AWS Well-Architected Framework

ตุลาคม 2015

ประกาศ

เอกสารนี้จัดทำขึ้นเพื่อวัตถุประสงค์ในการให้ข้อมูลเท่านั้น โดยจะแสดงเนื้อหาเกี่ยวกับแนวทางปฏิบัติและคุณสมบัติที่ AWS นำเสนอ ณ วันที่ออกเอกสารฉบับนี้ ซึ่งอาจเปลี่ยนแปลงได้โดยไม่ต้องแจ้งให้ทราบ ลูกค้าต้องรับผิดชอบต่อการประเมินข้อมูลในเอกสารฉบับนี้ และการใช้งานผลิตภัณฑ์หรือบริการใดๆ ของ AWS ด้วยตนเอง ซึ่งข้อมูล ผลิตภัณฑ์ และบริการต่างๆ นั้นมีให้ “ตามสภาพที่เป็นอยู่” โดยไม่มีการรับประกันในลักษณะใดๆ ไม่ว่าโดยแจ้งหรือโดยนัย เอกสารฉบับนี้ไม่ได้ก่อให้เกิดการรับประกัน การรับรอง ข้อผูกพันทางสัญญา เงื่อนไข หรือการประกันใดๆ จาก AWS รวมทั้งบริษัทในเครือ ผู้จัดหา หรือผู้อนุญาตให้ใช้สิทธิ์ของ AWS หน้าที่รับผิดชอบและความรับผิดของ AWS ต่อลูกค้าได้รับการควบคุมตามข้อตกลงของ AWS และเอกสารฉบับนี้ไม่ได้เป็นส่วนหนึ่งและไม่ได้เป็นการแก้ไขข้อตกลงใดๆ ระหว่าง AWS กับลูกค้า

สารบัญ

บทคัดย่อ

เอกสารฉบับนี้อธิบายเกี่ยวกับ **AWS Well-Architected Framework** ซึ่งเป็นเฟรมเวิร์กที่ช่วยให้ลูกค้าสามารถประเมินและปรับปรุงสถาปัตยกรรมในระบบคลาวด์และเข้าใจผลกระทบทางธุรกิจจากการตัดสินใจด้านการออกแบบได้ดียิ่งขึ้น เราจะพูดถึงหลักการออกแบบทั่วไป แนวทางปฏิบัติและคำแนะนำที่เฉพาะเจาะจง โดยแบ่งออกเป็น 4 แนวคิดที่เราถือเป็น *เสาหลัก* ของ Well-Architected Framework

ข้อมูลเบื้องต้น

Amazon Web Services (AWS) เข้าใจถึงประโยชน์ในการให้ความรู้แก่ลูกค้าของเราเกี่ยวกับแนวทางปฏิบัติเชิงสถาปัตยกรรมสำหรับการออกแบบระบบที่เชื่อถือได้ ปลอดภัย มีประสิทธิภาพ และคุ้มค่าในระบบคลาวด์ เพื่อให้ได้รับประโยชน์ดังกล่าว เราจึงได้พัฒนา AWS Well-Architected Framework ขึ้นมา ซึ่งจะช่วยให้คุณเข้าใจถึงข้อดีและข้อเสียจากการตัดสินใจสร้างระบบต่างๆ บน AWS เราเชื่อว่าระบบซึ่งมีโครงสร้างสถาปัตยกรรมที่ดีนั้นจะช่วยเพิ่มโอกาสความสำเร็จให้กับธุรกิจได้เป็นอย่างมาก

สถาปนิกด้านโซลูชันของ AWS มีประสบการณ์ที่ยาวนานในด้านการออกแบบสถาปัตยกรรมโซลูชันให้กับธุรกิจเฉพาะทางและกรณีการใช้งานที่หลากหลาย และเราได้ช่วยออกแบบและประเมินระบบสถาปัตยกรรมบน AWS ให้กับลูกค้ามาแล้วหลายพันราย ประสบการณ์ที่ยาวนานนี้ทำให้เราระบุได้ถึงแนวทางปฏิบัติและกลยุทธ์หลักในการออกแบบสถาปัตยกรรมระบบในคลาวด์ AWS Well-Architected Framework จัดทำเอกสารชุดคำถามพื้นฐานที่จะช่วยให้คุณทราบว่าระบบสถาปัตยกรรมที่เฉพาะเจาะจงนั้นสอดคล้องกับแนวทางปฏิบัติของคลาวด์หรือไม่ เฟรมเวิร์กดังกล่าวจะนำเสนอแนวทางที่สอดคล้องในการประเมินผลระบบเทียบกับคุณภาพที่คุณคาดหวังจากระบบคลาวด์ที่ทันสมัย และการปรับปรุงที่อาจต้องดำเนินการเพื่อให้ได้คุณภาพตามที่กำหนด ขณะที่แพลตฟอร์ม AWS พัฒนาไปเรื่อยๆ และเราเรียนรู้จากการทำงานกับลูกค้ามากขึ้น เราจะสามารถปรับปรุงคำจำกัดความของระบบที่มีการออกแบบสถาปัตยกรรมที่ดีได้อย่างต่อเนื่อง

เอกสารฉบับนี้มีไว้สำหรับบุคคลซึ่งมีบทบาทด้านเทคโนโลยี เช่น ประธานเจ้าหน้าที่ฝ่ายเทคโนโลยี (CTO) สถาปนิก นักพัฒนา และสมาชิกทีมปฏิบัติงาน หลังจากอ่านเอกสารฉบับนี้ คุณจะเข้าใจถึงแนวทางปฏิบัติและกลยุทธ์ที่ AWS จะใช้ในการออกแบบสถาปัตยกรรมระบบคลาวด์ เอกสารฉบับนี้ไม่มีรายละเอียดการดำเนินการหรือรูปแบบเชิงสถาปัตยกรรม แต่จะมีการอ้างอิงถึงแหล่งข้อมูลที่เหมาะสมสำหรับข้อมูลดังกล่าว

คำจำกัดความของ AWS Well-Architected Framework

ทุกๆ วันผู้เชี่ยวชาญที่ AWS ให้ความช่วยเหลือลูกค้าในการออกแบบสถาปัตยกรรมระบบให้ได้รับประโยชน์จากแนวปฏิบัติในระบบคลาวด์ เราทำงานร่วมกับคุณเพื่อวิเคราะห์หาข้อดีข้อเสียเชิงสถาปัตยกรรมพร้อมกับที่งานออกแบบของคุณพัฒนาไปเรื่อยๆ ในขณะที่คุณติดตั้งใช้งานระบบเหล่านี้กับสภาพแวดล้อมการใช้งานจริง เราได้เรียนรู้ประสิทธิภาพการทำงานของระบบ และผลสืบเนื่องที่เกิดจากข้อดีข้อเสียดังกล่าว

จากสิ่งที่ได้เรียนรู้ เราจึงได้จัดทำ AWS Well-Architected Framework ซึ่งเป็นชุดคำถามที่คุณสามารถใช้ประเมินว่าสถาปัตยกรรมสอดคล้องกับแนวทางปฏิบัติของ AWS ได้ดีเพียงใด

AWS Well-Architected Framework ประกอบด้วยสี่เสาหลัก ซึ่งได้แก่ความปลอดภัย ความน่าเชื่อถือ ประสิทธิภาพการทำงาน และการเพิ่มประสิทธิภาพต้นทุน ซึ่งมีคำจำกัดความดังต่อไปนี้:

ความปลอดภัย	ความสามารถในการปกป้องข้อมูล ระบบ และสินทรัพย์ต่างๆ พร้อมกับนำเสนอประโยชน์ทางธุรกิจโดยใช้กลยุทธ์การประเมินและการลดความเสี่ยง
ความน่าเชื่อถือ	ความสามารถของระบบในการกู้คืนจากความล้มเหลวของโครงสร้างพื้นฐานหรือบริการต่างๆ ปรับเปลี่ยนการใช้งานทรัพยากรการประมวลผลให้ตรงกับความต้องการ และลดปัญหาขัดข้อง เช่น การกำหนดค่าผิดพลาดหรือปัญหาเกี่ยวกับเครือข่ายที่เกิดขึ้นชั่วคราว
ประสิทธิภาพการทำงาน	ความสามารถในการใช้ทรัพยากรการประมวลผลอย่างมีประสิทธิภาพเพื่อให้ตรงกับความต้องการของระบบ และรักษาความมีประสิทธิภาพดังกล่าวไว้ได้ดังเดิมเมื่อความต้องการเปลี่ยนแปลงไปและเมื่อเกิดการพัฒนาทางเทคโนโลยี
การเพิ่มประสิทธิภาพต้นทุน	ความสามารถในการหลีกเลี่ยงหรือขจัดค่าใช้จ่ายที่ไม่จำเป็นหรือทรัพยากรซึ่งมีคุณสมบัติต่ำกว่าระดับที่เหมาะสม

หลักการออกแบบทั่วไป

Well-Architected Framework กำหนดชุดหลักการออกแบบทั่วไปเพื่อเอื้อให้เกิดการออกแบบที่เหมาะสมในระบบคลาวด์ไว้ดังนี้

- **เลิกคาดเดาความต้องการด้านความจุ:** ไม่ต้องคาดเดาความต้องการด้านความจุของโครงสร้างพื้นฐาน เมื่อตัดสินใจด้านความจุก่อนที่จะติดตั้งใช้งานระบบ สิ่งที่คุณได้อาจเป็นทรัพยากรที่มีราคาแพงและไม่ได้ใช้งาน หรือต้องจัดการกับปัญหาประสิทธิภาพที่มาพร้อมกับความจุที่จำกัด การประมวลผลบนระบบคลาวด์ทำให้ปัญหานี้หมดไป คุณสามารถใช้ความจุได้มากหรือน้อยเท่าที่ต้องการ และปรับเพิ่มลดได้โดยอัตโนมัติ
- **ทดสอบระบบในขอบเขตการใช้งานจริง**: ในสภาพแวดล้อมแบบเดิมที่ไม่ได้อยู่ในระบบคลาวด์ การสร้างระบบสภาพแวดล้อมคู่กันเพื่อการทดสอบเพียงอย่างเดียวมักจะทำได้ยากด้วยเหตุผลด้านต้นทุน ดังนั้นสภาพแวดล้อมการทดสอบส่วนใหญ่จึงไม่ได้ทดสอบในระดับความต้องการใช้งานจริง แต่ในระบบคลาวด์คุณสามารถสร้างสภาพแวดล้อมแบบเดียวกันได้ตามต้องการ ดำเนินการทดสอบจนเสร็จสมบูรณ์ แล้วจึงค่อยปลดการใช้งานทรัพยากรเหล่านั้น เนื่องจากคุณจ่ายเงินสำหรับสภาพแวดล้อมทดสอบก็ต่อเมื่อมีการใช้งาน คุณจึงสามารถจำลองสภาพแวดล้อมการใช้งานจริงได้โดยใช้ต้นทุนเพียงส่วนเดียวสำหรับการทดสอบภายในองค์กร
- **ลดความเสี่ยงด้านการเปลี่ยนแปลงสถาปัตยกรรม:** เนื่องจากคุณสามารถใช้ระบบอัตโนมัติเพื่อสร้างสภาพแวดล้อมในการทดสอบที่เลียนแบบการกำหนดค่าการใช้งานจริง คุณจึงสามารถดำเนินการทดสอบได้ง่าย นอกจากนี้ คุณยังสามารถย้ายการซีเรียลไลซ์การทดสอบที่เกิดขึ้นในสภาพแวดล้อมภายในองค์กรที่ใช้งานออกได้ในกรณีที่ทีมต้องจัดคิวการใช้ทรัพยากรทดสอบ
- **ใช้ระบบอัตโนมัติเพื่อให้การทดลองทางสถาปัตยกรรมทำได้ง่ายขึ้น:** ระบบอัตโนมัติช่วยให้คุณสามารถสร้างและจำลองระบบได้โดยใช้ต้นทุนต่ำ (ไม่ต้องลงมือทำเอง) คุณสามารถติดตามการเปลี่ยนแปลงของระบบอัตโนมัติ ตรวจสอบผลกระทบ และแปลงกลับพารามิเตอร์ที่ใช้ไปก่อนหน้านั้นได้เมื่อจำเป็น
- **รองรับสถาปัตยกรรมเชิงวิวัฒนาการ:** ในสภาพแวดล้อมแบบดั้งเดิม การตัดสินใจเชิงสถาปัตยกรรมมักดำเนินในรูปแบบครั้งเดียวและคงที่ในลักษณะเดิม โดยมีเวอร์ชันสำคัญของระบบเพียงไม่กี่เวอร์ชันในระหว่างอายุการใช้งาน เมื่อธุรกิจและบริบททางธุรกิจเปลี่ยนแปลงอยู่ตลอดเวลา การตัดสินใจเมื่อเริ่มแรกจึงอาจเป็นอุปสรรคที่ทำให้ระบบไม่สามารถรองรับความต้องการทางธุรกิจที่เปลี่ยนแปลงไปได้ แต่ในระบบคลาวด์ ความสามารถในการใช้งานระบบอัตโนมัติและการทดสอบตามต้องการจะลดความเสี่ยงของผลกระทบที่เกิดจากการเปลี่ยนแปลงด้านการออกแบบ ซึ่งช่วยให้ระบบพัฒนาได้อย่างต่อเนื่องเพื่อให้ธุรกิจต่างๆ สามารถใช้ประโยชน์จากนวัตกรรมใหม่ๆ มาเป็นแนวทางปฏิบัติมาตรฐาน

สี่เสาหลักของ Well-Architected Framework

การสร้างระบบซอฟต์แวร์มีหลายอย่างที่เหมือนกับการสร้างอาคาร หากรากฐานไม่มั่นคง ก็อาจเกิดปัญหาทางโครงสร้างที่ทำลายความสมบูรณ์และฟังก์ชันการใช้งานของตัวอาคาร ในการออกแบบสถาปัตยกรรมของโซลูชันเทคโนโลยี หากคุณละเลยสี่เสาหลัก ซึ่งได้แก่ ความปลอดภัย ความน่าเชื่อถือ ประสิทธิภาพการทำงาน และการเพิ่มประสิทธิภาพต้นทุน ก็อาจกลายเป็นอุปสรรคต่อการสร้างระบบที่จะสนองตอบความคาดหวังและความต้องการของคุณ การประกอบเสาหลักทั้งสี่ด้านเข้ากับสถาปัตยกรรมจะช่วยคุณสร้างระบบที่มีความเสถียรและประสิทธิภาพ คุณจึงมีโอกาสที่จะมุ่งเน้นในแง่มุมอื่นๆ ของการออกแบบได้ เช่น ความต้องการด้านการใช้งาน

หัวข้อนี้จะอธิบายถึงเสาหลักสี่ด้าน และมีคำจำกัดความ แนวทางปฏิบัติ คำถาม ข้อควรพิจารณา และบริการ AWS หลักๆ ที่เกี่ยวข้อง

เสาหลักด้านความปลอดภัย

เสาหลักด้าน **ความปลอดภัย** เกี่ยวข้องกับความสามารถในการปกป้องข้อมูล ระบบ และสินทรัพย์ต่างๆ พร้อมกับนำเสนอประโยชน์ทางธุรกิจ โดยใช้กลยุทธ์การประเมินและการลดความเสี่ยง

หลักการออกแบบ

ในระบบคลาวด์นั้น มีหลักการต่างๆ มากมายที่สามารถช่วยคุณเสริมความแข็งแกร่งให้ระบบปลอดภัยยิ่งขึ้น

- **ใช้ระบบความปลอดภัยในทุกระดับ:** นอกจากการใช้งานเครื่องมือรักษาความปลอดภัย (เช่น ไฟร์วอลล์) ในขอบเขตโครงสร้างพื้นฐานแล้ว ให้ใช้ไฟร์วอลล์และการควบคุมความปลอดภัยอื่นๆ สำหรับทรัพยากรทั้งหมดของคุณด้วย (เช่น เซิร์ฟเวอร์เสมือนทุกเซิร์ฟเวอร์ โหลดบาลานเซอร์ และซับเน็ตของเครือข่าย)
- **ใช้งานความสามารถในการติดตาม:** บันทึกล็อกและตรวจสอบการดำเนินการและการเปลี่ยนแปลงทั้งหมดที่เกิดขึ้นในสภาพแวดล้อม
- **ใช้ระบบอัตโนมัติเพื่อรับมือเหตุการณ์ด้านความปลอดภัย:** ตรวจสอบและทริกเกอร์การตอบสนองต่อการแจ้งเตือนที่เปลี่ยนไปตามเหตุการณ์หรือตามเงื่อนไขโดยอัตโนมัติ
- **มุ่งเน้นการรักษาความปลอดภัยในระบบ:** คุณสามารถใช้ โมเดลความรับผิดชอบในการรักษาความปลอดภัยร่วมกันของ AWS เพื่อมุ่งเน้นการรักษาความปลอดภัยให้แอปพลิเคชัน ข้อมูล และ

ระบบปฏิบัติการ ขณะที่ AWS ที่ให้โครงสร้างพื้นฐานและบริการที่ปลอดภัย

- **ใช้ระบบอัตโนมัติควบคุมแนวทางปฏิบัติด้านความปลอดภัย:** กลไกความปลอดภัยที่ใช้ซอฟต์แวร์ช่วยให้คุณสามารถปรับขยายการใช้งานได้อย่างปลอดภัย รวดเร็ว และคุ้มค่ามากขึ้น สร้างและบันทึกอิมเมจพื้นฐานแบบกำหนดเองของเซิร์ฟเวอร์เสมือน แล้วใช้อิมเมจนั้นโดยอัตโนมัติกับเซิร์ฟเวอร์ใหม่แต่ละเครื่องที่คุณเปิดใช้งาน สร้างโครงสร้างพื้นฐานทั้งระบบตามที่กำหนดและจัดการในเทมเพลต

คำจำกัดความ

การรักษาความปลอดภัยในระบบคลาวด์มีอยู่ด้วยกันสี่ด้าน:

1. การป้องกันข้อมูล
2. การจัดการสิทธิ์การใช้งาน
3. การป้องกันโครงสร้างพื้นฐาน
4. การควบคุมเชิงตรวจสอบ

โมเดลความรับผิดชอบในการรักษาความปลอดภัยร่วมกันของ AWS ช่วยให้องค์กรที่ใช้ระบบคลาวด์บรรลุเป้าหมายด้านความปลอดภัยและการปฏิบัติตามข้อกำหนด เนื่องจาก AWS ดูแลความปลอดภัยทางกายภาพให้กับโครงสร้างพื้นฐานที่รองรับบริการในระบบคลาวด์ของเรา ลูกค้า AWS จึงหันไปมุ่งเน้นให้กับการใช้งานบริการเพื่อให้บรรลุตามเป้าหมายได้อย่างเต็มที่ นอกจากนี้ ระบบคลาวด์ของ AWS ยังช่วยให้เข้าถึงข้อมูลความปลอดภัยและวิธีการที่ใช้ระบบอัตโนมัติเพื่อตอบสนองเหตุการณ์ด้านความปลอดภัยได้มากขึ้น

แนวทางปฏิบัติ

การป้องกันข้อมูล

ก่อนที่จะออกแบบสถาปัตยกรรมให้กับระบบใดๆ ควรมีแนวทางปฏิบัติพื้นฐานเพื่อกำหนดทิศทางด้านความปลอดภัยอยู่ ตัวอย่างเช่น *การจำแนกข้อมูล* จะกำหนดวิธีจัดประเภทข้อมูลองค์กรตามระดับความสำคัญ เช่น *สิทธิ์ระดับน้อยที่สุด* จะจำกัดการเข้าถึงให้อยู่ในระดับต่ำสุดโดยที่ยังอนุญาตให้ใช้ฟังก์ชันต่างๆ ได้ตามปกติ และ *การเข้ารหัส* จะช่วยปกป้องข้อมูลด้วยวิธีประมวลผลข้อมูลที่ไม่สามารถอ่านได้ให้เป็นการเข้าถึงที่ไม่ได้รับอนุญาต เครื่องมือและเทคนิคเหล่านี้ล้วนมีความสำคัญในการสามารถรองรับเป้าหมายด้านต่างๆ ได้ เช่น การป้องกันผลขาดทุนทางการเงินหรือการปฏิบัติตามระเบียบข้อบังคับ

การป้องกันข้อมูลเป็นการใช้การควบคุมและรูปแบบที่ออกแบบมาเพื่อเก็บรักษาข้อมูลเป็นความลับพร้อมกับรักษาความสมบูรณ์ของข้อมูล และตรวจสอบถึงความพร้อมใช้งานเมื่อคุณต้องการ

ใน AWS แนวทางปฏิบัติดังต่อไปนี้จะอำนวยความสะดวกให้กับการป้องกันข้อมูล:

- ลูกค้า AWS มีสิทธิ์ควบคุมข้อมูลของตนได้เต็มที่เช่นเดิม
- AWS ช่วยให้คุณเข้ารหัสข้อมูลและจัดการคีย์ต่างๆ ได้ง่ายขึ้น รวมทั้งการหมุนเวียนคีย์ตามกำหนดการ ซึ่งลูกค้าสามารถปรับปรุงได้เองหรือใช้ระบบอัตโนมัติของ AWS ที่สะดวกต่อการใช้งาน
- ในระบบมีการบันทึกล็อกโดยละเอียดซึ่งประกอบด้วยเนื้อหาสำคัญ เช่น การเข้าถึงไฟล์และการเปลี่ยนแปลงต่างๆ
- AWS ออกแบบระบบจัดเก็บข้อมูลเพื่อให้มีความยืดหยุ่นในระดับสูงสุด ตัวอย่างเช่น Amazon Simple Storage Service (S3) ออกแบบมาให้มีความคงทนในระดับ 99.999999999% (เช่น ถ้าคุณจัดเก็บออบเจ็กต์ 10,000 รายการโดยใช้ Amazon S3 อัตราการสูญเสียหนึ่งออบเจ็กต์ที่อาจเกิดขึ้นโดยเฉลี่ยจะอยู่ที่หนึ่งครั้งทุกๆ 10,000,000 ปี)
- การกำหนดเวอร์ชันซึ่งอาจใช้เป็นส่วนหนึ่งในกระบวนการจัดการวงจรการใช้งานของข้อมูลขนาดใหญ่จะช่วยป้องกันการเขียนทับหรือการลบข้อมูลโดยไม่ตั้งใจ รวมทั้งความเสียหายในลักษณะเดียวกัน
- AWS จะไม่ใช้การเคลื่อนย้ายข้อมูลระหว่างภูมิภาค เนื้อหาที่อยู่ในภูมิภาคใดภูมิภาคหนึ่งจะยังคงอยู่ที่เดิม เว้นแต่ลูกค้าจะเปิดใช้งานคุณสมบัติหรือใช้ประโยชน์จากบริการที่มีฟังก์ชันดังกล่าวโดยชัดแจ้ง

คำถามต่อไปนี้จะมุ่งเน้นถึงข้อควรพิจารณาเกี่ยวกับการรักษาความปลอดภัยของข้อมูล (โปรดดูรายการคำถาม คำตอบ และแนวทางปฏิบัติด้านความปลอดภัยจากภาคผนวก):

SEC 1. คุณจะเข้ารหัสและป้องกันข้อมูลในที่จัดเก็บอย่างไร

SEC 2. คุณจะเข้ารหัสและป้องกันข้อมูลที่อยู่ระหว่างการรับส่งอย่างไร

AWS มีวิธีการเข้ารหัสสำหรับข้อมูลในที่จัดเก็บและข้อมูลที่อยู่ระหว่างการรับส่งอยู่หลายวิธี เราสร้างคุณสมบัติรวมไว้ในผลิตภัณฑ์และบริการต่างๆ เพื่อช่วยให้คุณเข้ารหัสข้อมูลได้ง่ายขึ้น เช่น การนำวิธีการ Server Side Encryption (SSE) มาใช้กับ Amazon S3 เพื่อช่วยให้จัดเก็บข้อมูลในรูปแบบเข้ารหัสได้ง่ายขึ้น

คุณยังสามารถกำหนดให้ Elastic Load Balancing เป็นตัวจัดการการเข้าและถอดรหัส HTTPS (โดยทั่วไปเรียกว่า

SSL Termination) ได้ทั้งกระบวนการ

การจัดการสิทธิ์การใช้งาน

การจัดการสิทธิ์การใช้งานเป็นส่วนสำคัญของโปรแกรมรักษาความปลอดภัยของข้อมูล เพื่อดูแลให้ผู้ใช้ที่ได้รับอนุญาตและผ่านการรับรองความถูกต้องเท่านั้นที่สามารถเข้าใช้ทรัพยากรต่างๆ ตามวัตถุประสงค์ที่เหมาะสม ตัวอย่างเช่น รายการควบคุมการเข้าถึง (ACL) เป็นรายการอนุญาตการเข้าถึงที่เชื่อมโยงกับออบเจ็กต์ การควบคุมการเข้าถึงตามบทบาท (RBAC) เป็นชุดสิทธิ์ที่สอดคล้องกับบทบาทหรือหน้าที่ของผู้ใช้ และการจัดการรหัสผ่านมีข้อกำหนดที่ซับซ้อนและช่วงเวลาสำหรับการเปลี่ยนรหัส องค์ประกอบการจัดการสิทธิ์การใช้งานเหล่านี้สำคัญต่อสถาปัตยกรรมด้านความปลอดภัยของข้อมูล เนื่องจากเป็นแนวคิดหลักของการอนุญาตและรับรองความถูกต้องของผู้ใช้

โดยหลักแล้ว การจัดการสิทธิ์การใช้งานใน AWS จะสนับสนุนโดยบริการ AWS Identity and Access Management (IAM) ซึ่งอนุญาตให้ลูกค้าควบคุมการเข้าถึงบริการและทรัพยากร AWS สำหรับผู้ใช้ คุณสามารถนำนโยบายที่มีรายละเอียดแยกย่อยไปใช้ได้ ซึ่งจะกำหนดสิทธิ์ให้กับผู้ใช้ กลุ่ม บทบาท หรือทรัพยากร และยังสามารถกำหนดให้ใช้แนวทางปฏิบัติเกี่ยวกับรหัสผ่านที่คาดเดายาก เช่น ความซับซ้อน การนำกลับมาใช้ใหม่ และการรับรองความถูกต้องแบบหลายปัจจัย (MFA) โดยคุณสามารถใช้เชื่อมต่อภายนอกกับบริการไดเรกทอรีที่ใช้งานอยู่ได้อีกด้วย

คำถามต่อไปนี้จะเน้นถึงข้อควรพิจารณาเกี่ยวกับการจัดการสิทธิ์การใช้งานเพื่อรักษาความปลอดภัย:

SEC 3. คุณจะป้องกันการเข้าถึงและการใช้งานข้อมูลประจำตัวบัญชีหลักของ AWS อย่างไร

SEC 4. คุณจะกำหนดบทบาทและความรับผิดชอบสำหรับผู้ใช้ในระบบเพื่อควบคุมการเข้าถึงของบุคคลใน AWS Management Console และ API อย่างไร

SEC 5. คุณจะจำกัดการเข้าถึงแบบอัตโนมัติ (เช่น จากแอปพลิเคชัน สคริปต์ และเครื่องมือหรือบริการภายนอก) ไปยังทรัพยากรของ AWS อย่างไร

SEC 6. คุณจะจัดการคีย์และข้อมูลประจำตัวอย่างไร

การดูแลข้อมูลประจำตัวของบัญชีระดับรูทให้ปลอดภัยอยู่เสมอนั้นเป็นสิ่งสำคัญ เพื่อให้บรรลุตามวัตถุประสงค์นี้ AWS แนะนำให้เชื่อมต่อ MFA กับบัญชีระดับรูทและล็อกข้อมูลประจำตัวด้วย MFA ในตำแหน่งที่ตั้งทางกายภาพที่มีการรักษาความปลอดภัย บริการ IAM ช่วยให้คุณสร้างและจัดการสิทธิ์สำหรับผู้ใช้อื่นๆ (ที่ไม่ใช่ระดับรูท) รวมทั้งกำหนดระดับการเข้าถึงทรัพยากรต่างๆ ได้

การป้องกันโครงสร้างพื้นฐาน

การป้องกันโครงสร้างพื้นฐานประกอบด้วยระเบียบวิธีด้านการควบคุมที่จำเป็นต่อการปฏิบัติตามแนวทางที่กำหนดและข้อผูกพันของอุตสาหกรรมและกฎระเบียบ เช่น การรับรองความถูกต้องแบบละเอียดและใช้หลายปัจจัย การใช้ระเบียบวิธีเหล่านี้สำคัญต่อการดำเนินงานแบบต่อเนื่องให้ประสบความสำเร็จไม่ว่าจะเป็นแบบภายในองค์กรหรือบนระบบคลาวด์

ใน AWS คุณสามารถใช้การตรวจสอบแพคเก็ตแบบเก็บสถานะและไม่เก็บสถานะได้โดยใช้เทคโนโลยีของ AWS โดยตรงหรือใช้ผลิตภัณฑ์และบริการจากบริษัทคู่ค้าที่มีให้บริการผ่าน AWS Marketplace นอกจากนี้ คุณยังสามารถใช้ Amazon Virtual Private Cloud (VPC) เพื่อสร้างสภาพแวดล้อมแบบส่วนตัว ที่มีความปลอดภัย และปรับขยายได้ เพื่อกำหนดโทโพโลยี ซึ่งได้แก่ เกตเวย์ ตารางกำหนดเส้นทาง และซับเน็ตแบบสาธารณะและ/หรือแบบส่วนตัว

คำถามต่อไปนี้จะเน้นถึงข้อควรพิจารณาเกี่ยวกับการป้องกันโครงสร้างพื้นฐานเพื่อรักษาความปลอดภัย:

SEC 7. คุณจะบังคับใช้การป้องกันในขอบเขตเครือข่ายและระดับโฮสต์อย่างไร

SEC 8. คุณจะบังคับใช้การป้องกันในระดับบริการ AWS อย่างไร

SEC 9. คุณจะปกป้องความสมบูรณ์ของระบบปฏิบัติการใน Amazon EC2 Instance อย่างไร

การป้องกันแบบหลายชั้นเป็นวิธีที่แนะนำในสภาพแวดล้อมทุกรูปแบบ และในกรณีของการป้องกันโครงสร้างพื้นฐาน ก็สามารถใช้แนวคิดและวิธีการที่หลากหลายสำหรับรูปแบบระบบคลาวด์และแบบในองค์กร การบังคับใช้การป้องกันขอบเขต การตรวจสอบจุดสื่อสารขาเข้าและขาออก และการบันทึกล็อกแบบสมบูรณ์ การตรวจสอบและแจ้งเตือนล้วนเป็นสิ่งสำคัญต่อแผนการรักษาความปลอดภัยของข้อมูลที่มีประสิทธิภาพ

ตามที่ได้กล่าวไปแล้วในหัวข้อ *หลักการออกแบบ*ข้างต้น ลูกค้า AWS สามารถปรับแต่งหรือเสริมความปลอดภัยให้กับการกำหนดค่า EC2 instance และนำการกำหนดค่านี้ไปใช้กับ Amazon Machine Image (AMI) ที่ไม่มีการเปลี่ยนแปลง หลังจากนั้น เซิร์ฟเวอร์ใหม่แบบเสมือน (อินสแตนซ์) ทั้งหมดที่เปิดใช้งานด้วย AMI นี้ก็จะได้รับการกำหนดค่าที่เข้มงวดมากขึ้นนี้ด้วย ไม่ว่าจะทริกเกอร์โดยคุณสมบัติ Auto Scaling หรือเริ่มต้นโดยผู้ใช้

การควบคุมเชิงตรวจสอบ

คุณสามารถใช้การควบคุมเชิงตรวจสอบเพื่อตรวจหาหรือระบุถึงการละเมิดความปลอดภัย การควบคุมนี้เป็นส่วนพื้นฐานของเฟรมเวิร์กการกำกับดูแล และสามารถใช้เพื่อรองรับกระบวนการด้านคุณภาพ ข้อผูกพันในการปฏิบัติตามกฎหมาย และ/หรือการดำเนินการเพื่อระบุและตอบสนองต่อภัยคุกคาม การควบคุมเชิงตรวจสอบมีอยู่ด้วยกันหลาย

แบบ ตัวอย่างเช่น สินทรัพย์รายการข้อมูลและแอททริบิวต์แบบละเอียดจะช่วยเสริมประสิทธิภาพในการตัดสินใจ (และการควบคุมวงจรการใช้งาน) เพื่อช่วยกำหนดเกณฑ์พื้นฐานในการดำเนินงาน หรือคุณสามารถใช้การตรวจสอบภายใน ซึ่งเป็นการตรวจสอบการควบคุมที่เกี่ยวข้องกับระบบข้อมูล เพื่อทำให้แน่ใจว่าแนวทางปฏิบัติต่างๆ เป็นไปตามนโยบายและข้อกำหนด และคุณได้ตั้งค่าการแจ้งเตือนอัตโนมัติอย่างถูกต้องตามเงื่อนไขที่กำหนด การควบคุมเหล่านี้เป็นปัจจัยโต้ตอบสำคัญที่ช่วยให้องค์กรสามารถระบุและเข้าใจถึงขอบเขตของกิจกรรมที่ผิดปกติในระบบ บริการที่สนับสนุนการควบคุมเชิงตรวจสอบใน AWS มีดังนี้:

- **AWS CloudTrail** – บริการบนเว็บที่บันทึกล็อกการเรียกใช้ API รวมทั้งข้อมูลเฉพาะตัวในการติดต่อ เวลาติดต่อ ที่อยู่ IP ต้นทาง พารามิเตอร์ และองค์ประกอบการตอบสนอง
- **Amazon CloudWatch** – บริการตรวจสอบทรัพยากร AWS ที่บันทึกล็อกองค์ประกอบต่างๆ เช่น CPU, ดิสก์ และกิจกรรมเครือข่ายของ Amazon Elastic Compute Cloud (EC2), อินสแตนซ์ฐานข้อมูล Amazon Relational Database Service (RDS), ไดรฟ์ข้อมูล Amazon Elastic Block Store (EBS) และองค์ประกอบอื่นๆ CloudWatch มีความสามารถในการแจ้งเตือนตามตัววัดผลเหล่านี้และอื่นๆ
- **AWS Config** – บริการประวัติการกำหนดค่าและรายการข้อมูลที่ให้ข้อมูลเกี่ยวกับการกำหนดค่าและการเปลี่ยนแปลงโครงสร้างพื้นฐานในช่วงเวลาต่างๆ
- **Amazon Simple Storage Service (S3)**– ลูกค้าสามารถใช้การตรวจสอบการเข้าถึงข้อมูลของ Amazon S3 ในการกำหนดค่าบัคเก็ต Amazon S3 เพื่อบันทึกรายละเอียดคำขอเข้าถึง ซึ่งได้แก่ ประเภท ทรัพยากร วันที่และเวลา
- **Amazon Glacier**– ลูกค้าสามารถใช้คุณสมบัติการล็อคชุดเก็บข้อมูลเพื่อรักษาข้อมูลที่สำคัญต่อการดำเนินงานด้วยตัวควบคุมการปฏิบัติตามข้อกำหนดที่ออกแบบมาให้รองรับการเก็บรักษาข้อมูลในระยะยาว

คำถามต่อไปนี้จะเน้นถึงข้อควรพิจารณาเกี่ยวกับตัวควบคุมเชิงตรวจสอบเพื่อรักษาความปลอดภัย:

SEC 10. คุณจะเก็บข้อมูลและวิเคราะห์ล็อก AWS อย่างไร

การจัดการล็อกสำคัญต่อการออกแบบตามหลักสถาปัตยกรรมที่เหมาะสมทั้งในด้านการวิเคราะห์/การรักษาความปลอดภัยไปจนถึงข้อกำหนดด้านกฎระเบียบและกฎหมาย AWS มีฟังก์ชันที่ช่วยให้ทำการจัดการล็อกได้ง่ายขึ้น โดยให้ผู้ใช้สามารถกำหนดวงจรสำหรับการเก็บรักษาข้อมูล หรือกำหนดที่ตั้งในการเก็บรักษา แยกเก็บถาวร และ/หรือลบออกในท้ายที่สุด ซึ่งช่วยลดความซับซ้อนในการจัดการข้อมูลที่คาดการณ์และเชื่อถือได้ รวมทั้งมีความคุ้มค่า

บริการ AWS ที่สำคัญ

บริการ AWS ที่สำคัญต่อการรักษาความปลอดภัยได้แก่ AWS Identity and Access Management (IAM) ซึ่งช่วยควบคุมการเข้าถึงบริการและทรัพยากร AWS ของผู้ใช้ได้อย่างปลอดภัย บริการและคุณสมบัติต่อไปนี้รองรับขอบเขตความปลอดภัยสี่ด้าน:

การป้องกันข้อมูล: บริการอย่างเช่น Elastic Load Balancing, Amazon Elastic Block Store (EBS), Amazon Simple Storage Service (S3) และ Amazon Relational Database Service (RDS) มาพร้อมกับความสามารถในการเข้ารหัสเพื่อป้องกันข้อมูลระหว่างการรับส่งและเมื่ออยู่ในที่จัดเก็บ AWS Key Management Service (KMS) ช่วยให้ลูกค้าสร้างและควบคุมคีย์ที่ใช้ในการเข้ารหัสได้ง่ายขึ้น

การจัดการสิทธิ์การใช้งาน: IAM ช่วยให้คุณควบคุมการเข้าถึงบริการและทรัพยากร AWS ได้อย่างปลอดภัย การรับรองความถูกต้องโดยใช้หลายปัจจัย (MFA) ช่วยเพิ่มการป้องกันในระดับพิเศษที่นอกเหนือจากชื่อผู้ใช้และรหัสผ่าน

การป้องกันโครงสร้างพื้นฐาน: Amazon Virtual Private Cloud (VPC) ช่วยคุณเตรียมใช้งานเซกชันแบบแยกที่เป็นส่วนตัวในระบบคลาวด์ของ AWS ซึ่งคุณสามารถเปิดใช้ทรัพยากร AWS ในเครือข่ายเสมือน

การควบคุมเชิงตรวจสอบ: AWS CloudTrail บันทึกการเรียกใช้ AWS API ส่วน AWS Config จะระบุรายการข้อมูลโดยละเอียดของทรัพยากรและการกำหนดค่า AWS และ Amazon CloudWatch จะตรวจสอบบริการสำหรับทรัพยากรของ AWS

แหล่งข้อมูล

โปรดดูแหล่งข้อมูลต่อไปนี้เพื่อศึกษาเพิ่มเติมเกี่ยวกับแนวทางปฏิบัติสำหรับการรักษาความปลอดภัยของเรา

เอกสารประกอบและบล็อก

- ศูนย์ความปลอดภัยของ AWS
- การปฏิบัติตามข้อกำหนดของ AWS
- บล็อกความปลอดภัยของ AWS

รายงาน

- ภาพรวมการรักษาความปลอดภัยของ AWS
- แนวทางปฏิบัติด้านความปลอดภัยของ AWS
- ความเสี่ยงและปฏิบัติตามข้อกำหนดของ AWS

วิดีโอ

- ระบบความปลอดภัยของ AWS Cloud
- ภาพรวมเกี่ยวกับหน้าที่รับผิดชอบร่วมกัน

เสาหลักด้านความน่าเชื่อถือ

เสาหลักด้าน **ความน่าเชื่อถือ** เป็นความสามารถของระบบในการกู้คืนจากข้อขัดข้องของโครงสร้างพื้นฐานหรือบริการ ปรับเปลี่ยนการใช้ทรัพยากรการประมวลผลเพื่อให้บรรลุตามความต้องการและลดปัญหาขัดข้องต่างๆ เช่น การกำหนดค่าผิดพลาดหรือปัญหาชั่วคราวของเครือข่าย

หลักการออกแบบ

ระบบคลาวด์มีหลักการต่างๆ มากมายที่สามารถช่วยคุณเพิ่มความน่าเชื่อถือของระบบ:

- **ทดสอบขั้นตอนการกู้คืน:** ในสภาพแวดล้อมภายในองค์กร การทดสอบมักดำเนินการเพื่อพิสูจน์ว่าระบบทำงานได้ในสถานการณ์เฉพาะ โดยทั่วไปแล้ว การทดสอบนี้ไม่ได้ใช้เพื่อตรวจสอบความถูกต้องของกลยุทธ์การกู้คืน แต่ในระบบคลาวด์ คุณสามารถทดสอบว่าระบบล้มเหลวได้อย่างไร และตรวจสอบความถูกต้องของขั้นตอนการกู้คืนได้ คุณสามารถใช้ระบบอัตโนมัติเพื่อจำลองความล้มเหลวต่างๆ หรือสร้างสถานการณ์จำลองที่เป็นสาเหตุของความล้มเหลวก่อนหน้านั้นขึ้นมาอีกครั้งได้ การทดสอบนี้จะแสดงเส้นทางที่นำไปสู่ความล้มเหลว ซึ่งคุณสามารถทดสอบและแก้ไข*ก่อน*ที่เหตุการณ์ดังกล่าวจะเกิดขึ้นจริง จึงช่วยลดความเสี่ยงที่จะเกิดความล้มเหลวขึ้นกับคอมโพเนนต์ที่ยังไม่เคยทดสอบมาก่อน
- **กู้คืนระบบจากความล้มเหลวโดยอัตโนมัติ:** คุณสามารถใช้วิธีติดตามตัวบ่งชี้ประสิทธิภาพหลัก (KPI) ของระบบเพื่อทริกเกอร์ระบบอัตโนมัติได้เมื่อค่าของตัวบ่งชี้เกินเกณฑ์ขั้นต่ำ วิธีนี้จะช่วยให้ระบบแจ้งข้อมูลและติดตามความล้มเหลวโดยอัตโนมัติและดำเนินกระบวนการกู้คืนเพื่อซ่อมแซมและแก้ไขความล้มเหลวดังกล่าว ระบบอัตโนมัติที่ความทันสมัยมากขึ้นจะช่วยให้คาดคะเนและแก้ไขความล้มเหลวได้ก่อนที่จะเกิดขึ้นจริง

- **ปรับเพิ่มลดขนาดเพื่อเพิ่มความพร้อมใช้งานโดยรวมของระบบ:** แทนที่ทรัพยากรเดียวขนาดใหญ่ด้วยหลายๆ ทรัพยากรขนาดเล็กเพื่อลดผลกระทบจากความล้มเหลวในระบบโดยรวม กระจายคำขอไปยังหลายๆ ทรัพยากรที่มีขนาดเล็กลงเพื่อให้มั่นใจได้ว่าทรัพยากรเหล่านั้นไม่มีจุดที่ผิดพลาดในลักษณะเดียวกัน
- **เลิกคาดเดาเรื่องความจุ:** สาเหตุหนึ่งของความล้มเหลวที่พบบ่อยในระบบแบบในสถานที่คือ การอิ่มตัวของทรัพยากร เมื่อมีความต้องการใช้งานเพิ่มขึ้นในระบบจนเกินความจุ (ซึ่งมักเป็นเป้าหมายในการโจมตีเพื่อให้ระบบหยุดการทำงาน) แต่ในระบบคลาวด์ คุณสามารถตรวจสอบความต้องการและการใช้งานระบบ รวมทั้งเพิ่มหรือย้ายทรัพยากรออกได้โดยอัตโนมัติเพื่อรักษาขนาดให้อยู่ในระดับที่เหมาะสมตามความต้องการ โดยไม่มีการจัดเตรียมใช้งานที่มากหรือน้อยเกินไป

คำจำกัดความ

ความน่าเชื่อถือในระบบคลาวด์ประกอบด้วยสามส่วนดังนี้:

1. รากฐาน
2. การจัดการการเปลี่ยนแปลง
3. การจัดการความล้มเหลว

เพื่อให้บรรลุเป้าหมายด้านความน่าเชื่อถือ ระบบต้องมีรากฐานที่วางแผนไว้อย่างเหมาะสมและการตรวจติดตามอยู่ตลอดเวลาด้วยกลไกสำหรับจัดการความเปลี่ยนแปลงด้านความต้องการหรือข้อกำหนด ระบบควรได้รับการออกแบบให้ตรวจหาความล้มเหลวและแก้ไขได้เองโดยอัตโนมัติ

แนวทางปฏิบัติ

รากฐาน

ก่อนการออกแบบสถาปัตยกรรมระบบใดๆ ควรมีการจัดเตรียมข้อกำหนดด้านรากฐานที่ส่งผลต่อความน่าเชื่อถือให้พร้อม เช่น คุณต้องมีแบนด์วิดท์เครือข่ายที่เพียงพอกับศูนย์ข้อมูล บางครั้งข้อกำหนดเหล่านี้ก็ถูกละเลย (เนื่องจากอยู่นอกเหนือขอบเขตของโครงการเดียว) การละเลยในเรื่องนี้อาจส่งผลกระทบที่สำคัญต่อความสามารถในการให้บริการระบบที่น่าเชื่อถือ ในสภาพแวดล้อมแบบภายในองค์กร ข้อกำหนดเหล่านี้อาจใช้เวลาดำเนินการยาวนานเนื่องจากต้องอ้างอิงกับหลายๆ ส่วนและรวมไว้ในระบบในระหว่างการวางแผนเบื้องต้น

เมื่อใช้ AWS ข้อกำหนดด้านรากฐานเหล่านี้จะรวมไว้ในระบบอยู่แล้วหรือสามารถระบุได้ตามที่จำเป็น ระบบคลาวด์ออกแบบมาให้ก้าวข้ามขีดจำกัดต่างๆ ดังนั้น AWS จึงมีหน้าที่ในการรองรับความต้องการด้านระบบเครือข่ายและความจุของระบบประมวลผลให้เพียงพอ ขณะที่คุณสามารถเปลี่ยนขนาดและการจัดสรรทรัพยากร เช่น ขนาดของอุปกรณ์จัดเก็บข้อมูลได้อย่างอิสระตามความต้องการ

คำถามต่อไปนี้จะเน้นถึงข้อควรพิจารณาด้านรากฐานเพื่อความน่าเชื่อถือของระบบ (โปรดดูรายการคำถาม คำตอบ และแนวทางปฏิบัติทั้งหมดเกี่ยวกับความน่าเชื่อถือจากภาคผนวก):

REL 1. คุณจะจัดการกับค่าจำกัดของบริการ AWS สำหรับบัญชีอย่างไร

REL 2. คุณจะวางแผนโทโพโลยีเครือข่ายบน AWS อย่างไร

REL 3. คุณมีพาธการเลื่อนระดับเพื่อรองรับปัญหาทางเทคนิคหรือไม่

AWS กำหนดค่าจำกัดของบริการ (ค่าจำกัดสูงสุดเกี่ยวกับจำนวนทรัพยากรแต่ละอย่างที่ทีมของคุณสามารถขอได้) เพื่อไม่ให้คุณต้องจัดเตรียมทรัพยากรที่มากเกินไปโดยไม่ตั้งใจ คุณจะต้องมีกระบวนการและการกำกับดูแลเพื่อตรวจสอบและเปลี่ยนแปลงค่าจำกัดเหล่านี้เพื่อให้ตรงกับความต้องการทางธุรกิจ เมื่อคุณใช้งานระบบคลาวด์ คุณอาจต้องวางแผนการรวมระบบเข้ากับทรัพยากรแบบภายในองค์กรที่ใช้งานอยู่ (แนวทางแบบผสมผสาน) รูปแบบผสมผสานจะเอื้อต่อการเปลี่ยนผ่านอย่างค่อยเป็นค่อยไปสู่แนวทางระบบคลาวด์เต็มรูปแบบเมื่อเวลาผ่านไปดังนั้น การออกแบบลักษณะการทำงานของ AWS และทรัพยากรแบบภายในองค์กรในฐานะของโทโพโลยีเครือข่ายจึงเป็นสิ่งสำคัญ ประการสุดท้าย คุณอาจต้องตรวจสอบว่าทีมงานฝ่ายเทคโนโลยีสารสนเทศได้รับการฝึกอบรมและมีกระบวนการที่เป็นปัจจุบันเพื่อรองรับการใช้งานระบบคลาวด์แบบสาธารณะ และดูว่าคุณมีข้อตกลงร่วมกับคู่ค้าหรือข้อตกลงเกี่ยวกับการสนับสนุนอยู่พร้อมตามความเหมาะสม

การจัดการการเปลี่ยนแปลง

การตระหนักถึงผลกระทบจากการเปลี่ยนแปลงที่มีต่อระบบจะช่วยคุณในการวางแผนเชิงรุก และการตรวจสอบจะระบุแนวโน้มที่อาจก่อให้เกิดปัญหาด้านความจุหรือการละเมิด SLA ได้อย่างรวดเร็ว ในสภาพแวดล้อมแบบดั้งเดิม กระบวนการควบคุมการเปลี่ยนแปลงมักดำเนินการโดยผู้ใช้และต้องมีการประสานกับการตรวจสอบอย่างระมัดระวัง เพื่อให้ควบคุมบุคคลและเวลาที่ทำการเปลี่ยนแปลงได้อย่างมีประสิทธิภาพ

เมื่อใช้ AWS คุณสามารถตรวจสอบลักษณะการทำงานของระบบและใช้ระบบอัตโนมัติในการตอบสนองต่อ KPI เช่น การเพิ่มเซิร์ฟเวอร์อื่นๆ เมื่อระบบมีผู้ใช้เพิ่มเติม คุณสามารถควบคุมบุคคลที่มีสิทธิ์เปลี่ยนแปลงข้อมูลในระบบและตรวจสอบประวัติการเปลี่ยนแปลงเหล่านั้นได้

คำถามต่อไปนี้จะเน้นถึงข้อควรพิจารณาเกี่ยวกับการจัดการการเปลี่ยนแปลงเพื่อความน่าเชื่อถือ:

REL 4. ระบบของคุณมีวิธีรองรับความต้องการใช้งานที่เปลี่ยนแปลงไปอย่างไร

REL 5. คุณจะตรวจสอบทรัพยากร AWS ได้อย่างไร

REL 6. คุณจะจัดการกับการเปลี่ยนแปลงอย่างไร

เมื่อคุณออกแบบระบบให้เพิ่มและย้ายทรัพยากรออกโดยอัตโนมัติเพื่อตอบสนองต่อความต้องการใช้งานที่เปลี่ยนแปลงไป การดำเนินการนี้ไม่เพียงแต่เพิ่มความน่าเชื่อถือเท่านั้น แต่ยังเป็นการดูแลให้เส้นทางสู่ความสำเร็จของธุรกิจเป็นไปอย่างราบรื่น การตรวจสอบที่พร้อมใช้งานช่วยให้ทีมของคุณได้รับการแจ้งเตือนโดยอัตโนมัติเมื่อ KPI เบี่ยงเบนไปจากค่าบรรทัดฐานที่คาดไว้ การบันทึกล็อกการเปลี่ยนแปลงด้านสภาพแวดล้อมโดยอัตโนมัติจะช่วยให้สามารถตรวจสอบและระบุการดำเนินการที่อาจส่งผลต่อความน่าเชื่อถือได้อย่างรวดเร็ว ตัวควบคุมด้านการจัดการการเปลี่ยนแปลงช่วยคุณบังคับใช้กฎต่างๆ เพื่อทำให้เกิดความน่าเชื่อถือในแบบที่คุณต้องการ

การจัดการความล้มเหลว

ในทุกระบบที่มีความซับซ้อนอย่างสมเหตุสมผล ความล้มเหลวต่างๆ มีโอกาสเกิดขึ้นได้ การทราบถึงวิธีตรวจหาความล้มเหลว รับมือ และป้องกันไม่ให้เกิดขึ้นอีกจะเป็นประโยชน์กับคุณ

ใน AWS เราสามารถใช้ประโยชน์จากระบบอัตโนมัติเพื่อตอบสนองต่อการตรวจติดตามข้อมูล เช่น เมื่อตัววัดผลตัวใดตัวหนึ่งมีค่าเกินเกณฑ์ขั้นต่ำ คุณสามารถทริกเกอร์การดำเนินการอัตโนมัติเพื่อแก้ไขปัญหาดังกล่าวได้ และแทนที่จะพยายามวิเคราะห์และแก้ไขทรัพยากรที่มีข้อผิดพลาดที่อยู่ในสภาพแวดล้อมการใช้งานจริง คุณยังสามารถ

แทนที่ทรัพยากรเดิมด้วยทรัพยากรใหม่และดำเนินการวิเคราะห์ทรัพยากรที่มีข้อผิดพลาดแยกต่างหากได้ เนื่องจากระบบคลาวด์ช่วยให้คุณกำหนดเวอร์ชันชั่วคราวของทั้งระบบได้โดยมีค่าใช้จ่ายต่ำ คุณจึงสามารถดำเนินการทดสอบโดยอัตโนมัติเพื่อตรวจสอบกระบวนการกู้คืนทั้งหมดได้

คำถามต่อไปนี้จะเน้นถึงข้อควรพิจารณาเกี่ยวกับการจัดการความล้มเหลวเพื่อความน่าเชื่อถือของระบบ:

REL 7. คุณจะสำรองข้อมูลได้อย่างไร

REL 8. ระบบของคุณจะรับมือกับความล้มเหลวที่เกิดขึ้นกับคอมโพเนนต์อย่างไร

REL 9. คุณจะวางแผนการกู้คืนระบบอย่างไร

สำรองข้อมูลของคุณเป็นประจำและทดสอบไฟล์ข้อมูลสำรอง เพื่อให้มั่นใจว่าคุณสามารถกู้คืนจากข้อผิดพลาดทั้งในแบบลอจิคัลและกายภาพ หลักสำคัญในการจัดการกับความล้มเหลวคือ การใช้ระบบอัตโนมัติเพื่อทดสอบความล้มเหลวในระบบและตลอดจนการกู้คืนเป็นประจำ (ควรดำเนินการตามกำหนดการปกติและทริกเกอร์หลังจากที่มีการเปลี่ยนแปลงสำคัญในระบบ) ติดตาม KPI อย่างต่อเนื่อง เช่น เวลากู้คืนระบบที่ยอมรับได้ (RTO) และระยะเวลาสูงสุดที่ยอมให้ข้อมูลเสียหาย (RPO) เพื่อประเมินความสมบูรณ์ของระบบ (โดยเฉพาะอย่างยิ่งในสภาวะการทดสอบความล้มเหลว) และเพื่อช่วยให้สามารถระบุและลดจุดที่เกิดความล้มเหลว เป้าหมายคือการทดสอบกระบวนการกู้คืนระบบโดยละเอียด เพื่อให้แน่ใจได้ว่าคุณจะสามารถกู้คืนข้อมูลทั้งหมดและให้บริการลูกค้าต่อไปได้แม้ในขณะที่ปัญหายังคงดำเนินอยู่ กระบวนการกู้คืนควรดำเนินการในลักษณะเดียวกับกระบวนการใช้งานตามปกติ

บริการ AWS ที่สำคัญ

บริการ AWS ที่สำคัญต่อการตรวจสอบความน่าเชื่อถือของระบบ คือ Amazon CloudWatch ซึ่งจะตรวจสอบตัววัดผลขณะทำงาน บริการและคุณสมบัติอื่นๆ ที่รองรับความน่าเชื่อถือในสามส่วนมีดังต่อไปนี้

รากฐาน: AWS Identity and Access Management (IAM) ช่วยให้คุณควบคุมการเข้าถึงบริการและทรัพยากร AWS ได้อย่างปลอดภัย Amazon VPC จะช่วยคุณเตรียมใช้งานเซกชันแบบแยกส่วนที่เป็นส่วนตัวในระบบคลาวด์ของ AWS ที่คุณสามารถเปิดใช้ทรัพยากร AWS ในเครือข่ายเสมือน

การจัดการการเปลี่ยนแปลง: AWS CloudTrail บันทึกการเรียก AWS API สำหรับบัญชีและส่งมอบล็อกไฟล์เพื่อให้คุณใช้ในการตรวจสอบ AWS Config จะแสดงรายการข้อมูลโดยละเอียดของทรัพยากรและการกำหนดค่า AWS และบันทึกการเปลี่ยนแปลงการกำหนดค่าอย่างต่อเนื่อง

การจัดการความล้มเหลว: AWS CloudFormation สามารถใช้เพื่อสร้างเทมเพลตทรัพยากร AWS และเตรียมการใช้งานในลักษณะที่เป็นลำดับและคาดการณ์ได้

แหล่งข้อมูล

โปรดดูแหล่งข้อมูลต่อไปนี้เพื่อศึกษาเพิ่มเติมเกี่ยวกับแนวทางปฏิบัติที่เกี่ยวข้องกับความน่าเชื่อถือ

วิดีโอและรายงานจากนักวิเคราะห์

- การเตรียมพร้อมรับความล้มเหลว: การสร้างข้อบกพร่องและความน่าเชื่อถือของบริการ
- การเทียบวัดมาตรฐานความพร้อมใช้งานและความน่าเชื่อในระบบคลาวด์

เอกสารประกอบและบล็อก

- เอกสารประกอบเกี่ยวกับค่าจำกัดของบริการ
- โพสต์ในบล็อกรายงานเกี่ยวกับค่าจำกัดของบริการ

รายงาน

- รายงานแนวทางการเก็บถาวรและคืนค่าข้อมูลสำรองโดยใช้ AWS
- รายงานการจัดการโครงสร้างพื้นฐาน AWS ตามขนาดที่เหมาะสม
- รายงานการกู้คืนจากความเสียหายของ AWS
- รายงานตัวเลือกการเชื่อมต่อ Amazon VPC ของ AWS

การสนับสนุนของ AWS

- AWS Premium Support
- Trusted Advisor

เสาหลักด้านประสิทธิภาพการทำงาน

เสาหลักด้าน **ประสิทธิภาพการทำงาน** จะมุ่งเน้นการใช้ทรัพยากรการประมวลผลอย่างมีประสิทธิภาพเพื่อให้บรรลุตามความต้องการ และรักษาความมีประสิทธิภาพดังกล่าวไว้ได้ดังเดิม เมื่อความต้องการเปลี่ยนแปลงไปและเมื่อเกิดการพัฒนาทางเทคโนโลยี

หลักการออกแบบ

ในระบบคลาวด์มีหลักการต่างๆ มากมายที่สามารถช่วยให้คุณบรรลุเป้าหมายด้านประสิทธิภาพการทำงาน:

- **เปิดให้ใช้งานเทคโนโลยีขั้นสูงอย่างทั่วถึง:** เทคโนโลยีที่นำไปใช้ได้ยากอาจใช้งานได้สะดวกขึ้นด้วยการรวบรวมองค์ความรู้และข้อมูลที่ซับซ้อนเหล่านั้นไว้ในโดเมนของผู้จัดจำหน่ายระบบคลาวด์ แทนที่จะให้ทีมงานฝ่ายเทคโนโลยีสารสนเทศศึกษาวิธีโฮสต์และเรียกใช้เทคโนโลยีใหม่ๆ คุณสามารถใช้งานเทคโนโลยีในรูปแบบของบริการแทนได้ เช่น ฐานข้อมูล NoSQL, การแปลงรหัสสื่อ และการเรียนรู้กลไกล้วนเป็นเทคโนโลยีที่ต้องอาศัยความเชี่ยวชาญและไม่ได้มีการกระจายไปยังคอมมูนิตีทางเทคนิคได้อย่างทั่วถึง แต่ในระบบคลาวด์ เทคโนโลยีเหล่านี้มีอยู่ในรูปแบบบริการที่ทีมงานของคุณสามารถใช้งานได้ควบคู่กับการมุ่งเน้น

การพัฒนาผลิตภัณฑ์มากกว่าการเตรียมใช้งานและการจัดการทรัพยากร

- **รองรับการใช้งานทั่วโลกโดยใช้เวลาน้อย:** ติดตั้งใช้งานระบบของคุณได้ง่ายๆ ในหลายภูมิภาคทั่วโลกในไม่กี่คลิก ซึ่งทำให้เสียเวลาน้อยลง คุณจึงมอบประสบการณ์ที่ดียิ่งขึ้นให้กับลูกค้าได้อย่างเรียบง่ายและเสียค่าใช้จ่ายน้อยที่สุด
- **ใช้สถาปัตยกรรม*แบบไม่มีเซิร์ฟเวอร์*:** สถาปัตยกรรมแบบไม่มีเซิร์ฟเวอร์ในระบบคลาวด์ทำให้คุณไม่จำเป็นต้องเรียกใช้และซ่อมบำรุงเซิร์ฟเวอร์เพื่อดำเนินการประมวลผลในแบบเดิม เช่น บริการจัดเก็บข้อมูลสามารถทำหน้าที่เป็นเว็บไซต์แบบคงที่ ทำให้ไม่จำเป็นต้องมีเว็บเซิร์ฟเวอร์ และบริการเหตุการณ์สามารถโฮสต์โค้ดต่างๆ ให้กับคุณได้ สถาปัตยกรรมในรูปแบบนี้ไม่เพียงแต่ลดภาระการดำเนินงานจากการจัดการเซิร์ฟเวอร์เท่านั้น แต่ยังช่วยลดต้นทุนการทำธุรกรรม เนื่องจากบริการที่ได้รับการจัดการเหล่านี้ดำเนินงานในระดับคลาวด์
- **ทดลองได้บ่อยขึ้น:** ทรัพยากรแบบเสมือนและรองรับระบบอัตโนมัติช่วยให้คุณดำเนินการทดสอบเชิงเปรียบเทียบได้อย่างรวดเร็วโดยใช้อินสแตนซ์ ที่เก็บข้อมูล และการกำหนดค่าประเภทต่างๆ กัน

คำจำกัดความ

ประสิทธิภาพการทำงานในระบบคลาวด์ประกอบด้วยสี่ส่วนดังนี้:

1. การประมวลผล
2. การจัดเก็บข้อมูล
3. ฐานข้อมูล
4. การแลกกันของพื้นที่กับเวลา

ข้อควรพิจารณาเกี่ยวกับขอบเขตแต่ละด้านเหล่านี้ประกอบด้วย 1) วิธีเลือกแนวทางและทรัพยากรที่เหมาะสมที่สุด 2) วิธีปรับปรุงแนวทางที่ใช้ให้สอดคล้องกับความสามารถของระบบคลาวด์ที่พัฒนาอย่างต่อเนื่อง 3) วิธีตรวจสอบประสิทธิภาพขณะทำงานเทียบกับที่คาดไว้ และ 4) วิธีปรับขนาดทรัพยากรตามความต้องการใช้งาน

แนวทางปฏิบัติ

การประมวลผล

โครงสร้างเซิร์ฟเวอร์ที่เหมาะสมที่สุดกับสถาปัตยกรรมเฉพาะอาจแตกต่างกันไปตามการออกแบบแอปพลิเคชัน

รูปแบบการใช้งาน และการตั้งค่าโครงสร้าง หลายๆ ระบบใช้โครงสร้างเซิร์ฟเวอร์ที่ต่างกันสำหรับคอมโพเนนต์ต่างๆ และใช้คุณสมบัติที่แตกต่างเพื่อปรับปรุงประสิทธิภาพ การเลือกโครงสร้างเซิร์ฟเวอร์ที่ไม่เหมาะกับรูปแบบการใช้งานอาจเป็นสาเหตุให้ประสิทธิภาพการทำงานลดลง

ใน AWS เซิร์ฟเวอร์ได้รับการจำลองเสมือน คุณจึงสามารถเปลี่ยนแปลงความสามารถต่างๆ ได้ง่ายเพียงคลิกปุ่มหรือเรียก API เนื่องจากไม่มีข้อจำกัดตายตัวในการตัดสินใจเรื่องทรัพยากรอีกต่อไป คุณจึงสามารถทดลองกับประเภทเซิร์ฟเวอร์ที่ต่างกันได้ ใน AWS *อินสแตนซ์*ของเซิร์ฟเวอร์เสมือนเหล่านี้มีขนาดและกลุ่มผลิตภัณฑ์ที่แตกต่างกัน จึงให้ความสามารถที่หลากหลาย เช่น SSD และ GPU และคุณยังสามารถดำเนินการประมวลผลแบบไม่ใช้เซิร์ฟเวอร์ได้อีกด้วย เช่น AWS Lambda ให้คุณสามารถเรียกใช้โค้ดได้โดยไม่ต้องเรียกใช้การทำงานอินสแตนซ์

คำถามตัวอย่างต่อไปนี้จะเน้นถึงข้อควรพิจารณาเกี่ยวกับระบบประมวลผล (โปรดดูรายการคำถาม คำตอบ และแนวทางปฏิบัติทั้งหมดเกี่ยวกับประสิทธิภาพการทำงานจากภาคผนวก):

PERF 1. คุณจะเลือกประเภทอินสแตนซ์ที่เหมาะกับระบบของคุณอย่างไร

PERF 2. คุณมีวิธีการอย่างไรในการตรวจสอบว่าคุณใช้ประเภทอินสแตนซ์ที่เหมาะสมที่สุดอยู่ขณะที่มีประเภทและคุณสมบัติใหม่ๆ ของอินสแตนซ์ออกมาให้บริการ

PERF 3. คุณมีวิธีการอย่างไรในการตรวจสอบอินสแตนซ์หลังจากเปิดใช้งานเพื่อให้แน่ใจว่าอินสแตนซ์ทำงานได้ตามที่คาดหวัง

PERF 4. คุณมีวิธีการอย่างไรในการตรวจสอบว่าจำนวนอินสแตนซ์เป็นไปตามความต้องการ

เมื่อเลือกประเภทอินสแตนซ์ที่จะใช้ คุณต้องมีข้อมูลสำหรับทดสอบที่แสดงประเภทอินสแตนซ์ (หรือวิธีการแบบไม่ใช้เซิร์ฟเวอร์) ที่เหมาะกับเวิร์กโหลด*นั้นๆ* มากที่สุด การทดสอบเหล่านี้ต้องสามารถทำซ้ำได้ (ควรเป็นส่วนหนึ่งของไปป์ไลน์การส่งมอบที่ต่อเนื่อง (CD)) เพื่อให้คุณทดสอบประเภทหรือความสามารถใหม่ๆ ของอินสแตนซ์ได้ง่ายเมื่อประเภทหรือความสามารถเหล่านั้นพร้อมใช้งาน ในแง่ของการดำเนินงาน คุณควรจัดเตรียมการตรวจสอบให้พร้อมสำหรับการแจ้งเตือนคุณเมื่อประสิทธิภาพการทำงานเสื่อมถอยลง

การจัดเก็บข้อมูล

โซลูชันการจัดเก็บข้อมูลที่เหมาะสมที่สุดกับระบบใดระบบหนึ่งจะแตกต่างกันไปตามวิธีการเข้าถึง (บล็อค ไฟล์ หรือออบเจ็กต์) รูปแบบการเข้าถึง (แบบสุ่มหรือแบบตามลำดับ) ข้อมูลประมวลผลที่ต้องการ ความถี่ในการเข้าถึง (ออนไลน์ ออฟไลน์ หรือเก็บถาวร) ความถี่ในการอัปเดต (แบบบันทึกครั้งเดียวอ่านหลายครั้ง แบบไดนามิก) และข้อจำกัดด้านความพร้อมใช้งานและความคงทน ระบบสถาปัตยกรรมที่เหมาะสมจะใช้โซลูชันการจัดเก็บข้อมูลหลาย

โซลูชันและใช้งานคุณสมบัติที่แตกต่างกันเพื่อปรับปรุงประสิทธิภาพให้ดียิ่งขึ้น

พื้นที่จัดเก็บข้อมูลใน AWS เป็นแบบการจำลองเสมือนและพร้อมให้ใช้งานในหลากหลายประเภท ทำให้สามารถเลือกวิธีจัดเก็บที่ตรงกับความต้องการของคุณได้ง่ายขึ้นและยังมีตัวเลือกการจัดเก็บที่หาได้ยากในโครงสร้างพื้นฐานแบบภายในองค์กร เช่น Amazon S3 ที่ได้รับการออกแบบให้มีความคงทนในระดับ 99.999999999% คุณยังสามารถเปลี่ยนจากฮาร์ดไดรฟ์แม่เหล็ก (HDD) มาใช้โซลิดสเทตไดร์ฟ (SSD) ได้ และย้ายไดรฟ์เสมือนจากอินสแตนซ์ไปยังอีกอินสแตนซ์หนึ่งได้ง่ายภายในไม่กี่วินาที

คำถามตัวอย่างต่อไปนี้จะเน้นที่ข้อควรพิจารณาเกี่ยวกับการจัดเก็บข้อมูลเพื่อประสิทธิภาพการทำงาน:

PERF 5. คุณจะเลือกโซลูชันการจัดเก็บข้อมูลที่เหมาะกับระบบของคุณอย่างไร

PERF 6. คุณมีวิธีการอย่างไรในการตรวจสอบว่าคุณใช้โซลูชันการจัดเก็บข้อมูลที่เหมาะสมที่สุดอยู่ขณะที่มีโซลูชันและคุณสมบัติใหม่ๆ ในการจัดเก็บข้อมูลออกมาให้บริการ

PERF 7. คุณมีวิธีการอย่างไรในการตรวจสอบโซลูชันการจัดเก็บข้อมูลเพื่อให้แน่ใจว่าโซลูชันนั้นทำงานได้ตามที่คาดหวัง

PERF 8. คุณมีวิธีการอย่างไรในการตรวจสอบว่าความจุและอัตราความเร็วของโซลูชันการจัดเก็บข้อมูลนั้นตรงกับความต้องการ

เมื่อเลือกโซลูชันการจัดเก็บข้อมูล คุณควรมีข้อมูลสำหรับทดสอบที่แสดงโซลูชันการจัดเก็บข้อมูล ซึ่งสามารถนำเสนอขอบเขตต้นทุน/ผลที่ได้ตามที่จำเป็นสำหรับเวิร์กโหลด*นั้นๆ* การทดสอบเหล่านี้ต้องสามารถทำซ้ำได้ (ควรเป็นส่วนหนึ่งของไปป์ไลน์ CD) เพื่อให้คุณทดสอบโซลูชันหรือความสามารถใหม่ๆ ในการจัดเก็บข้อมูลได้ง่ายเมื่อมีให้บริการ ประเภทของอุปกรณ์เก็บข้อมูล (EBS เทียบกับการจัดเก็บข้อมูลของอินสแตนซ์ หรือ HDD เทียบกับ SSD) ที่ใช้สำหรับอินสแตนซ์ต่างๆ อาจทำให้ประสิทธิภาพการทำงานของระบบเปลี่ยนแปลงไปได้มาก ในแง่ของการดำเนินงาน คุณควรจัดเตรียมการตรวจสอบให้พร้อมสำหรับการแจ้งเตือนคุณเมื่อประสิทธิภาพการทำงานเสื่อมถอยลง

ฐานข้อมูล

โซลูชันฐานข้อมูลที่เหมาะสมที่สุดกับระบบใดระบบหนึ่งอาจแตกต่างกันไปตามความต้องการด้านความสอดคล้องของข้อมูล ความพร้อมใช้งาน การทนต่อการจัดเก็บที่แยกออกเป็นเครือข่าย และเวลาแฝง หลายๆ ระบบใช้โซลูชันฐานข้อมูลที่ต่างกันสำหรับระบบย่อยต่างๆ และใช้คุณสมบัติที่แตกต่างเพื่อปรับปรุงประสิทธิภาพให้ดียิ่งขึ้น การเลือก

โซลูชันและคุณสมบัติฐานข้อมูลที่ไม่เหมาะกับระบบอาจเป็นสาเหตุให้ประสิทธิภาพการทำงานลดลง

Amazon Relational Database Service (RDS) ใน AWS มีฐานข้อมูลเชิงสัมพันธ์ที่มีการจัดการเต็มรูปแบบ เมื่อใช้ Amazon RDS คุณสามารถปรับขนาดทรัพยากรการจัดเก็บและการประมวลผลของฐานข้อมูลได้โดยที่ระบบไม่หยุดชะงัก และเรายังมีโซลูชันฐานข้อมูลและการจัดเก็บข้อมูลอื่นๆ อีกด้วย Amazon DynamoDB เป็นฐานข้อมูล NoSQL ที่มีการจัดการเต็มรูปแบบและมีค่าเวลาแฝงไม่ถึง 10 มิลลิวินาทีสำหรับทุกขนาดการใช้งาน Amazon Redshift เป็นคลังข้อมูลในระดับเพตาไบต์ที่ได้รับการจัดการ ซึ่งช่วยให้คุณสามารถเปลี่ยนแปลงจำนวนหรือประเภทโหนดได้เมื่อความต้องการด้านประสิทธิภาพหรือความจุเปลี่ยนแปลงไป

คำถามตัวอย่างต่อไปนี้จะเน้นที่ข้อควรพิจารณาเกี่ยวกับฐานข้อมูลเพื่อประสิทธิภาพการทำงาน:

PERF 9. **คุณจะเลือกโซลูชันฐานข้อมูลที่เหมาะกับระบบของคุณอย่างไร**

PERF 10. **คุณมีวิธีการตรวจสอบอย่างไรว่าคุณใช้โซลูชันและคุณสมบัติฐานข้อมูลที่เหมาะสมที่สุดอยู่ขณะที่มีโซลูชันและคุณสมบัติใหม่ๆ ของฐานข้อมูลออกมาให้บริการ**

PERF 11. **คุณมีวิธีการอย่างไรในการตรวจสอบฐานข้อมูลเพื่อให้แน่ใจว่าฐานข้อมูลนั้นทำงานได้ตามที่คาดหวัง**

PERF 12. **คุณมีวิธีการอย่างไรในการตรวจสอบว่าความจุและอัตราความเร็วของฐานข้อมูลนั้นตรงกับความต้องการ**

ถึงแม้ว่าแนวทางเกี่ยวกับฐานข้อมูลขององค์กร (RDBMS, NoSQL เป็นต้น) จะส่งผลกระทบสำคัญต่อประสิทธิภาพการทำงานของระบบ แต่ก็มักเป็นส่วนที่เลือกใช้ตามสถานะพื้นฐานขององค์กรมากกว่าที่จะเลือกผ่านการประเมินระหว่างการสร้างและปรับใช้โซลูชันฐานข้อมูล ให้ดำเนินการกับฐานข้อมูลในรูปแบบของโค้ดเพื่อรองรับการพัฒนาที่ต่อเนื่องมากกว่าการตัดสินใจครั้งเดียวแบบตายตัว ใช้ข้อมูลสำหรับทดสอบเพื่อค้นหาโซลูชันฐานข้อมูลที่เหมาะกับแต่ละเวิร์กโหลดมากที่สุด การทดสอบเหล่านี้ต้องสามารถทำซ้ำได้ (ควรเป็นส่วนหนึ่งของไปป์ไลน์ CD) เพื่อให้คุณทดสอบโซลูชันหรือความสามารถใหม่ๆ ของฐานข้อมูลได้ง่ายเมื่อมีให้บริการ ตัวอย่างเช่น ประเมินว่าการทำซ้ำแบบอ่านอย่างเดียวจะช่วยเพิ่มประสิทธิภาพการทำงานโดยไม่เป็นการละเมิดความต้องการอื่นๆ ที่ไม่ใช่หน้าที่หลักหรือไม่ ในแง่ของการดำเนินงาน คุณควรจัดเตรียมการตรวจสอบให้พร้อมสำหรับการแจ้งเตือนคุณเมื่อประสิทธิภาพการทำงานเสื่อมถอยลง

การแลกกันของพื้นที่กับเวลา

เมื่อออกแบบสถาปัตยกรรมโซลูชัน มีการใช้การแลกกันพื้นที่ (หน่วยความจำหรือที่เก็บข้อมูล) เพื่อลดเวลาในการ

ดำเนินการ (การประมวลผล) หรือมีการใช้เวลาในการลดพื้นที่ คุณสามารถจัดตำแหน่งของทรัพยากรหรือข้อมูลที่แคชให้อยู่ใกล้ผู้ใช้มากขึ้นเพื่อลดเวลาได้เช่นกัน

เมื่อใช้ AWS คุณสามารถเข้าถึงพื้นที่ทั่วโลกได้ในไม่กี่นาทีและปรับใช้ทรัพยากรในหลายที่ตั้งทั่วโลกเพื่อให้อยู่ในตำแหน่งที่ใกล้ผู้ใช้มากขึ้น คุณยังสามารถปรับเพิ่มการทำซ้ำแบบอ่านอย่างเดียวในที่เก็บข้อมูลต่างๆ เช่น ระบบฐานข้อมูล เพื่อลดภาระงานในฐานข้อมูลหลัก

ใช้โครงสร้างพื้นฐานที่เป็นสากลของ AWS เพื่อลดค่าเวลาแฝงและเพิ่มปริมาณงานที่ได้ และดูแลให้ข้อมูลของคุณอยู่ในภูมิภาคที่คุณระบุเท่านั้น โซลูชันระบบเครือข่าย เช่น AWS Direct Connect ออกแบบมาให้มีค่าเวลาแฝงที่คาดการณ์ได้ระหว่างเครือข่ายภายในองค์กรกับโครงสร้างพื้นฐานของ AWS นอกจากนี้ AWS ยังมีโซลูชันการแคชข้อมูล เช่น Amazon ElastiCache ที่จะช่วยปรับปรุงประสิทธิภาพการทำงาน และ Amazon CloudFront ที่สามารถแคชสำเนาเนื้อหาแบบคงที่ให้อยู่ใกล้ผู้ใช้มากขึ้น

คำถามตัวอย่างต่อไปนี้จะเน้นที่การแลกกันของพื้นที่กับเวลาเพื่อประสิทธิภาพการทำงาน:

PERF 13. **คุณจะเลือกโซลูชันการแคชข้อมูลและการตรวจหาอุปกรณ์ระยะใกล้เคียงสำหรับระบบของคุณอย่างไร**

PERF 14. **คุณมีวิธีการอย่างไรในการตรวจสอบว่าคุณใช้โซลูชันการแคชข้อมูลและการตรวจหาอุปกรณ์ระยะใกล้เคียงที่เหมาะสมที่สุดอยู่ขณะที่มีโซลูชันใหม่ๆ ออกมาให้บริการ**

PERF 15. **คุณมีวิธีการอย่างไรในการตรวจสอบโซลูชันการแคชข้อมูลและการตรวจหาอุปกรณ์ระยะใกล้เคียงเพื่อให้แน่ใจว่าประสิทธิภาพเป็นไปตามที่คาดหวัง**

PERF 16. **คุณมีวิธีการอย่างไรในการตรวจสอบว่าโซลูชันการแคชข้อมูลและการตรวจหาอุปกรณ์ระยะใกล้เคียงที่คุณมีอยู่เป็นไปตามความต้องการ**

การแลกกันของพื้นที่กับเวลานั้นจำเป็นต่อการให้ประสิทธิภาพการทำงาน และคุณควรมีข้อมูลสำหรับทดสอบที่แสดงการแลกเปลี่ยนกันที่เหมาะกับเวิร์กโหลด*นั้นๆ* มากที่สุด การทดสอบเหล่านี้ต้องสามารถทำซ้ำได้ (ควรเป็นส่วนหนึ่งของไปป์ไลน์ CD) เพื่อให้คุณทดสอบวิธีการหรือความสามารถใหม่ๆ ได้ง่ายเมื่อมีให้บริการ ตัวอย่างเช่น ทดสอบเพื่อดูว่าการใช้ Amazon ElastiCache เป็นแคชสำรองการเขียน (Write-aside) จะช่วยเพิ่มประสิทธิภาพการทำงานโดยไม่เป็นการละเมิดความต้องการอื่นๆ ที่ไม่ใช่หน้าที่หลักหรือไม่ ในแง่ของการดำเนินงาน คุณควรจัดเตรียมการตรวจสอบให้พร้อมสำหรับการแจ้งเตือนคุณเมื่อประสิทธิภาพการทำงานเสื่อมถอยลง สถาปัตยกรรมนี้ควรปรับขนาดให้เหมาะกับความต้องการและรักษาขอบเขตดังกล่าวไว้เสมอ

บริการ AWS ที่สำคัญ

บริการ AWS ที่สำคัญต่อประสิทธิภาพการทำงาน คือ Amazon CloudWatch ซึ่งจะตรวจสอบทรัพยากรและระบบเพื่อแสดงผลข้อมูลเกี่ยวกับประสิทธิภาพโดยรวมและความสมบูรณ์ในการดำเนินงาน บริการที่สำคัญต่อขอบเขตทั้งสี่ส่วนของประสิทธิภาพการทำงานมีดังนี้:

การประมวลผล: Auto Scaling เป็นคุณสมบัติสำคัญที่ช่วยให้แน่ใจได้ว่าคุณมีอินสแตนซ์เพียงพอกับความต้องการและรักษาการตอบสนองที่สม่ำเสมอ

การจัดเก็บข้อมูล: Amazon EBS มีตัวเลือกการจัดเก็บข้อมูลที่หลากหลาย (เช่น SSD และ PIOPS) ที่ช่วยให้คุณสามารถปรับเปลี่ยนให้เหมาะกับกรณีการใช้งาน Amazon S3 มีการจัดเก็บที่ลดความซ้ำซ้อนของข้อมูล (Reduced Redundancy Storage) และนโยบายวงจรการใช้งานสำหรับ Amazon Glacier (การแยกเก็บถาวร) และการส่งมอบเนื้อหาแบบไม่ใช้เซิร์ฟเวอร์

ฐานข้อมูล: Amazon RDS มีคุณสมบัติเกี่ยวกับฐานข้อมูลที่หลากหลาย (เช่น IOPS ที่ผ่านการเตรียมใช้งาน และการจำลองการอ่าน) ที่ช่วยให้คุณสามารถปรับเปลี่ยนให้เหมาะกับกรณีการใช้งาน Amazon DynamoDB มีค่าเวลาแฝงที่น้อยกว่า 10 มิลลิวินาทีสำหรับทุกขนาดการใช้งาน

การแลกกันของพื้นที่กับเวลา: AWS ให้บริการในหลายภูมิภาคทั่วโลก คุณจึงสามารถเลือกที่ตั้งที่เหมาะสมสำหรับทรัพยากร ข้อมูล และการประมวลผลของคุณได้ ใช้ Amazon CloudFront เพื่อแคชเนื้อหาที่อยู่ใกล้กับผู้ใช้ของคุณมากกว่า

แหล่งข้อมูล

โปรดดูแหล่งข้อมูลต่อไปนี้เพื่อศึกษาเพิ่มเติมเกี่ยวกับแนวทางปฏิบัติที่เกี่ยวข้องกับประสิทธิภาพการทำงาน

วิดีโอ

- Performance Channel
- การเทียบวัดมาตรฐานของประสิทธิภาพการทำงานใน AWS

เอกสารประกอบ

- เอกสารประกอบเกี่ยวกับการเพิ่มประสิทธิภาพการทำงานด้วย Amazon S3

- เอกสารเกี่ยวกับประสิทธิภาพของไดรฟ์ข้อมูล Amazon EBS

เสาหลักด้านการเพิ่มประสิทธิภาพต้นทุน

ใช้เสาหลักด้าน **การเพิ่มประสิทธิภาพต้นทุน** เพื่อประเมินความสามารถในการหลีกเลี่ยงหรือลดต้นทุนที่ไม่จำเป็น หรือทรัพยากรที่มีคุณภาพต่ำกว่ามาตรฐาน และใช้ต้นทุนส่วนที่ประหยัดเพื่อให้เกิดประโยชน์ที่สร้างความแตกต่างให้กับธุรกิจ ระบบที่เพิ่มประสิทธิภาพต้นทุนช่วยให้คุณสามารถบรรลุเป้าหมายทางธุรกิจและดำเนินการตามเสาหลักด้านอื่นๆ ของโครงสร้าง Well-Architected ได้ตามหรือเกินความต้องการ โดยใช้ต้นทุนจำนวนน้อยที่สุด คุณสามารถบรรลุผลตามการเพิ่มประสิทธิภาพต้นทุนได้โดยใช้เทคนิคต่างๆ เพื่อเลือกสถาปัตยกรรมที่เหมาะสม ลดทรัพยากรที่ไม่ได้ใช้ และเลือกวิธีการที่คุ้มค่ามากที่สุด

หลักการออกแบบ

คุณสามารถปฏิบัติตามหลักการต่างๆ ในระบบคลาวด์ที่จะช่วยให้บรรลุผลตามการเพิ่มประสิทธิภาพต้นทุนได้ดังนี้:

- **กำหนดค่าใช้จ่ายอย่างโปร่งใส:** ระบบคลาวด์ช่วยให้ระบุต้นทุนของระบบและกำหนดให้เจ้าของธุรกิจแต่ละรายเป็นผู้ดูแลต้นทุนด้านเทคโนโลยีสารสนเทศได้ง่ายขึ้น จึงช่วยให้ระบุผลตอบแทนจากการลงทุนได้ และส่งผลให้เจ้าของธุรกิจมีแรงจูงใจที่จะเพิ่มประสิทธิภาพของทรัพยากรและลดต้นทุนเหล่านั้น
- **ใช้บริการที่ได้รับการจัดการเพื่อลดต้นทุนในการเป็นเจ้าของ:** ในระบบคลาวด์ บริการที่ได้รับการจัดการจะช่วยลดภาระการดำเนินงานจากการที่ต้องบำรุงรักษาเซิร์ฟเวอร์เพื่อทำงานต่างๆ เช่น ส่งอีเมลหรือจัดการฐานข้อมูล นอกจากนี้ เนื่องจากบริการเหล่านี้ดำเนินงานในระดับของระบบคลาวด์ จึงมีต้นทุนต่อบริการหรือต่อธุรกรรมที่ต่ำกว่า
- **เปลี่ยนค่าใช้จ่ายในการลงทุนเป็นค่าใช้จ่ายการดำเนินงาน:** แทนที่จะลงทุนด้วยเงินจำนวนมากไปกับศูนย์ข้อมูลและเซิร์ฟเวอร์ก่อนที่คุณจะทราบลักษณะการใช้งาน คุณสามารถจ่ายเงินเพื่อใช้เฉพาะทรัพยากรการประมวลผลที่คุณต้องการและในเวลาที่ต้องการได้ ตัวอย่างเช่น โดยทั่วไปแล้ว สภาพแวดล้อมการพัฒนาและการทดสอบใช้เวลาเพียงแปดชั่วโมงต่อวันเท่านั้นในระหว่างวันทำงาน คุณจึงสามารถหยุดใช้ทรัพยากรเหล่านี้ได้เมื่อไม่ใช้งาน เพื่อให้ประหยัดต้นทุนได้ถึง 75% (40 ชั่วโมงเทียบกับ 168 ชั่วโมง)
- **ใช้ประโยชน์จากการประหยัดต่อขนาด:** การใช้ระบบประมวลผลในระบบคลาวด์จะช่วยลดต้นทุนผันแปรให้คุณได้มากขึ้น เนื่องจาก AWS มีอัตราการประหยัดที่สูงกว่าเมื่อคำนวณจากพื้นที่ ในระบบคลาวด์ของ AWS มีลูกค้ารวมอยู่หลายหมื่นราย ซึ่งเมื่อคำนวณแล้วจะได้ราคาการจ่ายเท่าที่ใช้งานเป็นจำนวนที่น้อยกว่า
- **หยุดใช้เงินไปกับการดำเนินการศูนย์ข้อมูล:** AWS จัดเตรียมการติดตั้ง ประกอบ และเพิ่มประสิทธิภาพของ

เซิร์ฟเวอร์ คุณจึงสามารถให้ความสำคัญกับลูกค้าและโครงการธุรกิจได้เต็มที่แทนที่จะมุ่งเน้นกับโครงสร้างพื้นฐานด้านเทคโนโลยีสารสนเทศ

คำจำกัดความ

การเพิ่มประสิทธิภาพต้นทุนในระบบคลาวด์ประกอบด้วยสี่ส่วนดังนี้:

1. อุปสงค์และอุปทานที่สอดคล้องกัน
2. ทรัพยากรที่คุ้มค่า
3. การรับรู้ค่าใช้จ่าย
4. การเพิ่มประสิทธิภาพอย่างต่อเนื่อง

การเพิ่มประสิทธิภาพต้นทุนมีข้อดีข้อเสียที่ต้องพิจารณาเช่นเดียวกับเสาหลักด้านอื่นๆ เช่น ควรเพิ่มประสิทธิภาพเพื่อความรวดเร็วในการเข้าสู่ตลาดหรือเพื่อต้นทุน ในบางกรณี ตัวเลือกที่ดีที่สุดคือเพิ่มประสิทธิภาพเพื่อความรวดเร็ว กล่าวคือ เข้าสู่ตลาดอย่างรวดเร็ว นำเสนอคุณสมบัติใหม่ๆ หรือดำเนินงานให้เสร็จตามกำหนดเวลามากกว่าการลงทุนเพื่อเพิ่มประสิทธิภาพต้นทุนล่วงหน้า การตัดสินใจด้านการออกแบบจะดำเนินไปอย่างเร่งรีบในบางครั้งแทนการศึกษาข้อมูลเชิงประจักษ์ เนื่องจากมีสิ่งกระตุ้นให้เติมเต็มส่วนที่ขาดในลักษณะของการ "เผื่อไว้" มากกว่าการใช้เวลาเทียบวัดมาตรฐานเพื่อการปรับใช้ที่เหมาะกับต้นทุนมากที่สุดอยู่เสมอ ซึ่งแนวทางนี้มักทำให้การปรับใช้มีจำนวนที่มากเกินไปและมีความเหมาะสมน้อยกว่าที่กำหนด หัวข้อต่อไปนี้จะมีเทคนิคและคำแนะนำเชิงกลยุทธ์สำหรับการเพิ่มประสิทธิภาพต้นทุนในเบื้องต้นและแบบต่อเนื่องสำหรับการปรับใช้ระบบ

แนวทางปฏิบัติ

อุปสงค์และอุปทานที่สอดคล้องกัน

การจับคู่ความต้องการใช้งาน (อุปสงค์) กับการตอบสนองความต้องการ (อุปทาน) อย่างเหมาะสมจะทำให้ใช้ต้นทุนสำหรับระบบน้อยที่สุด แต่จะต้องมีการตอบสนองความต้องการที่เพียงพอเผื่อไว้ด้วย ในกรณีที่เกิดข้อผิดพลาดเกี่ยวกับเวลาจัดเตรียมและทรัพยากรเฉพาะอย่าง ความต้องการอาจกำหนดตายตัวหรือแปรผันได้ จึงต้องมีตัววัดผลและระบบอัตโนมัติเพื่อตรวจสอบให้มั่นใจได้ว่าการบริหารจัดการไม่ได้ก่อให้เกิดต้นทุนที่มากจนเกินไป

ใน AWS คุณสามารถเตรียมใช้งานทรัพยากรโดยอัตโนมัติให้สอดคล้องกับความต้องการได้ Auto Scaling และแนวทางปรับใช้ตามเวลา ตามเหตุการณ์ และตามคิวช่วยให้คุณสามารถเพิ่มและย้ายทรัพยากรออกได้ตามที่จำเป็น หากคุณสามารถคาดคะเนถึงความต้องการที่เปลี่ยนแปลงไป คุณจะสามารถประหยัดต้นทุนได้มากขึ้นและมั่นใจได้

ว่าทรัพยากรต่างๆ สอดคล้องกับความต้องการของระบบ

คำถามตัวอย่างต่อไปนี้จะเน้นถึงอุปสงค์และอุปทานที่สอดคล้องกันสำหรับการเพิ่มประสิทธิภาพต้นทุน (โปรดดูรายการคำถาม คำตอบ และแนวทางปฏิบัติทั้งหมดเกี่ยวกับการเพิ่มประสิทธิภาพต้นทุนจากภาคผนวก):

COST 1. คุณมีวิธีการอย่างไรในการตรวจสอบว่าประสิทธิภาพความจุสอดคล้องกับความต้องการแต่ไม่มากจนเกินไป

COST 2. คุณจะเพิ่มประสิทธิภาพการใช้งานบริการ AWS อย่างไร

เครื่องมือการตรวจสอบและการเทียบวัดมาตรฐานอย่างสม่ำเสมอจะช่วยให้คุณใช้ประโยชน์จากทรัพยากรได้อย่างมีประสิทธิภาพยิ่งขึ้น ความยืดหยุ่นของการประมวลผลตามความต้องการ คุณสมบัติ Auto Scaling และกลไกการปรับใช้แบบอัตโนมัติอื่นๆ จะช่วยเพิ่มประสิทธิภาพให้เหมาะสมยิ่งขึ้น เพื่อให้คุณเตรียมใช้งานเฉพาะทรัพยากรที่จำเป็นและปรับเพิ่มลดขนาดที่จะใช้งานได้

ทรัพยากรที่คุ้มค่า

การใช้อินสแตนซ์และทรัพยากรที่เหมาะกับระบบของคุณเป็นกุญแจสำคัญสู่การประหยัดต้นทุน ตัวอย่างเช่น กระบวนการรายงานข้อมูลอาจใช้เวลาห้าชั่วโมงบนเซิร์ฟเวอร์ขนาดเล็กกว่า แต่เซิร์ฟเวอร์ที่ใหญ่กว่าซึ่งมีราคาแพงกว่าสองเท่าสามารถดำเนินการได้ในหนึ่งชั่วโมง งานทั้งสองอย่างนี้ให้ผลลัพธ์เดียวกัน แต่เซิร์ฟเวอร์ขนาดเล็กกว่าจะมีต้นทุนสูงกว่าเมื่อเทียบกับเวลา

ระบบสถาปัตยกรรมที่เหมาะสมจะใช้ทรัพยากรที่ประหยัดต้นทุนมากที่สุด ซึ่งจะให้ผลในเชิงบวกที่สำคัญทางด้านการเงิน นอกจากนี้ คุณยังสามารถใช้บริการที่ได้รับการจัดการเพื่อลดต้นทุนได้อีกด้วย ตัวอย่างเช่น แทนที่จะต้องดูแลรักษาเซิร์ฟเวอร์เพื่อส่งอีเมล คุณสามารถใช้บริการที่คิดค่าบริการตามจำนวนของข้อความได้

AWS มีตัวเลือกด้านราคาที่ยืดหยุ่นและประหยัดต้นทุนครอบคลุมสำหรับการใช้บริการ Amazon EC2 Instance ในลักษณะที่เหมาะกับความต้องการของคุณ *อินสแตนซ์ตามการใช้งานจริง (On-Demand Instance)* จะคิดค่าบริการความจุของระบบประมวลผลตามจำนวนชั่วโมง โดยไม่มีข้อผูกมัดเกี่ยวกับจำนวนขั้นต่ำ *อินสแตนซ์แบบเหมาจ่าย (Reserved Instance: RI)* ช่วยให้คุณสามารถสำรองความจุและประหยัดค่าใช้จ่ายได้ถึง 75 เปอร์เซ็นต์เมื่อเทียบกับราคาตามการใช้งานจริง เมื่อใช้ *อินสแตนซ์แบบประมูลราคา (Spot Instance)* คุณสามารถประมูลความจุ Amazon EC2 ที่ไม่มีการใช้งานได้โดยมีส่วนลดให้ อินสแตนซ์แบบประมูลราคาเหมาะกับการใช้งานในกรณีที่ระบบทนต่อการใช้งานกลุ่มเซิร์ฟเวอร์ที่แต่ละเซิร์ฟเวอร์อาจเกิดการเปลี่ยนแปลงได้ตลอด เช่น เมื่อใช้ HPC หรือ Big Data

คำถามตัวอย่างต่อไปนี้จะเน้นที่การเลือกทรัพยากรที่คุ้มค่าเพื่อการเพิ่มประสิทธิภาพต้นทุน:

COST 3. **คุณเลือกประเภททรัพยากรที่เหมาะสมเพื่อให้บรรลุตามเป้าหมายด้านต้นทุนหรือไม่**

COST 4. **คุณเลือกรูปแบบราคาที่เหมาะสมเพื่อให้บรรลุตามเป้าหมายด้านต้นทุนหรือไม่**

COST 5. **มีบริการที่ได้รับการจัดการ (บริการในระดับที่สูงกว่า Amazon EC2, Amazon EBS และ Amazon S3) ที่คุณสามารถใช้เพื่อปรับปรุง ROI หรือไม่**

การใช้เครื่องมือต่างๆ เช่น AWS Trusted Advisor เพื่อตรวจสอบการใช้งาน AWS เป็นประจำจะช่วยให้คุณติดตามการใช้ประโยชน์และแก้ไขการปรับใช้ได้อย่างต่อเนื่องและสอดคล้องกัน คุณยังสามารถใช้ประโยชน์จากบริการ AWS ที่ได้รับการจัดการ เช่น Amazon RDS, Amazon Elastic MapReduce (EMR) และ Amazon DynamoDB ซึ่งจะช่วยลดต้นทุนการจัดการและการใช้งานต่อหนึ่งรายการได้ พิจารณาเลือกโซลูชัน CDN เช่น Amazon CloudFront เพื่อลดต้นทุนที่เกี่ยวข้องกับปริมาณข้อมูลในเครือข่าย

การรับรู้ค่าใช้จ่าย

ความยืดหยุ่นและความคล่องตัวที่เพิ่มมากขึ้นจากระบบคลาวด์ช่วยเสริมนวัตกรรม รวมทั้งการพัฒนาและการปรับใช้ที่รวดเร็ว ทั้งยังลดกระบวนการที่ต้องดำเนินการเองและเวลาที่เกี่ยวข้องกับการเตรียมใช้งานในโครงสร้างพื้นฐานภายในองค์กร ซึ่งรวมถึง

การระบุข้อกำหนดเฉพาะของฮาร์ดแวร์ การเจรจาต่อรองราคา การจัดการใบสั่งซื้อ การกำหนดเวลาจัดส่ง และการปรับใช้ทรัพยากร อย่างไรก็ตาม การใช้งานที่สะดวกและความจุตามความต้องการแบบไม่จำกัดในระบบเสมือนอาจต้องใช้แนวทางใหม่ในการพิจารณาเกี่ยวกับค่าใช้จ่าย

หลายๆ ธุรกิจมีระบบที่ใช้งานหลายระบบซึ่งดูแลโดยทีมต่างๆ ความสามารถที่จะกำหนดให้เจ้าของธุรกิจแต่ละรายหรือเจ้าของผลิตภัณฑ์เป็นผู้ดูแลต้นทุนทรัพยากรจะส่งเสริมให้เกิดการใช้งานที่มีประสิทธิภาพและช่วยลดความสูญเสีย การกำหนดต้นทุนที่ถูกต้องยังช่วยให้คุณทราบถึงผลิตภัณฑ์ที่สามารถทำกำไรได้อย่างแท้จริง และช่วยให้คุณตัดสินใจจากข้อมูลที่ครบถ้วนว่าจะจัดสรรงบประมาณให้กับส่วนใดบ้าง

คำถามตัวอย่างต่อไปนี้จะเน้นที่การรับรู้ค่าใช้จ่ายเพื่อการเพิ่มประสิทธิภาพต้นทุน:

COST 6. **คุณมีการควบคุมการเข้าถึงและขั้นตอนใดบ้างในการกำกับดูแลต้นทุน AWS**

COST 7. **คุณมีวิธีการอย่างไรในการตรวจสอบการใช้งานและค่าใช้จ่าย**

COST 8. **คุณมีวิธีการอย่างไรในการปลดการใช้งานทรัพยากรที่ไม่ต้องการใช้อีกต่อไปหรือหยุดใช้ทรัพยากรที่ไม่จำเป็นชั่วคราว**

COST 9. **คุณมีวิธีการอย่างไรในการพิจารณาค่าใช้จ่ายสำหรับการถ่ายโอนข้อมูลในการออกแบบสถาปัตยกรรม**

คุณสามารถใช้แท็กการจัดสรรต้นทุนเพื่อแยกประเภทและติดตามต้นทุนของ AWS เมื่อคุณใช้แท็กกับทรัพยากรของ AWS (เช่น Amazon EC2 Instance หรือบัคเก็ต Amazon S3) AWS จะสร้างรายงานการจัดสรรต้นทุนที่มีข้อมูลการใช้งานและต้นทุนที่รวมไว้โดยแท็กของคุณ คุณสามารถใช้แท็กที่แสดงถึงหมวดหมู่ทางธุรกิจ (เช่น ศูนย์ต้นทุน ชื่อระบบ หรือเจ้าของ) เพื่อจัดการต้นทุนในหลายๆ บริการได้

ข้อมูลแสดงผลต้นทุนกับทรัพยากรที่ติดแท็กไว้จะช่วยเพิ่มความสะดวกในการระบุทรัพยากรที่ไม่ได้อยู่ในส่วนใดหรือโครงการที่ไม่ได้สร้างมูลค่าให้กับธุรกิจอีกต่อไปและควรเลิกดำเนินการ คุณสามารถตั้งค่าแจ้งเตือนการเรียกเก็บเงินเพื่อให้แจ้งเตือนคุณเมื่อมีการใช้จ่ายเกินจำนวนที่คาดการณ์ไว้ โดยที่ AWS Simple Monthly Calculator สามารถช่วยคุณคำนวณต้นทุนการถ่ายโอนข้อมูล

การเพิ่มประสิทธิภาพอย่างต่อเนื่อง

เมื่อ AWS นำเสนอบริการและคุณสมบัติใหม่ๆ ออกสู่ตลาด แนวทางที่ควรปฏิบัติก็คือ การประเมินระบบสถาปัตยกรรมที่ตัดสินใจเลือกอีกครั้งเพื่อให้แน่ใจว่าระบบดังกล่าวยังคงมีประสิทธิภาพที่คุ้มต้นทุนมากที่สุด เมื่อความต้องการเปลี่ยนแปลงไป ควรพิจารณาปลดการทำงานของทรัพยากร บริการทั้งบริการ หรือระบบที่คุณไม่จำเป็นต้องใช้อีก

บริการที่ได้รับการจัดการจาก AWS มักจะช่วยให้โซลูชันทำงานได้อย่างเหมาะสมยิ่งขึ้น คุณจึงควรทราบถึงบริการที่ได้รับการจัดการใหม่ๆ เมื่อมีให้บริการ เช่น การเรียกใช้ฐานข้อมูล Amazon RDS จะมีค่าใช้จ่ายที่น้อยกว่าการเรียกใช้ฐานข้อมูลของคุณเองบน Amazon EC2

คำถามตัวอย่างต่อไปนี้จะเน้นที่การประเมินค่าใช้จ่ายซ้ำเพื่อการเพิ่มประสิทธิภาพต้นทุน:

COST 10. **คุณจะจัดการและ/หรือพิจารณานำบริการใหม่ๆ มาใช้งานอย่างไร**

การประเมินสิ่งที่คุณปรับใช้อย่างต่อเนื่องเป็นประจำจะช่วยให้คุณสามารถใช้ประโยชน์จากบริการใหม่ๆ ของ AWS เพื่อลดต้นทุนได้ นอกจากนี้ คุณควรประเมินการประยุกต์ใช้งานบริการที่ใหม่กว่าเพื่อช่วยประหยัดค่าใช้จ่ายให้กับคุณ เช่น AWS RDS สำหรับ Aurora สามารถช่วยลดต้นทุนในการใช้ฐานข้อมูลเชิงสัมพันธ์

บริการ AWS ที่สำคัญ

คุณสมบัติสำคัญของ AWS ที่รองรับการเพิ่มประสิทธิภาพต้นทุนได้แก่ แท็กการจัดสรรต้นทุน ซึ่งจะช่วยให้คุณเข้าใจถึงต้นทุนของระบบ บริการและคุณสมบัติที่สำคัญต่อการเพิ่มประสิทธิภาพต้นทุนในสี่ด้านมีดังนี้:

อุปสงค์และอุปทานที่สอดคล้องกัน: Auto Scaling ช่วยให้คุณสามารถเพิ่มหรือย้ายทรัพยากรออกเพื่อให้สอดคล้องกับความต้องการโดยไม่เกิดการใช้เกินจำนวน

ทรัพยากรที่คุ้มค่า: คุณสามารถใช้ Reserved Instances และความจุแบบเหมาจ่ายล่วงหน้าเพื่อลดต้นทุนได้ AWS Trusted Advisor สามารถใช้เพื่อตรวจสอบสภาพแวดล้อม AWS และค้นหาโอกาสในการประหยัดค่าใช้จ่าย

การรับรู้ค่าใช้จ่าย: ระบบแจ้งเตือนของ Amazon CloudWatch และระบบแจ้งข้อมูลของ Amazon Simple Notification Service (SNS) จะเตือนให้คุณทราบเมื่อคุณมียอดค่าใช้จ่ายหรือคาดการณ์ว่าจะมียอดค่าใช้จ่ายเกินงบประมาณที่กำหนด

การเพิ่มประสิทธิภาพอย่างต่อเนื่อง: บล็อกของ AWS และหัวข้อ *What's New* บนเว็บไซต์ AWS เป็นแหล่งข้อมูลสำหรับการเรียนรู้คุณสมบัติและบริการใหม่ๆ ที่มีการเปิดตัว AWS Trusted Advisor จะตรวจสอบสภาพแวดล้อม AWS ของคุณ และค้นหาโอกาสในการประหยัดค่าใช้จ่ายโดยการตัดทรัพยากรที่ไม่ได้ใช้งานหรือไม่มีผลต่อการใช้งาน หรือปรับไปสู่การใช้ความจุในรูปแบบ Reserved Instance

แหล่งข้อมูล

โปรดดูแหล่งข้อมูลต่อไปนี้เพื่อศึกษาเพิ่มเติมเกี่ยวกับแนวทางปฏิบัติของ AWS สำหรับการเพิ่มประสิทธิภาพต้นทุน

วิดีโอ

- การเพิ่มประสิทธิภาพต้นทุนใน AWS

เอกสารประกอบ

- AWS Economics Center

เครื่องมือ

- เครื่องมือคำนวณต้นทุนรวมในการเป็นเจ้าของ (TCO) ของ AWS
- รายงานการเรียกเก็บเงินโดยละเอียดของ AWS
- AWS Simple Monthly Calculator
- AWS Cost Explorer

บทสรุป

AWS Well-Architected Framework นำเสนอแนวทางปฏิบัติด้านสถาปัตยกรรมที่ครอบคลุมเสาหลักสี่ด้านสำหรับการออกแบบระบบที่น่าเชื่อถือ ปลอดภัย มีประสิทธิภาพ และคุ้มค่าใช้จ่ายในระบบคลาวด์ เฟรมเวิร์กนี้ยังมาพร้อมกับชุดคำถามที่คุณสามารถใช้เพื่อประเมินสถาปัตยกรรมที่มีอยู่หรือได้รับการนำเสนอ รวมทั้งชุดแนวทางปฏิบัติของ AWS สำหรับเสาหลักแต่ละด้าน การใช้เฟรมเวิร์กนี้ในสถาปัตยกรรมของคุณจะช่วยให้สามารถสร้างระบบที่ทำงานเสถียรและมีประสิทธิภาพ เพื่อที่คุณสามารถให้ความสำคัญกับสิ่งที่เป็นความต้องการเชิงการใช้งานได้อย่างเต็มที่

ผู้ร่วมจัดทำ

บุคคลและหน่วยงานดังต่อไปนี้เป็นผู้ร่วมจัดทำเอกสารฉบับนี้:

- Philip Fitzsimons, สถาปนิกโซลูชันโปรแกรมจัดการ, Amazon Web Services
- Erin Rifkin, ผู้จัดการโปรแกรมอาวุโส, Amazon Web Services
- Callum Hughes, สถาปนิกโซลูชัน, Amazon Web Services
- Max Ramsay, สถาปนิกโซลูชันหลักด้านความปลอดภัย, Amazon Web Services
- Scott Paddock, สถาปนิกโซลูชันความปลอดภัย, Amazon Web Services

ประวัติของเอกสาร

20 พฤศจิกายน 2015 อัปเดตภาคผนวกที่มีข้อมูล Amazon CloudWatch Log ปัจจุบัน

ภาคผนวก: คำถาม คำตอบ และแนวทางปฏิบัติเกี่ยวกับระบบ Well-Architected

ภาคผนวกนี้ประกอบด้วยรายการคำถามและคำตอบทั้งหมดเกี่ยวกับระบบ Well-Architected รวมทั้งแนวทางปฏิบัติต่างๆ โดยแยกตามเสาหลักแต่ละด้าน:

เสาหลักด้านความปลอดภัย (SEC)

SEC 1. คุณจะเข้ารหัสและป้องกันข้อมูลในที่จัดเก็บอย่างไร

การควบคุมความปลอดภัยแบบดั้งเดิมจะใช้วิธีเข้ารหัสข้อมูลในที่จัดเก็บ AWS รองรับคุณสมบัตินี้โดยใช้การเข้ารหัสทั้งทางฝั่งไคลเอนต์ (ได้แก่ รองรับ SDK, รองรับระบบปฏิบัติการ, Windows Bitlocker, dm-crypt, Trend Micro SafeNet เป็นต้น) และทางฝั่งเซิร์ฟเวอร์ (ได้แก่ Amazon S3) และคุณยังสามารถใช้ Server-Side Encryption (SSE) และ Amazon Elastic Block Store Encrypted Volumes ได้อีกด้วย

แนวทางปฏิบัติ:

- ข้อมูลในที่จัดเก็บจะได้รับการเข้ารหัสโดยใช้ตัวควบคุมสำหรับบริการ AWS โดยเฉพาะ (ได้แก่ Amazon S3 SSE, ไดรฟ์ข้อมูลที่เข้ารหัสโดย Amazon EBS, Amazon Relational Database Service (RDS) Transparent Data Encryption (TDE) เป็นต้น)
- ข้อมูลในที่จัดเก็บจะได้รับการเข้ารหัสโดยใช้วิธีทางเทคนิคของฝั่งไคลเอ็นต์
- โซลูชันจาก AWS Marketplace หรือ APN Partner

SEC 2. คุณจะเข้ารหัสและป้องกันข้อมูลที่อยู่ระหว่างการรับส่งอย่างไร

แนวทางปฏิบัติที่ดีที่สุดคือ ใช้การเข้ารหัสป้องกันข้อมูลที่อยู่ระหว่างการรับส่ง AWS รองรับคุณสมบัตินี้โดยใช้จุดเชื่อมต่อที่มีการเข้ารหัสสำหรับ API บริการ นอกจากนี้ ลูกค้าสามารถใช้เทคนิคอื่นๆ ภายใน Amazon EC2 Instance ได้

แนวทางปฏิบัติ:

- ใช้ AWS API ที่มีการใช้งาน SSL อย่างเหมาะสม

- ใช้ SSL หรือระบบที่เทียบเท่าในการติดต่อสื่อสาร
- โซลูชันที่มีการใช้งาน VPN
- ระบบเชื่อมต่อแบบส่วนตัว (เช่น AWS Direct Connect)
- ใช้โซลูชันจาก AWS Marketplace

SEC 3. คุณจะป้องกันการเข้าถึงและการใช้งานข้อมูลประจำตัวบัญชีหลักของ AWS อย่างไร

ข้อมูลประจำตัวบัญชีรูทของ AWS นั้นคล้ายกับรูทหรือผู้ดูแลระบบภายในของระบบปฏิบัติการอื่นๆ และควรใช้งานอย่างจำกัด แนวทางปฏิบัติในปัจจุบัน ได้แก่ การสร้างผู้ใช้ในระบบ AWS Identity และ Access Management (IAM) เชื่อมโยงผู้ใช้เข้ากับกลุ่มผู้ดูแลระบบ และใช้ผู้ใช้ IAM เพื่อจัดการบัญชีดังกล่าว บัญชีรูทของ AWS ไม่ควรมีคีย์ API แต่ควรมีรหัสผ่านที่คาดเดายาก และเชื่อมโยงกับอุปกรณ์ที่มีการรับรองความถูกต้องแบบหลายปัจจัย (MFA) สำหรับฮาร์ดแวร์ วิธีนี้จะบังคับให้มีการใช้ข้อมูลระบุตัวตนในระดับรูทผ่าน AWS Management Console เท่านั้น และไม่อนุญาตให้ใช้กับการเรียกใช้อินเทอร์เฟซโปรแกรมแอปพลิเคชัน (API) โปรดทราบว่าตัวแทนจำหน่ายบางรายหรือบางภูมิภาคจะไม่มีให้บริการหรือรองรับข้อมูลประจำตัวบัญชีรูทของ AWS

แนวทางปฏิบัติ:

- ข้อมูลประจำตัวบัญชีรูทของ AWS ใช้สำหรับการใช้งานขั้นต่ำที่จำเป็นเท่านั้น
- มีอุปกรณ์ฮาร์ดแวร์ MFA ที่เชื่อมโยงกับบัญชีรูทของ AWS
- ใช้โซลูชันจาก AWS Marketplace

SEC 4. คุณจะกำหนดบทบาทและความรับผิดชอบสำหรับผู้ใช้ในระบบอย่างไร เพื่อควบคุมการเข้าถึงโดยบุคคลใน AWS Management Console และ API

แนวทางปฏิบัติในปัจจุบัน ได้แก่ การให้ลูกค้าแยกส่วนบทบาทและความรับผิดชอบที่กำหนดของผู้ใช้ในระบบโดยการสร้างกลุ่มผู้ใช้ กลุ่มผู้ใช้สามารถกำหนดได้โดยใช้หลายเทคโนโลยีที่แตกต่างกันเช่น: กลุ่ม Identity and Access Management (IAM), บทบาท IAM สำหรับการเข้าถึงแบบข้ามบัญชี, ข้อมูลประจำตัวบนเว็บ, ผ่านการรวมระบบ Security Assertion Markup Language (SAML) (ได้แก่ การกำหนดบทบาทใน Active Directory) หรือการใช้โซลูชันภายนอก (ได้แก่ Okta, Ping Identity หรือเทคนิคอื่นๆ ที่กำหนดเอง) ซึ่งโดยทั่วไปจะใช้ร่วมกันได้ผ่าน SAML หรือ AWS Security Token Service (STS) ทั้งนี้ ไม่แนะนำให้ใช้บัญชีแบบใช้ร่วมกัน

แนวทางปฏิบัติ:

- ผู้ใช้และกลุ่ม IAM
- การรวมระบบ SAML
- การเชื่อมโยงข้อมูลประจำตัวบนเว็บ
- AWS Security Token Service (STS)
- บทบาท IAM สำหรับการเข้าถึงแบบข้ามบัญชี
- โซลูชันจาก AWS Marketplace (เช่น Okta และ Ping Identity) หรือจาก APN Partner
- กำหนดและบังคับใช้นโยบายวงจรการทำงานของพนักงาน
- กำหนดผู้ใช้ กลุ่ม และบทบาทอย่างชัดเจน และให้สิทธิ์การใช้งานขั้นต่ำที่จำเป็นต่อการดำเนินงานทางธุรกิจให้สำเร็จลุล่วงเท่านั้น

SEC 5. คุณจะจำกัดการเข้าถึงแบบอัตโนมัติไปยังทรัพยากรของ AWS อย่างไร (เช่น จากแอปพลิเคชัน สคริปต์ และ/หรือเครื่องมือหรือบริการภายนอก)

การเข้าถึงแบบเป็นระบบควรได้รับการกำหนดในลักษณะที่คล้ายกับการสร้างกลุ่มผู้ใช้สำหรับการเข้าถึงโดยบุคคล สำหรับ Amazon EC2 Instance กลุ่มเหล่านี้จะเรียกว่าบทบาท IAM สำหรับ EC2 แนวทางปฏิบัติในปัจจุบัน ได้แก่ การใช้บทบาท IAM สำหรับ EC2 และ AWS SDK หรือ CLI ซึ่งรองรับการเรียกข้อมูลบทบาท IAM สำหรับข้อมูลประจำตัว EC2 ได้ในตัว โดยทั่วไปแล้ว ข้อมูลประจำตัวของผู้ใช้จะได้รับการแทรกลงใน EC2 Instance แต่ไม่แนะนำให้ใช้การเข้ารหัสข้อมูลประจำตัวแบบตายตัว (Hard-coding) ลงในสคริปต์และซอร์สโค้ด

แนวทางปฏิบัติ:

- บทบาทของ IAM สำหรับ Amazon EC2
- ใช้ข้อมูลประจำของตัวผู้ใช้ IAM แต่ไม่เข้ารหัสตายตัวลงในสคริปต์และแอปพลิเคชัน
- การรวมระบบ SAML
- AWS Security Token Service (STS)
- ใช้ตัวควบคุมเฉพาะระบบปฏิบัติการสำหรับ EC2 Instance
- ใช้โซลูชันจาก AWS Marketplace

SEC 6. คุณจะจัดการคีย์และข้อมูลประจำตัวอย่างไร

คีย์และข้อมูลประจำตัวเป็นข้อมูลลับที่ควรได้รับการป้องกันและควรมีการกำหนดและใช้นโยบายการหมุนเวียนที่เหมาะสม แนวทางปฏิบัติในเรื่องนี้ก็คือ ไม่ควรเข้ารหัสข้อมูลแบบตายตัวลงในสคริปต์และแอปพลิเคชันโปรแกรมจัดการ แต่การดำเนินการนี้มักจะเกิดขึ้นบ่อย

แนวทางปฏิบัติ:

- ใช้นโยบายการเปลี่ยนคีย์และข้อมูลประจำตัวอย่างเหมาะสม
- ใช้ AWS CloudHSM
- ใช้เทคนิคทางฝั่งเซิร์ฟเวอร์ของ AWS กับคีย์ที่จัดการโดย AWS (เช่น Amazon S3 SSE และไดรฟ์ข้อมูลที่มีการเข้ารหัสของ Amazon EBS)
- โซลูชันจาก AWS Marketplace (เช่น SafeNet และ TrendMicro)

SEC 7. คุณจะบังคับใช้การป้องกันในขอบเขตเครือข่ายและระดับโฮสต์อย่างไร

ในศูนย์ข้อมูลแบบภายในองค์กร วิธี DMZ จะใช้ไฟร์วอลล์เพื่อแยกระบบออกเป็นโซนที่เชื่อถือได้และที่ไม่น่าเชื่อถือ ใน AWS จะมีการใช้งานไฟร์วอลล์ทั้งแบบเก็บสถานะ (Stateful Firewall) และไม่เก็บสถานะ (Stateless Firewall) ไฟร์วอลล์แบบเก็บสถานะจะเรียกว่า กลุ่มความปลอดภัย และไฟร์วอลล์แบบไม่เก็บสถานะจะเรียกว่า รายการควบคุมการเข้าถึง (ACL) ที่ป้องกันซับเน็ตใน Amazon Virtual Private Cloud (VPC) แนวทางปฏิบัติในปัจจุบัน ได้แก่ การเรียกใช้ระบบใน VPC และกำหนดความปลอดภัยอิงตามบทบาทในกลุ่มความปลอดภัย (เช่น ระดับเว็บและระดับแอป) และความปลอดภัยอิงตามที่ตั้งใน ACL เครือข่าย (เช่น ระดับ Elastic Load Balancing ในหนึ่งซับเน็ตต่อหนึ่ง Availability Zone และระดับเว็บในอีกหนึ่งซับเน็ตต่อหนึ่ง Availability Zone)

แนวทางปฏิบัติ:

- ใช้กลุ่มความปลอดภัยที่มีการให้อนุญาตในระดับต่ำสุดเพื่อบังคับใช้การเข้าถึงตามบทบาท
- ระบบทำงานใน VPC อย่างน้อยหนึ่งรายการ
- การเข้าถึง VPC ที่เชื่อถือได้ดำเนินการผ่านกลไกแบบส่วนตัว (ได้แก่ เครือข่ายส่วนตัวเสมือน (VPN), ช่องทางการเชื่อมต่อ IPsec, AWS Direct Connect, โซลูชันจาก AWS Marketplace เป็นต้น)
- ใช้ซับเน็ตและ ACL เครือข่ายอย่างเหมาะสม
- ใช้ไฟร์วอลล์ตามโฮสต์ที่มีการอนุญาตในระดับต่ำสุด

- ใช้ตัวควบคุมการเข้าถึงเฉพาะบริการ (เช่น นโยบายบัคเก็ต)
- ใช้การเชื่อมต่อแบบส่วนตัวไปยัง VPC (ได้แก่ VPN, AWS Direct Connect, การเชื่อมต่อแบบเพียร์ของ VPC เป็นต้น)
- ใช้เทคนิค Bastion Host (โฮสต์ที่มีความเสี่ยงสูงต่อการถูกโจมตี) เพื่อจัดการอินสแตนซ์
- ดำเนินการทดสอบความปลอดภัยอย่างสม่ำเสมอ
- ประเมินการตรวจสอบของ AWS Trusted Advisor อย่างสม่ำเสมอ

SEC 8. คุณจะบังคับใช้การป้องกันในระดับบริการ AWS อย่างไร

แนวทางปฏิบัติอีกอย่างหนึ่งได้แก่ การควบคุมการเข้าถึงทรัพยากร AWS Identity and Access Management (IAM) ช่วยให้คุณสามารถกำหนดการควบคุมทรัพยากรในระดับต่างๆ ได้ (เช่น การใช้การเข้ารหัส เวลาในวัน และ IP ต้นทาง) และบริการที่แตกต่างกันเอื้อต่อการใช้เทคนิคเพิ่มเติม (เช่น นโยบายบัคเก็ตของ Amazon S3) นอกจากนี้ ลูกค้าสามารถใช้เทคนิคอื่นๆ ภายใน Amazon EC2 Instance ได้

แนวทางปฏิบัติ:

- กำหนดค่าข้อมูลประจำตัวโดยใช้สิทธิ์การใช้งานระดับต่ำสุด
- การแบ่งแยกหน้าที่
- การตรวจสอบสิทธิ์เป็นระยะ
- กำหนดความต้องการทรัพยากรสำหรับการเรียกใช้ API ที่ไวต่อความปลอดภัย เช่น การกำหนดให้มีการรับรองความถูกต้องแบบ MFA และการเข้ารหัส
- ระบุและใช้งานข้อกำหนดเฉพาะบริการ
- ใช้โซลูชันจาก AWS Marketplace

SEC 9. คุณจะปกป้องความสมบูรณ์ของระบบปฏิบัติการใน Amazon EC2 Instance อย่างไร

การควบคุมแบบดั้งเดิมอีกอย่างหนึ่ง ได้แก่ การปกป้องความสมบูรณ์ของระบบปฏิบัติการ ซึ่งสามารถทำได้ง่ายใน EC2 ด้วยเทคนิคการใช้โฮสต์แบบดั้งเดิม (ได้แก่ OSSEC, Tripwire, Trend Micro Deep Security เป็นต้น)

แนวทางปฏิบัติ:

- ใช้ตัวควบคุมความสมบูรณ์ของไฟล์สำหรับ EC2 Instance

- ใช้ตัวควบคุมการตรวจหาการบุกรุกผ่านโฮสต์สำหรับ EC2 Instance
- ใช้โซลูชันจาก AWS Marketplace หรือ APN Partner
- ใช้ AMI แบบกำหนดเองหรือเครื่องมือจัดการการกำหนดค่า (เช่น Puppet หรือ Chef) ซึ่งมีการรักษาความปลอดภัยตามค่าเริ่มต้น

SEC 10. คุณจะเก็บข้อมูลและวิเคราะห์ล็อก AWS อย่างไร

การเก็บบันทึกล็อกนั้นสำคัญต่อการตรวจสอบเหตุการณ์ต่างๆ ตั้งแต่ประสิทธิภาพการทำงานไปจนถึงความปลอดภัย แนวทางปฏิบัติในปัจจุบัน ได้แก่ การย้ายล็อกออกจากต้นทางไปยังระบบประมวลผลล็อกเป็นระยะ (เช่น CloudWatch Logs, Splunk, Papertrail) หรือจัดเก็บในบัคเก็ต Amazon S3 เพื่อประมวลผลต่อไปตามความต้องการของธุรกิจ แหล่งที่มาโดยทั่วไปของล็อก ได้แก่ ล็อกที่เกี่ยวข้องกับผู้ใช้และ API ของ AWS (เช่น AWS CloudTrail), ล็อกเฉพาะบริการ AWS ที่เกี่ยวข้อง (เช่น Amazon S3, Amazon CloudFront), ล็อกที่สร้างขึ้นโดยระบบปฏิบัติการ และล็อกเฉพาะแอปพลิเคชันภายนอก คุณสามารถใช้ Amazon CloudWatch Logs เพื่อตรวจสอบ จัดเก็บ และเข้าถึงไฟล์ล็อกได้จาก Amazon EC2 Instance, AWS CloudTrail หรือแหล่งที่มาอื่นๆ

แนวทางปฏิบัติ:

- AWS CloudTrail
- ล็อกจาก Amazon CloudWatch
- ล็อกจาก Elastic Load Balancing (ELB)
- ล็อกฟิลเตอร์จาก Amazon Virtual Private Cloud (VPC)
- ล็อกบัคเก็ตจาก Amazon S3
- แหล่งข้อมูลล็อกอื่นๆ เฉพาะบริการ AWS
- ล็อกจากระบบปฏิบัติการหรือแอปพลิเคชันภายนอก
- ใช้โซลูชันจาก AWS Marketplace

เสาหลักด้านความน่าเชื่อถือ (REL)

REL 1. คุณจะจัดการค่าจำกัดบริการ AWS สำหรับบัญชีของคุณอย่างไร

บัญชี AWS ได้รับการจัดเตรียมโดยใช้ค่าจำกัดบริการเริ่มต้นเพื่อป้องกันไม่ให้ผู้ใช้ใหม่เตรียมใช้งานทรัพยากร

เกินความต้องการโดยที่ไม่ได้ตั้งใจ ลูกค้า AWS ควรประเมินความต้องการบริการ AWS ของตนและขอเปลี่ยนแปลงค่าจำกัดต่างๆ ตามความเหมาะสมกับแต่ละภูมิภาคที่ใช้งาน

แนวทางปฏิบัติ:

- **ตรวจสอบและจัดการค่าจำกัด** ประเมินการใช้งานที่จะเกิดขึ้นใน AWS เพิ่มค่าจำกัดตามภูมิภาคอย่างเหมาะสม และรองรับการใช้งานที่จะเติบโตขึ้นตามแผนที่วางไว้
- **กำหนดการตรวจสอบแบบอัตโนมัติ** ใช้งานเครื่องมือ เช่น SDK เพื่อให้แจ้งเตือนคุณเมื่อมีจำนวนที่ใกล้ถึงค่าจำกัด
- **ตระหนักถึงค่าจำกัดบริการแบบตายตัว** คำนึงถึงค่าจำกัดบริการแบบที่ไม่สามารถเปลี่ยนแปลงได้ และออกแบบสถาปัตยกรรมโดยพิจารณาข้อมูลเหล่านี้

REL 2. คุณจะวางแผนโทโพโลยีเครือข่ายบน AWS อย่างไร

แอปพลิเคชันสามารถปรากฏในสภาพแวดล้อมมากกว่าหนึ่งระบบ ไม่ว่าจะเป็น EC2 Classic, VPC หรือ VPC ตามค่าเริ่มต้น ข้อควรพิจารณาเกี่ยวกับเครือข่าย เช่น การเชื่อมต่อระบบ การจัดการ EIP/ที่อยู่ IP สาธารณะ, การจัดการ VPC/ที่อยู่ส่วนตัว และการแปลงชื่อเป็นสิ่งสำคัญต่อการใช้ประโยชน์จากทรัพยากรในระบบคลาวด์ การปรับใช้ที่วางแผนอย่างดีและจัดทำเป็นเอกสารจะช่วยลดความเสี่ยงจากพื้นที่ทับซ้อนและการแย่งช่องสัญญาณ

แนวทางปฏิบัติ:

- **การเชื่อมต่อที่มีความพร้อมใช้งานสูงไปยัง AWS** หลายวงจร DX, หลายช่องทางการเชื่อมต่อ VPN และเครื่องมือต่างๆ จาก AWS Marketplace
- **การเชื่อมต่อที่มีความพร้อมใช้งานสูงไปยังระบบ** โหลดบาลานซ์และ/หรือพร็อกซีแบบพร้อมใช้งานสูง โซลูชันที่ใช้ DNS และเครื่องมือต่างๆ จาก AWS Marketplace เป็นต้น
- **ช่วง IP ส่วนตัวที่ไม่ทับซ้อนกัน** การใช้ช่วงที่อยู่ IP และซับเน็ตในระบบคลาวด์ส่วนตัวแบบเสมือนไม่ควรทับซ้อนกัน และไม่ควรทับซ้อนกับสภาพแวดล้อมคลาวด์อื่นๆ หรือสภาพแวดล้อมภายในองค์กรของคุณ
- **การจัดสรรซับเน็ตของ IP** ช่วงที่อยู่ IP ของ Amazon VPC ควรใหญ่พอที่จะรองรับความต้องการของแอปพลิเคชัน ซึ่งรวมถึงการแยกองค์ประกอบเพื่อขยายการใช้ในอนาคต และการจัดสรรที่อยู่ IP ให้กับซับเน็ตระหว่าง Availability Zone ต่างๆ

REL 3. คุณมีพาธการเลื่อนระดับเพื่อรองรับปัญหาทางเทคนิคหรือไม่

ลูกค้าควรใช้ประโยชน์จาก AWS Support หรือคู่ค้า AWS การติดต่อรับบริการเป็นประจำจะช่วยในการระบุและป้องกันปัญหาซึ่งเป็นที่ทราบกันดี ปิดช่องโหว่ด้านองค์ความรู้ และข้อกังวลด้านการออกแบบ ซึ่งจะลดความเสี่ยงจากความล้มเหลวในการนำไปใช้งาน รวมทั้งความขัดข้องที่เกิดขึ้นในวงกว้าง

แนวทางปฏิบัติ:

- **วางแผน**ความร่วมมือ/ความสัมพันธ์ที่ต่อเนื่องกับ AWS Support หรือ APN Partner
- **ใช้ประโยชน์จาก API ของ AWS Support** รวมระบบ API ของ AWS Support เข้ากับระบบตรวจสอบและแจ้งปัญหาของคุณ

REL 4. ระบบของคุณมีวิธีรองรับความต้องการใช้งานที่เปลี่ยนแปลงไปอย่างไร

ระบบที่ปรับขนาดใช้งานได้จะมีความยืดหยุ่นในการเพิ่มและนำทรัพยากรออกโดยอัตโนมัติเพื่อให้มีปริมาณสอดคล้องกับความต้องการมากที่สุดตามเวลาที่กำหนด

แนวทางปฏิบัติ:

- **การปรับขนาดใช้งานอัตโนมัติ** ใช้บริการที่สามารถปรับขนาดใช้งานได้โดยอัตโนมัติ เช่น Amazon S3, Amazon CloudFront, Auto Scaling, Amazon DynamoDB, AWS Elastic Beanstalk เป็นต้น
- **ทดสอบโหลด** ปรับใช้ระเบียบวิธีการทดสอบปริมาณงานเพื่อประเมินว่าการปรับขนาดใช้งานเป็นไปตามข้อกำหนดของแอปพลิเคชันหรือไม่

REL 5. คุณจะตรวจสอบทรัพยากร AWS ได้อย่างไร

ล็อกและตัววัดผลต่างๆ เป็นเครื่องมือที่มีประสิทธิภาพในการรวบรวมข้อมูลเชิงลึกเกี่ยวกับความสมบูรณ์ของแอปพลิเคชัน คุณสามารถกำหนดค่าระบบให้ตรวจสอบล็อกและตัววัดผล แล้วส่งการแจ้งข้อมูลเมื่อมีค่าที่เกินเกณฑ์หรือเมื่อเกิดเหตุการณ์สำคัญ ตามหลักการแล้ว เมื่อประสิทธิภาพของระบบมีค่าถึงเกณฑ์ประสิทธิภาพต่ำหรือเกิดความล้มเหลว ระบบจะได้รับการออกแบบให้ซ่อมแซมตัวเองโดยอัตโนมัติหรือปรับขนาดเพื่อตอบสนองต่อเหตุการณ์

แนวทางปฏิบัติ:

- **การตรวจสอบ** ตรวจสอบแอปพลิเคชันของคุณด้วย Amazon CloudWatch หรือเครื่องมือจากบริษัทอื่นๆ
- **การแจ้งข้อมูล** วางแผนรับการแจ้งเตือนเมื่อเกิดเหตุการณ์สำคัญ
- **การตอบสนองอัตโนมัติ** ใช้ระบบอัตโนมัติเพื่อดำเนินการเมื่อตรวจพบข้อผิดพลาด เช่น แทนที่คอมโพเนนต์ที่เกิดความล้มเหลว
- **การประเมิน** ดำเนินการประเมินความถี่ของระบบตามเหตุการณ์สำคัญเพื่อประเมินผลสถาปัตยกรรมระบบ

REL 6. คุณจะจัดการกับการเปลี่ยนแปลงอย่างไร

การจัดการกับการเปลี่ยนแปลงทรัพยากรและแอปพลิเคชัน AWS ที่เตรียมใช้งานเป็นสิ่งจำเป็นเพื่อให้มั่นใจได้ว่าแอปพลิเคชันและสภาพแวดล้อมปฏิบัติการใช้งานซอฟต์แวร์ที่ระบบรู้จัก และสามารถแก้ไขหรือแทนที่ในลักษณะที่ควบคุมได้

แนวทางปฏิบัติ:

- **CM อัตโนมัติ** ใช้ระบบอัตโนมัติในการปรับใช้/แก้ไขความบกพร่อง

REL 7. คุณจะสำรองข้อมูลได้อย่างไร

สำรองข้อมูล แอปพลิเคชัน และสภาพแวดล้อมปฏิบัติการ (ในฐานะระบบปฏิบัติการที่กำหนดค่าด้วยแอปพลิเคชัน) เพื่อให้เป็นไปตามข้อกำหนดเวลาเฉลี่ยในการกู้คืน (MTTR) และเวลากู้คืนที่ระบบยอมรับได้ (RPO)

แนวทางปฏิบัติ:

- **มีการสำรองข้อมูล** สำรองข้อมูลสำคัญโดยใช้ Amazon S3, สแนปช็อต Amazon EBS หรือซอฟต์แวร์จากบริษัทอื่นๆ เพื่อให้เป็นไปตามเกณฑ์ RPO
- **สำรองข้อมูลด้วยระบบอัตโนมัติ** ใช้คุณสมบัติของ AWS, โซลูชันจาก AWS Marketplace หรือซอฟต์แวร์ของบริษัทอื่นๆ เพื่อสำรองข้อมูลโดยอัตโนมัติ

- **รักษาความปลอดภัยและ/หรือเข้ารหัสข้อมูลสำรอง** ดูรายงานแนวทางปฏิบัติการรักษาความปลอดภัยของ AWS
- **ทดสอบการกู้คืนเป็นระยะ** ตรวจสอบว่าการดำเนินขั้นตอนการสำรองข้อมูลเป็นไปตามเกณฑ์ RTO และ RPO โดยใช้การทดสอบการกู้คืน

REL 8. ระบบของคุณจะรับมือกับความล้มเหลวที่เกิดขึ้นกับคอมโพเนนต์อย่างไร

แอปพลิเคชันของคุณมีความต้องการโดยนัยหรือโดยแจ้งเกี่ยวกับความพร้อมใช้งานสูงและเวลาเฉลี่ยต่ำในการกู้คืน (MTTR) หรือไม่ ถ้ามี ให้ออกแบบสถาปัตยกรรมแอปพลิเคชันให้มีความยืดหยุ่นและกระจายแอปพลิเคชันเหล่านี้เพื่อรับมือกับข้อขัดข้อง เพื่อให้มีความพร้อมใช้งานในระดับที่สูงขึ้น การกระจายนี้จะต้องครอบคลุมตำแหน่งที่ตั้งทางกายภาพที่ต่างกัน ออกแบบแต่ละเลเยอร์ (เช่น เว็บเซิร์ฟเวอร์หรือฐานข้อมูล) เพื่อความยืดหยุ่น ซึ่งรวมถึงการตรวจสอบระบบซ่อมแซมตัวเอง และการแจ้งเตือนข้อขัดข้องและความล้มเหลวที่เป็นเหตุการณ์สำคัญ

แนวทางปฏิบัติ:

- **โหลดบาลานซ์** ใช้โหลดบาลานซ์เซอร์หน้าพูลทรัพยากร
- **การใช้งานแบบหลาย AZ/ภูมิภาค** กระจายแอปพลิเคชันในหลาย Availability Zone/ภูมิภาค
- **การซ่อมแซมอัตโนมัติ** ใช้ความสามารถในการตรวจจับความล้มเหลวและดำเนินการแก้ไขด้วยระบบอัตโนมัติ
- **การตรวจสอบ** ตรวจสอบความสมบูรณ์ของระบบอย่างต่อเนื่อง
- **การแจ้งข้อมูล** วางแผนรับการแจ้งเตือนเหตุการณ์สำคัญ

REL 9. คุณจะวางแผนการกู้คืนระบบอย่างไร

การกู้คืนข้อมูลมีความสำคัญอย่างยิ่งหากต้องการคืนค่าข้อมูลจากวิธีการสำรองข้อมูล ข้อกำหนดและการดำเนินการของคุณเกี่ยวกับวัตถุประสงค์ ทรัพยากร ที่ตั้ง และฟังก์ชันของข้อมูลนี้จะต้องสอดคล้องกับวัตถุประสงค์ของ RTO และ RPO

แนวทางปฏิบัติ:

- **กำหนดวัตถุประสงค์** กำหนด RTO และ RPO

- **การกู้คืนเมื่อเกิดความเสียหาย** กำหนดกลยุทธ์ DR
- **การเปลี่ยนแปลงค่าที่กำหนด (Configuration Drift)** ตรวจสอบว่า Amazon Machine Images (AMIs) และการกำหนดค่าระบบเป็นปัจจุบันเสมอในไซต์/ภูมิภาค DR
- **ค่าจำกัดบริการ** ขอเพิ่มค่าจำกัดบริการสำหรับไซต์ DR เพื่ออำนวยความสะดวกให้กับการทำหน้าที่แทนของระบบสำรอง
- **ทดสอบและตรวจสอบความถูกต้องของ DR** ทดสอบระบบสำรองสำหรับ DR เป็นประจำเพื่อตรวจสอบค่า RTO และ RPO ให้เป็นไปตามที่กำหนด
- **ดำเนินการกู้คืนด้วยระบบอัตโนมัติ** ใช้ AWS และ/หรือเครื่องมืออื่นๆ เพื่อกู้คืนระบบโดยอัตโนมัติ

เสาหลักด้านประสิทธิภาพ

PERF 1. คุณจะเลือกประเภทอินสแตนซ์ที่เหมาะกับระบบของคุณอย่างไร

Amazon EC2 มีอินสแตนซ์หลากหลายประเภทที่ผ่านการปรับประสิทธิภาพให้รองรับกรณีการใช้งานที่แตกต่างกัน อินสแตนซ์ประเภทต่างๆ ประกอบด้วยชุดค่าที่ต่างกันของ CPU, หน่วยความจำ พื้นที่จัดเก็บ และความจุระบบเครือข่าย เพื่อให้คุณเลือกทรัพยากรที่ผสมผสานกันสำหรับแอปพลิเคชันได้อย่างยืดหยุ่นและเหมาะสม แต่ละประเภทอินสแตนซ์ประกอบด้วยขนาดอินสแตนซ์หนึ่งหรือหลายขนาด รองรับการปรับขนาดใช้งานทรัพยากรให้เหมาะกับความต้องการของเวิร์กโหลดเป้าหมาย AWS รองรับสถาปัตยกรรมแบบไม่ใช้เซิร์ฟเวอร์ เช่น AWS Lambda ซึ่งสามารถเปลี่ยนแปลงประสิทธิภาพการทำงานของเวิร์กโหลดอย่างเห็นได้ชัด

แนวทางปฏิบัติ:

- **นโยบาย/สถาปัตยกรรมอ้างอิง** เลือกประเภทและขนาดอินสแตนซ์โดยอิงตามความต้องการทรัพยากรที่คาดการณ์ไว้กับมาตรฐานการกำกับดูแลภายในองค์กร
- **ต้นทุน/งบประมาณ** เลือกประเภทและขนาดอินสแตนซ์โดยอิงตามความต้องการทรัพยากรที่คาดการณ์ไว้กับการควบคุมต้นทุนภายในองค์กร
- **การเทียบวัดมาตรฐาน** ทดสอบปริมาณเวิร์กโหลดที่ทราบใน AWS และใช้การทดสอบนี้เพื่อประมาณตัวเลือกที่ดีที่สุด กล่าวคือ ทดสอบมาตรฐานประสิทธิภาพที่ทราบเทียบกับเวิร์กโหลดที่ทราบ
- **คำแนะนำจาก AWS หรือจากสมาชิกของ AWS Partner Network (APN)** ตัดสินใจโดยอาศัยคำแนะนำตามแนวทางปฏิบัติที่ดีที่สุด
- **ทดสอบโหลด** ปรับใช้เวอร์ชันล่าสุดของระบบคุณใน AWS โดยใช้ประเภทและขนาดอินสแตนซ์ที่

ต่างกัน ใช้การตรวจสอบเพื่อบันทึกข้อมูลตัววัดประสิทธิภาพ แล้วตัดสินใจเลือกโดยอาศัยการคำนวณประสิทธิภาพ/ต้นทุน

PERF 2. คุณมีวิธีการอย่างไรในการตรวจสอบว่าคุณใช้ประเภทอินสแตนซ์ที่เหมาะสมที่สุดอยู่ขณะที่มีประเภทและคุณสมบัติใหม่ๆ ของอินสแตนซ์ออกมาให้บริการ

AWS รับฟังข้อคิดเห็นจากลูกค้าและพัฒนาประเภทและขนาดอินสแตนซ์ใหม่ๆ อย่างต่อเนื่อง รวมทั้งรองรับองค์ประกอบที่ผสมผสานใหม่ๆ สำหรับ CPU, หน่วยความจำ, พื้นที่จัดเก็บ และความจุระบบเครือข่าย ซึ่งหมายความว่า AWS อาจนำเสนอประเภทอินสแตนซ์ใหม่ๆ เพื่อมอบประสิทธิภาพการทำงานที่ดียิ่งขึ้นกว่าประเภทอินสแตนซ์ที่คุณเลือกเมื่อเริ่มแรก

แนวทางปฏิบัติ:

- **ประเมิน** เลือกประเภทและขนาดอินสแตนซ์ใหม่อีกครั้งตามวงจรการใช้งานโดยอิงจากความต้องการทรัพยากรที่คาดการณ์ไว้
- **การเทียบวัดมาตรฐาน** หลังจากมีการนำเสนออินสแตนซ์ใหม่แต่ละประเภท ให้ดำเนินการทดสอบปริมาณเวิร์กโหลดที่ทราบใน AWS และใช้การทดสอบนี้เพื่อประมาณตัวเลือกที่ดีที่สุด
- **การทดสอบโหลด** หลังจากมีการนำเสนออินสแตนซ์ใหม่ๆ แต่ละประเภท ให้ปรับใช้เวอร์ชันล่าสุดของระบบใน AWS และใช้การตรวจสอบเพื่อบันทึกข้อมูลตัววัดประสิทธิภาพ แล้วตัดสินใจเลือกโดยอาศัยการคำนวณประสิทธิภาพ/ต้นทุน

PERF 3. คุณมีวิธีการอย่างไรในการตรวจสอบอินสแตนซ์หลังจากเปิดใช้งานเพื่อให้แน่ใจว่าอินสแตนซ์ทำงานได้ตามที่คาดหวัง

ประสิทธิภาพของระบบอาจเสื่อมถอยลงเมื่อเวลาผ่านไปเนื่องมาจากปัจจัยภายในและ/หรือปัจจัยภายนอก การตรวจสอบประสิทธิภาพของระบบจะช่วยให้คุณระบุถึงการเสื่อมถอยลงนี้ได้ และแก้ไขปัจจัยภายในหรือภายนอก (เช่น ระบบปฏิบัติการหรือโหลดงานของแอปพลิเคชัน)

แนวทางปฏิบัติ:

- **การตรวจสอบโดย Amazon CloudWatch** ใช้ CloudWatch เพื่อตรวจสอบอินสแตนซ์ต่างๆ
- **การตรวจสอบโดยเครื่องมือภายนอก** ใช้เครื่องมืออื่นๆ ในการตรวจสอบระบบ

- **การประเมินตามระยะ** ตรวจดูแดชบอร์ดการตรวจสอบของคุณเป็นระยะ
- **การแจ้งให้ทราบตามข้อมูลแจ้งเตือน** รับการแจ้งเตือนอัตโนมัติจากระบบตรวจสอบเมื่อค่าของตัววัดผลอยู่นอกช่วงที่ปลอดภัย
- **การดำเนินการตามทริกเกอร์** การแจ้งเตือนจะทำให้เกิดการดำเนินการอัตโนมัติเพื่อแก้ไขหรือเลื่อนระดับปัญหา

PERF 4. คุณมีวิธีการอย่างไรในการตรวจสอบว่าจำนวนอินสแตนซ์เป็นไปตามความต้องการ

ปริมาณความต้องการที่เพิ่มในระบบมักแตกต่างกันไปตามวงจรที่ต่างกัน: วงจรผลิตภัณฑ์ เช่น การเปิดตัวหรือการเติบโต วงจรชั่วคราว เช่น เวลาของวัน วันในสัปดาห์หรือเดือน หรือวงจรที่ไม่สามารถคาดการณ์ได้ เช่น ทัศนวิสัยของโซเชียลมีเดีย และวงจรที่คาดการณ์ได้ เช่น ตอนในรายการโทรทัศน์ อินสแตนซ์ที่ไม่เพียงพอกับเวิร์กโหลดอาจทำให้ผู้ใช้ได้รับประสิทธิภาพลดลง และผลที่แย่ที่สุดคืออาจทำให้ระบบล้มเหลว

แนวทางปฏิบัติ:

- **วางแผน** วางแผนโดยอาศัยตัววัดผลและ/หรือเหตุการณ์ที่วางแผนไว้
- **ระบบอัตโนมัติ - สคริปต์** ใช้เครื่องมือสำหรับการจัดการอัตโนมัติ
- **ระบบอัตโนมัติ - Auto Scaling** ใช้ Auto Scaling สำหรับการจัดการอัตโนมัติ

PERF 5. คุณจะเลือกโซลูชันการจัดเก็บข้อมูลที่เหมาะกับระบบของคุณอย่างไร

AWS ได้รับการออกแบบให้มีพื้นที่เก็บข้อมูลต้นทุนต่ำ ที่มาพร้อมกับความทนทานและความพร้อมใช้งานในระดับสูง AWS มีตัวเลือกพื้นที่เก็บข้อมูลสำหรับการสำรองข้อมูล การเก็บถาวร และการกู้คืนจากความเสียหายรุนแรง ตลอดจนพื้นที่เก็บข้อมูลบล็อค ไฟล์ และออบเจ็กต์

แนวทางปฏิบัติ:

- **นโยบาย/สถาปัตยกรรมอ้างอิง** เลือกโซลูชันและคุณสมบัติการจัดเก็บข้อมูลโดยอิงตามความต้องการทรัพยากรที่คาดการณ์ไว้กับมาตรฐานการกำกับดูแลภายในองค์กร
- **ต้นทุน/งบประมาณ** เลือกโซลูชันและคุณสมบัติการจัดเก็บข้อมูลโดยอิงตามความต้องการทรัพยากรที่คาดการณ์ไว้กับการควบคุมต้นทุนภายในองค์กร

- **การเทียบวัดมาตรฐาน** ทดสอบปริมาณเวิร์กโหลดที่ทราบใน AWS และใช้การทดสอบนี้เพื่อประมาณตัวเลือกที่ดีที่สุด กล่าวคือ ทดสอบมาตรฐานประสิทธิภาพที่ทราบเทียบกับเวิร์กโหลดที่ทราบ
- **คำแนะนำจาก AWS หรือ APN Partner** เลือกโซลูชันโดยอาศัยคำแนะนำตามแนวทางปฏิบัติที่ดีที่สุด
- **การทดสอบโหลด** ปรับใช้เวอร์ชันล่าสุดของระบบคุณใน AWS โดยใช้โซลูชันการจัดเก็บข้อมูลที่ต่างกัน ใช้การตรวจสอบเพื่อบันทึกข้อมูลตัววัดประสิทธิภาพ แล้วตัดสินใจเลือกโดยอาศัยการคำนวณประสิทธิภาพ/ต้นทุน

PERF 6. คุณมีวิธีการอย่างไรในการตรวจสอบว่าคุณใช้โซลูชันการจัดเก็บข้อมูลที่เหมาะสมที่สุดอยู่ขณะที่มีโซลูชันและคุณสมบัติใหม่ๆ ในการจัดเก็บข้อมูลออกมาให้บริการ

AWS รับฟังข้อคิดเห็นจากลูกค้าและพัฒนาโซลูชันและคุณสมบัติใหม่ในการจัดเก็บข้อมูลอย่างต่อเนื่อง รวมทั้งรองรับองค์ประกอบที่ผสมผสานใหม่ๆ สำหรับความจุ อัตราความเร็ว และความคงทน ซึ่งหมายความว่า AWS อาจนำเสนอโซลูชันการจัดเก็บข้อมูลใหม่ๆ เพื่อมอบประสิทธิภาพการทำงานที่ดียิ่งขึ้นกว่าโซลูชันที่คุณเลือกเมื่อเริ่มแรก

แนวทางปฏิบัติ:

- **ประเมิน** เลือกโซลูชันและคุณสมบัติการจัดเก็บข้อมูลใหม่อีกครั้งตามวงจรการใช้งานโดยอิงจากความต้องการทรัพยากรที่คาดการณ์ไว้
- **การเทียบวัดมาตรฐาน** หลังจากมีการนำเสนอโซลูชันและคุณสมบัติใหม่ๆ ในการจัดเก็บข้อมูลใหม่ ให้ดำเนินการทดสอบปริมาณเวิร์กโหลดที่ทราบใน AWS และใช้การทดสอบนี้เพื่อประมาณตัวเลือกที่ดีที่สุด
- **การทดสอบโหลด** หลังจากมีการนำเสนอโซลูชันใหม่ๆ ในการจัดเก็บข้อมูล ให้ปรับใช้เวอร์ชันล่าสุดของระบบใน AWS และใช้การตรวจสอบเพื่อบันทึกข้อมูลตัววัดประสิทธิภาพ แล้วตัดสินใจเลือกโดยอาศัยการคำนวณประสิทธิภาพ/ต้นทุน

PERF 7. คุณมีวิธีการอย่างไรในการตรวจสอบโซลูชันการจัดเก็บข้อมูลเพื่อให้แน่ใจว่าโซลูชันนั้นทำงานได้ตามที่คาดหวัง

ประสิทธิภาพของระบบอาจเสื่อมถอยลงเมื่อเวลาหรือช่วงเวลาผ่านไปเนื่องมาจากปัจจัยภายในหรือภายนอก การตรวจสอบประสิทธิภาพของระบบจะช่วยให้คุณระบุถึงการเสื่อมถอยลงนี้และแก้ไขปัจจัยภายในหรือภายนอก

แนวทางปฏิบัติ:

- **การตรวจสอบโดย Amazon CloudWatch** ใช้ CloudWatch เพื่อตรวจสอบระบบจัดเก็บข้อมูล
- **การตรวจสอบโดยเครื่องมือภายนอก** ใช้เครื่องมืออื่นๆ ในการตรวจสอบระบบจัดเก็บข้อมูล
- **การประเมินตามระยะ** ตรวจดูแดชบอร์ดการตรวจสอบของคุณเป็นระยะ
- **การตรวจสอบตามข้อมูลแจ้งเตือน** วางแผนให้ระบบตรวจสอบแจ้งเตือนคุณโดยอัตโนมัติเมื่อค่าของตัววัดผลอยู่นอกช่วงที่ปลอดภัย
- **การดำเนินการตามทริกเกอร์** วางแผนการแจ้งเตือนเพื่อให้เกิดการดำเนินการอัตโนมัติเพื่อแก้ไขหรือเลื่อนระดับปัญหา

PERF 8. คุณมีวิธีการอย่างไรในการตรวจสอบว่าความจุและอัตราความเร็วของโซลูชันการจัดเก็บข้อมูลนั้นตรงกับความต้องการ

ปริมาณความต้องการที่เพิ่มในระบบมักแตกต่างกันไปตามวงจรที่ต่างกัน: วงจรผลิตภัณฑ์ เช่น การเปิดตัวหรือการเติบโต วงจรชั่วคราว เช่น เวลาของวัน วันในสัปดาห์หรือเดือน หรือวงจรที่ไม่สามารถคาดการณ์ได้ เช่น ทัศนวิสัยของโซเชียลมีเดีย และวงจรที่คาดการณ์ได้ เช่น ตอนในรายการโทรทัศน์ ความจุหรืออัตราความเร็วของระบบจัดเก็บข้อมูลที่ไม่เพียงพอกับเวิร์กโหลดอาจทำให้ผู้ใช้ได้รับประสิทธิภาพลดลง และผลที่แย่ที่สุดคืออาจทำให้ระบบล้มเหลว

แนวทางปฏิบัติ:

- **ตอบสนองเชิงรับ** จัดการด้วยตนเองโดยอาศัยตัววัดผลต่างๆ
- **วางแผน** วางแผนความจุและอัตราความเร็วในอนาคตตามตัววัดผลและ/หรือเหตุการณ์ที่วางแผนไว้
- **ระบบอัตโนมัติ** ดำเนินการอัตโนมัติตามตัววัดผลต่างๆ

PERF 9. คุณจะเลือกโซลูชันฐานข้อมูลที่เหมาะกับระบบของคุณอย่างไร

โซลูชันฐานข้อมูลที่เหมาะสมที่สุดกับระบบใดระบบหนึ่งอาจแตกต่างกันไปตามความต้องการด้านความสอดคล้องของข้อมูล ความพร้อมใช้งาน การทนต่อการจัดเก็บที่แยกออกเป็นเครือข่าย และเวลาแฝง หลายๆ ระบบใช้โซลูชันฐานข้อมูลที่ต่างกันสำหรับระบบย่อยต่างๆ และใช้คุณสมบัติที่แตกต่างเพื่อปรับปรุงประสิทธิภาพ การเลือกโซลูชันและคุณสมบัติของฐานข้อมูลที่ไม่เหมาะกับเวิร์กโหลดของระบบอาจเป็นสาเหตุ

ให้ประสิทธิภาพการทำงานลดลง

แนวทางปฏิบัติ:

- **นโยบาย/สถาปัตยกรรมอ้างอิง** เลือกโซลูชันและคุณสมบัติของฐานข้อมูลโดยอิงตามความต้องการทรัพยากรที่คาดการณ์ไว้กับมาตรฐานการกำกับดูแลภายในองค์กร
- **ต้นทุน/งบประมาณ** เลือกโซลูชันและคุณสมบัติของฐานข้อมูลโดยอิงตามความต้องการทรัพยากรที่คาดการณ์ไว้กับการควบคุมต้นทุนภายในองค์กร
- **การเทียบวัดมาตรฐาน** ทดสอบปริมาณเวิร์กโหลดที่ทราบใน AWS และใช้การทดสอบนี้เพื่อประมาณตัวเลือกที่ดีที่สุด กล่าวคือ ทดสอบมาตรฐานประสิทธิภาพที่ทราบเทียบกับเวิร์กโหลดที่ทราบ
- **คำแนะนำจาก AWS หรือ APN Partner** เลือกโซลูชันโดยอาศัยคำแนะนำตามแนวทางปฏิบัติที่ดีที่สุด
- **การทดสอบโหลด** ปรับใช้เวอร์ชันล่าสุดของระบบใน AWS โดยใช้โซลูชันและคุณสมบัติฐานข้อมูลที่ต่างกัน ใช้การตรวจสอบเพื่อบันทึกข้อมูลตัววัดประสิทธิภาพ แล้วตัดสินใจเลือกโดยอาศัยการคำนวณประสิทธิภาพ/ต้นทุน

PERF 10. คุณมีวิธีการตรวจสอบอย่างไรว่าคุณใช้โซลูชันและคุณสมบัติฐานข้อมูลที่เหมาะสมที่สุดอยู่ขณะที่มีโซลูชันและคุณสมบัติใหม่ๆ ของฐานข้อมูลออกมาให้บริการ

AWS รับฟังข้อคิดเห็นจากลูกค้าและพัฒนาโซลูชันและคุณสมบัติของฐานข้อมูลใหม่อย่างต่อเนื่อง รวมทั้งรองรับองค์ประกอบที่ผสมผสานใหม่ๆ เพื่อความสอดคล้องของข้อมูล ความพร้อมใช้งาน การทนต่อการจัดเก็บที่แยกออกเป็นเครือข่าย และเวลาแฝง ซึ่งหมายความว่า AWS อาจนำเสนอโซลูชันหรือคุณสมบัติของฐานข้อมูลใหม่ๆ เพื่อมอบประสิทธิภาพการทำงานที่ดียิ่งขึ้นกว่าโซลูชันที่คุณเลือกเมื่อเริ่มแรก

แนวทางปฏิบัติ:

- **ประเมิน** เลือกโซลูชันและคุณสมบัติของฐานข้อมูลใหม่อีกครั้งตามวงจรการใช้งานโดยอิงจากความต้องการทรัพยากรที่คาดการณ์ไว้
- **การเทียบวัดมาตรฐาน** หลังจากมีการนำเสนอโซลูชันหรือคุณสมบัติของฐานข้อมูลใหม่ๆ แต่ละอย่าง ให้ดำเนินการทดสอบปริมาณเวิร์กโหลดที่ทราบใน AWS และใช้การทดสอบนี้เพื่อประมาณตัวเลือกที่ดีที่สุด
- **การทดสอบโหลด** หลังจากมีการนำเสนอโซลูชันหรือคุณสมบัติฐานข้อมูลใหม่ๆ แต่ละอย่าง ให้

ปรับใช้เวอร์ชันล่าสุดของระบบใน AWS และใช้การตรวจสอบเพื่อบันทึกข้อมูลตัววัดประสิทธิภาพ แล้วตัดสินใจเลือกโดยอาศัยการคำนวณประสิทธิภาพ/ต้นทุน

PERF 11. คุณมีวิธีการอย่างไรในการตรวจสอบฐานข้อมูลเพื่อให้แน่ใจว่าฐานข้อมูลนั้นทำงานได้ตามที่คาดหวัง

ประสิทธิภาพของระบบอาจเสื่อมถอยลงเมื่อเวลาผ่านไปเนื่องมาจากปัจจัยภายในหรือภายนอก การตรวจสอบประสิทธิภาพของระบบจะช่วยให้คุณระบุถึงการเสื่อมถอยลงนี้และแก้ไขปัจจัยภายในหรือภายนอก

แนวทางปฏิบัติ:

- **การตรวจสอบโดย Amazon CloudWatch** ใช้ CloudWatch เพื่อตรวจสอบฐานข้อมูลต่างๆ
- **การตรวจสอบโดยเครื่องมือภายนอก** ใช้เครื่องมืออื่นๆ ในการตรวจสอบฐานข้อมูล
- **การประเมินตามระยะ** ตรวจดูแดชบอร์ดการตรวจสอบของคุณเป็นระยะ
- **การแจ้งให้ทราบตามข้อมูลแจ้งเตือน** วางแผนให้ระบบตรวจสอบแจ้งเตือนคุณโดยอัตโนมัติเมื่อค่าของตัววัดผลอยู่นอกช่วงที่ปลอดภัย
- **การดำเนินการตามทริกเกอร์** วางแผนให้การแจ้งเตือนดำเนินการอัตโนมัติเพื่อแก้ไขหรือเลื่อนระดับปัญหา

PERF 12. คุณมีวิธีการอย่างไรในการตรวจสอบว่าความจุและอัตราความเร็วของฐานข้อมูลนั้นตรงกับความต้องการ

ปริมาณความต้องการที่เพิ่มในระบบมักแตกต่างกันไปตามวงจรที่ต่างกัน: วงจรผลิตภัณฑ์ เช่น การเปิดตัวหรือการเติบโต วงจรชั่วคราว เช่น เวลาของวัน วันในสัปดาห์หรือเดือน วงจรที่ไม่สามารถคาดการณ์ได้ตามที่เห็นในโซเชียลมีเดีย และวงจรที่คาดการณ์ได้ เช่น ตอนในรายการโทรทัศน์ ความจุและอัตราความเร็วของฐานข้อมูลที่ไม่เพียงพอกับเวิร์กโหลดอาจทำให้ผู้ใช้ได้รับประสิทธิภาพลดลง และผลที่แย่ที่สุดคืออาจทำให้ระบบล้มเหลว

แนวทางปฏิบัติ:

- **วางแผน** วางแผนความจุและอัตราความเร็วในอนาคตตามตัววัดผลและ/หรือเหตุการณ์ที่วางแผนไว้

- **ระบบอัตโนมัติ** ดำเนินการอัตโนมัติตามตัววัดผลต่างๆ

PERF 13. คุณจะเลือกโซลูชันการแคชข้อมูลและการตรวจหาอุปกรณ์ระยะใกล้เคียงสำหรับระบบของคุณอย่างไร

ระยะจริง ระยะเครือข่าย หรือคำขอที่ใช้เวลาดำเนินการนานอาจทำให้ระบบเกิดความล่าช้า ค่าเวลาแฝงที่ไม่สามารถระบุได้อาจหน่วงเวลาทรัพยากรระบบนานเกินความจำเป็น และเป็นสาเหตุให้ประสิทธิภาพทั้งภายในและภายนอกลดลง ในการลดเวลาแฝงนี้ ให้พิจารณาถึงประสิทธิภาพแบบครบวงจรของทั้งระบบในแง่มุมของผู้ใช้ และมองหาโอกาสในการปรับเปลี่ยนระยะทางกายภาพของทรัพยากรหรือโซลูชันการแคชข้อมูล

แนวทางปฏิบัติ:

- **นโยบาย/สถาปัตยกรรมอ้างอิง** เลือกโซลูชันการแคชข้อมูลและการตรวจหาอุปกรณ์ระยะใกล้เคียงโดยอิงตามความต้องการทรัพยากรที่คาดการณ์ไว้กับมาตรฐานการกำกับดูแลภายในองค์กร
- **ต้นทุน/งบประมาณ** เลือกโซลูชันการแคชข้อมูลและการตรวจหาอุปกรณ์ระยะใกล้เคียงโดยอิงตามความต้องการทรัพยากรที่คาดการณ์ไว้กับการควบคุมต้นทุนภายในองค์กร
- **การเทียบวัดมาตรฐาน** ทดสอบปริมาณเวิร์กโหลดที่ทราบใน AWS และใช้การทดสอบนี้เพื่อประมาณตัวเลือกที่ดีที่สุด กล่าวคือ ทดสอบมาตรฐานประสิทธิภาพที่ทราบเทียบกับเวิร์กโหลดที่ทราบ
- **คำแนะนำจาก AWS หรือ APN Partner** เลือกโซลูชันการแคชข้อมูลและการตรวจหาอุปกรณ์ระยะใกล้เคียงโดยอาศัยคำแนะนำตามแนวทางปฏิบัติที่ดีที่สุด
- **การทดสอบโหลด** ปรับใช้เวอร์ชันล่าสุดของระบบคุณใน AWS โดยใช้โซลูชันการแคชข้อมูลและการตรวจหาอุปกรณ์ระยะใกล้เคียงที่ต่างกัน ใช้การตรวจสอบเพื่อบันทึกข้อมูลตัววัดประสิทธิภาพ แล้วดำเนินการตัดสินใจโดยอาศัยการคำนวณประสิทธิภาพ/ต้นทุน

PERF 14. คุณมีวิธีการตรวจสอบอย่างไรว่าคุณใช้โซลูชันการแคชข้อมูลและการตรวจหาอุปกรณ์ระยะใกล้เคียงที่เหมาะสมที่สุดอยู่ขณะที่มีโซลูชันใหม่ๆ ออกมาให้บริการ

AWS รับฟังข้อคิดเห็นจากลูกค้าและพัฒนาโซลูชันใหม่ๆ ในการแคชข้อมูลและการตรวจหาอุปกรณ์ระยะใกล้เคียงอย่างต่อเนื่อง รวมทั้งรองรับองค์ประกอบที่ผสมผสานใหม่ๆ สำหรับการตรวจหาอุปกรณ์ระยะใกล้เคียง การแคชข้อมูล และเวลาแฝง ซึ่งหมายความว่า AWS อาจนำเสนอโซลูชันใหม่ๆ ในการแคชข้อมูลและการ

ตรวจหาอุปกรณ์ระยะใกล้เคียง เพื่อมอบประสิทธิภาพการทำงานที่ดียิ่งขึ้นกว่าโซลูชันที่คุณเลือกเมื่อเริ่มแรก หาโอกาสในการลดค่าเวลาแฝงและเพิ่มประสิทธิภาพการทำงานทั่วทั้งระบบ ตัวอย่างเช่น คุณดำเนินการปรับประสิทธิภาพของระบบให้เหมาะสมแบบครั้งเดียว หรือปรับประสิทธิภาพแบบต่อเนื่องตามความต้องการที่เปลี่ยนแปลงไปตามเวลา

แนวทางปฏิบัติ:

- **ประเมิน** เลือกโซลูชันการแคชข้อมูลและการตรวจหาอุปกรณ์ระยะใกล้เคียงใหม่อีกครั้งตามวงจรการใช้งานโดยอิงจากความต้องการทรัพยากรที่คาดการณ์ไว้
- **การเทียบวัดมาตรฐาน** หลังจากมีการนำเสนอโซลูชันการแคชข้อมูลและการตรวจหาอุปกรณ์ระยะใกล้เคียง ให้ดำเนินการทดสอบปริมาณเวิร์กโหลดที่ทราบใน AWS และใช้การทดสอบนี้เพื่อประมาณตัวเลือกที่ดีที่สุด
- **การทดสอบโหลด** หลังจากมีการนำเสนอโซลูชันการแคชข้อมูลและการตรวจหาอุปกรณ์ระยะใกล้เคียง ให้ปรับใช้เวอร์ชันล่าสุดของระบบใน AWS และใช้การตรวจสอบเพื่อบันทึกข้อมูลตัววัดประสิทธิภาพ แล้วเลือกโดยอาศัยการคำนวณประสิทธิภาพ/ต้นทุน
- **การตรวจสอบเชิงรุก – การตรวจสอบโดย Amazon Cloud Watch** ใช้ Amazon CloudWatch เพื่อตรวจติดตามโซลูชันการแคชข้อมูลและการตรวจหาอุปกรณ์ระยะใกล้เคียง
- **การตรวจสอบเชิงรุก – การตรวจสอบโดยเครื่องมืออื่นๆ** ใช้เครื่องมืออื่นๆ เพื่อตรวจติดตามโซลูชันการแคชข้อมูลและการตรวจหาอุปกรณ์ระยะใกล้เคียง
- **การแจ้งให้ทราบตามข้อมูลแจ้งเตือน** วางแผนให้ระบบตรวจสอบแจ้งเตือนคุณโดยอัตโนมัติเมื่อค่าของตัววัดผลอยู่นอกช่วงที่ปลอดภัย
- **การดำเนินการตามทริกเกอร์** วางแผนการแจ้งเตือนเพื่อให้เกิดการดำเนินการอัตโนมัติเพื่อแก้ไขหรือเลื่อนระดับปัญหา

PERF 15. คุณมีวิธีการอย่างไรในการตรวจสอบโซลูชันการแคชข้อมูลและการตรวจหาอุปกรณ์ระยะใกล้เคียงเพื่อให้แน่ใจว่าประสิทธิภาพเป็นไปตามที่คาดหวัง

ประสิทธิภาพของระบบอาจเสื่อมถอยลงเมื่อเวลาผ่านไปเนื่องมาจากปัจจัยภายในหรือภายนอก การตรวจสอบประสิทธิภาพของระบบจะช่วยให้คุณระบุถึงการเสื่อมถอยลงนี้และแก้ไขปัจจัยภายในหรือภายนอก

แนวทางปฏิบัติ:

- **การตรวจสอบโดย Amazon CloudWatch** ใช้ CloudWatch เพื่อตรวจสอบอินสแตนซ์ต่างๆ
- **การตรวจสอบโดยเครื่องมือภายนอก** ใช้เครื่องมืออื่นๆ ในการตรวจสอบระบบ
- **การประเมินตามระยะ** ตรวจดูแดชบอร์ดการตรวจสอบของคุณเป็นระยะ
- **การแจ้งให้ทราบตามข้อมูลแจ้งเตือน** วางแผนให้ระบบตรวจสอบแจ้งเตือนคุณโดยอัตโนมัติเมื่อค่าของตัววัดผลอยู่นอกช่วงที่ปลอดภัย
- **การดำเนินการตามทริกเกอร์** วางแผนการแจ้งเตือนเพื่อให้เกิดการดำเนินการอัตโนมัติเพื่อแก้ไขหรือเลื่อนระดับปัญหา

PERF 16. คุณมีวิธีการอย่างไรในการตรวจสอบว่าโซลูชันการแคชข้อมูลและการตรวจหาอุปกรณ์ระยะใกล้เคียงที่คุณมีอยู่เป็นไปตามความต้องการ

ปริมาณความต้องการที่เพิ่มในระบบมักแตกต่างกันไปตามวงจรที่ต่างกัน: วงจรผลิตภัณฑ์ เช่น การเปิดตัวหรือการเติบโต วงจรชั่วคราว เช่น เวลาของวัน วันในสัปดาห์หรือเดือน วงจรที่ไม่สามารถคาดการณ์ได้ตามที่เห็นในโซเชียลมีเดีย และวงจรที่คาดการณ์ได้ เช่น ตอนในรายการโทรทัศน์ โซลูชันการแคชข้อมูลและการตรวจหาอุปกรณ์ระยะใกล้เคียงที่ไม่เหมาะสมกับเวิร์กโหลดอาจทำให้ผู้ใช้ได้รับประสิทธิภาพลดลง และผลที่แย่ที่สุดคืออาจทำให้ระบบล้มเหลว ซึ่งจะเกิดขึ้นได้โดยเฉพาะอย่างยิ่งถ้าคุณมีหรือวางแผนที่จะมีฐานผู้ใช้ทั่วโลก

แนวทางปฏิบัติ:

- **วางแผน** วางแผนโซลูชันการแคชข้อมูลหรือการตรวจหาอุปกรณ์ระยะใกล้เคียงในอนาคตตามตัววัดผลและ/หรือเหตุการณ์ที่วางแผนไว้
- **ตรวจสอบ** ตรวจสอบความต้องการและการใช้งานแคชอย่างต่อเนื่อง
- **การประเมินตามระยะ** ประเมินความต้องการและการใช้งานแคชอย่างต่อเนื่อง

เสาหลักด้านการเพิ่มประสิทธิภาพต้นทุน (COST)

COST 1. คุณมีวิธีการอย่างไรในการตรวจสอบว่าประสิทธิภาพความจุสอดคล้องกับความต้องการแต่ไม่มากจนเกินไป

เพื่อให้ระบบสถาปัตยกรรมมีความสมดุลในแง่ของค่าใช้จ่ายและประสิทธิภาพ ให้ตรวจสอบว่ามีการใช้งานในสิ่งที่จ่ายไปอย่างครบถ้วน และหลีกเลี่ยงอินสแตนซ์ที่มีการใช้งานน้อยกว่าความเป็นจริงอย่างเห็นได้ชัด ตัววัดการใช้งานที่ผิดจากความจริงไม่ว่าในทิศทางใดจะส่งผลกระทบที่ไม่พึงประสงค์ต่อธุรกิจในแง่ของต้นทุนการดำเนินงาน (ประสิทธิภาพที่ลดลงเนื่องจากการใช้งานที่มากเกินไป) หรือค่าใช้จ่ายของ AWS ที่สูญเสียไป (เนื่องจากการเตรียมใช้งานที่มากเกินไป)

แนวทางปฏิบัติ:

- **แนวทางที่อิงตามความต้องการ** ใช้ Auto Scaling เพื่อตอบสนองความต้องการที่แตกต่างกัน
- **แนวทางที่อิงตามคิว** เรียกใช้คิวจาก Amazon Simple Queue Service (SQS) ของคุณเอง และเร่งความเร็ว/ปิดระบบอินสแตนซ์ตามความต้องการใช้งาน
- **แนวทางที่อิงตามเวลา** เช่น: ตามเวลาดวงอาทิตย์ขึ้น-ตก ปิดอินสแตนซ์การพัฒนา/ทดสอบในช่วงสุดสัปดาห์ พิจารณากำหนดเวลารายไตรมาสหรือรายปี (เช่น ช่วง Black Friday)
- **เตรียมใช้งานอย่างเหมาะสม** จัดเตรียมข้อมูลประมวลผล การกำหนดขนาด และพื้นที่จัดเก็บสำหรับบริการต่างๆ เช่น Amazon DynamoDB, Amazon EBS (IOPS ที่มีการเตรียมใช้งาน), Amazon RDS และ Amazon EMR

COST 2. คุณจะเพิ่มประสิทธิภาพการใช้งานบริการ AWS อย่างไร

ถ้าคุณใช้บริการในระดับแอปพลิเคชัน ให้ตรวจสอบว่ามีการใช้งานที่เหมาะสม ตัวอย่างเช่น ใช้นโยบายวงจรการใช้งานเพื่อควบคุมการใช้ Amazon S3 และใช้ประโยชน์จากบริการ เช่น Amazon RDS และ Amazon DynamoDB เพื่อขยายความยืดหยุ่นให้เพิ่มมากขึ้น ตรวจสอบการใช้งานที่เหมาะสม ได้แก่ ตรวจสอบความถูกต้องในการปรับใช้หลาย AZ สำหรับ Amazon RDS หรือตรวจสอบว่ามีการใช้งาน IOPS ที่ผ่านการจัดเตรียมไว้ในตาราง Amazon DynamoDB

แนวทางปฏิบัติ:

- **การเพิ่มประสิทธิภาพเฉพาะบริการ** เช่น การย่อขนาด I/O สำหรับ Amazon EBS การหลีกเลี่ยงการอัปโหลดไฟล์ขนาดเล็กเป็นจำนวนมากเกินไปลงใน Amazon S3 และการขยายการใช้อินสแตนซ์แบบประมูลราคาสำหรับ Amazon EMR

COST 3. คุณเลือกทรัพยากรที่เหมาะสมเพื่อให้บรรลุตามเป้าหมายด้านต้นทุนหรือไม่

ตรวจสอบว่า Amazon EC2 Instance ที่คุณเลือกเหมาะกับปริมาณงานที่ทำอยู่ AWS สนับสนุนให้ใช้การประเมินเทียบวัดมาตรฐานเพื่อตรวจสอบว่าประเภทอินสแตนซ์ที่คุณเลือกมีประสิทธิภาพที่เหมาะสมกับเวิร์กโหลด

แนวทางปฏิบัติ:

- **จับคู่โปรไฟล์อินสแตนซ์โดยอิงตามความต้องการ**เช่น จับคู่ตามคำอธิบายเวิร์กโหลดและอินสแตนซ์ โดยเน้นถึงระบบประมวลผล หน่วยความจำ หรือพื้นที่เก็บข้อมูล
- **ผลิตภัณฑ์ภายนอก** เช่น ใช้ผลิตภัณฑ์จากบริษัทอื่น เช่น CopperEgg หรือ New Relic เพื่อกำหนดประเภทอินสแตนซ์ที่เหมาะสม
- **Amazon CloudWatch** ใช้ CloudWatch เพื่อกำหนดโหลดตัวประมวลผล
- **ตัววัดผลแบบกำหนดเอง** โหลดสคริปต์หน่วยความจำแบบกำหนดเองและใช้ CloudWatch ตรวจสอบการใช้หน่วยความจำ
- **แอปพลิเคชันตามรายการข้อมูล** จัดทำไฟล์รายการข้อมูลของแอปพลิเคชัน เพื่อให้ทราบว่าจะใช้ Amazon EBS ประเภทใด (แม่เหล็ก, การใช้งานทั่วไป (SSD), IOPS ที่มีการเตรียมใช้งาน) และเมื่อใดใช้อินสแตนซ์ที่ผ่านการเพิ่มประสิทธิภาพโดย EBS เมื่อจำเป็นเท่านั้น

COST 4. คุณเลือกรูปแบบราคาที่เหมาะสมเพื่อให้บรรลุตามเป้าหมายด้านต้นทุนหรือไม่

ใช้รูปแบบราคาที่เหมาะกับเวิร์กโหลดของคุณมากที่สุดเพื่อลดค่าใช้จ่ายให้เหลือน้อยที่สุด การปรับใช้ที่เหมาะสมอาจเป็นได้ทั้งอินสแตนซ์ตามการใช้งานจริงทั้งหมด ผสมระหว่างอินสแตนซ์ตามการใช้งานจริงกับอินสแตนซ์แบบเหมาจ่าย หรือคุณอาจรวมอินสแตนซ์แบบประมูลราคา หากสามารถใช้ได้

แนวทางปฏิบัติ:

- **คอยสังเกต** ใช้อินสแตนซ์แบบประมูลราคา (Spot Instance) สำหรับเวิร์กโหลดที่เลือก
- **วิเคราะห์การใช้งาน** วิเคราะห์การใช้งานเป็นประจำและซื้ออินสแตนซ์แบบเหมาจ่าย (Reserved Instance) ให้สอดคล้องกัน
- **ขายอินสแตนซ์แบบเหมาจ่าย** เมื่อความต้องการเปลี่ยนแปลงไป คุณสามารถขายอินสแตนซ์แบบเหมาจ่ายที่ไม่ต้องใช้อีกต่อไปบน Reserved Instances Marketplace และซื้ออินสแตนซ์อื่น
- **การดำเนินการแบบอัตโนมัติ** ขออนุญาตสถาปนิกระบบปิดอินสแตนซ์ที่ไม่ได้ใช้งาน (เช่น ใช้ Auto Scaling

เพื่อปรับลดขนาดการใช้งานช่วงนอกเวลาทำการ)

- **พิจารณาต้นทุน** แยกต้นทุนสำหรับการเลือกภูมิภาค

COST 5. มีบริการที่ได้รับการจัดการ (บริการในระดับที่สูงกว่า Amazon EC2, Amazon EBS และ Amazon S3) ที่คุณสามารถใช้เพื่อปรับปรุง ROI หรือไม่

Amazon EC2, Amazon EBS และ Amazon S3 เป็นบริการ AWS ระดับ “Building-Block” ส่วนบริการที่ได้รับการจัดการ เช่น Amazon RDS และ Amazon DynamoDB เป็นบริการ AWS “ระดับสูงกว่า” การใช้บริการที่ได้รับการจัดการเหล่านี้จะช่วยคุณลดหรือตัดค่าใช้จ่ายด้านการดูแลระบบหรือการดำเนินงานได้มาก และทำงานในแอปพลิเคชันและกิจกรรมที่เกี่ยวข้องกับธุรกิจได้อย่างอิสระ

แนวทางปฏิบัติ:

- **วิเคราะห์บริการ** วิเคราะห์บริการระดับแอปพลิเคชันเพื่อดูว่าแอปพลิเคชันใดที่คุณสามารถใช้ได้
- **พิจารณาฐานข้อมูลที่เหมาะสม** ใช้ Amazon Relational Database Service (RDS) (Postgres, MySQL, SQL Server, Oracle Server) หรือ Amazon DynamoDB (หรือที่เก็บค่าคีย์อื่นๆ ทางเลือก NoSQL) ตามความเหมาะสม
- **พิจารณาบริการระดับแอปพลิเคชันอื่นๆ** ใช้ Amazon Simple Queue Service (SQS), Amazon Simple Notification Service (SNS) และ Amazon Simple Email Service (SES) ตามความเหมาะสม
- **พิจารณา AWS CloudFormation, AWS Elastic Beanstalk หรือ AWS Opsworks** ใช้เทมเพลต AWS CloudFormation / AWS Elastic Beanstalk /AWS OpsWorks เพื่อใช้ประโยชน์ในการกำหนดมาตรฐานและการควบคุมต้นทุน

COST 6. คุณมีการควบคุมการเข้าถึงและขั้นตอนใดบ้างในการกำกับดูแลการใช้งาน AWS

กำหนดนโยบายและกลไกเพื่อดูแลให้ต้นทุนมีจำนวนที่เหมาะสมและเป็นไปตามวัตถุประสงค์ การใช้แนวทางตรวจสอบและถ่วงดุลด้วยวิธีการติดแท็กและการควบคุม IAM จะช่วยให้คุณพัฒนาต่อไปได้โดยไม่เกิดการใช้จ่ายที่มากเกินไป

แนวทางปฏิบัติ:

- **กำหนดกลุ่มและบทบาท** (เช่น: พัฒนา/ทดสอบ/ใช้งานจริง) ใช้กลไกกำกับดูแลของ AWS เช่น IAM เพื่อ

ควบคุมว่าบุคคลใดที่สามารถเปิดใช้อินสแตนซ์และทรัพยากรในแต่ละกลุ่ม (แนวทางนี้สามารถใช้ได้กับบริการ AWS หรือโซลูชันภายนอก)

- **ติดตามวงจรของโครงการ** ติดตาม วัดผล และตรวจสอบวงจรของโครงการ ทีม และระบบสภาพแวดล้อมเพื่อหลีกเลี่ยงการใช้งานและการจ่ายเงินไปกับทรัพยากรที่ไม่จำเป็น

COST 7. คุณมีวิธีการอย่างไรในการตรวจสอบการใช้งานและค่าใช้จ่าย

กำหนดนโยบายและขั้นตอนเพื่อตรวจสอบ ควบคุม และจัดสรรต้นทุนอย่างเหมาะสม ใช้ประโยชน์จากเครื่องมือที่ AWS มีให้บริการเพื่อให้ทราบข้อมูลการใช้งานของบุคคล และจำนวนต้นทุนที่ใช้ วิธีนี้จะช่วยให้คุณเข้าใจถึงความต้องการของธุรกิจและการดำเนินงานของทีมได้ละเอียดขึ้น

แนวทางปฏิบัติ:

- **ติดแท็กทรัพยากรทั้งหมด** เพื่อให้สามารถเชื่อมโยงการเปลี่ยนแปลงจำนวนเงินที่ต้องจ่ายกับการเปลี่ยนแปลงโครงสร้างพื้นฐานและการใช้งาน
- **ตรวจสอบรายงานการเรียกเก็บเงินแบบละเอียด** กำหนดกระบวนการมาตรฐานเพื่อโหลดและตีความรายงานการเรียกเก็บแบบละเอียด
- **ระบบสถาปัตยกรรมที่คุ้มค่า** กำหนดแผนสำหรับทั้งการใช้งานและค่าใช้จ่าย (แบบต่อหน่วย เช่น ผู้ใช้และกิกะไบต์ข้อมูล)
- **การตรวจสอบ** ตรวจสอบการใช้งานและค่าใช้จ่ายเป็นประจำโดยใช้ Amazon CloudWatch หรือผู้ให้บริการอื่นๆ (เช่น: Cloudability, CloudCheckr)
- **การแจ้งข้อมูล** แจ้งให้สมาชิกหลักในทีมทราบเมื่อค่าใช้จ่ายขยับไปอยู่นอกค่าจำกัดที่กำหนดไว้อย่างเหมาะสม
- **ใช้ AWS Cost Explorer**
- **วิธีเรียกเก็บเงินตามการใช้งานจริงโดยอาศัยระบบการเงิน** ใช้วิธีนี้เพื่อจัดสรรอินสแตนซ์และทรัพยากรให้กับศูนย์ต้นทุน (เช่น การติดแท็ก)

COST 8. คุณมีวิธีอย่างไรในการปลดการใช้งานทรัพยากรที่ไม่ต้องการใช้อีกต่อไปหรือหยุดใช้ทรัพยากรที่ไม่จำเป็นชั่วคราว

ตรวจสอบว่าคุณจ่ายเงินสำหรับบริการที่ใช้งานอยู่เท่านั้น ใช้การควบคุมการเปลี่ยนแปลงและการจัดการทรัพยากรตั้งแต่เริ่มจนถึงสิ้นสุดโครงการ เพื่อให้ระบุการเปลี่ยนแปลงหรือการเพิ่มเติมกระบวนการที่จำเป็นได้

ตามความเหมาะสม ทำงานร่วมกับ AWS Support เพื่อขอคำแนะนำในการเพิ่มประสิทธิภาพโครงการให้เหมาะกับเวิร์กโหลดของคุณ เช่น ควรใช้ Auto Scaling, AWS OpsWorks, AWS Data Pipeline หรือแนวทางการเตรียมใช้งาน Amazon EC2 อื่นๆ เมื่อใด

แนวทางปฏิบัติ:

- ออกแบบระบบให้ดูแลการยุติการใช้งานอินสแตนซ์ได้อย่างสมบูรณ์ เมื่อคุณพบและต้องการปลดการใช้งานอินสแตนซ์ที่ไม่สำคัญหรือไม่จำเป็นหรือทรัพยากรที่มีการใช้งานในระดับต่ำ
- กำหนดกระบวนการเพื่อระบุและปลดการใช้งานทรัพยากรที่ไม่ได้อยู่ในส่วนใด
- ปรับปรุงยอดทรัพยากรที่ปลดการใช้งานตามระบบหรือกระบวนการ

COST 9. คุณพิจารณาค่าใช้จ่ายในการถ่ายโอนข้อมูลในการออกแบบสถาปัตยกรรมของคุณหรือไม่

ตรวจสอบค่าใช้จ่ายในการถ่ายโอนข้อมูลอยู่เสมอเพื่อให้สามารถดำเนินการตัดสินใจเชิงสถาปัตยกรรมเพื่อลดต้นทุนเหล่านี้ลงบางส่วน ตัวอย่างเช่น ถ้าคุณเป็นผู้ให้บริการเนื้อหาและให้บริการเนื้อหาจากบัคเก็ต Amazon S3 โดยตรงแก่ผู้ใช้ คุณอาจลดต้นทุนได้ถ้าคุณพุชเนื้อหาไปยัง Amazon CloudFront CDN จำไว้ว่าการเปลี่ยนแปลงเชิงสถาปัตยกรรมที่มีประสิทธิภาพแม้เพียงเล็กน้อยก็สามารถลดต้นทุนในการดำเนินงานได้อย่างเห็นผล

แนวทางปฏิบัติ:

- ใช้ CDN
- ออกแบบสถาปัตยกรรมเพื่อเพิ่มประสิทธิภาพในการถ่ายโอนข้อมูล (การออกแบบแอปพลิเคชัน, การเร่งประสิทธิภาพของ WAN เป็นต้น)
- วิเคราะห์สถานการณ์และใช้ AWS Direct Connect เพื่อประหยัดค่าใช้จ่ายและปรับปรุงประสิทธิภาพการทำงาน
- ปรับสมดุลต้นทุนการถ่ายโอนข้อมูลของระบบสถาปัตยกรรมกับความต้องการด้านความพร้อมใช้งานสูง (HA) และความน่าเชื่อถือ

COST 10. คุณจะจัดการและ/หรือพิจารณานำบริการใหม่ๆ มาใช้งานอย่างไร

AWS มีเป้าหมายที่จะช่วยคุณออกแบบสถาปัตยกรรมให้มีความเหมาะสมและคุ้มค่ากับต้นทุนมากที่สุด บริการและคุณสมบัติใหม่ๆ อาจช่วยลดต้นทุนของคุณได้โดยตรง ที่เห็นได้ชัดก็เช่น Amazon Glacier ซึ่งนำเสนอ

โซลูชันการจัดเก็บข้อมูลแบบ “แช่เย็น” สำหรับข้อมูลที่ไม่ได้เข้าใช้งานบ่อยๆ แต่ยังต้องเก็บรักษาด้วยเหตุผลทางธุรกิจและกฎหมาย ตัวอย่างอื่นๆ เช่น Reduced Redundancy Storage สำหรับ Amazon S3 ที่ให้คุณสามารถเลือกสำเนาออบเจ็กต์ Amazon S3 น้อยลง (ความซ้ำซ้อนในระดับที่น้อยลง) ในราคาที่ต่ำกว่า คุณควรพิจารณาข้อสังเกตต่างๆ เมื่อดำเนินการตัดสินใจเหล่านี้เช่น: “การมีสำเนาข้อมูลที่น้อยลงจะส่งผลอย่างไร” หรือ “ฉันจำเป็นต้องเข้าถึงข้อมูลนี้มากกว่าที่คิดไว้หรือไม่”

แนวทางปฏิบัติ:

- พบปะกับสถาปนิกด้านโซลูชัน ที่ปรึกษา หรือทีมดูแลลูกค้าของ AWS เป็นประจำ และพิจารณาบริการหรือคุณสมบัติใหม่ๆ ที่คุณสามารถนำไปใช้เพื่อประหยัดค่าใช้จ่าย