

你是一名 SRE,今日發生系統服務中斷狀況,造成客戶極度不滿,要求你提供事件報告。 我會提供你事件真實發生的時間點及處置狀況,請撰寫一份事件報告,內容須包含事件描述、影響範圍及改善計畫。

改善計畫的部分請盡可能條列多項並詳細說明。

- 09:00 接獲多家客戶反映服務無法登入。
- 09:03 從監控數據發現多台機器有 CPU 過高的狀況,使用率皆超過 99%。
- 09:05 因未設定自動擴展機制,先以人工方式升級機器規格,並調高機器數量。
- 09:10 自動擴展生效後,機器 CPU 過高現象減緩,服務恢復可用狀態。
- 09:15 通知客戶服務已可正常使用。



事件報告

事件描述:

2023 年 3 月 30 日,09:00,我們收到多家客戶反映無法登入我們的系統服務。經調查後發現,多台機器的 CPU 使用率超過 99%,造成系統無法正常運作。由於未設定自動擴展機制,我們在09:05 以人工方式升級機器規格並調高機器數量。在 09:10,自動擴展生效後,機器 CPU 過高現象減緩,服務恢復可用狀態。最後,在 09:15,我們通知客戶服務已可正常使用。

影響範圍:

此次事件影響了多家客戶,導致他們在約 15 分鐘的時間內無法正常使用我們的系統服務。

改善計畫:

為防止類似問題再次發生,我們提出以下改善措施:

- 1. 設定自動擴展機制:自動根據負載狀況擴展或縮減機器數量,以確保服務的穩定性。
- 2. 優化系統性能:對系統進行性能分析,找出可能的瓶頸,並進行優化以減少資源消耗。
- 3. 實施監控告警:設定合適的告警閾值,對 CPU、內存等重要指標進行實時監控,一旦超出閾值即時通知相關人員。
- 4. 定期檢查:進行定期檢查,確保系統配置正確且運行良好,提前預防潛在問題。
- 5. 建立風險評估機制:對現有系統進行風險評估,確保具備充足的容量以應對意外情況。
- 6. 加強與客戶溝通:建立更有效的客戶溝通渠道,及時了解客戶需求,適時調整服務策略。

我們將持續監控並改善系統,以提供客戶更穩定可靠的服務。再次為此次服務中斷事件向所有受 影響的客戶表示誠摯歉意。