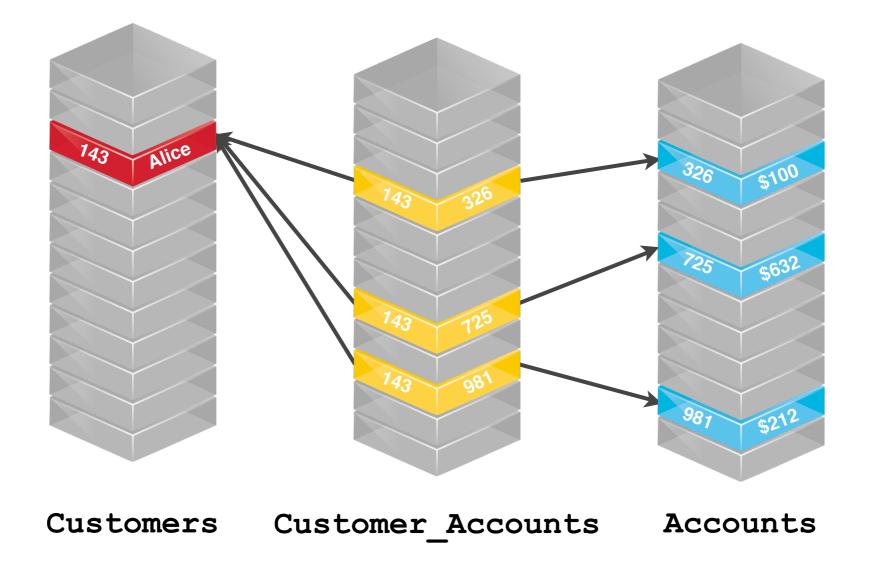
# What's a Graph Database?



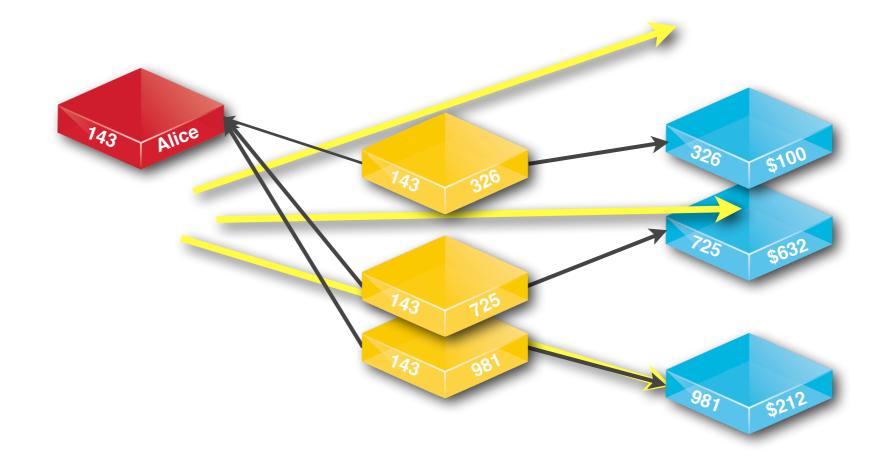




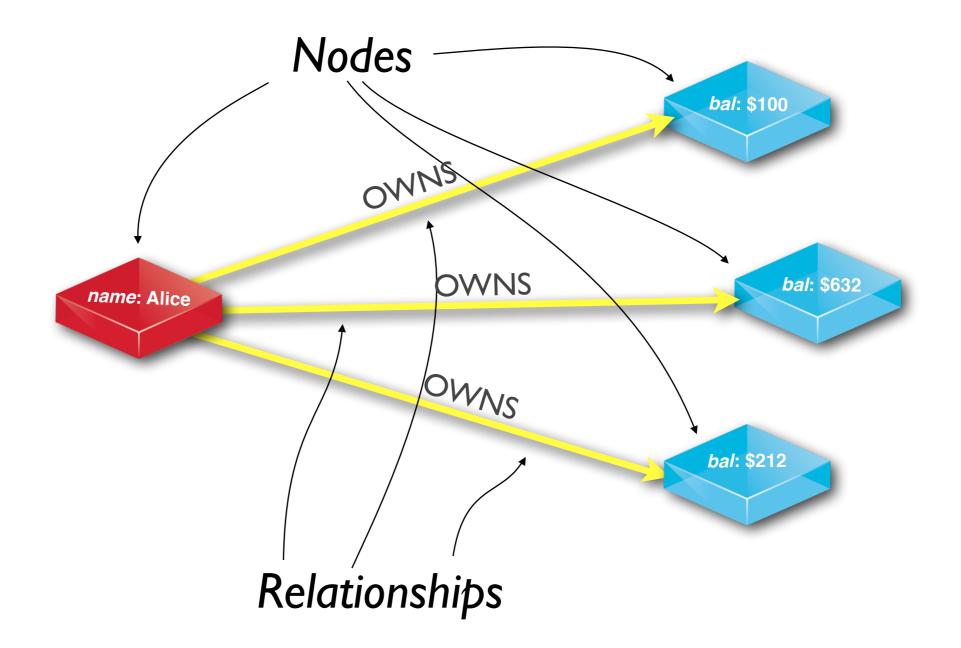














# Quick Demo

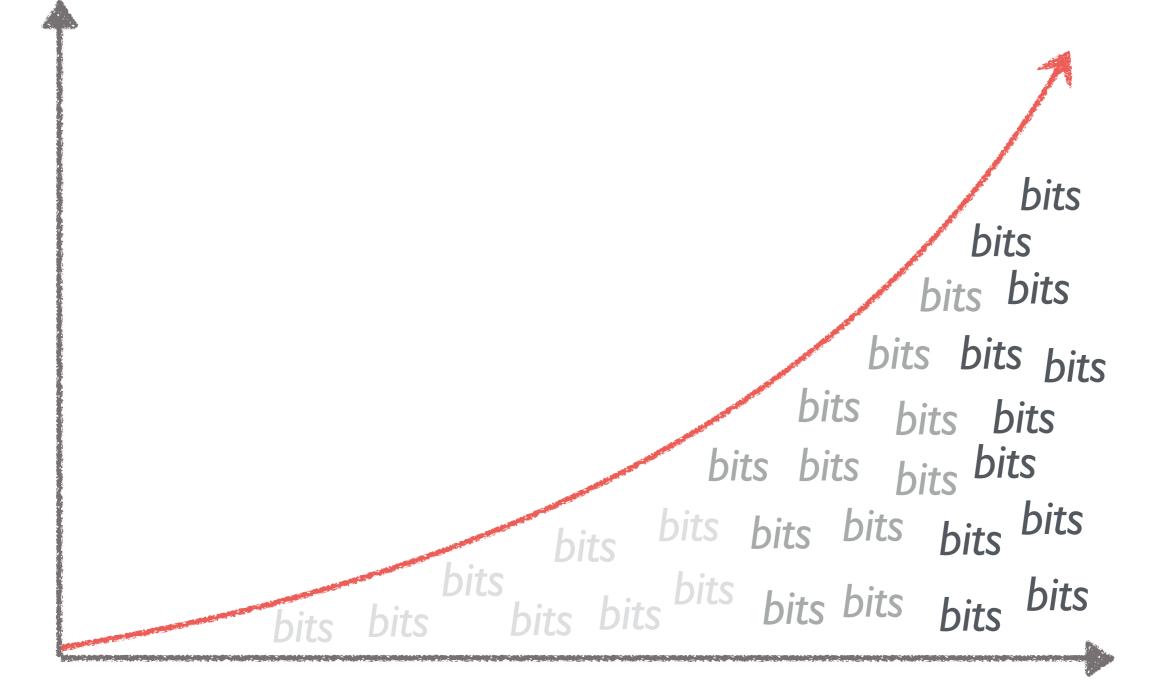


# Graph history, benefits & differentiators

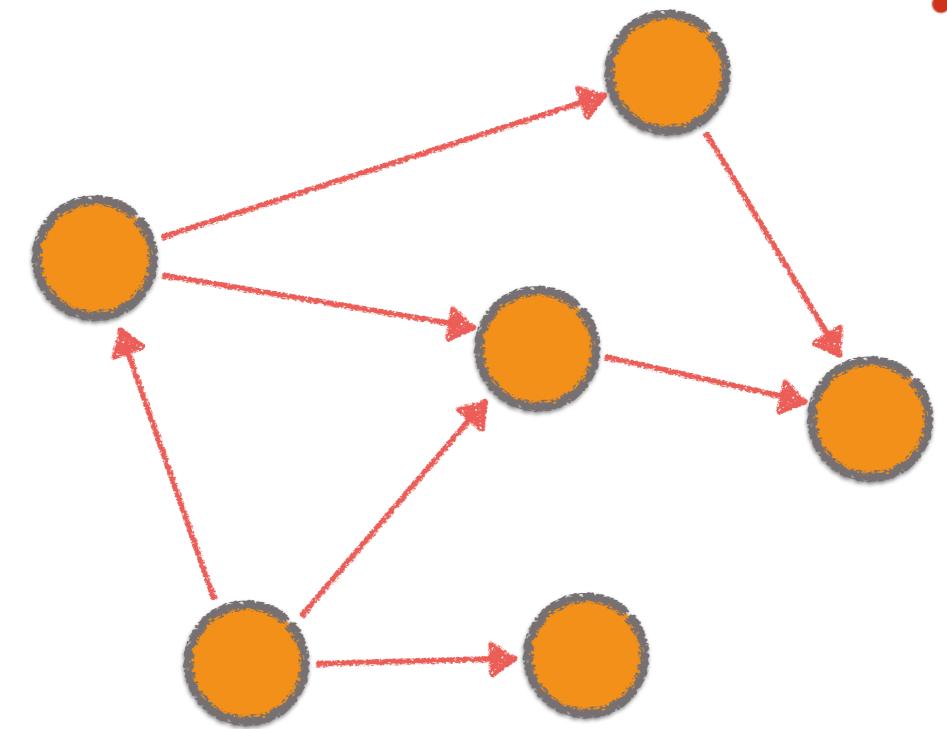


# Not Only SQL



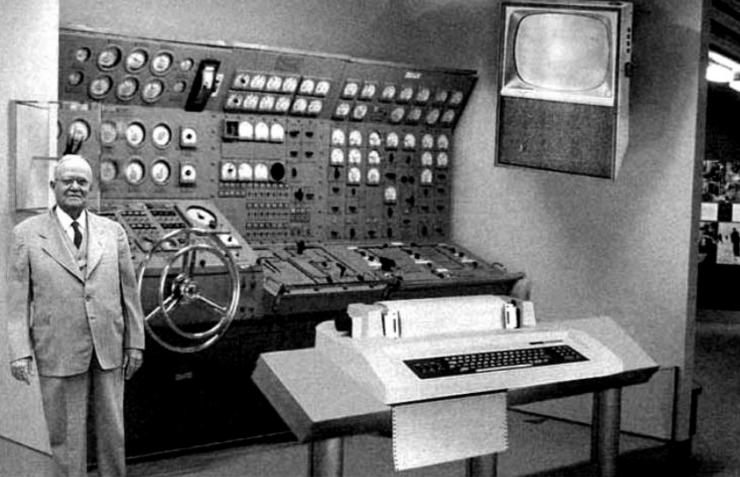


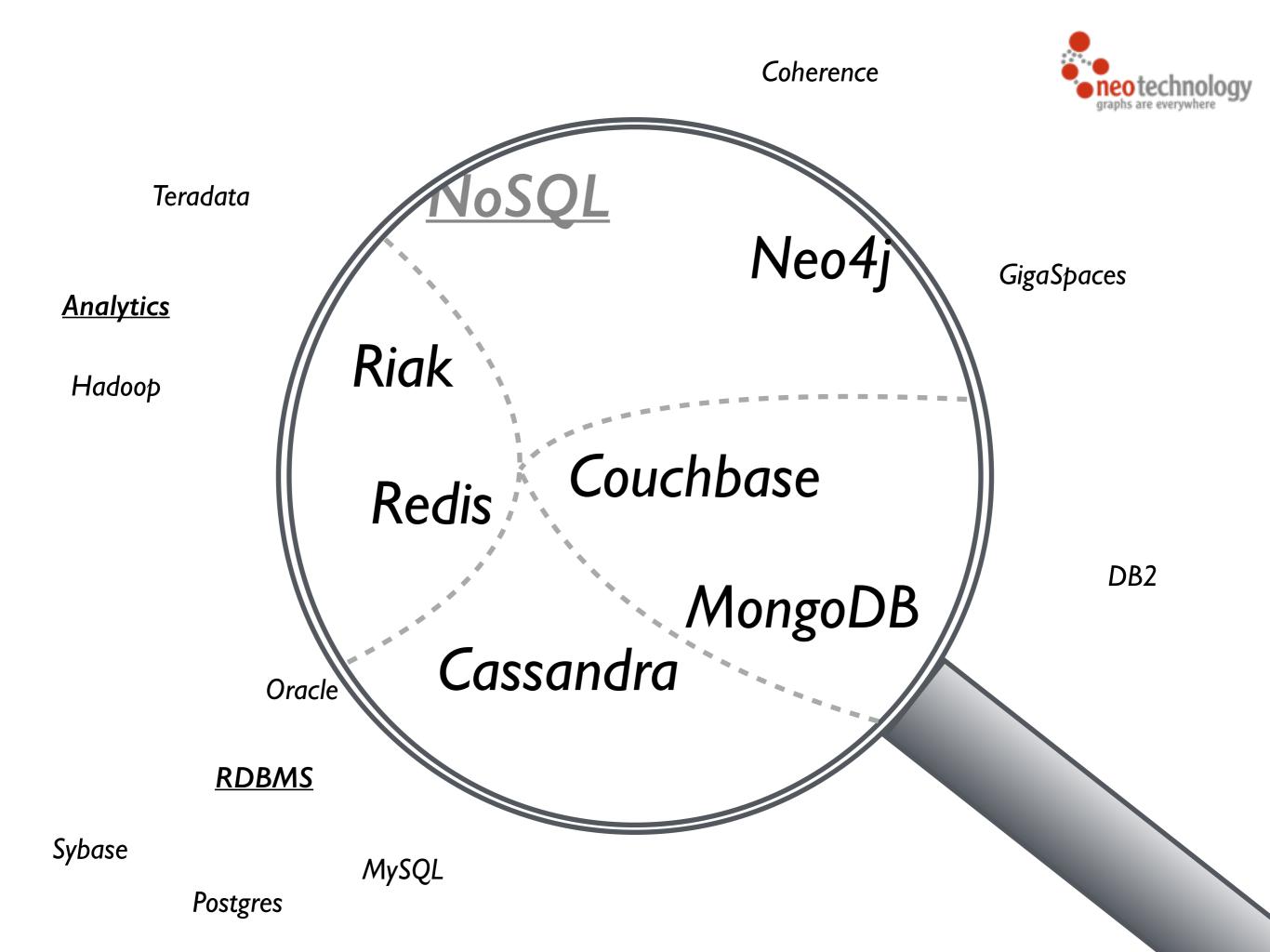




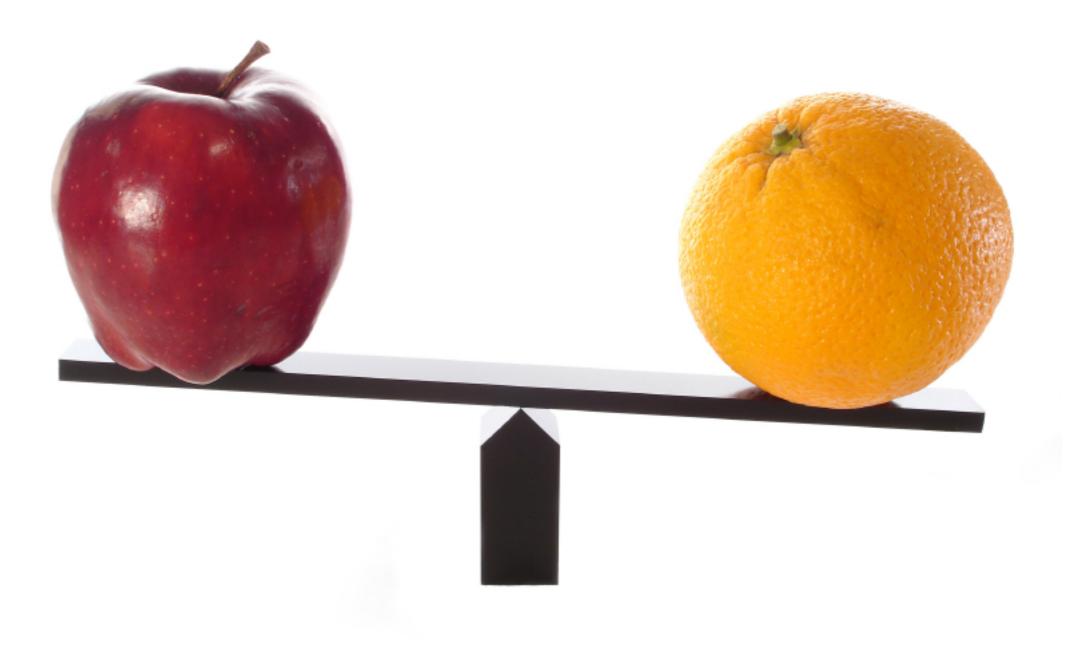












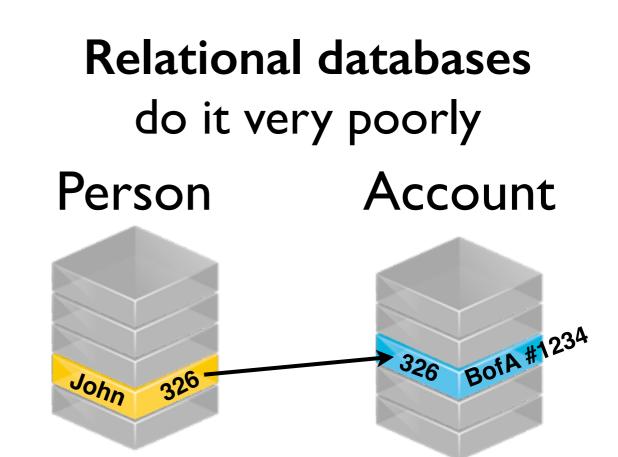
## Built ground-up for graphs



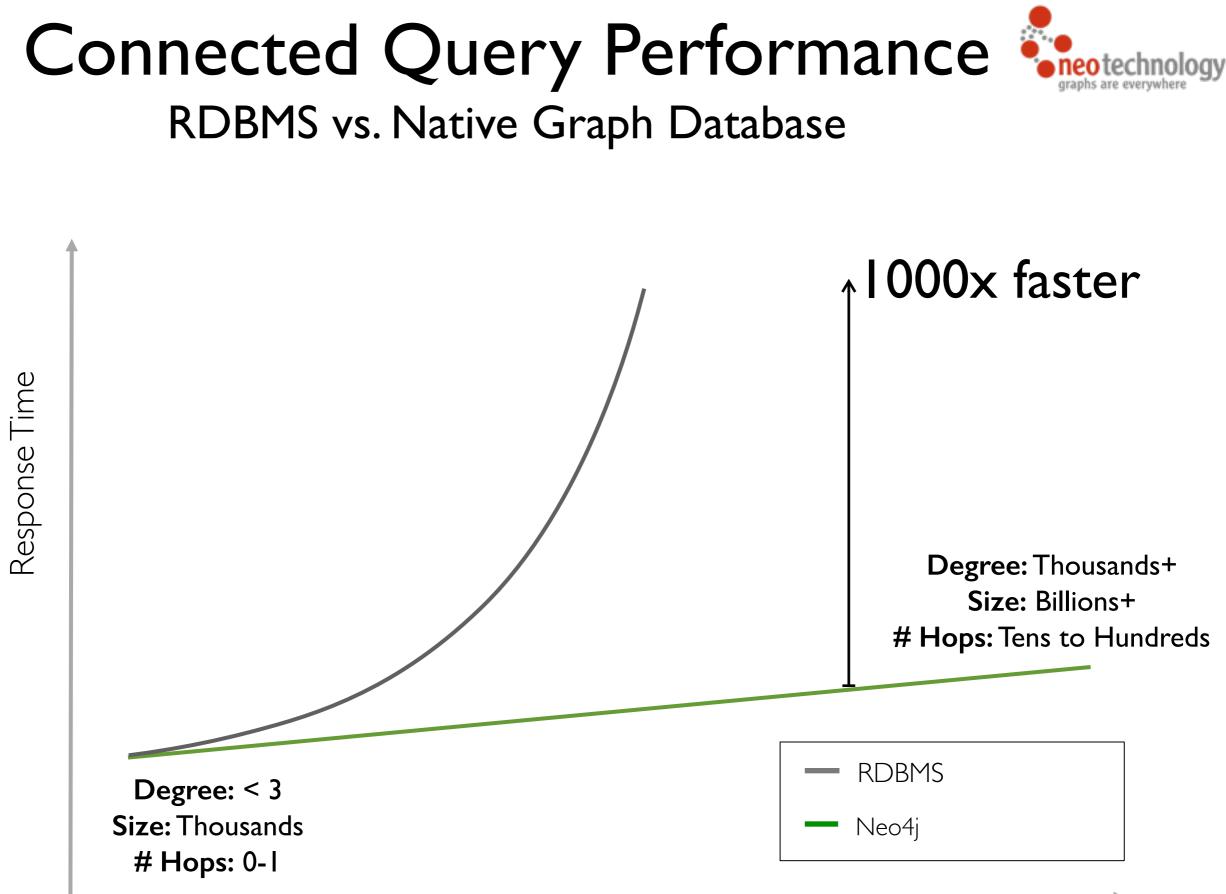
From the storage layer to the query language, graphs are native to Neo4j.

#### Other NoSQL databases don't do it at all





Rigid schema & costly joining of IDs required every lookup



Connectedness of Data Set

### Cypher vs SQL



#### MATCH (boss)-[:MANAGES\*0..3]->(sub), (sub)-[:MANAGES\*1..3]->(report) WHERE boss.name = "John Doe" RETURN sub.name AS Subordinate, count(report) AS Total

(SELECT T.directReportees AS directReportees, sum(T.count) AS count FROM ( SELECT manager.pid AS directReportees, 0 AS count FROM person reportee manage WHERE manager.pid = (SELECT id FROM person WHERE name = "fName IName") UNION SELECT manager.pid AS directReportees, count(manager.directly manages) AS count FROM person reportee manager WHERE manager.pid = (SELECT id FROM person WHERE name = "fName IName") GROUP BY directReportees UNION SELECT manager.pid AS directReportees, count(reportee.directly\_manages) AS count FROM person reportee manager JOIN person\_reportee reportee ON manager.directly manages = reportee.pid WHERE manager.pid = (SELECT id FROM person WHERE name = "fName IName") GROUP BY directReportees UNION SELECT manager.pid AS directReportees, count(L2Reportees.directly manages) AS count FROM person reportee manager JOIN person reportee L1Reportees ON manager.directly manages = L1Reportees.pid JOIN person reportee L2Reportees ON L1Reportees.directly manages = L2Reportees.pid WHERE manager.pid = (SELECT id FROM person WHERE name = "fName IName") GROUP BY directReportees ) AS T GROUP BY directReportees) UNION (SELECT T.directReportees AS directReportees, sum(T.count) AS count FROM ( SELECT manager.directly\_manages AS directReportees, 0 AS count FROM person reportee manager WHERE manager.pid = (SELECT id FROM person WHERE name = "fName IName") UNION SELECT reportee.pid AS directReportees, count(reportee.directly manages) AS count FROM person reportee manager JOIN person reportee reportee ON manager.directly\_manages = reportee.pid WHERE manager.pid = (SELECT id FROM person WHERE name = "fName IName") GROUP BY directReportees UNION

(continued from previous page ... ) SELECT depth1Reportees.pid AS directReportees, count(depth2Reportees.directly manages) AS count FROM person reportee manager JOIN person reportee L1Reportees ON manager.directly\_manages = L1Reportees.pid JOIN person reportee L2Reportees ON L1Reportees.directly\_manages = L2Reportees.pid WHERE manager.pid = (SELECT id FROM person WHERE name = "fName IName") GROUP BY directReportees ) AS T GROUP BY directReportees) UNION (SELECT T.directReportees AS directReportees, sum(T.count) AS count FROM SELECT reportee.directly\_manages AS directReportees, 0 AS count FROM person reportee manager JOIN person reportee reportee ON manager.directly manages = reportee.pid WHERE manager.pid = (SELECT id FROM person WHERE name = "fName IName") GROUP BY directReportees UNION SELECT L2Reportees.pid AS directReportees, count(L2Reportees.directly\_manages) AS count FROM person\_reportee manager JOIN person\_reportee L1Reportees ON manager.directly manages = L1Reportees.pid JOIN person reportee L2Reportees ON L1Reportees.directly manages = L2Reportees.pid WHERE manager.pid = (SELECT id FROM person WHERE name = "fName IName") GROUP BY directReportees ) AS T GROUP BY directReportees) UNION (SELECT L2Reportees.directly manages AS directReportees, 0 AS count FROM person reportee manager JOIN person reportee L1Reportees ON manager.directly manages = L1Reportees.pid JOIN person reportee L2Reportees ON L1Reportees.directly manages = L2Reportees.pid WHERE manager.pid = (SELECT id FROM person WHERE name = "fName IName")



 Forrester estimates that over 25% of enterprises will be using graph databases by 2017 to support the next-generation applications that need connected data sets.

- Forrester Research (TechRadar: Enterprise DBMS, Q1 2014)

66 ... they are the solution that can deliver truly new insights from data.

– Svetlana Sicular, Research Director, Gartner



## 4 Case Studies

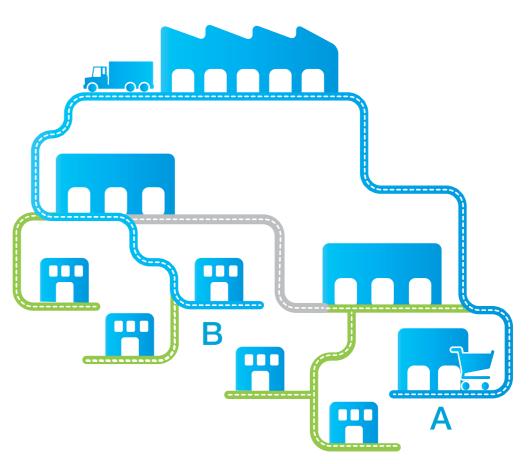


# Logistics Real-time Logistics Routing

#### **Challenge**



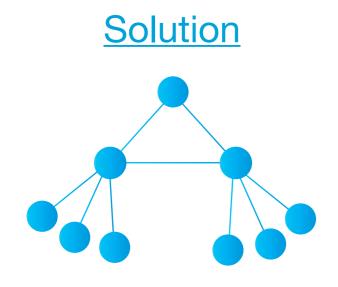
eCommerce Delivery Changing Network Dynamics



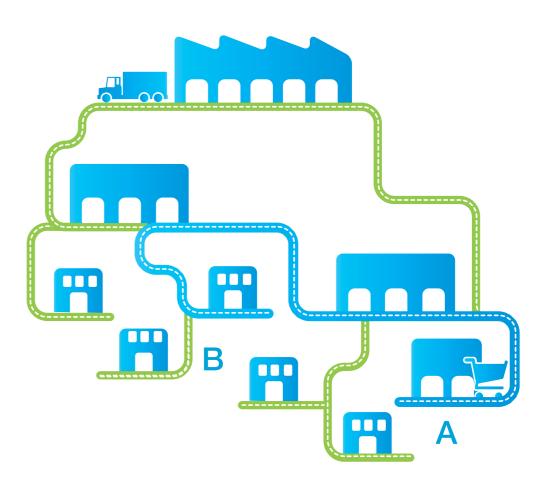
Hierarchical Routing System Did Not Support Point-to-Point Deliveries



# Logistics Real-time Logistics Routing



Model the Logistics Network as a Graph



5M Packages Per Day. 3K Per Second.

Other examples





PeoplePost



### Content Management

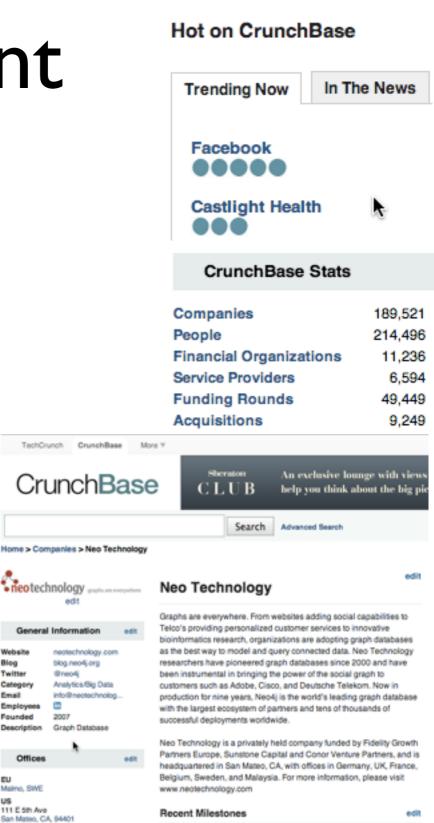


- Next-gen site required 360° deep view of any entity in the system
- RDBMS environment slow, difficult to manage and grow

#### Results

- Next-gen site deployed on Neo4j
- Statistics and drill-downs are easily created & customized





Twitte

Email

EU Maimo, SWE

US 111 E 5th Ave

URA

See nearby companies

Calepoor

Employee

Founded

Description

Neo Technology received \$111M in Series B funding. (11/2/12) Posted 11/2/12 at 4:20em via technology.com

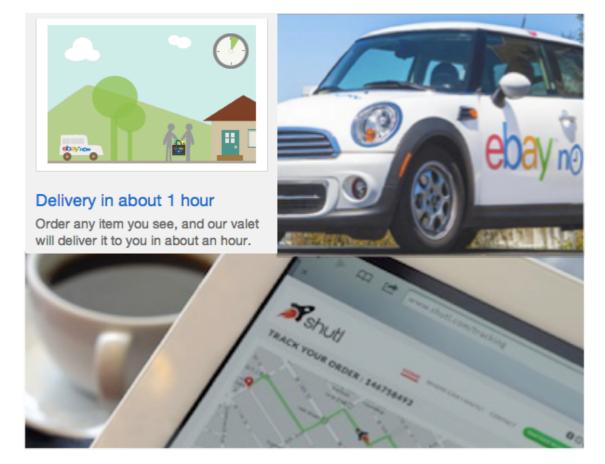


### Route Planning



### <u>Challenge</u>

- Maintain large network of routes covering many carriers and couriers
- MySQL-based solution not fast enough for real-time use



#### <u>Results</u>

- 50x less code, 2000x faster calculations
- Complete ownership of data, and flexibility to modify algorithms



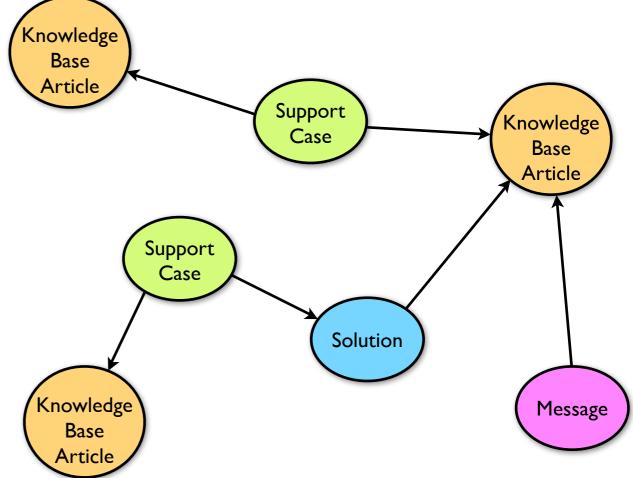
### **CISCO** Support Case Avoidance

<u>Challenge</u>

- Support cost & resolution times too high
- RDBMS infrastructure did not support expansion

#### <u>Results</u>

 Faster answers for customers, with lower reliance on support



Relational databases have a hard
time dealing with the complexities
of connected data.

– Prem Malhotra, Director Enterprise Architecture



MDM / **Recommendations** 



- Constructing a 360° view of the customer for the sales team
- IBM DB2 system not able to meet performance requirements

#### <u>Results</u>

- Flexibly search for insurance policies and associated personal data
- Migration and deployment was easy







# Patient transition & referral

#### <u>Challenge</u>

- Real-time search on Oracle not fast enough for next gen product
- Handling 15% of all transitions nationwide in the US

#### <u>Results</u>

 Real-time deep recommendations on widely heterogeneous data



Neo Technology, Inc Confidential



# Come talk to us about the graphs you see!

### Upcoming events:

- Stockholm Training on Oct 17
- Øredev Training on Nov 4