DataDirect TECHNOLOGIES

# Best Practices for Data Connectivity

**Rob Steward**
**Director, Research & Development**

**Enterprise Architect Summit**
**Key Biscayne**
**May, 2006**

## Agenda

- Data Connectivity Landscape
- Business Implications for Database Connectivity
- Technical Considerations for Database Connectivity
- Best Practice Approach for Data Connectivity

## Database Connectivity History

- Progression of data storage technologies
  - Indexed -> Relational ->Object-> XML
- Architecture advancements
  - Mainframe -> Client/Server -> Web/App server -> SOA
- Data connectivity progression
  - Static / code embedded -> proprietary API -> standards-based APIs

## Database Connectivity History

- Standards – historical and current view
  - Microsoft introduces ODBC
  - OLE DB
  - JDBC follows ODBC path
  - ADO.NET for .NET development
  - XML-based standards (e.g., XQuery)
  - SOA-based standards (e.g., WSDL)

## Mainframe Data Access History

- Mainframe Data
  - Initially dominated by ISAM, VSAM, etc.
  - IBM DB2 introduced relational database technology
  - Mainframe remains primary data repository for large organizations
  - Applications that reside on other Win/UNIX/Linux require access to mainframe data & business processing
- Connectivity Approach
  - Direct access to relational data
    - Native client access
    - Standard APIs (ODBC, JDBC, ADO.NET)
  - Message Oriented Middleware options (e.g., MQ Series)
  - SOA / Web Services

## Open Source Impact

- Open source alternatives have increased the number of technology options and increased the level of heterogeneity in IT shops
- Maturity of open source alternatives vary by architecture component (Operating system, RDBMS, web/app server, connectivity options)
- Open Source really means Mixed Source
  - BEA WebSphere accessing DB2 / Linux
  - JBoss accessing Oracle / UNIX
  - C++ application accessing MySQL / Linux
- Data connectivity strategy should encompass support for mixed source environments

JBoss redhat MySQL

## Impact of latest technologies

- Object-Relational Mapping (ORM)
  - Introduced to address impedance mismatch between relational database and object-oriented models
- Service Oriented Architecture (SOA)
  - Web Services-based programming paradigm introduces another architecture that requires access to existing data sources
- XQuery / XQJ
  - Nascent standard that is focused on providing XML-based approach for accessing XML, relational and non-relational data in the Java environment

DataDirect TECHNOLOGIES

## Impact of latest technologies

- While data is still stored in a database, you need middleware to connect you to it
- All of the new architectures/technologies still rely on a data connectivity layer to connect them to the data

DataDirect TECHNOLOGIES

## Agenda

- Data Connectivity Landscape
- Business Implications for Database Connectivity
- Technical Considerations for Database Connectivity
- Best Practice Approach for Data Connectivity

DataDirect TECHNOLOGIES

## Database Connectivity – Why is it so important?

- In a well-tuned environment, up to 75% of the data access time is spent in the connectivity layer
- There is a lot of literature on tuning your database server, but little on tuning your middleware
- A lot of time is spent deciding "major" architecture choices (DBMS, Environment, Language, Tools, etc.)
- In today's environments, the database connectivity layer is even more important than it used to be
- Success of the entire project can hinge on the database connectivity layer (more later…)

DataDirect TECHNOLOGIES

## Database Connectivity – Further Cost Implications

Goal – Manage the cost / complexity of development, test and deployment

- Eliminate development delays
  - Unplanned costs that crop up late in the development cycle (or early in production rollout phase) can be debilitating
- Eliminate QA delays
  - Standards-based approach can limit the cost associated with testing multiple proprietary interfaces
- Excessive deployment cost
  - Eliminating need for native database components can significantly reduce deployment cost
- Contain Network / System resource cost
  - Connectivity design significantly impacts the network and server resources required to support an application

DataDirect TECHNOLOGIES

## Database Connectivity – Risk Mitigation

Goal – Mitigate operational and data security risk while minimizing risk reduction investment

- Ensure data integrity and regulatory compliance
  - Sarbanes-Oxley, Basel II, etc.
  - Secure data / credential transmission
  - Single Sign-On
- Reduce IT Operational risk
  - Project delivery risk
  - Operational and system downtime risk
  - Ensure interoperability / flexibility (future proof)

DataDirect TECHNOLOGIES

## Database Connectivity – Revenue/ Profit / Shareholder Value

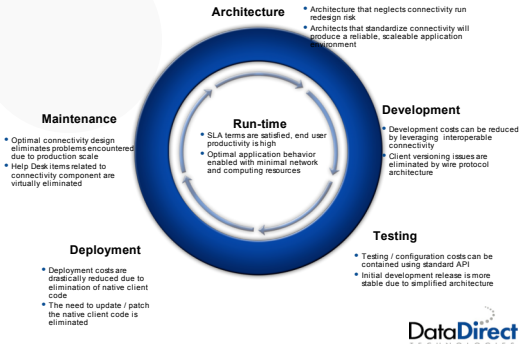
**Goal – Ensure success of critical applications that directly impact revenue, profit and shareholder value**

- Application performance, scalability, usability for critical applications directly impacts the bottom-line
- Superior Customer Interactions are necessary for corporate / IT success
- Service Level Agreements can be positively affected by superior database connectivity

Database Connectivity is often overlooked, but plays a key role in the success / failure of the application

DataDirect TECHNOLOGIES

## Database Connectivity – Impact is not limited to developers




DataDirect TECHNOLOGIES

## Agenda

- Data Connectivity Landscape
- Business Implications for Database Connectivity
- Technical Considerations for Database Connectivity
- Best Practice Approach for Data Connectivity

DataDirect TECHNOLOGIES

## Application Architecture strongly influences Data Connectivity requirements

- Data source types are closely associated with the architecture
- Method used to access data is strongly influenced by architecture type
- Most enterprises develop their new applications on the latest architectures, but "legacy" architectures and data co-exist
- Programmer experience is limited by application architecture

Mainframe Architecture (e.g., zOS)
Client / Server Architecture
Multi-Tier Architecture (e.g., Application Server)
Web-based applications (e.g., HTTP)
Web Service based applications (e.g., SOA)

DataDirect TECHNOLOGIES

## Development Environment largely dictates Database Connectivity approach

- Each development environment has a preferred methodology for data connectivity
- Standards adoption / level of database connectivity maturity varies by development environment
- Abstraction between data and programming style varies between development environment
- Programmer experience is directly related to development environment

COBOL
C / C++
Perl / PHP
Visual Basic
Java
.NET
HTML based
XML based

DataDirect TECHNOLOGIES

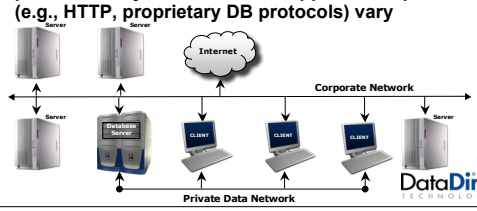
## Development Environment largely dictates Database Connectivity approach

- C, C++, Perl, COBOL, etc. = ODBC
- Java = JDBC
- .NET = ADO.NET

DataDirect TECHNOLOGIES

## Network Topology influences Database Connectivity requirements

- LAN / WAN topology – where the end user lives relative to the application / database influences data access approach
- Security considerations based on whether the network is private / public are critical
- Although low-level network protocol are predominantly TCP/IP-based, "application" protocols (e.g., HTTP, proprietary DB protocols) vary




DataDirect TECHNOLOGIES

## Data Storage Type dictates Database Connectivity approach

- Tools used to manage, develop, test, deploy are specific to the data storage type
- Format of data storage influences data access approach
- Proprietary API used to create / access data is provided by each storage type
- Standards-based access is available for most data
- Most organizations have a heterogeneous combination of data storage types and some applications require "distributed" access to multiple formats
  - Flat file format
  - Index-based (e.g., ISAM, VSAM)
  - Text-based documents
  - Relational (e.g., Oracle, SQL Server)
  - Object-Oriented
  - Tag-based (XML, HTML)

DataDirect TECHNOLOGIES

## Application Requirements influence Database Connectivity requirements

- Applications that directly influence the business success (revenue, cost, risk) drive the performance, scalability, reliability and availability requirements
- End user type (e.g. internal, external) dictates network topology
- Client form factor (desktop, laptop, PDA) and connection style (connected / wireless) influences query and security requirements
- Ability to meet Service Level Agreements (SLAs) or internal user expectations is paramount to IT success

DataDirect TECHNOLOGIES

## Database Connectivity – What you don't know can hurt you

- Myth – If I tune my database, network and application code, I have performance & scalability covered
  - Developer's don't realize that up to 75% of the data access time is spent in the middleware (vs. 25% in database engine)
- Myth – The database vendors are positioned to provide the best connectivity components
  - The database vendors treat connectivity as a check-list item
  - "Wire protocol" design passed them by
- Myth – Standards-based data access is slower than native access
  - This is a vestige of ODBC-based implementations that sit on top of the native database driver
  - "Wire protocol"-based drivers consistently outperform native database components

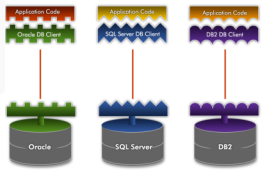
DataDirect TECHNOLOGIES

## Database Connectivity Options

- Native access
- Standards-based
  - Shim approach (native client wrapped with standards-based interface)
  - Direct approach (direct access to the database via the wire protocol)
- Web Services / SOA

DataDirect TECHNOLOGIES

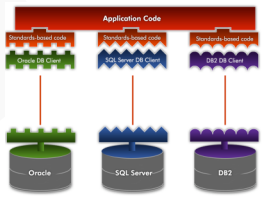
## Native Database Access




- Proprietary connectivity eliminates interoperability - Each database requires different application code, which adds to development costs
- Excessive deployment cost - Database clients must be deployed, which results in administration overhead.
- Performance & scalability limitations - Extra layer in communication stack degrades performance/scalability.

DataDirect TECHNOLOGIES

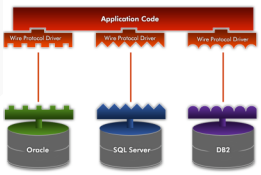
## Shim-based Standards API Approach




- Interoperability provided by standard API - Single version of application code for all databases cuts development costs.
- Excessive deployment cost - Database clients must be deployed, which results in administration overhead.
- Performance & scalability limitations - Extra layer in communication stack degrades performance/scalability.

DataDirect TECHNOLOGIES

## Wire Protocol Standards API Approach




- Interoperability provided by standard API - Single version of application code for all databases cuts development costs.
- Streamlined deployment - no database client eliminates administration overhead.
- Optimal performance & scalability - Direct communication with database results in better performance/scalability.

DataDirect TECHNOLOGIES

## Agenda

- Data Connectivity Landscape
- Business Implications for Database Connectivity
- Technical Considerations for Database Connectivity
- **Best Practice Approach for Data Connectivity**

DataDirect TECHNOLOGIES

## Pitfalls of ad-hoc data connectivity

NUCLEUS RESEARCH

**Planning and development**
-Learn new database versions
-Learn and test new database drivers

**Deployment**
-Plan and execute
-Check client libraries
-Debug

**Testing**
-Test multiple drivers for each application
-Risk performance problems that slow development cycles
-Increase troubleshooting time
-Increase tactical development costs

**Production**
-Support SLAs
-Manage performance issues
-Address user challenges

**Upgrade**
-Increase upgrade complexity
-Slow upgrade efficiency
-Increase upgrade costs

**Retire and rebuild**
-Identify new technologies
-Learn new database versions
-Learn and test new database drivers

DataDirect TECHNOLOGIES

## Benefits associated with a well planned connectivity strategy

NUCLEUS RESEARCH

**Planning and development**
-Improved design
-Improved development
-Reduced ramp-up time

**Deployment**
-Reduced time to execute
-Less risk of conflicts

**Testing**
-Reduced testing time and cost
-Increased testing predictability

**Production**
-Less support
-SLA compliance
-Better user performance

**Upgrade**
-Reduce cost and complexity

**Retire and rebuild**
-Leverage existing data
-Less "putting out fires"
-More focus on operations

Ongoing application performance and management

DataDirect TECHNOLOGIES

## Data Connectivity Best Practices – Strategic Considerations

- **Invest in your core competency, don't allocate your development resources to middleware components**
- **Think strategically, not tactically**
  - **Design in flexibility – technology will change, make sure that your approach is extensible**
  - **Design in interoperability – don't limit yourself to a single data source option**
  - **Don't sacrifice application performance, scalability, reliability**
- **Consider the "hidden" business implications related to database connectivity**
- **Your application architecture is as strong as the weakest component in the stack, don't neglect the data connectivity layer**

DataDirect TECHNOLOGIES

**Rob Steward**

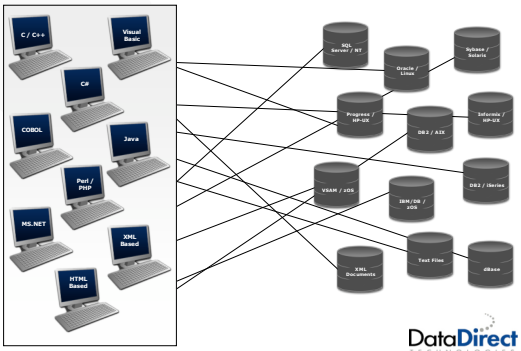
Enterprise Architect Summit 2006

## Data Connectivity Best Practices – Technical Considerations

- Satisfy application requirements
  - Performance, Scalability, Availability, Reliability
- Meet the needs of your heterogeneous environment
  - Server platforms, development strategies, databases
- Design for interoperability
  - Connectivity architecture should enable interoperability
- Don't limit application functionality
  - Connectivity solution should provide robust functionality
- Comply with security regulations
  - Satisfy your security / regulatory requirements
- Effectively leverage network & server resources
  - Design should ensure optimal network / server utilization

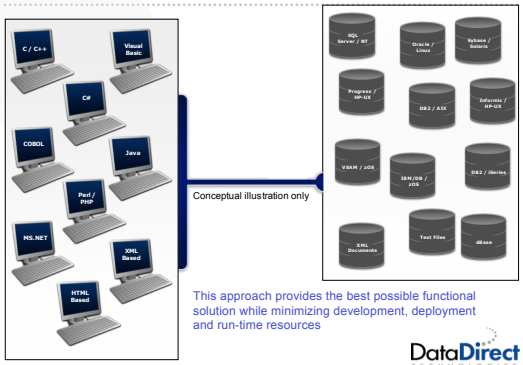
DataDirect TECHNOLOGIES

## Data Connectivity is complicated by heterogeneous environments




DataDirect TECHNOLOGIES

## Standards-based architecture simplifies database connectivity




This approach provides the best possible functional solution while minimizing development, deployment and run-time resources

DataDirect TECHNOLOGIES

DataDirect TECHNOLOGIES

## Questions / Discussion

**Rob Steward**
**Director, Research & Development**
**rob.steward@datadirect.com**