



Troubleshooting Apache® Ignite™



Stan Lukyanov Customer Solutions, GridGain

GridGain and Apache Ignite







Apache Ignite Support – Faster Time to Reliable Ignite

- Get up and running faster with 2 hours initial consultation
- Ensure fast, reliable Ignite with unlimited 9x5 global support
 - Unlimited web/e-mail support
 - Identify bugs, workarounds
 - Troubleshoot performance, reliability issues



https://www.gridgain.com/products/services/support/support-apache-ignite



Agenda



- Monitoring
 - Logging
 - JMX
 - GridGain Web Console
- Troubleshooting
 - Network
 - Storage
 - Performance
- Checklist



Monitoring









Setup it before something bad happens!





Agenda



- Monitoring
 - Logging
 - JMX
 - GridGain Web Console
- Troubleshooting
 - Network
 - Storage
 - Performance
- Checklist







What can you do with logs

- Manually check nodes state
- Identify issues with cluster configuration
- Add automatic parsing to report issues on the fly
 - With custom or third-party tools
- Provide to GridGain Support experts







Configuring GridGain logs

- https://apacheignite.readme.io/docs/logging
- Use log4j 2.x integration
 - Other options: j.u.logging, log4j 1.x, slf4j, custom integration
- Start Ignite in verbose mode
 - ignite.sh -v
 - java -DIGNITE_QUIET=false -cp ...



Logging



Quiet log

\$ ignite.sh
[17:51:02]
[17:51:02] / _/ _/ // // _/ _/
[17:51:02] _/ // (7 7 // / / / _/
[17:51:02] //_/_/ /_/ //
[17:51:02]
[17:51:02] ver. 2.7.2#20190206-sha1:5f8f5488
[17:51:02] 2019 Copyright(C) Apache Software Foundation
[17:51:02]
[17:51:02] Ignite documentation: http://ignite.apache.org
[17:51:02]
[17:51:02] Quiet mode.
[1/:51:02] A Logging to the c:\binaries\ignite\gridgain-ultimate-8.7.2\work\log\ignite-8abf2e4e.0.log
[1/:51:02] ∧ Lögging by 'JavaLogger [quiet=true, contig=null]'
[1/:31:02] ^ TO See ^^FULL^^ CONSOLE TOG HERE ADD -DIGNITE_QUIET=TAISE OF -V TO IGNITE.{Sh[bat}
[1/:11:02] VS: Windows 10 10.0 and Hot The Environment 1 & 0 151-b12 Oracle Connection lave HetCret(TM) 64-Bit Server VM 25 151-b
$[17,31,22]$ VM information. Sava(M) SE Kuntime Environment 1.8.0_ $131-012$ of attending available function of the vertice of 23.131-0
re [17:51:02] Please set system property '_Diava pet preferIPv4Stack=true' to avoid possible problems in mixed environments
[17:51:02] Configured Durins:
[17:51:02] A GridGain 8.7.2#20190205-sha1:75399045
[17:51:02] ^ 2019 Copyright (C) GridGain Systems
17:51:02]
[17:51:02] Configured failure handler: [hnd=StopNodeOrHaltFailureHandler [tryStop=false, timeout=0, super=AbstractFailureHandler [igno
redFailureTypes=[SYSTEM_WORKER_BLOCKED, SYSTEM_CRITICAL_OPERATION_TIMEOUT]]]]
[17:51:03] Message queue limit is set to 0 which may lead to potential OOMEs when running cache operations in FULL_ASYNC or PRIMARY_SY
NC modes due to message queues growth on sender and receiver sides.
<pre>[17:51:03] Security status [authentication=off, t]s/ss]=off]</pre>
[17:51:03] Rolling updates are disabled. GridGain version update will require full cluster restart. Consider changing 'GridGainConfigu
ration.rollingUpdatesEnabled' configuration property.
[17:51:14] Performance suggestions for grid (fix if possible)
[17:51:14] To disable, set -DIGNITE_PERFORMANCE_SUGGESTIONS_DISABLED=true
[17:51:14] A Enable GI Garbage Collector (add '-XX:+UseGIGC' to JVM options)
[17:51:14] ^ Set max direct memory size if getting 'OOME: Direct buffer memory' (add '-XX:MaxDirectMemorySize= <size>[g G m M k K]'</size>
to JVM options)
[17:51:14] A Disable processing of calls to System.gc() (add -XX:+DisableExplicitGC to JVM options)
[1/:51:14] Refer to this page for more performance suggestions: https://apacheignite.readme.io/docs/jvm-and-system-tuning
[17:31:14] [17:51:14] To start Concolo Management & Monitoning num ignitouisonend [ch[hat]
[17:51:14] Data Consider Management & Monitoring Fun Tgnitevisorchu, {snjbat}
[17:51:14] Data Regions Configured:
[17:51:47]
[17:51:14] Ionite node started OK (id=Sahf2ede)
[17:51:4] Tomology spanshot [ver=1]locNode=8ahf2e4e_servers=1_clients=0_state=4CTIVE_CPUs=8_offhean=3_2GR_hean=1_0GR]







Configuring GC logs

- <u>https://apacheignite.readme.io/docs/jvm-and-system-tuning</u>
- Crucial to troubleshoot a lot of issues
- To configure
 - -XX:+PrintGCDetails -XX:+PrintGCTimeStamps
 - -XX:+PrintGCDateStamps -XX:+UseGCLogFileRotation
 - -XX:NumberOfGCLogFiles=10 -XX:GCLogFileSize=100M
 - -Xloggc:/path/to/gc/logs/log.txt







How to manage logs

- Choose location for the log files
 - The default location is \${IGNITE_HOME}/work/log/ignite.log
 - Use a local disk with enough space (2 GB+)
 - Don't use /tmp!
 - Good idea to store GridGain, application and GC logs together
- Archive old log files periodically to save on storage space
- Try to save logs for the cluster's current uptime



Logging



Run-time configuration changes

- Helpful when you need to debug a running deployment
- Easy to do with log4j 2.x
 - Edit the log4j config file directly
 - <u>https://logging.apache.org/log4j/2.x/manual/c</u> onfiguration.html#AutomaticReconfiguration
 - Use JMX log4j has its own bean
 - <u>https://logging.apache.org/log4j/2.x/manual/j</u> <u>mx.html</u>

Java Monitoring & Management Console - p Connection Window Help	id: 8620 org.apache.ignite.e	xamples — 🗆 X				
Overview Memory Threads Classes VM Summary MBeans						
😥 🖳 JMImplementation	Attribute values					
com.sun.management	Name	Value				
iava.lang	Additive	true				
iava.util.logging	AppenderRefs	java.lang.String[2]				
	Filter	null				
org.apache.logging.log4j2	IncludeLocation	true				
	Level	INFO				
Attributes	Name					
Operations						
Appenders						
ContextSelector						
Attributes						
The second secon						
⊕ containazonawa						
⊕ (igree/pserjetty) ⊕ (igree/pserjetty.util.component						
		Refresh				



Agenda



- Monitoring
 - Logging
 - JMX
 - GridGain Web Console
- Troubleshooting
 - Network
 - Storage
 - Performance







What can you do with JMX

- Check state of various grid subsystems: storage, thread pools, etc
- Monitor metrics changing over time
- Add automatic warnings on important metrics
 - With custom or third-party tools







How to leverage JMX beans

- Standard way to monitor Java applications
 - A lot of tools on the market
 - JConsole and VisualVM are bundled with JDK

🕌 Java Monitoring & Management Console - p	id: 19000 org.apache.ignite.sta	rtup.cmdline.CommandLineStartup config/default-config — 🛛 🛛 🗙
🙆 Connection Window Help		_ 5 [,] ×
Overview Memory Threads Classes VM Summar	y MBeans	
	Attribute values	
ia com.sun.management ia java.lang	Name	Value
	CheckpointSpiFormatted	[NoopCheckpointSpi []]
iava.util.logging	CollisionSpiFormatted	NoopCollisionSpi [
org.apache	CommunicationSpiFormatted	TcpCommunicationSpi [connectGate=org.apache.ignite.spi.communication.tcp.TcpComm
- 764c12b6	Copyright	2018 Copyright(C) Apache Software Foundation
	DeploymentSpiFormatted	LocalDeploymentSpi [Isnr=org.apache.ignite.internal.managers.deployment.GridDeploy
. Clients	DiscoverySpiFormatted	TcpDiscoverySpi [addrRslvr=null, sockTimeout=5000, ackTimeout=5000, marsh=JdkMar
DataRegionMetrics	EventStorageSpiFormatted	org.apache.ignite.spi.eventstorage.NoopEventStorageSpi@1622f1b
- Kernal	ExecutorServiceFormatted	8
. ClusterLocalNodeMetricsMXBeanIr	FailoverSpiFormatted	[AlwaysFailoverSpi [maxFailoverAttempts=5, totalFailoverJobs=0]]
ClusterMetricsMXBeanImpl	FullVersion	2.4.12-20181011
□ · · · · · · · · · · · · · · · · · · ·	GridLoggerFormatted	GridLoggerProxy [igniteInstanceName=null, id8=37117504]
	IgniteHome	c:/binaries/ignite/gridgain-ultimate-fabric-8.4.12
Operations	InstanceName	
	JdkInformation	Java(TM) SE Runtime Environment 1.8.0_151-b12 Oracle Corporation Java HotSpot(TM)
. SPIs	LifecycleBeansFormatted	
	LoadBalancingSpiFormatted	[RoundRobinLoadBalancingSpi [balancer=RoundRobinGlobalLoadBalancer [ctx=org.apac
	LocalNodeId	37117504-c75b-4c50-8ad5-df8530ee422e
	MBeanServerFormatted	com.sun.jmx.mbeanserver.JmxMBeanServer@6f94fa3e
	OsInformation	Windows 10 10.0 amd64
	OsUser	gridgain
	PeerClassLoadingEnabled	false
	RebalanceEnabled	true
	StartTimestamp	1550431454134
	StartTimestampFormatted	17.02.2019 22:24:14
	UpTime	1486512
	UpTimeFormatted	00:24:46.512
	UserAttributesFormatted	0
	VmName	19000@einstein
< >		Refresh



GridGain Metrics



Name	Description	JMX Query	Default
Cluster metrics	Basic node information	<pre>org.apache:clsLdr=*,group=Kernal,name=ClusterMetricsMXBeanImpl org.apache:clsLdr=*,group=Kernal,name=ClusterLocalNodeMetricsMXBeanImpl</pre>	Enabled
Data region metrics	Memory information	<pre>org.apache:clsLdr=*,group=DataRegionMetrics,name=<region_name></region_name></pre>	Disabled
Data storage metrics	Persistent storage information	<pre>org.apache:clsLdr=*,group="Persistent Store",name=DataStorageMetrics</pre>	Disabled
Cache metrics	Cache statistics	<pre>org.apache:clsLdr=*,group=<cache_name>,name="org.apache.ignite.internal.proc essors.cache.CacheClusterMetricsMXBeanImpl" org.apache:clsLdr=*,group=<cache_name>,name="org.apache.ignite.internal.proc essors.cache.CacheLocalMetricsMXBeanImpl"</cache_name></cache_name></pre>	Disabled
Thread pool metrics	Thread pools information	<pre>org.apache:clsLdr=*,group="Thread Pools",name=<thread_pool_name></thread_pool_name></pre>	Enabled

- <u>https://apacheignite.readme.io/docs/cluster-groups#section-cluster-group-metrics</u>
- <u>https://apacheignite.readme.io/docs/memory-metrics</u>
- <u>https://apacheignite.readme.io/docs/cache-metrics</u>



Agenda



- Monitoring
 - Logging
 - JMX
 - GridGain Web Console
- Troubleshooting
 - Network
 - Storage
 - Performance



GridGain Tools

Web Console

- Basic version in Ignite
- Enhanced version in GridGain
- All-in-one solution
- Web-based
- The richest functionality

Visor GUI

- Available in GridGain
- All-in-one solution
- Runs on a user's PC

Command line tools

- Available in Ignite
- Multiple scripts
 - Visor CMD
 - control.sh
 - SQLLine





What can be done

- Manage multiple GridGain clusters through a web interface
 - Monitor metrics changing over time
 - Watch logs and thread dumps of the nodes
 - Execute and monitor running queries
- Generate cluster configurations and RDBMS integrations
- Available online to try at <u>https://console.gridgain.com/</u>
- On-premise installation for production using Docker or bare metal
 - <u>https://apacheignite-tools.readme.io/docs/docker-deployment</u>











≡ Gi	idGain 2		My Connected Clusters: 1	Start Demo	٠	Stanislav Lukyanov 👻
411 11	Configuration for node: 434A1217					
SQL	A Monitoring / Node Info					
	Filter					<u> </u>
\$	Property	Value				
	General 🕐					
ক্ষ	Node ID ③	434A1217-400/	A-4A0E-A03C-BC1717E6F846			
	Active on start ⑦	~				
	Client mode ⑦	×				
C ⁰	Enable update notifier (?)	~				
	Quiet mode (?)	\checkmark				
	Restart enabled ⑦	×				
. V.	Is daemon 💿	×				
	No discovery order (?)	×				
	Disable shutdown book	×				





≡ Gi	id Gain 201	My Connected Clusters: 1	Start Demo	٠	Stanislav Lukyanov 👻
*1: 1:1 1:1	Log for node: 43C06DC7 (Feb 20, 00:57:22)			<u>↓</u> Downle	pad 🗘 Refresh
SQL	C:\binaries\ignite\gridgain-ultimate-8.7.2\work\log\ignite-43c06dc7.0.log				٩
	[00:55:59,020][INFO][main][IgniteKernal]				1
63	<pre>>>> / // (7 // // // // // // // // // // // // //</pre>				
	<pre>[00:55:59,029][INFO][main][IgniteKernal] Config URL: file:/C:/binaries/ignite/gridgain-ultimate-8.7.2/config/defa [00:55:59,053][INFO][main][IgniteKernal] IgniteConfiguration [igniteInstanceName=null, pubPoolSize=8, svcPoolSize= [00:55:59,053][INFO][main][IgniteKernal] Deemon mode: off [00:55:59,053][INFO][main][IgniteKernal] OS: Windows 10 10.0 amd64 [00:55:59,053][INFO][main][IgniteKernal] OS user: gridgain [00:55:59,054][INFO][main][IgniteKernal] PID: 15740 [00:55:59,054][INFO][main][IgniteKernal] Language runtime: Java Platform API Specification ver. 1.8 [00:55:59,054][INFO][main][IgniteKernal] VM information: Java[IN] S5 Puntime Environment 1 8.0 151-102 Oracle Cont [00:55:59,054][INFO][main][IgniteKernal] Language runtime: Java Platform API Specification ver. 1.8</pre>	ult-config.xml =8, callbackPoolSize=8, striped poration lava HotSnot(TM) 64-86	PoolSize=8, sysPoolS	iize=8, mgmtPo	olSize=4, igfsPoolSiz
Q	<pre>[00:55:59,656][INFO][main][IgniteKernal] WH total memory: 0.966B [00:55:59,656][INFO][main][IgniteKernal] Remote Management [restart: on, REST: on, JMX (remote: on, port: 4916, s [00:55:59,657][INFO][main][IgniteKernal] Logger: Javalogger [udiet=true, configerul1] [00:55:59,657][INFO][main][IgniteKernal] Logger: Javalogger [udiet=true, configerul1] [00:55:59,657][INFO][main][IgniteKernal] VM arguments: [-XX:HaggressiveOpts, -Xxns], -Xx:MavMetaspaceSize [00:55:59,657][INFO][main][IgniteKernal] UM arguments: [-XX:HaggressiveOpts, -Xxns], -Xx:MavMetaspaceSize [00:55:59,677][INFO][main][IgniteKernal] Configured caches [in 'sysMemPlc' dataRegion: ['ignite-sys-cache']] [00:55:59,072][INANING][main][IgniteKernal] Please set system property '-Djava.net.preferIPv4Stacketrue' to avoid [00:55:59,073][INANING][main][IgniteKernal] 3-rd party license can be found at: C:binaries/ignite/gridgain-ultimat [00:55:59,409][INFO][main][IgniteKernal] -rd party license can be found at: C:binaries/ignite/gridgain-ultimat</pre>	puration sava hotspot(h+) 64-51 auth: off, ssl: off)] =256m, -DIGNITE_QUIET=true, -DI nfiguration.systemRegionInitial possible problems in mixed env 0 10.0 amd64. Our team will app e-8.7.2\libs\licenses	GNITE_SUCCESS_FILE=C Size property to cha ironments, reclate the feedback	::\binaries\ig inge the setti : if you exper	nite∖gridgain-ultimat ng. ience any problems ru
Ŵ	<pre>[00:55:59,469][INFO][main][IgnitePluginProcessor] ^ 2019 Copyright(C) GridGain Systems [00:55:59,469][INFO][main][IgnitePluginProcessor]</pre>				
()					





≡ Gi	ridGain 201	My Connected Clusters: 1	Start Demo	Stanislav Lukyanov 👻
	SQL queries notebook 🖉 🔒 Cluster 4ED54BC3 🕜 Cluster active 💶		Scroll to query - + Add que	ery + Add scan
SQL	✿ Notebooks / Notebook 'SQL queries notebook'			
₩ &	 With query notebook you can Create any number of queries Execute and explain SQL queries Execute scan queries View data in tabular form and as charts More info 	HAR(100)); nn'), (3, 'Emma');		
103				
	Select all			Execute 🝷
Q	1 SELECT * FROM PERSON		No caches	5
	Execute on selected node Explain Rows per page: 100 Max	د pages: <u>Unlimited</u> العن العام ال	 Allow non-collocated join Collocated Query (?) Enforce join order (?) Lazy result set (?) 	s 🕐
.≟ (



Troubleshooting



Agenda



- Monitoring
 - Logging
 - JMX
 - GridGain Web Console
- Troubleshooting
 - Network
 - Storage
 - Performance
- Checklist



Network troubleshooting

- Nodes not joining the cluster
- Nodes joining a wrong cluster
- Nodes taking a long time to connect
- Node is kicked out from the cluster



Nodes not joining the cluster

Symptom

- Started nodes don't join in a single cluster each has a cluster of its own instead ۲
- Topology snapshot is "ver=1, <...> servers=1" on both nodes



Nodes not joining the cluster

Possible cause and solution

- Connection issues, firewall, etc
 - Check via ping, check firewall settings
 - Make sure TCP ports are open: 47500-47509, 47100-47109, 11211, 10800
- IP Finder configuration
 - Use TcpDiscoveryVmIpFinder when you know all hosts in advance
 - Make sure to list all IPs and ports
 - A lot of other options: Kubernetes, cloud, etc
 - <u>https://apacheignite.readme.io/docs/kubernetes-deployment</u>
 - <u>https://apacheignite.readme.io/docs/tcpip-discovery</u>



Nodes joining a wrong cluster

Symptom

- Node started in a seemingly clean environment connects to some unknown cluster
- Topology snapshot is "ver=2, servers=2"





Nodes joining a wrong cluster

Possible cause and solution

- Multicast (TcpDiscoveryMulticastIpFinder) is used
 - Only use it in closed subnets, or for a "Hello, world!"
- Wrong IPs in the configuration
 - Check and fix IP finder configuration see previous slide



Nodes taking a long time to connect

Symptom

- Cluster is taking a long time to start
 - Or the first node is taking a long time to start
- After the nodes joined in the cluster performance is good



Nodes taking a long time to connect

Possible cause and solution

- Too many addresses in the IP finder
 - Ignite will try to connect to each know IP
 - Only list the addresses and ports you use
- Too high IgniteConfiguration.failureDetectionTimeout
 - Basic timeout for most network operations
 - Used for establishing a connection
 - Reduce the failureDetectionTimeout to scan through the IP list faster



Node is kicked out from the cluster

Symptom

• Cluster nodes report that one of them has failed

[...][WARN][disco-event-worker-#42][GridDiscoveryManager] Node FAILED: TcpDiscoveryNode [id=13169858-92db-48a3-bc4bdb369cab6457, addrs=[...], sockAddrs=[...], discPort=47501, order=2, intOrder=2, lastExchangeTime=1550680047403, loc=false, ver=2.7.2#20190206-sha1:5f8f5488, isClient=false]

• *"Local node segmented"* messages on the failed node

[...][WARN][disco-event-worker-#42][GridDiscoveryManager] Local node SEGMENTED: TcpDiscoveryNode [id=13169858-92db-48a3bc4b-db369cab6457, addrs=[...], sockAddrs=[...], discPort=47501, order=2, intOrder=2, lastExchangeTime=1550680109573, loc=true, ver=2.7.2#20190206-sha1:5f8f5488, isClient=false]



Node is kicked out from the cluster

Possible cause and solution

- Likely a GC pause
- To confirm
 - Check for "Possible too long JVM pause" messages in the logs
 - Analyze GC logs around the time of the issue
 - <u>https://gceasy.io/</u> can help

Node is kicked out from the cluster

Possible cause and solution

- To handle GC issues
 - Increase heap: -Xmx16g
 - Try G1 with a latency target: -XX:+UseG1GC -XX:MaxGCPauseMillis=200
 - Reduce heap pressure
 - Use smaller SQL queries
 - More about SQL later
 - Use IgniteCache.withKeepBinary() to avoid deserialization
 - https://ignite.apache.org/releases/latest/javadoc/org/apache/ignite/lgniteCache.html#withKeepBinary--
 - Use on-heap caching with CacheConfiguration.copyOnRead=false
 - https://apacheignite.readme.io/docs/memory-configuration#section-on-heap-caching
 - https://apacheignite.readme.io/docs/jvm-and-system-tuning
- Remedy: change failureDetectionTimeout to allow a longer inactivity period



Agenda



- Monitoring
 - Logging
 - JMX
 - GridGain Web Console
- Troubleshooting
 - Network
 - Storage
 - Performance
- Checklist



Storage troubleshooting

- "Out of memory" errors
- Lost data after node restart
- Incorrect/partial data returned by SQL
- Some nodes don't store data



Symptom

Out of memory

- Three kinds of "out of memory" conditions
 - Java's OutOfMemoryError exception
 - IgniteOutOfMemoryException exception
 - OS killing the Ignite's JVM



Out of memory



Possible cause and solution: OutOfMemoryError

- Heap is too small
 - Increase -Xmx
- Large SQL queries are running
 - Use lazy queries
 - <u>https://apacheignite-sql.readme.io/docs/performance-and-debugging#section-result-set-lazy-loading</u>
 - Avoid many large queries running concurrently
 - Split query to reduce data set size
 - Before:
 - SELECT * FROM PERSON ORDER BY AGE
 - After:
 - SELECT * FROM PERSON WHERE AGE < 40 ORDER BY AGE</p>
 - SELECT * FROM PERSON WHERE AGE >= 40 ORDER BY AGE



Out of memory



Possible cause and solution: IgniteOutOfMemoryException

- Data region is too small
 - IgniteOutOfMemoryException: Out of memory in data region [name=default, initSize=256,0 MiB, maxSize=476,8 MiB, persistenceEnabled=false] Try the following:
 - ^-- Increase maximum off-heap memory size (DataRegionConfiguration.maxSize)
 - ^-- Enable Ignite persistence (DataRegionConfiguration.persistenceEnabled)
 - ^-- Enable eviction or expiration policies
 - Increase DataRegionConfiguration.maxSize
 - <u>https://apacheignite.readme.io/docs/capacity-planning</u>
 - Use Native Persistence
 - <u>https://apacheignite.readme.io/docs/distributed-persistent-store</u>
 - Use DataRegionConfiguration.pageEvictionMode=RANDOM_2_LRU
 - <u>https://apacheignite.readme.io/docs/evictions</u>
 - Use an expiry policy
 - <u>https://apacheignite.readme.io/docs/expiry-policies</u>



Out of memory



Possible cause and solution: OOM killer

- Total size of the processes is greater than RAM
 - Check for message
 - Out of memory: Kill process <PID> (java) score <SCORE> or sacrifice child
 - Make sure that total size of the processes is within bounds
 - Disable overcommit to have a more graceful, in-process errors
 - sysctl -w vm.overcommit_memory=2
 - https://www.kernel.org/doc/Documentation/vm/overcommit-accounting
 - May affect other applications, use carefully



Lost data after node restart

Symptom

• Some or all data was lost after one or several node restarts





Lost data after node restart

Possible cause and solution

- In-memory cache with zero backups
 - Taking down a node will always lose data
- In-memory cache with one or more backups
 - Don't take down more nodes simultaneously than there are backups
 - Wait for rebalance after bringing a node back online (use Web Console for that)

Rolling Update	🔓 Cluster C71E746E 🤇 ?) Cluster active 🛑					⊘ 30 sec 🔻
My Nodes 🏾 🏟							
Node ID8	Node IP	Mode	Baseline Status	Version 💌	Uptime	Busy, %	Rebalance Status
3230D63C	10.0.75.1	SERVER	Node in baseline	2.7.2	4m 58s	1.47%	1%
DCC3851D	10.0.75.1	SERVER	Node in baseline	2.7.2	4m 42s	2.36%	1%



Incorrect/partial data returned by SQL

Symptom

• An SQL query with JOIN returns less data than expected

CREATE TABLE ORGANIZATION (ID INT PRIMARY KEY, NAME VARCHAR); CREATE TABLE PERSON (ID INT, ORGID INT, NAME VARCHAR, PRIMARY KEY (ID, ORGID));

SELECT * FROM PERSON P JOIN ORGANIZATION O ON **O.ORGID = O.ID;**



Incorrect/partial data returned by SQL

Possible cause and solution

- Data isn't collocated
 - https://apacheignite.readme.io/docs/affinity-collocation
 - Use affinity key to keep related data together CREATE TABLE ORGANIZATION (ID INT PRIMARY KEY, NAME VARCHAR); CREATE TABLE PERSON (ID INT, ORGID INT, NAME VARCHAR, PRIMARY KEY (ID, ORGID)) WITH "AFFINITY_KEY=ORGID";

SELECT * FROM PERSON P JOIN ORGANIZATION O ON O.ORGID = O.ID;

Use replicated tables – they're collocated with all others
 CREATE TABLE ORGANIZATION (ID INT PRIMARY KEY, NAME VARCHAR)
 WITH "TEMPLATE=REPLICATED";

CREATE TABLE PERSON (ID INT, **ORGID INT**, NAME VARCHAR, PRIMARY KEY (ID, ORGID)); SELECT * FROM PERSON P JOIN ORGANIZATION O ON **O.ORGID = O.ID**;

- Use distributedJoins=true to do JOINs without collocation (this is costly!)



Some nodes don't store data

Symptom

- Native persistence is enabled
- Data is only stored on one node/subset of nodes
- New nodes don't store data

Persistent storage	¢	
Node ID8 ▲ Filter by node ID	Data size	Siz
2EE66CD5	329.2MB	76
E37241C3	180KB	64
E99CE057	330.5MB	76

Showing: 3 rows





Some nodes don't store data

Possible cause and solution

- Baseline topology doesn't include some nodes
 - <u>https://apacheignite.readme.io/docs/baseline-topology</u>
 - Update the baseline after every long-term topology change (use Web Console for that)

Baseline Topology Cluster 3EB4DAED Cluster active			🗘 Refresh
Baseline Nodes 🔅 If you need to remove node from Baseline, you should to stop it f	irst. More info		
Consistent ID Filter	Node ID8 Filter	Status	Baseline
1d2cb769-a685-4d9c-9740-338852128e59	2EE66CD5	Online	•
cd2992de-2d50-4a79-8712-83c6a3edccf3	E99CE057	Online	-
f0954a14-7e1c-4d65-802d-553eeb5f1821	E37241C3	Online	
			Cancel ✓ Save Changes



Agenda



- Monitoring
 - Logging
 - JMX
 - GridGain Web Console
- Troubleshooting
 - Network
 - Storage
 - Performance
- Checklist



Performance troubleshooting

- Throughput of persistent cache updates drops to zero
- Low throughput of SQL access



Throughput of cache updates drops to zero

Symptom

- Native persistence is enabled
- Write performance is good but there are periodic intervals of 0 ops/sec





Throughput of cache updates drops to zero

Possible cause and solution

- Updates in RAM happen faster than on disk disk needs to catch up
 - Enable write throttling
 - DataStorageConfiguration.writeThrottlingEnabled=true
 - Sacrifice peak performance for stable latency and throughput
 - Increase checkpoint page buffer size
 - Allow more pending updates in RAM
- <u>https://cwiki.apache.org/confluence/display/IGNITE/Ignite+Persistent+Sto</u>
 <u>re+-+under+the+hood</u>



Low throughput of SQL access

Symptom

- SQL reads using indexed fields are slow
 - Even for simple queries like SELECT * FROM FOO WHERE ID = 123
- Writes to an SQL-enabled cache are slow
 - Even via key-value API



Low throughput of SQL access

Possible cause and solution

- A lot of possible causes performance tuning is hard!
 - <u>https://apacheignite-sql.readme.io/docs/performance-and-debugging</u>
- A common issue indexes not fitting into maximum inline size
 - Check for "Indexed columns of a row cannot be fully inlined into index" warnings

- Do what the warning says

[2019-02-19T23:36:26,848][WARN][main][H2TreeIndex] <CacheQueryExamplePersons> Indexed columns of a row cannot be fully inlined into index what may lead to slowdown due to additional data page reads, increase index inline size if needed (use INLINE_SIZE option for CREATE INDEX command, QuerySqlField.inlineSize for annotated classes, or QueryIndex.inlineSize for explicit QueryEntity configuration) [cacheName=PERSON, tableName=CacheQueryExamplePersons, idxName=PERSON_FIRSTNAME_IDX, idxCols=(FIRSTNAME, _KEY), idxType=SECONDARY, curSize=10, recommendedInlineSize=71]



Agenda



- Monitoring
 - Logging
 - JMX
 - GridGain Web Console
- Troubleshooting
 - Network
 - Storage
 - Performance
- Checklist









Be prepared

- Setup monitoring before something happens, not after!
 - Make sure logs are safely stored
 - GridGain Web Console is an easy and effective way to monitor
- Properly configure the cluster
 - Implement the suggestions from this presentation before you encounter the issues they help with
- Know the common problems and how to react
 - Logs often provide solutions





Avoiding common problems

- Network-related issues are often caused by IP configuration
 - Configure IP finder TcpDiscoveryVmIpFinder is usually a good fit
- GC pauses cause performance and stability issues
 - Use G1 GC, set a fitting heap size
 - Reduce GC pressure by using lazy SQL, withKeepBinary() and onheap cache
- Insufficient memory will result in a crash
 - Plan storage capacity in advance
 - Assess required heap size during testing
 - Account for other processes in the system





Avoiding common problems

- Implement cluster administration processes
 - For in-memory: wait for rebalance when starting or stopping nodes
 - For persistence: update baseline for long-term topology changes
- Write performance may be unstable when update speed is more than the disk speed
 - Enable write throttling
- Consider performance suggestions reported in the log
 - Index inline size suggestion is an example



Questions?



Apache Ignite Resources

- Apache Ignite documentation
 - https://apacheignite.readme.io/docs/
- Apache Ignite community resources
 - <u>user@ignite.apache.org</u> the mailing list
 - <u>https://ignite.apache.org/community/resources.html</u> other resources and instructions
 - <u>http://apache-ignite-users.70518.x6.nabble.com/</u> forum and archive
 - <u>https://stackoverflow.com/questions/tagged/ignite</u> StackOverflow questions
- Contact me!
 - <u>slukyanov@gridgain.com</u>
 - stanlukyanov@gmail.com



GridGain Resources



- White Papers
 - Visit https://www.gridgain.com/resources/papers
- Webinars
 - Visit <u>https://www.gridgain.com/resources/webinars</u>
- Videos
 - Visit <u>https://www.gridgain.com/resources/videos</u>
- Free 30-Day Ultimate, Enterprise or Professional Edition Trial
 - Visit https://www.gridgain.com/resources/download
- GridGain Support for Apache Ignite
 - Visit <u>https://www.gridgain.com/products/services/support/support-apache-ignite</u>

