Redesign ideas

Stephen Hicks

Introduction

- Web Services
- Streaming
- Refactoring

Web Services

- API's
  - User: e.g. insert(), close()
  - Factory: e.g. createConsumer()
  - System: e.g. startStreaming(), execute()
- WSDL
  - Subset of…
    - User & Factory API (public)
    - System API (private)

WS-Resource Framework

- Stateful services
  - Resources / Instances
  - Connection ID
    - SOAP Header
- WS-ResourceLifetime
  - Destroy
  - InitialTerminationTime
  - CurrentTime
  - SetTerminationTime
- WS-Addressing
  - EndpointReference
    - URL
    - Resource ID

Streaming

- All consumer-side streaming code in StreamingServer
  - Transport
    - TCP/IP
  - Data protocol
    - consumer ID first 4 bytes sent to socket
  - Data format
    - XML
  - Application logic
    - Push ResultSet onto stack
- Split into separate modules
  - Well-defined interfaces
  - Pluggable implementations

Streaming: producers

- Which thread should stream data for the producer?
  - HTTPProcessor*
    - Simple, no extra threads, user's insert() calls can block
  - One I/O thread
    - insert() returns quickly, slow socket blocks all I/O (→ no TCP flow control)
  - One thread per producer/consumer pair
    - insert() returns quickly, many threads, slow socket won't block I/O
  - One thread per machine connection
    - insert() returns quickly, slow socket won't block I/O, no TCP flow control (i.e. if stack.isFull…)

Streaming: consumers

- How should the consumer ID be sent?
  - With each Tuple / ResultSet
    - One data channel can be used for multiple consumers
      - “full” check no longer possible
    - More data sent
    - XML data must be parsed before being pushed onto stack
  - Once per “channel”*
    - One consumer per channel
- How many sockets should be used?
  - One per machine connection
    - Relatively few sockets
    - No TCP flow control
  - One per producer/consumer pair*
    - Many sockets

General refactoring

- Packaging
- SQL parsing
- Database access
- Testing
  - Refactor tests before most of code
  - Use tests to verify new code

Producers

- One producer
    - Just change API
    - Change API and implementation
- Define basic behaviour in interfaces
    - TupleStorer
        - Store tuple
        - Retrieve tuple
    - TupleStreamer
        - Stream tuple
- Plug in different implementations
    - MemoryTupleStorer, DatabaseTupleStorer
    - TCPTupleStreamer, JMSTupleStreamer