

# Project0

Shuyang Yao

## A. General Purpose:

System	Pros	Cons
1 Lotus Notes: <sup>1</sup>	<ul style="list-style-type: none"><li>♦ Better security control than its competitor</li><li>♦ Supports more platforms than its competitor</li><li>♦ Support hardware as well as software virtualization</li></ul>	<ul style="list-style-type: none"><li>♦ It's not very flexible in term of using because it's a little bit too secure</li></ul>
2 Actian Versant <sup>2</sup>	<ul style="list-style-type: none"><li>♦ It's Object Oriented! It's simple and straight forward to build complex data models which are hierarchical.</li><li>♦ No need to use a query language</li><li>♦ No primary keys</li><li>♦ Make developer's life easier by allowing database to process complex objects models without writing mappings.</li><li>♦ Allows you to change database schema while your service is on-line.</li><li>♦ Good Scalability</li><li>♦ Full support for transactions, logging and locking.</li></ul>	<ul style="list-style-type: none"><li>♦ Lack of solid theoretical basis.</li><li>♦ API is language dependent</li><li>♦ Schema change may need a system-wide compile(not in Versant's case)</li></ul>
3 InterSystems Cache <sup>3</sup>	<ul style="list-style-type: none"><li>♦ A fully persistent database with high throughput even comparing with in-memory database</li><li>♦ Good scalability and performance when hosted</li></ul>	<ul style="list-style-type: none"><li>♦</li></ul>

	on inexpensive machines.	
4 McObject:	<ul style="list-style-type: none"> <li>♦ Memory data base with good scalability.</li> <li>♦ ACID-Compliant</li> <li>♦ It's Object Oriented! It's simple and straight forward to build complex data models which are hierarchical.</li> <li>♦ No need to use a query language</li> <li>♦ No primary keys</li> </ul>	<ul style="list-style-type: none"> <li>♦ Lack of solid theoretical basis.</li> <li>♦ API is language dependent</li> <li>♦ Schema change may need a system-wide compile(not in Versant's case)</li> </ul>
5 ObjectStore:	<ul style="list-style-type: none"> <li>♦ It's Object Oriented! It's simple and straight forward to build complex data models which are hierarchical.</li> <li>♦ No need to use a query language</li> <li>♦ No primary keys</li> </ul>	<ul style="list-style-type: none"> <li>♦ Lack of solid theoretical basis.</li> <li>♦ API is language dependent</li> <li>♦ Schema change may need a system-wide compile(not in Versant's case)</li> </ul>
6 WakandaDB: <sup>4</sup>	<ul style="list-style-type: none"> <li>♦ Everything(schema, server-side processing, querying) can all be done in JavaScript</li> <li>♦ Open Source</li> </ul>	♦ Open Source and not mature
7 IBM IMS: <sup>5</sup>	<ul style="list-style-type: none"> <li>♦ It's a full function database</li> <li>♦ Optimized for high transection rates.</li> <li>♦ High Availability</li> <li>♦</li> </ul>	♦ Developers may need to write more code.
8 Adabas <sup>6</sup>	<ul style="list-style-type: none"> <li>♦ Is able to work close with previously existing database system</li> </ul>	♦ No access control in term of native network encryption
9 UniVerse: <sup>78</sup>	<ul style="list-style-type: none"> <li>♦ Multi-valued database</li> <li>♦ High Availability</li> <li>♦ Good Scalability</li> <li>♦ Intuitive database design</li> <li>♦ High Performance</li> <li>♦ Does not constrain by 1st Normal Form</li> </ul>	♦ The fact that it does not adhere to 1FN can be abused
10 UniData	<ul style="list-style-type: none"> <li>♦ Secure</li> <li>♦ Similar to UniVerse</li> </ul>	♦ Similar to UniVerse
11 Documentation xDB	<ul style="list-style-type: none"> <li>♦ XML-based Database</li> <li>♦ Allows the schema to be easily modified.</li> </ul>	♦ Not ACID-compliant

	<ul style="list-style-type: none"> <li>♦ Flexible schema compared to relational database</li> <li>♦ EMC^2's disaster-recovery options</li> </ul>	
12 Tamino XML Server	<ul style="list-style-type: none"> <li>♦ XML-based Database advantage described in Documentation xDB</li> </ul>	<ul style="list-style-type: none"> <li>♦ Not ACID-compliant like many other document based database</li> </ul>
13 Ipedo XML Database	<ul style="list-style-type: none"> <li>♦ XML-based Database advantage described in Documentation xDB</li> </ul>	<ul style="list-style-type: none"> <li>♦ Not ACID-compliant like many other document based database</li> </ul>
14 OrientDB: <sup>910</sup>	<ul style="list-style-type: none"> <li>♦ Document oriented database with graph database feature</li> <li>♦ Open Source</li> <li>♦ Can be queried using SQL</li> <li>♦ AICD</li> </ul>	<ul style="list-style-type: none"> <li>♦ Not Mature, API changes over time.</li> </ul>
15 SQLite	<ul style="list-style-type: none"> <li>♦ Easy to use</li> <li>♦ Consumes less resources</li> </ul>	<ul style="list-style-type: none"> <li>♦ Doesn't scale very well</li> </ul>
16 Firebird <sup>11</sup>	<ul style="list-style-type: none"> <li>♦ Free open source</li> <li>♦ Relational database</li> <li>♦ Well established and tested based on solid theoretical foundation</li> <li>♦ ACID</li> <li>♦ SQL as access language</li> <li>♦ Join!</li> <li>♦ Large Throughput</li> </ul>	<ul style="list-style-type: none"> <li>♦ Not very scalable in term of horizontal scaling</li> <li>♦ Difficult to model complex data model because it is table based</li> </ul>
17 SAP Sybase ASE: <sup>12</sup>	<ul style="list-style-type: none"> <li>♦ Relational database advantages as described in Firebird</li> </ul>	<ul style="list-style-type: none"> <li>♦ Relational database disadvantages described in Firebird</li> </ul>
18 SAP SQL Anywhere	<ul style="list-style-type: none"> <li>♦ Relational database advantages as described in Firebird</li> <li>♦ Embed: consume less resources.</li> <li>♦</li> </ul>	<ul style="list-style-type: none"> <li>♦ Relational database disadvantages described in Firebird</li> </ul>
19 Postgres-XL:	<ul style="list-style-type: none"> <li>♦ ACID</li> <li>♦ Open Source</li> <li>♦ Cluster-wide Consistency</li> <li>♦ Secure</li> <li>♦ SQL</li> <li>♦ Horizontal scaling</li> <li>♦ Rich feature set.</li> </ul>	<ul style="list-style-type: none"> <li>♦ Too complex for simple stuff</li> </ul>
20 Pecona	<ul style="list-style-type: none"> <li>♦ Same as MySql</li> </ul>	<ul style="list-style-type: none"> <li>♦ Same as MySql</li> </ul>

21 MySQL	<ul style="list-style-type: none"> <li>♦ Relational database advantages as described in Firebird</li> </ul>	<ul style="list-style-type: none"> <li>♦ Not very Scalable</li> </ul>
22 Oracle Database	<ul style="list-style-type: none"> <li>♦ Relational database advantages as described in Firebird</li> <li>♦ Commercial</li> </ul>	<ul style="list-style-type: none"> <li>♦ Price</li> <li>♦ Relational database disadvantages as described in Firebird</li> </ul>
23 IBM DB2	<ul style="list-style-type: none"> <li>♦ Relational database advantages as described in Firebird</li> <li>♦ Commercial</li> </ul>	<ul style="list-style-type: none"> <li>♦ Price</li> <li>♦ Relational database disadvantages as described in Firebird</li> </ul>

## B. Specialist analytic

System	Pros	Cons
25. Google BigQuery	<ul style="list-style-type: none"> <li>♦ Allows users to use SQL-like queries to query massive datasets</li> <li>♦ Rest API</li> <li>♦ Google's infrastructure makes the operation fast and economic.</li> <li>♦ Makes ad-hoc and trial-and-error interactive query on large dataset possible</li> </ul>	<ul style="list-style-type: none"> <li>♦ Not able to do complex data processing.</li> <li>♦ You can't update your data, only appending is allowed.</li> <li>♦ No large Join.</li> </ul>
26. InfluxDB <sup>13</sup>	<ul style="list-style-type: none"> <li>♦ Open Source</li> <li>♦ SQL-like language</li> <li>♦ Native HTTP API</li> <li>♦ Can process big amount of data.</li> <li>♦ Focus on time series.</li> </ul>	<ul style="list-style-type: none"> <li>♦ Maturity</li> </ul>
27. 1010data	<ul style="list-style-type: none"> <li>♦ Interactive analytical service.</li> <li>♦ A rich set of analytic functions are integrated.</li> <li>♦ Suit for very large data set.</li> <li>♦ Decent performance with high scalability.</li> </ul>	<ul style="list-style-type: none"> <li>♦ Price.</li> <li>♦ Should support more complex database operations.</li> </ul>
28. BitYota: <sup>14</sup>	<ul style="list-style-type: none"> <li>♦ Data warehouse as a service</li> <li>♦ SQL</li> </ul>	<ul style="list-style-type: none"> <li>♦</li> </ul>

	<ul style="list-style-type: none"> <li>♦ Real time data analysis</li> </ul>	
29. AWS RedShift <sup>15</sup>	<ul style="list-style-type: none"> <li>♦ Scaling is easy</li> <li>♦ You can use SQL to query</li> <li>♦ Can work seamlessly with other AWS services</li> </ul>	<ul style="list-style-type: none"> <li>♦ Because RedShift does not maintain the uniqueness of data, programmer are responsible for their data integrity.</li> </ul>
30. SpaceCurve: <sup>16</sup>	<ul style="list-style-type: none"> <li>♦ Focus mainly on spatial data.</li> <li>♦ Good at real-time location analysis</li> </ul>	<ul style="list-style-type: none"> <li>♦ Query: Some patent-pending strategies that optimize queries</li> <li>♦ View: Algorithm that make updating views more efficient</li> <li>♦ Supports serializable concurrent transaction</li> </ul>
31. LogicBlox:	<ul style="list-style-type: none"> <li>♦ Query: Some patent-pending strategies that optimize queries</li> <li>♦ View: Algorithm that make updating views more efficient</li> <li>♦ Supports serializable concurrent transaction</li> </ul>	<ul style="list-style-type: none"> <li>♦</li> </ul>
32. MonetDB: <sup>17</sup>	<ul style="list-style-type: none"> <li>♦ Column oriented store:</li> <li>♦ Good at OLAP scenario</li> <li>♦ Highly Compressed</li> </ul>	<ul style="list-style-type: none"> <li>♦ Increased disk seek time</li> <li>♦ Insertion costs more</li> </ul>
33. Pivotal GreenPlum: <sup>18</sup>	<ul style="list-style-type: none"> <li>♦ Supports both row and column-oriented storage</li> <li>♦ Highly scalable</li> </ul>	<ul style="list-style-type: none"> <li>♦ Column oriented store disadvantage</li> <li>♦ Increased disk seek time</li> <li>♦ Insertion costs more</li> <li>♦</li> </ul>
34. HP Verica	<ul style="list-style-type: none"> <li>♦ Column Oriented</li> <li>♦ Highly compressed</li> <li>♦ Good at log parsing</li> </ul>	<ul style="list-style-type: none"> <li>♦ Immaturity</li> <li>♦ Increased disk seek time</li> <li>♦ Insertion costs more</li> </ul>
35. SAP Sybase IQ <sup>19</sup>	<ul style="list-style-type: none"> <li>♦ Column Oriented</li> </ul>	<ul style="list-style-type: none"> <li>♦ Increased disk seek time</li> <li>♦ Insertion costs more</li> </ul>
36. ParStream: <sup>20</sup>	<ul style="list-style-type: none"> <li>♦ Real time analysis</li> <li>♦ Focus on IOT(Internet of Things) data</li> </ul>	
37. IBM InfoSphere	<ul style="list-style-type: none"> <li>♦ Real-time analytic platform</li> <li>♦ Merge diverse data</li> </ul>	<ul style="list-style-type: none"> <li>♦ Rapid deployment analysis</li> <li>♦ Column-oriented database</li> <li>♦ In memory</li> <li>♦ Highly compressed</li> </ul>
38. Kx Systems: <sup>21</sup>	<ul style="list-style-type: none"> <li>♦ Column store database advantages described in MonetDB</li> </ul>	<ul style="list-style-type: none"> <li>♦ Column store database disadvantages described in MonetDB</li> </ul>
39. LucidDB: <sup>22</sup>	<ul style="list-style-type: none"> <li>♦ Column store database</li> </ul>	<ul style="list-style-type: none"> <li>♦ Column store database</li> </ul>

	advantages described in MonetDB <ul style="list-style-type: none"> <li>♦ Open Source</li> <li>♦ Bitmap indexing</li> <li>♦ Hash join/aggregation</li> <li>♦ Multiversioning</li> </ul>	disadvantages described in MonetDB
40. Kognitio <sup>23</sup> :	<ul style="list-style-type: none"> <li>♦ In memory</li> <li>♦ Support SQL</li> <li>♦ integrated with HADOOP</li> </ul>	♦
41. Actian Vector	<ul style="list-style-type: none"> <li>♦ A high performance analytic frame built on Hadoop</li> <li>♦ Developer can use SQL to interact with the system</li> </ul>	♦
42. MetaMarkets Druid: <sup>24</sup>	<ul style="list-style-type: none"> <li>♦ A distributed real-time data store</li> <li>♦ Real time ingestion</li> <li>♦ Column-oriented storage's advantage</li> <li>♦ Bitmap indexing</li> <li>♦ Fault tolerance</li> </ul>	♦ Column-oriented storage's disadvantage
43. Teradata <sup>2526</sup>	<ul style="list-style-type: none"> <li>♦ A decent data warehouse system</li> <li>♦ Developers can choose to store the data either based on row or column</li> </ul>	♦ Price
44. SQream	<ul style="list-style-type: none"> <li>♦ Scalable SQL data base</li> <li>♦ GPU based database brings high parallel processing ability</li> <li>♦ Column oriented storage advantages</li> </ul>	♦ Column oriented storage disadvantages
45. RainStor	<ul style="list-style-type: none"> <li>♦ Can work with different data types</li> </ul>	♦
46. HPCC <sup>27</sup>	<ul style="list-style-type: none"> <li>♦ Introduced a new programming language: ECL</li> <li>♦ It is more complex than a key-value pair storage.</li> <li>♦ High availability, scalability and consistent</li> </ul>	♦ Still growing.
47. Teradata Aster: <sup>28</sup>	<ul style="list-style-type: none"> <li>♦ Allows users to write map reduce code that manipulate relational data base data.</li> </ul>	♦

	<ul style="list-style-type: none"> <li>♦ Graph analytics engine</li> <li>♦ Support massive parallel processing</li> </ul>	
48. SciDB <sup>29</sup>	<ul style="list-style-type: none"> <li>♦ Array data model</li> <li>♦ Supports complex mathematic processing on the arrays</li> <li>♦ Can model uncertainty</li> </ul>	♦ Focus mainly on Mathematic operations
49. Hadapt <sup>30</sup>	<ul style="list-style-type: none"> <li>♦ Brings SQL to Hadoop, which allows users to write SQL to query on massive amount of data</li> <li>♦ Uses a hybrid storage engine which stores structured data in a traditional relational database while unstructured data in HDFS.</li> </ul>	♦ No transections
50. JethroData <sup>31</sup>	<ul style="list-style-type: none"> <li>♦ Like Hadapt, it builds a layer on top of Hadoop that allows user to write SQL on Hadoop</li> <li>♦ Unique index strategy</li> <li>♦ Scalability that comes with HDFS</li> </ul>	♦ No transections
51. CitusDB: <sup>32</sup>	<ul style="list-style-type: none"> <li>♦ SQL on Hadoop</li> <li>♦ Also support semi-structured data</li> <li>♦ Optimized specially for time-series data.</li> </ul>	♦ No transections
52. Impala: <sup>3334</sup>	<ul style="list-style-type: none"> <li>♦ SQL on Hadoop</li> <li>♦ It's supported by Cloudera</li> </ul>	<ul style="list-style-type: none"> <li>♦ No transections</li> <li>♦ Data need to be stored in a specific data format</li> </ul>
53. IBM Big SQL <sup>35</sup>	<ul style="list-style-type: none"> <li>♦ SQL on Hadoop</li> </ul>	♦ No transections
54. Presto	<ul style="list-style-type: none"> <li>♦ Sql on Hadoop</li> <li>♦ Can query data from different source and bring them together</li> </ul>	♦ No transections
55. Apache Drill: <sup>36</sup>	<ul style="list-style-type: none"> <li>♦ SQL on Hadoop</li> <li>♦ Apache license</li> <li>♦ Can work with semi-structured or nested data</li> <li>♦ Low latency</li> </ul>	♦ No transections
56. Apache Hive: <sup>3738</sup>	<ul style="list-style-type: none"> <li>♦ Data warehouse built on</li> </ul>	♦ No update and delete

	Hadoop <ul style="list-style-type: none"> <li>♦ Use a SQL like language</li> <li>♦ Bitmap index</li> <li>♦ Supports different storage type.</li> <li>♦ Data are compressed</li> <li>♦</li> </ul>	operation <ul style="list-style-type: none"> <li>♦ No access control</li> <li>♦ The overhead brought by Map Reduce make it a little bit slow</li> </ul>
57. Apache Tajo <sup>39</sup>	<ul style="list-style-type: none"> <li>♦ Fully distributed SQL</li> <li>♦ Various query optimization</li> <li>♦ Supports ANSI/ISO SQL</li> <li>♦ Has a shell</li> </ul>	<ul style="list-style-type: none"> <li>♦ No transections</li> </ul>

## C. Big Tables

System	♦ Pros	♦ Cons
58. Google Cloud Datastore <sup>4041</sup>	<ul style="list-style-type: none"> <li>♦ No schema is needed, Aims at storing non-relational data</li> <li>♦ Write scale and read scale.</li> <li>♦ Supports transection</li> </ul>	<ul style="list-style-type: none"> <li>♦ No database layer caching</li> <li>♦ No Join</li> <li>♦ Filter results using a subquery is not supported</li> </ul>
59. Google App Engine Datastore <sup>42</sup>	<ul style="list-style-type: none"> <li>♦ Key-value pair store makes it more flexible</li> </ul>	<ul style="list-style-type: none"> <li>♦ Does not support ACID transactions</li> <li>♦ No join</li> </ul>
60. Cassandra.io <sup>4344</sup>	<ul style="list-style-type: none"> <li>♦ Linear scalability(All nodes are identical)</li> <li>♦ Fault-tolerance on inexpensive hardware</li> <li>♦ The language it uses(CQL3) is very similar to SQL</li> <li>♦ Constant-time writes</li> <li>♦ Integrated with Hadoop</li> </ul>	<ul style="list-style-type: none"> <li>♦ No Join</li> <li>♦ Does not support ACID</li> </ul>
61. Accumulo <sup>4546</sup>	<ul style="list-style-type: none"> <li>♦ Wode Column Store DB similar to Cassandra and HBase</li> <li>♦ Better Performance(can scan 800k entries per second per node) compare to HBase</li> <li>♦ Provides cell-level security</li> </ul>	<ul style="list-style-type: none"> <li>♦</li> </ul>
62. Hbase <sup>47</sup>	<ul style="list-style-type: none"> <li>♦ Works hand in hand with Hadoop</li> </ul>	<ul style="list-style-type: none"> <li>♦ No strict ACID</li> <li>♦ Because of its master and</li> </ul>



	<ul style="list-style-type: none"> <li>♦ Specially optimized for real time analysis</li> <li>♦ Also linear scalability</li> <li>♦ Consistent reads and writes</li> <li>♦ Row level Atomic</li> </ul>	slave architecture, Hbase has the problem of single point failure ♦ No join
63. HyperTable <sup>48</sup>	<ul style="list-style-type: none"> <li>♦ Implements using c++</li> <li>♦ Runs on haddop</li> <li>♦ SQL like language</li> <li>♦ Faster and smaller than HBase</li> </ul>	♦
64. DataStax Enterprise <sup>49</sup>	<ul style="list-style-type: none"> <li>♦ Built on Cassandra</li> <li>♦ Comercial</li> <li>♦ Similar to Cassandra</li> </ul>	♦ Similar to Cassandra

## D. Key value stores

System	Pros	Cons
65. AWS DynamoDB <sup>50</sup> <sup>51</sup>	<ul style="list-style-type: none"> <li>♦ Supports both document and key-value data</li> <li>♦ Low latency</li> <li>♦ Highly scalable</li> <li>♦ Highly available</li> <li>♦ Strong consistency on read</li> <li>♦ Supports atomic counters</li> <li>♦ Secure: fine access control</li> </ul>	<ul style="list-style-type: none"> <li>♦ Consume more resource because its stronger Consistent constrain</li> <li>♦ No join</li> <li>♦ No support for transection</li> <li>♦ No ACID</li> </ul>
66. AWS SimpleDB <sup>52</sup> <sup>53</sup>	<ul style="list-style-type: none"> <li>♦ Fit for smaller workloads</li> <li>♦ Automatically index all things</li> </ul>	<ul style="list-style-type: none"> <li>♦ Table cannot grow over 10 GB</li> <li>♦ Not as scalable as DynamoDB</li> <li>♦ No joins</li> </ul>
67. MagnetDB <sup>54</sup>	<ul style="list-style-type: none"> <li>♦ A key-value storage for open stack</li> <li>♦ Highly scalable</li> <li>♦ Supports both eventual and strong consistency reads</li> <li>♦ Fault tolerance</li> </ul>	<ul style="list-style-type: none"> <li>♦ No join</li> <li>♦ No support for transection</li> <li>♦ No ACID</li> </ul>
68. Redis Cloud <sup>55</sup>	<ul style="list-style-type: none"> <li>♦ In memory non-relational database</li> <li>♦ Scale out seamlessly</li> <li>♦ Zero Down time</li> <li>♦ Secure</li> </ul>	♦ No ACID
69. Redis Labs <sup>56</sup>	♦ Similar to Redis Cloud	♦ Similar to Redis Cloud
70. AWS ElastiCache <sup>57</sup>	♦ You can choose from two in memory cache options: Redos or Memcached	♦ Disadvantages are similar to Redis

	<ul style="list-style-type: none"> <li>♦ The advantages are similar to those two</li> </ul>	
71. Redis-to-go	<ul style="list-style-type: none"> <li>♦ A redis management tool</li> </ul>	<ul style="list-style-type: none"> <li>♦ Redis' disadvantages</li> </ul>
72. RedisGreen <sup>58</sup>	<ul style="list-style-type: none"> <li>♦ A redis hosting service</li> </ul>	<ul style="list-style-type: none"> <li>♦ Redis' disadvantages</li> </ul>
73. ObjectRocket Redis <sup>59</sup>	<ul style="list-style-type: none"> <li>♦ A redis hosting service</li> </ul>	<ul style="list-style-type: none"> <li>♦ Redis' disadvantages</li> </ul>
74. HyperDex <sup>60</sup>	<ul style="list-style-type: none"> <li>♦ Key-value storage</li> <li>♦ Strong consistency</li> <li>♦ Fault tolerance</li> <li>♦ ACID</li> </ul>	<ul style="list-style-type: none"> <li>♦ No Join</li> </ul>
75. LevelDB <sup>61</sup>	<ul style="list-style-type: none"> <li>♦ Key-value</li> <li>♦ Comparison function can be customized</li> <li>♦ Compressed</li> </ul>	<ul style="list-style-type: none"> <li>♦ No indexes</li> <li>♦ It only allows one process to access the database at a time</li> </ul>
76. BerkeleyDB <sup>62</sup>	<ul style="list-style-type: none"> <li>♦ Provides building blocks that can help you develop your own data management solution</li> </ul>	<ul style="list-style-type: none"> <li>♦</li> </ul>
77. Oracle NoSQL <sup>63</sup>	<ul style="list-style-type: none"> <li>♦ Key value storage with secondary indexes</li> <li>♦ ACID</li> <li>♦ Secure</li> </ul>	<ul style="list-style-type: none"> <li>♦ No Join</li> </ul>
78. Voldemort	<ul style="list-style-type: none"> <li>♦</li> </ul>	<ul style="list-style-type: none"> <li>♦</li> </ul>
79. Redis <sup>64</sup>	<ul style="list-style-type: none"> <li>♦ In memory non-relational database</li> <li>♦ Scale out seamlessly</li> <li>♦ Zero Down time</li> <li>♦ Secure</li> </ul>	<ul style="list-style-type: none"> <li>♦ No ACID</li> </ul>
80. Couchbase <sup>65</sup>	<ul style="list-style-type: none"> <li>♦ Key-value</li> <li>♦ Document(Json)</li> <li>♦</li> </ul>	<ul style="list-style-type: none"> <li>♦ No join</li> <li>♦ No transaction</li> </ul>
81. FatDB	<ul style="list-style-type: none"> <li>♦ Tight integration with SQL Server</li> </ul>	<ul style="list-style-type: none"> <li>♦</li> </ul>
82. Riak <sup>66</sup>	<ul style="list-style-type: none"> <li>♦ Buck key together</li> <li>♦ Strongly consistent</li> <li>♦ Non-key based query use map reduce to get the answer</li> </ul>	<ul style="list-style-type: none"> <li>♦ ACID</li> <li>♦ Join</li> </ul>
83. ArangoDB <sup>67</sup>	<ul style="list-style-type: none"> <li>♦ Multi-model database: Support documents, graphs and key-values data model</li> <li>♦ SQL-like</li> <li>♦ Joins like operation</li> <li>♦ Transactions</li> </ul>	<ul style="list-style-type: none"> <li>♦</li> </ul>

84. Aerospike <sup>68</sup>	<ul style="list-style-type: none"> <li>♦ Handle real time data</li> <li>♦ ACID</li> <li>♦ Flash as storage</li> <li>♦ Mainly key-value</li> <li>♦ Map reduce</li> </ul>	♦ Join
-----------------------------	---	--------

## E. Hadoop

System	Pros	Cons
85. GridGain <sup>69</sup>	<ul style="list-style-type: none"> <li>♦ In Memory data fabric</li> <li>♦ Can act as a cache layer to accelerate Hadoop</li> <li>♦ Realtime streaming</li> <li>♦ Linearly scale out</li> <li>♦ ACID transaction</li> </ul>	♦ history
86. ScaleOut <sup>70</sup> Software	♦ In memory storage	♦
87. Pivotal GenFire XD <sup>71</sup>	<ul style="list-style-type: none"> <li>♦ Helps SQL to scale out using Hadoop</li> <li>♦ High availability</li> <li>♦ In memory</li> <li>♦</li> </ul>	♦ Transaction and ACID
88. Sqrrl Enterprise	<ul style="list-style-type: none"> <li>♦ Based on Apache Accumulo</li> <li>♦ Advantages similar to Accumulo</li> </ul>	♦ disadvantages similar to Accumulo
89. LucidWorks Big Data <sup>72</sup>	♦ A big data platform brings together Hadoop solr and etc.	♦
90. Trafodion <sup>73</sup>	<ul style="list-style-type: none"> <li>♦ SQL on Hbase</li> <li>♦ ACID Transaction</li> <li>♦ scaling</li> </ul>	♦
91. Splice Machine <sup>74</sup>	<ul style="list-style-type: none"> <li>♦ Full function SQL on Hadoop</li> <li>♦ Scale out</li> <li>♦ Transaction</li> <li>♦ High concurrency</li> <li>♦ Real time updates</li> </ul>	♦
92. Apache Tajo - Pivotal HD	♦ Already described in Specialist analytic	♦ Already described in Specialist analytic

## F. Appliance

System	Pros	Cons
93. Oracle Big Data Appliance <sup>75</sup>	<ul style="list-style-type: none"> <li>♦ Cloudera Enterprise Technology software</li> <li>♦ Oracle NoSQL database</li> <li>♦ Integrated solution</li> </ul>	<ul style="list-style-type: none"> <li>♦ Integrated solution</li> <li>♦ price</li> </ul>
94. Oracle Exalytics <sup>76</sup>	<ul style="list-style-type: none"> <li>♦ In memory integrated solution</li> </ul>	<ul style="list-style-type: none"> <li>♦ Integrated solution</li> <li>♦ price</li> </ul>
95. Microsoft SQL Server PDW	<ul style="list-style-type: none"> <li>♦ Integrated solution</li> </ul>	<ul style="list-style-type: none"> <li>♦ Integrated solution</li> </ul>
96. SAP HANA <sup>77</sup>	<ul style="list-style-type: none"> <li>♦ In memory</li> <li>♦ Column oriented</li> <li>♦ Relational database</li> </ul>	<ul style="list-style-type: none"> <li>♦ Increased disk seek time</li> <li>♦ Insertion costs more</li> </ul>
97. IBM Pure Data	<ul style="list-style-type: none"> <li>♦ Integrated solution</li> </ul>	<ul style="list-style-type: none"> <li>♦ Integrated solution</li> </ul>
98. Oracle Exadata	<ul style="list-style-type: none"> <li>♦ Integrated solution</li> <li>♦ flash</li> </ul>	<ul style="list-style-type: none"> <li>♦ Integrated solution</li> </ul>

## G. Graph

System	Pros	Cons
99. infiniteGraph <sup>78</sup>	<ul style="list-style-type: none"> <li>♦ Distributed graph database</li> <li>♦ graph specific queries</li> <li>♦ Policy-driven consistent</li> <li>♦ Data visualization is integrated</li> </ul>	<ul style="list-style-type: none"> <li>♦ Does not support map reduce</li> <li>♦ Does not support data compression</li> <li>♦ Eventual consistency</li> </ul>
100. HypergraphDB <sup>79</sup>	<ul style="list-style-type: none"> <li>♦ Graph oriented</li> <li>♦ graph specific queries</li> <li>♦ Transaction</li> </ul>	<ul style="list-style-type: none"> <li>♦ Does not support map reduce</li> </ul>
101. Allegrograph <sup>80</sup>	<ul style="list-style-type: none"> <li>♦ ACID transaction</li> <li>♦ Automatic indexing</li> <li>♦ SOLR and MongoDB are integrated</li> <li>♦ Secure</li> <li>♦ Sharding</li> </ul>	<ul style="list-style-type: none"> <li>♦ Does not support map reduce</li> </ul>
102. Giraph <sup>81</sup>	<ul style="list-style-type: none"> <li>♦ Data analysis tool on graph data</li> <li>♦ Apache</li> <li>♦ Used by Facebook</li> <li>♦ Runs as map reduce jobs</li> </ul>	<ul style="list-style-type: none"> <li>♦</li> </ul>
103. SPARQLBASE <sup>82</sup>	<ul style="list-style-type: none"> <li>♦ Graph database</li> <li>♦ In memory</li> </ul>	<ul style="list-style-type: none"> <li>♦</li> </ul>

	<ul style="list-style-type: none"> <li>♦ Uses HDLC to store data</li> </ul>	
104. Trinity <sup>83</sup>	<ul style="list-style-type: none"> <li>♦ Embed or distributed graph storage</li> <li>♦ In memory</li> <li>♦ Data compressed</li> </ul>	♦ Not mature
105. Titan <sup>84</sup>	<ul style="list-style-type: none"> <li>♦ Distributed graph database</li> <li>♦ Support Transaction and eventual consistency</li> <li>♦ Can use Cassandra, HBase or BerkeleyDB to store data.</li> <li>♦ Support geo, numeric and text search</li> <li>♦ Map reduce</li> </ul>	♦
106. Objectivity: <sup>85</sup>	<ul style="list-style-type: none"> <li>♦ Graph NoSQL database</li> <li>♦ Good at exploring relationships in data.</li> <li>♦ Suits for areas like social networks.</li> </ul>	♦
107. Stardog	<ul style="list-style-type: none"> <li>♦ Graph database</li> <li>♦ ACID</li> </ul>	♦
108. FlockDB <sup>86</sup>	<ul style="list-style-type: none"> <li>♦ Graph database</li> <li>♦ Twitter uses it to store social graphs</li> <li>♦ Designs for websites</li> </ul>	♦ Fewer function cause it's simpler(maybe it's an advantage)
109. GrapheneDB	<ul style="list-style-type: none"> <li>♦ Cloud hosting Neo4j</li> </ul>	♦ Same as Neo4j
110. Sparksee <sup>87</sup>	<ul style="list-style-type: none"> <li>♦ Data compression(use bitmap to represent data)</li> </ul>	♦
111. Neo4j <sup>88</sup>	<ul style="list-style-type: none"> <li>♦ High Availability</li> <li>♦ Data compression</li> <li>♦ Fully ACID</li> <li>♦</li> </ul>	<ul style="list-style-type: none"> <li>♦ Does not support map reduce</li> <li>♦ Has Max size value limitation</li> </ul>
112. CortexDB <sup>89</sup>	<ul style="list-style-type: none"> <li>♦ Multiple data model: key-value, graph, multi value column</li> <li>♦ Distributed</li> </ul>	♦

## H. Data Caching

System	Pros	Cons
113. MemCachier <sup>90</sup>	<ul style="list-style-type: none"> <li>♦ In memory scalable key value pair cache</li> <li>♦ Better reliability and usability than memcached</li> </ul>	♦ No ACID

114. Redis	<ul style="list-style-type: none"> <li>♦ In memory non-relational database</li> <li>♦ Scale out seamlessly</li> <li>♦ Zero Down time</li> <li>♦ Secure</li> <li>♦ Persistent</li> </ul>	♦ No ACID
115. Redis Labs Memcached Cloud <sup>91</sup>	<ul style="list-style-type: none"> <li>♦ Cloud hosting Memcached</li> <li>♦ Similar as Memcached</li> </ul>	♦ Similar as Memcached
116. IronCache <sup>92</sup>	<ul style="list-style-type: none"> <li>♦ Key value cache</li> <li>♦ Cloud service</li> <li>♦ Can persist the data</li> </ul>	♦ No ACID
117. AWS ElastiCache	<ul style="list-style-type: none"> <li>♦ You can choose from two in memory cache options: Redis or Memcached</li> <li>♦ The advantages are similar to those two</li> </ul>	♦ Disadvantages are similar to Redis and Memcached
118. BigMemory <sup>93</sup>	<ul style="list-style-type: none"> <li>♦ In memory data store</li> <li>♦ Supports SQL</li> <li>♦ Runs Ehcache</li> </ul>	♦
119. Ehcache <sup>94</sup>	<ul style="list-style-type: none"> <li>♦ In memory data store</li> <li>♦ Schema less</li> <li>♦ ACID</li> <li>♦ Sharding and replication</li> </ul>	♦
120. InfiniSpan <sup>95</sup>	<ul style="list-style-type: none"> <li>♦ In memory key value data store</li> <li>♦ Support Map reduce</li> <li>♦ Support data compression</li> <li>♦ Support full text search, and graph data</li> <li>♦ Persistent</li> <li>♦ ACID transaction</li> </ul>	♦
121. RedHat JBoss Data Grid <sup>96</sup>	<ul style="list-style-type: none"> <li>♦ In memory distributed caching</li> <li>♦ Support map reduce</li> <li>♦ Supports replication</li> <li>♦ Transaction</li> <li>♦ Redhat support</li> </ul>	♦
122. Memcached <sup>97</sup>	<ul style="list-style-type: none"> <li>♦ In memory key value pair cache</li> <li>♦ Simpler than Redis makes it easier to scale out</li> <li>♦ ACID</li> </ul>	♦ Value is limited to 1MB

## I. Document

System	Pros	Cons
123. Informix <sup>98</sup>	<ul style="list-style-type: none"> <li>♦ Real time processing</li> <li>♦ Availability: zero down time</li> <li>♦ Supports SQL, JSON, and even time/special data</li> <li>♦ Support Rest API</li> </ul>	<ul style="list-style-type: none"> <li>♦ Comercial</li> </ul>
124. JumboDB <sup>99</sup>	<ul style="list-style-type: none"> <li>♦ Supports indexing on Json</li> <li>♦ Supports data compression</li> <li>♦ Supports complex data model</li> </ul>	<ul style="list-style-type: none"> <li>♦ No sharding and replication yet</li> <li>♦ Immaturity</li> <li>♦ Join</li> <li>♦ No ACID Transection</li> </ul>
125. RethinkDB <sup>100</sup>	<ul style="list-style-type: none"> <li>♦ Use Json as storage</li> <li>♦ Supports complex data model</li> <li>♦ Sharding and replication</li> <li>♦ Fault tolerance</li> <li>♦ MapReduce</li> <li>♦ Schema-less</li> </ul>	<ul style="list-style-type: none"> <li>♦ No Join</li> <li>♦ No ACID Transection</li> </ul>
126. CouchDB <sup>101</sup>	<ul style="list-style-type: none"> <li>♦ JSON as storage</li> <li>♦ Supports features that important to web development such as real time change notification</li> <li>♦ Supports complex data model</li> <li>♦ MapReduce</li> <li>♦ Eventual consistency</li> <li>♦ Schema-less</li> </ul>	<ul style="list-style-type: none"> <li>♦ No ACID Transection</li> <li>♦ No join</li> </ul>
127. RavenDB <sup>102</sup>	<ul style="list-style-type: none"> <li>♦ Schema-less</li> <li>♦ Data compression</li> <li>♦ ACID</li> <li>♦ MapReduce</li> </ul>	<ul style="list-style-type: none"> <li>♦</li> </ul>
128. TokuMX <sup>103</sup>	<ul style="list-style-type: none"> <li>♦ A high performance distribution of MongoDB</li> <li>♦ Better caching strategy</li> <li>♦ Optimized IO</li> <li>♦ Supports document-level locking allows better concurrency</li> </ul>	<ul style="list-style-type: none"> <li>♦ No ACID Transection</li> </ul>
129. MongoDB	<ul style="list-style-type: none"> <li>♦ Use Json as storage</li> <li>♦ Supports complex data</li> </ul>	<ul style="list-style-type: none"> <li>♦ No ACID</li> <li>♦ No Join</li> </ul>

	<ul style="list-style-type: none"> <li>♦ Supports immediate and strong consistency</li> <li>♦ Supports Sharding and replication</li> <li>♦ Schema-less</li> </ul>	
130. Compose	<ul style="list-style-type: none"> <li>♦ Cloud hosting mongodb</li> <li>♦ Similar as mongodb</li> </ul>	♦ Similar as mongo db
131. Iris Couch	<ul style="list-style-type: none"> <li>♦ Cloud hosting CouchDB</li> <li>♦ Similar as CouchDB</li> </ul>	♦ Similar as CouchDB
132. MongoLab	<ul style="list-style-type: none"> <li>♦ Cloud hosting mongodb</li> <li>♦ Similar as mongodb</li> </ul>	♦ Similar as mongo db
133. Object Rocket	<ul style="list-style-type: none"> <li>♦ Cloud hosting mongodb and redis</li> <li>♦ Similar as mongodb</li> </ul>	♦ Similar as mongodb
134. Azure DocumentDB	<ul style="list-style-type: none"> <li>♦ Use Json as storage</li> <li>♦ Supports complex data model</li> <li>♦ Schema-free</li> <li>♦ Supports different level of consistency</li> <li>♦ Transaction</li> </ul>	♦ No Join
135. Cloudant	<ul style="list-style-type: none"> <li>♦ Use Json as storage</li> <li>♦ Supports complex data model</li> <li>♦ Schema-free</li> <li>♦ Supports Full-text search</li> <li>♦ Supports spatial indexes</li> <li>♦ Data compression</li> </ul>	♦ No ACID

To be continued...

<sup>1</sup> Comparative Analysis of Microsoft Exchange and Lotus Notes

<http://www.brighthub.com/computing/windows-platform/articles/52715.aspx>

<sup>2</sup> Versant Object Database <http://www.actian.com/products/operational-databases/versant/>

<sup>3</sup> InterSystems Caché Benchmark [http://www.intersystems.com/assets/Cache\\_benchmark-1c69bf617b51d5a2dee145442deaa371.pdf](http://www.intersystems.com/assets/Cache_benchmark-1c69bf617b51d5a2dee145442deaa371.pdf)

<sup>4</sup> <http://wakandadb.org/>

<sup>5</sup> [http://en.wikipedia.org/wiki/IBM\\_Information\\_Management\\_System](http://en.wikipedia.org/wiki/IBM_Information_Management_System)

<sup>6</sup> <http://www.itqlick.com/adabas/feedback>

<sup>7</sup> <http://www.rocketsoftware.com/products/rocket-universe>

<sup>8</sup> <http://stackoverflow.com/questions/4219624/pros-and-cons-of-multi-value-databases>

<sup>9</sup> <http://www.orienttechnologies.com/orientdb/>

<sup>10</sup> <https://www.mail-archive.com/orient-database@googlegroups.com/msg03928.html>

<sup>11</sup> Firebrid.org

<sup>12</sup> [www.sap.com](http://www.sap.com)



---

13 [www.influxdb.net](http://www.influxdb.net)  
14 [www.bityota.com](http://www.bityota.com)  
15 AWS RedShift  
16 [www.spacecurve.com/](http://www.spacecurve.com/)  
17 Column Oriented Database Vs Row Oriented Databases  
18 <http://www.pivotal.io/>  
19 <http://www.sap.com/>  
20 <https://www.parstream.com/>  
21 [Kx.com](http://Kx.com)  
22 [Luciddb.com](http://Luciddb.com)  
23 [Kognitio.com](http://Kognitio.com)  
24 <http://druid.io/blog/2012/10/24/introducing-druid.html>  
25 Vertica vs Aster Data vs Greenplum vs Netezza vs Teradata from stackoverflow  
26 [site.teradata.com](http://site.teradata.com)  
27 <http://hpccsystems.com/>  
28 Teradata Aster gets graph database, HDFS-compatible file store  
<http://www.zdnet.com/article/teradata-aster-gets-graph-database-hdfs-compatible-file-store/>  
29 [Sicdb.org](http://Sicdb.org)  
30 SQL on Hadoop Landscape and Considerations  
31 [www.jethrodata.com](http://www.jethrodata.com)  
32 [www.citusdata.com/](http://www.citusdata.com/)  
33 <http://www.cloudera.com/content/cloudera/en/products-and-services/cdh/impala.html>  
34 8 SQL-on-Hadoop frameworks worth checking out  
35 <http://www-01.ibm.com/software/data/what-is/big-sql.html>  
36 <http://drill.apache.org/>  
37 <https://hive.apache.org/>  
38 Apache Hive Review <http://www.gise.cse.iitb.ac.in/wiki/images/2/26/Hive.pdf>  
39 <http://tajo.apache.org/>  
40 StackOverflow: GoogleApps Datastore Cons and Pros  
41 <https://cloud.google.com/datastore/docs>  
42 StackOverflow: GAE DataStore vs Google Cloud SQL for Enterprise Management Systems  
43 Cassandra vs MongoDB vs CouchDB vs Redis vs Riak vs HBase vs Couchbase vs OrientDB vs Aerospike vs Neo4j vs Hypertable vs Elasticsearch vs Accumulo vs VoltDB vs Scalaris comparison  
by Kristof Kovacs  
44 <http://cassandra.apache.org/>  
45 <http://apache-accumulo.1065345.n5.nabble.com/How-does-Accumulo-compare-to-HBase-td10464.html>  
46 Database Options: Beyond RDBMS <http://www.hcltech.com/blogs/engineering-rd-services/database-options-beyond-rdbms>  
47 HBase Architecture Analysis, CYANNY  
48 <http://hypertable.org/>  
49 <http://en.wikipedia.org/wiki/DataStax>  
50 <http://aws.amazon.com/dynamodb>  
51 <http://www.slideshare.net/saniyakhalsa/dynamo-db-pros-and-cons>  
52 <http://aws.amazon.com/simplifiedb/>  
53 <http://aws.amazon.com/dynamodb/faqs/>  
54 <https://wiki.openstack.org/wiki/Magnetodb>  
55 <https://redislabs.com/redis-comparison>  
56 <https://redislabs.com/redis-comparison>  
57 <http://aws.amazon.com/elasticache/>  
58 <http://www.redisgreen.net/>  
59 <http://objectrocket.com/redis>  
60 <http://hyperdex.org/>  
61 <https://github.com/google/leveldb>

---

62 <http://www.oracle.com/technetwork/database/database-technologies/berkeleydb/overview/index.html>

63 <http://www.oracle.com/us/products/database/nosql/overview/index.html>

64 <https://redislabs.com/redis-comparison>

65 [wikipedia](#)

66 <http://basho.com/riak/>

67 <https://www.arangodb.com/>

68 <http://www.aerospike.com/docs/architecture/>

69 <http://www.gridgain.com/products/in-memory-data-fabric/features/>

70 <http://www.scaleoutsoftware.com/>

71 <http://www.pivotal.io/big-data/pivotal-gemfire-xd>

72 <http://lucidworks.com/product>

73 <https://wiki.trafodion.org>

74 <http://www.splicemachine.com/>

75 <https://go.oracle.com/>

76 <https://go.oracle.com/>

77 <http://hana.sap.com/>

78 <http://vschart.com/compare/infinitegraph/vs/neo4j>

79 <http://www.hypergraphdb.org/>

80 <http://franz.com/agraph/allegrograph/>

81 [Scaling Apache Giraph to a trillion edges](#)

82 <http://sparqlcity.com/>

83 <http://research.microsoft.com/en-us/projects/trinity/>

84 [thinkaurelius.github.io/titan/](http://thinkaurelius.github.io/titan/)

85 [www.objectivity.com](http://www.objectivity.com)

86 <https://github.com/twitter/flockdb>

87 <http://sparsity-technologies.com/#sparksee>

88 <http://neo4j.com/>

89 <http://www.odbms.org/>

90 <https://www.memcachier.com/>

91 <https://redislabs.com/memcached-cloud>

92 <http://www.iron.io/cache>

93 <http://terracotta.org/>

94 <http://ehcache.org/>

95 <http://infinispan.org/about/>

96 <http://www.redhat.com/en/technologies/jboss-middleware/data-grid>

97 <http://memcached.org/>

98 <http://www-01.ibm.com/software/data/informix/>

99 <https://github.com/comsysto/jumbodb>

100 <http://rethinkdb.com/>

101 <http://docs.couchdb.org/en/latest/intro/why.html>

102 <http://ravendb.net/>

103 <http://www.tokutek.com/tokumx-for-mongodb/>