Neo Technology

The World’s Leading Graph Database

NOSQL Roadshow

Dirk Möller
dirk.moeller@neotechnology.com
Cell: +49 151 40136308
Agenda

1. About Neo Technology
2. Graph Momentum & Relevance
3. Graph Databases & The Neo4j Graph DB
4. Neo Customers
5. Q&A
Our Mission:

Help the world to make sense of data
Our Vision for Database use in 2020:
NOSQL Dominated by Graph DBMSs

Stores 50% of data
$45B market*  
Stores 50% of data
? market

* Source: Gartner
Neo Technology, Inc. - Confidential
Neo Technology
Fact Sheet

• **Numbers:**
  • 50 people / Nine countries / Four continents
  • $24M raised (seed round Oct 2009 + $10M series A led by **Fidelity** Oct 2011 + $11M in Series B led by **Sunstone** Nov 2012)
  • 30,000+ downloads per month
  • ~100 commercial customers

• **Team:**
  • Selected Leadership: Emil Eifrem (CEO) Rod Johnson (Chairman - SpringSource/VMW), Lars Nordwall (COO - ex SugarCRM, Pentaho), Philip Rathle (Sr Dir of Products - ex Accenture, Embarcadero), Jim Webber (Chief Scientist - ex ThoughtWorks)

• **Product:**
  • Development of Neo4j started in **2000** in Sweden
  • Put in 24/7 production in **2003**
  • Open sourced in **2007**
  • **Today the leading graph database**
Neo Technology story

- Idea!
- 24/7 production
- Neo Tech founded
- 1.0 OSS version released
- Seed round ($2.5M)
- Series A ($10.6M)
- Series B ($11M)
- First F100 customer
- Rod Johnson investor & advisor

Timeline:
- 2000
- 2001
- 2002
- 2003
- 2004 - 2006
- 2007
- 2008
- 2009
- 2010
- 2011
- 2012
Graph Momentum & Relevance
The Early Adopters
Facebook
Introducing Graph Search

Photos of my friends

https://www.facebook.com/about/graphsearch
Google
Google

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

See it in action
Discover answers to questions you never thought to ask, and explore collections and lists.

http://www.google.com/insidesearch/features/search/knowledge.html
When you search, you’re not just looking for a webpage. You’re looking to get answers, understand or explore.

Continue ➤
So we’re building a massive graph of real-world things and their connections, to bring you more meaningful results.

Introducing the Knowledge Graph ➤

http://www.google.com/insidesearch/features/search/knowledge.html
http://www.google.com/insidesearch/features/search/knowledge.html
Twitter
Who to follow
Twitter accounts suggested for you based on who you follow and more.

[Search using a person’s full name or @username]

Lessig
@lessig

law professor, reformer.

Followed by Manya Ellenberg, William Cook and Juan Pablo Buritica.

Jeremy Kemper
@bitsweat

Building it: 37signals, Ruby on Rails

Sarah Mei
@sarahmei

Ruby developer at Pivotal Labs, RailsBridge co-founder, dancer, gamer, friend, mom...etc.

Followed by Jeff Casimir, Mike Gehard and Vanessa Hurst.

Tom Preston-Werner
@mojombo

GitHub Co-founder - Social Coding

(http://github.com)

Followed by Derek P., Corey Haines and Paul Dix.
For the Facebook Graph Question:

What sushi restaurants in NYC do my friends like?
What the Graph Looks Like:

What sushi restaurants in NYC do my friends like?
What the Search Looks Like:

What sushi restaurants in NYC do my friends like?
What drugs will bind to protein X and not interact with drug Y?
Beyond Buzz

Some Actual Neo4j Graphs

Content Management & Access Control

Insurance Risk Analysis

Geo Routing (Public Transport)

Network Cell Analysis

Network Asset Management

BioInformatics

Bio4j
Beyond Buzz
Some Actual Neo4j Graphs

Web Browsing

Portfolio Analytics

Gene Sequencing

Mobile Social Application
# Early Adopter Segments

<table>
<thead>
<tr>
<th>Core Industries &amp; Use Cases:</th>
<th>Web / ISV</th>
<th>Finance &amp; Insurance</th>
<th>Datacom / Telecom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network Management</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geo</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Actual Commercial Demand

<table>
<thead>
<tr>
<th>Core Industries &amp; Use Cases:</th>
<th>Web / ISV</th>
<th>Finance &amp; Insurance</th>
<th>Datacom / Telecom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network Management</td>
<td>AXON ACTIVE</td>
<td>Bloomberg</td>
<td>SFR</td>
</tr>
<tr>
<td>MDM</td>
<td>Pitney Bowes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social</td>
<td>viadeo</td>
<td>careerbuilder</td>
<td>Deutsche Telekom</td>
</tr>
<tr>
<td>Geo</td>
<td>glassdoor.com</td>
<td>ice</td>
<td>Justdial.com</td>
</tr>
</tbody>
</table>

Neo Technology, Inc Confidential
# Actual Commercial Demand

<table>
<thead>
<tr>
<th>Core Industries &amp; Use Cases</th>
<th>Web / ISV</th>
<th>Financial Services</th>
<th>Communications</th>
<th>Logistics</th>
<th>Life Sciences</th>
<th>Mining &amp; Publishing</th>
<th>Education, Not-for-Profit</th>
<th>Government, Aerospace, Gaming, Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network Management</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CEP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authorization &amp; Access Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content Management</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recommendations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraud Detection, Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Web / ISV**: AXON ACTIVE, SERENA, Pitney Bowes, viaduct, careerbuilder, ice, Justdial.com
- **Financial Services**: Bloomberg, SFR, Deutsche Telekom, Zephyr Health Inc.
- **Communications**: Cisco
- **Logistics**: maaii
- **Life Sciences**: indiatimes, Squidoo, zeebox
- **Mining & Publishing**: LifeChurch.tv, DSOB New Media GmbH, justdial.com
- **Education, Not-for-Profit**: Deutscher Olympischer Sport-Bund
- **Government, Aerospace, Gaming, Other**: laurate international university, teachscape, lockheed martin

- **Accenture**: Justdial.com, telenor, Seven Bridges, fuseworks, CHIP, CHIP, Lockeed Martin, Compete, Research Now, DataLink, SA-NT, DataLink, Lockeed Martin
Why Customers Choose Neo4j

1. Order-of-magnitude improvements in \textit{query performance} for complex, connected data

2. Drastically accelerated \textit{application development cycles}

3. \textit{Maintainability} and \textit{extensibility} of the data model

4. Relative \textit{maturity} of the community and product
Benefits of a commercial relationship

• Commercial license vs open-source
• Service offerings, OEM
• Features
  • HA, DR, reporting
  • Turbo Cache
  • Online backup
  • Monitoring
## Overview of Editions & Licenses

<table>
<thead>
<tr>
<th>Edition</th>
<th>Features</th>
<th>Primary Licensing Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community</td>
<td>Full-featured graph database</td>
<td>Open Source (GPLv3)</td>
</tr>
</tbody>
</table>
| Advanced  | Full-featured graph database + Advanced Monitoring                       | Commercial License available through Neo Technology  
*Includes 5X10 Support*                                                                                                                                  |
| Enterprise| Full-featured graph database + Advanced Monitoring + Robust, fault-tolerant replicated cluster for demanding production use  
+ Online backups  
+ Turbo (aka GCR) cache | Commercial License available through Neo Technology  
*Includes 7X24 Support*                                                                                                                                  |

### (Special Cases)

- Commercial (OEM)
- Open Source (AGPLv3) - avail for Open Source Projects
- Open Source (AGPLv3) - avail for Open Source Projects

---

**Any edition can be run as Server or Embedded**

Neo Technology, Inc Confidential
Use Cases & Industries
Use Cases

Network/Cloud Management: Management & tracking of physical or virtual computing resources outside of a data center, connected by wired or wireless networking elements.

Master Data Management: Storage and management of master data, esp. that which is hierarchical and variable in structure, such as organization and product.

Social: Systems that enable users to express social behavior, such as ratings, reviews, and discovery of content and other users.

Geo: Use of Neo4j to solve problems related to geographic locations, such as best and shortest path routing and nearest neighbor
Use Cases

Resource Authorization & Access Control: Storing information about resources and parties, as well as rules governing access to resources. Facilitating efficient execution of those rules.

Content Management: Processes and technologies that support the collection, managing, and publishing of information assets.

Recommendations: Methods for recommending relevant content to a user, based on known information about users and content, including other users’ preferences and activities.

Data Center Management: Management & tracking of physical and virtual computing assets inside the data center, and how they are connected.

Fraud Detection: Refers to any one of various methods for detecting fraudulent behavior: either as it occurs, or after the fact.
Neo Customers
Adobe

social networking, recommendations, access control

Description
Identifies which collections a user has access to
Finds third-party assets that are like a user’s assets
Infers professional relations based on user-generated content

Background
Creative Cloud, announced 2011, is a cloud-based offering for professional users of Adobe’s creative suite
Collaborative Cloud is the social element of the Creative Cloud, connecting professional users around the world

Benefits
Fit
Graph model is a natural fit for social network
Collaborative user experience adds competitive advantage to Adobe offering

Flexibility
Datamodel can be easily evolved to support permissions and more sophisticated recommendation strategies

Performance
Sub-second results for large, densely-connected datasets
Cisco

*master data management, sales compensation management, online customer support*

**Description**
Real-time conflict detection in sales compensation management.
Business-critical “P1” system. Neo4j allows Cisco to model complex algorithms, which still maintaining high performance over a large dataset.

**Background**
Neo4j replaces Oracle RAC, which was not performant enough for the use case.

**Benefits**

*Performance: “Minutes to Milliseconds”*
Outperforms Oracle RAC, serving complex queries in real time

*Flexibility*
Allows for Cisco to model interconnected data and complex queries with ease

*Robustness*
With 9+ years of production experience, Neo4j brings a solid product.

**Architecture**

3-node Enterprise cluster with mirrored disaster recovery cluster
Dedicated hardware in own datacenter
Embedded in custom webapp

**Sizing**

35 million nodes
50 million relationships
600 million properties
Use case description - Cisco

**Response Time**
Real time planning and resource management is possible when the application responds to ‘what if scenario’ in less than 5 seconds

**Volume of Data**
Sales account management identifies every Cisco customer with one sales representative for a given date. This association is query intensive operation spanning across more than 15 database tables scanning through 15-100 M records

**Performance**
Oracle database performance degrades when multiple tables are joined even after accounting for query tuning, partitioning and building multiple indexes

**Concurrency**
Oracle performance further degrades when multiple read / write operations are performed. Such behavior seriously impacts year round planning where users can do many changes during ‘what if analysis’ scenario in less than 5 seconds
SFR
Industry: Datacom/Telecom
Use case: Network Management

Background

- Second largest Telco in France
- Part of Vivendi Group, partnering with Vodaphone

Business problem

- Need for flexible network inventory management, aggregation, and troubleshooting
- Impact analysis of planned and unplanned network outages, so that affected services can be notified or receive increased redundancy
- Highly volatile network structure changing daily, with business requirements changing as well

Solution & Benefits

- Neo4j Enterprise with a highly available cluster
- Dynamic system allowing for new applications to tie into network structure data
- Near 1:1 mapping of real world to graph, greatly reducing modeling work
- High adaptability to changing business requirements
**Telenor**
*organizational modeling, access control*

**Description**
Identifies which customers, accounts and subscriptions an (administrative) user has access to
Identifies which users can modify a customer, account or subscription

**Background**
Current Sybase solution is too slow and too difficult and costly to replace
Big self-service customers require better user experience and improved performance
Telenor on track to exceed batch window - sought a move to real time

**Benefits**

**Cost**
Low cost compared to alternatives

**Flexibility**
Extensible data model allows new information to be attached to the organisational structure

**Performance**
Sub-second results for complex queries across broad and deep graphs

[http://www.slideshare.net/verheughe/how-nosql-paid-off-for-telenor](http://www.slideshare.net/verheughe/how-nosql-paid-off-for-telenor)
Network Management

• Modeling of physical and logical networks
  ▶ impact analysis
  ▶ configuration management
  ▶ network inventory
Case study: Social networking
Viadeo

<table>
<thead>
<tr>
<th>Problem definition</th>
<th>Benefits &amp; time frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real-time recommendation imperative to attract new user and to maintain positive user retention</td>
<td>Scalable solution with real-time end-user experience</td>
</tr>
<tr>
<td>Clustered MySQL solution not scalable and fast enough to support real-time requirements</td>
<td>Low maintenance and reliable architecture with High-Availability (HA) and master failover</td>
</tr>
<tr>
<td>8M nodes; 35M relationships</td>
<td>8 week implementation</td>
</tr>
<tr>
<td>Complex batch calculations with week old data</td>
<td>Three technical resources involved part-time</td>
</tr>
<tr>
<td>Real-time traversals of the social graph not possible</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Solution</th>
<th>Company</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clustered Neo4j Enterprise architecture</td>
<td>The French LinkedIn</td>
</tr>
<tr>
<td>Solr for search; Tomcat; Memcache; etc.</td>
<td>35 million members</td>
</tr>
<tr>
<td>MySQL for general data storage; neo4j for social graph relationship characteristics</td>
<td>30,000 new members daily</td>
</tr>
</tbody>
</table>
Get involved in the community
http://stackoverflow.com/questions/tagged/neo4j
Neo4j Google Group

Share your experiences and expertise with fellow graphistas.

Join now »

http://groups.google.com/group/neo4j
Encountered an issue with Neo4j?
Submit it here.

https://github.com/neo4j/neo4j/issues
Meetups / User Groups

Neo4j meetups are worldwide. Make a connection or start a new group.

Join a Meetup »

http://neo4j.meetup.com/
Graphistas World Map

Add yourself to the graphistas world map and let it become a smaller place.

Add yourself »

http://www.neo4j.org/participate/contributors#map
How to reach us:

Web:
  • www.neo4j.org,
  • www.neotechnology.com

Twitter: @Neo4j, @Neo4jDE
Email: dirk.moeller@neotechnology.com
Phone: +49 800 723 6231
Q&A

http://www.neo4j.org
http://groups.google.com/group/neo4j
http://www.neo4j.org/develop
Thank you!