Structured Data Storage

Xgen Congress Short Course 2010

Adam Kraut

BIOTEAM
Enabling Science

BioTeam Inc.

- **Independent Consulting Shop:**
  Vendor/technology agnostic

  Staffed by:
  - Scientists forced to learn High Performance IT to conduct research

- **Our specialty:**
  Bridging the gap between Science & IT

BIOTEAM
Enabling Science

Data Management Buzzwords

- Linked Data
- NoSQL
- Distributed Database
- Non-Relational (Schema-free)
- Document-based
- Object-based
- Key-value
- Partitioning
- Fault Tolerance

BIOTEAM
Enabling Science

One Size Does Not Fit All

- RDBMS have become ubiquitous
  - Often synonymous with the term database
- Databases precede the implementation relational systems
- Structured storage extends far beyond the relational realm
- 90% of applications are using 10% of the features of modern RDBMS

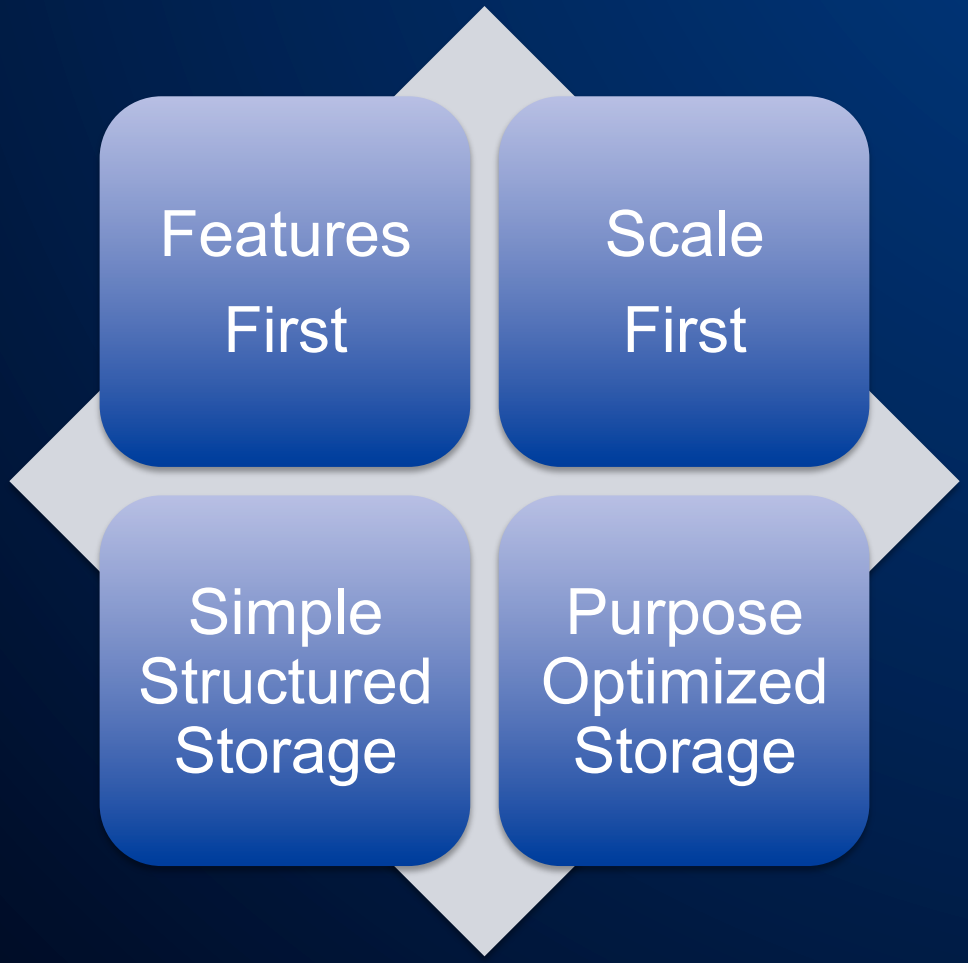
BIOTEAM
Enabling Science

Scaling RDBMS

- "An infinitely scalable relational database is an engineering impossibility" – Werner Vogels

BIOTEAM
Enabling Science

Database taxonomies



BIOTEAM
Enabling Science

Feature-first

- Oracle
- SQL Server
- PostgreSQL
- MySQL

- Even in large enterprises, a single database instance can support the entire workload

BIOTEAM
Enabling Science

Scale-first

- Scale is more important than features
- When a single RDBMS won't cut it
  - Shard the data over a large number of systems
- Full relational model is impossible to scale
  - Cross-instance joins
  - Aggregations
  - Globally unique secondary indexes
  - Global stored procedures
- Examples
  - DB2 Parallel Edition
  - Oracle Real Application Clusters (RAC)

BIOTEAM
Enabling Science

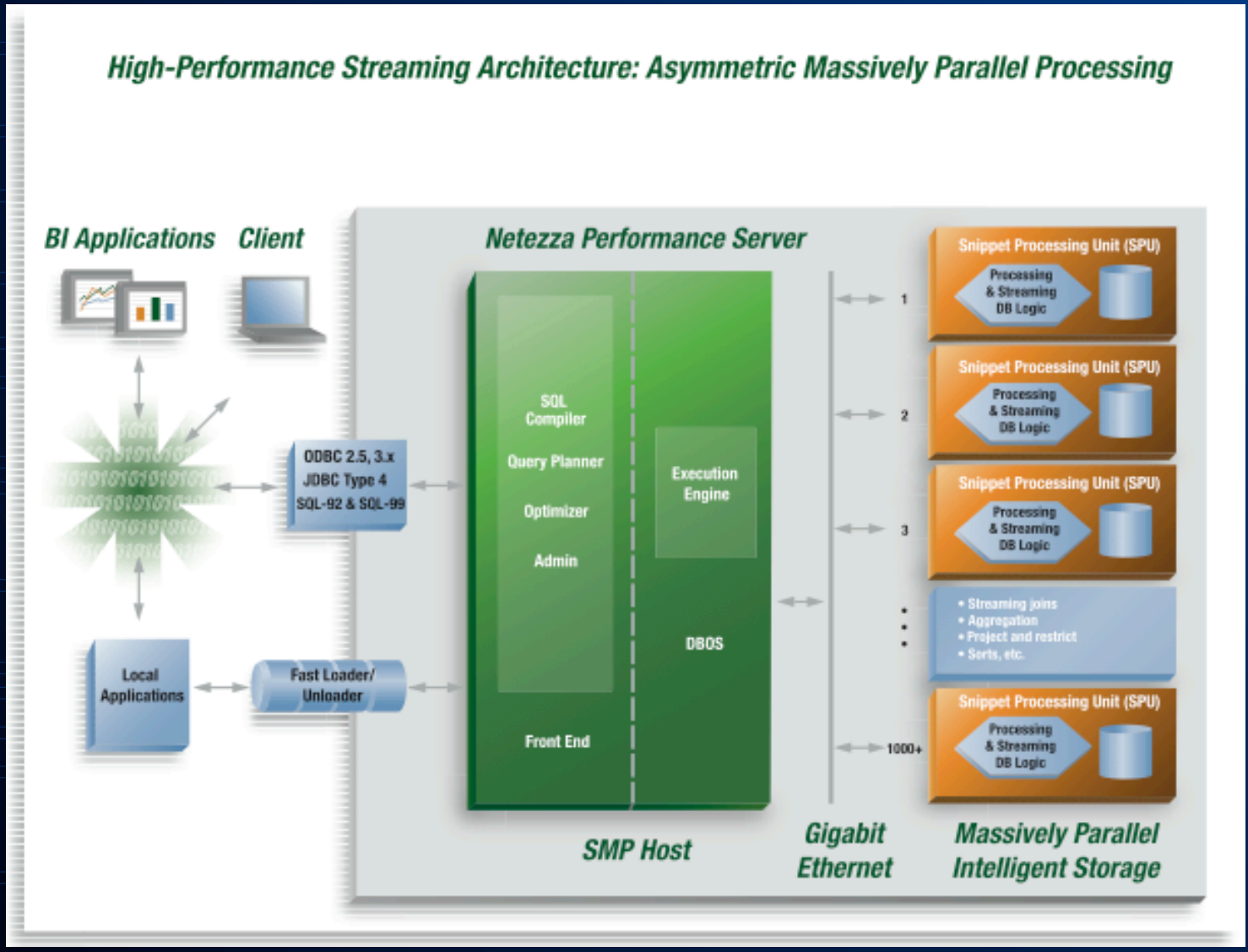
Purpose Optimized Storage

- Special purpose
- Often designed to beat commercial RDBMS on specific benchmarks

- Event stream processing
- Data warehousing products

- Examples
  - Aster Data, Netezza, Greenplum

BIOTEAM
Enabling Science



BIOTEAM
Enabling Science

Simple Structured Storage

- Simple, cheap, fast
- Low operational burden

- Examples
  - BerkeleyDB
  - SimpleDB
  - MongoDB

BIOTEAM
Enabling Science

Alternative Database Engines

- BerkeleyDB - http://www.oracle.com/database/berkeley-db
- memcached - http://memcached.org
- BigTable - http://labs.google.com/papers/bigtable.html
- HBase - http://hadoop.apache.org/hbase
- CouchDB - http://couchdb.apache.org
- MongoDB - http://www.mongodb.org
- Tokyo Cabinet - http://1978th.net/tokyocabinet
- Redis - http://code.google.com/p/redis
- Riak - http://riak.basho.com
- Cassandra - http://incubator.apache.org/cassandra

BIOTEAM
Enabling Science

MongoDB

- Document-oriented storage (JSON-like schema)
- Written in C++
- Fast, in-place updates
- Replication, and fail-over support
- Auto-sharding
- MapReduce for aggregations
  - Written in Javascript

BIOTEAM
Enabling Science

Redis

- Fast, in memory key-value store
- STRING, LIST, SET, and ZSET data types
- Persistence via async snapshots
- Perfect Data Structures/State/Cache Server

BIOTEAM
Enabling Science

SimpleDB

- Hierarchical structure, not a table
- Schema-less
  - Attributes only exist when associated with a value
  - No NULL values
- Limited query capability
  - No SQL
  - No joins
- All data is stored as text
  - No data types
- Limited Attribute Sizes (1024 bytes)
- Eventual consistency model
  - Information may be slightly out of date

BIOTEAM
Enabling Science

NoSQL Hype vs. Reality

- Schema-free
- Scalable
- Fast
- Hierarchical data structures

- No general-purpose query language
  - Yet another language to learn
- Many-to-many relationships are problematic
- Lacking tool support
- Lacking library support

BIOTEAM
Enabling Science

BioTeam's Use of NoSQL

- Today…
- WikiLIMS
  - Semantic MediaWiki (RDF Triple store)
- Configuration Management Framework
  - Chef (CouchDB)
- Amazon Web Services Workflows
  - SimpleDB to store state

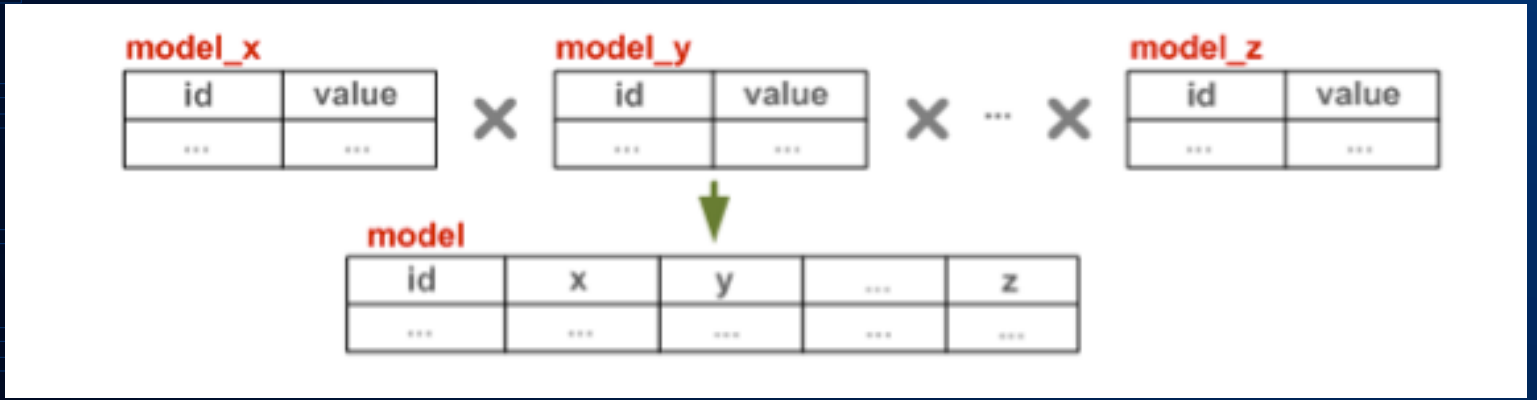
BIOTEAM
Enabling Science

Schema-Free

ID	Category	Subcat.	Name	Color	Size	Make	Model
Item_01	Clothes	Sweater	Cathair Sweater	Siamese	Small, Medium, Large		
Item_02	Clothes	Pants	Designer Jeans	Paisley Acid Wash	30x32, 32x32, 32x34		
Item_03	Clothes	Pants	Sweatpants	Blue, Yellow, Pink	Large		
Item_04	Car Parts	Engine	Turbos			Audi	S4
Item_05	Car Parts	Emissions	02 Sensor			Audi	S4
Item_06	Motorcycle Parts	Bodywork	Fender Eliminator	Blue		Yamaha	R1
Item_07	Motorcycle Parts, Clothing	Clothing	Leather Pants	Black	Small, Medium, Large		

http://aws.amazon.com/

Schema-Free MySQL



BIOTEAM
Enabling Science

Properties of Distributed Systems

- Design for failure
  - Disks will fail
  - When is RAID6 unacceptable (2019?)
  - Nodes will fail
- Must maintain data consistency
  - Is it useful otherwise?
- Network partitioning

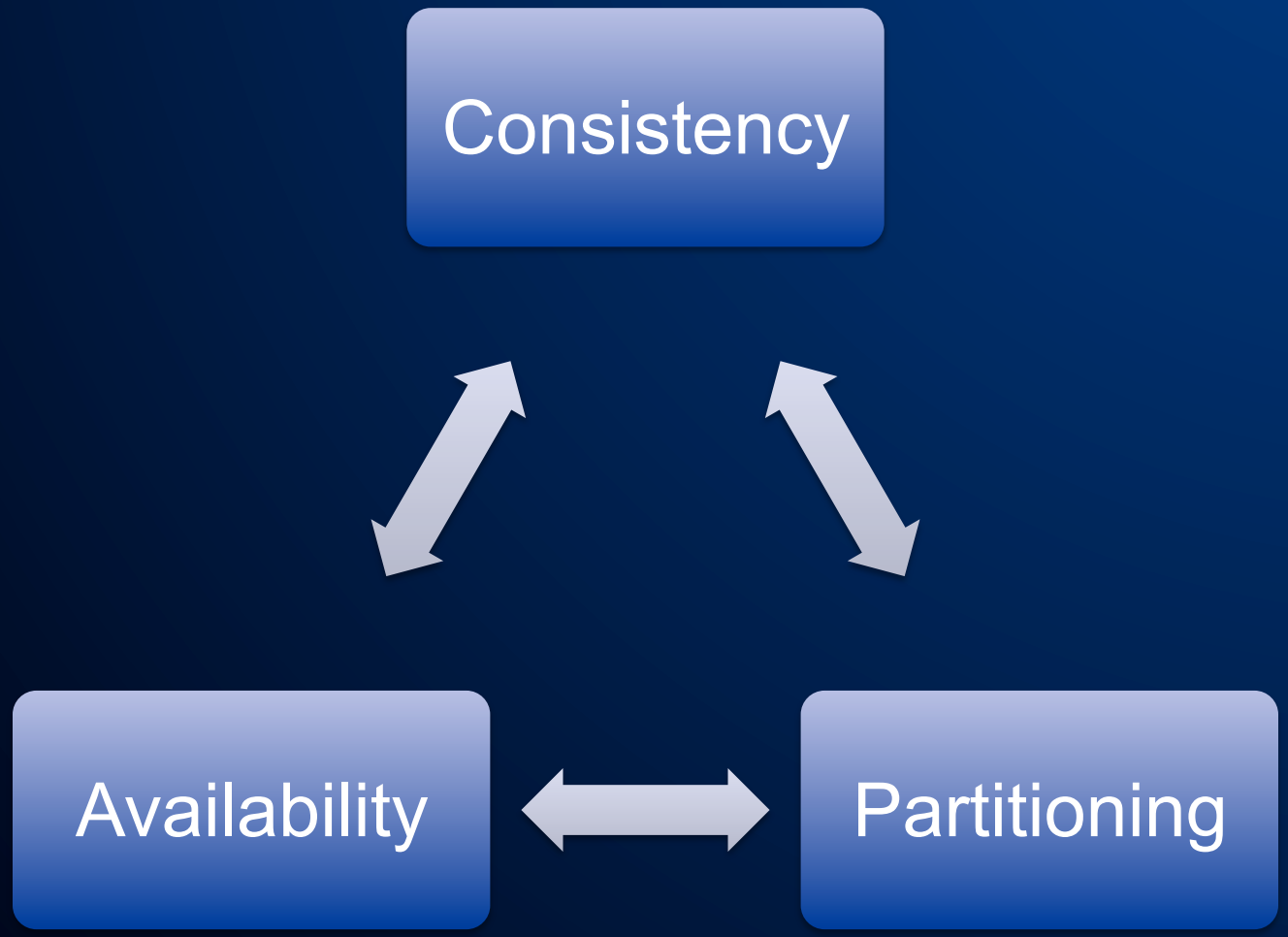
BIOTEAM
Enabling Science

Eventual Consistency

- “when no updates occur for a long period of time, eventually all updates will propagate through the system and all the replicas will be consistent”
- Eventually all clients will see the updates

BIOTEAM
Enabling Science

CAP Theorem

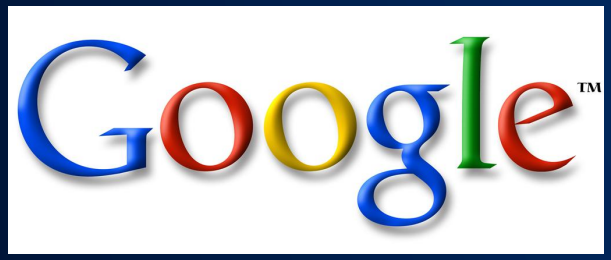


BIOTEAM
Enabling Science

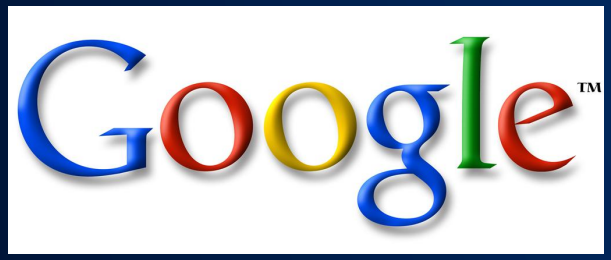
Balancing the tradeoffs…

- Three properties of shared-data systems
  - Consistency of the data
  - Availability of the system
  - Partition tolerance
- Only two can be achieved at any given time
- Network partitions are a given
- See also: Project Triangle
  - Good, Fast, Cheap… pick two.

BIOTEAM
Enabling Science



BIOTEAM
Enabling Science



Map Reduce

BigTable

Dynamo

Thrift

Cassandra

BIOTEAM
Enabling Science

Things to consider…

- Nested data structures
- Document-model
- BLOBS
- Natural partitions
- Client access patterns
- Eventual Consistency

BIOTEAM
Enabling Science

NoSQL in Practice

- Choose the right storage system for your data
- De-normalize your data
- No ACID guarantees
- Do JOINs in your application code
- Less well suited for
  - Highly-transactional systems
  - Traditional BI systems
  - Problems that require SQL

BIOTEAM
Enabling Science

Example: SNPs in SimpleDB

```
~/Projects/AWS-Training/SimpleDB-Demo-1 % ruby find_snps.rb
New Aws::SdbInterface using per_request-connection mode
Opening new HTTPS connection to sdb.amazonaws.com:443
Closing HTTPS connection to sdb.amazonaws.com:443
{"name"=>["rs2228714"], "seq5"=>["ATGAAACGAGTAGCGGCAGCAATAGTTCTCGAGAGGATATATCAAGTAGA"], "id"=>"2228714", "seq3"=>["GTTAGTAGTAGTATCGTTATACTA
ATTATGTTTTGGATCCCCCAAACTTA"], "observed"=>["A/G"]}
{"name"=>["rs2228728"], "seq5"=>["TAGTCTCAAAGGTCACAGCAGCAGCAGCAACAATCAGAGCAACAATGGAG"], "id"=>"2228728", "seq3"=>["TCCCGATAGCTACCAAAATATCTT
GATGATGCGTCACGGTGATCGCATTG"], "observed"=>["A/G"]}
{"name"=>["rs2228751"], "seq5"=>["GATTTTCTTTATTTAGCTTTTGAATTTCTTTTTTATTTTTTTAATCAACT"], "id"=>"2228751", "seq3"=>["GTATATCACATTTTCGGTGACATT
CACTGATGGCAGTCCAGATCAGTTTT"], "observed"=>["A/G"]}
{"name"=>["rs2228769"], "seq5"=>["TAAATGACTAAAACTACAATTAATAGCTGCAACTCCAAATGCGTGACCAA"], "id"=>"2228769", "seq3"=>["CTTGGATTGATCAAAAGAAAACAC
ATGAGAAGAAGATAAAGATTTAAGCG"], "observed"=>["A/G"]}
{"name"=>["rs2228789"], "seq5"=>["AATAACTATTAACTAATTTCAAAAATTATTAGTCCCTTTGAAGACCTCCT"], "id"=>"2228789", "seq3"=>["AATTATTTTTCCAAAACCGGAAAT
AATTTACGCCGGATAACTTAAAAATC"], "observed"=>["A/G"]}
{"name"=>["rs2228805"], "seq5"=>["AGAAAAGAGAGAGAGTCAAAGCAGTAATGAATTACCTGTGTATGAAATGT"], "id"=>"2228805", "seq3"=>["CCTGTGACATTTTATGAGGATATG
TGGCTATTATCCAGCGGTACCACGAA"], "observed"=>["A/G"]}
{"name"=>["rs2228811"], "seq5"=>["AATCTTGTTCGTTTGTCAATTTGTCTCGATATGTTTTTTTTATGGAGATT"], "id"=>"2228811", "seq3"=>["TCAACTTCAGAGATGGGAAAAGCT
AGCCATTAGTTTCCGATTTACGCTGA"], "observed"=>["A/G"]}
{"name"=>["rs2228822"], "seq5"=>["AGCTAGACCTTTGAGCACGGCGAGCTTGGTGAATTTGGCGATCAGTCATA"], "id"=>"2228822", "seq3"=>["TGCAACCTCTTCGTTATCAATAGA
GGATCTTCTAGCCCGAGATGAACTAT"], "observed"=>["A/G"]}
{"name"=>["rs2228823"], "seq5"=>["AGCTGTCCATCGGAATATGTCGAGAGCTAGATCTGGATCGGATTGAGCTC"], "id"=>"2228823", "seq3"=>["TCGGAGTTTGTGAATCCTGGTTTT
AGATTCTGAATCCATGTTTCGAATTG"], "observed"=>["A/G"]}
{"name"=>["rs2228824"], "seq5"=>["CTGGTTTTAGATTCTGAATCCATGTTTCGAATTGAGTTTCGAGTGGTGTT"], "id"=>"2228824", "seq3"=>["AGGGGAAGTGTGTGATTCCGGTGA
AGATGAGACGGAAGAGAATGATAGAGA"], "observed"=>["A/G"]}
{"name"=>["rs2228862"], "seq5"=>["GTAAAGTGTTTGGAGGTATACCTACTTTCCGTTAGTCGAATGAAAATAAA"], "id"=>"2228862", "seq3"=>["ATTACTGGTTCCACTAAACACCTA
ATCCATGCACCAACAACCAAGTGACT"], "observed"=>["A/G"]}
{"name"=>["rs2228870"], "seq5"=>["AAAACTCCAGGAAATCCCGCAATTTTCTGTAGACTTGATGAACTCTCTGT"], "id"=>"2228870", "seq3"=>["GTTGGTCGAGTAAAATGATGATTT
AAAAGTTTTTGGGTTTTAAACATCAC"], "observed"=>["A/G"]}
{"name"=>["rs2228881"], "seq5"=>["GGGAGTTGGAATCAGTTCGAGCCGATGAAGTTAAGTACCCAGAGAAGGCT"], "id"=>"2228881", "seq3"=>["CCAACCTGACCTGGTTGAACTCAC
TTCTCTCCTGGACGACACCCCTCCCTA"], "observed"=>["A/G"]}
{"name"=>["rs2228891"], "seq5"=>["ATATGAAACGAGTAGCGGCAGCAATAGTTCTCGAGAGGATATATCAAGTAGA"], "id"=>"2228891", "seq3"=>["GTTAGTAGTAGTATCGTTATAC
TAATTATGTTTTGGATCCCCCAAACTTA"], "observed"=>["A/G"]}
~/Projects/AWS-Training/SimpleDB-Demo-1 %
```

BIOTEAM
Enabling Science

Example: Storing short-reads

- 2.8 million unique reads loaded
- 0.5 million reads retrieved
- Single node system

Database	Load time	Retrieval time	File size
Tokyo Cabinet/Tyrant	12 minutes	3 1/2 minutes	24MB
CouchDB	5 1/2 minutes	14 1/2 minutes	236MB
MongoDB	3 minutes	4 minutes	192-960MB

http://bcbio.wordpress.com/

BIOTEAM
Enabling Science

NoSQL use cases

- Real-time analytics
  - Fast real-time inserts, updates, and queries
- Problems requiring high scalability
  - Tens or hundreds of servers
  - Replication/sharding built-in
- Persistent object store
  - Think persistent memcached
- Document or key-value oriented schemas
  - JSON-like data schemas

BIOTEAM
Enabling Science

That's it

- Thanks!

- kraut@bioteam.net
- www.twitter.com/adamkraut
- www.friendfeed.com/adamk
- blog.bleedingedgebiotech.com

BIOTEAM
Enabling Science