

October 27, 2003

> **Avaki Data Grid:**
Enterprise Information Integration, Powered by Grid
pidra61m



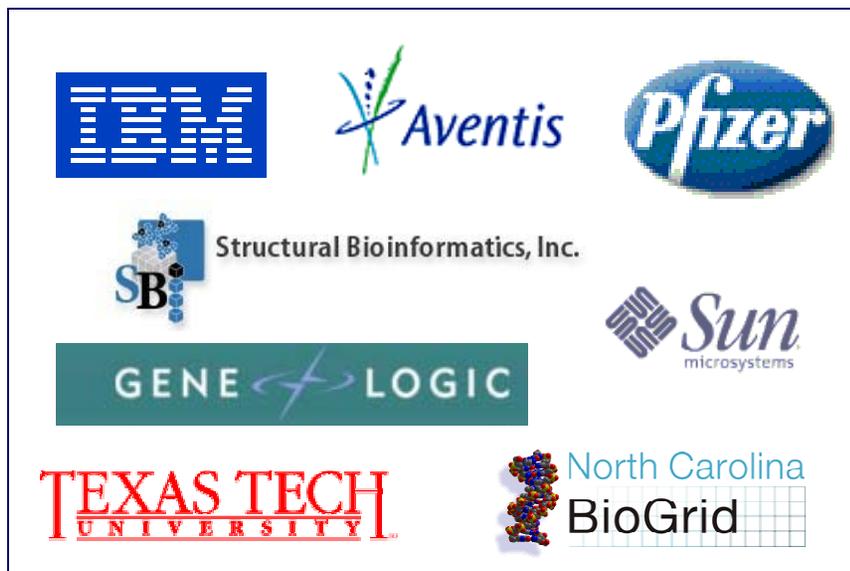
Avaki Background
Background on Grids
Avaki Grid Software
Standards
Wrap Up

Avaki Company Background

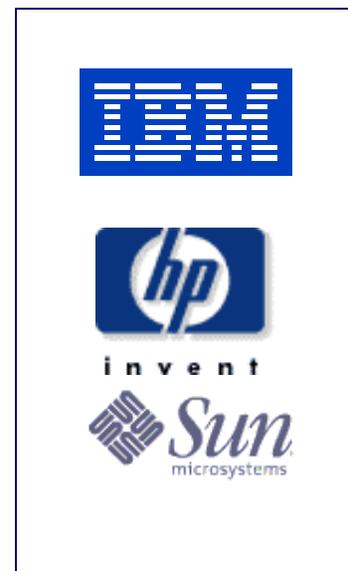


- > Award winning software for data provisioning, access and integration, powered by grid technology
- > Focus on distributed enterprises and their partners
- > Founded in 2001, top-tier investors

Customers



Partners



Awards



Avaki Background

Background on Grids

Avaki Grid Software

Standards

Wrap Up

Data Chaos

- > Huge amounts of dispersed data and data silos
- > Users in multiple locations
- > Constant data updates causing version control nightmares

Increased Computational Needs

- > Compute intensive and high-throughput applications
- > Spikes in demand for computing power
- > Complex staging requirements

Distributed and Virtual Organizations

- > Users forced to be IT experts
- > Complex administration of heterogeneous and dynamic IT resources across domains
- > Must protect intellectual property

A grid is all about gathering together resources and making them accessible to users and applications.

A grid enables users to collaborate securely by sharing *processing, applications, and data* across heterogeneous systems and administrative domains for collaboration, faster application execution and easier access to data.

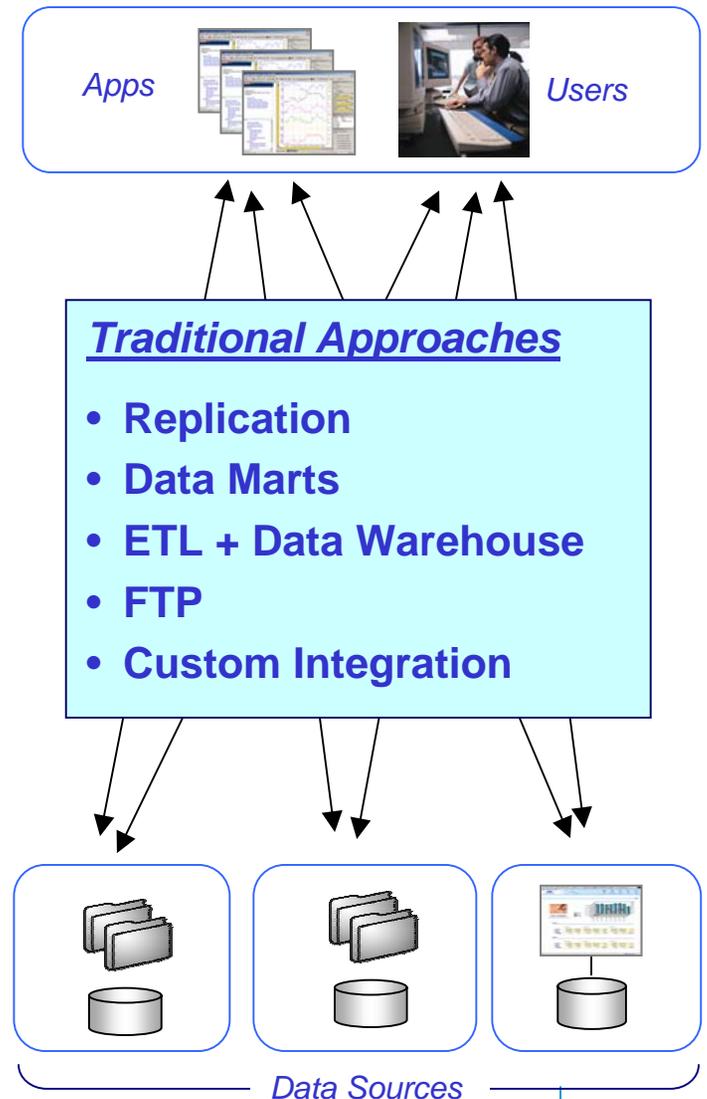
- > Compute Grids
- > Data Grids

- **Provide access to processing, data, and applications across domains**
- **Incorporate heterogeneous hardware, operating systems, and system configurations**
- **Support secure grid-wide sharing of resources, while maintaining local, diverse security policies**
- **Are resilient in the face of change and system failure**
- **Account for resources consumed**
- **Provide for arbitrarily scalability**

Data Challenges in Enterprise Integration

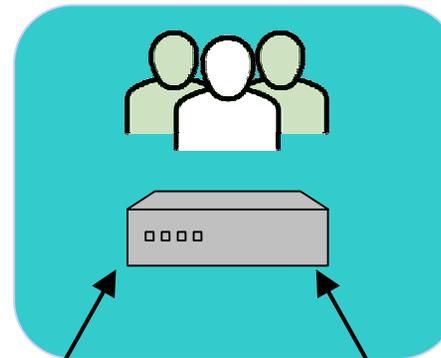


- **Data is not current**
- **Does not address all data types**
 - > Relational, files, XML, application data
- **Complex, costly to deploy & maintain**
 - > Infrastructure, programming, maintenance
- **No control over data provisioning**
 - > No enterprise layer – Point solutions
- **Little or no wide area support**



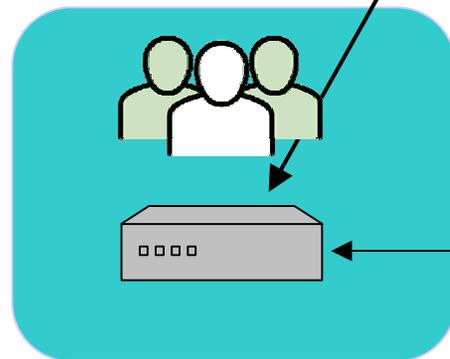
Replication Drawbacks:
NAS filers expensive
Brittle across multiple networks,
administrative domains
Data replicated that's not needed
Requires multiple backups, significant
maintenance

Center of Excellence

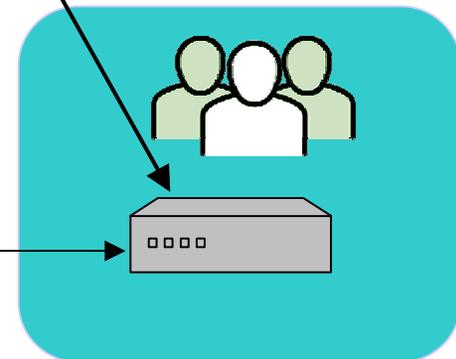


Other Potential Approaches:

- > FTP
- > DMZ
- > VPN



Company Headquarters



Center of Excellence

Alternative Data Management Approach: FTP

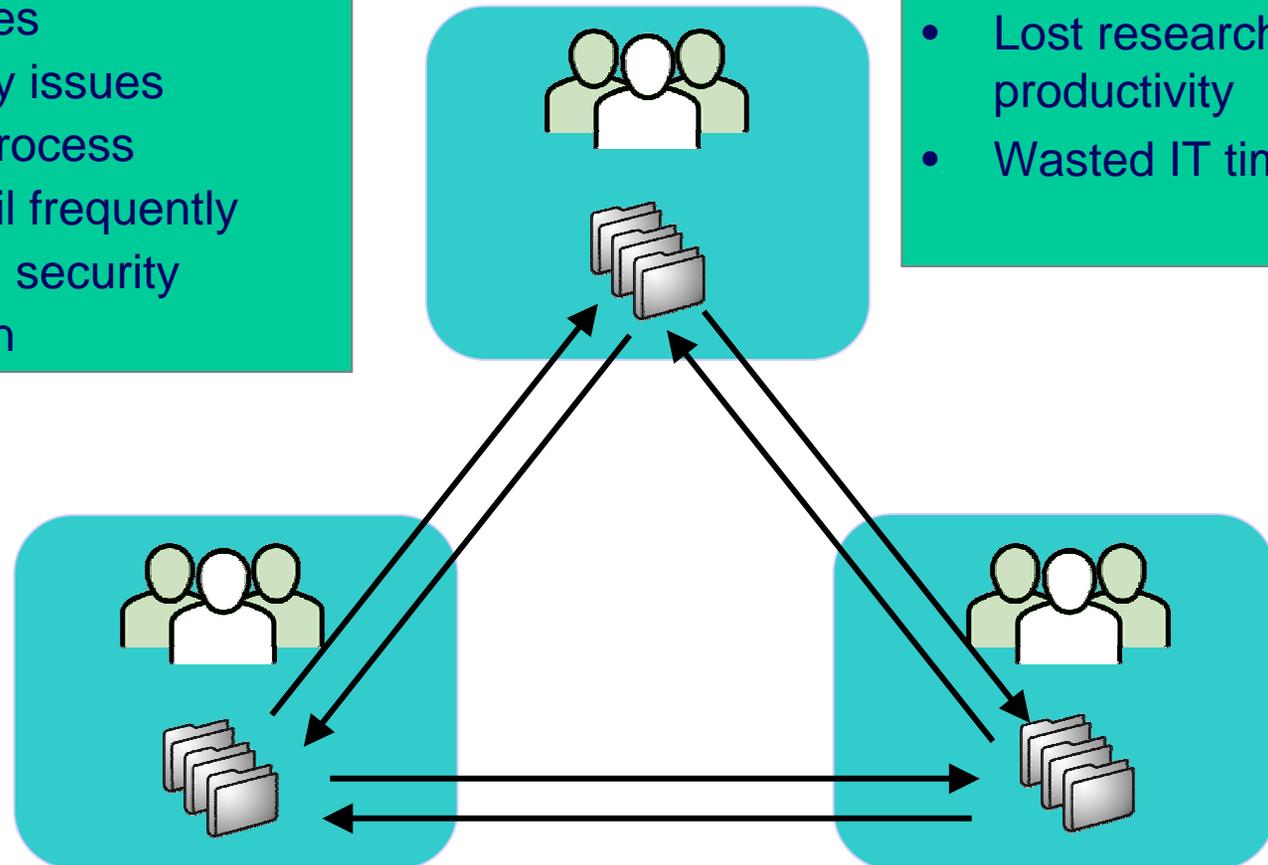


Limitations/drawbacks:

- Out of date copies
- Data consistency issues
- Highly manual process
- FTP transfers fail frequently
- Lacks integrated security
- Costly replication

Result:

- Lost researcher productivity
- Wasted IT time



“15-20% of researchers time is spent looking for the right data or tools”
“40% of the time, research has to be redone due to data or tools inconsistencies”

Avaki Background

Background on Grids

Avaki Grid Software

Standards

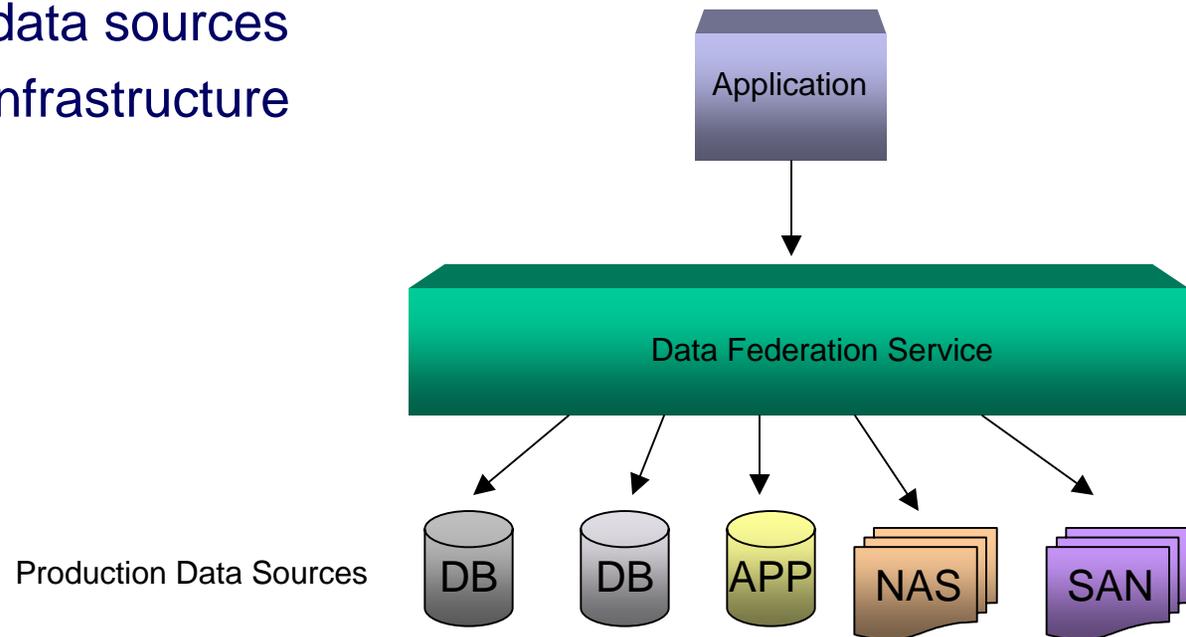
Wrap Up

Enterprise Information Integration (EII)

Data Federation for Timely, Accurate Data

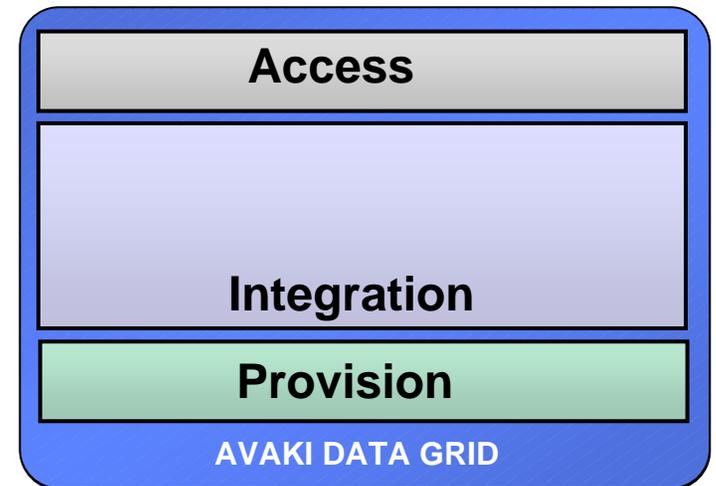


- Access distributed data from multiple sources in real time (federated data)
- Serve data from production sources – *no replicas*
- Integrate disparate data sources
- Reduce expensive infrastructure

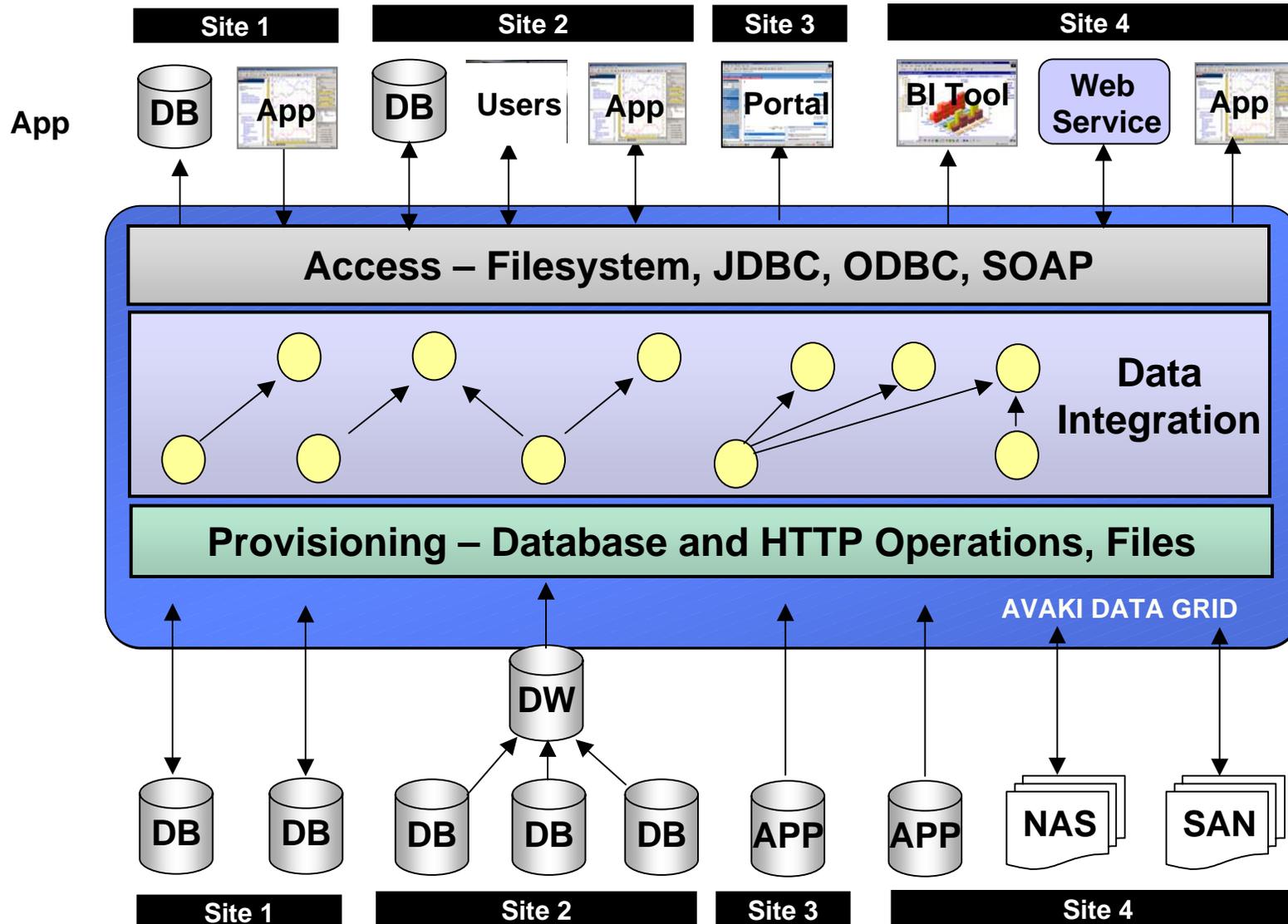


A single, packaged system to *provision, access and integrate* distributed data

- > Real-time access to federated data
 - No copies
 - Relational data & flat files
- > Protects production data sources
- > Data integration using XML
- > Comprehensive access control & security
- > Deliver data via efficient caching
- > Grid architecture – scalability, flexibility, resilience



Avaki Data Grid 4.0 - Unified Data Access & Provisioning Platform



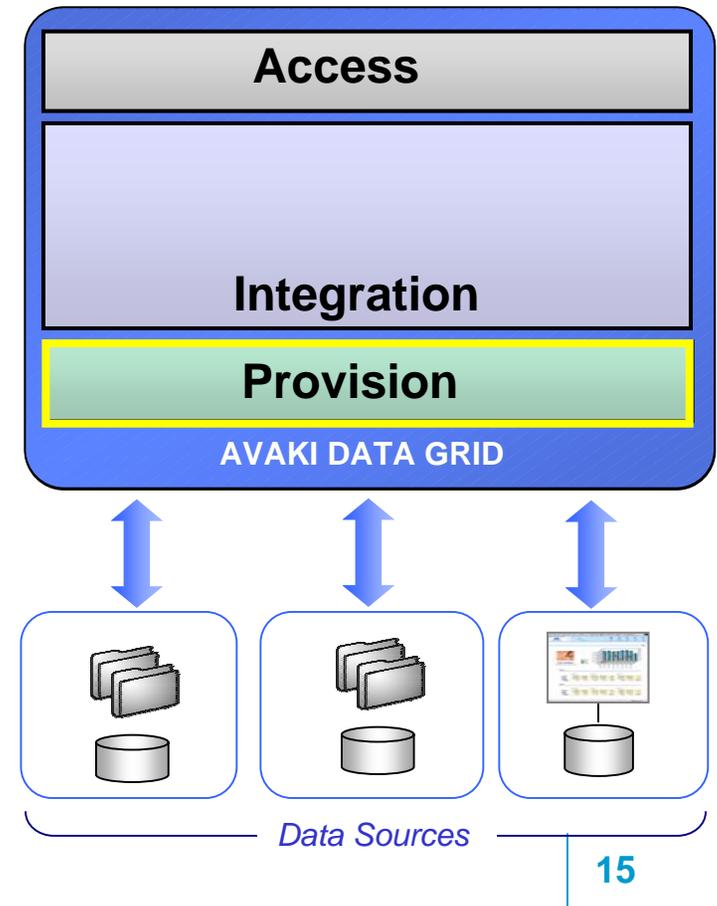
Data Provisioning with Avaki Data Grid AVAKI

> “Publish” any RDBMS data, files, XML, and application data sources (structured and unstructured data)

- Unified data catalog
- Data stays where it is, managed by local owners
- SQL statements, stored procedures
- No custom connectors

> Control access with one unified system

- Data owners control access rights at data object level
- Integration with multiple, heterogeneous directory services (LDAP, Microsoft Active Directory, NIS)



Data Access with Avaki Data Grid



> Simplified, real-time access to federated data

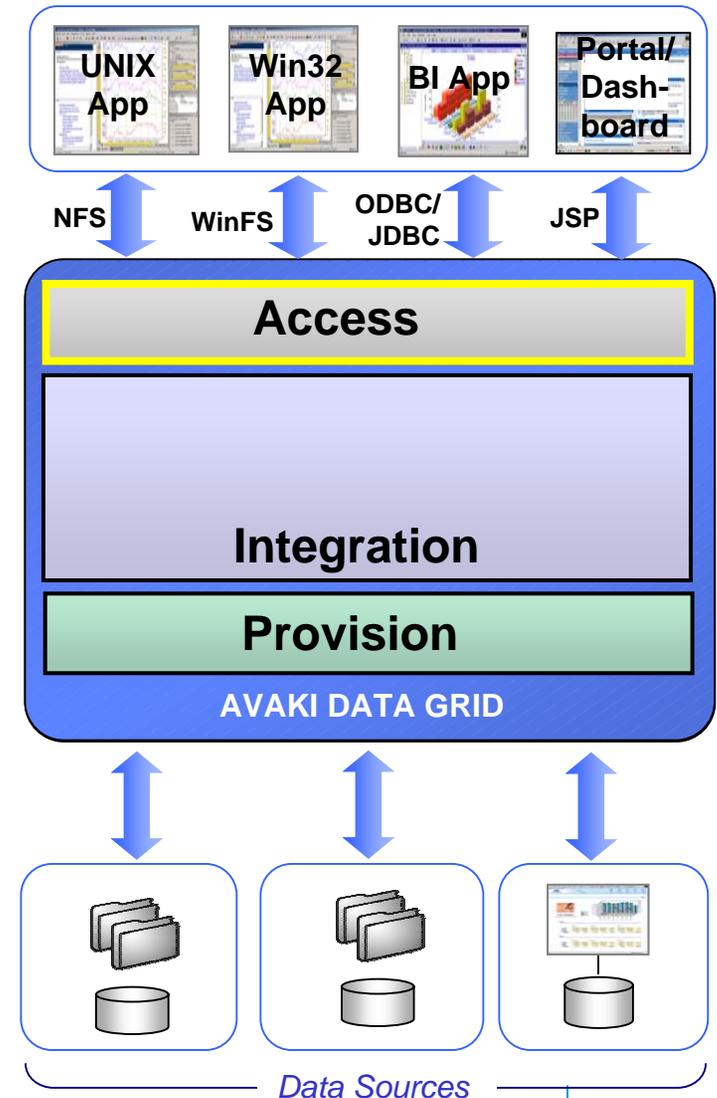
- SQL result sets, XML views, files, application data sources
- Access data anywhere as if it were local
- Read & write
- Caching options speed performance

> Unified view of all available data (data catalog)

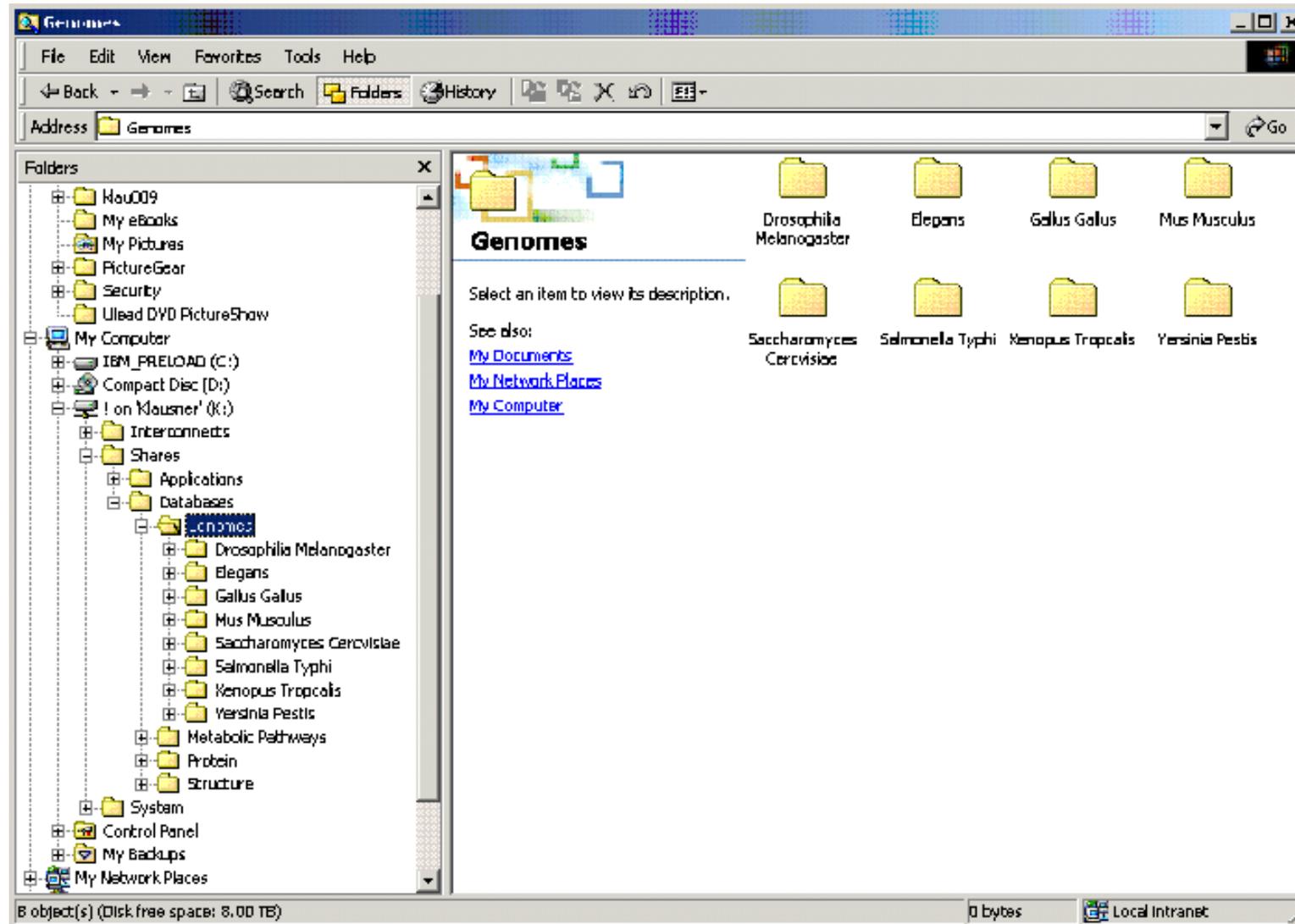
- Single service for retrieving all data objects
- Secure

> Standard Access

- File operations – NFS, WinFS
- ODBC
- JDBC
- Web Services/SOAP
- JSP/Tag library



WinFS Data Access



Data Integration with Avaki Data Grid



> Built-in transformations for relational and file data

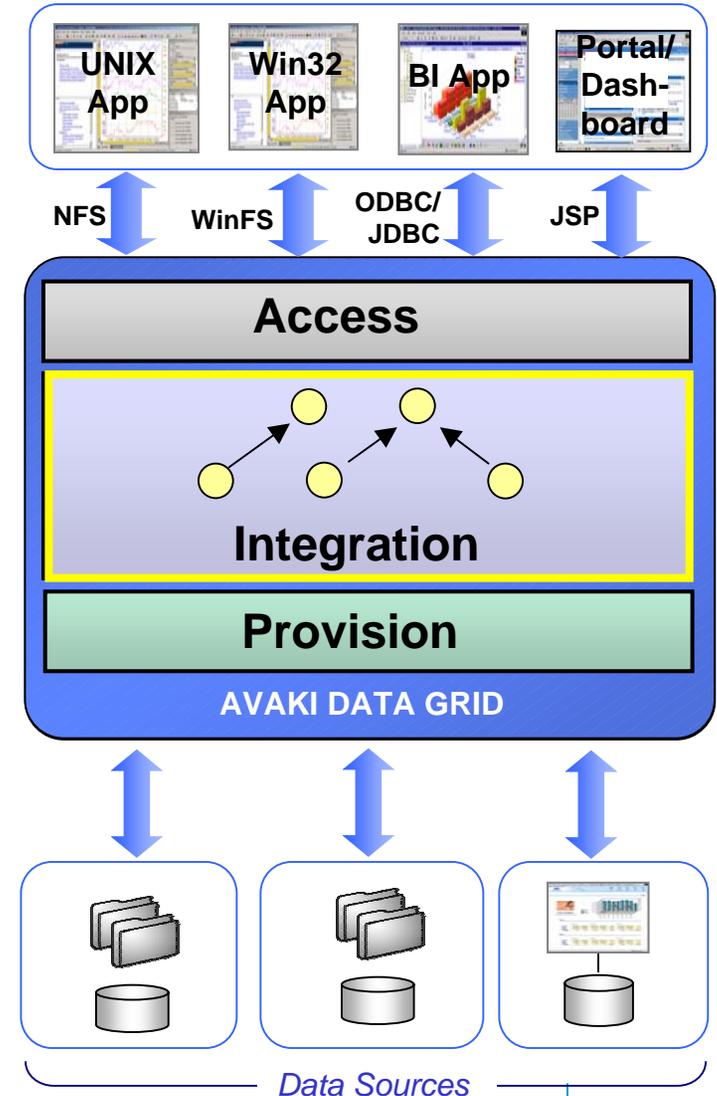
- Convert to XML, CSV, structured files
- Reduced programming

> Data transformation capabilities

- Transform data format and structure
- XSLT capabilities out of the box, with support for 3rd party tools
- Many unique “views” from same result set
- Aggregated views from multiple data sources
- XSLT style sheets reusable across the enterprise
- Integrate custom transformations if needed

> Data flows with event logic

- Triggers and schedules



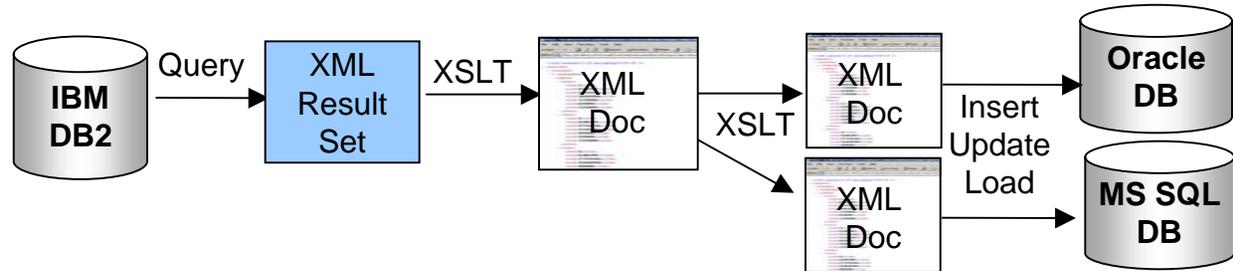
Avaki Data Flows – Typical Patterns



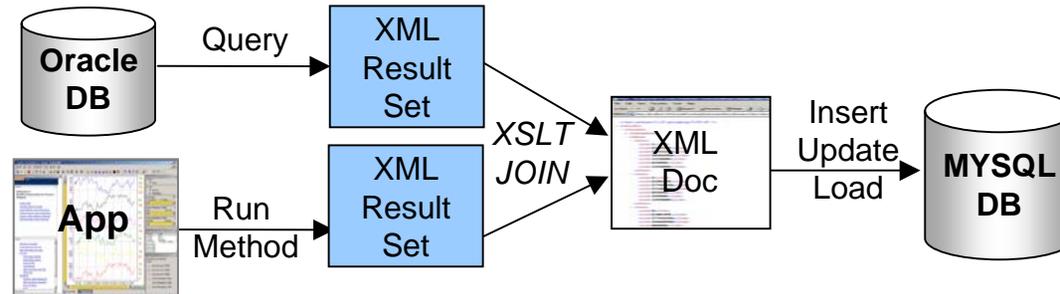
1. Format DB Data for an Application



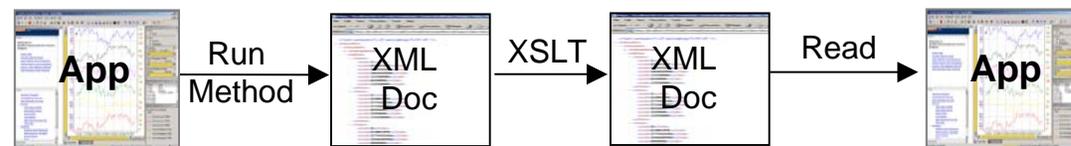
2. Copy data from a system of record and import in to other remote, customer facing databases and applications



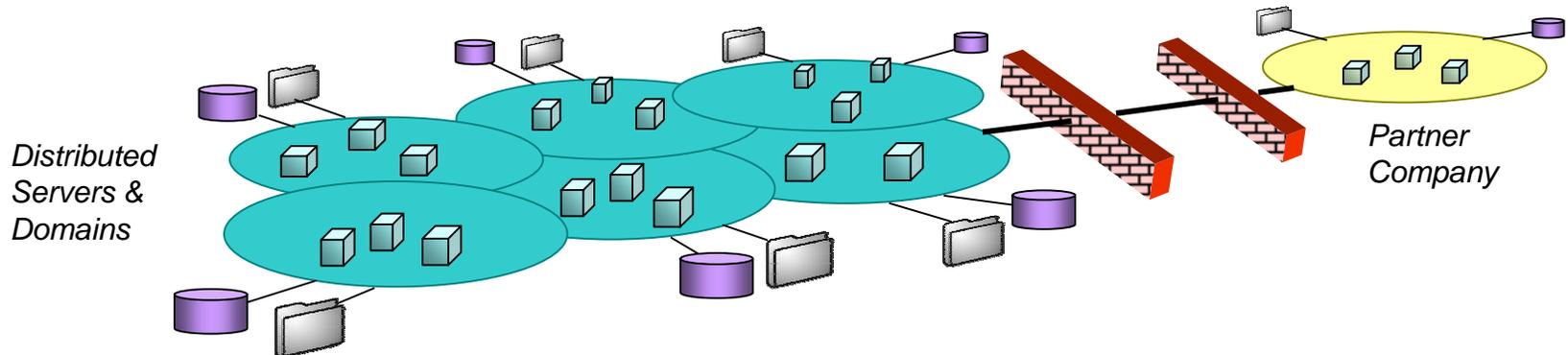
3. Combine DB and application results, then load into a remote database



4. Run a application, transform its output to the format required by another application

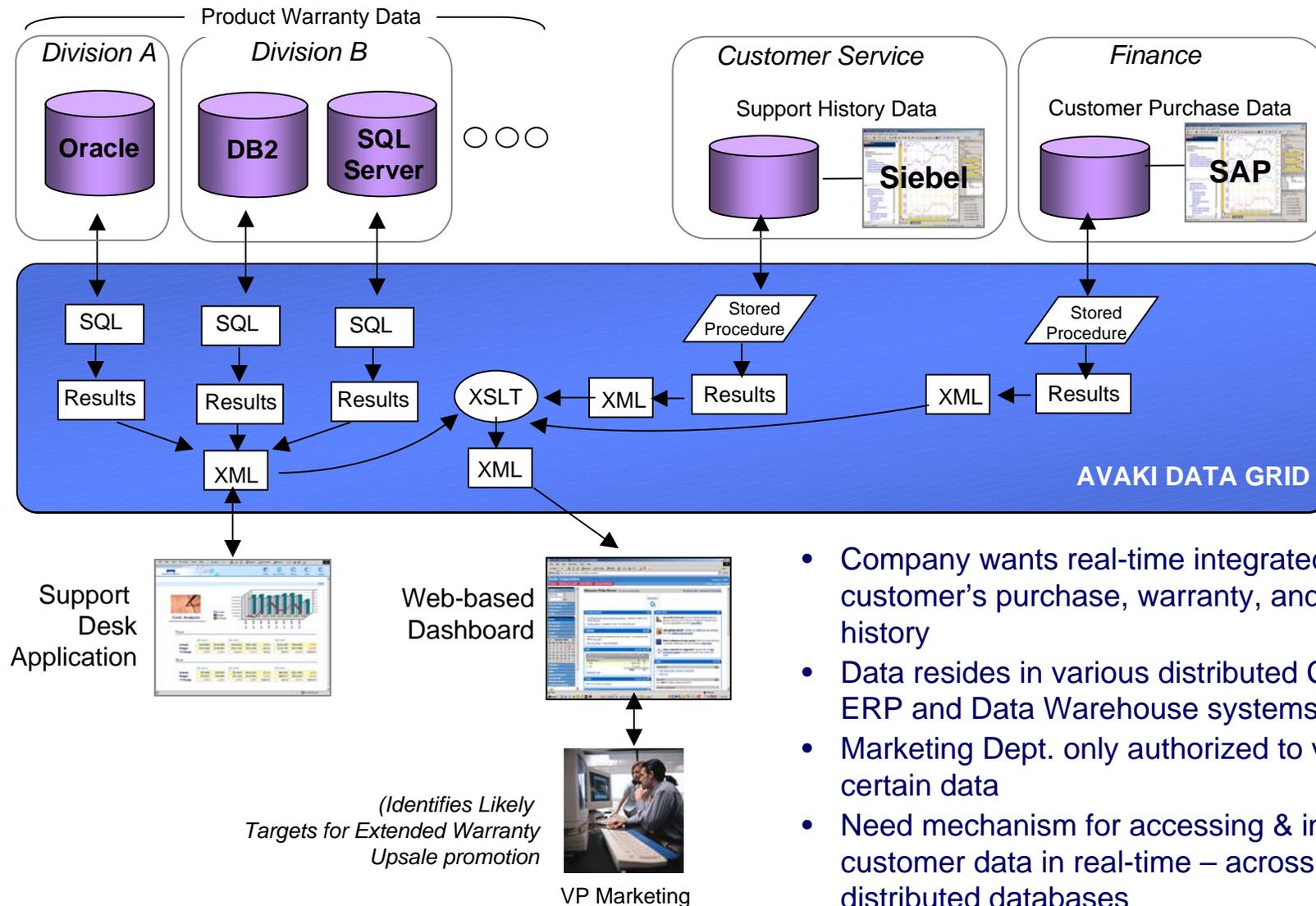


- > **Distributed, extensible architecture enables secure data sharing cross-enterprise**
 - Flexible deployment; Easy organic growth
 - Resilient, no single point of failure
- > **Provides for global sharing with local autonomy & control**
 - Local security, data management, infrastructure administration
 - Share across partner companies or other peer organizations



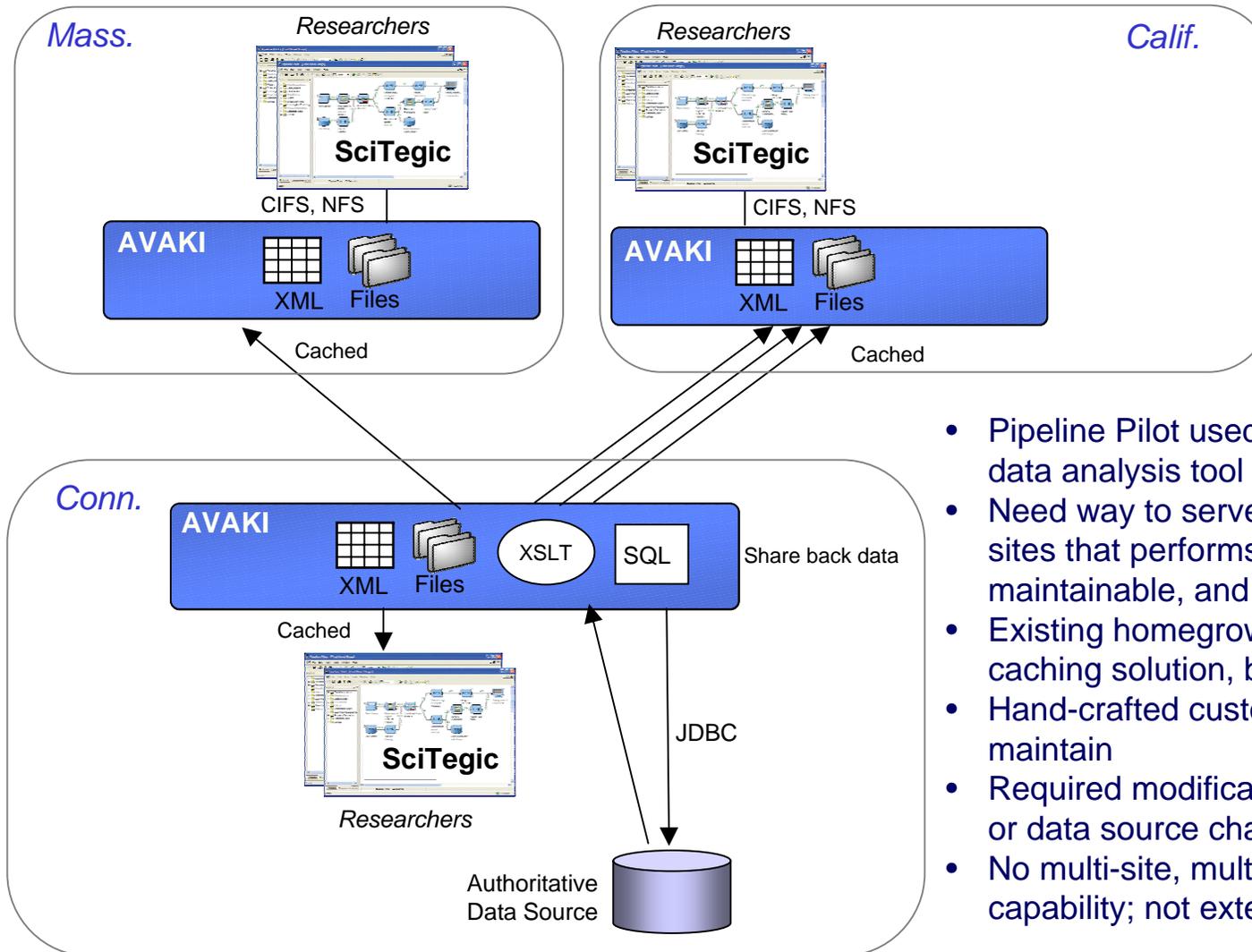
Avaki Data Grid: Use Cases

Unified View of the Customer, in Real Time



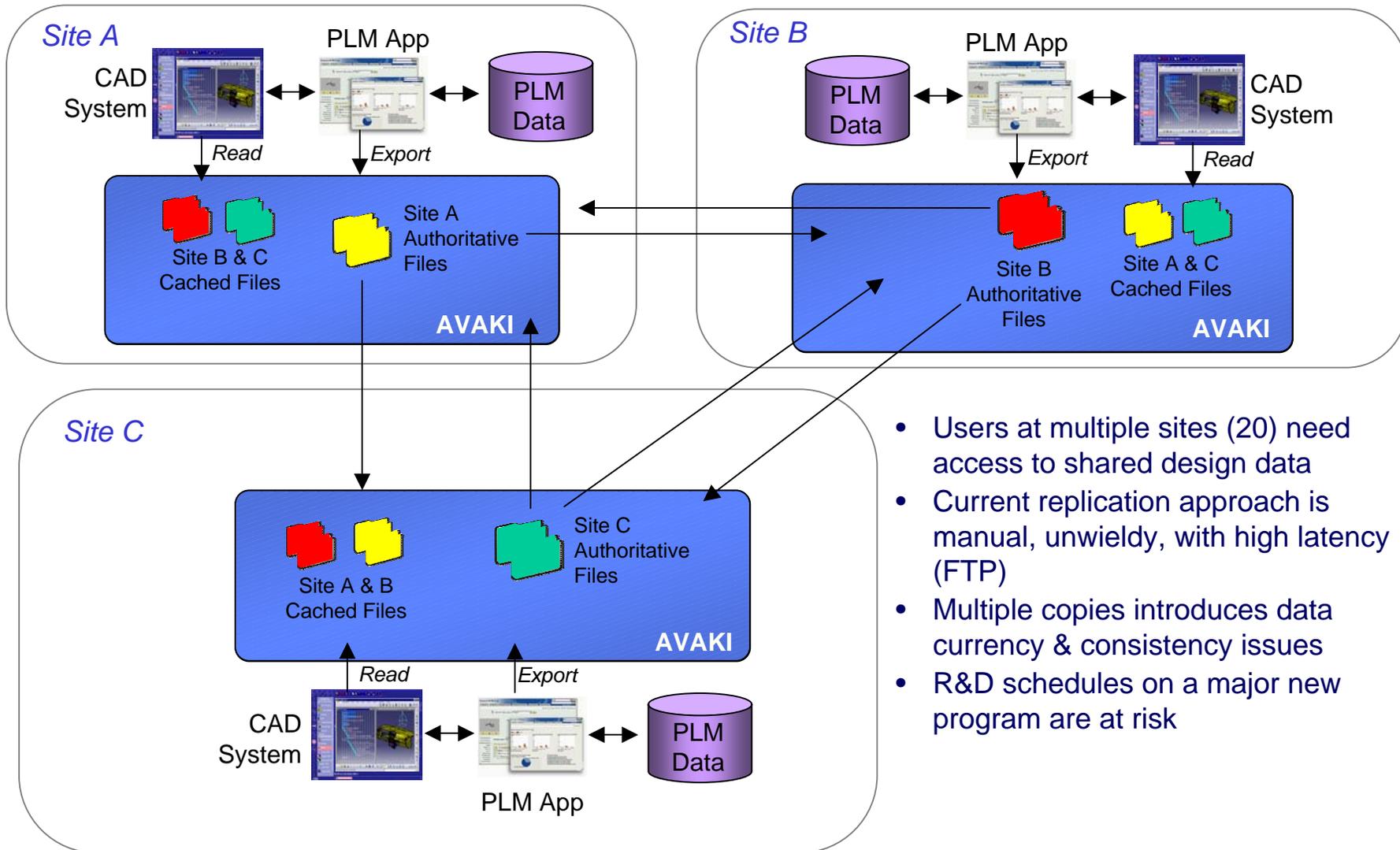
- Company wants real-time integrated view of customer's purchase, warranty, and support history
- Data resides in various distributed CRM, ERP and Data Warehouse systems
- Marketing Dept. only authorized to view certain data
- Need mechanism for accessing & integrating customer data in real-time – across many distributed databases

Top 5 Pharma – Data Provisioning for Informatics Applications



- Pipeline Pilot used at 5 sites – key data analysis tool
- Need way to serve data to remote sites that performs well, is maintainable, and extensible
- Existing homegrown transform/caching solution, but...
- Hand-crafted custom code, hard to maintain
- Required modification for each app or data source change (3 week effort)
- No multi-site, multi-database capability; not extensible

Aerospace Manufacturer – Product Design Collaboration



- Users at multiple sites (20) need access to shared design data
- Current replication approach is manual, unwieldy, with high latency (FTP)
- Multiple copies introduces data currency & consistency issues
- R&D schedules on a major new program are at risk

Avaki Background
Background on Grids
Avaki Grid Software
Standards
Wrap Up

Background:

Grid standards are now being developed at the Global Grid Forum (GGF)

In-development standard, *Open Grid Services Infrastructure* (OGSI) will extend Web Services (SOAP/XML, WSDL, etc.)

- > **Names and a two level name scheme**
- > **Factories and lifetime management**
- > **Mandatory set of interfaces, e.g., discovery interfaces**

OGSA – The Open Grid Services Architecture

- > **OGSA will define the architecture of the grid, but is a work in progress**

AVAKI taking a visible, active role at the GGF

- > **AVAKI CTO Andrew Grimshaw on GGF Steering Committee**
- > **AVAKI active in OGSA and numerous other GGF Working Groups**

Contributed Secure Grid Naming Protocol (SGNP) to OGSI WG

- > **Spec for scalable naming of grid entities, capabilities for secure, reliable communication in spite of object migration, replication, failure, etc.**

Commitment:

Respect for the standards: **AVAKI will deliver an early commercial implementations of OGSI compliant components**

Respect for customer investments: **AVAKI will interoperate with any other OGSI-compliant apps (following ratification), including future versions of the Globus toolkit**

Met last week in Hinxton, UK

Major effort is around naming and access via web services to named data objects

The name scheme developed is LSID – Life Science Identifiers

Two level name scheme, very similar to Grid Service Handles in OGSF

LSID implementations are available, and have been deployed for several different data sources.

Avaki Background
Background on Grids
Avaki Grid Software
Standards
Wrap Up

- > **Real-time access to data**
- > **Global data provisioning and access, with local control**
- > **Read *and* write capabilities**
- > **Multi-site, multi-domain**
- > **Flexible, reusable “data layer” for entire enterprise**
- > **Standards-based data integration and data flows**
- > **Support for multiple data types**
- > **Grid architecture**

For follow-up:

office: 781-272-3331
©2003 Avaki Corporation



 **AVAKI**