

IBM® developerWorks WebSphere®

Online chat transcript, February 19, 2009, 12:00 pm-1:00 pm U.S. ET

Topic: *IBM WebSphere MQ high availability and disaster recovery*

Hosted by:

T Rob Wyatt (T Rob)
Christopher Frank (chrisfra)
Peter Potkay (PeterPotkay)
Paul Faulkner (pfaukn)
Ian Vanstone (Ian or Ivans)

tomyoung: Hello everyone, thanks for attending today's Meet the experts chat. I'm Tom Young with developerWorks, and I'm here with T.Rob Wyatt, Christopher Frank, Peter Potkay, Paul Faulkner, and Ian Vanstone to answer your questions about WebSphere MQ high availability and disaster recovery. I'd like give the experts a chance to introduce themselves, and then the floor will be open for questions. We'll be ending the chat at 1 p.m. Eastern. We'll post a link to the transcript after the chat is over, at http://www.ibm.com/developerworks/websphere/library/chats/0902_mqchat/0902_mqchat.html

T Rob: Hello! I'm T.Rob and I specialize in WMQ security, clustering and HA/DR.

chrisfra: Hi, this is Chris Frank, I'm a technical product specialist for MQ and Message Broker.

Ian or Ivans: Hi, this is Ian Vanstone. I work for IBM in the Hursley Lab, UK. I work in the Software Group Federated Integration Test (FIT) team, where I do integration testing with a multi-platform, multi-product environment - and WMQ plays a key role in that environment. As well as being a WMQ user, up until last year I was part of the WMQ development team in Hursley, working mainly on WMQ clustering and intercommunications features.

PeterPotkay: Hello, I'm Peter Potkay and I've done a lot of work recently with MSCS and VCS clusters for H.A., along with getting a DR MQ farm up for a client. I have a good amount of experience with MQ clustering as well, which can play a role in H.A. and DR for MQ.

T Rob: Please feel free to ask any questions you might have regarding High Availability features of WMQ, disaster recovery, failover, etc.

pfaukn: Hello, I am Paul Faulkner. I work for IBM Software Services for WebSphere. I specialize in ESB integration with WebSphere products and have experience implementing solutions with WMQ Security and High Availability.

chrisfra: We are eager to answer your questions - who wants to go first?

mfrost: I'm interested in the possibility of putting MQ instances behind an F5 load balancer. Ideally to provide a redundant environment where we can perform maintenance without incurring app downtime. That is, using an appliance rather than MQ clustering for load balancing. Has anyone done anything with that type of configuration?

T Rob: @mfrost - that can work with client connections, but it does not work for QMr-to-QMgr or where XA transactions are required.

T Rob: These uses require the ability to reconnect to specific instances of WMQ.

T Rob: However, clients that can connect to any instance can use this kind of load balancing.

T Rob: Peter has more info on this.

PeterPotkay: I have done this for MQClients coming into z/Linux images. The Load balancer will flood your MQ error logs with messages, so suppress them with this tip: <http://www-01.ibm.com/support/docview.wss?uid=swg21161118>

mfrost: I'm hoping to have a scenario where an app could connect to one or more (identical) qmgrs without knowing it. Would this potentially cause issues with SSL or anything other than qmgr-qmgr connections?

PeterPotkay: Works well. But not possible for QM to QM connections, only MQ clients.

PeterPotkay: mfrost, MQ Channel Tables will help you.

PeterPotkay: Build a channel table with 5 entries to it, and your QM connects to a QM name of *. The table will get your connection to one of the 5. In MQ 7.0, the channel tables have more capabilities related to load balancing the incoming connections.

T Rob: Looks like some folks are getting past the login problems. Welcome! Please feel free to toss out any questions you might have.

mfrost: Most of our apps are JMS. I didn't think I could use the channel tables for apps that don't directly use the MQI. Or am I incorrect about that?

chrisfra: JMS MQ does support the CCDT.

T Rob: You can specify the channel table in the managed objects.

PeterPotkay: Channel Tables are possible with Java and JMS.

Ian or Ivans: @mfrost. Here's a link to a JMS and MQ Chl table article:

http://www.ibm.com/developerworks/websphere/library/techarticles/0506_barrago/0506_barrago.html

mfrost: OK. I'll have to look into that more. Thanks!

BarbaraH: I was interested in DR. Currently, we don't try to "save" any of the data since everything I've read says that it is difficult. Our storage guys say that with the current SAN replication technology that they could replicate the important files and it would be similar to our current failover scenario where we just swing the file systems (UNIX).

T Rob: @mfrost, you will probably like the channel table option better when you get to v7.0 and above. Many enhancements to load balancing and channel protocol.

PeterPotkay: When it comes to DR, I think you have to consider 2 things: a) How will I get my MQ infrastructure available for the apps in the DR datacenter, and b) How / If I will get any messages in queues over to the DR datacenter.

mfrost: T-Rob, we are using MQ 7. I will check the tables out.

PeterPotkay: The problem with asynchronous data replication between 2 sites is that it opens up the likelihood of missing messages or duplicate messages.

T Rob: @mfrost - Oh, you are the one using WMQ v7.0! 😊

MichaelATMQSystems: What's the best way to test if your HA setup (Win/UNIX) is OK, i.e. what process to kill to test if the failover is working?

PeterPotkay: You want to ask yourself: Is it more important to have my queues, channels and QMs up, empty and ready for all new traffic, or start dealing with async replication and missing messages. Or duplicate messages. A DR scenario is bad enough without dealing with those complexities, which most apps are NOT prepared for.

T Rob: @Michael - What I like to do is kill the QMgr under heavy load. In a "normal" DR test where the primary QMgr is quiesced everything always works as expected. I like to "pull the plug" on the network cable, kill WMQ processes, shut down the server, etc.

T Rob: All under heavy load.

PeterPotkay: Back up QMs, new in MQ 6.0, offer a Log Shipping style of data replication built into the product. But the data only gets shipped over every time a new linear log is cut, so again you are dealing with async replication and missing messages or duplicate messages.

T Rob: This tests things such as replication. I have had some instances where the replication was interrupted and the DR site could not come up. You don't run into this if you allow the primary site to quiesce normally.

cedolny: T Rob, Peter - Does MQ do anything special during quiesce/shutdown to the file systems that would not be available through the replicated file systems if we "pull the plug"? We have been unsuccessful in "pull the plug" type exercises.

T Rob: I meant QMgr there, not site.

PeterPotkay: One thing to do to test H.A. is to issue endmqm. You are "gracefully" yanking the QM out from under the clustering software. It should fail over. On the other extreme, just start killing random MQ PIDs, as your H.A. software should handle that too.

pfaulkn: Michael, one thing to consider is how your HA environment is configured. It is important to test under heavy load as T-Rob indicated to allow for the stabilization configuration to be tested. This prevents issues where the HA environment my flip-flop back and forth while MQ is starting.

T Rob: @cedolny - the issue you are seeing is that block-level replication does not faithfully replicate data structures that occupy more than one block.

BarbaraH: Peter - you are saying exactly what I've been telling people. Right now, we use saveqmgr and amqoamd information to rebuild the QManager at time of disaster - plus I save my crontab and user information. Has worked well in our DR tests. Glad to hear that I wasn't off-base. Thanks for confirming my methodology.

T Rob: So the problem is that when you interrupt the QMgr under load with replication, the files on the remote side may be left in a corrupt state.

PeterPotkay: Barbara, we set up a MQ Farm at the DR site, and all the apps connect to the main prod site via VIPs. When DR strikes, we repoint the VIPS from the PROD QMs to the DR QMs and instantly all new traffic goes thru the DR MQ Farm

T Rob: BarbaraH - you also want to save the qm.ini file, the contents of any exit directories and your SSL keys.

BarbaraH: T.Rob - Good idea. Currently I just "manually" keep that information in sync, but rsync'ing them would be safer.

T Rob: @cedolny - When databases use disk replication, they have a protocol that reconciles the replicated blocks back to actual transactions. They don't apply blocks, but wait until they can apply the transaction. Block replication for ANY application does not have this level of transactionality.

DepremCA: Is there a way to gradually drain work from one queue manager and redirect it to another queue manager?

T Rob: DepremCA - If you have a cluster, you can suspend the QMgr from the cluster.

T Rob: Any in-flight messages continue, but new ones are redirected.

T Rob: That's an MQ cluster, not a VCS, HACMP, or other hardware cluster, by the way.

T Rob: If you don't use a WMQ cluster, you can PUT-disable the queue. The consuming app can still get messages, but no new ones can be placed there. This only works if the PUTting app is local though.

DepremCA: Understood. Would there be an adverse impact if queue managers were at different versions?

PeterPotkay: I see these ways to get your MQ "infrastructure" (not messages) over to a DR site that is too far for sync replication:

PeterPotkay: MQ clustering, backup QMs, manual rebuilding, duplicate infrastructure fronted by VIPS.

T Rob: @DepremCA - Not as long as the functions used were that of the lowest version. For example, you can't publish v7.0 topics to a v6.0 QMgr.

T Rob: And of course, this assumes "lowest version" is 6.0 or higher.

T Rob: If it's 5.3, then there might be other issues.

T Rob: Everyone is off of 5.3 by now, yes?

DepremCA: Yes.

T Rob: [Anyone says "no" will be booted from the chat!!!]

BarbaraH: Almost 8-)

T Rob: BarbaraH - Just kidding.

tom in indy 42: We recently had an issue with failover in a Solaris VCS clustering environment. Failover occurred successfully... however, we were not able to revert back to the primary node later due to a hanging process (amqzfuma) on the primary node. Is there a best practice for making sure everything is cleaned up (shut down/killed) on the primary after a failover? We are considering adding logic to look for and kill hanging processes before VCS initiates startup but need to be sure we don't kill an active environment. Are there examples on how to handle this out there somewhere?

T Rob: What's keeping you on 5.3? Vendor code not certified on 6.0? App is sunsetting and not worth trouble to upgrade?

BarbaraH: Then I won't mention the 5.1 MQ clients...

PeterPotkay: tom, multiple QMs on that VCS cluster?

tom in indy 42: Not sure...

BarbaraH: 5.3 works for the small amount of data that is going through it. And they keep saying they are going to rewrite it to not use MQ...and they don't have time to test on a new version. (I'm not even sure that anyone remembers how to test the application.)

PeterPotkay: tom, what version of MQ? VCS? MC91? I'm surprised that there would be orphaned processes left behind with current software.

T Rob: @Tom - I usually include in the shutdown script commands to clean up orphan processes. Some people do this on start-up but, as you say, there is often an issue in determining whether the QMgr is legitimately running.

T Rob: @BarbaraH - rewrite the app to not use WMQ? Traitors! This sounds like what I've heard in several cases. App is low-priority and near end of life or is pending rewrite/refactoring.

PeterPotkay: I think any app that is designed to send MQ messages into an MQ cluster that has QMs across multiple sites, and connects to its QMs with MQ Client channel tables makes for a very easy MQ DR scenario. If one site goes down, all new traffic automatically goes to new QMs via new QMs. Messages still on queues at the dead site are N/A, but MQ queues should always be empty, and MQ is not a D/B, right? Anyone out there so lucky to have this type of setup?

tom in indy 42: Sorry to say that this is one of our few remaining MQ 5.3 environments.... it supports SAP interfaces and the application folks are in the middle of a migration and don't have time to test their interfaces with a new version of MQ.

T Rob: The clients on 5.1 concern me more though. Should be able to use a more modern client, even with an older QMgr. The Java class packaging changes some, but the API remains compatible. LOTS of issues with Java/JMS/Security/Function fixed in later clients.

PeterPotkay: Or is anyone using backup QMs with their "Log Shipping" style of replication?

T Rob: @BarbaraH - I would characterize use of code that old as an availability or DR problem. It is much more likely to be unstable and if it goes down, you get no help in fixing it so take an extended outage.

DepremCA: Are there any best practices for managing client channel tables for large numbers of clients and many different queue managers?

T Rob: DpremsCA - With v7.0, you can specify that connections are randomly assigned to one of a set of QMgrs.

tom in indy 42: We purchased extended support for 5.3 since the app is critical.... but just wanted to know how others handled the issue. Our VCS folks tell me that there is no way to look for and kill hanging processes on the primary node after failover to the secondary... so checking in the startup script seems to be the way to go.

Ian or Ivans: What angle does everyone approach DR from? What is the focus? A product (e.g. MQ, MB, DB)? A system/systems (e.g. a site of AIX box, or a sysplex)? A site? Something else?

T Rob: Prior to that you had to provide different channel tables where the QMgrs were ordered differently.

BarbaraH: The MQ clients are scattered throughout the organization (to connect to the 5.3 QManager). Again ... no one has time to test. And it doesn't seem to be critical enough to matter... it can have problems for days before anyone has a concern. (Not that MQ has problems for days -- just other pieces of the app).

PeterPotkay: tom, VCS allows for processes to be asynchronously kicked off at shutdown and start up, check out the MC91 Support Pack. You could kick off a script to clean up.

PSherman: Question: We have a Web service in the DMZ that was sending msgs into a QMgr on our internal network. But, the Web service wasn't always reliable and we occasionally lost msgs. We put a QMgr out in the DMZ - connected to internal QMgr by SSL - and have no problems with losing msgs, but apparently this is a "no-no". Any recommendations for reconfiguring this setup?

PeterPotkay: But it sounds like something else is wrong.

PeterPotkay: If VCS ends the QM, it should really end it. Maybe the script it calls to endmqm is lacking something.

T Rob: @PSherman - not sure why this configuration is a no-no. I usually recommend a gateway QMgr to isolate the internal QMgrs from direct connection of external things.

BarbaraH: Ian -- We use SunGard for our DR. We have some servers permanently there (including my MQ servers) and then rent the large machines, including mainframe. So we do site DR.

PeterPotkay: PSherman, we use MQIPT in the DMZ. Not having QM in the DMZ means not having messages sitting on disk in the DMZ - a good thing.

T Rob: In this case, it is your own app, but still, why was it discouraged?

chrisfra: @PSherman - An alternative approach might use MQ IPT in the DMZ.

PSherman: Right - the problem is that the msgs exist on the disk on the DMZ server - so it's actually a security breach.

PSherman: ok - we'll look at MQ IPT. Thanks!

T Rob: @PSherman - Ah, ok. It sounds like the connection might have been getting severed by the firewall. Having two QMgrs alleviates that because they have built-in detection and correction of the lost connection. If the app is using a client, it needs to tune that better.

T Rob: WMQ v7.0 might help out if that was the problem causing lost messages.

T Rob: Better channel tuning and detection of lost connections.

PSherman: Ah, ok, Rob ... I understand.

A J Aronoff: @PeterPotkay: You mentioned used clusters in DR.

T Rob: WMQ IPT might help as well. No messages go to disk and you can bridge between older WMQ client and newer WMQ server, or vice versa.

T Rob: You can pass the SSL credentials through MQIPT, or use it as an SSL proxy.

A J Aronoff: @PeterPotkay. We have a customer who is doing a lot of DR via clusters. Are there any best practices to recommend? Are there any cluster properties that you recommend setting so that the traffic stays on the primary system until a failover (cluster rank, etc)?

T Rob: What it does NOT do, however, is relieve you of any of the normal WMQ hardening you would do with a direct connection. Still need MCAUSER everywhere, still need to authenticate connections, etc.

tom in indy 42: Thanks Peter.... I think the stop script is probably OK since it has been in place for some time and we have only seen this occur once just recently..... I'll discuss the possibility of using functionality from the MC91 support pack to kick off a cleanup process with our Unix folks.

lan or ivans: @Barbara. Thanks. Sounds good. I always assume DR is about a site, but I've come across situations where a product or system team has done their own DR. I guess I'm trying to work out what the reasons are behind not starting from a site level.

lan or ivans: A J Aronoff: Clusters and DR. Are you describing having two active sites and routing only to one?

PeterPotkay: I mentioned missing messages and duplicate messages with asynchronous replication. Here's why. MQPUT happens at 12:00:00. You are 20 seconds behind for replication. DR strikes at 12:00:10. The MQPUT worked, BUT the message is missing in DR land. Or you MQGET at 01:00:00, DR strikes at 01:00:10, and the message is there again in DR land. Duplicate messages. It gets very hairy, very fast, and most apps aren't designed for that. As soon as you say you are replicating data, the head honchos in the corner office assume everything will be there in DR land. I don't go down that path. I rather promise a stable functional ready DR MQ infrastructure, and sorry if you had any messages sitting in a queue. We had a disaster after all. Your databases should have your critical data, riiiiight? 😊

A J Aronoff: The customer's DR site is currently passive.

A J Aronoff: @peterpotkay: The customer's DR site is currently passive. There is no automated method of starting up the D site. This seems weak to me. Is it possible to have two active sites, but set cluster properties so that traffic would only go to the DR site in the event of a failure? Would this be a recommended practice?

lan or ivans: A J Aronoff: If the site is passive, the channel to it will not be running. If the channel to the main site is healthy, then WMq clustering will load balance all messages to the healthy system (based on channel status).

lan or ivans: A J Aronoff: So does "passive" mean that the site is not available?

PeterPotkay:

AJ, the problem is MQ clustering sees any channel status other than Inactive or Running as bad, and will decide to go to the "alternate" route. That makes sense if the channels are down for a legit DR reason. But what if it's just binding, initializing, etc, for a second on its way to a normal start? I prefer to control MQ cluster to DR sites by not clustering the DR queues until a DR, and having a FR in the DR site and a FR in the primary site.

A J Aronoff: @lan or ivans. I believe that clustering has rank and priority properties. That should allow work load balancing to only use the D site when needed.

PeterPotkay: If you have a passive DR site, you have to be real careful not to allow any work to go there accidentally. A normally starting channel can't be allowed to do that to you.

T Rob: AJ - I have a number of customers who can have active/active apps. For example, they have sites in New York, San Fran, and London. Any site can process any message, but one is considered "primary". This works great with clusters because the backup or DR site can get a message at any time, even if it is weighted really low.

T Rob: But you have to design the app ahead of time to do this.

A J Aronoff: @T.Rob. Awesome! Do you have a write up, documentation, best practice link? Thank You!!!

lan or ivans: A J Aronoff: Yes, you can use CLWLRANK. But that will mean that to fail over requires a manual change (i.e. change the CLWLRANK value).

A J Aronoff: @PeterPotkay. OK. It is the lack of automation at the customer site that worries me. Do you have some recommendations best practices, etc, to determine if the primary site is down? Then automatically do what is required on the DR site?

PeterPotkay: It's sorta what I posted up above, that a cluster spanning data centers is the best way, assuming some do half of the work can go to the DR site at ant time.

T Rob: A.J. - I will have to add that to my list of topics for the Mission:Messaging column. Nothing I can share at the moment as it was all customer contracted work.

PeterPotkay: Most sites will not allow DR to be determined automatically. It has to be a human decision, typically at the CIO or CEO level.

Ian or Ivans: Whilst in the subject of WMQ clustering. When using WMQ clustering, the strong recommendation is to use bind-not-fixed routing. This means that every message is workload balanced, and importantly, that messages on the SYSTEM.CLUSTER.TRANSMIT.QUEUE can be automatically load balanced again if the channel fails before it can get the messages off the SYSTEM.CLUSTER.TRANSMIT.QUEUE.

tomyoung: Hi all, we only have about 3 minutes left, so if anyone has any final questions or comments, please jump in.

A J Aronoff: @PeterPotkay. None of the work can go to the DR site, unless there is a failover (licensing). Your comment about the automatic failover makes sense. Thank you!

A J Aronoff: @ T Rob. Thank you very much!!

A J Aronoff: @Ivan, I believe that with the CLWLRANK property, if the primary is down, then the traffic would flow to the DR site.

tomyoung: Okay, we're going to end the discussion here. Thanks to everyone for participating. We'll post a transcript of today's chat here:

http://www.ibm.com/developerworks/websphere/library/chats/0902_mqchat/0902_mqchat.html

Ian or Ivans: A J Aronoff: No.

BarbaraH: Thanks.

Ian or Ivans: CLWLRANK is a very powerful feature.

DepremCA: Thanks all.

A J Aronoff: @ian, thank you very much. I will reread the documentation.

PeterPotkay: Thanks everyone. Feel free to post question on the MQ list serve and on <http://www.mqseries.net/>. I frequent both and love to discuss MQ clustering, DR, and H.A.

mfrost: Thank you.

Ian or Ivans: A J Aronoff. If you set CLWLRANK, then availability is ignored.

T Rob: Thanks everyone!

A J Aronoff: @Ivan - thank you Ivan!