Next Generation Cloud Leadership

Cloud Operations

Develop Intelligence

Allen R. Sanders
Senior Technology Instructor

Objectives

- Provide AVP & VP level management with the skills & knowledge they need to be more Cloud conversant both from a practical and strategic perspective
- Ultimate objective is for some participants to get certified as AWS Cloud Practitioners

Plan-of-Attack

CLOUD OPERATIONS

Examine key considerations associated to managing and operating Cloud

- Managing data and implementing analytics on Cloud
- Security & compliance
- Role of DevOps and DevSecOps in Cloud computing
- Cost management - CAPEX vs. OPEX
- Managing application & infrastructure performance in the Cloud
- Evaluate case studies - determining operational concerns for a Cloud project

Operations

Metrics, Measurement & Assessment

"Every line is the perfect length if you don't measure it."

- Marty Rubin

"What gets measured gets managed."

- Pearl Zhu

"If you don't collect any metrics, you're flying blind. If you collect and focus on too many, they may be obstructing your field of view."

- Scott M. Graffius

"What science has failed to notice is that the measurement has become more real than the thing being measured."

- R.A. Delmonico

Metrics, Measurement & Assessment

"That which cannot be measured cannot be proven."

- Anthony W. Richardson

"All conflict in the world is essentially about our differences in measurement."

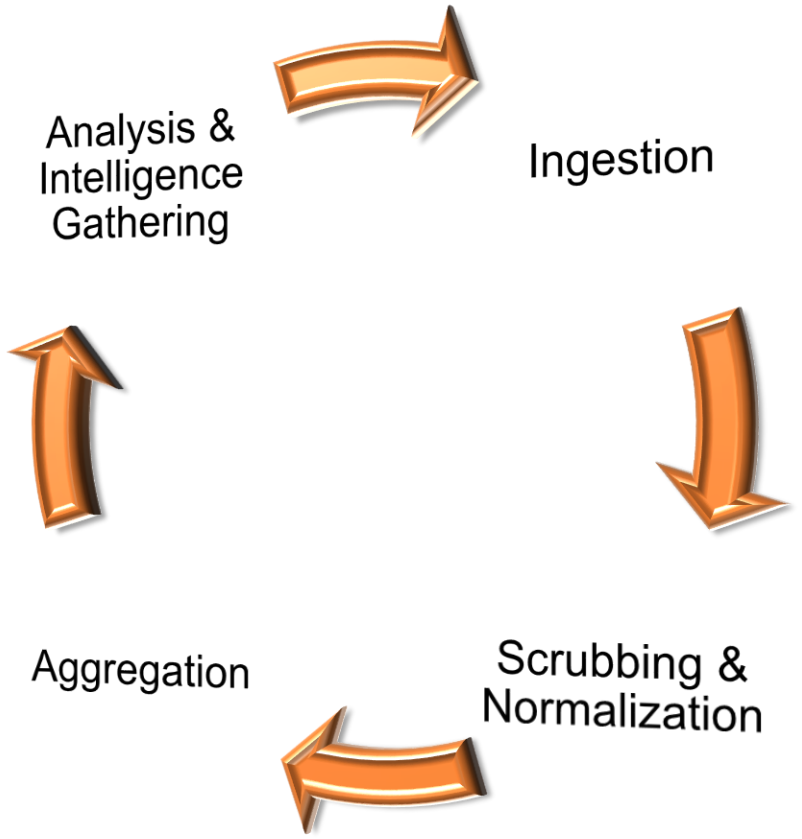
- Joseph Rain

"It is impossible to escape the impression that people commonly use false standards of measurement – that they seek power, success and wealth for themselves and admire them in others, and that they underestimate what is of true value in life."

- Sigmund Freud

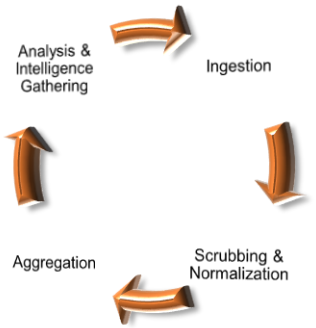
Data Management

Data Management – Stages



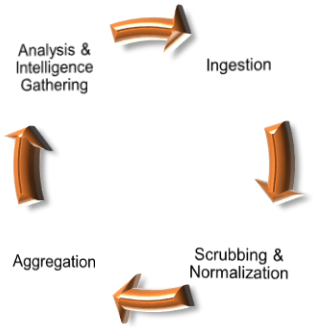
Data Ingestion

- Could be via message exchange or streaming
- Depending on size/scope, may translate to LARGE amounts of incoming data



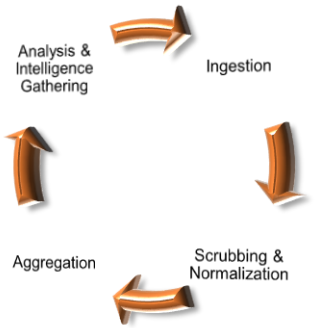
Data Ingestion

- Because of potential scale, bandwidth may be a concern
- Depending on application, latency may also be a concern
- Data may require translation (e.g., from low-level bytes to object or JSON)



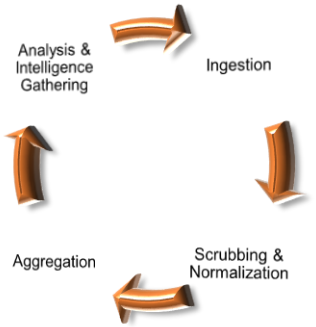
Data Ingestion

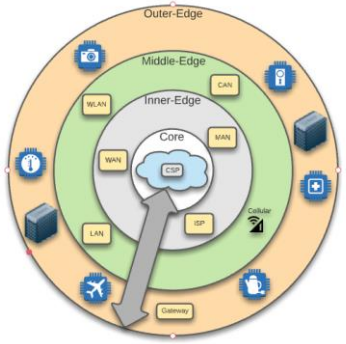
- Event hubs or streaming analytics platforms support ingestion at scale
- Provide time and context-aware processing for correct sequencing
- Data may flow through intermediate storage on way to final processing
- Depending on sensitivity of data, could require robust security at each stop



Data Ingestion – What About the Edge?

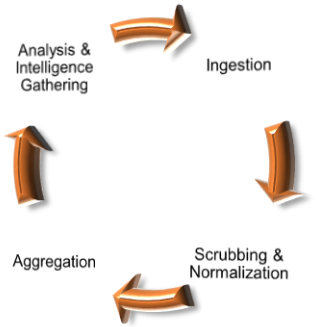
- Edge components (e.g., gateways) can help optimize
- Preliminary processing at the edge can be used to filter what really matters
- Potential for bundling or compressing data for transmit to cloud
- Can help with bandwidth or latency issues





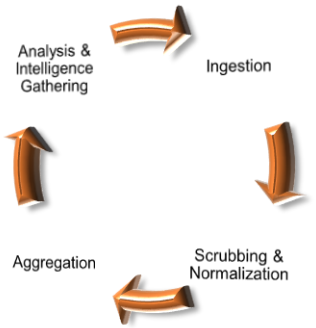
Data Scrubbing & Normalization

- Depending on payload, some portions of the data may not be needed
- Or some portions might contain sensitive detail
- Those parts not needed or sensitive can be “scrubbed” to exclude



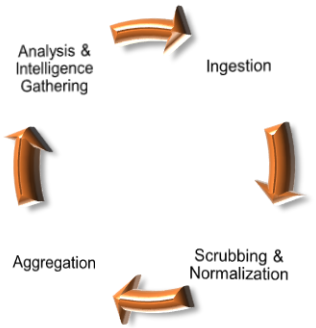
Data Scrubbing & Normalization

- Represents another potential optimization that can preserve storage
- In other cases, similar data may be coming in multiple, disparate formats



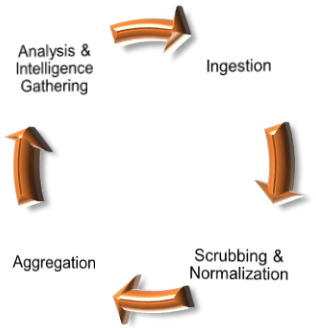
Data Scrubbing & Normalization

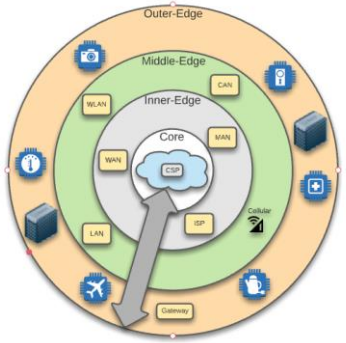
- Normalization can bring consistency to the disparate content
- By normalizing, becomes a single dataset for comprehensive analysis
- Normalization may happen as part of ingestion or as part of a separate step



Data Scrubbing & Normalization – What About the Edge?

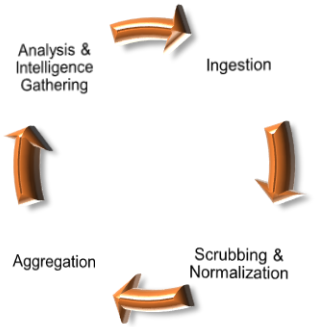
- Depending on complexity, may execute faster closer to the data
- Might involve proprietary algorithms best kept within full control
- Allows addressing of sensitive data before routed to Cloud
- Can also provide additional optimization (relative to bandwidth)





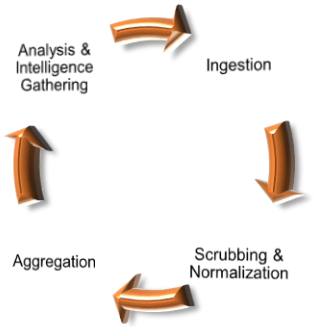
Data Aggregation

- Helps provide full picture of data from multiple streams
- May also be used to enrich with info from other data sources
- Data will be stored in persistent storage for downstream analysis & reporting



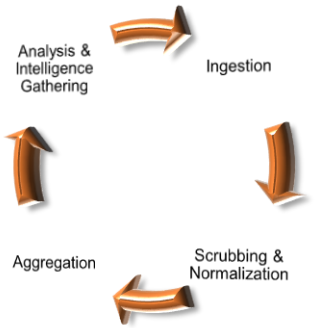
Data Aggregation

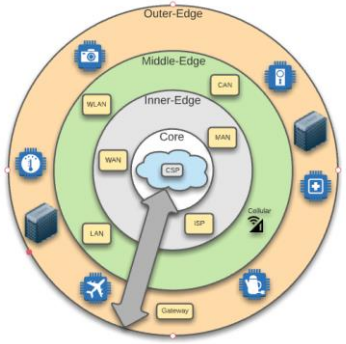
- In statistical analysis, the larger the sample size, the more accurate the inference
- To manage costs, large sets of data may leverage different types of storage:
  - Hot storage – most recent data and most relevant for current analysis
  - Cool storage – data not actively used but potentially relevant (short-term trends)
  - Cold or archive storage – data kept for historical purposes and long-term trending
- Security of the stored data and encryption at rest become critical



Data Aggregation – What About the Edge?

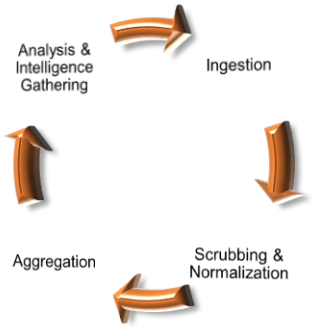
- Provides an additional layer of storage
- Data not transmitted to Cloud (due to optimizations) may still be valuable to keep
- Enables storage of sensitive data in “raw” format in controlled environment
- Can help balance costs against short to mid-term retention requirements





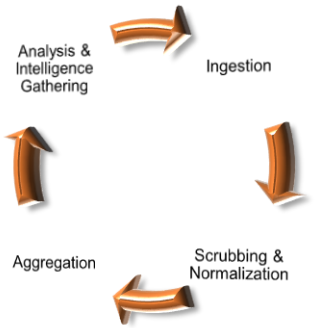
Data Analysis & Intelligence Gathering

- In the digital age, data is the competitive edge
- Companies that manage their data as a critical asset succeed
- Keys:
  - Aggregating efficiently
  - Analyzing effectively



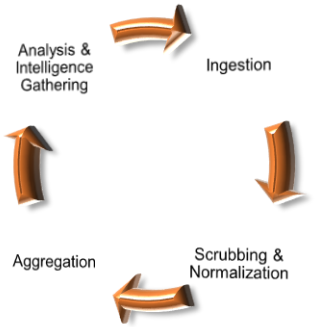
Data Analysis & Intelligence Gathering

- Goal is to identify and leverage the most important data points
- Importance is measured by business value-driven decision-making
- What can I learn about today's customers, scenarios, or business cases?
- What can I effectively predict about tomorrow?



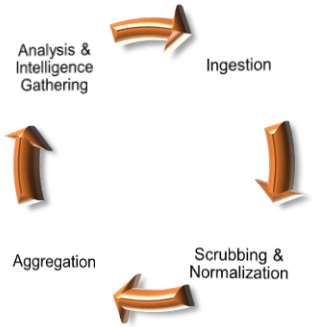
Data Analysis & Intelligence Gathering

- Requires balancing of competing concerns:
  - To increase quality of intelligence, more data is required (sometimes MUCH more)
  - But massive datasets can be complex to manage and process

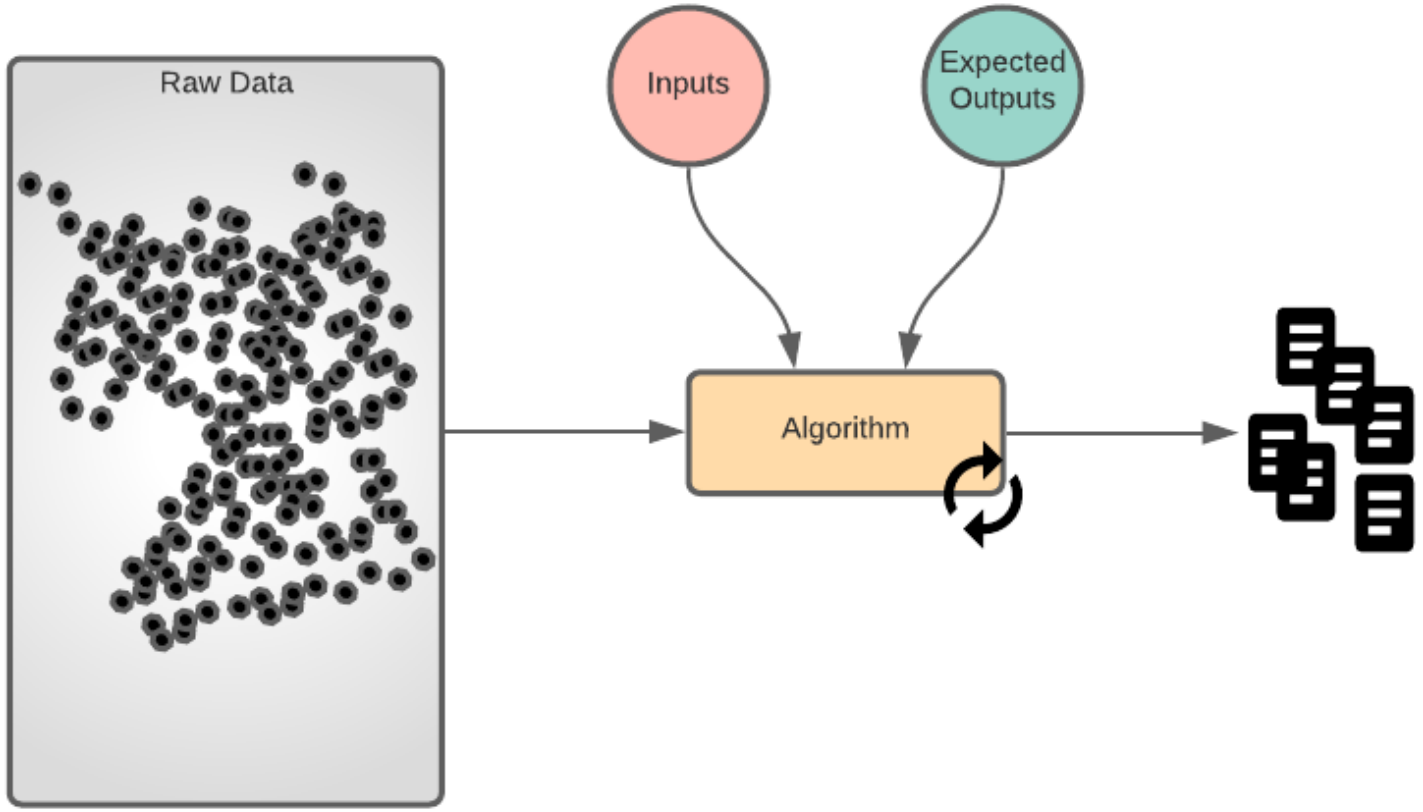


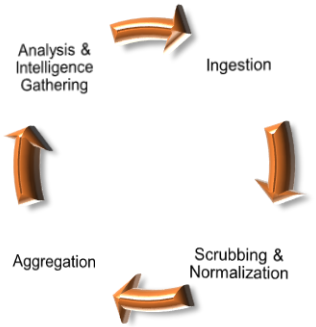
Data Analysis & Intelligence Gathering

- Enter ML / AI:
  - Algorithms are used to build mathematical models from existing data
  - Results in a mathematical “trajectory” (and confidence level)
  - Algorithms can be configured to learn and improve over time
- Hyperscale available in the Cloud brings near-limitless power to bear

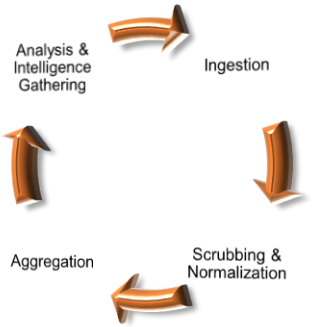


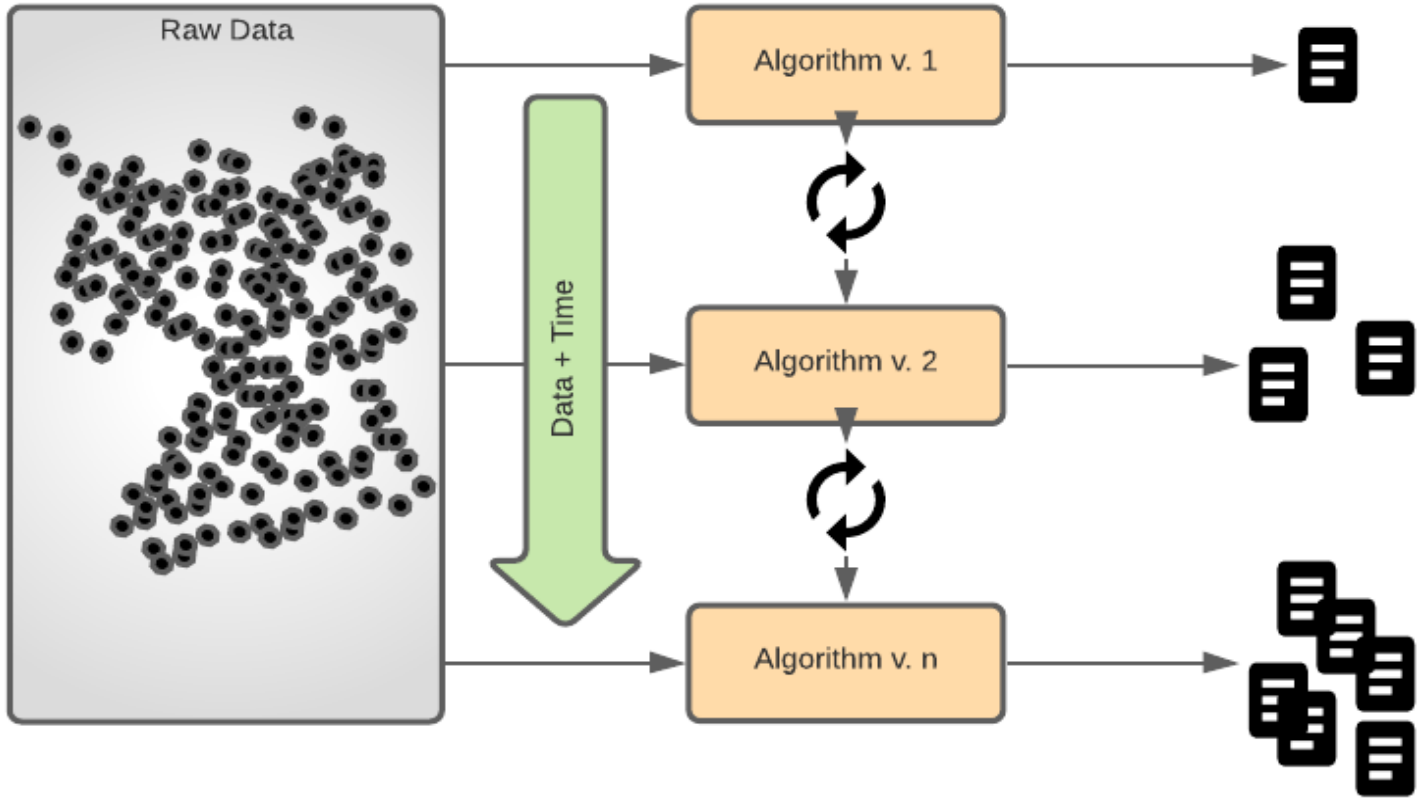
Supervised Learning



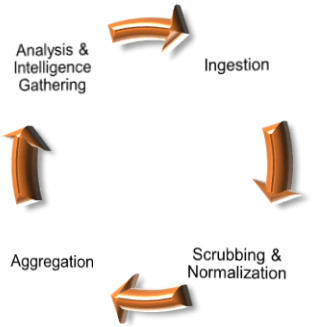


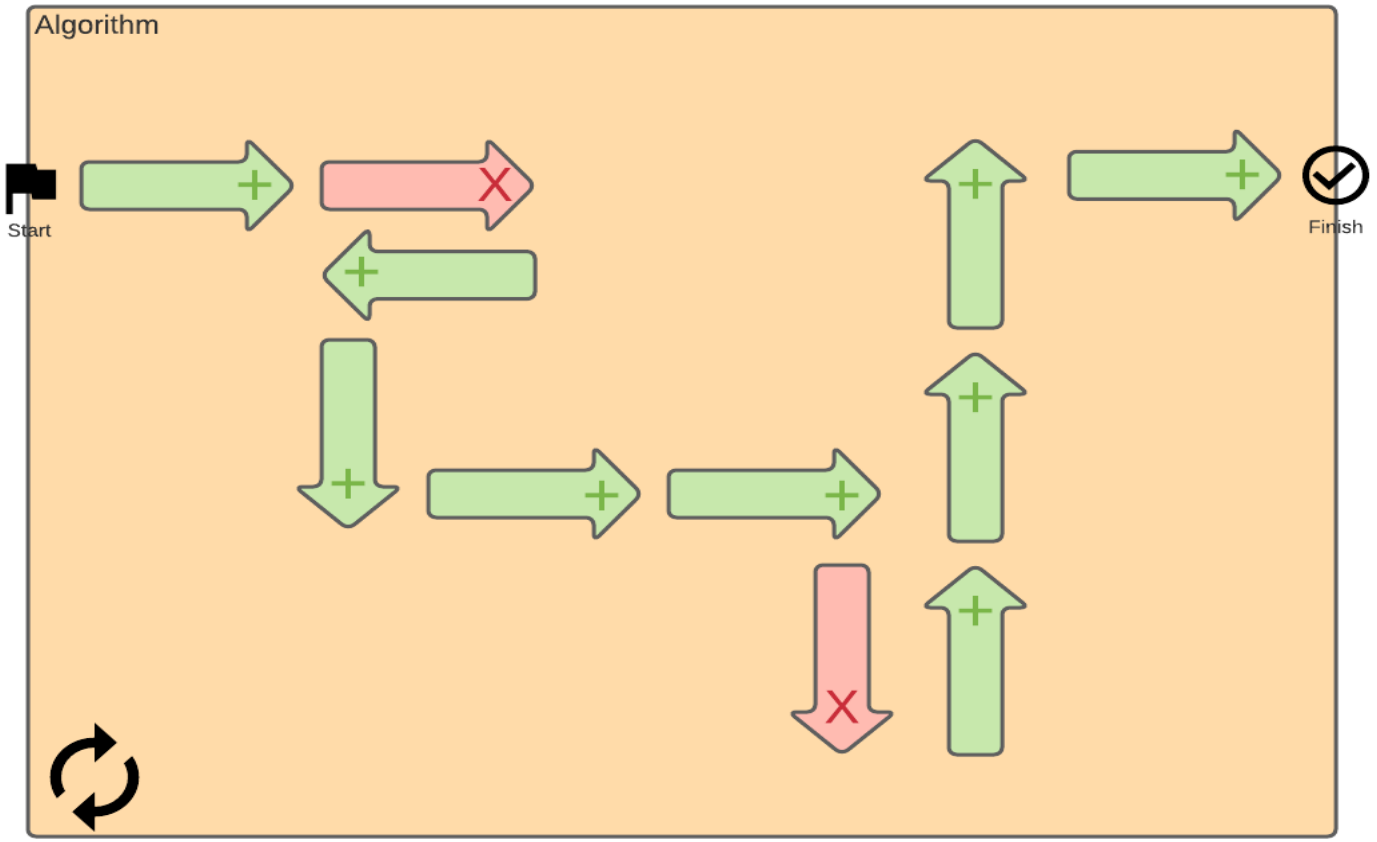
Unsupervised Learning





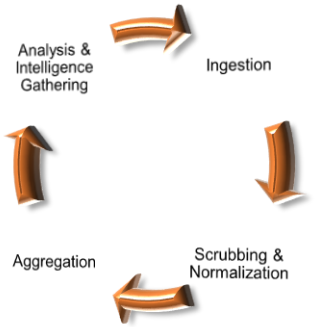
Reinforcement Learning





Data Analysis & Intelligence Gathering – What About the Edge?

- With powerful systems at the Edge, sophisticated analysis can be localized
- Can also provide controlled environment for validating algorithms
- Potential options for additional layers of optimization prior to Cloud transmission
- With K8S and containerization at Edge, supports “smart workflow”



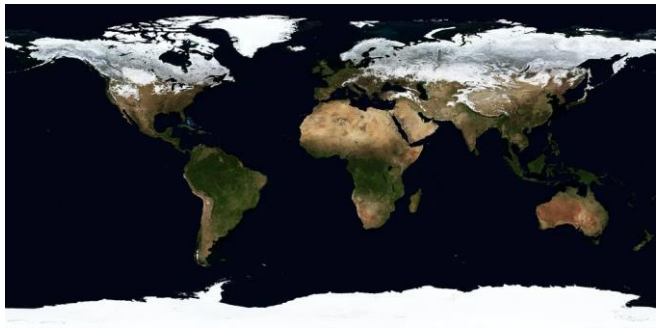
Security & Compliance

Standards & Compliance Categories

- Can include:
  - By geographical region
  - By industry
  - By technology

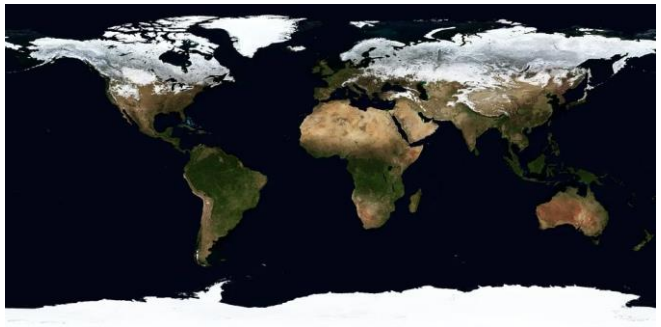
Standards & Compliance by Region

- Standards and compliance enforcement can vary by area of the world
- For example, the EU likely has different requirements than the US:
  - Federal Communications Commission (FCC) certification in the US
  - General Data Protection Regulation (GDPR) in the EU
- Other likely apply regardless of locality
  - PCI
  - SOX



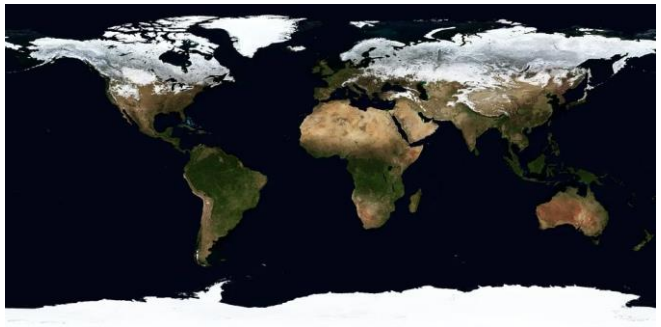
Standards & Compliance by Region

- Can include considerations for:
  - How data is transmitted
  - How data is secured, managed, and used
  - Physical or systems security of the device or Edge component itself



Standards & Compliance by Region

- Failure to adhere can limit ability to do business in the region
- Or can result in significant penalties and/or reputational damage
- Can add permutations to approach to build out of the tech



Standards & Compliance by Industry

- Different industries may have different regulations
- There can also be a difference in physical requirements
- Think remote oil field vs. data center vs. nuclear power plant

Standards & Compliance by Industry

- Regulations often driven by types of data being gathered
- Medical devices likely subject to HIPAA regulations
- Point-of-Sale (POS) devices may require PCI compliance

Standards & Compliance by Industry

- Depending on the industry, failure to comply may have devastating impact
- Think potential exposure for autonomous vehicles, for example

Standards & Compliance by Technology

- Azure Security → https://invidgroup.com/how-secure-is-microsoft-azure/
- AWS Security → https://docs.aws.amazon.com/whitepapers/latest/aws-overview/security-and-compliance.html
- GCP Security → https://medium.com/google-cloud/is-my-data-safe-in-cloud-41608c1d1f89

Assessing Security Risks

- To secure a solution, attack surfaces and potential threats must be identified
- Common practice utilizes something called threat modeling
- Includes modeling and analyzing possible attack vectors based on application

Assessing Security Risks

- Risk assessment should account for different “zones” of execution
- Security requirements for device in remote oil field different from secure data center
- And, ideally, threat modeling would be executed during design & dev phases

Securing Data in Motion

- Security required as data flows through the ether between producer and consumer
- If attacker able to intercept information flowing between the two:
  - Potentially exposes sensitive information contained within header or payload
  - Could allow insertion of alternate, damaging detail or control instruction

Securing Data in Motion

Develop Intelligence

- Certificate/secrets-based Transport Layer Security (TLS) can be used to protect
- Highlights need to protect security keys
- Impact can range from trivial to devastating (depending on application)

Securing Data at Rest

- Aggregated data stored in plain text can create a vulnerability
- In previous topic on data management, goal is gained intelligence from the data
- If the data at rest has been compromised:
  - May lead to inaccurate conclusions from analysis
  - Could provide competitor or bad actor access to a company's competitive advantage
- As with "in motion", certificate-based encryption in storage is key

DevOps & Automation

Develop Intelligence

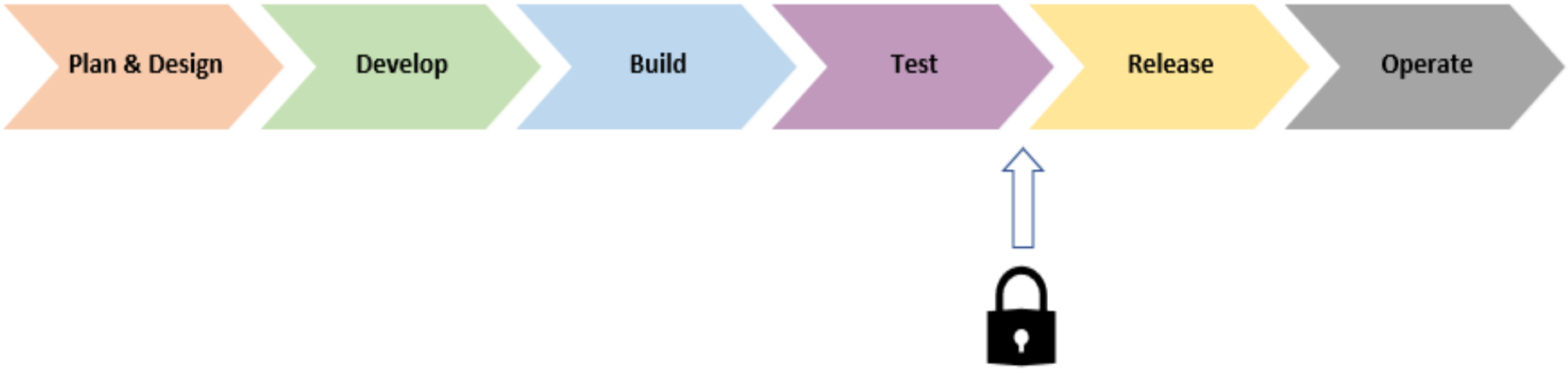
- As with other software applications, DevOps can provide lift
- Key principle of DevOps is automation
- Continuous Integration & Continuous Delivery can be applied as well

DevOps & Automation

- Presents opportunities to apply scripting to:
  - Automate onboarding, offboarding, and configuration
  - Deployment & configuration of Edge components
  - Deployment & configuration of Cloud services used to aggregate & analyze data
- Practicing principles of DevSecOps helps ensure security is “shifted left”

DevOps & Automation

- What is often done:



What are the potential challenges with taking this approach?

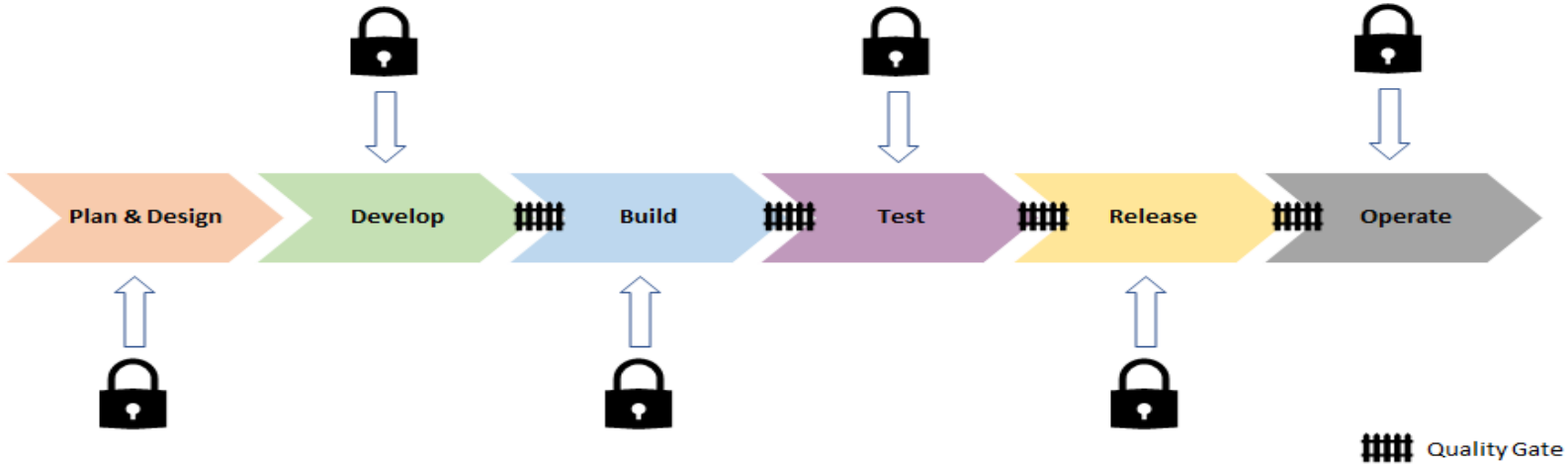
DevOps & Automation

Develop Intelligence

- Potential challenges:
  - At this point, there may not be enough time in schedule to absorb change
  - Activities required to remediate may be complex
  - May require revisit of one or more previous phases to properly address

DevOps & Automation

- Better approach:



DevOps & Automation

Develop Intelligence

- Plan & Design – Threat modeling, data protection, and risk assessment
- Develop – SAST (Static Application Security Testing) tools
- Build – SAST and SCA (Software Composition Analysis) tooling
- Test – DAST (Dynamic Application Security Testing) tools, passive/active scans and “fuzzing”
- Release – Additional security-specific scanning, port scans, and log validation
- Operate – Monitoring & alerting, RCA (Root Cause Analysis)

DevOps & Automation

Develop Intelligence

- Quality gates guard against moving security defects forward
- In true DevOps fashion:
  - Information gathered from early phases feeds into later phases
  - Lessons learned feed continuous improvement of overall process

Data Visualization & Dashboarding

- Computers are great at “crunching” large amounts of raw data
- For humans, sometimes a “picture is worth a million bytes”
- With dashboarding & graphical visualization, it can be much easier to see trends
- Data science and forecasting can help with extrapolating for the future

Data Visualization & Dashboarding

- Tools like Power BI or Tableau (among others) provide powerful options
- Able to integrate through connectors with multiple data sources
- Visualizations & charts can be layered onto large datasets to provide insight

Data Visualization & Dashboarding

- Often the tools support regular data refresh for fuller picture over time
- Tools leverage defined authentication/authorization against data sources
- Helps ensure ongoing, end-to-end security of key data

Cost Management

- CAPEX vs. OPEX
- Azure → https://azure.microsoft.com/en-us/pricing/calculator/
- AWS → https://calculator.aws/
- GCP → https://cloud.google.com/products/calculator/
- Key is – ensuring accounting for ALL components
- Drives ROI (Return on Investment) & CBA (Cost-Benefit Analysis)

Evaluate Case Studies – Operational Concerns

THANK YOU

Develop
Intelligence


Develop
Intelligence