Oracle Coherence
2c/2p on Monitoring, Profiling, Right-Sizing

Santiago Martin-Romani
santiago_martin@yahoo.com
Software Architect, Deutsche Bank

Reasons to use Oracle Coherence

- Application layer friendly
- Single holistic view!
- Scalable!
- Fast, in-memory speeds, data oriented, event driven!
- Search, Analyze and Process data in parallel!
  - Java, C#/.NET, C++ objects
  - Or anything that can fit in a POF stream really..

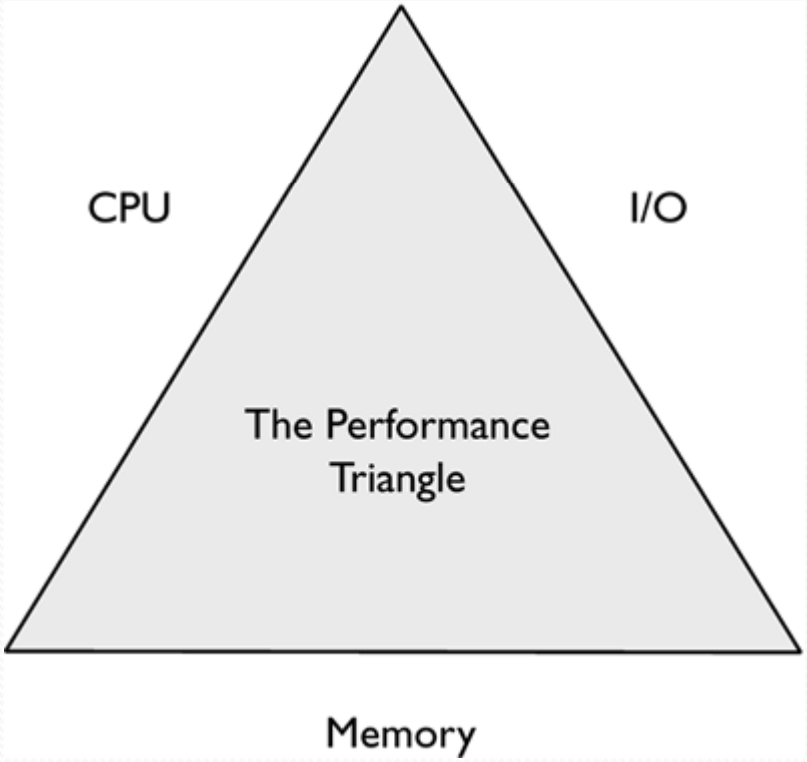
Available Operations

- Key-based operations: put/putAll, get/getAll
- Ad-hoc Queries, CQs, Aggregations, EntryProcessors
- Real-time event subscription
- Ordered inserts/updates, relaxed ACID properties for performance

Oracle Coherence - sizing building blocks

- Cluster, Distributed Cache Services, caches
- Storage enabled JVMs, Storage disabled JVMs
- Proxy JVMs, thread pools
- Extend clients (e.g. Java, C#/.NET, C++), REST API

Performance Triangle – where your application logic fits in..



Qualify your Access Pattern first, then build, test, and right-size, profile, monitor…

- How much data, what type of data, data model dependencies?
  - Data loading/refreshing
  - Expiration/Eviction
  - Latency, Throughput
- Who are the publishers, consumers?
  - How many? Where? Java, C#.NET, C++? REST API?
- Operations per unit of time? What operations (over time)?
  - Data cycle
  - Compute cycle
  - Statistics: Average, Standard Deviation, Percentile targets?
- Application life cycle, load requirements for R1, R2, R3..?
- What SLA, DR/BCP strategy, RTO?
  - What budget, timelines?

Oracle Coherence – Stability Risks Factors

- User and data load causing problem; excessive load!?
- Latency, bottlenecks, timeouts
- Did Coherence survey my code/config changes?
- Long Garbage Collection, Low Memory (before an OOME)
- Network Communication failures
- Cache sizing errors ← easier to manage
- Hotspots (logical, systemic)
- Serialization
- Un-indexed queries, query resultsets too big
- “Rogue” users
- Swapping ← avoid always!

What is RTView OCM? How does it help?

- RTView OCM – Oracle Coherence Monitor
  - RTView OCM provides over 50 out-of-the-box views of all monitoring metrics (e.g. cluster, service, cache, node, JVM, storage, network, etc.), organized in ways especially useful for troubleshooting or cluster analysis, in both development and operational scenarios. It can be extended. It provides near-real time alerts out of the box.

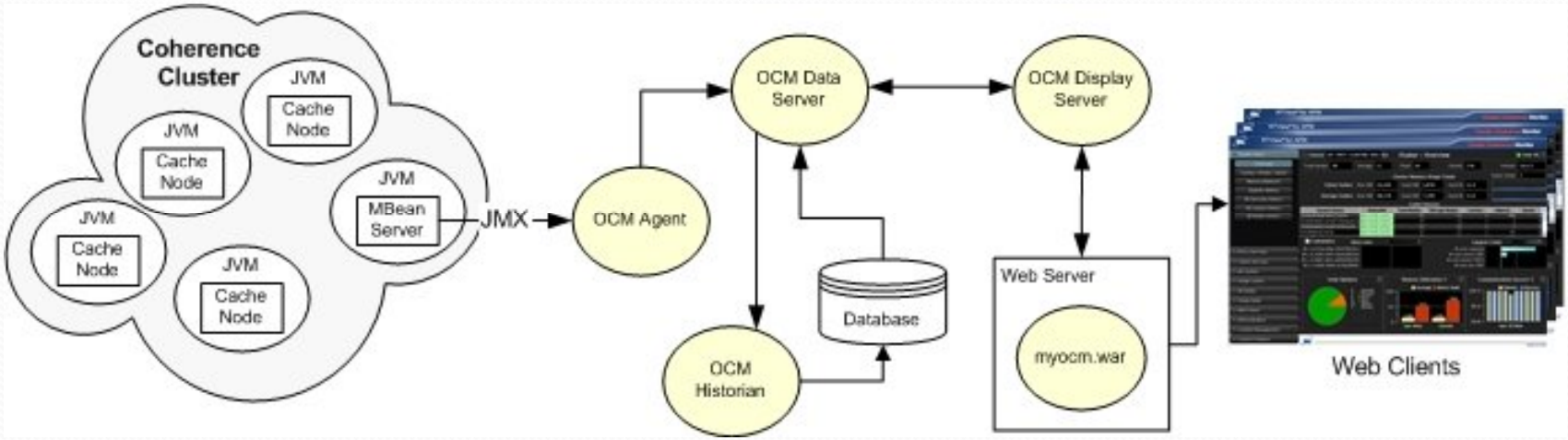
What is RTView OCM?

- Monitoring, profiling, right-sizing, troubleshooting software vendor tool.
- Real-time and historical data visualization and analysis of metrics
- Lightweight and scalable solution for monitoring the health and performance of Coherence clusters in ALL environments.
- Makes available ALL monitoring data exposed by Coherence through comprehensive views in a variety of forms.
- Efficient JMX data collection.
- Generates alerts on exceptional conditions.
- Improves productivity by helping to understand the behavior of Coherence and the effect of implementation changes.

What is it not?

- A view to Java objects in caches
- Java object modifier
- A controller/manager to the cluster
- It is not like ITRS Geneos, Nagios – RTView OCM only monitors Coherence!
- A JVM profiler
- Can't be used to monitor multiple clusters through one RTView instance

RTView OCM - Architecture



RTView OCM – Web client

Intelligent Alerting in RTView OCM

- What are the benefits?
  - Alerts notify users of possible problems and/or trends on near real-time.
  - A custom alert can notify users of meaningful events. For example:
    - Can be used to do application instrumentation propagating any valuable information/condition/event (e.g. your code Exceptions, Errors, performance measurements, any info or event such as SLA/QoS related and more – MANY POSIBILITIES!)
    - Or simply communicating: “Life is good” (e.g. app doing what is supposed to be doing, app meeting SLAs).

Alerting in RTView OCM

RTView comes with 20+ pre-defined alerts for caches, nodes and clusters, most are enabled by default...

- AvailableMemoryLowCluster
- AvailableMemoryLowNode
- AvailableMemoryLowNodeSpike
- BadCommunicationCluster
- BadCommunicationNode
- BadCommunicationNodesIntimeRange
- CapacityLimitAllCaches
- CapacityLimitCache
- DepartedNode
- DepartedNodesPercentage
- EndangeredAllCaches
- HighPendingRequestNode
- HighGCDutyCycleNode
- HighTaskBacklogNode
- HighThreadAbandonedNode
- LongGCDurationNode
- ObjectCountDeltaUpCache
- ObjectCountDeltaDownCache

Demo

Q&A