# 日志

Understanding Solr log

【https://support.lucidworks.com/hc/en-us/articles/201298187-Understanding-Solr-log】

[http://lucene.apache.org/core/4\_8\_1/core/org/apache/lucene/codecs/lucene46/package-summary.html#package\_description]

# Lucene index file formats

The fundamental concepts in Lucene are index, document, field and term.

A document is a sequence of fields.

A field is a named sequence of terms.

A term is a sequence of bytes.（字节序）

### Inverted Indexing

反向索引，在自然关系中，根据document可以list出terms，反向索引与其相反，根据term可以列出documents。

### Types of Fields

In Lucene, fields may be stored, in which case their text is stored in the index literally, in a non-inverted manner. Fields that are inverted are called indexed. A field may be both stored and indexed.

The text of a field may be tokenized into terms to be indexed, or the text of a field may be used literally as a term to be indexed

(field中的text可以被tokeniized )

### Segments

每个segment包含多个documents

Lucene indexes may be composed of multiple sub-indexes, or segments. Each segment is a fully independent index

Segment为独立单元

Index evolve by （）

Creating new segments for newly added documents

Merging existing segments

（在segment合并时丢弃deleted document）

Deleted documents are dropped when segments are merged

### Filed 数据

保存所有stored filed 的数据信息

域(Field)的数据信息(.fdt， ，， ，.fdx)

能快速根据文档号提取相应的文档数据（正向数据）

### Term vector数据

### Term vector filed 文件（tvf文件 词向量域文件）

词典索引文件（内存）

词典索引条目指向词典文件相应词的位置

词典文件（全部词）

词典中条目 指向 词频文件 和 词位置文件

# Solr性能

【Solr Performance Data】

【<http://wiki.apache.org/solr/SolrPerformanceData>】

http://wiki.apache.org/solr/SolrPerformanceProblems

<http://blog.thetaphi.de/2012/07/use-lucenes-mmapdirectory-on-64bit.html>

[http://www.kafka0102.com/2010/10/366.html]

[http://wiki.apache.org/solr/SolrPerformanceFactors]

非常好的文章

【http://h3x.no/2011/05/10/guide-solr-performance-tuning】

### mergeFactor

多少个segment文件 进行合并 default 10

当文档数量到达 maxBufferedDocs 时将会在磁盘创建一个segment

solrcache

【<http://wiki.apache.org/solr/SolrCaching】>

# Solr synonym

http://nolanlawson.com/2012/10/31/better-synonym-handling-in-solr/

# Jboss 添加访问日志

<http://chandank.com/application-server/tomcat/jboss-7-access-log-configuration>

## jboss debug 启动

--debug 4343

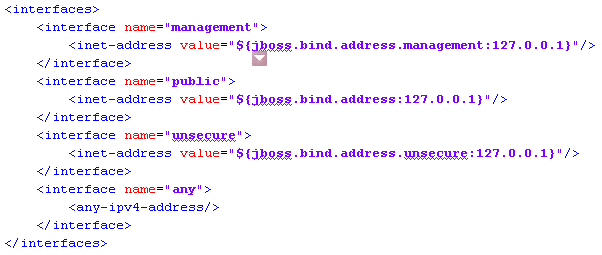
# Jboss jconsole 配置

检查configuration/standalone.xml 文件中配置

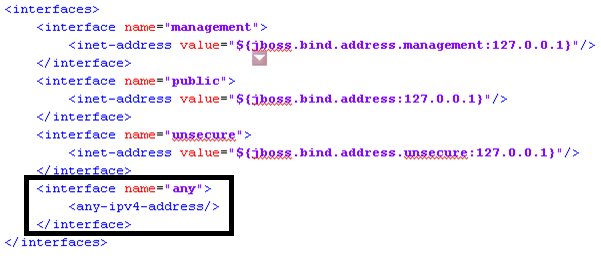




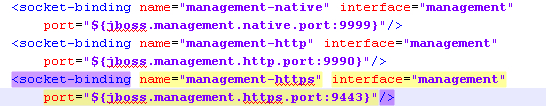
特别注意 inteface 所对应的inet-address ，如上jboss 将只会bind在127.0.0.1这个IP 上。



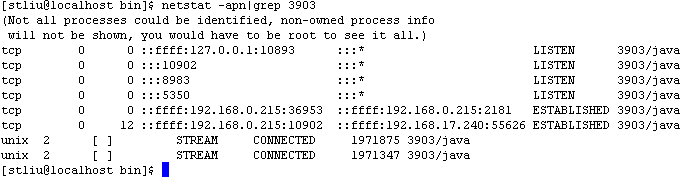
增加any interface，bind在任意ipv4 地址上



2检查linux 防火墙是否打开， 如果打开需要过滤如下地址

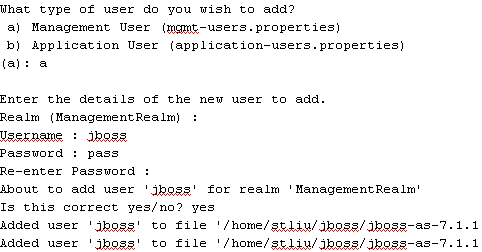


3.检查linux 系统中 jboss 地址bind情况



确认bind 正确

4. 添加用户，运行jboss\_home/bin 下 add-user.sh



5. 启动 jconsole

位于jboss\_home/bin 目录下



注意后边的主机及端口号以jboss中配置为准，点击链接即可。

# Jboss jvisualvm

[https://github.com/johnaoahra80/jboss-as-tool-integration/blob/master/visualvm/visualvm.bat]

# jboss thread cpu

<http://www.mastertheboss.com/jboss-monitoring/how-to-monitor-jboss-cpu-usage>

# increase max client connection

https://community.jboss.org/thread/199845

# jboss performance tuning

<http://www.mastertheboss.com/jboss-performance/jboss-performance-tuning-part-1>

<http://www.mastertheboss.com/jboss-performance/jboss-as-7-performance-tuning>

# Jboss installing and tuning

<https://support.eapps.com/index.php?/Knowledgebase/Article/View/418/52/user-guide---jboss-7-installation-and-tuning#jboss-7-administration-console>

<http://www.mastertheboss.com/jboss-performance/jboss-as-7-performance-tuning/page-2>

# JBoss access log

<https://community.jboss.org/thread/201861>

# monitor jboss graphically

http://www.mastertheboss.com/jboss-monitoring/how-to-monitor-jboss-graphically

## 设置max-connection

[connector name="http" protocol="HTTP/1.1" scheme="http" socket-binding="http" **max-connections="1000"**/]

## Linux 内存查看

http://blog.csdn.net/tianlesoftware/article/details/6459044

# Linux Tune Network Stack (Buffers Size) To Increase Networking Performance

<http://www.cyberciti.biz/faq/linux-tcp-tuning/>

# TCP initial window

<https://developers.google.com/speed/protocols/tcpm-IW10>

<http://atifans.net/articles/adjust-initial-window-size-under-linux/\>

sysctl -a | grep tcp\_slow\_start\_after\_idle

sysctl -w net.ipv4.tcp\_slow\_start\_after\_idle=0

### HTTP Management Endpoint

### CLI Management Endpoint

## Performance tuning a CentOS LAMP web server for high traffic volumes

http://blog.williamjamieson.me/2010/08/17/performance-tuning-a-centos-lamp-web-server-for-high-traffic-volumes/

# Performance Tuning a CentOS Web Server

http://beginlinux.com/blog/2008/11/performance-tuning-a-centos-web-server/

## IT-Efficiency:Things to check during your coffee break

http://henkvandervalk.com/it-efficiencythings-to-check-during-your-coffee-break

查看接受字节数/s、发送字节数/s 停留在一条线上表示有bottleneck

# [How do I migrate the connector thread settings in JBoss 5.1.0](https://community.jboss.org/message/747232#747232)

Connector 设置

https://community.jboss.org/thread/202409

# JVM 参数详解

<http://www.cnblogs.com/redcreen/archive/2011/05/04/2037057.html>

# JVM GC日志详解

<https://blogs.oracle.com/poonam/entry/understanding_cms_gc_logs>

# [linux 内核参数调整说明](http://blog.csdn.net/radkitty/article/details/3009522)

<http://blog.csdn.net/radkitty/article/details/3009522>

<https://publib.boulder.ibm.com/infocenter/clresctr/vxrx/index.jsp?topic=%2Fcom.ibm.cluster.pe.v1r3.pe200.doc%2Fam101_tysfbpjp.htm>

<http://jaseywang.me/2013/12/08/10g82599eb-%E6%B5%8B%E8%AF%95%E4%BC%98%E5%8C%96kernel/>

# unix 磁盘效能查看

<http://unix.stackexchange.com/questions/55212/how-can-i-monitor-disk-io>

## 非常好的性能文章

<http://www.thegeekstuff.com/2011/03/linux-performance-monitoring-intro/>

## Run Queue

Run queue indicates the total number of active processes in the current queue for CPU.

When CPU is ready to execute a process, it picks it up from the run queue based on the priority of the process.

Please note that processes that are in sleep state, or i/o wait state are not in the run queue.

So, a higher number of processes in the run queue can cause performance issues.

## Cpu Utilization

This indicates how much of the CPU is currently getting used.

This is fairly straight forward, and you can view the CPU utilization from the [top command](http://www.thegeekstuff.com/2010/01/15-practical-unix-linux-top-command-examples/).

100% CPU utilization means the system is fully loaded.

So, a higher %age of CPU utilization will cause performance issues.

Load Average

## 理解CPU LOAD

http://blog.scoutapp.com/articles/2009/07/31/understanding-load-averages

0.7 < load < 1: 此时是不错的状态，如果进来更多的汽车，你的马路仍然可以应付。

load = 1: 你的马路即将拥堵，而且没有更多的资源额外的任务，赶紧看看发生了什么吧。

load > 5: 非常严重拥堵，我们的马路非常繁忙，每辆车都无法很快的运行

This indicates the average CPU load over a specific time period.

On Linux, load average is displayed for the last 1 minute, 5 minutes, and 15 minutes. This is helpful to see whether the overall load on the system is going up or down.

For example, a load average of “0.75 1.70 2.10″ indicates that the load on the system is coming down. 0.75 is the load average in the last 1 minute. 1.70 is the load average in the last 5 minutes. 2.10 is the load average in the last 15 minutes.

Please note that this load average is calculated by combining both the total number of process in the queue, and the total number of processes in the uninterruptable task status.

Remember the 80/20 rule — 80% of the performance improvement comes from tuning the application, and the rest 20% comes from tuning the infrastructure components.

### 1. CPU

#### Context Switch

#### Run Queue

#### Cpu Utilization

#### Load Average

### 2. Network

### 3. I/O

<http://www.slashroot.in/linux-system-io-monitoring>

There are several system monitoring tools in linux, that can be used to monitor, I/O, usage. they are mentioned below.

* sar
* vmstat
* top
* iostat

every thing is file

* Block devices(Hard-disks,Compact Disk's,Floppy,Flash Memory)
* Character devices or serial devices (Mouse, keyboard)
* Network Devices

Vmstat

**Wa** wait on io

**"bi"** column in the vmstat ouput shows the blocks that are read in to RAM from the disk

### 4. Memory

## Tools

top, free, ps, iostat, vmstat, mpstat, sar, tcpump, netstat, iozone,

vmstat 2 5

# Performance Monitoring : Vmstat – to find performance bottleneck

http://unixadminschool.com/blog/2011/05/performance-monitoring-vmstat-to-find-performance-bottleneck/

诊断CPU

: CPU, disk I/O’s, memory, and network I/O’s. The vmstat can help to find problems within the first three categories

r column

如果r 的值 > 2倍 CPUs ， CPU 存在瓶颈

There are several schools of thought on the maximum number is appropriate here, but most people agree that more than 2 to 5 times the number of CPUs on the system shows a bottleneck (This estimate is different when using multi-core CPUs). The way to resolve this is to run fewer applications or to add CPU modules to the system.

诊断DISK IO

b column

被Block，等待 IO 的线程数量， 如果一直固定值

**Disk / IO Performance**

The **vmstat** utility cannot tell us which disks have a bottleneck, but it can tell us if there is an IO problem overall. The important column in the output is the **b** column. It is the second column in the output, and stands for “Blocked.” This refers to the number of threads that were blocked, or waiting, for IO in the past interval. Over time, this column should contain 0′s the majority of the time. If there is constantly a number in that column you can check **mpstat** to see if **wt** (percent wait time) column has a high number too. That in conjunction with the **vmstat** **b** column indicate a system blocked for disk I/O and  it is advisable to get an extended **iostat -xpn** or **iostat -xtc** output and examine it thoroughly to try to detect the bottlenecks.

## 诊断 MEMORY

Virtual Meomory Statistics

**sr** column

If the value in this column is greater than zero then the page scanner is scanning memory pages to put them back on the free list to be reused.

r b swpd free buff cache si so bi bo in cs us sy id wa st

0 0 18864 475100 265804 960860 0 0 0 3 4 5 0 0 100 0 0

Procs（进程）：

r: 运行队列中进程数量

b: 等待IO的进程数量

Memory（内存）：

swpd: 使用虚拟内存大小

free: 可用内存大小

buff: 用作缓冲的内存大小

cache: 用作缓存的内存大小

Swap：

si: 每秒从交换区写到内存的大小

so: 每秒写入交换区的内存大小

IO：（现在的Linux版本块的大小为1024bytes）

bi: 每秒读取的块数

bo: 每秒写入的块数

系统：

in: 每秒中断数，包括时钟中断。

cs: 每秒上下文切换数。

CPU（以百分比表示）：

us: 用户进程执行时间(user time)

sy: 系统进程执行时间(system time)

id: 空闲时间(包括IO等待时间)

wa: 等待IO时间

* **VSS** - Virtual Set Size 虚拟耗用内存（包含共享库占用的内存）
* **RSS** - Resident Set Size 实际使用物理内存（包含共享库占用的内存）
* **PSS** - Proportional Set Size 实际使用的物理内存（比例分配共享库占用的内存）
* **USS** - Unique Set Size 进程独自占用的物理内存（不包含共享库占用的内存）

**ps -e -o pid,rss,args**

* -si:交换内存使用，由磁盘调入内存  
  -so:交换内存使用，由内存调入磁盘  
  内存够用的时候，这2个值都是0，如果这2个值长期大于0时，系统性能会受到影响，磁盘IO和CPU资源都会被消耗。

## 显示网络情况

## dmstat -n

<http://www.cyberciti.biz/tips/top-linux-monitoring-tools.html>

<http://www.51testing.com/html/48/202848-242043.html>

## netstat show network traffic

<http://www.petri.co.il/netstat-command-monitor-network-traffic.htm>

# eclipse plugin architecture

<http://www.eclipse.org/articles/Article-Plug-in-architecture/plugin_architecture.html>

# JS 创建CLASS 几种方式

<http://www.cnblogs.com/tiwlin/archive/2009/08/06/1540161.html>

Options

|  |  |  |
| --- | --- | --- |
| AreaChart | $jit.AreaChart |  |
| Canvas |  |
| Margin |  |
| Label |  |
|  |  |  |
|  |  |  |

# 数据结构

树

|  |  |  |
| --- | --- | --- |
| 名称 | 特点 | 时间 |
| Tree |  |  |
| **binary search tree** (**BST**) | * The left [subtree](http://en.wikipedia.org/wiki/Tree_(data_structure)#Subtree) of a node contains only nodes with keys less than the node's key. * The right subtree of a node contains only nodes with keys greater than the node's key. * The left and right subtree each must also be a binary search tree. * Each node can have up to two successor nodes. * There must be no duplicate nodes.(无重复节点)   左子树小于 < 父节点 < 右子树 |  |
| **AVL tree** | 任意节点左右子树高度相差最大为1 | 基于BST对insertion和deletion做平衡 |
| R-B tree | 红黑树 1）每个结点要么是红的，要么是黑的。  2）根结点是黑的。  3）每个叶结点，即空结点（NIL）是黑的。  4）如果一个结点是红的，那么它的俩个儿子都是黑的。  5）对每个结点，从该结点到其子孙结点的所有路径上包含相同数目的黑结点。 |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

# JVM 命令

jstat -gccapacity pid

http://www.oracle-base.com/articles/misc/monitoring-java-garbage-collection-using-jstat.php

# GIT 操作

## Git操作模式

<http://nvie.com/posts/a-successful-git-branching-model/>

## diffference between tag and branch

【http://alblue.bandlem.com/2011/04/git-tip-of-week-tags.html】

# CODE review

http://www.devx.com/enterprise/Article/31658