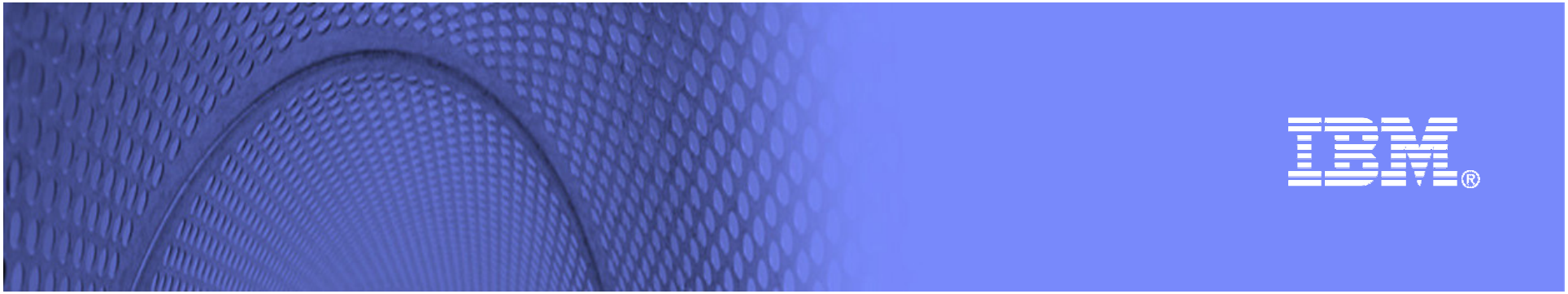
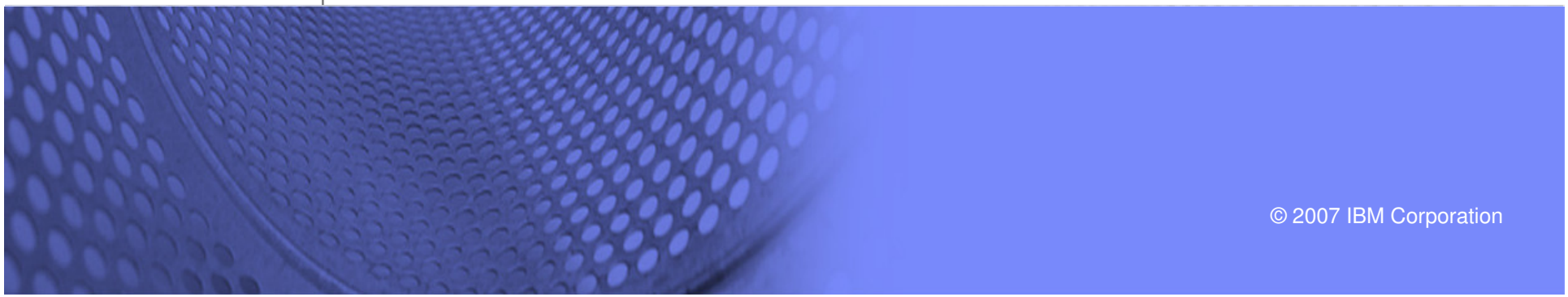


Java 5.0 Reliability, Availability and Serviceability

Java 5 Diagnostic Tools and Capabilities

Chris Bailey baileyc@uk.ibm.com
Trent Gray-Donald trent@ca.ibm.com



Agenda

- History
- 5.0 VM Problem Determination recap
- Strategic direction
- Tools dive
- Education / IBM Support Assistant
- Discussion / Questions

History

- Fragmented tooling story
  - Different tools for different folks, find tools in various places
    - In the JDK itself
    - alphaWorks
    - developerWorks
    - From Java|WAS support
  - Tools JVM level specific
- Substantial technology changes in underlying JVM implementation between 1.4.2 and 5.0
  - Significant robustness improvements (better compaction / fragmentation support, enhanced FFDC)
  - Fundamental PD data produced in same format

Java Problem Determination Strategy

- Centralization of Tools
  - Central customer visible repository of supported, maintained tools
  - Extensible, open tools, with programmable extension interfaces
- Tools must be usable everywhere
  - GUI mode for interactive use
  - Report generation for headless environments
- Documentation
  - Improvements to problem determination doc
  - Aggregated search in IBM Support Assistant
- Iterate!
  - Tools being deployed very regularly – looking for customer feedback.

Recap of JDK 5.0 Problem Determination facilities

- -Xtrace
  - Always on FFDC trace for major components
  - Separate buffers for GC – allows ‘flight recorder’
  - Method parameters available (if enabled)
- -Xdump
  - New triggers (thread start/stop, GC, heap expand, etc..)
  - More naming configuration allowed (date/time/etc..)
- java.lang.Management
  - Java level APIs to introspect into running system

Java Trace Engine

- Thousands of tracepoints throughout VM native code
- Can capture Java Methods data also
- Many different operation modes to aid debugging and problem diagnosis
- Limited “Flight Recorder” trace set is always on in Java 5.0
  - Key vm tracepoints constantly traced into per thread wrapping memory buffers
  - GCLogger tracepoints are stored in separate buffer to ensure they are not overwritten by the high frequency tracepoints

Controlling Trace

- The Trace Engine can be controlled through a number of mechanisms:
  - Through the -Xtrace command-line option
  - Using a trace properties file
  - Dynamically using Java code through the com.ibm.jvm.Trace API
  - Using trace trigger events
  - From an external agent using the C-based JVM RAS Interface (JVMRI)
- Primary way is via the -Xtrace option on the command line.

Basics of Activating Trace

- The first thing that you need to determine is the destination to which the trace output should be directed
  - `minimal` Trace identifier and timestamp only to in-core buffer.
  - `maximal` Trace identifier, timestamp and associated data to in-core buffer.
  - `count` Report the number of times a selected tracepoint is called
  - `print` Trace selected tracepoints to stderr with no indentation.
  - `iprint` Trace selected tracepoints to stderr with indentation.
  - `external` Route selected tracepoints to a JVMRI listener.
  - `exception` Trace selected tracepoints to an in-core buffer reserved for exceptions.
- The value of each keyword is then set to the trace points required
  - `-Xtrace:maximal=all` traces all of the information available from all JVM trace points to internal wrapping buffers.
  - `-Xtrace:iprint=awt` traces all of the JVM internal AWT trace points to stderr, with Indentations on entry and exit.
  - `-Xtrace:iprint=mt` activates method trace and set the output to stderr with indentations

Putting Trace Into Files

- Data can be written to file as an extension to in-storage trace
- To specify that the output of `minimal` or `maximal` trace options should be written to a file, the `output` keyword should be used
  - `-Xtrace:maximal=all,output=trace.out`
    traces into a file called trace.out.
  - `-Xtrace:maximal=all,output={trace.out,5m}`
    traces into a file called trace.out and wraps within the file once it has reached 5MB in size.
  - `-Xtrace:maximal=all,output={trace#.out,5m,5}`
    traces sequentially into five files, each 5MB in size, with # substituted for the file iteration number.
- It is also possible to put the following substitutions into the file name:
  - `%p:` The ID for the Java process.
  - `%d:` The current date, in yyyymmdd format.
  - `%t:` The current time, in hhmmss format.

Method Trace

- Instrumentation free tracing of:
  - Any Java Methods:
    - Core Java API, Middleware, 3rd party code, Customer code
  - Detailed information:
    - Entry and Exit points, with thread information and microsecond time stamps
- Two part invocation:
  - 1) Add `methods` keyword as a token to –Xtrace
  - 2) Use of `mt` as value to destination keyword (eg. maximal, print)
- Method selection by class name or method name
  - use of wildcards, along with the not operator, !, allowing for complex selection criteria.
    - `-Xtrace:print=mt,methods={*.*,!java/lang/*.*}`
      Write method trace to stderr for all methods and for all classes except those in the java.lang package.
    - `-Xtrace:maximal=mt,output=trace.out,methods={tests/mytest/*.*}`
      Write method trace to file for all methods in the tests.mytest package. (Note that this option selects only the methods that are to be traced.)

Triggering

- Trace can produce significant amounts of data
- The trace engine provides the ability to trigger on trace events
  - Provides ability to create targeted trace output
    - Reduces volume of data and performance impact
  - Provides ability to also generate dumps on trace points
- Following actions available as value of `trigger` keyword

`suspend`	Suspend *all* tracing
`resume`	Resume *all* tracing (except for threads suspended by resumecount property and Trace.suspendThis() calls).
`suspendthis`	Increment the suspend count for this thread. A non-zero suspend count prevents tracing for the thread.
`resumethis`	Decrement the suspend count for this thread if it is greater than zero. If the suspend count reaches zero, tracing for this thread will be resumed.
`sysdump`	Produce a non-destructive system dump.
`javadump`	Produce a Java dump.
`heapdump`	Produce a heap dump.
`snap`	Snap all active trace buffers to a file in the current working directory.

Triggering Examples

- Following format used to specify triggers in method events:
- `trigger=method{method spec,entry action,exit action,delay count,match count}`
  - On entering any method that matches the method spec indicated, the entry action is executed.
  - When exiting the method, the exit action is performed.
  - If the delay count is specified, the entry and exit actions are only carried out when entry and exit have occurred more times than the delay count.
  - If the match count is specified, the actions are only carried out a maximum of that many times.
- Examples:
  - `-Xtrace:trigger=method{java/lang/StackOverflowError*, sysdump}`
    create a system dump on the first (and only the first) instance of a StackOverflowError method being called - which is the <clinit> method.
  - `-Xtrace:resumecount=1`
  - `-Xtrace:trigger=method{HelloWorld.main,resume,suspend}`
    trace all threads once HelloWorld.main() is called and stop tracing when HelloWorld.main() returns.

Triggering and Method Trace in Action

-Xtrace:print=mt,methods={myapp/MyTime*},resumecount=1,trigger=method{myapp/MyTime.main,resume,suspend}

```
21:05:47.992*0x806cb00      mt.3         > myapp/MyTime.main([Ljava/lang/String;)V Bytecode static method
21:05:47.994 0x806cb00      mt.19         - Static method arguments: ([L@55D8CB98)
21:05:47.994 0x806cb00      mt.0          > myapp/MyTime.<init>()V Bytecode method, This = 809baec
21:05:47.994 0x806cb00      mt.18          - Instance method receiver: myapp/MyTime@55D8CBA8 arguments: ()
21:05:47.994 0x806cb00      mt.6          < myapp/MyTime.<init>()V Bytecode method
21:05:47.994 0x806cb00      mt.0          > myapp/MyTime.test()V Bytecode method, This = 809baf0
21:05:47.994 0x806cb00      mt.18          - Instance method receiver: myapp/MyTime@55D8CBA8 arguments: ()
21:05:48.079 0x806cb00      mt.6          < myapp/MyTime.test()V Bytecode method
21:05:48.079 0x806cb00      mt.9         < myapp/MyTime.main([Ljava/lang/String;)V Bytecode static method
```

Only real time (79ms) is in the call to MyTime.test()

Could now drill down into MyTime.test():
extend scope of methods traced, and reduce scope of tracing into MyTime.test()

-Xtrace:print=mt,methods={myapp/*},resumecount=1,trigger=method{myapp/MyTime.test,resume,suspend}

```
21:07:14.968*0x806cb00      mt.0         > myapp/MyTime.test()V Bytecode method, This = 809baf0
21:07:14.970 0x806cb00      mt.18         - Instance method receiver: myapp/MyTime@55D8CBA8 arguments: ()
21:07:15.067 0x806cb00      mt.3          > myapp/MyTimer.getTime()V Bytecode static method
21:07:15.067 0x806cb00      mt.19          - Static method arguments: ()
21:07:15.067 0x806cb00      mt.9          < myapp/MyTimer.getTime()V Bytecode static method
21:07:15.069 0x806cb00      mt.6         < myapp/MyTime.test()V Bytecode method
```

Java Dump Engine

- 5/6 available dump types:

console	Basic thread dump to stderr
system	Capture raw process image
tool	Run command line program
java	Write application summary
heap	Capture raw heap image
snap	Take a snap of the trace buffers

- Large range of dump triggers
  - 14 available triggers
  - Dump events extended by the use of filters
- User defined dump labels
  - Ability to include: time, date, pid, uid, jre info
- Ability to configure number of dumps generated
- Ability to execute tool on dump event

Trigger Events

gpf	an unexpected crash, such as a SIGSEGV or a SIGILL
user	SIGQUIT signal (Ctrl-Brk on Windows, Ctrl-\ on Linux, Ctrl-V on z/OS)
abort	a controlled crash, as triggered by the abort() system call
vmstart	the VM has finished initialization
vmstop	the VM is about to shutdown
load	a new class has been loaded
unload	a classloader has been unloaded
throw	a Java exception has been thrown
catch	a Java exception has been caught
uncaught	a Java exception was not handled by the application
thrstart	a new thread has started
thrstop	an old thread has stopped
blocked	a thread is blocked entering a monitor
fullgc	garbage collection has started

Dump filters / ranges

- class events (load, throw, catch, uncaught)
  - exact name `filter=java/lang/OutOfMemoryError`
  - prefix `filter=java/lang/Out*`
  - substring `filter=*OutOfMemory*`
- vmstop event
  - exit code(s) `filter=#129..192#-42#255`
- ranges can be used to remove “noise” or save disk space
  - bounded `range=1..4`
  - open-ended `range=8..0`

Dump Labels

- Can use any combination of tokens and text
- Tokens expanded at time of event
  - Usual date and time: %Y, %y, %m, %d, %H, %M, %S
  - High precision time: %tick (msec), %seq (dump counter)
  - Process info: %pid, %uid (plus %job on z/OS)
  - JRE info: %home, %last (last snapped dump label)
- VM will try to create intermediate directories, for example:

  /mnt/archive/dumps/%Y%m%d/%pid/javacore.%tick.txt

Dump Tools

- Spawns external command
  - -Xdump:tool:exec=<command string>
- Command string can contain tokens
  - “%home\bin\jextract core.%Y%m%d.%H%M%S.%pid.dmp”
- Default tool attaches platform debugger to VM
  - Windows: windbg
  - Linux: gdb
  - AIX, z/OS: dbx

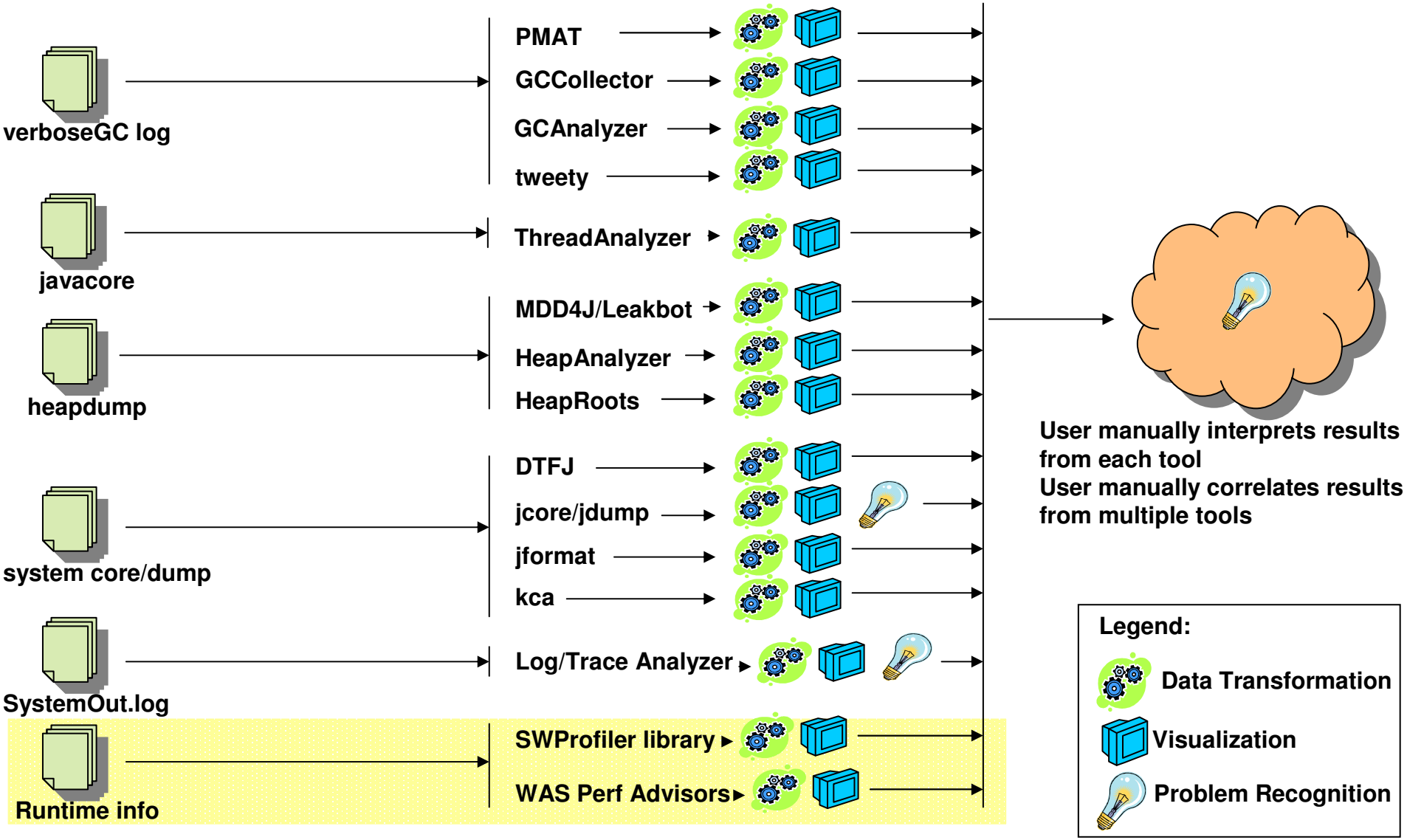
Strategic Direction - DTFJ

DTFJ (Diagnostic Tooling Framework for Java) is a new technology within the IBM JDK to analyze and diagnose problems in Java applications

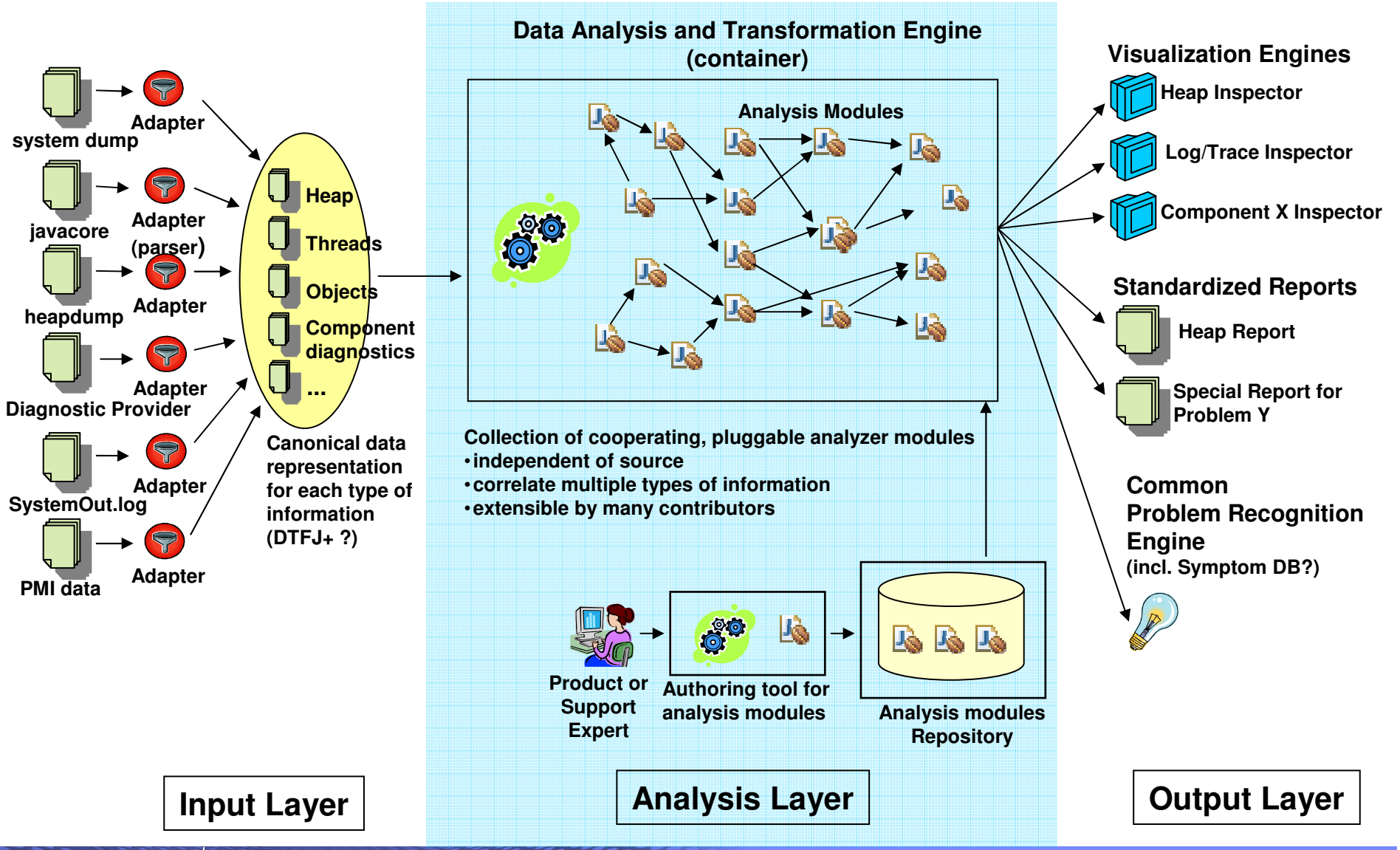
- Read RAS artifacts from a JVM (e.g. a core file) and extract all kinds of useful information from that dump
- Not just one tool: an extensible framework for building many different tools

- Components of the DTFJ family
  - **jextract**: a tool to capture information from a JVM system dump (e.g. core file) and package it into a platform-independent format
  - **DTFJ library proper or core library**: a library that parses the contents of the system dump file packaged by jextract, and provides access to its contents in a standardized manner, through a standard API
  - **DTFJ-based tools**: a collection of end-user tools that call the DTFJ library through the DTFJ API, to present and analyze information in various ways useful to the users

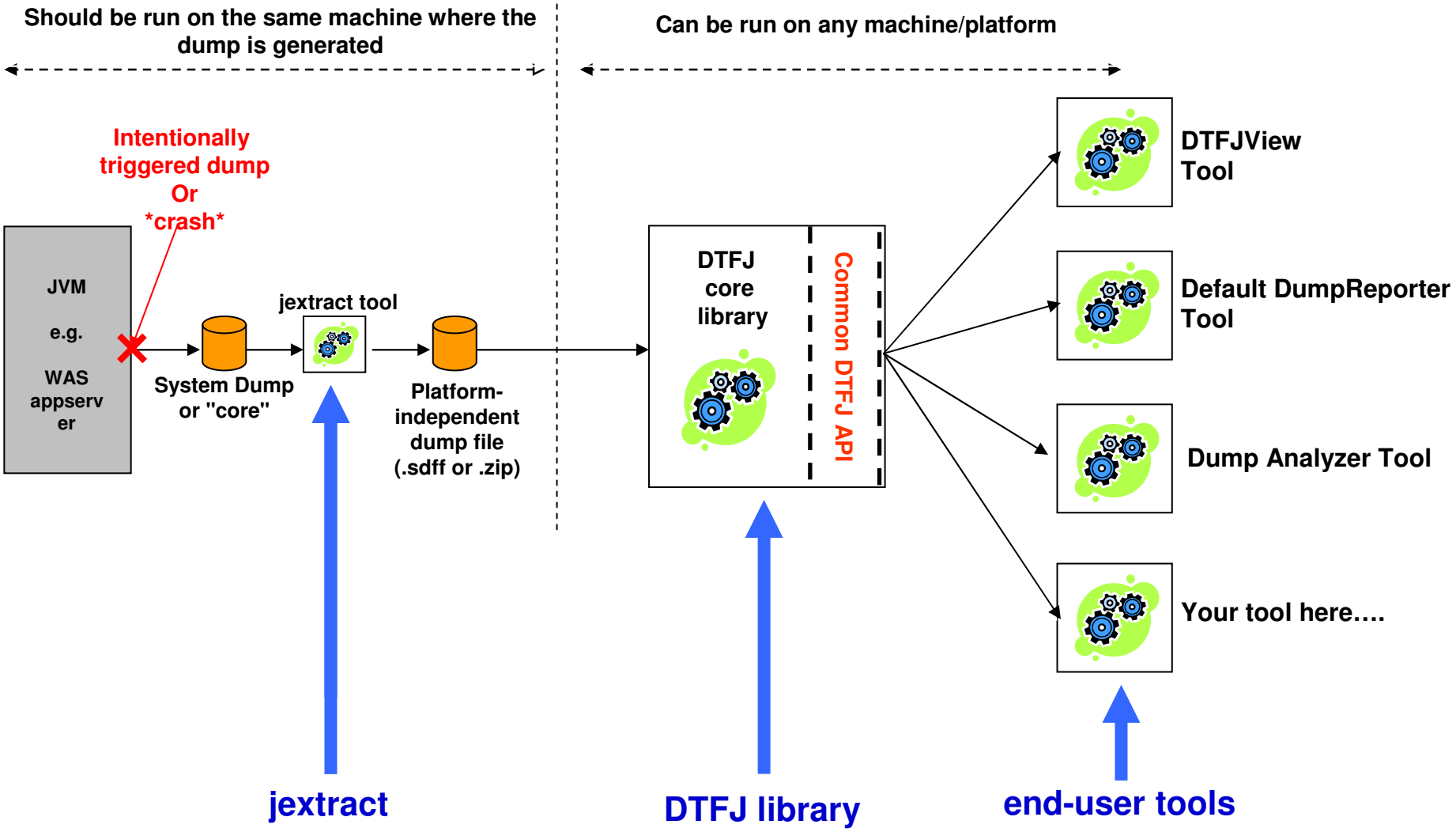
Roadmap: Where we are coming from



Roadmap: Composable Tools Architecture – where we are going



Using the DTFJ components - example

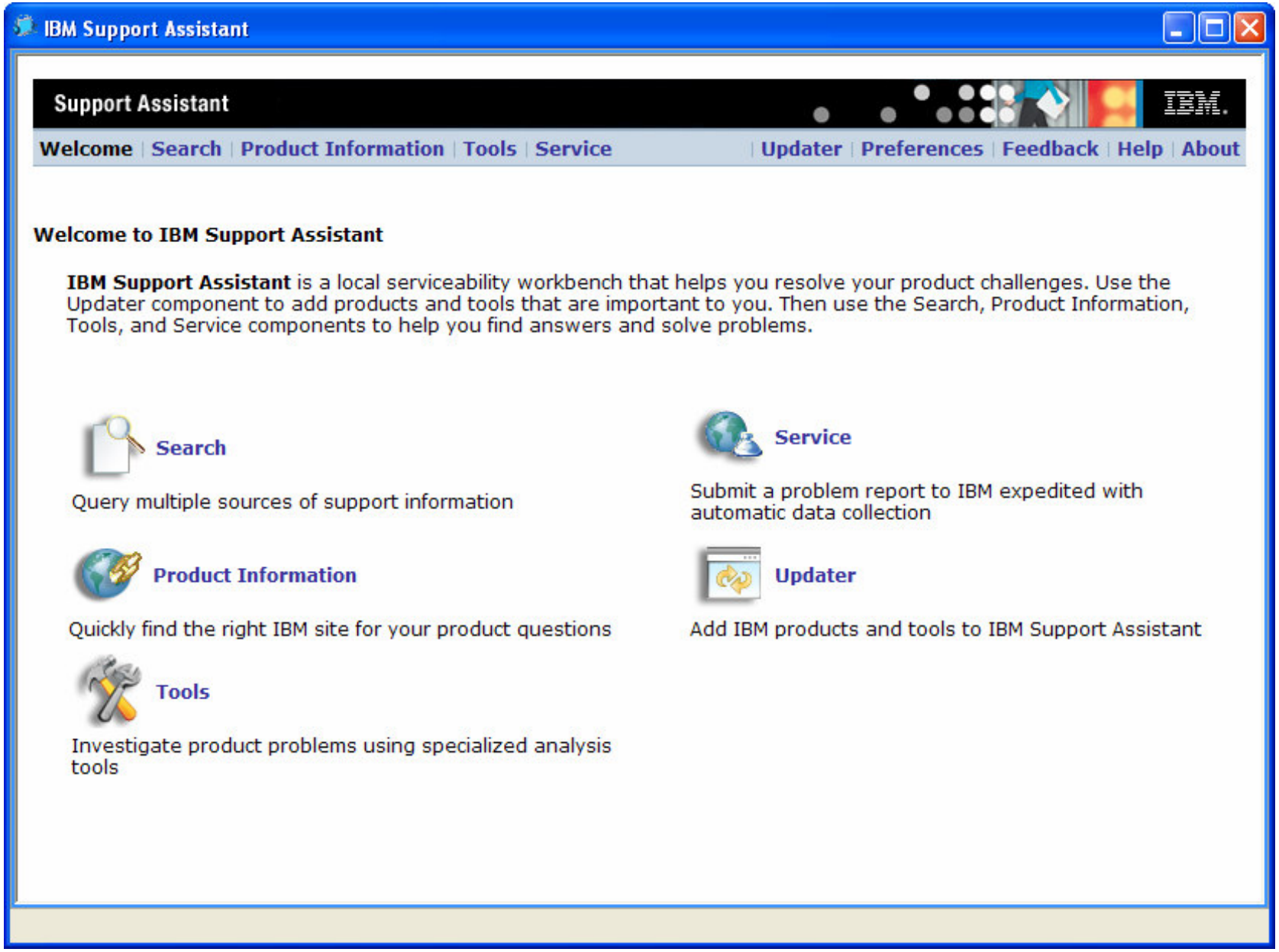


Common Problem Recognition

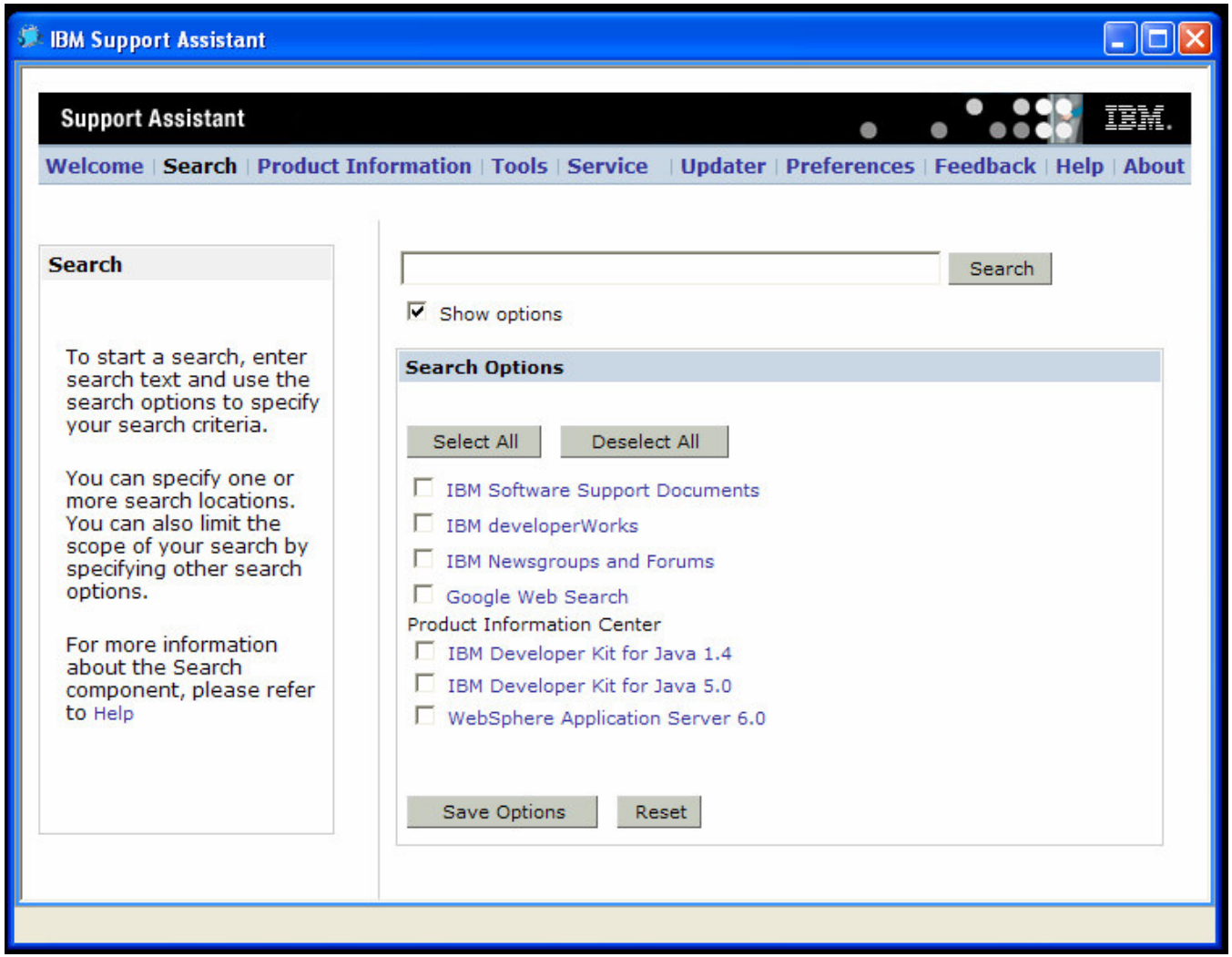
- Goal: reduce service turnaround time by providing a comprehensive suite of self-help tools and documentation
  - Causal analysis of PMRs in both WAS and JDK to detect common idiom (across all platforms)
  - Documentation revisions and enhancements
    - IBM Education Assistant
    - IBM Guided Activity Assistant

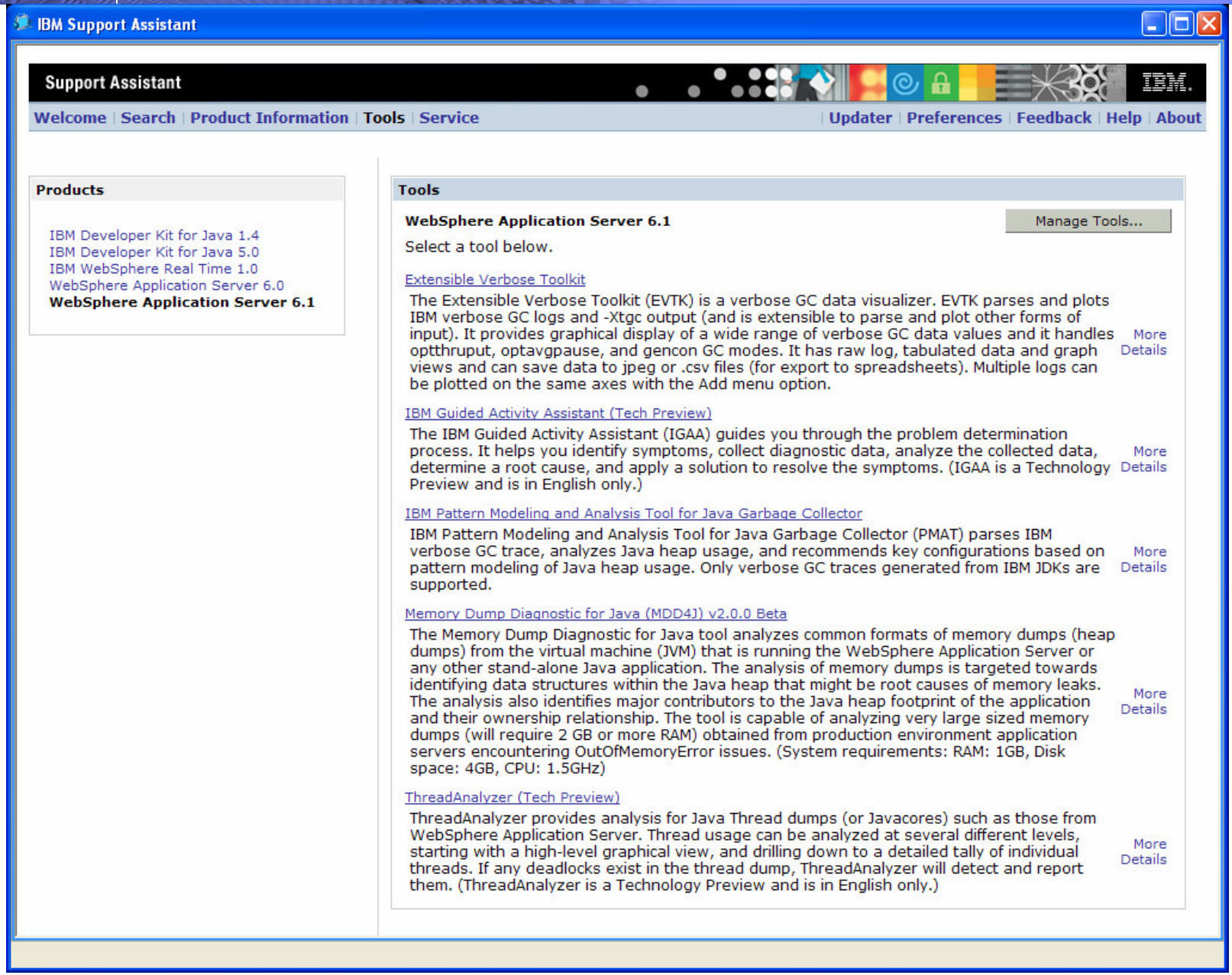
IBM Support Assistant

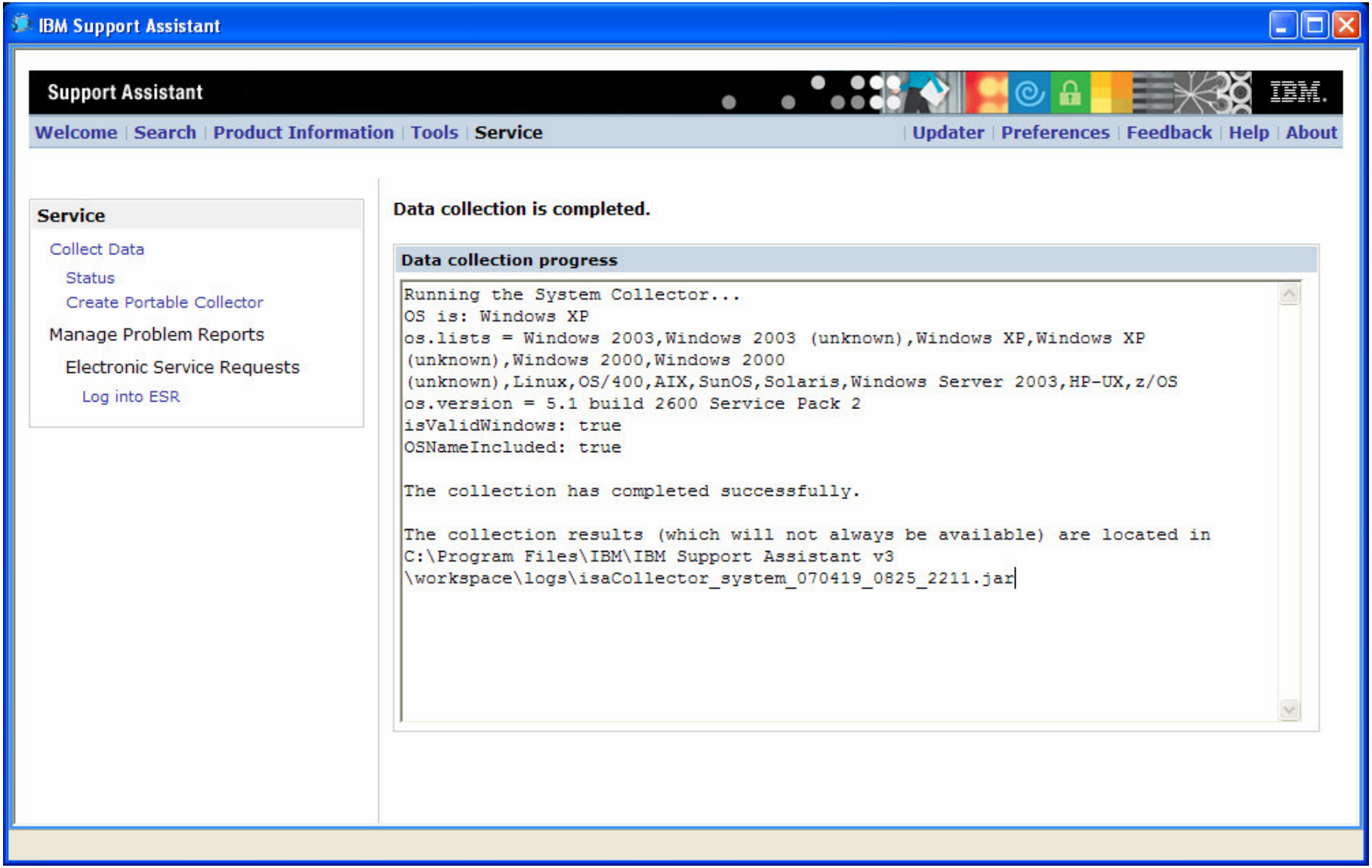
- Hosting for Serviceability Tools across product families
- Automatic PD data gathering
- Assist with opening PMRs and working with IBM Support
- Documentation
  - Aggregated Search across sources
  - Regular updates to Diagnostics Guide



ISA - Search







Common Support Concerns – what we heard

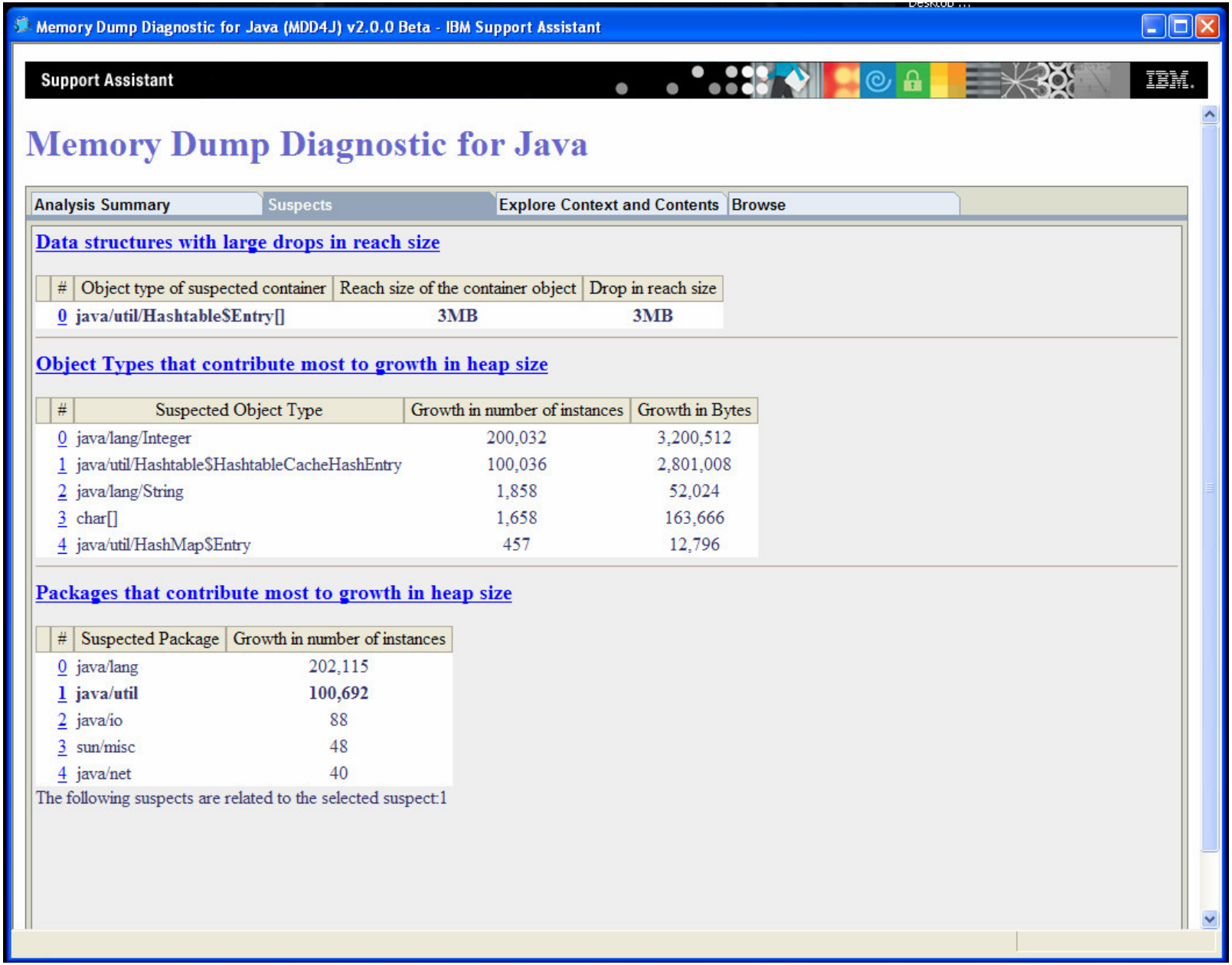
- OutOfMemoryError / heap Size Tuning
  - It's hard to tune the right GC parameters, and figure out where memory leaks come from.
- Deadlocks / hangs / spins
  - Need ability to introspect on a running JVM to determine what's happening at the moment – in a report based way.
- General analysis tools
  - Need ability to examine JVM data – classloaders, threads, monitors, etc.. to do general PD tasks.

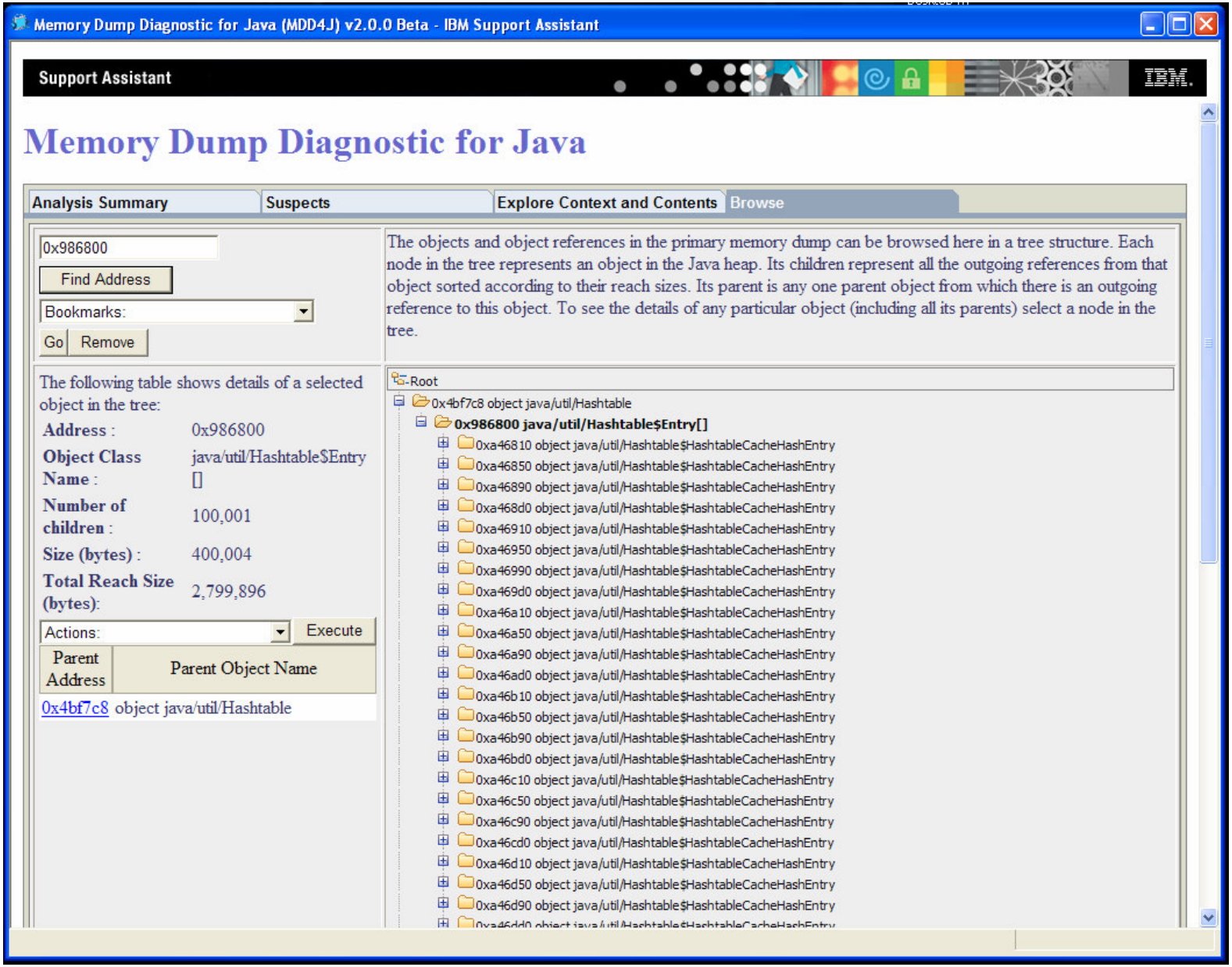
Tools

- IBM Support Assistant “umbrella”
  - MDD4J (Memory Dump Diagnostics For Java)
  - EVTK (Extensible Verbose ToolKit)
  - DumpAnalyser
  - ThreadAnalyser
  - Java Lock Monitor (JLM)
  - IBM Guided Activity Assistant (IGAA)
- RAD (Rational Application Developer) & family

MDD4J

- Java Memory Analysis tool
  - Help explain / track down OutOfMemoryError
  - Performance problems when object use
- 2 modes of use
  - Single snapshot – to visualize a given heap
  - Delta mode – to track growth between 2 points in time
- Input data types supported
  - IBM Portable Heap Dump (heapdump.phd)
  - IBM Text heap dump (heapdump.txt)
  - HPROF heap dump format (hprof.txt)
  - 2007 workplan – DTFJ adaptors (thus direct svcdump consumption)
- Aims to replace HeapRoots, FindRoots, etc..
  - Compatible imputs with those tools
  - Full replacement only when true full functional superset





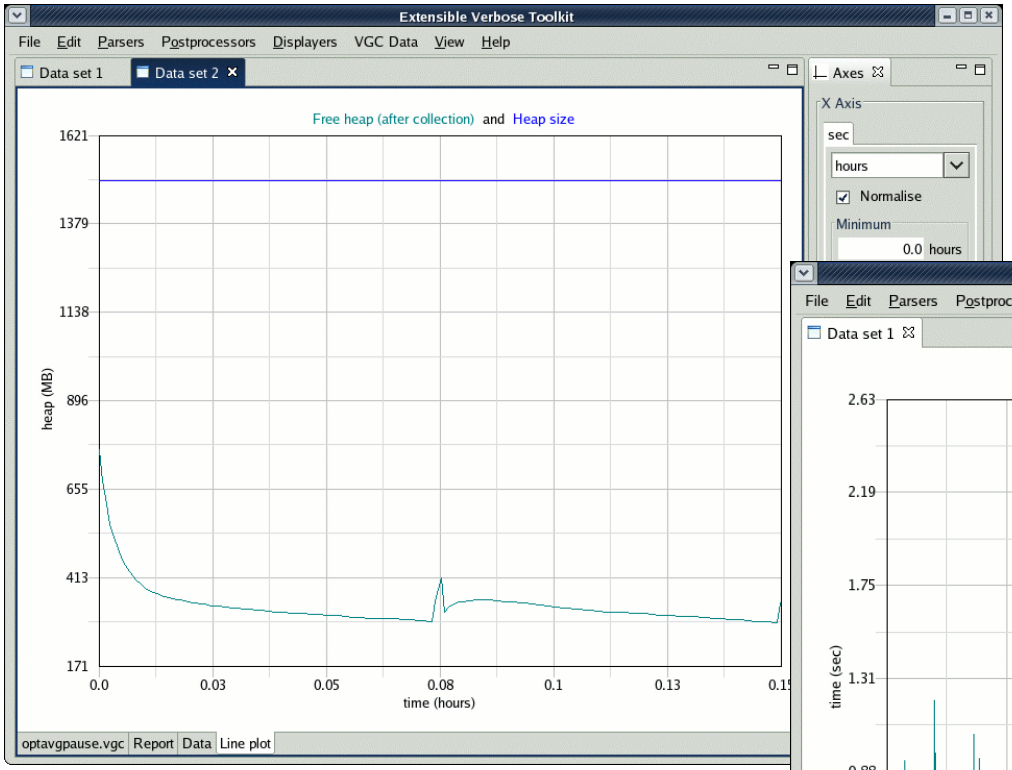
Extensible Verbose Toolkit (eVTK)

- EVTK is a verbose GC analysis tool
- Handles verbose GC from all versions of IBM JVMs
  - 1.4.2 and lower
  - 1.5.0 and higher
  - WebSphere real time
  - Intel, PowerPC, Z-Series,
- … and Solaris platforms

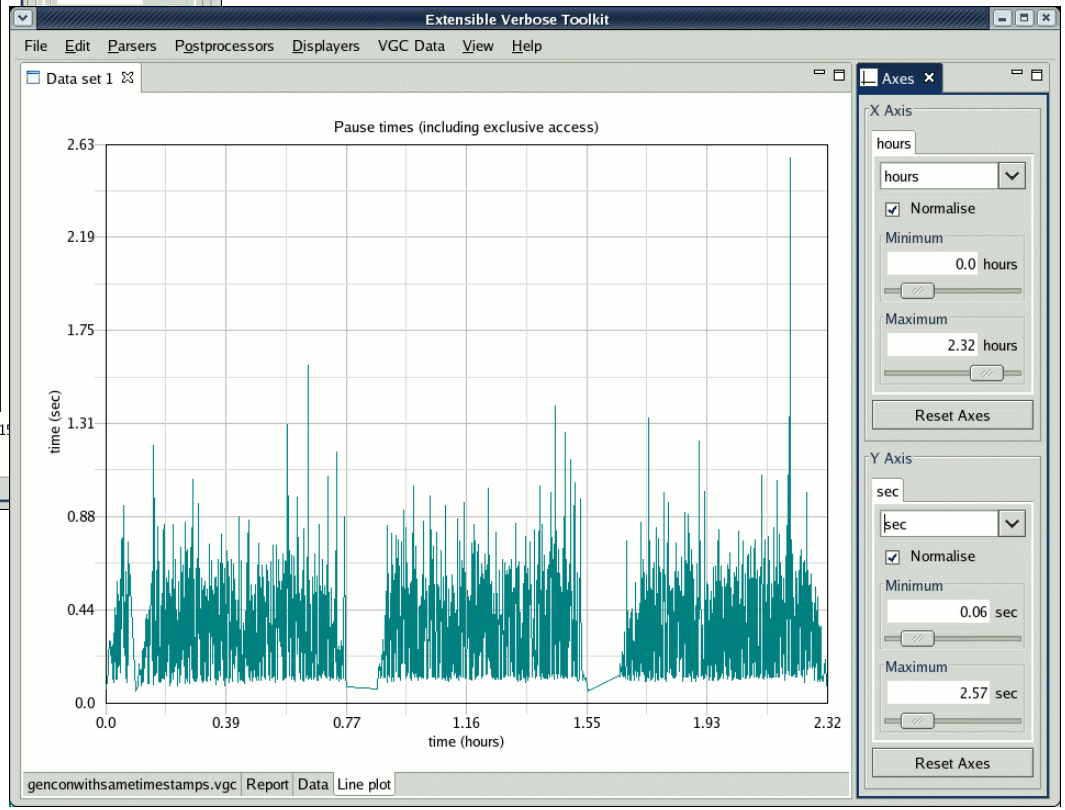
EVTK capabilities

- Analyses heap usage, heap size, pause times, and many other properties
- Compare multiple logs in the same plots and reports
- Many views on data
  - Reports
  - Graphs
  - Tables
- Can save data to
  - HTML reports
  - JPEG pictures
  - CSV files

EVTK – Heap Visualization

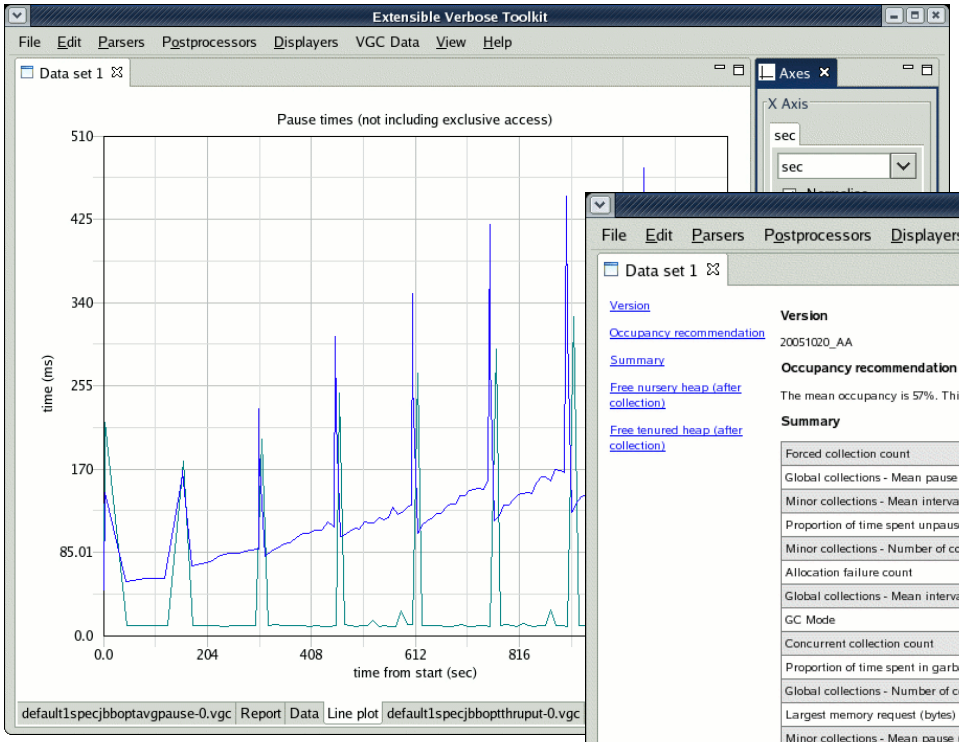


Heap occupancy

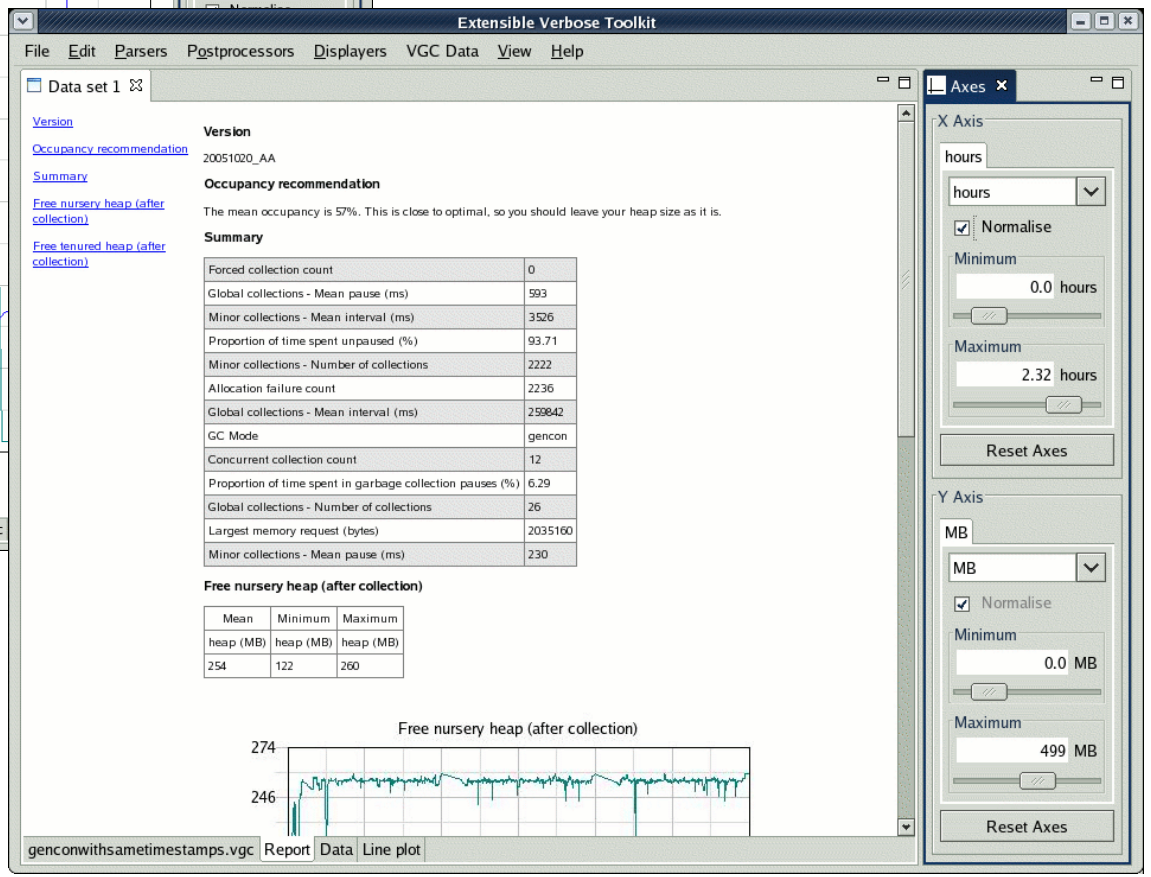


Pause times

EVTK - Comparison & Advice



Compare runs…



Performance advisor…

DumpAnalyzer **initial** release

- Aim to diagnose
  - Deadlock in Java code
    - Report thread names / locations etc.
  - Out of memory condition
    - Report populations and large collections etc.
    - Summarise the native memory usage
    - Recommend further analysis using eg MDD4J
    - Analysis generally requires multiple heap dumps
  - Internal error (gpf etc.)
    - Is failure in non-IBM native code ?
      - Probably user coding error, report location etc.
      - If on J9 recommend running with -Xcheck:jni
    - Otherwise open PMR (use ISA to assist with data gathering)

DumpAnalyzer deadlock example

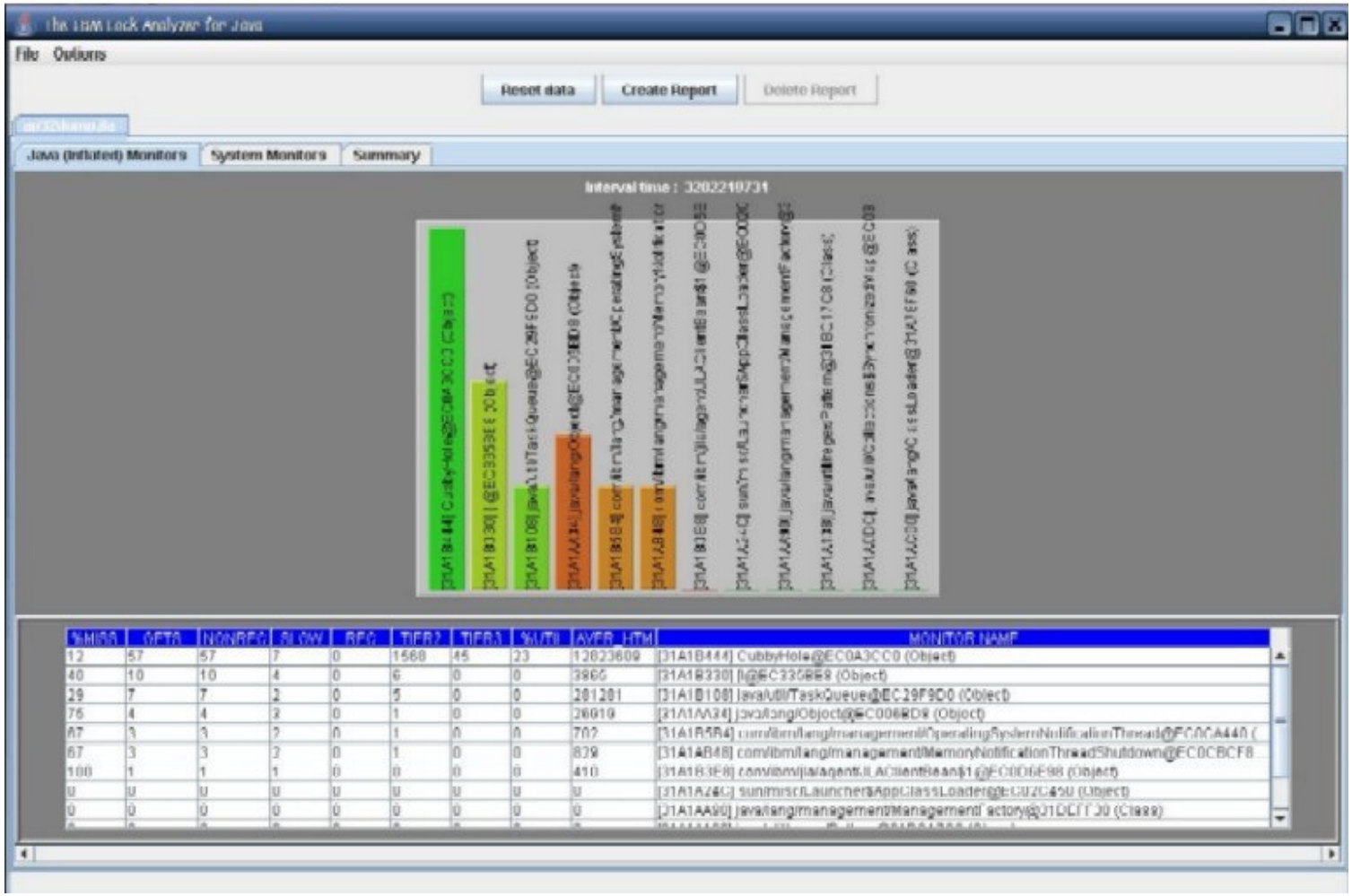
```
Start analysis of C:\a_tsm\workspaces\DumpAnalyzer\bin\core.20070315.101244.5436.dmp.zip
Execution time : 3925ms

Deadlocks detected - now trying to determine cause

Run report on analyzer: com.ibm.dtfj.analyzer.DTFJDeadlock rule: default

Error : Deadlock cycle detected on monitor com.ibm.dtfj.java.j9.JavaMonitor@7b3c48
Error : Deadlock cycle detected on monitor com.ibm.dtfj.java.j9.JavaMonitor@7b3c88
Error : Deadlock cycle detected on monitor com.ibm.dtfj.java.j9.JavaMonitor@7b35a8

Deadlock cycle detected
0 monitor <com.ibm.dtfj.java.j9.JavaMonitor@7b3c48> is owned by thread <Thread 1> which is waiting
     on monitor ...
1 monitor <com.ibm.dtfj.java.j9.JavaMonitor@7b3c88> is owned by thread <Thread 2> which is waiting
     on monitor ...
2 monitor <com.ibm.dtfj.java.j9.JavaMonitor@7b35a8> is owned by thread <Thread 3> which is waiting
     on monitor ...
3 monitor <com.ibm.dtfj.java.j9.JavaMonitor@7b3c48> is owned by thread <Thread 1> which is waiting
     on monitor ...
.
.

Thread Thread 1 : owned monitors and top 2 frames on stack
  owns com.ibm.dtfj.java.j9.JavaMonitor@7b3c48
  frame 1 com/ibm/monitor/test/TestDeadlock$AThread::grabLocks ()V lev 0
  frame 2 com/ibm/monitor/test/TestDeadlock$AThread::run ()V lev 0
Thread Thread 3 : owned monitors and top 2 frames on stack
  owns com.ibm.dtfj.java.j9.JavaMonitor@7b35a8
  frame 1 com/ibm/monitor/test/TestDeadlock$AThread::grabLocks ()V lev 0
  frame 2 com/ibm/monitor/test/TestDeadlock$AThread::run ()V lev 1
Thread Thread 2 : owned monitors and top 2 frames on stack
  owns com.ibm.dtfj.java.j9.JavaMonitor@7b3c88
  frame 1 com/ibm/monitor/test/TestDeadlock$AThread::grabLocks ()V lev 0
  frame 2 com/ibm/monitor/test/TestDeadlock$AThread::run ()V lev 1
```

```
============================ Deployed applications ============================

   DefaultApplication                   STARTED           DefaultApplication.ear/deployments/DefaultApplication
   ivtApp                               STARTED           ivtApp.ear/deployments/ivtApp
   query                                STARTED           query.ear/deployments/query
   SamplesGallery                       STARTED           SamplesGallery.ear/deployments/SamplesGallery
   PlantsByWebSphere                    STARTED           PlantsByWebSphere.ear/deployments/PlantsByWebSphere
   IBMUTC                               STARTED           IBMUTC.ear/deployments/IBMUTC
   ver/systemApp                        STARTED           e:/WAS61/AppServer/systemApps/ManagementEJB.ear
   ver/systemAp                         STARTED           e:/WAS61/AppServer/systemApps/filetransfer.ear
   ver/systemApps/Sch                   STARTED           e:/WAS61/AppServer/systemApps/SchedulerCalendars.ear
   isclite                              STARTED           isclite.ear/deployments/isclite
================================ EJB Container ================================

   Number of EJB modules defined: 6
   Number of EJB homes defined: 20
   EJB Container: com/ibm/ws/runtime/component/EJBContainerImpl@00F1EA88
   Number of EJBs currently in the EJB cache: 0
   Number of EJB wrappers currently in the wrapper cache: 0
   Local JNDI naming context name: ejb
================================ Web Container ================================
   Number of web modules defined: 14
   Number of servlets defined: 130
   Number of servlet requests currently active: 2
   Web Container: com/ibm/ws/webcontainer/component/WebContainerImpl@0104A9C8
   Number of requests: 0
================================ Thread Pools =================================
   Thread Pool:
      Name: PluginConfigService
      Min: 5                Max: 20               Grow-as-needed: false
      Current size: 0
      Current active threads: 0
   Thread Pool:
…..
```

Java Lock Analyser (JLA)

- JLA provides profiling data on monitors used in Java applications and the JVM:
  - Counters associated with contended locks
  - Total number of successful acquires
  - Recursive acquires
  - Frequency with which a thread had to block waiting on the monitor
  - Cumulative time the monitor was held.
  - For platforms that support 3 Tier Spin Locking the following are also collected
    - Number of times the requesting thread went through the inner (spin loop) while attempting to acquire the monitor.
    - Number of times the requesting thread went through the outer (thread yield loop) while attempting to acquire the monitor.

Java Lock Analyser (JLA)



JLA Data Provided

Column name	Description
%MISS	Pecentage of the total Gets (acquires) where the requesting thread was blocked waiting on this monitor
GETS	Total number of successful acquires
NONREC	Total number of non recursive aquires. This number includes SLOW gets
SLOW	Total number of non recursive acquires which caused the requesting thread to block waiting for the monitor to become unowned. This number is included in NONREC
REC	Total number of recursive aquires. A recursive acquire is one where the requesting thread already owns the monitor
TIER2	Total number of Tier 2 (inner spin loop) iterations on platforms that support 3-Tier spin locking
TIER3	Total number of Tier 3(outer thread yield loop) iterations on platforms that support 3-Tier spin locking
%UTIL	Monitor hold time divided by Interval Time. Hold time accounting must be switched on
AVER-HTM	Average amount of time the monitor was held. Recursive acquires are not included because the monitor is already owned when acquired recursively
MONITOR NAME	Monitor name or NULL (blank) if name is not known

Discussion / Questions