

java.com.sun/javaone

LinkedIn: A Professional Social Network Built with Java™ Technologies and Agile Practices

Nick Dellamaggiore, Principal Software Engineer
Eishay Smith, Senior Software Engineer





Learn how Java™ and agile practices are employed in building large-scale consumer internet sites like LinkedIn, the world's largest professional network.



GOAL



Agenda

- LinkedIn(troduction)
- Agile engineering process
- Architecture
- Building LinkedIn with Java™
- Questions/Discussion

[People](#) | [Jobs](#) | [Answers](#) | [Service Providers](#)[Account & Settings](#) | [Help](#) | [Sign Out](#)[Advanced Search](#)[People](#)[Search](#)

- [Home](#)
- [Profile](#)
 - [Edit My Profile](#)
 - [View My Profile](#)
- [Contacts](#)
 - [Connections](#)
 - [Imported Contacts](#)
 - [Network Statistics](#)
- [Inbox \(2\)](#)
 - [Messages](#)
 - [InMail](#)
 - [Introductions \(1\)](#)
 - [Invitations \(1\)](#)
 - [Profiles](#)
 - [Q&A](#)
 - [Jobs](#)
 - [Recommendations](#)
- [Groups](#)
 - [My Groups](#)
 - [Groups Directory](#)
 - [Create a Group](#)
- [**Add Connections**](#)



Nick
Dellamaggiore

Principal Software
Engineer at LinkedIn

What are you working on?

[Find out why more small businesses choose UnitedHealthcare®](#)

Inbox

- | | | |
|--|---|---------------------|
| Ed Campana | FW: Job Referral for... | May 6 archive |
| Ashish Mantri
(toashishmantri@gmail.com) | Invitation to connect | May 6 take action ▾ |

LinkedIn News

- [Most Activity](#)
 - [Company](#)
 - [Other Companies](#)
 - [Industry](#)
- [A billion-dollar valuation for LinkedIn?](#) 4 comments
Webware.com | May 06, 2008
- [LinkedIn may — and should — use its new funding to go after the business software market](#) 4 comments
Venture Beat | May 06, 2008
- [What's happening at LinkedIn? Is it getting bought?](#) 6 comments
Venture Beat | May 05, 2008
- [Allen & Co. Pitching LinkedIn at \\$1 Billion](#) [Start discussion](#)
Internet Sector and Stocks Analysis from Seeking Alpha | May 05, 2008

[Submit a new article](#)

[See more News »](#)

Network Updates

Today

[Mike Speiser](#) is now connected to [Sean Mountcastle](#) and [Catherine Davis Teitelbaum](#)

[Adam Nash](#) commented on the "A billion-dollar valuation for LinkedIn?" article. [Discuss with LinkedIn »](#)

People you may know

- | | |
|---|--|
| Eric Olson
Business Owner, IT Consulting/Staffing Firm | invite X |
| Deepak Kumar Suri
Operations East Coast - Business Development at EC-Manage | invite X |
| Krista Bradford
Executive Search & Human Capital Intelligence Consultant (krista.bradford@thegoodsearch.net) | invite X |
- [See more »](#)

EXPLORE...

Save \$100 if you register before May 12



TIECON 2008
[www.tiecon.org](#)

The World's Largest Conference for Entrepreneurs
May 16-17
Santa Clara Convention Center

Profile Views

Your profile has been viewed by **14 people** in the last 44 days. In the last day, your profile was viewed in





Your professional identity on the web

➤ The site

- Launched in 2003
- 3 degree network
- Strong focus on privacy

➤ The company

- Profitable since 2006
- 250 employees
- Based in Mountain View, CA
- We're hiring!

➤ The numbers

- 22 million members
- 4+ million unique visitors/month
- 40 million page views/day
- 2 million searches/day
- 250K invitations sent/day
- 1 million answers posted



Engineering Process

- Release often
 - Short, 2-4 week development cycles
- All dev tasks are broken down into small, manageable engineering cards
- Strong focus on testing
- Minimize meetings, status (standup meetings as needed)



Engineering Cards

Context: Assigned to me Filters Sort by Natural ▾

User: Nicholas Dellamaggiore Refresh: Au

HOME BROWSE PROJECT FIND ISSUES CREATE NEW ISSUE | PLANNING BOARD TASK BOARD CHART BOARD BOARD QUERIES

ANET HOME PAGES ▾

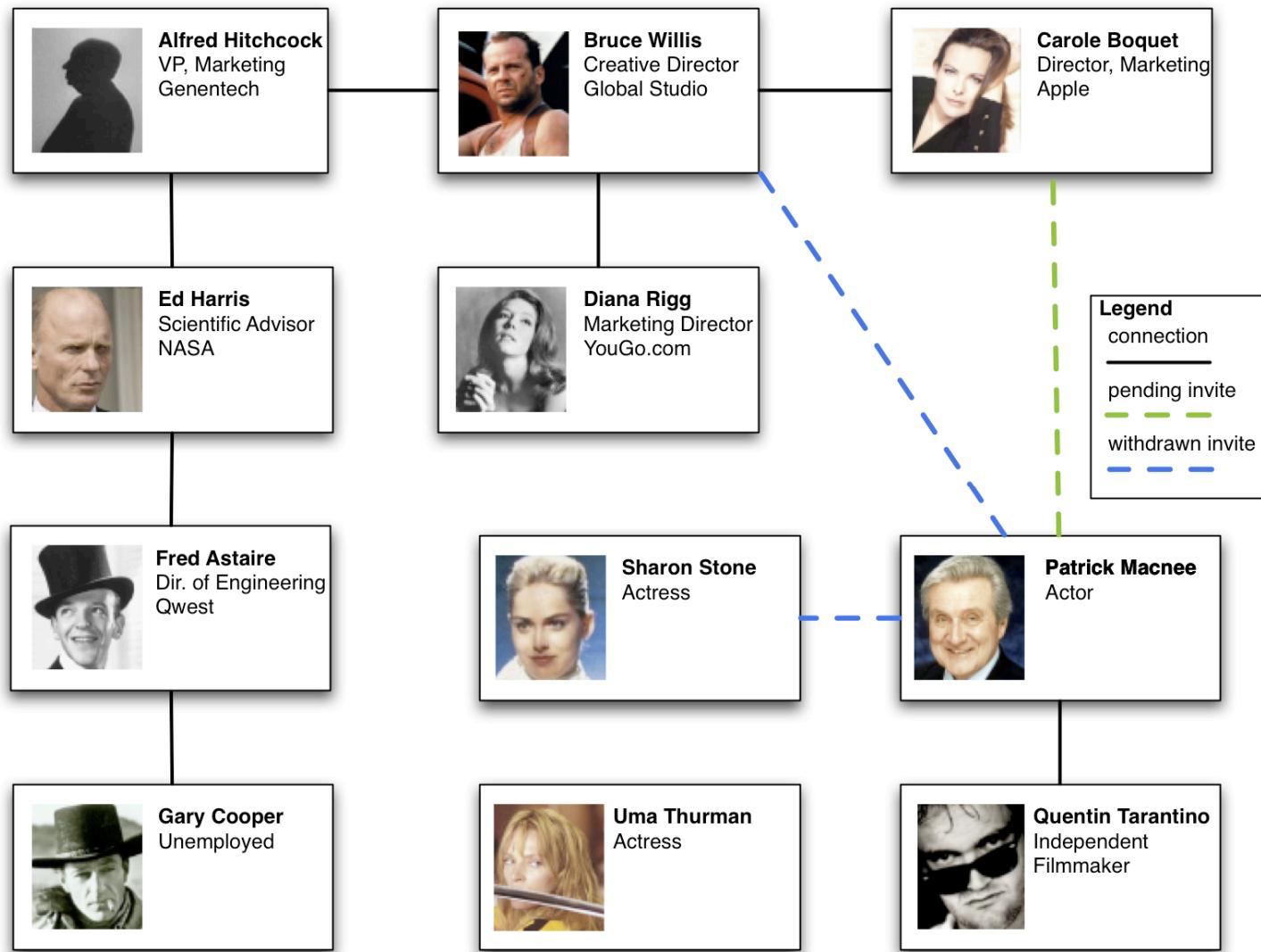
TO DO (16)	Issue Navigator	IN PROGRESS (4)	Issue Navigator	DONE (26)
CDN-526 ↑ Articles in Most Activity shouldn't be more than 7 ... Status: Open	CDN-527 ↑ 16 Subs ANet Phase 2 Status: Open	CDN-534 ↑ ANet P2 - Synchronous ANetConfirmationService Status: In Progress Version: ANet Home Pages Component: anet Original estimate: 4 days Time spent: 1 d, 3 h, 30 m Remaining: 2 d, 2 h, 30 m Assignee: dellamag	Comment	CDN-528 ↑ ANet P2 - ANetService uses surrogate ANetID Status: Closed
CDN-538 ↑ CDN-527 ANet P2 - Synchronous ANetConfirmationService: Hea ... Status: Open	CDN-550 ↑ CDN-527 ANet P2 - Shorten anet_currency column to CHAR(1) Status: Open	CDN-535 ↑ CDN-527 ANet P2 - CompanyDirectory cache invalidation Status: In Progress	CDN-536 ↑ CDN-527 ANet P2 - CompanyDirectory supports soft deletes Status: In Progress	CDN-530 ↑ CDN-527 ANet P2 - Discussions uses surrogate anetID Status: Closed
CDN-562 ↑ CDN-527 ANet P2 - Investigate/fix "contest currency" logic ... Status: Open	CDN-601 ↑ CDN-527 ANet P2 - Ordering for anet_mail_domains Status: Open	CDN-540 ↑ Add I18N support to NUS Status: In Progress Version: ANet Home Pages Component: nus Original estimate: 2 days Time spent: 1 hour Remaining: 1 day, 5 hours Assignee: sdawson	Comment	CDN-532 ↑ CDN-527 ANet P2 - UI uses surrogate anetID (especially URL ... Status: Closed
CDN-602 ↑ CDN-527 ANet P2 - Implement primary ANet selection in ANet ... Status: Open	CDN-583 ↑ Add auth check for adding article in bl Status: Open	CDN-537 ↑ CDN-527 ANet P2 - ANet membership impacted by ANet state (... Status: Closed		CDN-465 ↑ Create Engineering Specs for ANet Answers Status: Closed
CDN-587 ↑ Optimize batch delivery loop Status: Open	CDN-600 ↑ Cache news search keywords and start refactoring f ... Status: Open	CDN-520 ↑ CDN-519 BL changes Status: Closed		CDN-522 ↑ Enable users to withdraw requests to join a group Status: Closed
CDN-604 ↑ Add "common" email domains for anets Status: Open	CDN-460 ↓ Support batch look up of multiple item ids within ... Status: Open			
CDN-461 ↓ RecommendedItemStateCollectorProcess: support bat ... Status: Open	CDN-543 ↓ News Ajax Status: Open			



Testing

- 6500+ unit and integration tests
- 500 HtmlUnit tests
- Large-scale Hudson deployment
 - 20+ nodes
 - Continuous integration build/smoke test on SVN commit
 - Nightly test run (full test suite)

“Celebrities” Test Network



Hudson

Hudson

Hudson » [CDN Test Branches](#)

[People](#)[Build History](#)

Build Queue

- [BR_REL_819](#)
- [BR_ENG_SEARCH NEWS LIMITED RELEASES](#)
- [BR_ENG_MNY_CLOSE_ACCOUNT](#)
- [BR_ENG_SEARCH AYN LIMITED RELEASES](#)

Build Executor Status

Phoenix

1	Idle	gatest201
1	Idle	gatest202
1	Idle	gatest203
1	Building	BR_ENG_ADS_SAS3_BB #17
1	Building	trunk2 #6
1	Building	BR_ENG_SEARCH_PPL LIMITED RELEASES #34
gatest206 - Team Money		

All	Branches - R814	Branches - R815	Branches - R816	Branches - R817	Branches - R818	Branches - R819
S	W	Job ↓				
		BR_ENG_CDN_ABOOK_MVP1				1 day
		BR_ENG_CDN_ABOOK_MVP1_CI				1 day
		BR_ENG_CDN_ANET_HP_3				6 hours
		BR_ENG_CDN_ANET_HP_3_CI				N/A
		BR_ENG_CDN_CPROF_CS				4 days
		BR_ENG_CDN_CPROF_CS_CI				3 days
		BR_ENG_CDN_NEWS_V3				1 day
		Build stability: 1 out of the last 5 builds failed.				
		Test Result: 25 tests failing out of a total of 7092 tests.				
		BR_REL_CDN_ANET_HP_2_BB				1 day
		BR_REL_CDN_ANET_HP_2_CI				3 days

Testing...

Lessons Learned

- Test suite too focused on integration
 - Long running time
 - Long-running setup/teardown routine
 - Exacerbated by move to SOA
 - Tests become redundant due to overlapping integration tests
 - Better: use mocks (EasyMock works great)

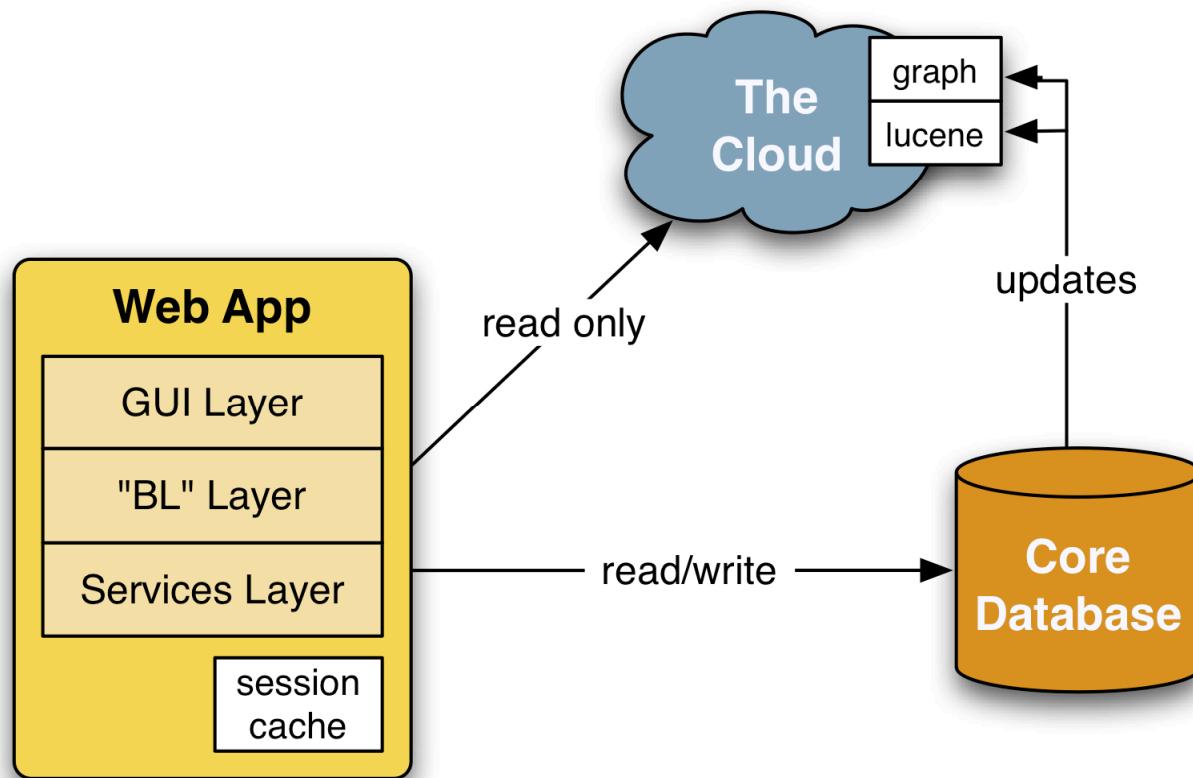


JavaOne®

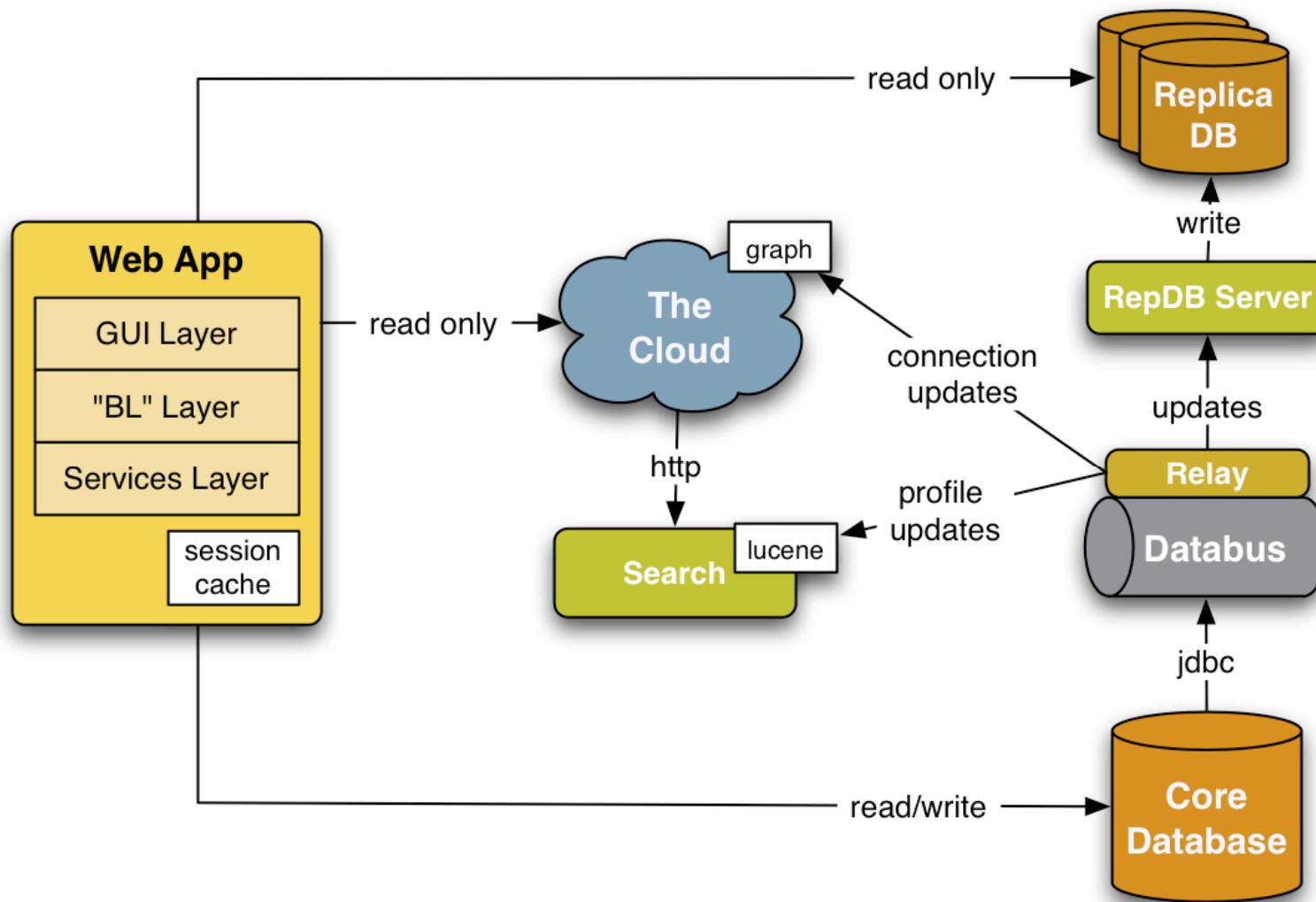
Architecture



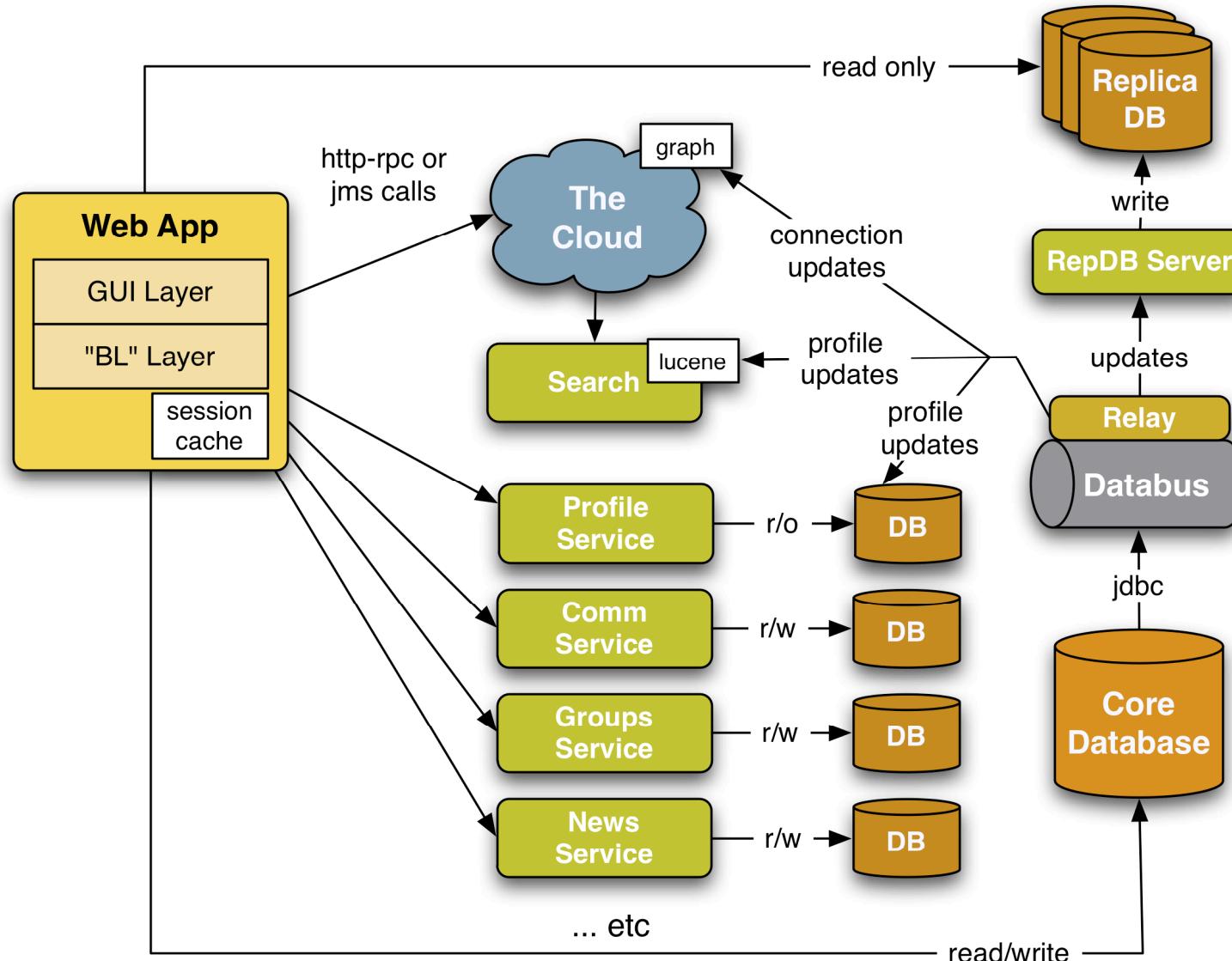
LinkedIn Architecture: 2003-2005



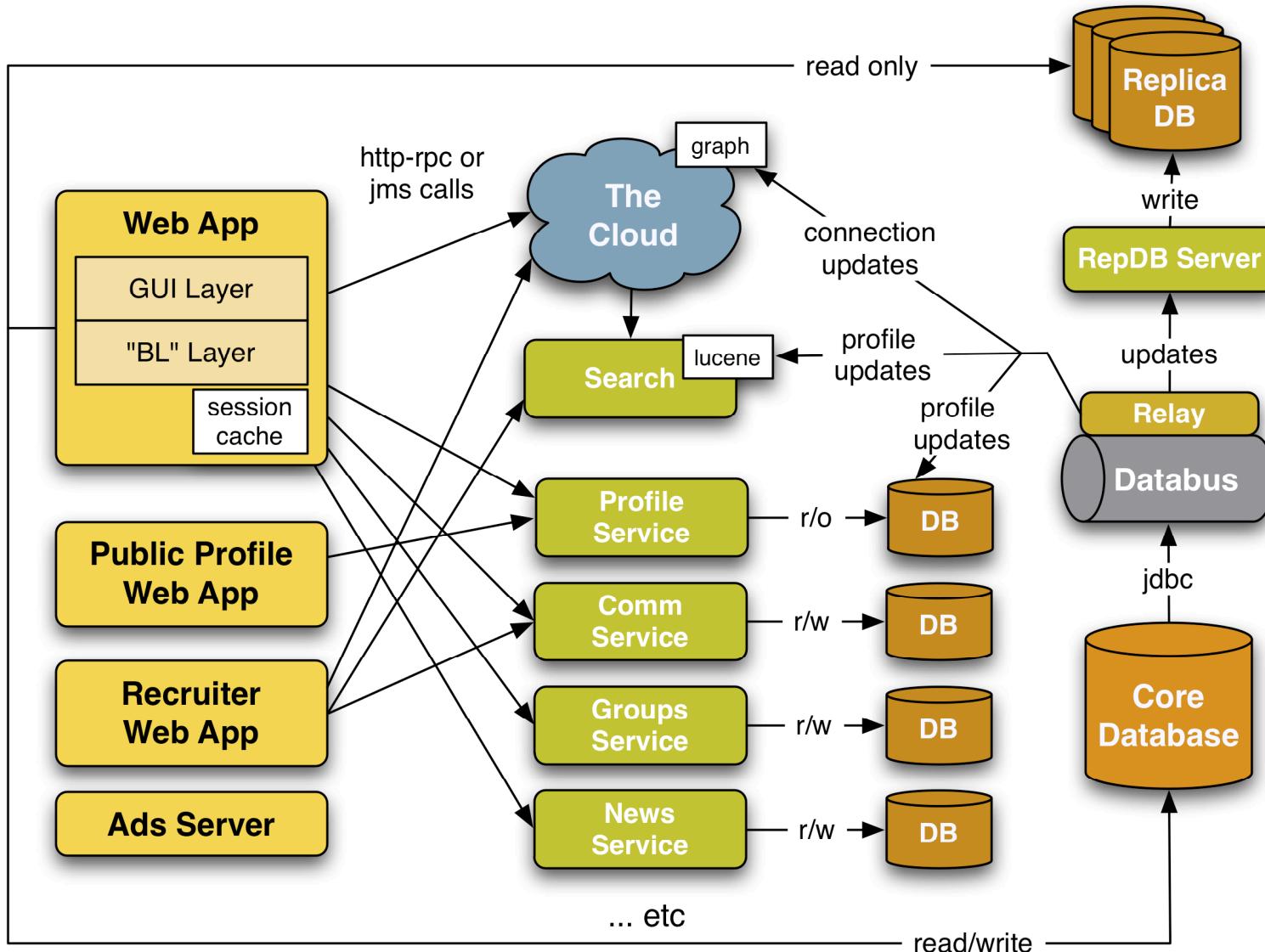
LinkedIn Architecture: 2006



LinkedIn Architecture: Today



LinkedIn Architecture: Today



Service-oriented Architecture

➤ Benefits

- Each service can scale independently
- Encourages decoupling, reduces code+deployment dependencies
- Graceful degradation of functionality

➤ Caveats

- Diminished data integrity, consistency
- No cross-service transactions, referential integrity or joins
- Backward compatibility
- “8 Fallacies of Distributed Computing”

➤ Takeaways

- Not applicable to small or simple sites, startups
- Build services with interfaces, POJOs, dependency injection
- Be conscious of API design, granularity



JavaOneSM

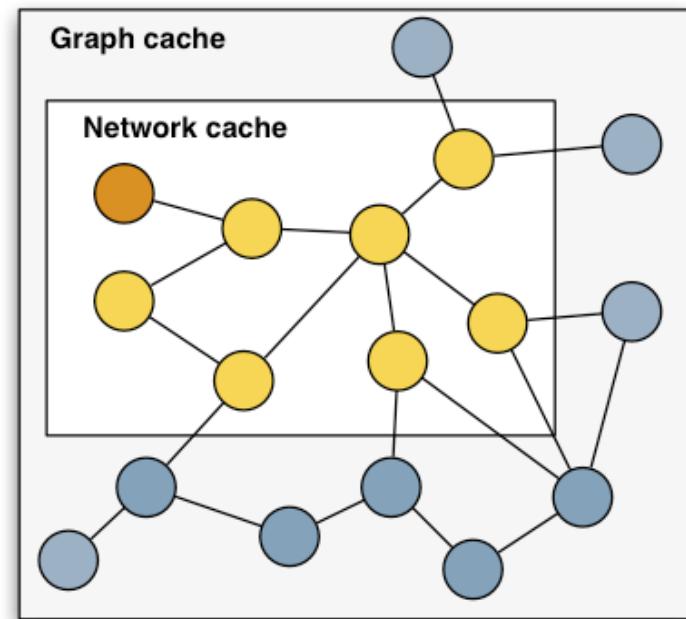
Services



The Cloud

Backend server caching the entire LinkedIn Network

- Graph cache
 - Updated via Databus
 - Persisted to disk on shutdown
- Transient network cache
 - Computationally intensive to build
 - Bound to member session
 - Sticky load balancing to Cloud
- Graph operations:
 - `findRoute(m1, m2)`
 - `visit(visitor, deg)`
 - `visit(visitor, deg, since)`





The Cloud

The numbers

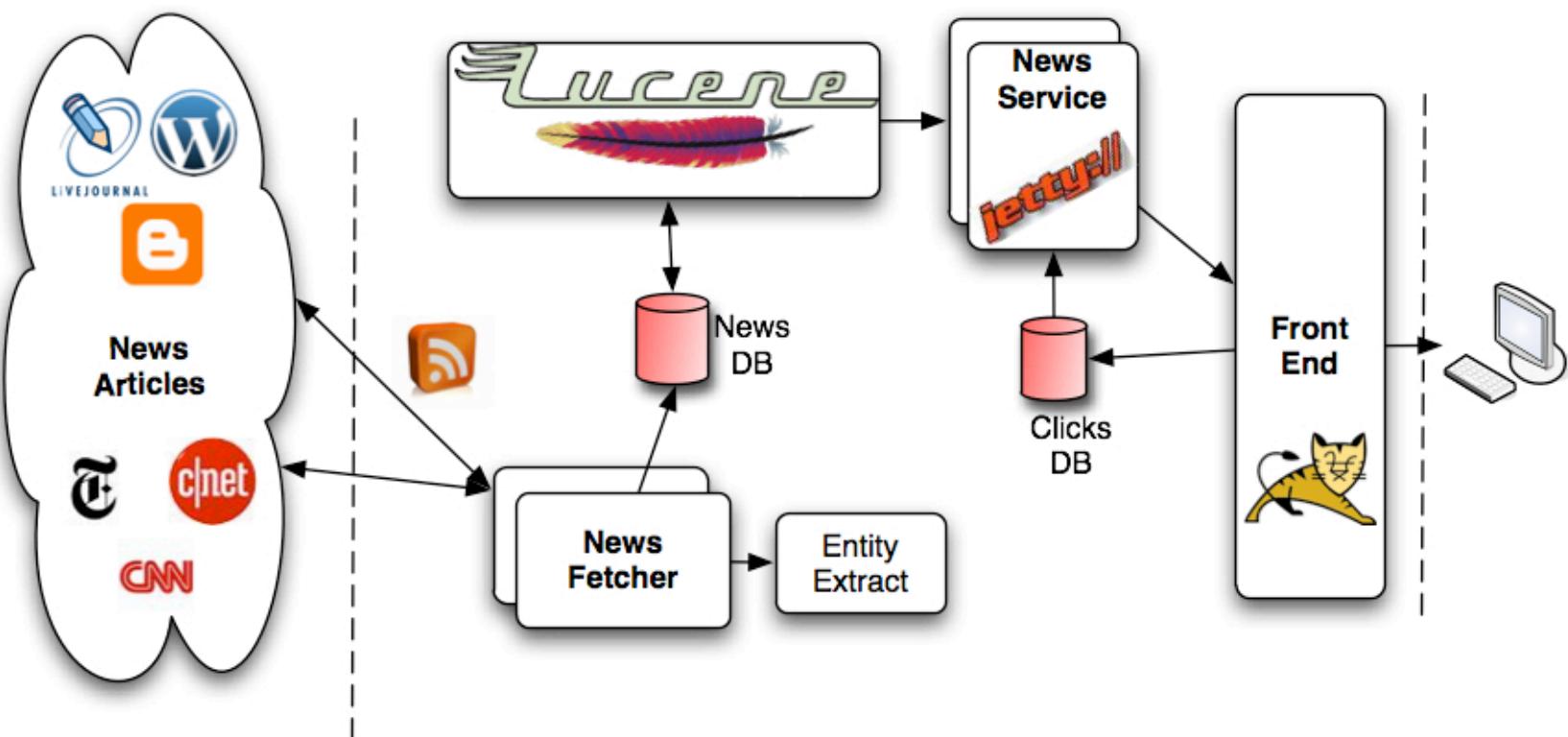
- 22M nodes, 120 million edges
- 12 GB JVM heap
 - In-memory caches implemented in C++, accessed via JNI
- 40 Cloud server instances deployed in production



LinkedIn News

- Part of a new trend in LinkedIn
- Crawling the web for news
 - Distributed task management
 - Parsing technologies: StAX, Rome
- Indexing and searching
 - Serving the right news to the right people
- Duplications Duplications Duplications
 - Real time problem, fast solution
 - 500 articles batch = 125K compares

News Service Architecture





LinkedIn: 99% Pure Java™

Plumbing	Services Tier	Web Tier	Tools
Spring Remoting	LinkedIn Spring	Louis (in-house fwk)	Hudson (CI)
ActiveMQ	Eh-cache	Spring MVC	Eclipse+Mylyn
Quartz	Lucene	Grails	JIRA/Greenhopper
HttpClient	Jetty	DWR	JUnit, HtmlUnit



DWR





Why we love Java™ at LinkedIn



- Static typing is a lifesaver
 - Huge codebase, 1M+ lines of code
 - Refactoring with confidence
 - Code navigation (via IDE)
 - 20 active branches, compiler helps with merging
 - Near-excessive use of generics
- Scaling the engineering organization
 - Currently 50+ engineers, 8 teams... and growing
 - Excellent talent pool of Java™ engineers to hire from
- Community

THANK YOU

LinkedIn: A Professional Social Network Built
with Java™ Technologies and Agile Practices

Nick Dellamaggiore (<http://www.linkedin.com/in/nick>)

Eishay Smith (<http://www.linkedin.com/in/eishay>)

We're hiring!

