

ApacheCon

Open source source Adoption

Ravi Saraswathi
Ravisar@gmail.com
Information Technology Executive



Leading the Wave
of Open Source

ApacheCon

- Open Source Adoption Model
- Middleware Standards
- Open Source Migration Strategy
- How we did the migration to open source ?
- Lessons Learned



Leading the Wave
of Open Source

ApacheCon

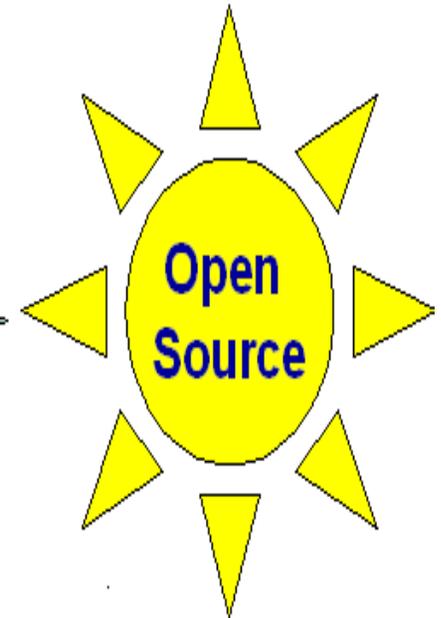
Build Open Source Java Applications

Less Complex, Low TCO,
High-Performance, Scalable



Leading the Wave
of Open Source

Open Source Success Story



Think as a Business Owner

Uses both ?

Success Story

Reduced Commercial Software inventory and saved over \$MM in 9 months.

Migrated to Apache Open Source.





Fact

As per one estimation,

10 dedicated resource for an year

==

Maintenance cost of one commercial software for an year !

Pricing sheet?

Some of the Web Startups
the software license cost is Zero !

(Eg: LAMP – Linux(Cent OS), Apache, MySQL, PHP)

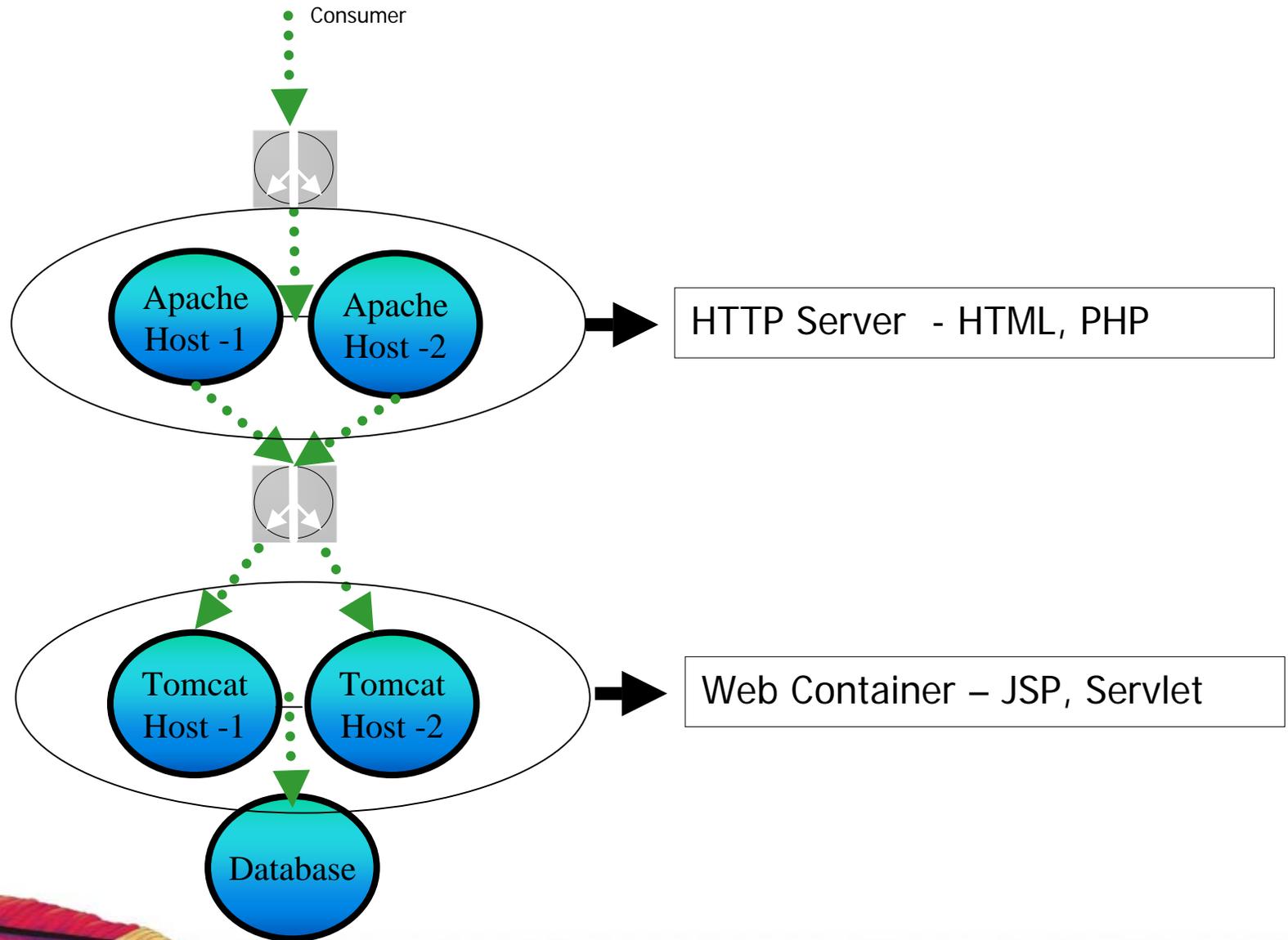


For Simple Web Applications:

- Why do we need EJBs ?
- Complex propriety Frameworks ? (MVC rocks !)
- Vendor APIs ?



A Typical Web application System Architecture





Middleware Standards

<u>Technologies</u>	<u>Standard</u>
• Static web pages (HTML)	✓ Apache
• HTML, JSP, Servlet	✓ Tomcat
• Messaging	✓ Active MQ
• Search	Lucene via Solr
• Database	MySQL
• ETL	Talend



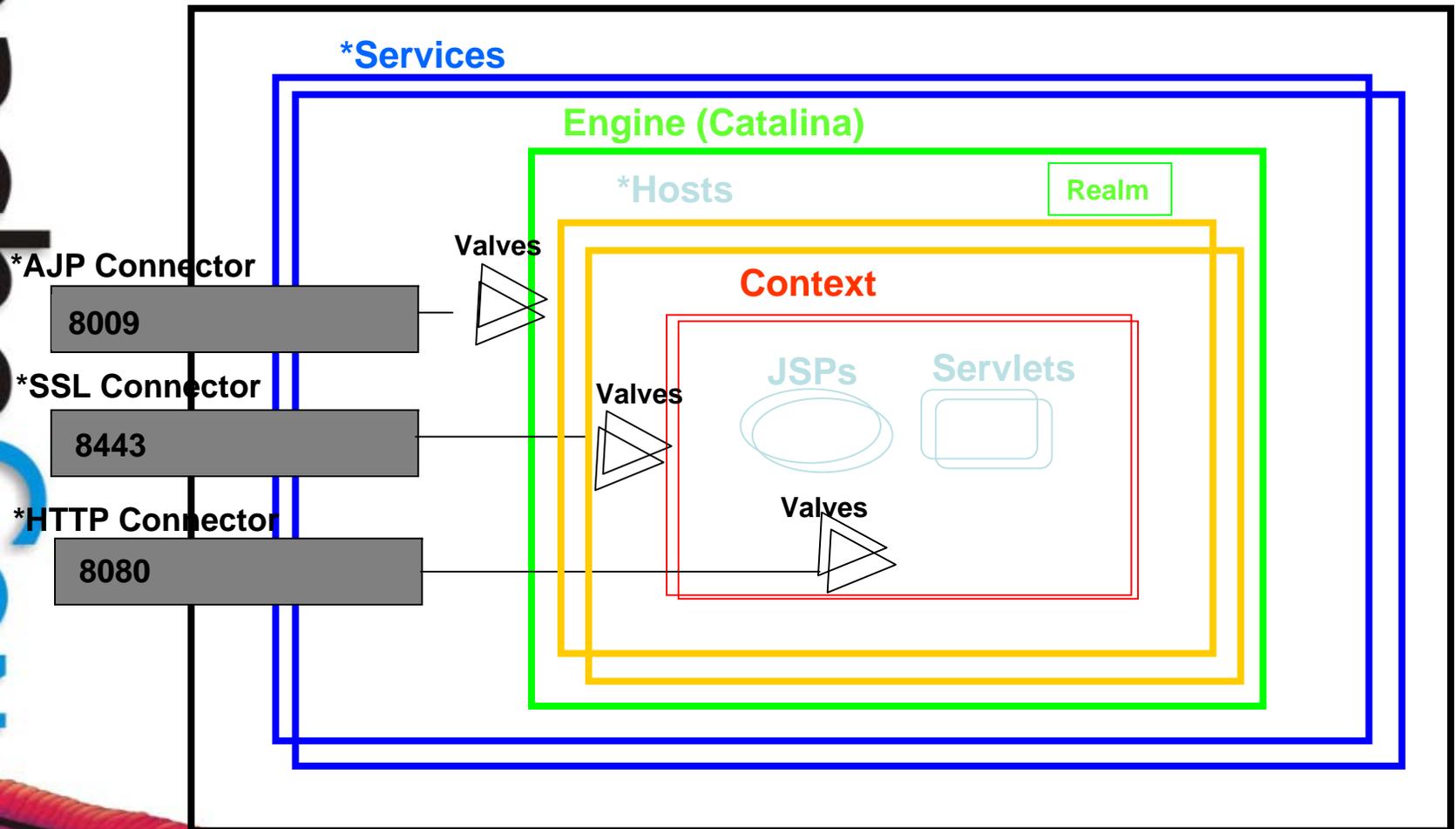
Middleware Standards



<i>ESB</i>	<ul style="list-style-type: none">✓ Messaging,✓ Web services,✓ Data transformation,✓ Intelligent routing
<i>ActiveMQ</i>	<ul style="list-style-type: none">✓ Message Oriented Middleware
<i>CXF</i>	<ul style="list-style-type: none">✓ Web Service Implementation Framework
<i>Camel</i>	<ul style="list-style-type: none">✓ Routing✓ Transformation✓ Mediation

Tomcat Server Topology

Tomcat – The Server





Set Guidelines E.g: Mod_JK vs. Mod_Proxy

Mod_jk

- Mod_JK = AJP Connector
- Web server plug-in/
Tomcat Redirector
- Supports load balancing
- Native mod_jk download
available for different OS.
- Very Fast
- Complex Configuration
- Let Apache serve static
pages

Mod_Proxy

- Connectors available for FTP, HTTP,
AJP or HTTPS
- mod_proxy, mod_proxy_balancer,
mod_proxy_http, and mod_proxy_ajp
- Easy to Configure
- Mod_proxy modules comes as part of
apache
- Lacks sticky session load balancing.
mod_proxy_ajp supports load
balancing



Web Applications - Software

Category	Open Source Software Products
Servlet Container	Tomcat, Jetty
EJB Container	JBoss, Glassfish, Resin, Geronimo, JOnAS
Portal	Liferay, JBoss Portal
Enterprise Service Bus	Mule, Apache ServiceMix, Synapse, OpenESB
Registry	Juddi
IDE	Eclipse, Netbeans, JEdit, DrJava, J, Cube-J
Content management system	OpenCMS, Apache Lenya, Dspace, InfoGlue
Cache Solutions	JCS, Terracotta, OSCache, EHCACHE, JCache



Evaluation Criteria

- **Operations Parameters**
 - License & Support Model
 - Architecture & Platform Support
 - New Technology Adoption
 - Performance, Scalability, Maintainability and Security
- **Ready for production use ?**
 - Have benchmarked it and used in your organization ?
 - Tools available for configure, standardize and deploy
 - Available documentations
 - Stable versions available: binary versions
 - Training options: on-site and regular open enrollment
 - Support options, previous versions and SLAs
 - Release cycle, patches, new OS support.
 - Integration with other (commercial) products.
 - Enterprise class, good standards support

Evaluation Criteria

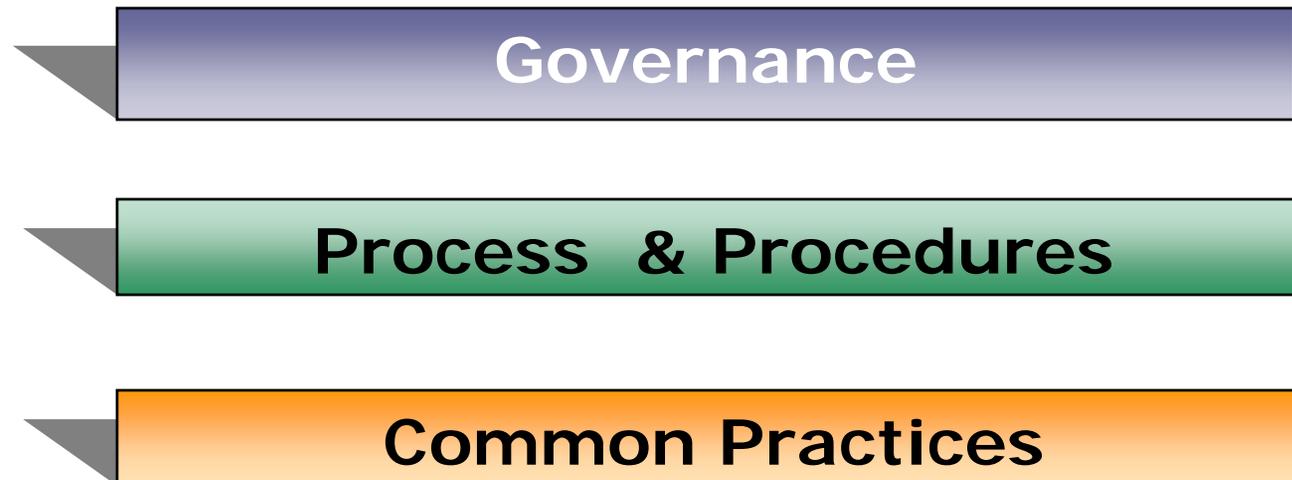
- **Implementation by major companies**
 - As they move companies, do they stay with the technology?
 - Is the user group and contributor base increasing?
- **Commitment to technology**
 - Participate in standards efforts
 - Put out papers positioning points of view.
 - Accommodate new approaches. (Adaptable)?
- **Sponsor and commitment of user base**
 - Long term strategic deployment
 - Funding or donating new features.
 - Vendor involvement - cooperation by SW and HW companies

Open Source Adoption Model



Resistance within Organization ?

Enterprise Open Source Adoption Model





What Open Source Adoption Means ?

Governance

- Open Source implementation Vision
- Policies , Roadmap & Architecture
- Communications Plan
- Training Plan
- Staffing Strategy and Fulfillment

What Open Source Implementation Means ?

Process & Procedures

- Test Strategy & Plans
- Change Mgmt
- Deployment & Release
- Production Execution
- Ops Support Model
- Adoption Timeline





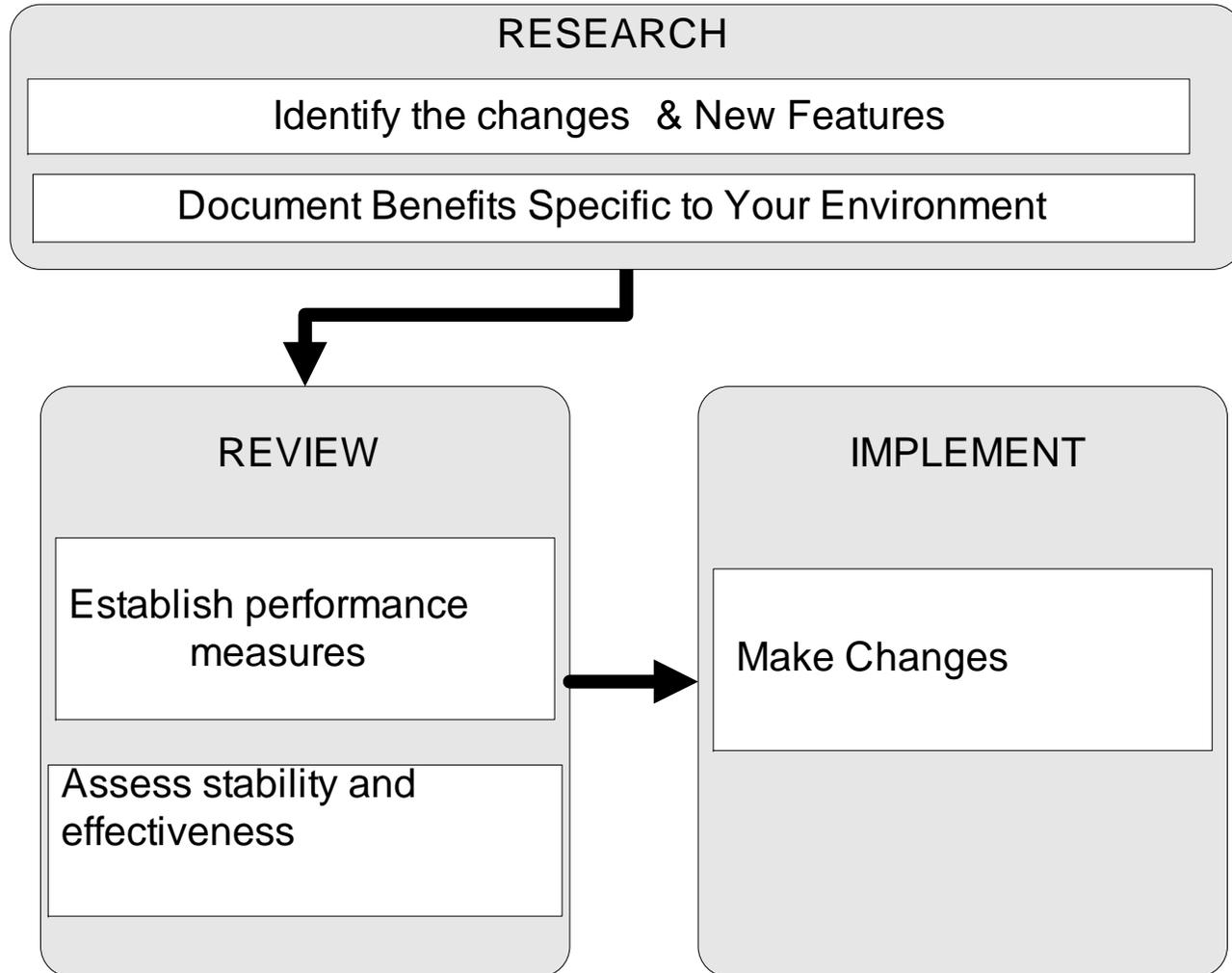
What Open Source Implementation Means ?

Common Practices

- Procurement & Build
- Infrastructure and Eng.
- Installation Guidelines
- Security & Performance
- Capacity Tuning
- Integration Approach



Adopt a Process



Challenges, Benefits, Approach





Benefits

- Standardization & Simplicity
- Reduced TCO
- Reduced dependence on vendors
- Easier customization
- Increase Developer productivity



Challenges

- Changing the mind set
- Execute support
- Education



Implementation approach

- Create success stories
- Finalize implementation methodology
- Use similar approach for every project

Sample Infrastructure Catalog

WEBSERVER
Apache, IIS

APPLICATION SERVER
Tomcat, WebLogic, Jboss, WebLogic
Integration, Portal, Aqualogic Service Bus,
ColdFusion, WebSphere

COTS PRODUCTS
SAP, PeopleSoft

BATCH PROCESSING
Auto Sys, Windows Task Scheduler

DATAWAREHOUSE - ETL
Ab initio, Perl

DATABASE
Sybase, Oracle, MySQL, Netezza

OPERATING SYSTEM
Linux, Windows, HP-UX, Solaris,

HARDWARE
Dell, HP, Sun, Egenera

LOAD BALANCER
Netscaler, Foundry Switches



How we did Migration from commercial software to Apache open Source ?



Migration Strategy

- License Inventory
- Resources & Plan
- Tracking Progress.



Approach

- Delete unwanted licenses
- Downgrade licenses
- Consolidate Environments
- Migrate to open source
 - Target simple applications first
 - Select one mission critical application



Step 1 - License Inventory

Use Discovery Tools & Perl Scripts

Products	Licenses	CPU	Cost
Commercial Software 1	466	(1151)	\$xx
Commercial Software 2	82	(233)	\$xx
Commercial Software 3	116	(283)	\$xx
Commercial Software 4	32	(62)	\$xx

Total :	1495	(3029)	\$xx



Step 2 - Report by Group

Use Discovery Tools & Perl Scripts

OPS-Group	Licenses	cpus	Total cost
Operation group – 1	166	342	\$X,XXXX,000
Operation group – 2	368	639	\$X,XXXX,000
Operation group – 3	69	198	\$X,XXX,000
Operation group – 4	101	222	\$X,XXX,000
Operation group – 5	102	237	\$X,XXXX,000
Operation group – 6	39	91	\$X,XXX,000
Operation group – 7	47	109	\$X,XXXX,000
Operation group – 8	65	131	\$X,XXX,000



Standardize - Software

Technology	Default/Approved Product
Java Application Server	Tomcat 5.5.25-1 with JDK1.5.0_11
Webserver	Apache httpd 2.2.6
Database	MySQL 5.0.67
Alternate Application Server	PHP 5.2.5 running on Apache.
Operating System	CentOS 5.2



Implementation issues - JVM

- **Memory Leak**
- **Out of Memory Error**
- **Memory exceptions** (StackOverflowError etc.)
- **JVM Crashes**
 - A fault in JVM
 - A fault in native JNI code run by JVM
- **JVM hangs – Stop Responding**
 - ✓ Application in infinite loop
 - ✓ A dead lock in application
- **Poor Performance**
 - Other process consumes lot of memory/CPU
 - I/O bottlenecks

Resources used for JVM trouble shooting

➤ Log files messages

- ✓ -verbose:gc, -XX:+PrintGC, -XX:+PrintGCDetails, -XX:+PrintGCTimeStamps, -XX:+HeapDumpOnOutOfMemoryError

➤ Java VisualVM

- ✓ Available from JDK 6. Federates several existing tools, including JConsole, jstat, jinfo, jstack, and jmap

➤ Other Tools

- ✓ Argus, Jprobe, Jconsole, HPROF - Heap Profiler, Eclipse Memory Analyzer, IBM Heap & thread analyzer

➤ JVM Utilities

- ✓ Jps, Jmap, Jinfo, Jhat, Jstack

➤ Unix Tools

- Top, vmstat, ulimit -a , sar -r , mpstat, netstat

Use right tools

Technology	Default/Approved Product
Java Application Server	Tomcat 5.5.25- with JDK1.5.0_11
Webserver	Apache httpd 2.2.6
Database	MySQL 5.0.67
Alternate Application Server	PHP 5.2.5 running on Apache.
Operating System	CentOS 5.2



Standardize – E.g. JVM Parameters

patches
CON

- Xms1024m \ - Minimum Heap
- Xmx1024m \ - Maximum Heap
- Xss512k \ - Set maximum native stack size for any thread
- XX:+UseConcMarkSweepGC
- XX:GCTimeRatio=99 \ - The ratio of GC time to application time
- verbose:gc \
- Xloggc:\$CATALINA_BASE/logs/gc_log \ - Log GC actions to file.
- XX:+PrintGCDetails \
- XX:+PrintGCTimeStamps \
- XX:MaxGCPauseMillis=20 \ - Pause times of 20ms desired.
- XX:MaxNewSize=384m \
- XX:MaxPermSize=512m \
- XX:NewSize=384m \ - Minor GC (Young Generation)
- XX:PermSize=256m \ - Permanent Size
- XX:SurvivorRatio=6 \
- XX:+HeapDumpOnOutOfMemoryError \

Tuning – E.g. JVM

JVM Parameters	Guidelines	Guidelines
-XX:NewSize=128m (Young Generation -Min)	High means - Less often Minor collection occur	Set -XX:NewSize to be one-fourth the size of the max heap size.
-XX:MaxNewSize=128m (Young Generation-Max)		Larger Young Generation means smaller Tenured
-Xms512m & -Xmx512m	60 % of available memory	Set min heap (-Xms) equal to the max heap size (-Xmx) to minimize garbage collections.
Very High Heap size	Full GC is slower & GC Occurs less frequently.	Setting too high can cause wasted memory. 32-bit OS cap the heap size t between 1.5 and 2.5GB

Minimize Software & Versions

- **Efficiency of day-to-day operations**
- **Quality and speed of deployments**
- **Lower cost interoperability**

- **Operational best practices**
- **managing and monitoring web applications**
- **Ensure optimal leverage of staff**





Lessons learned

- **Identify and implement an adoption strategy**
- **Standardization**
- **Management support**
- **Rewards and Recognition.**
- **Guidelines and best practices.**

Why Open Source ?

- Open Source Financial Advantage
- Open source Business advantage
 - Open source community & Strategy
- Open Source Technical advantage
 - Easy integration & Manageability
 - Actual developers.



Cost savings

Think as a Business Owner

- ✓ License & Support
- ✓ Data center cost (Hardware, power etc)
- ✓ Resource cost savings



Key Success Factors

Conduct detailed due diligence at planning stage

- Identify all dependencies
- Validate all assumptions before doing effort estimation
- Prepare realistic schedule and validate with customer
- Account for all parties' effort and activities

Gain understanding expectations:

- Deployment coordination
- Systems integration
- Performance testing

Identify and Assess Risks

- Communicate pro-actively
- Escalate problems in a timely manner

The Path Forward

- Building a Team (Goals & Analysis)
- Communication (TCO, Evaluation)
- Implementation (Resources)





Thank You !

Ravi Saraswathi

Ravisar@gmail.com

www.ravisar.com

www.ravisaraswathi.com