Databases in a Wireless World

David Yach
Senior VP Software
Research In Motion Limited
Agenda

- Problem Statement
- Wireless Background
- Alternatives and Analysis
- Going Forward
Problem Statement

How should an application access a corporate database from a Wireless Device

Database

Firewall

Internet

Wireless Network

Wireless Device
Problem Statement

How should an application access a corporate database from a Wireless Devices" bunch of"
Problem Statement

How should an application access a corporate database from a whole bunch of wireless devices?
Problem Statement

How should an application access a corporate database from a Wireless Devices?
Potential Solutions in a Wired World

- Browser/Thin Client
- Client/Server
- Distributed database
- Replication
What’s Different About Wireless?

- Portability
  - Routing
  - In and out of coverage
  - Not reliable

- Constraints, constraints, constraints!
  - Battery Life
  - Bandwidth
  - Latency

- Asymmetric
  - Constraints
  - Managability and Control
Wireless Basics

• Radio is essentially low frequency light
  – Travels at $c=3 \times 10^8$ m/second
  – Information packed into waveforms
  – Never receive a clean signal
  – Limited spectrum range practical
• Usually only the last leg is wireless
  – Rest of network is wired
  – Dynamic routing to find the device
• Wireless bandwidth costs money
Wireless Basics

• Radio is essentially low frequency light
  - Travels at $c=3\times10^8$ m/second
  - Information packed into waveforms
  - Never
  - Limited

• Usually
  - Rest
  - Dynamic

• Wireless
Wireless Basics

- Radio is essentially low frequency light
  - Travels at $c = 3 \times 10^8$ m/second
  - Information packed into waveforms
  - Limited power
- Usually
  - Rest
  - Dynamics
- Wireless

**Base facts**

$c = 300,000 \text{ km/second}$

**Data spectrum**

$\sim 800 \text{ Mhz} - 5 \text{ Ghz}$

**Currently** $\sim $2/MB
Battery Life

- Transmit requires more power than receive
- Receive is on more than transmit
- Radio, screen/backlight, radio, audio, ...
- Not enough power to listen all the time
  - Device frequently suspends
Battery Life

- Transmit requires more power than receive
- Receive is on more than transmit
- Radio, sensors, backlight, radios, radio
- Not enough
  - Device

**Approximations**

Transmit/bit = 10 x Receive/bit

Listen 5% of the time (50ms every second)

1 battery = 5000 packets
Bandwidth

• The myth of wireless bandwidth
  – Quoted at peak, not at capacity
  – Measure # bits/second in available spectrum
  – Can’t pack signals too close together

• Increasing capacity via smaller cells
  – When size is 1/n, cost is n^2

• Radio power roughly proportional to # bits
  – More data = more battery used
Bandwidth

- The myth of wireless bandwidth
  - Quoted at peak, not at capacity
  - Measure # bits/second in available spectrum
  - Can’t guarantee
- Increase
  - When
- Radio per spectrum
  - More

Approximations

#bits/spectrum for 3G technology ~ 2x 2.5G
Latency

- High bandwidth does not imply low latency
- Possible impacts of dynamic routing, battery life
- Speed of light can be a factor
- Devices sleep frequently
  - Heuristics can reduce this after first contact
Latency

- High bandwidth does not imply low latency
- Possible impacts of dynamic routing, battery life
- Speed of light can be a factor

Approximations

Avg push latency in 2.5G or 3G ~0.5 seconds for ONE device

Other latency ~50ms

Half way around the world takes ~100ms at speed of light
Not Good Things in Wireless

- Polling
- Multi-pass protocol
- Streaming
Not Good Things in Wireless

- Polling
- Multi-pass protocol
- Streaming

**Polling**

- Checking every 5 minutes
- $12 \times 24 = 288$ transmits per day
- $\frac{288}{5000} \approx 5\%$ of battery/day with no info retrieved!
Not Good Things in Wireless

- Polling
- Multi-pass protocol
- Streaming

Multi-pass protocol
Add 50ms for each pass
N passes means 1/N battery life
Not Good Things in Wireless

- Polling
- Multi-pass protocol
- Streaming

**Streaming**

1KB every 10 seconds

6x60x24 = 8.64MB/day

8x$2/MB=$16/day

$16x30=$480/month
Good Things in Wireless

• Push
• Transferring data only once
• Don’t count on coverage
  – Complicates Push
• Efficient data transfers
  – E.g. binary instead of XML
Why Push Changes Everything

• Device
  – Device must be “Always On, Always Connected”
  – Application must be ready to receive data at any time
  – Must have proper background processing

• Infrastructure
  – Device may not be in coverage
  – Requires queuing
  – Reliability and retries are used frequently
    • Must be efficient!
    • Bad implementations can kill the network
Potential Solutions in a Wired World

- Browser/Thin Client
- Client/Server
- Distributed database
- Replication
Potential Solutions in a Wired World

- Browser/Thin Client
- Client/Server
- Distributed database
- Replication
Browser/Thin Client

Pro:
Easy to develop, deploy, manage

Database -> Application Server -> Browser

Request -> Response
Browser/Thin Client

Pro:
Easy to develop, deploy, manage

Cons:
1. May send same data again
2. Send decoration every time
3. Useless when no coverage
Browser/Thin Client

**Pro:**
Easy to develop, deploy, manage

**Cons:**
1. May send same data again
2. Send decoration every time
3. Useless when no coverage

- Database
  - Application Server
    - Caching?
  - Browser
Browser/Thin Client

Pro:
Easy to develop, deploy, manage

Cons:
1. May send same data again
2. Send decoration every time
3. Useless when no coverage
Browser/Thin Client

**Pro:**
Easy to develop, deploy, maintain

**Example:** Price quote
- 1 KB response
- 10 bytes changed data

Sending $1000/10 = 100 \times$

Using 100X battery, no benefit

2. Send decoration every time
3. Useless when no coverage

**Caching?**

**How?**

**Request**

**Response**
Client/Server

Database → Application
Client/Server

Pro:
Always live data
Decoration is on device

Database request

Database response

Database

Application
Client/Server

**Pro:**
- Always live data
- Decoration is on device

**Database request**

**Cons:**
1. May send same data again
2. May be many passes
3. Useless when no coverage

**Database response**

Application
Client/Server

**Pro:**
- Always live data
- Decoration is on device

**Cons:**
1. May send same data again
2. May be many passes
3. Useless when no coverage

How to know?
**Client/Server**

**Database request**

**Pro:**
- Always live data
- Decoration is on device

**Database response**

**Application**

**Cons:**
1. May send same data again
2. May be many passes
3. Useless when no coverage

**How to know?**
- Batch operations?
Client/Server

Pro: Always live data
Decoration

Database request

Database

Example: Price list

Fetching 1000 rows
1000 requests
1000/5000 = 20% of battery
And usually no changes!

Application

How to know?
Batch operations?

2. May be many passes
3. Useless when no coverage
Client/Server

Database request

Pro:
Always live data

Decoration

Database

Batched Price list
1000 rows × 100 bytes
10X/day, 1000 sales reps
1MB/day × 1000 × 20
days/month × $2/MB

$40,000/month!

3. Useless when no coverage

How to know:
Batch operations?
Distributed Database
Distributed Database

Distributed updates

Distributed queries

Database

Application
Distributed Database

**Pro:**

ACID

**Distributed updates**

**Distributed queries**

Database

Application

BlackBerry™
Distributed Database

Pro: ACID

Distributed updates
Distributed queries

Cons:
1. Data on unreliable media
2. Too many nodes
3. Problems when no coverage

Application
Distributed Database

Pro:

ACID

Distributed updates

Distributed queries

Example: Distributed Query

1000 nodes

5 distributed queries/day

5000/day = 1 battery!

2. Too many nodes

3. Problems when no coverage

Database

Application

media
Distributed Database

**Example: 2PC**

- 1000 nodes
- 1 node will have maximum latency = 1s
- All distributed commits = 1s

**Pro:** ACID

**Distributed updates**

**Distributed queries**

1. Too many nodes
2. Problems when no coverage
3. Lack of media

Database
Replication

Subset data

Replicate changes

Database

Application
Replication

**Pro:**

- Subset data
- Can be push based
- Minimizes data
- Works out of coverage

**Replicate changes**
Replication

Pro: 
- Subset data
- Can be push based
- Minimizes data
- Works out of coverage

Subnet data
Replicate changes

Cons:
1. Not ACID
2. Conflict Resolution
3. Device has limited capacity
4. Fire, expired updates
Replication

Pro:       Subset data
Can be push based
Minimizes data
Works out of coverage

Cons:
1. Not ACID
2. Conflict Resolution
3. Device has limited capacity
4. Fire-erased updates

Application specific?
Replication

**Pro:**
- Subset data
- Can be push based
- Minimizes data
- Works out of coverage

**Cons:**
1. Not ACID
2. Conflict Resolution
3. Device has limited capacity
4. Fire sprained updates
Replication

Pro:
- Subset data
- Can be push based
- Minimizes data
- Works out of coverage

Cons:
1. Not ACID
2. Conflict Resolution
3. Device has limited capacity
4. Fire旷ioned updates
Replication

Pro:
- Subset data
- Can be push based
- Minimizes data
- Works out of coverage

Cons:
- Not ACID
- Conflict Resolution
- Device has limited capacity
- Fine-grained updates

Application specific?
Replication

Pro: 

Can be push based
Minimizes
Works out

Subset data

Replicate changes

Database

Data capacity

50 GB database
1% required locally

50,000 x 1% = 500MB

Application

Big verbs?

2. Conflict Resolution
3. Device has limited capacity
4. Fire expired updates

Application specific?
Proposed Solution Characteristics

- Push based, optimize data transfers
- Local utility when out of coverage
  - Read
  - Changes
- Data might not be on device
- Preserve integrity of Enterprise data
- Usable
- Maintainable
- Scalable
- Affordable
Proposed Solution

Remotely manage cache

Push changes

Big verbs

Database

Application

BlackBerry
Proposed Solution

Pro:
- Push based
- Data can be remote
- Works out of coverage

Remotely manage cache
- Push changes
- Big verbs

Database

Application
Proposed Solution

**Pro:**
- Push based
- Data can be remote
- Works out of coverage

**Cons:**
1. Server has lots of state
2. Optimization heuristics
3. When is missing data okay?
Challenges in Wireless Database World

- Server based remote cache management
- Verb aggregation
- Transaction semantics with old data
- Distributed query optimization on a wireless device
Summary

- Wireless is constrained in many dimensions
• Wireless is constrained in many dimensions
• Push is everything, and changes everything
• Wireless is constrained in many dimensions
• Push is everything, and changes everything
• Difficult to find a solution that satisfies all constraints
Summary

• Wireless is constrained in many dimensions
• Push is everything, and changes everything
• Difficult to find a solution that satisfies all constraints
• Do the math!
Databases in a Wireless World

David Yach
Senior VP Software
Research In Motion Limited