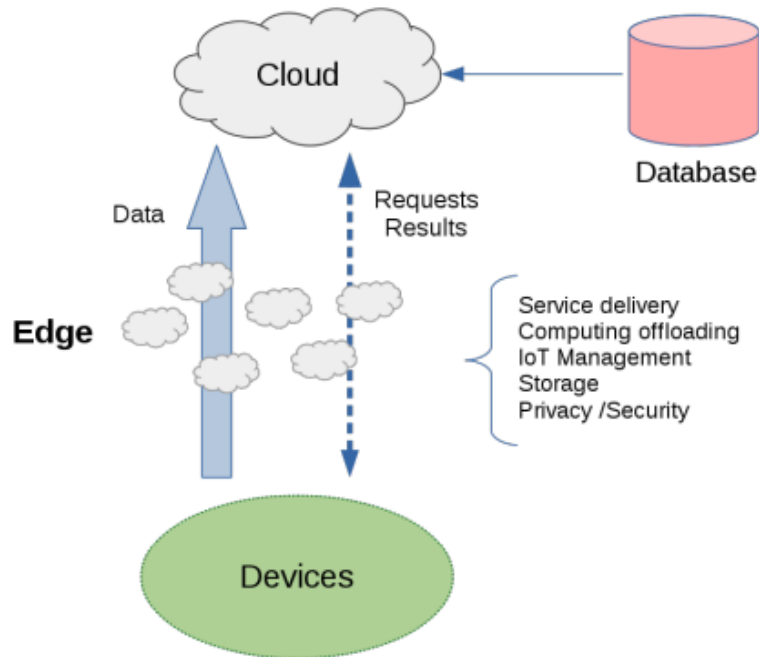


# Unit 8 Edge Computing

# Edge Computing คือ ? (edge)

- 🧑‍🔧 ช่วยในการแบ่งเบาภาระของ cloud service หลักของระบบโดยการสร้างระบบย่อยที่อยู่ใกล้อุปกรณ์ปลายทาง เช่น pc, iot
- 🧑‍🔧 ระบบย่อย ทำหน้าที่ เป็นหน่วยประมวลผลย่อยที่มีลักษณะคล้าย cloud service จะเชื่อมต่อกับ end-user ได้อย่างรวดเร็วส่งผลให้การสื่อสารไปอย่างรวดเร็ว หมายความว่าความล่าช้าที่ต่ำ (latency) โดยมักจะใช้ในระบบที่ต้องการตอบสนองแบบทันทีทันใด เช่น รถยนต์อัตโนมัติ
- 🧑‍🔧 บางระบบจะเป็นการที่ end-user เลือกเชื่อมต่อไปที่ cloud ที่อยู่ใกล้ที่สุดก็คือ Edge
- 🧑‍🔧 การเพิ่มขึ้นของ things ที่เชื่อมต่ออินเทอร์เน็ตได้ส่งผลให้มีข้อมูลปริมาณมากถูกขนส่งเครือข่าย ส่งผลให้มีเข้าถึงข้อมูลจำนวนมาก ร้องขอ เก็บ แก้ไขข้อมูลในระบบฐานข้อมูลและรองรับการเชื่อมต่อจำนวนมากในเวลาเดียวกัน ปัญหาเรื่องระยะทางในการรับส่งก็เป็นปัญหาสื่อสารตั้งแต่ live streaming, content delivery. เล่นเกมข้ามทวีปไปจนถึงความซับซ้อน ผลกระทบต่อสังคม เช่น ระบบอัตโนมัติต่างในโรงงานที่เป็น IOT
- 🧑‍🔧 เป็นการนำการประมวลผลเข้าใกล้แหล่งข้อมูลและไม่จำเป็นต้องส่งไปยังระบบคลาวด์ระยะไกลหรือระบบส่วนกลางอื่น
- 🧑‍🔧 โดยการลดระยะทางและเวลาที่มันต้องใช้ในการส่งข้อมูลไปยังแหล่งข้อมูลส่วนกลางสามารถปรับปรุงความเร็วและประสิทธิภาพของการส่งข้อมูลรวมถึงอุปกรณ์และแอปพลิเคชันที่อยู่บริเวณขอบเขต





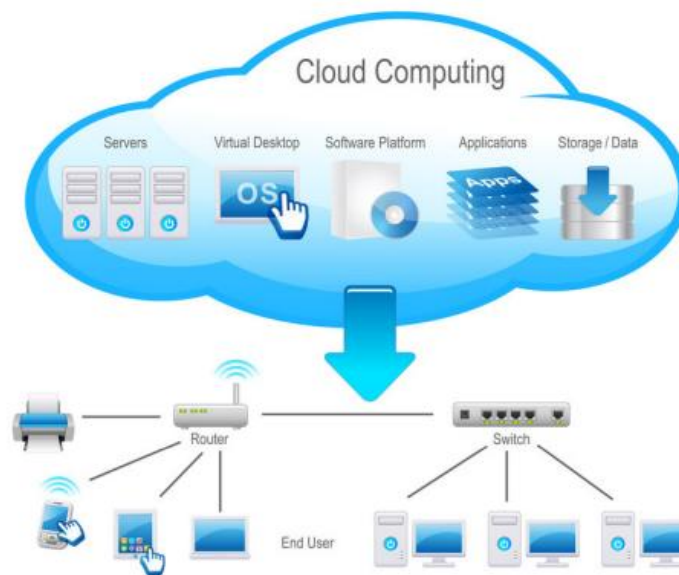
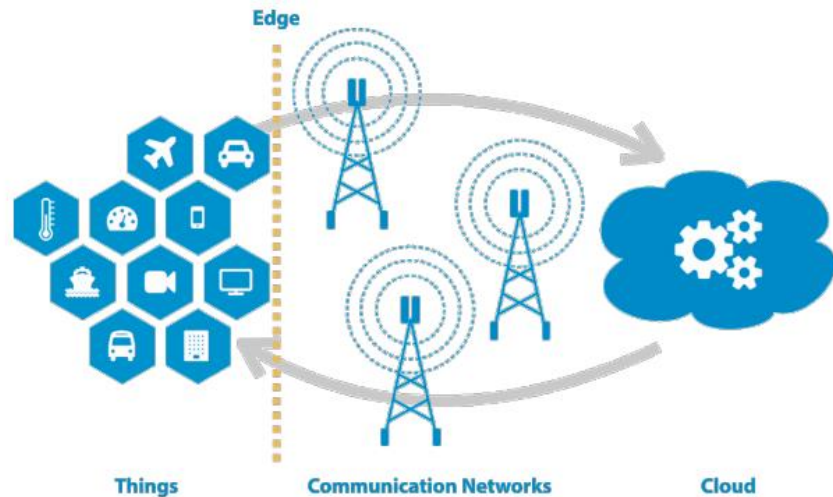
- จากภาพจะเห็นว่า edge จะอยู่ระหว่าง devices และ cloud ซึ่งในลักษณะนี้ edge จะทำหน้าที่ติดต่อกับ devices โดยตรงและจะติดต่อกับ cloud ด้วย เพื่อส่งต่อข้อมูลจาก devices ไปยัง cloud การที่นำ edge มาติดตั้งในลักษณะนี้จะช่วยทำให้การติดต่อ service ของ devices นั้นมีความเสถียรมากกว่าการติดต่อไปยัง cloud หาก cloud อยู่ไกล เช่น คนละประเทศ หรือคนละทวีป
- ในระบบ เช่น autonomous vehicles นั้น edge computing อาจอยู่ในตัวยานพาหนะเองเพื่อให้การบังคับพาหนะนั้นเป็นแบบ real-time โดยต้องมีค่า latency ต่ำที่สุดเท่าที่เป็นไปได้เพื่อป้องกันอุบัติเหตุที่อาจเกิดบนท้องถนน เพราะถือเป็นระบบที่ sensitive มาก

# แนวความคิด Edge Computing

- 📺 ปี 2021 มีอุปกรณ์ IOT = 13,000 ล้านอุปกรณ์ทำงานอยู่บน cloud
- 📺 ส่งผลให้มีข้อมูลจำนวนมหาศาลถูกรับ-ส่งอยู่ในระบบอินเทอร์เน็ต ส่งผลให้การจัดการกับข้อมูลที่มีอยู่อย่างมหาศาลจากจำนวนอุปกรณ์ที่เพิ่มมากขึ้นอย่างทวีคูณเป็นไปได้ยาก
- 📺 เนื่องจากข้อมูลบางข้อมูลที่ถูกรับบันทึกไว้ในคลาวด์มีความสำคัญน้อยมากและไม่ค่อยได้ใช้จึงกลายเป็นการสิ้นเปลืองทรัพยากรและพื้นที่เก็บข้อมูล
- 📺 ควรมีการสร้างระบบเก็บข้อมูลที่ถูกระบุใช้บ่อย หรือข้อมูลที่สำคัญมาเก็บไว้ใกล้ผู้ใช้หรือบริเวณอุปกรณ์นอกสุด (end users หรือ end devices) ส่งผลให้การเชื่อมต่อไปยังระบบ cloud ลดลงช่วยลดภาระและค่าใช้จ่ายในการดูแลระบบ cloud
- 📺 การแบ่งข้อมูลบางส่วนมาเก็บไว้ด้านนอกนี้เรียกว่า edge computing เพื่อเป็นการลดเวลาและปริมาณข้อมูลที่ต้องการส่งระหว่างอุปกรณ์นอกสุดกับคลาวด์








# การประมวลผลแบบคลาวด์ (Cloud Computing)

- เป็นการให้บริการทรัพยากรตามความต้องการของผู้ใช้ โดยเฉพาะอย่างยิ่งการจัดเก็บข้อมูลและการประมวลผลข้อมูล
- ผู้ใช้ไม่จำเป็นต้องรู้ว่าระบบมีฮาร์ดแวร์และซอฟต์แวร์อะไรบ้าง รู้เพียงการใช้งานเท่านั้น
- ตัวอย่าง cloud computing เช่น บริการเก็บข้อมูล Google Drive หรือ Dropbox แทนการเก็บข้อมูลลงบนคอมพิวเตอร์ส่วนตัว










## ข้อดีของการประมวลผลแบบ edge

-  เพิ่มความปลอดภัยของข้อมูลและความเป็นส่วนตัว
-  ประสิทธิภาพของแอปพลิเคชันที่ดีขึ้น ตอบสนองเร็วและแข็งแกร่งขึ้น
-  ลดต้นทุนการดำเนินงาน
-  ปรับปรุงประสิทธิภาพและความน่าเชื่อถือของธุรกิจ
-  ขยายขีดความสามารถไม่จำกัด
-  อนุรักษ์เครือข่ายและทรัพยากรการคำนวณ
-  เวลาแฝงที่ลดลง (latency)

## ความท้าทายที่เกิดขึ้นเมื่อมีการใช้

-  ประเด็นหลักๆ ที่ใช้พิจารณาเมื่อเลือกใช้อาจจะส่งผลกระทบต่อความเป็นส่วนตัวของผู้ใช้ ความปลอดภัยของข้อมูล การขยายตัวของระบบ ความน่าเชื่อถือ ความมีเสถียรภาพของระบบ
- 1. Privacy & Security ความเป็นส่วนตัวและความปลอดภัย
- 2. Scalability การปรับขนาดของระบบ
- 3. Reliability ความน่าเชื่อถือ

# 1. ความเป็นส่วนตัวและความปลอดภัย (Privacy and security)






-  การใช้งาน edge ส่งผลให้ระบบโดยรวมมีโอกาสถูกโจมตีได้มากขึ้น เนื่องจากการเพิ่มโหนดของการเชื่อมต่อระหว่าง user และ cloud ดังนั้นตัว edge เองอาจจะถูกโจมตีได้หรืออาจถูกโจมตีระหว่างการส่งข้อมูล (man-in-the-middle-attack)
-  การกระจายตัวกันทำงานของ cloud service หลัก micro-service ที่เป็น edge อาจมีมากกว่า 1 ระบบย่อย ทำให้การวางแผนการรักษาความปลอดภัยนั้นทำได้ยากยิ่งขึ้นและทำให้การเฝ้าระวังทำได้ไม่ทั่วถึงหากมีบุคลากรจำกัด
-  ส่งผลให้เกิดการใช้งบประมาณและทรัพยากรในการบริหารมากยิ่งขึ้น หน่วยงานที่ใช้ edge แต่มีงบประมาณจำกัดและต้องการลดต้นทุน
-  ส่วนนี้จะส่งผลให้ระบบขาดความปลอดภัยต่อการถูกโจมตีได้ไม่เพียงแต่ข้อมูลที่ถูกเข้ารหัส แต่ควรใช้กลไกการเข้ารหัสที่แตกต่างกัน ตั้งแต่ข้อมูลถูกส่งระหว่างโหนดแบบกระจายที่แตกต่างกันเชื่อมผ่านอินเทอร์เน็ตก่อนที่จะถูกคลาวด์

# 1. ความเป็นส่วนตัวและความปลอดภัย (Privacy and security)

- 🏠 Edge node อาจเป็นอุปกรณ์ที่ถูกจำกัดทรัพยากรเป็นตัวเลือกที่ถูกจำกัดในแง่ของวิธีการรักษาความปลอดภัย
- 🏠 การติดตั้ง edge เป็นการเปลี่ยนการทำงานแบบรวมศูนย์จากบนลงล่างไปเป็นโครงสร้างพื้นฐานแบบกระจายที่มี edge เป็น node ทำให้ต้องการระบบรักษาความปลอดภัยรวมถึงรูปแบบการส่งข้อมูลที่น่าเชื่อถือและปลอดภัย หากมีการเก็บข้อมูลไว้ที่ edge อาจเป็นไปได้ว่าข้อมูลอาจถูก hack โดยที่ maincloud service ไม่รู้สิ่งที่เกิดขึ้นหากไม่มีการวางแผนระบบรักษาความปลอดภัยที่ดี
- 🏠 ปัจจุบัน IoT ถือว่าเป็นเป้าหมายที่ค่อนข้างง่ายต่อ hacker เพราะมีอยู่ทั่วไปอีกทั้งยังมี protocol มากมายที่ถูกใช้งานและยังขาดมาตรฐานสากลที่เป็นที่ยอมรับทั่วโลกที่ดีมารองรับ



# ความเป็นไปได้ในการรักษาความปลอดภัย

-  เข้ารหัสข้อมูล (data encryption) ทั้งขณะที่ทำการส่งและการเก็บข้อมูล
-  ใช้เทคนิคการเข้ารหัสที่หลากหลาย และมีการเปลี่ยนวิธีการเข้ารหัสอยู่เสมอ
-  ใช้ระบบกรองข้อมูล (data filtering) และการเชื่อมต่อจากโหนดต่างๆ (connection filtering)
-  มีการยืนยันตนสำหรับการกรองข้อมูล (authentication identification)
-  มีการตรวจสอบและทดสอบระบบรักษาความปลอดภัยอย่างสม่ำเสมอ

## 2. การขยายตัวหรือการปรับขนาดของระบบ (Scalability)



การขยายตัวในเครือข่ายแบบกระจายต้องคำนึงถึงหลายปัจจัยเพื่อให้ระบบสามารถทำงานได้อย่างราบรื่นในกรณีที่มีการเชื่อมต่อมากยิ่งขึ้น เช่น การที่มี device เชื่อมต่อเข้ามาที่ edge มากขึ้นซึ่งจะส่งผลต่อการเชื่อมต่อ ดังนั้นผู้พัฒนาต้องมองไปยังอนาคตถึงแนวโน้มในการเพิ่มขึ้นของ device หรือ user ที่จะมีมากขึ้นในอนาคตและติดตั้งโครงสร้างพื้นฐาน (infrastructure) เช่น สายสัญญาณหรือจำนวน node ที่สามารถรับการเชื่อมต่อได้พร้อมๆ กัน อีกทั้งการเพิ่มขึ้นของการใช้พลังงานและพื้นที่ในการติดตั้งระบบเมื่อเกิดการขยายตัวอีกด้วย

### การรับมือกับการขยายตัวของระบบ



เลือกใช้อุปกรณ์เครือข่ายที่สามารถรองรับการเชื่อมต่อที่คาดว่าจะเพิ่มขึ้น



สำรองงบประมาณไว้สำหรับการลงทุนในติดตั้งส่วนขยายเพิ่มเติม เช่น infrastructure



การตรวจสอบความเสถียรของระบบโดยรวมเมื่อมีการขยายตัวเกิดขึ้นซึ่งอาจส่งผลต่อการสื่อสารของข้อมูลเพื่อให้ทราบถึงประสิทธิภาพของการทำงาน

### 3. ความน่าเชื่อถือ (Reliability)



เป็นเรื่องสำคัญของการบริการ การสร้างความน่าเชื่อถือคือการลดจำนวนการเกิดความผิดพลาดให้น้อยที่สุด และทำให้ user เกิดความพึงพอใจมากที่สุด ในมุมมองของผู้ใช้ในส่วนในระบบจะเป็นการทำงานประสานกันของส่วนประกอบย่อยต่างๆ เพื่อให้การใช้งานนั้นราบรื่นและมีประสิทธิภาพ



การสร้างความน่าเชื่อถือแบ่งได้ 2 ประเด็น คือ

1. ด้านการทำงานของระบบ
2. ด้านการบริหารงานบุคคล



การจัดการกับข้อผิดพลาดหรือความล้มเหลวเป็นสิ่งสำคัญที่ทำให้ระบบยังสามารถทำงานได้อยู่ ถ้ามีโหนดหยุดทำงานและไม่สามารถติดต่อได้ควรมีวิธีการเชื่อมต่อแบบฉุกเฉินรองรับเพื่อให้การเข้าถึงบริการไม่หยุดชะงักหรือเกิดความเสียหายมากยิ่งขึ้น ยิ่งกว่านั้นระบบที่น่าเชื่อถือควรที่จะต้องมีการกู้คืนข้อมูลที่สูญหายไปและแจ้งเตือนผู้ใช้เกี่ยวกับเหตุการณ์ที่เกิดขึ้น



เพื่อจุดประสงค์นี้อุปกรณ์แต่ละเครื่องจะต้องรักษาโครงสร้างเครือข่ายของระบบกระจายทั้งหมดเพื่อให้การตรวจจับข้อผิดพลาดและการกู้คืนสามารถนำไปใช้ได้อย่างง่ายด้ายปัจจัยอื่นๆ ที่อาจมีผลต่อแง่มุมนี้คือเทคโนโลยีการเชื่อมต่อที่ใช้ งานซึ่งอาจให้ระดับความน่าเชื่อถือที่แตกต่างกันและความถูกต้องของข้อมูลที่ผลิตที่ขอบซึ่งอาจไม่น่าเชื่อถือเนื่องจากสภาพแวดล้อมเฉพาะ



## ความน่าเชื่อถือด้านการทำงานของระบบ





เป็นสิ่งที่ user คาดหวังมากที่สุด การทำงานของระบบนั้นถือเป็นสิ่งที่สำคัญที่สุดไม่ว่าจะเป็นในมุมมองของ user หรือผู้ดูแลระบบเป็นความน่าเชื่อถือโดยรวมของระบบไม่ว่าจะเป็นการทำงานของระบบในส่วน of device, cloud, และ end-use application ซึ่งการที่จะทำให้แต่ละส่วนทำงานได้อย่างราบรื่นนั้นก็ขึ้นอยู่กับ การเลือกใช้อุปกรณ์ โปรโตคอล ระบบโครงสร้างพื้นฐานต่างๆ

## ความน่าเชื่อถือด้านการบริหารงานบุคคล




ทำได้ด้วยการวางแผนการทำงานของบุคลากร การระบุหน้าที่ให้บุคลากร การเลือกใช้ซึ่งเป็นการวางตัวบุคคล ให้เหมาะสมกับงาน เช่น การพัฒนาระบบ การบำรุงดูแลรักษาระบบ การติดต่อกับผู้ใช้ตั้งแต่บุคลากรระดับสูงที่ ดำเนินการด้านบริหาร ไปจนถึงพนักงานระดับปฏิบัติการไม่ว่าจะเป็นช่างเทคนิคหรือพนักงานบริการหลังการขายที่ติดต่อกับลูกค้าโดยตรง

## Ex. Autonomous Vehicles ยานพาหนะอัตโนมัติ

-  การตัดสินใจหยุดรถเพื่อให้คนเดินถนนข้ามด้านหน้าของยานพาหนะอัตโนมัติสามารถทำให้สะดวกขึ้นได้ โดยใช้Edge computing จะต้องดำเนินการในทันที(instant หรือ real-time reaction) ไม่สามารถรอการตัดสินใจที่ตอบกลับจากเซิร์ฟเวอร์ที่อยู่ระยะทางไกลได้
-  ยานพาหนะแบบอัตโนมัติสามารถใช้edge technology ในการโต้ตอบกันได้อย่างมีประสิทธิภาพมากขึ้น เพราะสามารถสื่อสารระหว่างยานพาหนะด้วยกันทำให้สามารถส่งข้อมูลเกี่ยวกับอุบัติเหตุสภาพอากาศ การจราจร หรือออกนอกเส้นทาง โดยไม่ต้องติดต่อกับเซิร์ฟเวอร์ก่อน

## Ex. Healthcare Devices อุปกรณ์ด้านสุขภาพ

-  เครื่องตรวจสอบสุขภาพและอุปกรณ์ดูแลสุขภาพที่สวมใส่ได้อื่น ๆ สามารถจับตามองสภาวะเรื้อรังสำหรับผู้ป่วยและช่วยชีวิตด้วยการแจ้งเตือนผู้ดูแลทันทีที่ต้องการความช่วยเหลือ นอกจากนี้หุ่นยนต์ที่ให้ความช่วยเหลือในการผ่าตัดจะต้องสามารถวิเคราะห์ข้อมูลได้อย่างรวดเร็วเพื่อช่วยให้ปลอดภัยรวดเร็วและแม่นยำ หากอุปกรณ์เหล่านี้พึ่งพาการส่งข้อมูลไปยังคลาวด์ก่อนตัดสินใจผลที่ได้อาจเป็นอันตรายถึงชีวิต

## Ex. Security Solutions ระบบทางด้านการปลอดภัย



เนื่องจากจำเป็นต้องตอบสนองต่อภัยคุกคามภายในไม่กี่วินาทีระบบเฝ้าระวังความปลอดภัยจึงสามารถได้รับประโยชน์จากเทคโนโลยีการคำนวณที่ทันสมัย ระบบความปลอดภัยสามารถระบุภัยคุกคามที่อาจเกิดขึ้นและแจ้งเตือนผู้ใช้ถึงกิจกรรมที่ผิดปกติแบบเรียลไทม์

## Ex. Smart Speakers ลำโพงอัจฉริยะ



ช่วยให้ลำโพงอัจฉริยะทำงานได้ดีขึ้น โดยเพิ่มความสามารถในการตีความคำสั่งเสียงภายในเพื่อเรียกใช้คำสั่งพื้นฐานเช่นการเปิดหรือปิดไฟหรือปรับการตั้งค่าเทอร์โมสแตท แม้ว่าการเชื่อมต่ออินเทอร์เน็ตจะล้มเหลว



## Ex. Video Conferencing การประชุมทางไกลผ่านระบบวิดีโอ



คุณภาพวิดีโอไม่ดี ความล่าช้าของเสียง หน้าจอค้าง - การเชื่อมโยงไปยังคลาวด์ช้าอาจทำให้การประชุมทางวิดีโอที่น่าผิดหวังหลายอย่าง วิธีการแก้ไขโดยการวางแผนซอฟต์แวร์การประชุมทางวิดีโอไว้ที่เซิร์ฟเวอร์ใกล้กับผู้เข้าร่วมประชุม ปัญหาของคุณภาพจะลดลง

## Ex. Content Delivery การโหลดเนื้อหา



เว็บไซต์หรือแอปพลิเคชันที่มีผู้ใช้จำนวนมากและมีข้อมูลจำนวนมากจำเป็นที่จะต้องรองรับการเชื่อมต่อที่เพิ่มขึ้น พร้อมกันได้อย่างมีประสิทธิภาพ โดยสามารถใช้หลักการ content delivery network (CDN) ซึ่งเป็นการกระจาย server เพื่อรองรับการใช้งานในหลายพื้นที่ที่มีผู้ใช้อยู่ในพื้นที่ต่างๆ เพื่อให้ผู้ใช้แต่ละรายสามารถเข้าถึงเนื้อหาได้อย่างรวดเร็ว มีค่า latency ที่ต่ำและเข้าถึงเนื้อหาได้อย่างต่อเนื่อง