

# SPEC Enterprise Java Benchmarks State of the Art and Future Directions

Samuel Kounev

Release Manager, SPEC Java Subcommittee

Chair, SPECjms Working Group

Kai Sachs

SPECjms2007 Lead Developer

*Databases and Distributed Systems Group, TU-Darmstadt, Germany*

*SPEC Benchmarking Joint US/Europe Colloquium  
Dresden, June 22, 2007*



## Agenda



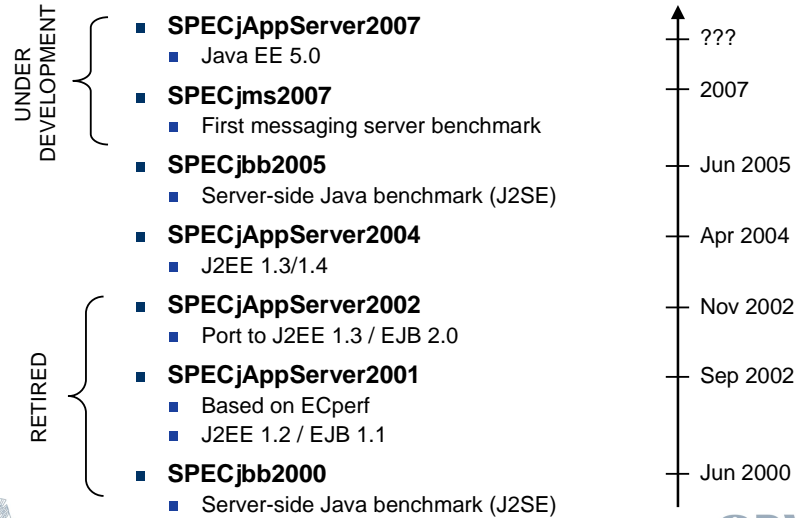
- I. Overview of enterprise Java benchmarks
- II. SPECjAppServer2004 after 3 years
- III. Towards SPECjAppServer2007
- IV. SPECjms2007
- V. Summary



2



## SPEC Enterprise Java Benchmarks



## Workload Requirements for Enterprise Java Benchmarks

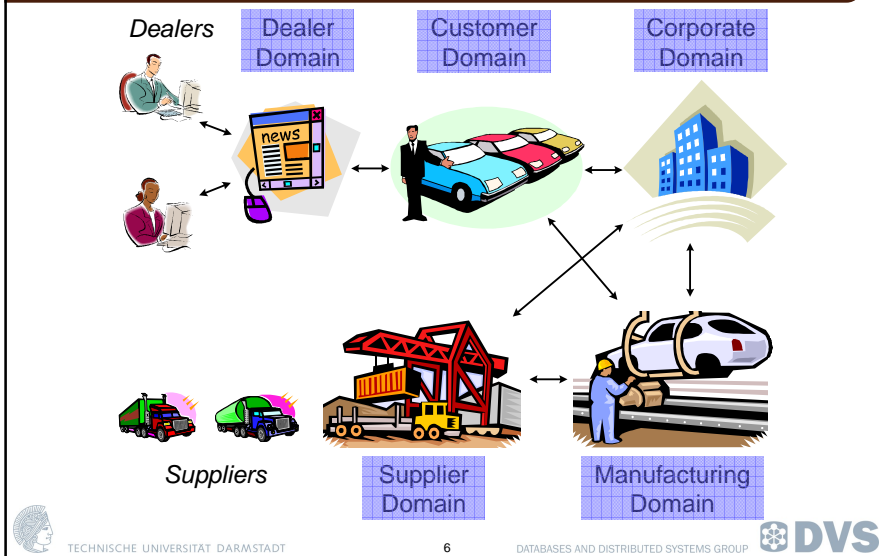
- **Representativeness**
  - Should be representative of real-life applications
- **Comprehensiveness**
  - Should exercise all platform services typically used
- **Focus**
  - Should be focused on the middle-tier
- **Scalability**
  - Must not have any inherent scalability limitations
- **[Configurability]**
  - Should be customizable/configurable



## SPEC OSG Java Subcommittee



## SPECjAppServer2004 Business Model



## Exercised Components / Services

- Software
  - ❖ J2EE application server
    - The Web container (servlets and JSPs)
    - The EJB container (session beans and entity beans)
    - Container-managed persistence
    - JMS and message driven beans
    - Transaction management
    - Database connectivity
  - ❖ DBMS software
  - ❖ OS, JVM software, JDBC drivers
- Hardware
  - ❖ Server systems
  - ❖ Storage systems
  - ❖ Network components
  - ❖ Load-balancers

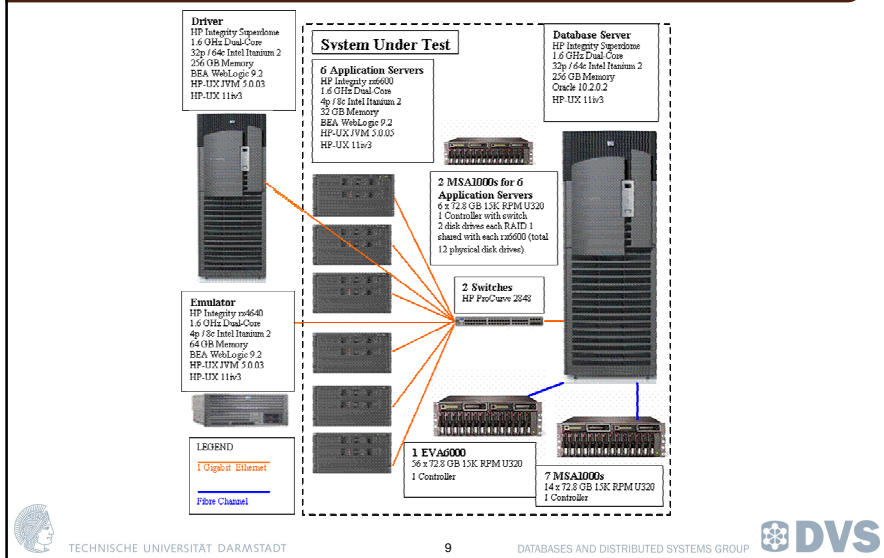


## SPECjAppServer2004

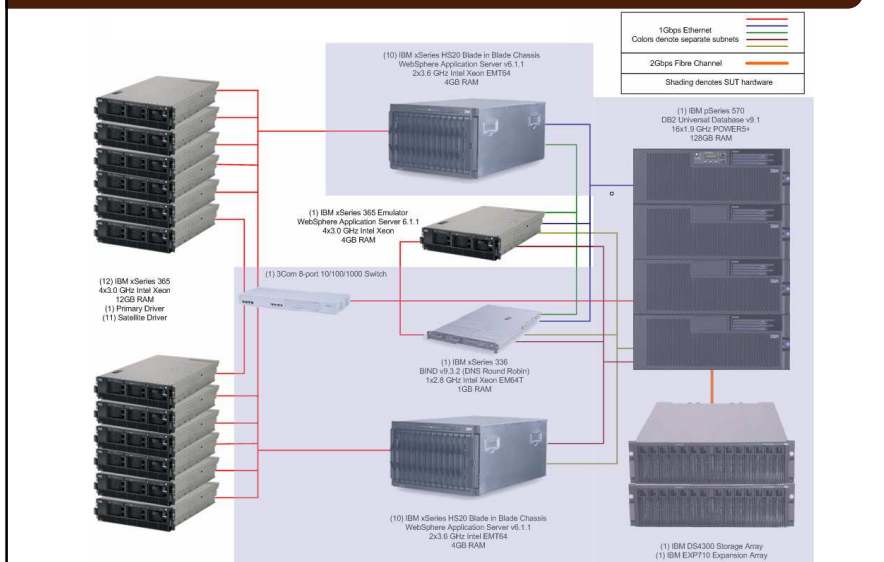
- 45 published results as of June 16, 2007
- Software platforms tested:
  - BEA WebLogic Server 9.0, 9.1, 9.2
  - Oracle Application Server 10g
  - IBM WebSphere Application Server 5.1, 6.0, 6.1
  - Sun Java System Application Server Platform Edition 8.1, 8.2, 9.0
  - Sybase EAServer
  - DBMS: Oracle Database 10g, IBM DB2 Universal Database, MySQL 5.0
- Hardware platforms tested:
  - IBM xSeries 335/365, OpenPower 720, p5 550, xSeries Blade Center
  - HP DL380, ProLiant BL685c G1, HP-UX rx2660, rx4640, rx3600, rx6600
  - Sun Fire T2000, X4100, X4600, E2900, V20z; Sun Blade 8000
  - Fujitsu SPARC Enterprise T2000
  - Inspur NF380D System



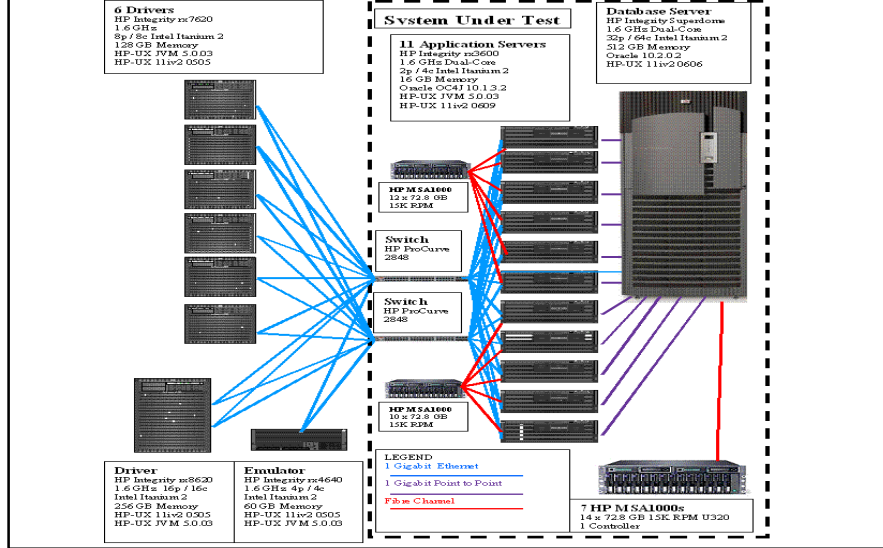
## BEA WebLogic Server 9.2 on HP Integrity rx6600 Cluster



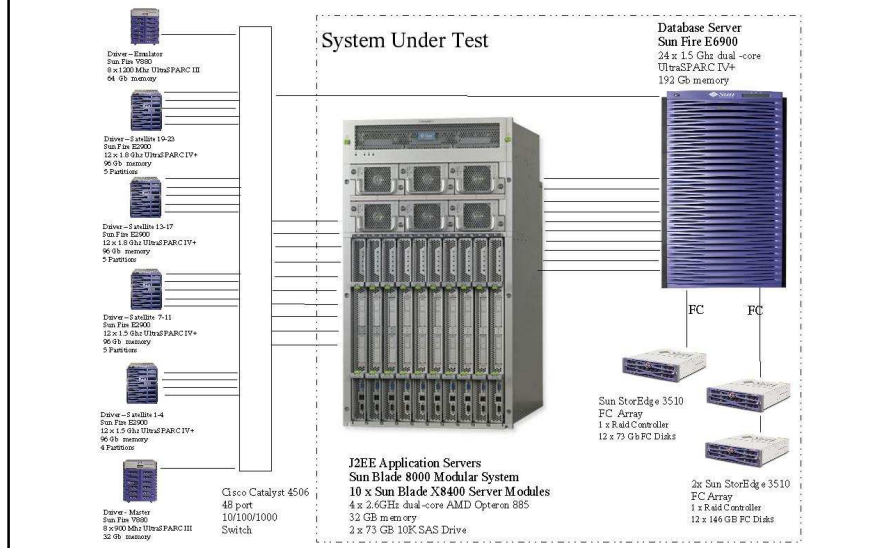
## IBM WebSphere 6.1 Application Server on xSeries Blade Center



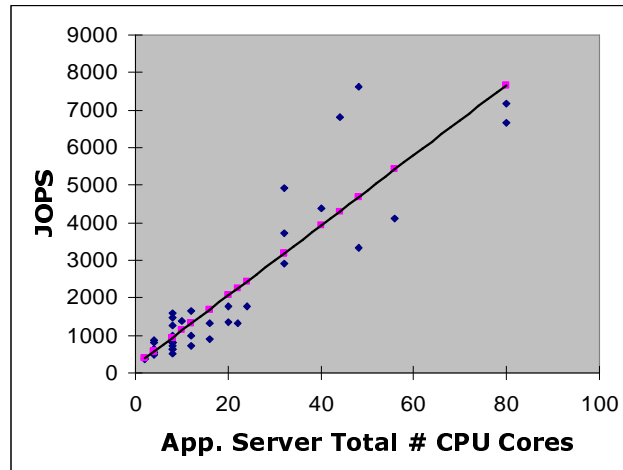
## Oracle Application Server 10g Release 10.1.3.2 on HP Integrity rx3600 Cluster



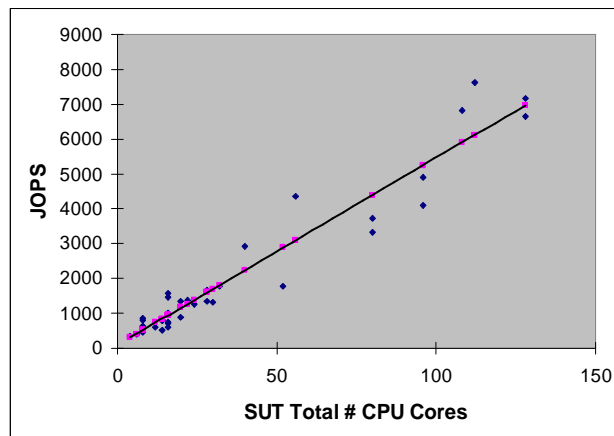
## BEA WebLogic Server 9.2 on Sun Blade 8000 Modular System



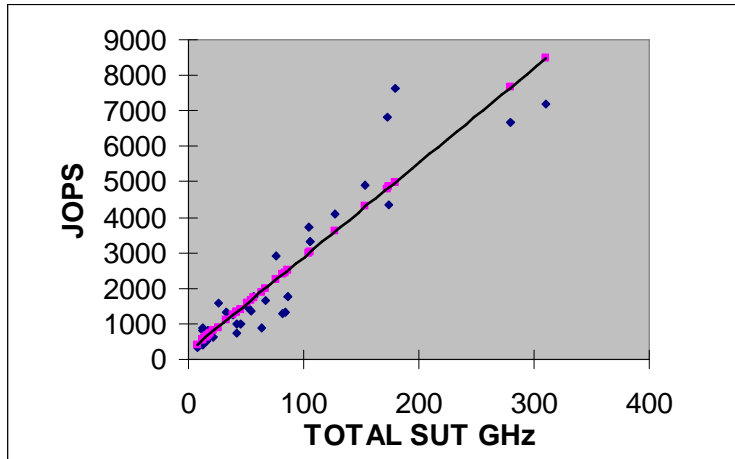
## SPECjAppServer2004 Scalability



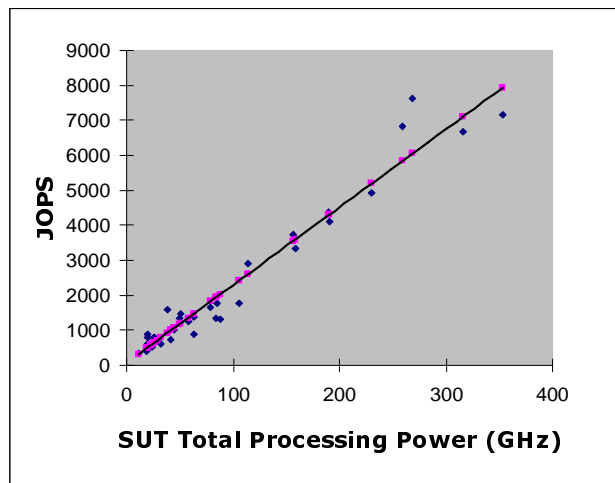
## SPECjAppServer2004 Scalability



## SPECjAppServer2004 Scalability



## SPECjAppServer2004 Scalability





## SPECjAppServer2007 Goals

- Simplify the deployment and running of the benchmark
- Improve driver
  - Simplify configuration and result collection
  - Reduce driver CPU consumption by 50%
- Migrate workload to Java EE 5.0 APIs
- Introduce new Java EE 5.0 features
  - Web services in manufacturing and supplier domains
- Improve run repeatability & predictability
- Reduce database contention

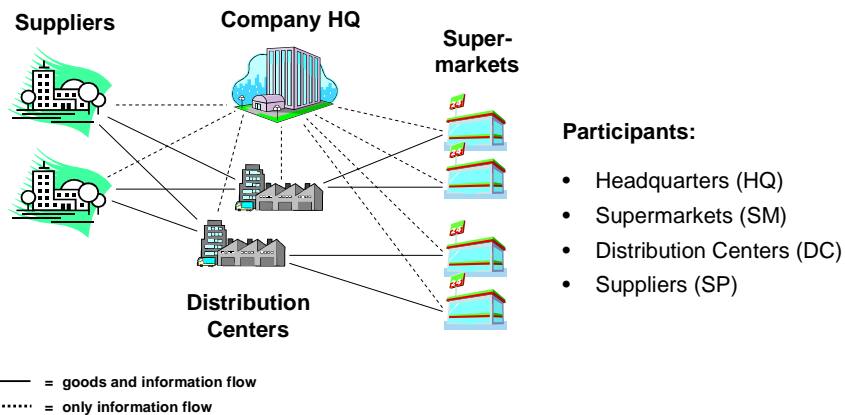


## SPECjms2007

- **World's 1st industry standard benchmark for MOM**
- Developed by the SPECjms Working Group with the participation of:
  - IBM
  - TU Darmstadt
  - Sun
  - BEA
  - Oracle
  - Sybase
  - Apache
  - JBoss



## Supermarket Supply Chain Scenario



## Modeled Interactions

The following interactions are part of the scenario:

1. Order / shipment handling btw. SM and DC
2. Order / shipment handling btw. DC and SP
3. Price updates
4. SM inventory management
5. Sales statistics collection
6. New product announcements
7. Credit card hot lists



## Back to Workload Requirements

### ■ Representativeness

- Based on a realistic workload scenario
- Features stressed weighted according to their usage in real-life systems

#### **Goal**

- Allow users to relate the observed behavior to their own applications and environments

### ■ Focus on the middle-tier

- Messaging server (software and hardware components)
- Minimize the impact of other components



## Comprehensiveness

### ■ Workload should consider the following dimensions:

- Transactional vs. non-transactional messages.
- Persistent vs. non-persistent messages.
- Different message types, e.g. TextMessages, ObjectMessages, StreamMessages or MapMessages.
- Messages of different sizes (small, medium, large).
- Point-to-point vs. pub/sub communication.
- One-to-one, one-to-many and many-to-many interactions.
- Durable vs. non-durable subscriptions.
- Ratio of message producers to message consumers.



## Scalability

- The workload can be scaled along two dimensions:
  - **Horizontal scaling / topology**
    - Increase the number of destinations (queues and topics)
    - Keep the traffic per destination fixed
  - **Vertical scaling / topology**
    - Increase traffic per destination
    - Keep the number of destinations fixed
- Additionally: Support for **freeform scaling / topology**
  - user defines number of destinations and traffic per destination
- Scaling factor used as metric
  - SPECjms2007@Horizontal
  - SPECjms2007@Vertical



## Configurability

- Goals
  - Provide a flexible framework for performance analysis of MOM
  - Allow user to customize the workload to stress specific aspects
  - Use interactions as building blocks
- Interactions must be decoupled
  - Standard mode: Ensures that dependencies between interactions are preserved as defined in the application scenario
  - Freeform mode: Run interactions in different combinations depending on the desired transaction mix
- Over 150 configuration parameters available to customize the workload



## SPECjms2007 Status

- Project started in Sept. 2005
- 18 releases
- Currently tested on the following platforms
  - IBM WebSphere AS
  - IBM WebSphere MQ
  - Sun Java System MQ
  - Sun JMS Grid
  - BEA WebLogic Server
  - Oracle Application Server
  - Apache Active MQ
  - Sybase EAServer



## Summary & Conclusions

- Main challenges in enterprise Java benchmarking
  - Ensuring benchmark scalability
  - Providing maximum
    - ease-of-use
    - flexibility
    - configurability
  - Defining representative metrics
- SPECjAppServer2004
  - 45 results published since April 2004
  - Results exhibit linear scalability
  - 75+ copies sold to non-members



## Summary & Conclusions (cont.)

- Increasing use of benchmarks for “in-house” performance analysis and research in academic environments
  - Validation of product performance and scalability
  - Performance tuning and optimization
  - In-depth performance analysis
  - Research in performance engineering
- SPECjms2007 about to become the world’s 1st industry standard benchmark for messaging platforms
  - In addition to standard metrics and workload, provides a full-blown performance analysis framework for messaging platforms



## Thank You For Your Attention!

### QUESTIONS

For copy of the slides go to

<http://www.dvs1.informatik.tu-darmstadt.de/~skounev>

