

Optimised WLM PodCast script

HOST: Hi, and welcome to the “Did you say Mainframe?” podcast series. This is where we regularly interview IBM technical experts who can help you to understand important IBM mainframe hardware and software issues. I'm your host Nick Garrod.

Today we're going to talk about [Performance improvements that customers can realize in CICS Transaction Server 4.1, by implementing Workload Management from CICSplex System Manager]

Our guest today is [Dave Williams] from [the CICS Systems Management development team in Hursley]. Hi Dave, it's great to have you here.

SME: Hi Nick, thanks for the invitation.

HOST: Before we begin, I'd like to mention to our listeners that there is more information available in [a short demo.] which I'll be giving details of at the end of this podcast.

Q1 So Dave, we have heard a lot about CICS TS V4.1 and the wonderful new capabilities it has delivered in the web 2.0 and Event space and how the CICS Explorer can help productivity of administrators and developers, but this title particularly intrigues me - it looks as though it offers benefit for early adopter's of new technology and the traditional workload customer alike, could you elaborate a little?

A1 Yes, we'd been doing some work on CICSplex SM's workload manager function in response to customer reports that it seemed to go awry when under throughput stress – which of course is the time that the system should be performing at its optimum level.

Q2. And you found that wasn't the case?

A2 Well being a developer, I usually treat reports of failure in the system with a little scepticism, so I created a sizable test case to verify that all was well, and to my horror I discovered that the customer reports were correct – For certain types of workload, WLM was not doing the best job that it could.

Q3 Certain types of workload, could you elaborate?

A3 Yes, since its inception, WLM has functionally evolved to cope with the additional types of Dynamic Routing requests that have become available with each CICS TS release. I found that for Synchronous types of request - such as the traditional Dynamic Transaction routing requests or Dynamic Program Links - the system seemed to behave perfectly well. However, for Asynchronous dynamic requests like those arising from an EXEC CICS START API command, then anomalies started to appear.

Q4 Anomalies? That sounds like a bug!

- A4 From my perspective, the word bug means that a program has been coded incorrectly. In this instance, WLM is actually working as designed. The problem is that it was designed to perform best when executing within the bounds of a single LPAR. Now in the early 90's this was a fairly common environment, but as the scope of customers' systems has expanded over the last 15 years – with parallel sysplexes becoming commonplace – then WLM has not kept up to exploit the newer technologies.
- Q5 [So what new technology did you use to address this issue.](#)
- A5 Well its not quite so new now, but is still very useful to us – we used the z/OS Coupling Facility, which we exploited using the existing CICS Coupling Facility Data Table function.... And it proved to be highly beneficial.
- Q6 [So exactly what was the problem that it resolved?](#)
- A6 Well Nick, I'm not looking at anyone in particular, but lets paint a picture for the less technically minded!
- Q7 [Please do!](#)
- A7 Ok - lets imagine a 1950's style office, with a manager sat at his desk at the front, and he directs work to his staff in the office for processing. So lets say that this is a quiet day, and he has a couple of staff waiting for work from him – Bill and George – very reliable. So the company messengers periodically appear at his office door with work requests which they deliver to him, and its his job to see all these requests get cleared down as quickly as possible – his Christmas bonus is riding on it! So for each work item that he has to distribute to his workers in-trays he has to make a decision about who to give it to.
- Q8 [Surely he'll give it to his fastest worker?](#)
- A8 Yes, maybe he should – but what if that worker already has 100 items in his in-tray, and because you've already given him so much to do he's broken his pencil?
- Q9 [What, and you're waiting for him to sharpen it?](#)
- A9 Exactly! The manager must evaluate a combination of conditions before giving the work out
- Who is my fastest worker?
 - How much work has he got currently?
 - Is he in a state to do this job?
 - How healthy is he?
- The answers to all of these questions must be factored together before he can decide who to route the work to. And on top of that, there might be other special considerations – Form XYZ123 must be completed with a red pencil, and George is the only worker who owns one, so he has special rules to apply to particular work items.
- Q10 [So the decision is not quite as straightforward as it would first seem then?](#)
- A10 No, in fact its very important that he makes well considered decisions, because otherwise work in the office would be backing up – and his bonus would be at risk! However, so long as he maintains the decision making rationale, closely monitoring the state of his workers before passing work out, then all should be OK.
- Q11 [Well this is a very vivid picture of 1950's office life, so how does it relate to WLM?](#)

- A11 Well its pretty simple – The Office manager is the user’s Routing region, which we know traditionally as the Terminal Owning region or TOR, and the workers are the Target regions, otherwise known as the Application Owning regions or AORs. Work gets queued to the Router for redirecting to the most appropriate Target, according the conditions prevalent at the time of the decision. And in a nutshell, that is how WLM works.
- Q12 [Well that sounds fine to me. Where is the problem?](#)
- A12 Ahhh. Well we need to revisit the office. Bill and George are admirable workers, but on busy days they can get swamped. So when the executive management see that the office cannot cope with the extra load, what do you think they’ll do?
- Q13 [Ummmm, get another worker in?](#)
- A13 Yes, let’s assume that the executives understand their business, and home in on the problem straight away. They assign Fred to join Bill and George in the office. And in doing so – without changing any of their in-house processes, they have increased the work capacity of the office by 50%.... and the Christmas bonus is looking good again!
- Q14 [Good news!](#)
- A14 Yes, good news, but Fred joined the office in September and now we are three weeks away from Christmas and the office workload has maxed out again, and it looks like the workload will increase still further. Uh-oh – the bonus is at risk again!
- Q15 [No problem, get another worker in!](#)
- A15 Great idea! But there’s a problem, the office hasn’t got room for another desk! Time to put in another executive call!
- Q16 [So how do they resolve that then?](#)
- A16 Easy, they move Tom into the office next door, where he’ll join the team remotely.
- Q17 [How will that work?](#)
- A17 Well the manager will have a couple of operational issues. The first is that it will take him longer to pass the work to Tom next door, rather than to Bill, George or Fred who share his office with him. So he has an additional routing factor, which means in terms of locality, Bill, George and Fred are preferable to Tom. But there is one other more critical factor.
- Q18 [Yes?](#)
- A18 Bill, George and Fred are in the manager’s view. He can see how much work they have in their in-trays at any time, and when he’s deciding who to give work to, the existing content of the workers’ in-trays is a very significant factor in making his routing decision.
- Q19 [But Tom has been put in the next office. How can he see how busy he is?](#)
- A19 Well he can’t! So he assigns a company messenger to report the size of Tom’s workload to him every five minutes. So every five minutes, he gets updated on the volume of work that is in Tom’s in-tray.
- Q20 [So the capacity problem is solved then?](#)

- A20 It would seem so wouldn't it? So you tell me, when Tom has just joined the workload, what will his workload status be? Its not a trick question!
- Q21 Well, it will be nothing won't it?
- A21 Yes, nothing. So when the next work item comes in, the manager will compare the workloads of Bill, George, Fred who are already busy with that of Tom. Who's he going to choose?
- Q22 Well that latest state reported by Tom was that he has nothing to do, so I guess it goes to Tom?
- A22 Absolutely. An empty in tray will generally outweigh all other routing factors, so the work item goes to Tom. And what happens with the next work item that quickly follows in? Who gets selected?
- Q23 Tom again?
- A23 Yes, Tom again....because all the manager knows about him is the latest messengers report that says he has nothing to do. So where do all the other work items go that come through the door in the next five minutes?
- Q24 It doesn't sound like Tom's lucky day!
- A24 You could say that. So finally, five minutes pass and a then messenger appears with Tom's latest status report. What do you think its going to say.
- Q25 I imagine it would make interesting reading!
- A25 Yeah, it certainly wouldn't say Merry Christmas! It's going to say STOP! I'm completely overloaded, stop sending stuff – in big bold letters!!! So where will next work item go?
- Q26 Anywhere but Tom?
- A26 Yep. Anywhere but Tom. For the next five minutes Bill, George and Fred are going to get all the work because the manager's last message from Tom was that he is overloaded. Now this is a shame because it took 2 minutes for Tom to clear down his backlog, so for three minutes he idling with the Times' crossword because he has no work. So what is Tom's next status message going to be to the manager?
- Q27 Well if he was smart he'd say nothing, but I guess in this instance he's going to say that he has no work.
- A27 Indeed, and so we're now back in the situation where Tom's getting all the work, and Bill, George and Fred are idle. And it'll stay like that for the next five minutes, and until Tom's next scream of desperation arrives, stating that he is swamped again!
- Q28 Now just a minute – it sounds to me like you have said that the company had a workload capacity issue, and so they added another worker to alleviate it, and it has made things worse!
- A28 Yes I'm afraid I did - and I can make things even more worse! Let's assume that our red hot company executives have been monitoring the situation, and the say "Ahhh, Tom's in-tray is filling up every five minutes. We need another worker – we'll put Phil in the office with Tom!" So Phil joins Tom as another remote worker. When the next work item comes in, who's going to get it?
- Q29 Phil, because he has nothing to do?

- A29 Quite likely, because the managers last report from Tom was that he is swamped again, and in the meantime work has been shared between Bill, George and Fred locally.... but what we have now is a nightmare scenario where either Tom or Phil will be saying every five minutes that they have nothing to do, and then five minutes later that they are complete maxed out. And the spin off effect from that is that Bill, George and Fred in the local office – who are the most convenient workers for the manager to pass work to - now only see a fraction of the total workload. So the net result of adding two remote workers to the workload is that the overall workload throughput is now even slower than when the workers comprised of just Bill, George and Fred.
- Q30 That doesn't sound good! How does it translate to WLM?
- Q31 Well for office, think LPAR. When the capacity of the local LPAR has run to its maximum, then the Systems Administrators have no choice but to run additional Application regions in an adjacent LPAR in the Sysplex. And regrettably, exactly the same performance issues arise as those I have just described in our virtual office.
- Q32 Tell me this is the problem you've solved!
- A32 We have indeed. Instead of the Manager looking at the workers in-trays to determine how much work they have, a huge billboard was erected on the opposite side of the road, and every worker posts his work state onto that board whenever his workload crosses a predefined load boundary. When the manager has decide who to give work to, he looks at the billboard and gets a good relative comparison of all his workers (local and remote) from a single source. And then net result of that is that his Christmas bonus is intact!
- Q33 Hooray.
- A33 Yep! Celebration time. And in the case of WLM, that billboard is the z/OS Coupling Facility. So we are now back in a regime where routers are making more appropriate routing decisions, regardless of the volume of transaction throughput at the time.
- Q34 I thought these changes were all about performance?
- A34 Well indirectly they are. When we did the first graphical comparisons between these Sysplex Optimised workloads, and those in "traditional" mode, what hit us immediately was the shape of the graphs – no workload batching in evidence anywhere with the optimised workloads – no throughput surging to MaxTask and back to idle. Just a nice smooth even distribution over all the target regions. It was a little later when I was admiring them that it suddenly hit me that there were fewer time periods showing for the optimised workloads – they were running through 40% faster! So in resolving a deep, technical functional problem we had achieved a significant performance improvement as a very fortunate spin-off!
- Q35 I suspect some customers will be concerned that we have laid another high performance function on top of an already overloaded Coupling facility.
- A36 Yes, and I share your suspicion. Finding a processing algorithm which solved our problem and had the least impact of the CF was one of the central tenets of the solution. So we've made numerous definitional controls available which will allow users to tune the impact to their coupling facilities, but I have to say that the default values that we have chosen were based on many, many hours of testing time, and I advise caution before changing them in a live environment
- Q37 So users of WLM have this in store for them when they upgrade to CICS TS 4.1?

A37 Yes, but there's other good WLM stuff as well. We've reworked the WLM Viewset structure in the WUI. Users can now perform proper generic drill-down queries of workloads routers and targets, and the target region status reporting is hugely improved. Users will now be able to tell how what their target region states are, and how much they are being exploited – this is all from new WLM runtime views. So we put lots of effort in to upgrading the runtime processing, and the status and utilisation reporting through the WUI. There's lots of new stuff to check out!

Q38 [Where do customers go from here?](#)

A38 The first place of reference is always the CICS Information Centre. However, I did publish a technote that condenses the Sysplex Optimised Workloads function into a concise document. However, you can see a real life comparison of Optimised and non-Optimised workloads on YouTube entitled "Performance comparison between CICS TS V3.2 and CICS TS V4.1 with WLM". This is recommended viewing if you are interested because "talk is cheap", and this is live, like for like comparison. It's about 10 minutes long..

Host [Thanks Dave, that was very interesting.](#)

SME Hey, no problem, I'll talk to anyone if they offer me a beer!

Host [Well, that wraps up this podcast discussion. To find out more about \[CICS performance improvements and the various information that is out there to help you understand and implement some of the new function\], please go to the description for this podcast at: <http://www.ibm.com/software/os/systemz/podcasts/websphereonz/>](#)

[Join us next time as we talk about another important mainframe topic. For now, this is Nick Garrod saying "Thanks for listening".](#)

Title: CICS TS V4.1 WLM Sysplex Optimised Workloads

Sub Title: CICS TS V4.1 Enhances and performance improvements to CPSM WLM.

Summary: Hear about how Sysplex Optimisation addresses dynamic routing problems across Parallel Sysplexes, and the performance gains that can result from it.

Keywords: CICS, CICSplex SM, Workload Manager, Sysplex Optimised Workloads, Coupling facility

Speaker Name: Dave Williams

Speaker Bio:

Dave Williams is a Senior Developer in the CICSplex SM development team.

Headshot from bluepages:

To be provided.

Offer: Demo on youTube (<http://www.youtube.com/watch?v=RpS2h4jZdsU>)

References:

Technote:

<http://www-01.ibm.com/support/docview.wss?uid=swg21384973>