A Scalable and Highly Available Networked Database Architecture

Roger Bamford
Rafiul Ahad
Angelo Pruscino
Outline

• Scalability, Availability, Mobility Challenges
• Client Technologies
• Server Technologies
• Open Issues
Pre-Internet Era Database Usage

End Users

Intermediaries

End Users

1. App
2. DB
3. App
4. App

End Users

End Users
Challenges

• **Client Technologies**
  – Supporting disconnected-mode/slow connections
  – Transparent install/management of SW on client
  – Catering diverse device capabilities and preferences

• **Server Technologies**
  – Scalability
  – High availability
  – Manageability => lows cost of operation/ownership
Server Technologies

Centralized Management Console

High Speed Switch or Interconnect

Clustered Database Servers

Hub or Switch Fabric

Mirrored Disk Subsystem

Internet/Intranet

Net8

IIOP or RMI

HTML/XML

SQL

Java/Corba

NT/Unix/mainframe

SAN
Open Issues

- **Effective Load Balancing across nodes**
  - Use CPU and memory and I/O and health of node (is node running out of resources)

- **HA for long running Transaction**
  - How do we save app state for LRT and restart transaction on another node (checkpoint of application state)

- **Affinity processing in cluster environment**
  - To Optimize Locking and cache coherency (dynamically route like-work)

- **Effective support for unlike speed node**

- **Scalable replication**