

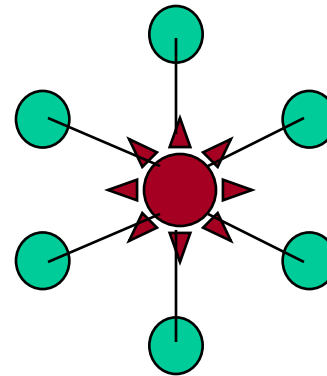
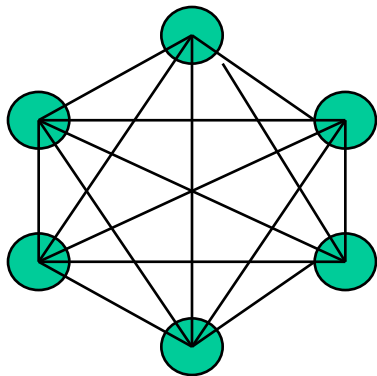
We Lose

Joe Hellerstein

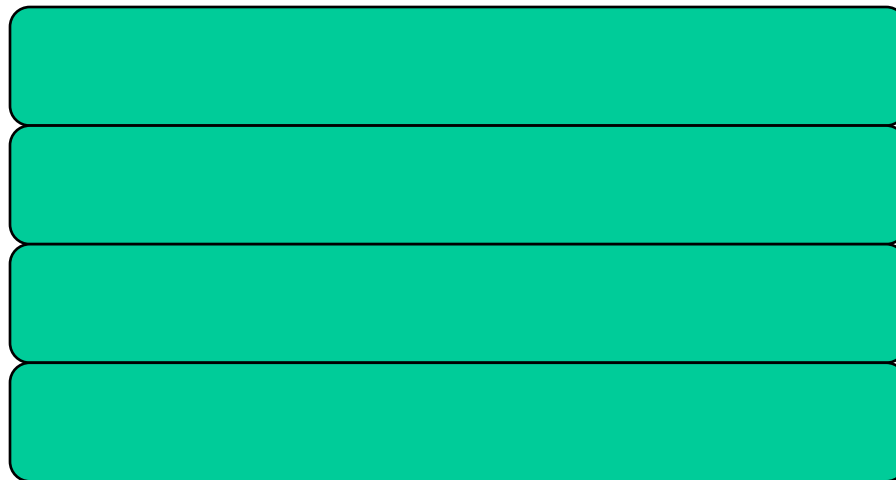
UC Berkeley

HPTS 2001

History



Generic.com, HPTS 1999



Everyone, et al., HPTS 2001

My Background

- I am a database true believer
- I am an academic
 - Enjoy elegance, beauty, conceptual richness, tweed, etc.
- I am a geek
 - I like to hack
 - I like to play sysadmin (latest fun: OS X)
- Net: I am a intellectual schizophrenic
 - And it's tearing me apart!
- Been thinking about this last couple years
 - As a result of team-teaching grad students with Eric Brewer

We Have a Beautiful Tradition

- Top-Down design of relational systems, TP
 - Semantics first
 - Codd gives a data model and declarative languages
 - System R crowd defines serializability, etc.
 - Implementation wizardry later
 - And we can rise to any challenge -- just watch us!
- We love to set high goals
 - General-purpose storage/query
 - No sacrifice on performance or availability or distribution or parallelization or load-balancing or....
 - AND still give you beautiful semantics
- Care about the user and their data

But We Lose!

- Grassroots* use Filesystems, not DBs
- Grassroots use App servers, not ORDBs
- Grassroots write Java, PERL, Python, PHP, etc. etc. etc. NOT SQL!
 - or XQuery ...

Hanging Out with OS Folks

- OS folk have a beautiful tradition too
 - Simple, narrow interfaces and tools
 - Strong ties to the PL world
 - Care about the programmers and their tools
- Bottom-up elegance
 - KISS
 - The art of engineering first. Semantics later.
- [Ahhh... DB folk can hack better than they can!]
 - Not the point!

Why We Lose

- While we've been thinking about users...
- The OS/Bottom-Uppers have been targeting programmers
- Without programmers, it's damn hard to reach users
 - THIS is why we missed the first waves of the Internet
 - THIS is why ORDBMS lost to App Servers
 - Like trying to sell drugs without the Mafia

Why We're Supposed to Win

- Eventually they'll come crying to us
 - When they realize they should have had data independence
 - And we give 'em the best server platform
- Maybe.
 - Maybe that will be too late.
 - Certainly their server platforms have been catching up
 - Databases commoditized and cornered?
 - To slow-moving, evolving, structure-intensive apps that *require* schema evolution
- Maybe we should reach out more?

Tools and Community

- We need to work on tools
 - Query debuggers
 - Data cleaners
 - Workflow builders/debuggers
 - Right direction: GUIs + logic + a little AI
 - Semi-automatic tools
 - One possibility: our “programmers” are *content managers* (see Stonebraker/Hellerstein SIGMOD ‘01)
- We need to foster a vibrant grassroots community
 - A la the USENIX/OpenSource world
 - How the hell do we do this? Databases are boring!
 - Even open source databases are boring!
 - Far more Linux/BSD buffs than Postgres/Sleepycat/MySQL

Fun With Our Stuff

- Query processing
 - It's not boring, it's the coolest thing in computing!
 - Text search, P2P. QP over *their* data, not DB data.
 - Or “continuous” QP
 - Pub/sub for fun (and profit?)
- A toolkit for QP, Cont. QP (pub/sub), etc.
 - Bottom-up: more like query algebra than SQL
 - Dataflow diagrams or “pipe” scripting
 - Plus infrastructure to make it work well over the Internet, etc.
 - You wanted it to be fun, right?
 - Deal with wide-area federation, adaptivity
 - Telegraph project at Berkeley

My Gut

- We lost the cool new Internet enterprise space
- Only the standard enterprise folks will come to us tail-between-legs
- We need to look for a new opportunity
 - Maybe P2P QP
 - Search interface today is just string-match. How to do more?
 - Maybe Internet QP
 - Web search can't do the most obvious things
 - Needs query/wrapper tools desperately
 - Maybe ubiquitous computing/sensornet QP?
 - See Intel Research lablets @ Berkeley/CMU/Washington
 - None of this seems like a business model (yet)
 - But it's cool – just what we're missing!

P2P Example: TeleNap


The screenshot displays the Telegraph Napster Client interface. On the left, a 'Server List' pane shows various server addresses. The main window, titled 'Telegraph Napster Client', has a menu bar with 'File' and 'Edit'. Below the menu bar are three tabs: 'Search', 'Downloads', and 'Console'. The 'Search' tab is active, showing a search bar with 'beatles' and a dropdown menu set to 'V'. To the right, another search bar contains 'white album' with a dropdown menu also set to 'V'. A checked 'Album Search' option is visible. A 'Clear Table' button is located below the search bars. The main area contains a table with the following columns: Song, Size, Server, Nick, Bitrate, Length, and LinkT... The table lists various Beatles songs and their details.

Song	Size	Server	Nick	Bitrate	Length	LinkT...
beatles(white album)-dear prudence.mp3	3784...	planet...	colma...	128	236	0
Beatles, The_02_Dear Prudence.mp3	3793...	planet...	cipollo...	128	237	8
Beatles - Continuing Story Of Bungalow Bill.....mp3	7798...	planet...	1	320	195	0
The Beatles - The BEATLES (White Album) Disc 1 - 06. The Continuing Sto	3105...	planet...	peach...	128	194	1
Beatles - The Continuing Story Of Bungalow Bill.mp3	2229...	planet...	rickde...	96	185	1
Beatles, The_06_The Continuing Story Of Bungalow Bill.mp3	2979...	planet...	cipollo...	128	186	8
Beatle - The Continuing Story Of Bungalow Bill.mp3	2229...	planet...	rickde...	96	185	4
beatles_-_white_album_1_-_track_06_-_the_continuing_story_of_bungalow	2957...	planet...	^^Silv...	128	184	5
The Beatles - The continuing story of Bungalow Bill.mp3	3104...	planet...	sgfrd...	128	194	1
Beatles - The Beatles (White Album) Disc 1 - 06 - The Continuing Story Of	3105...	planet...	colma...	128	194	0
Beatles, The_06_The Continuing Story Of Bungalow Bill.mp3	2979...	planet...	cipollo...	128	186	8
Save Ferris (Live) - Ob-la-di Ob-la-da (Beatles Cover).mp3	3670...	planet...	20giu...	128	229	4
ob-la-di ob-la-da.mp3	2336...	planet...	linkhir...	96	195	4
TheBeatles 1967 - 1970 CD2 03 Ob - La - Di,Ob - La - Da.mp3	3879...	planet...	brooklyn	64	485	8
The Beatles - 1967-1970 (Disc 2) - 03. Ob-la-di, Ob-la-da.mp3	3061...	planet...	peach...	128	191	1
Beatles - OB-la-di, OB-la-da.....mp3	7653...	planet...	1	320	191	0
The Beatles - The BEATLES (White Album) Disc 1 - 04. Ob-La-Di, Ob-La-	3022...	planet...	peach...	128	188	1
The Beatles - 1967-1970 (Disc 2) - 03. Ob-la-di, Ob-la-da.mp3	3061...	planet...	peach...	128	191	1
(Beatles) - Ob-la-di Ob-la-da.mp3	2296...	planet...	ej3452	96	191	8
Beatles - Ob-La-Di, Ob-La-Da.mp3	3778...	planet...	jersey...	160	188	7
30 Days - Disc 01 - Track 11 (Ob-La-Di, Ob-La-Da).mp3	1867...	planet...	lstoutl...	160	93	8
Beatles - Ob-La-Di Ob-La-Da.mp3	2969...	planet...	RedBa...	128	186	8
The Marmalade-Ob-La-Di, Ob-La-Da.mp3	3071...	planet...	Anjuli	128	191	8
Beatles, The_04_Ob-La-Di, Ob-La-Da.mp3	3030...	planet...	cipollo...	128	189	8
Beatles - Ob-la-di Ob-la-da.mp3	4558...	planet...	mamo...	192	189	4


Web Query Example: Telegraph FFF

Federatded Facts and Figures - Microsoft Internet Explorer

Address f.cs.berkeley.edu/demo5.html Go Back Forward File

Query Finished 

Results



Name	Occupation	Address	Amount
Smits, Jimmy	Self employed	Los Angeles, ...	250.00
Somers, Suzanne	Self	Valencia, CA ...	1,000.00
Stamp, Terence	Info Requested	Sanbornville, ...	1,000.00
Stone, Sharon	Self employed/Actress	Los Angeles, ...	1,000.00
Streisand, Barbra	Self employed/Singer / Prod...	Santa Monica...	1,000.00
Taylor, Elizabeth	Not employed/Homemaker	Tampa, FL 33...	250.00
Thomas, Heather	CIGNA Healthcare/New Busi...	Nashville, TN ...	250.00
Thomas, Michelle		Washington, ...	300.00
Thomas, Olive	National Council of Church...	Maryville, TN ...	1,000.00
Thomas, Olive	National Council of Church...	Maryville, TN ...	1,000.00
Tomlin, Lily	Self employed/Actress	Los Angeles, ...	250.00
Tripplehorn, Jeanne	Self employed/Actress	Los Angeles, ...	1,000.00
Wagner, Robert	Self employed/Doctor	McLean, VA 2...	500.00

134 -

13