XtremeData Inc.
dbX Analytics System Introduction
Founded in 2004

Locations: Schaumburg, IL (HQ / Engineering/ Sales); Bangalore, India (Engineering); St. Louis, MO (Engineering); South Bend, IN (Sales)

Current business and products
- Patented “In-Socket Accelerator” (ISA) that enables companies to create higher performance, lower power appliances. Branded as XD2000F and XD2000I in the High-Performance Computing market.

Management:
- Ravi Chandran; CEO and Founder
- Faisal Shah; Chief Technologist, Database Appliance
  - Former Cofounder Knightsbridge Solutions LLC
- Jay Desai; Sr. VP of Strategy
  - Former Cofounder Knightsbridge Solutions LLC
- Susan Clarke; VP of Finance
- Geno Valente; VP of Sales and Marketing
Analytic queries - an unmet need

**Data Sources**
- Operational Systems
- Normative Data
- Archived Data
- Extended Enterprise

**Enterprise Information Continuum**
- Information Hub
- Operational Data Store
- Data Warehouse

<table>
<thead>
<tr>
<th>Major Trends Driving BI</th>
<th>Consequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Becoming more real time &amp; pervasive &amp; even Mission Critical</td>
<td>Managing service levels for BI applications is increasingly becoming a challenge</td>
</tr>
<tr>
<td>Enterprise wide in scope with increasing data volume and data access needs</td>
<td>Educated users want more data - want ad hoc access/ unconstrained exploration of larger data sets</td>
</tr>
<tr>
<td>Customer/partner focused - business process differentiation; allowing outside users access to previously internal data</td>
<td>Pre-Defined queries no longer work; outside users can not or will not tell you how they want to use/access the data. They just want to use it.</td>
</tr>
</tbody>
</table>
Addressing Key Market Problems

- Solutions today “restrict” access to data based on pre-defined query patterns to get performance.

- Analysts and Data researchers are demanding “unrestricted” ad hoc access to find the business value inside their data – reference “Competing on Analytics”

- Customers are looking for cost take-out and low TCO. Major vendors today are Expensive.

- Customers are looking to bring longer time-series, archived data, and low-information density data into play for analytics, at a justifiable ROI.
Legacy solutions are not designed for analytic (ad hoc) access to large scale data

Existing large-scale BI solutions are too expensive (>$100K/TB)

Cheaper solutions (<$50K/TB) cannot perform analytics with acceptable execution speed
Key Difference

Only XtremeData offers an appliance purpose built for ad hoc analysis of large data sets.
XtremeData Approach

- Start with a commodity **Linux Cluster**
  - Add direct-attached distributed storage
  - Add high-speed interconnect network (InfiniBand)
  - Add a re-engineered open-source DB engine (postgresql) that:
    - Supports a shared-nothing, parallel query execution model
    - Supports hardware acceleration via “SQL on a Chip”

![Diagram showing Price and Performance relationship]
...and it morphs into an **SQL Super-Computer** delivering:

- Scalable, shared-nothing MPP architecture
- Hardware accelerated parallel SQL processing
- Efficient, zero-copy, data exchange on high bandwidth network
- Dynamic load balancing at run-time
Performance Results

Sampling of performance results from recent Proof-of-Concept engagements:

- Financial institution, household credit database: large tables, N-way Joins, Aggregations.
  - **Result:** On 8xNode system (MSRP $600,000): 33 min 42 sec. \(~16\times\) faster than existing solution, with estimated MSRP of $1,500,000.

- National Laboratory, Graph traversal to discover linkages: single narrow long table, N-way self-Joins.
  - **Result:** On 8xNode system (MSRP $600,000): 1 hour 25 mins. \(~5\times\) faster than closest competitor, system configuration and MSRP unknown.

- National Laboratory, Web page ranking (PageRank): rank computation, iterative traversal with updates.
  - **Result:** On 8xNode system (MSRP $600,000); \(2\times-8\times\) faster than closest competitor, with estimated MSRP of $1,200,000.
Differentiation

Utility
➢ Only XtremeData's dbX cost-effectively enables wide, deep, and unrestricted querying of structured data.

Price
➢ Only XtremeData's dbX enables Petabyte-scale analytic solutions that were previously cost-prohibitive - An appliance priced at $20K/Usable TB

Performance
➢ Only XtremeData's dbX delivers 4-30x faster results than competing products at 1/3 the investment

Green
➢ Only XtremeData's dbX provides power and cooling advantages with an unmatched carbon footprint.
The four components of a data analysis system are:

**Hardware (3x):** Storage, Computing (Servers), Interconnect Network

**Software (1x):** Relational DataBase Management System (RDBMS)

Unique among all vendors in the market today: the dbX architecture leverages the best of all industry trends...We can sustain our price/performance advantage into the future.
Engineered from first principles for large database analytics

Breakthrough price/performance (16xNodes): $20K/TB at 0.4TB/min query processing

Provides clear value vs. competitive solutions (Teradata, Netezza, Oracle, GreenPlum)

- Query performance (SQL processing)
- Load performance (1-4 TB/hour)
- Processing density (storage & performance/sq. foot)
- Carbon footprint (performance/watt)
### Under the hood

<table>
<thead>
<tr>
<th>dbX System</th>
<th>Data Node Features</th>
<th>SQL on a Chip</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head Node</td>
<td>HP DL385 with SAS Storage</td>
<td>Approved HP Accelerator</td>
</tr>
<tr>
<td>IB Switch</td>
<td>Quad-Core CPU; IB Connection</td>
<td>Contains SQL in Hardware</td>
</tr>
<tr>
<td>Data Nodes</td>
<td>Patented In-Socket Accelerator</td>
<td>Plugs into CPU Socket</td>
</tr>
</tbody>
</table>

**dbX: DB Appliance**

**Data x4**

**Head**

**Data x4**

**Node from Appliance**

**FPGA In-Socket Accelerator (ISA) from Node**
Patented “SQL on a Chip”

- Replaces CPU in COTS Tier 1 Servers... Qualified by HP (Accelerator Program)
- SQL operations accelerated “under the hood” in silicon
  - Big table data movement: Loads/Unloads/Scans
  - Big operators: Joins, Sorts, GroupBy, OrderBy,...
- 10x the Performance at 1/3\textsuperscript{rd} the power of CPU
- Leverages rapidly advancing cutting edge FPGA technology in patented way
### Configurations

<table>
<thead>
<tr>
<th>Model</th>
<th>dbX 1008</th>
<th>dbX 1016</th>
<th>dbX 1028</th>
<th>dbX 1060</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rack (Standard 24x40)</td>
<td>1 (50%)</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Data Nodes (Opteron + SQL on a Chip)</td>
<td>8</td>
<td>16</td>
<td>28</td>
<td>60</td>
</tr>
<tr>
<td>User Data (1TB drives)</td>
<td>30TB</td>
<td>60TB</td>
<td>105TB</td>
<td>225TB</td>
</tr>
</tbody>
</table>

Scalable to 1024 Nodes Over 5PB of User Data (15PB total disk space)
Applications for dbX

- Adjunct to existing environments
- Complement legacy investments in DW infrastructure
- Shift analytic query workload away from the current environment - better leverage legacy environment for operational BI applications
- “Sand box” for research, analysis, ad hoc.....
- SQL based analytics
Filling clear market needs for

Usability
XtremeData's dbX cost-effectively enables wide, deep, and unrestricted querying of structured data.

Price
An Appliance priced at $20TB/Usable TB, only XtremeData's dbX enables Petabyte-scale analytic solutions that were previously cost-prohibitive.

Performance
XtremeData's dbX delivers 10x faster results than competing products at 1/3 the investment.

Environment (energy-efficiency)
XtremeData's dbX provides power and cooling advantages with an unmatched competitive carbon footprint.