

PIR - Microsoft central outage impacting multiple services hosted within Northern Europe

M&S Digital&Tech *Post Incident Review*



Incident/ Problem Reference: 9069849

Incident Date: 01/04/2025

PIR Date: 08/04/2025

Author: Rehan Siddiqui

Attendees:

Charley Webster; Emmanuel George; Hisham Khalid; Valliammai Ramasamy; Vivekananda Gaikwad; Ezhilarasi Baghavan; Ramachandrarao rvs; Rehan Siddiqui; Saloni Chowdhury; Sri Balaji Seenu; Naveen Rayner Ernest; Gokula Krishnan.

Technical Owner: Emmanuel George

👉 *Analyses the root cause and provide technical recommendations to prevent incident recurrence.*

👉 *Develop and oversee the implementation of an action plan to address the root causes.*

👉 *Work closely with cross-functional teams to gather information and insights.*

1. Incident Summary & Impact:

👉 *Provide a concise overview of the incident with brief update on issue, impact, root cause and resolution.*

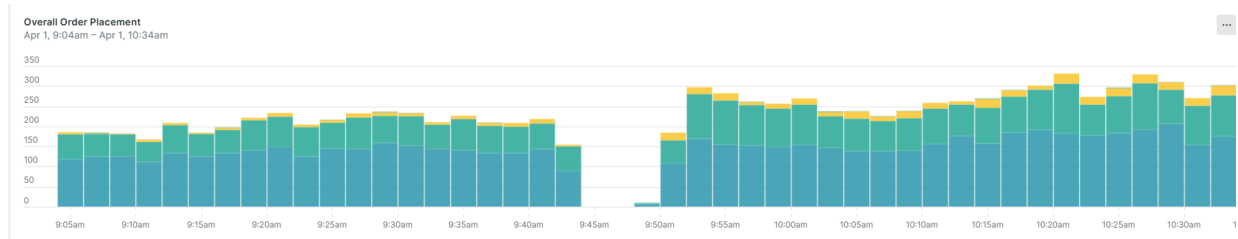
Multiple applications/services hosted on Microsoft Azure North Europe region were impacted between 09:40 and 11:30. Customers were unable to place orders on the UK and Ireland website for 4 mins resulting in an order drop and loss of sales. Also, Scan & Shop Mobile, Retail In-day Sales Dashboard, Stock Movement Tool, Power BI Reports, and SAP POSDTA application were impacted. The root cause was attributed to an unplanned power outage at Microsoft's North Europe Data Centre. Services were recovered after the power was restored at Microsoft datacentre.

Impact:

No orders were taken between 09:40 to 09:44, following this we started getting a few orders and around 09:48 we started seeing normal flow of orders.

Scan and Shop Mobile application was inaccessible to customers between 09:50 and 10:45

Retail in day dashboard was not showing up to date sale between 10:30 and 11:45



x`

	Product Portfolio	Impact Statement
1	Selling Experience	WCS Live DB VM impacted as part of MS Outage, application failed over to Secondary DB automatically.
2	Selling Experience	WCS AUTH - Primary and Secondary DB VM unavailable as part of MS Outage
3	Service Experience	Scan & Shop Mobile (MPG) - Customers were unable to use Scan & Shop Mobile between 09:50 and 10:45
4	Data	<p>BEAM:- Retail in-day dashboard used by store users were not getting updated with latest data between 10:30 and 11:45</p> <p>Stock movement tool used by suppliers and Voice of customer data were updated between 10:30 and 11:56</p> <p>Power BI reports were not available for the colleagues with latest data between 10:30 and 11:30</p>
5	Foods	ASO - Storage account Issues; no impact to the VM / production application

2. Incident Chronology :

👉 Include a record of the data/time of key events. e.g. "07:36 Incident started, notified by ..." **keep fact based**.

09:40 - Order volume begins to rapidly decline. Sharp increase in exceptions in WCS Logs indicating a connectivity issue between WCS and the primary database. WCS Primary Database experiences an issue and begins failover process to secondary.

09:43 - Order dip reported on TechOps channel

09:45 - Database failover completes to secondary which now becomes primary. Orders volumes start to increase.

09:48 - Order volumes are back to BAU - Service Restored.

09:49 - Cloud Ops team received alerts stating that Virtual Machines in North Europe would have observed connection failures when trying to access some Virtual Machines. These Virtual Machines may have also restarted unexpectedly.

09:50 - Incident bridge opened on Incident channel

09:55 - MS Support ticket 2504010030006257

10:10 - Cloud Ops team advise that multiple warnings received on the dotcom Azure subscription indicating an issue on Service Health

10:18 - Microsoft azure support on-boarded to incident bridge for further updates

10:38 - WCS Auth Primary DB is also impacted by this issue and is not able to failover

10:59 - WCS confirmed that the impacted DB VM had recovered

12:01 - **Latest Update from MS**

Summary of Impact: Between 08:51 and 10:15 UTC on 01 April 2025, we identified customer impact resulting from a power event in the North Europe region which impacted Virtual Machines, Storage, CosmosDB, PostgreSQL, Azure ExpressRoute and Azure NetApp Files. We can confirm that all affected services have now recovered.

A power maintenance event led to temporary power loss in a single datacenter, in Physical Availability Zone 2, in the North Europe region affecting multiple racks and devices. The power has been fully restored and services are seeing full recovery.

12:02 - WCS Live Standby DB is in sync with production

12:04 - Data team updated that the Retail in-day dashboard is now fully updated and reflecting data to 11AM. Performance with PowerBI has improved and we are seeking user confirmation of any ongoing slowness

3. Incident Resolution

👉 *The actions taken to restore services and mitigate impact.*

The WCS database takes approximately four minutes to fail over from primary to secondary due to its high availability configuration. While this setup helped mitigate the overall impact, it did result in a four-minute service gap.

Mobile Scan & Shop does not currently have a formal disaster recovery (DR) solution; however, workarounds are in place using handheld scanners, belted tills, and self-checkout options.

Regarding Power BI and data services, the impact falls under an accepted risk (M&S reference: 1584970835), as the decision was made not to subscribe to a secondary failover region.

4. Root Cause

👉 *Refers to the fundamental reason (s) why an incident occurred including problem description, analysis techniques, contributing factors, evidence with concise root cause statements*

A Microsoft power maintenance event led to temporary power loss in a single datacentre, in physical availability Zone 2, in the North Europe region affecting multiple racks and devices. The power was fully restored.

5. Findings

👉 *Refer to the key insights and conclusions drawn from the incident analysis - what went well/what went wrong highlighting gaps, failures, or inefficiencies.*

5.1. Observability –

👉 *Evaluating monitoring, logging, and alerting effectiveness in detecting the incident.*

In WCS, application alerts flagged an order drop, but the cloud team did not receive the corresponding VM alert until 09:49, as it only triggers after a VM has been down for five consecutive minutes. The PIR identified that the current generic alerting across the estate is insufficient. Enhancements are needed to improve visibility of impact, enabling quicker response and timely engagement of the right resources.

5.2. Technical Challenges -

👉 *Identifying system, infrastructure, or architectural issues that complicated resolution.*

The WCS database takes approximately four minutes to fail over from primary to secondary due to its high availability setup. While this is a positive aspect, the current solution is active-passive. To ensure truly seamless service, an active-active configuration would be more effective.

The WCS disaster recovery (DR) capability was last tested several years ago and now requires a refreshed and properly planned DR test.

Mobile Scan & Shop currently does not have a formal DR solution; however, a workaround is in place using handheld scanners or belted/self-checkout tills.

Power BI also lacks a DR solution. This has been documented and is considered an accepted risk as M&S have chosen not to subscribe to it - risk number: 1584970835

5.3. MIM involvement & impact assessment -

👉 *Timely engagement of MIM and analysing the impact on business operations based on top 10 business processes*

Due to insufficient alerting across the estate, it was difficult to quickly determine the full scope of the impact, which in turn delayed the engagement of the appropriate subject matter experts.

5.4. Incident handling/ stakeholder management –

👉 *Reviewing how teams collaborated to drive the incident to resolution and engaged with key stakeholders.*

Insufficient alerting across the estate led to delays in engaging the appropriate subject matter experts.

5.5. Business Communication –

👉 *Clear, concise & timely communication to stakeholders, focusing on business impact and recovery actions.*

The major incident text was initially missed by the .com service leads, which led to delayed communications from the MIM team. Access issues with the text tool have since been resolved, and it has been reinforced to both the service leads and the MIM team that all communications must be timely and transparent moving forward.

5.6. Supplier Engagement –

👉 *Assessing the responsiveness and support provided by external vendors and internal product teams.*

Although Microsoft joined the call promptly, as is often the case with this vendor, they were unable to provide details regarding the issue, its root cause, the estimated time to recovery, or the resolution steps taken.

Mitigation Actions:

👉 *Agree actions for each of the findings to prevent recurrence, strengthen system reliability & stability, and improve response efficiency for future incidents*

	≡ Action	≡ Owner	📅 Due Date	≡ Remarks	📁 Action Category
1	The WCS database disaster recovery (DR) test has not been conducted. A DR test plan should be developed and scheduled as a priority.	⊕ Webster, Charly	Tue, May 6, 2025		Problem Avoidance
2	WCS database failover currently takes approximately four minutes. Review the feasibility of implementing an active-active solution to reduce or eliminate service disruption.	⊕ Webster, Charly	Tue, May 6, 2025		Problem Avoidance
3	Review the current alerting configuration for VMs, which is set to trigger only after five minutes of consistent unresponsiveness, to ensure earlier detection of issues related to service impact.	⊕ George, Emmanuel 👤 Chowdhury, Saloni 👤 Siddiqui, Rehan	Tue, Apr 15, 20...		Problem Identification & Alerting

4	Request and obtain the full Root Cause Analysis (RCA) statement from Microsoft.	 Siddiqui, Rehan	Tue, Apr 15, 20...		Root Cause
5	As there is currently no disaster recovery (DR) solution in place for the Scan & Shop mobile application, an accepted risk should be formally raised to document this gap.	 Baghavan, Ezhilarasi  Ramasamy, Valliammai	Tue, Apr 15, 20...		Solution & Recovery
6	Define and document clear expectations for the information Microsoft should provide when joining a major incident call. Engage vendor management to ensure these expectations are communicated, agreed upon, and consistently met.	 Siddiqui, Rehan	Tue, Apr 15, 20...		Supplier Engagement

Please refer to  Major_ Significant and Key Incident Tracker v1.0.xlsx

for progress updates on Mitigations

<Mandatory participants: Put it in the meeting invite. >

Technical Details and Useful Documents: