Production Support – SRE in real world

Nidhi Singh

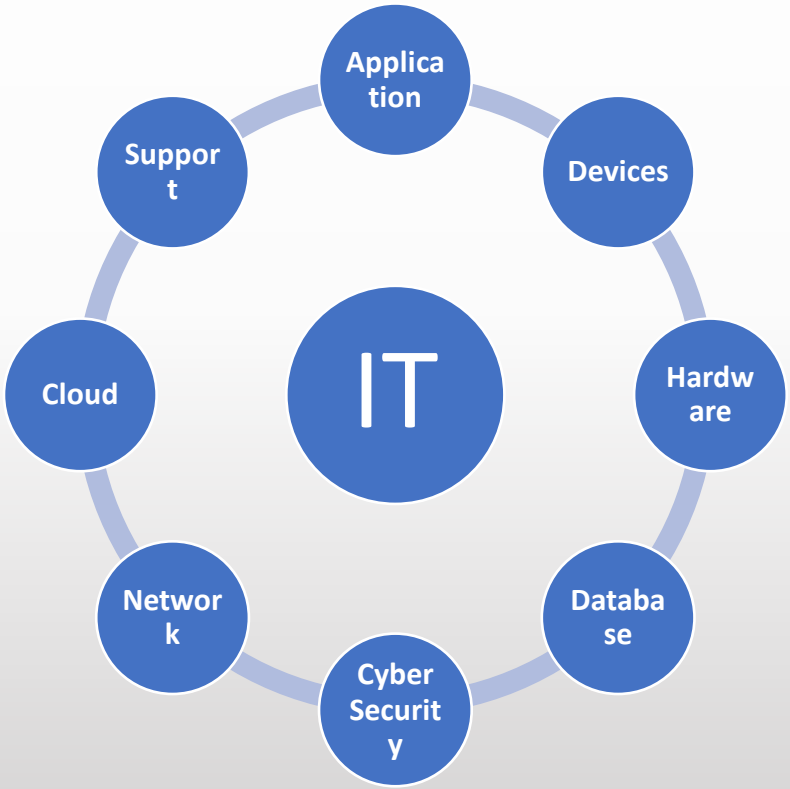
Content

Section -1

What is IT (information technology) System

What is IT (information technology) System

- Information Technology is the application of technology to solve specific business or organizational problems on a large scale or to serve a purpose
- Purpose can be of entertainment, ecommerce, medical services, telecom
- Productivity and efficiency improvement
- Globalization and bridge cultural gaps



Types of IT companies

- **Product company**

  Create and deliver the product to wide range of customer

- **Consulting services**

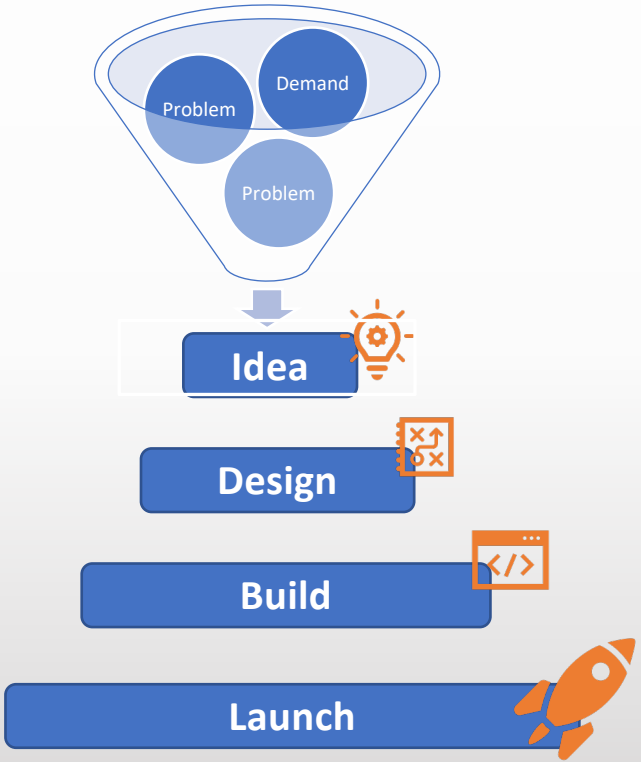
  Provide feed back/support in design and build of product

- **Service implementer**

  Build s/w or product only when customer approaches

Product based company

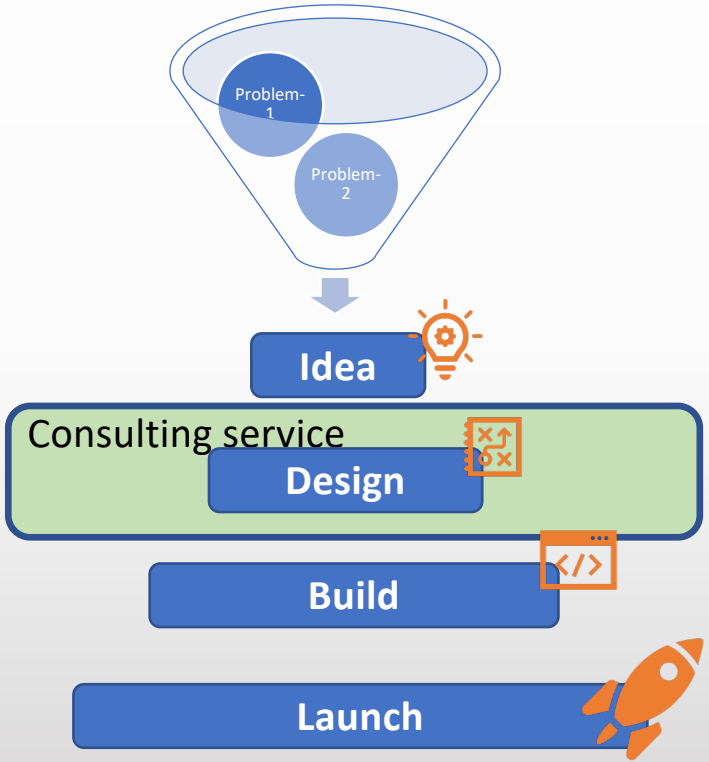
- **Product based companies** identify a problem statement or need of the market and create or design a **product** to fulfill and solve the problem
- Once the **product** is built or application is developed, it is opened to the market
- Product based companies make their product before the demand of the people
- Quality of the product is king



Google
Microsoft
amazon
SAMSUNG
TESLA
facebook
IBM
UBER

Consulting company

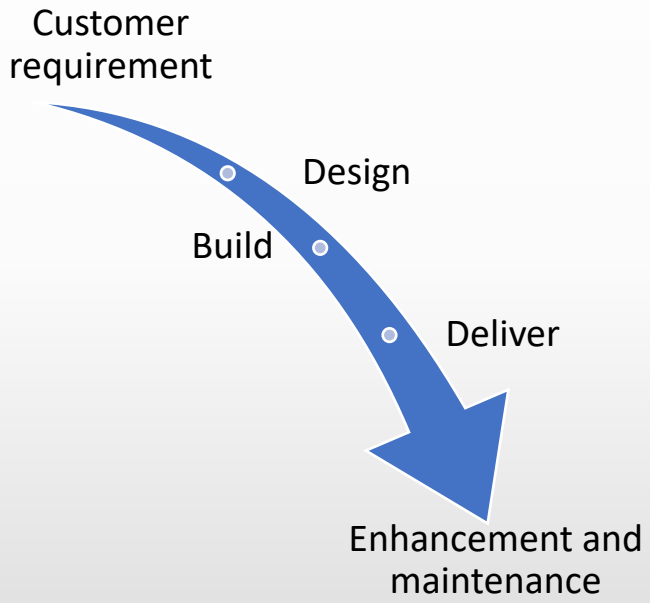
- A **consulting firm** is a **business** comprised of industry-specific experts who offer professional advice, guidance, and actionable solutions to businesses experiencing issues
- It **helps** organizations to improve their business performance in terms of operations, profitability, management, structure and strategy.
- Focus on designing and architecting of specific business requirement from client



accenture
KPMG
HITACHI
Deloitte.
pwc
WIPRO
BAIN & COMPANY
MARSH & McLENNAN COMPANIES
Infosys
bmc
TATA
SAPIENT CONSULTING

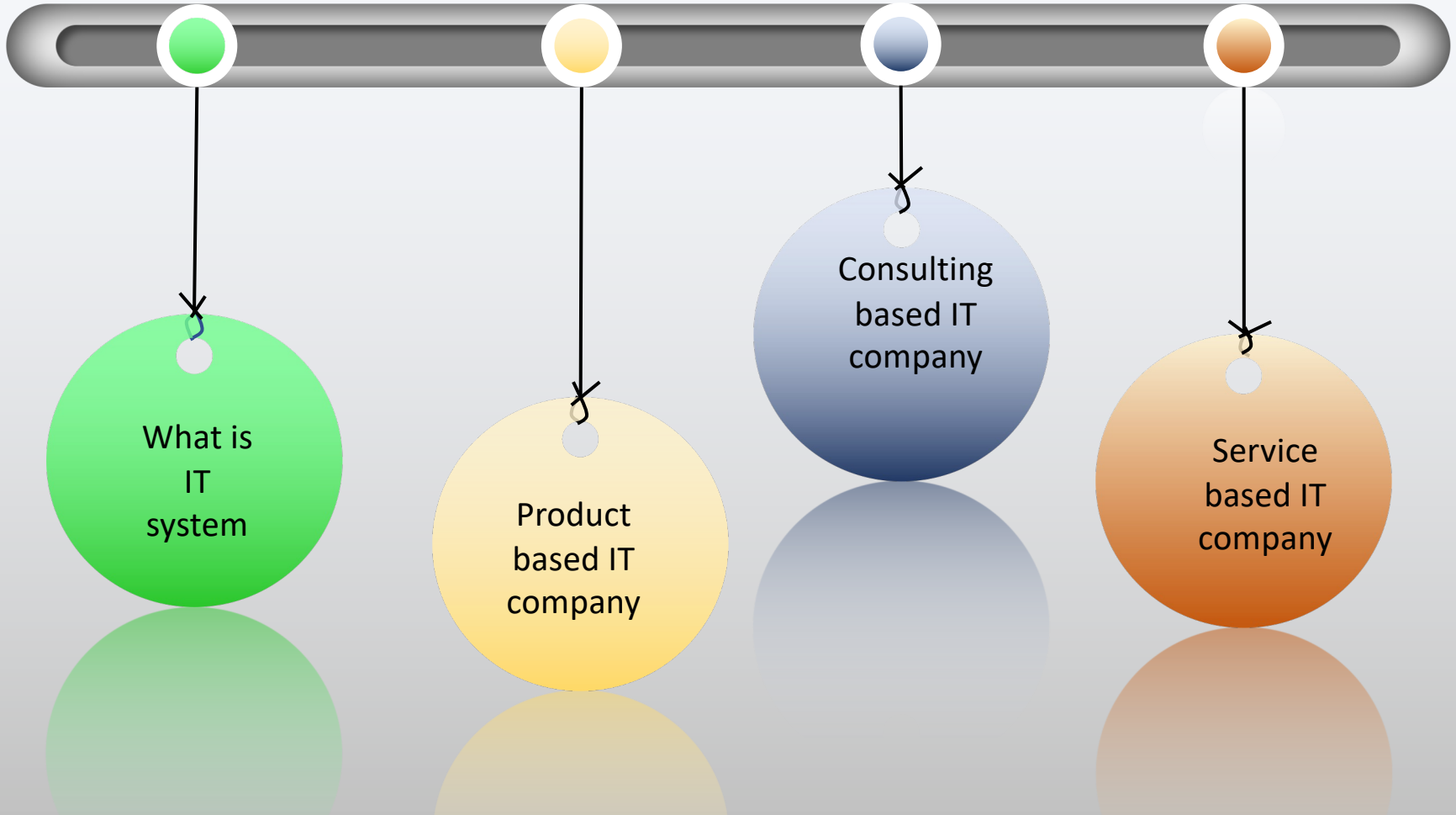
Service company

- A **service company** builds s/w or application **exclusive** to the customer **requirement**
- After delivery of product they provide **warranty** of code and support for any bug fix.
- Customer satisfaction is the **main motive**



IBM
ORACLE
accenture
TATA CONSULTANCY SERVICES
Capgemini
Cognizant
Infosys
wipro

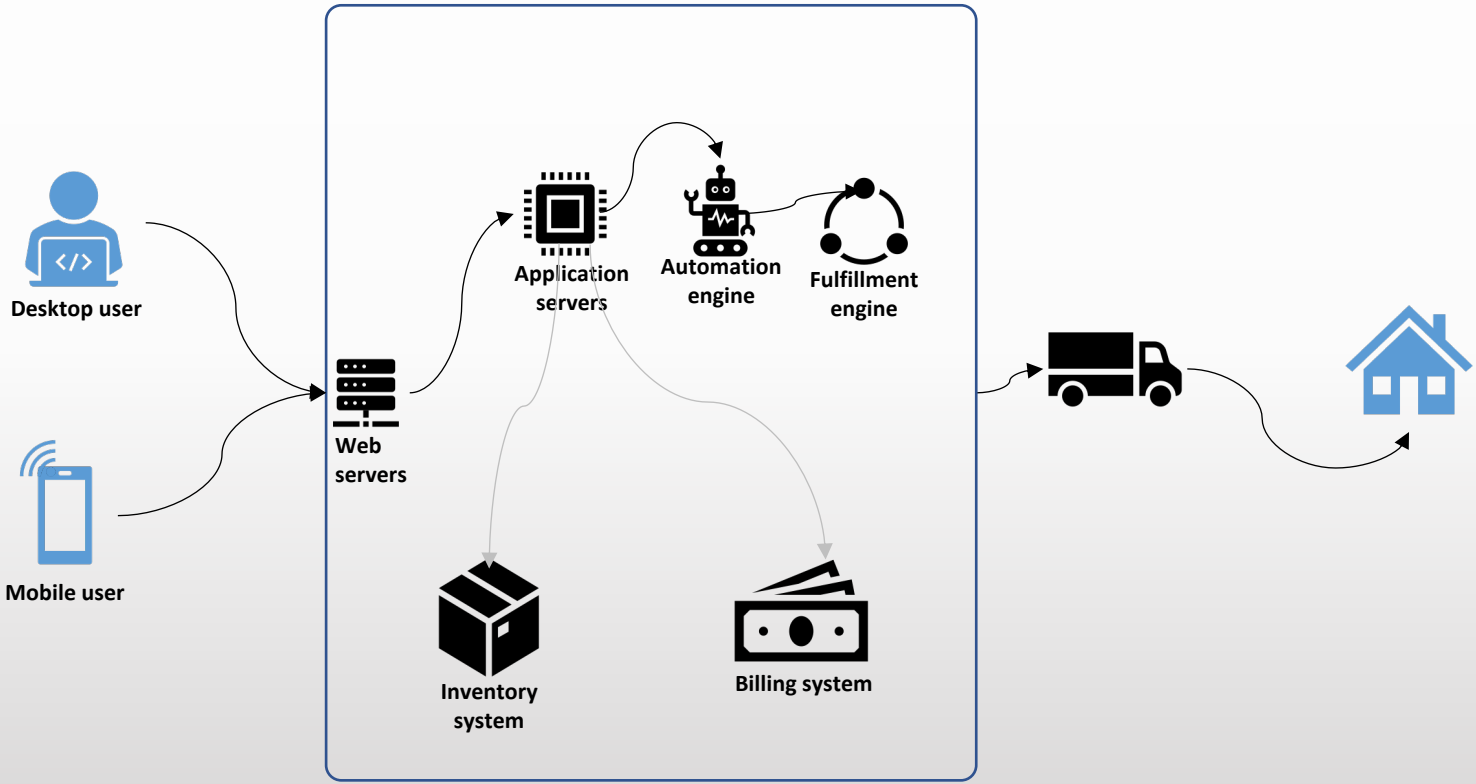
Section 1 - Summary



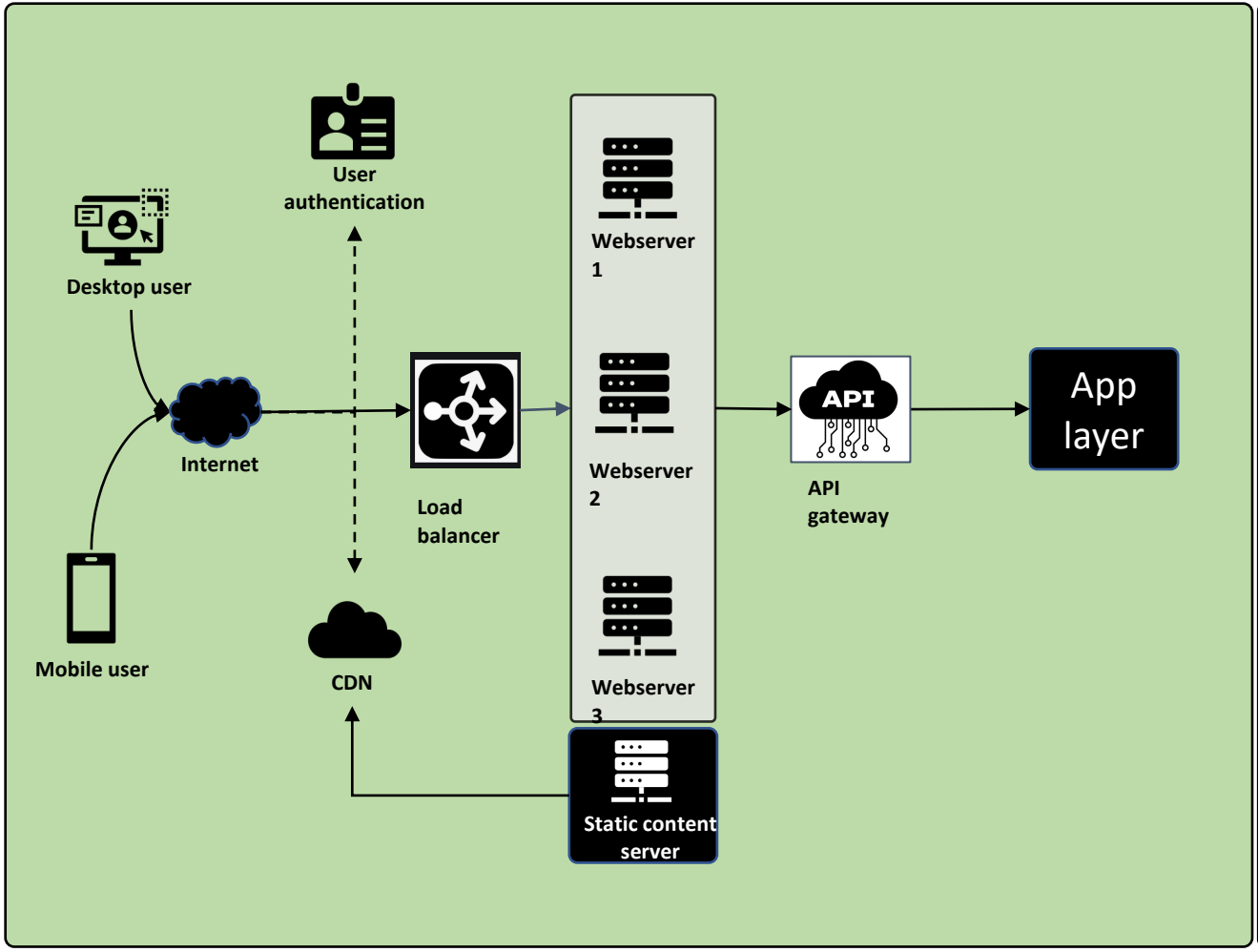
Section - 2

Understanding an application

Understanding an application

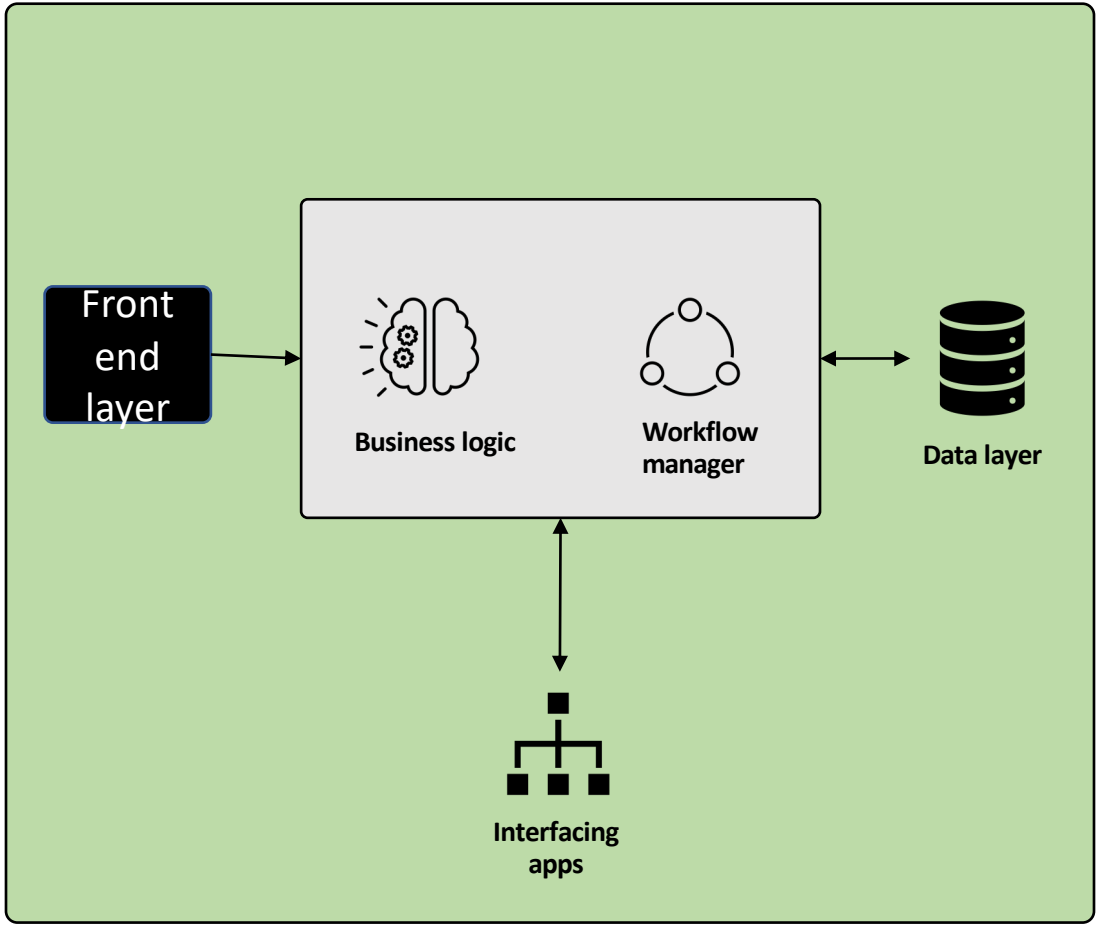


Frontend layer



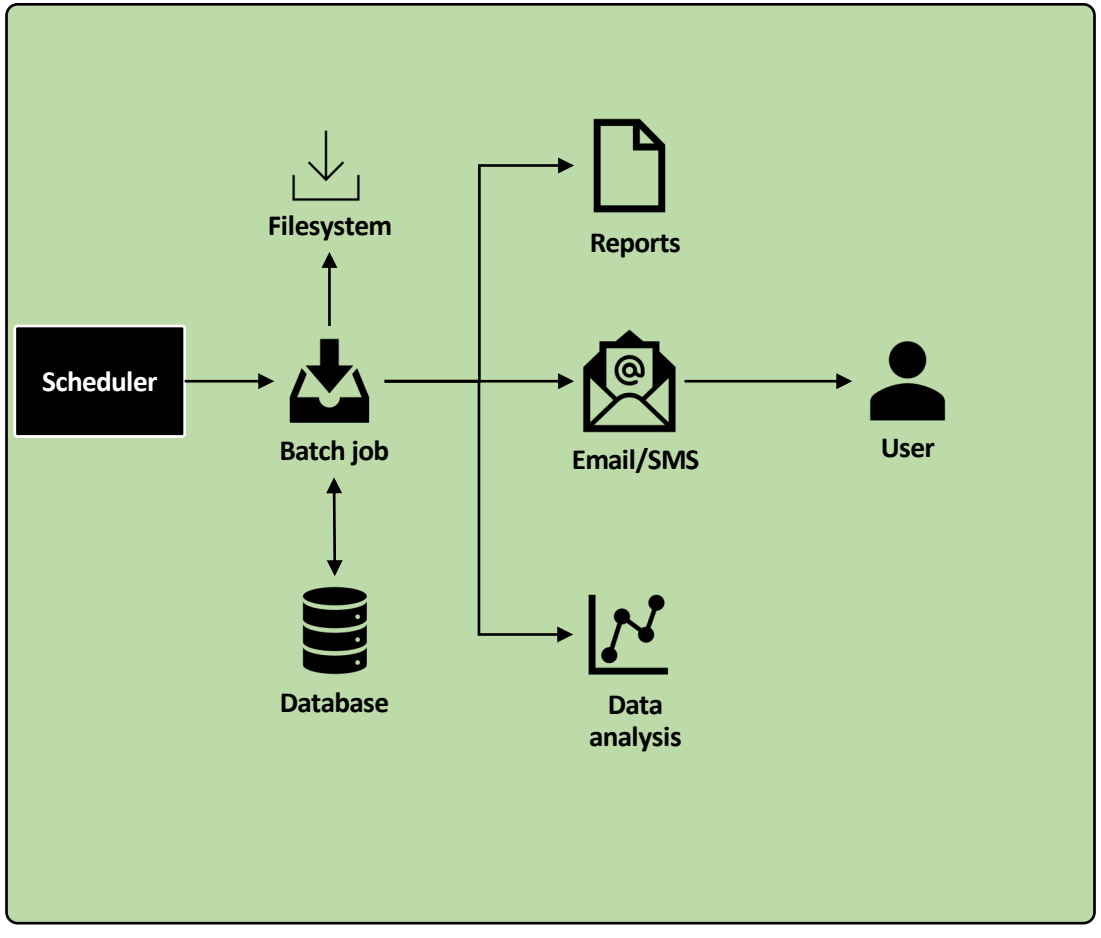
- Direct customer interaction
- Internet or intranet accessible
- User authentication to block fraudsters
- CDN(Content Delivery Network) for high performance
- Distributed traffic hits webservers through load balancer
- API gateway passes on the customer data based on system design

Application layer



- Collect data from frontend and process it based on business logic
- Any certificates or mechanized id/pwd used in application are configured in application layer
- Workflow manager acts based on business logic
- Accordingly interaction with interfacing application takes place
- The processed customer data is stored in data layer

Backend jobs



- Scheduled jobs run to collect customer data for further analysis or processing
- Send the output back to customer either by same flow how the data flew in or depending upon architecture
- Store the customer data
- Several types of reports are generated with processed data for analysis and business improvements
- Data analysis for several purposes

Examples of application

Online ticket booking portal – user view

Online portal

- # of passengers
- Date of travel
- Deals

Select flight

- Different flight options
- Flexibility to change date

Traveler details

- Name
- Age
- Email address
- Residential address

Payment

- Credit card details
- Billing address

Ticket confirmation

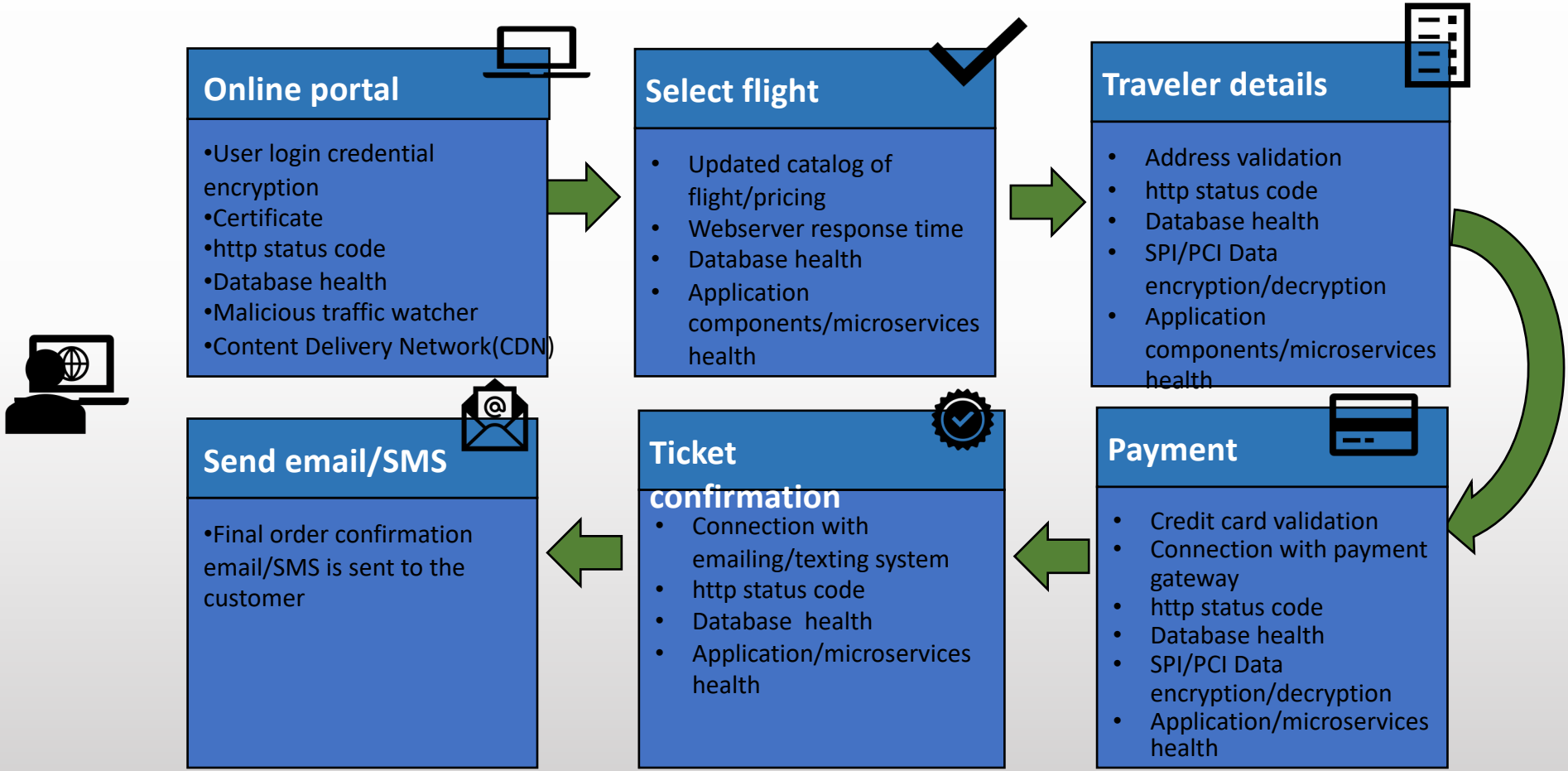
- Ticket confirmation page

Send email/SMS

- Final order confirmation email/SMS is sent to the customer

Examples of application

Online ticket booking portal – SRE view



Examples of application

Online shopping website – user view

Online portal

- Product Listing
- Customer rating

Select the product

- Product details
- Delivery options
- Customer review

Checkout

- Selected items list preview

Delivery detail

- Name/address
- Preferred timing of delivery (if available)

Payment

- Card details
- Billing address

Order confirmation

- Order confirmation page
- Final order confirmation email/SMS is sent to the customer

Examples of application

Online shopping website – SRE view

Online portal

- http response code
- Certificate
- Product catalog update
- User data encryption
- Database health
- Malicious traffic watcher

Select the product

- http response code
- Database health
- Application/microservices health

Checkout

- http response code
- Database health
- Application components/ microservices health

Delivery detail

- http response code
- Database health
- Address validation
- SPI/PCI Data encryption/decryption
- Application/ microservices health

Payment

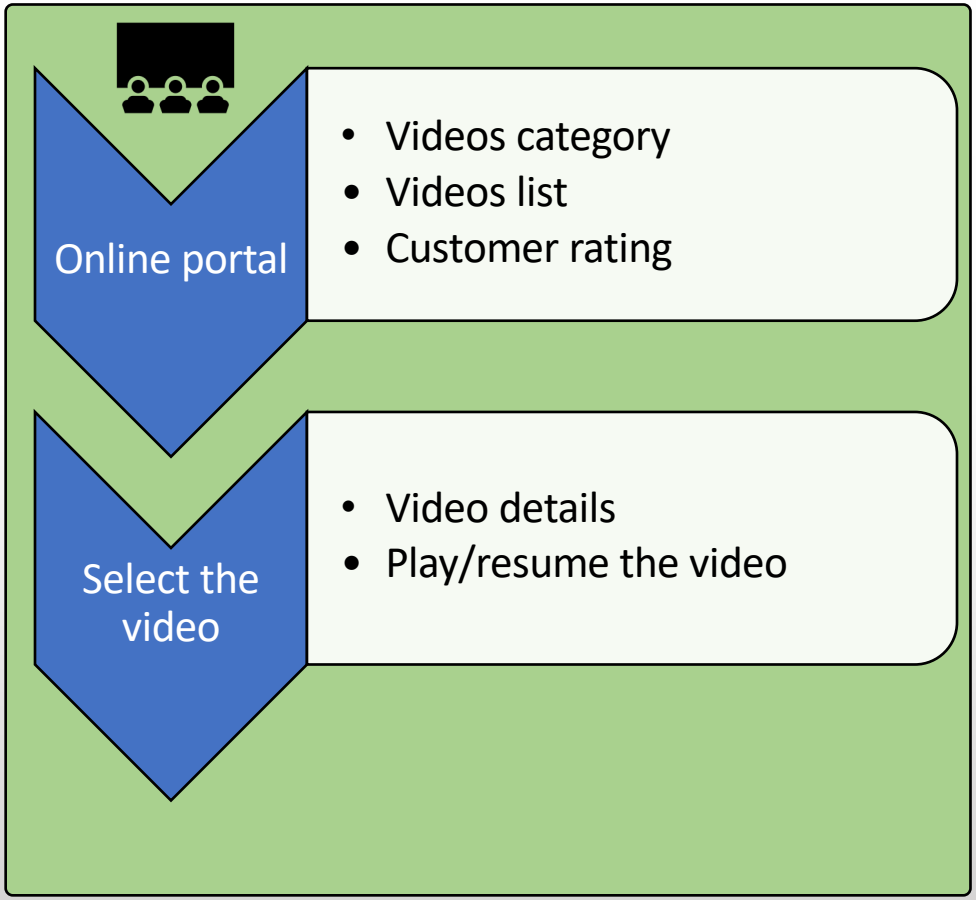
- Card validation
- Address validation
- http response code
- Database health
- SPI/PCI Data encryption/decryption
- Application/ microservices health

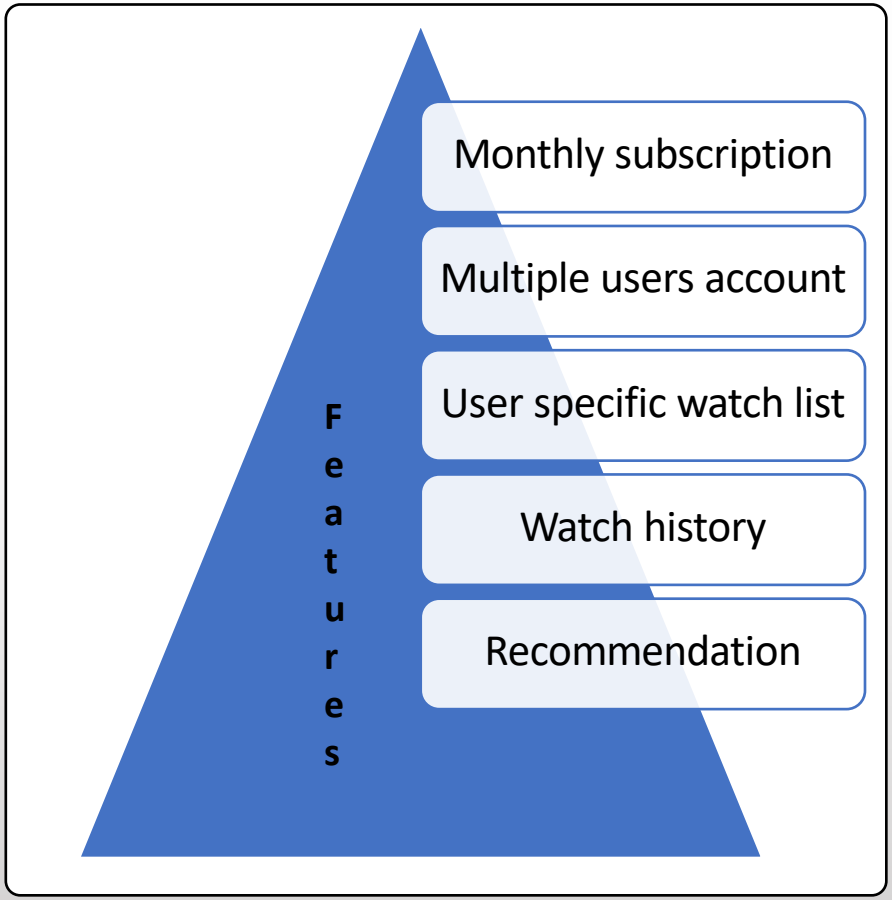
Order confirmation

- Connection with email/text service provider
- http response code
- Database health
- Application/ microservices health

Examples of application

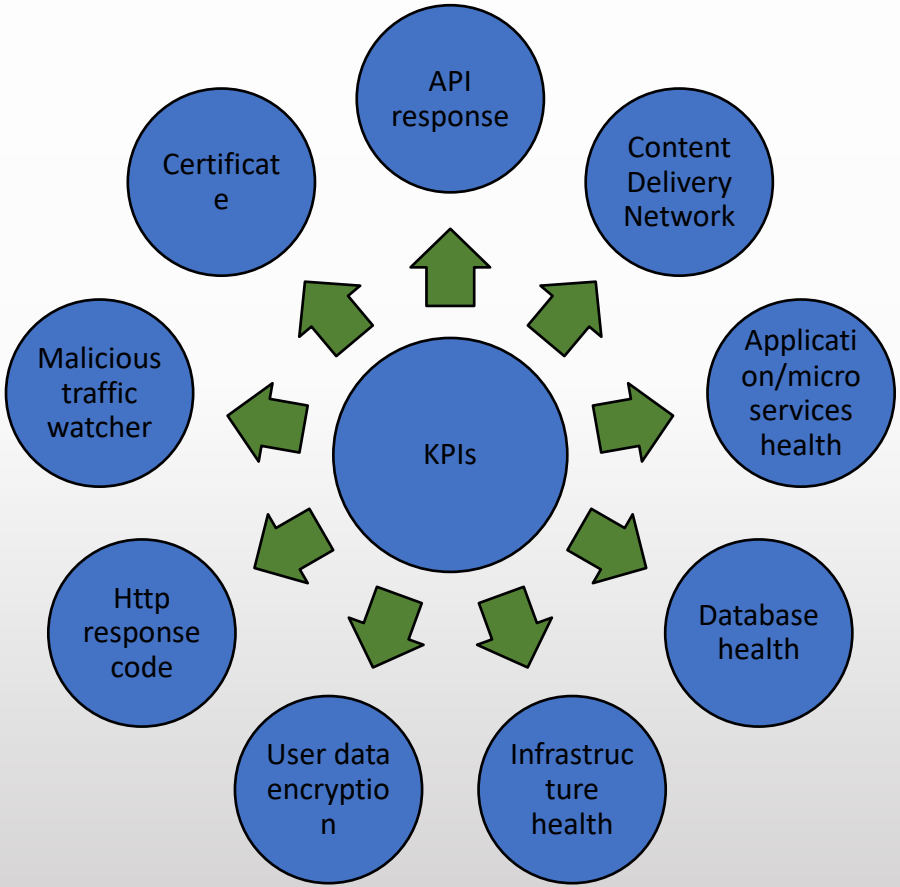
Entertainment site – user view





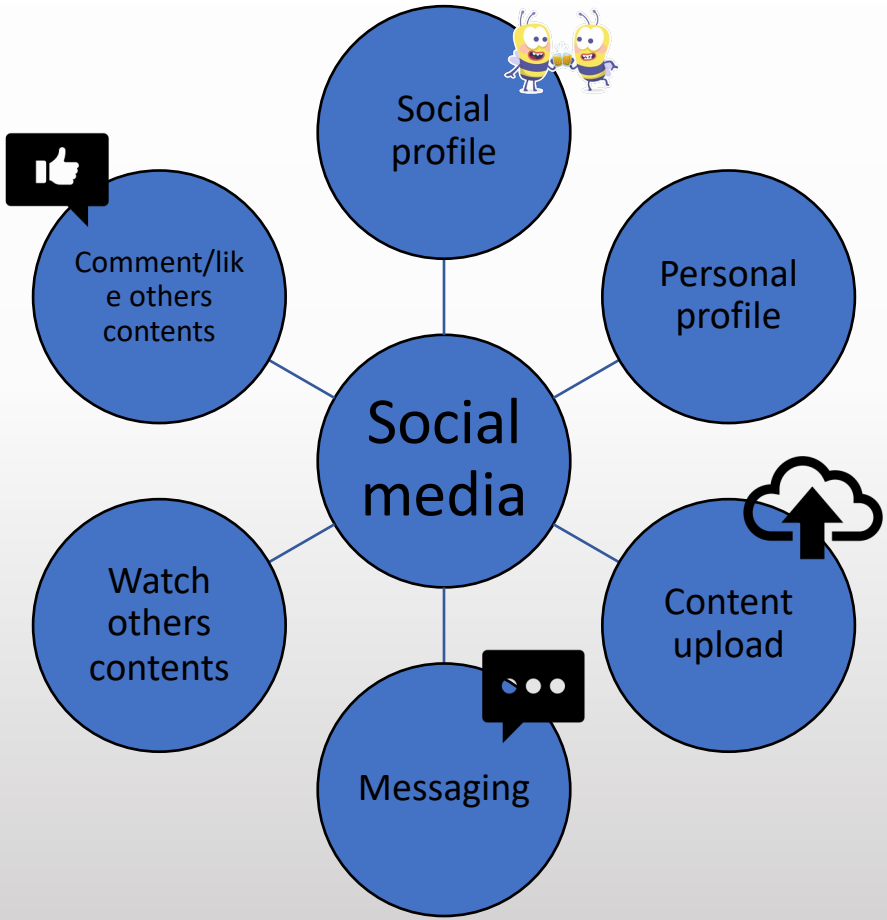
Examples of application

Entertainment site – SRE view



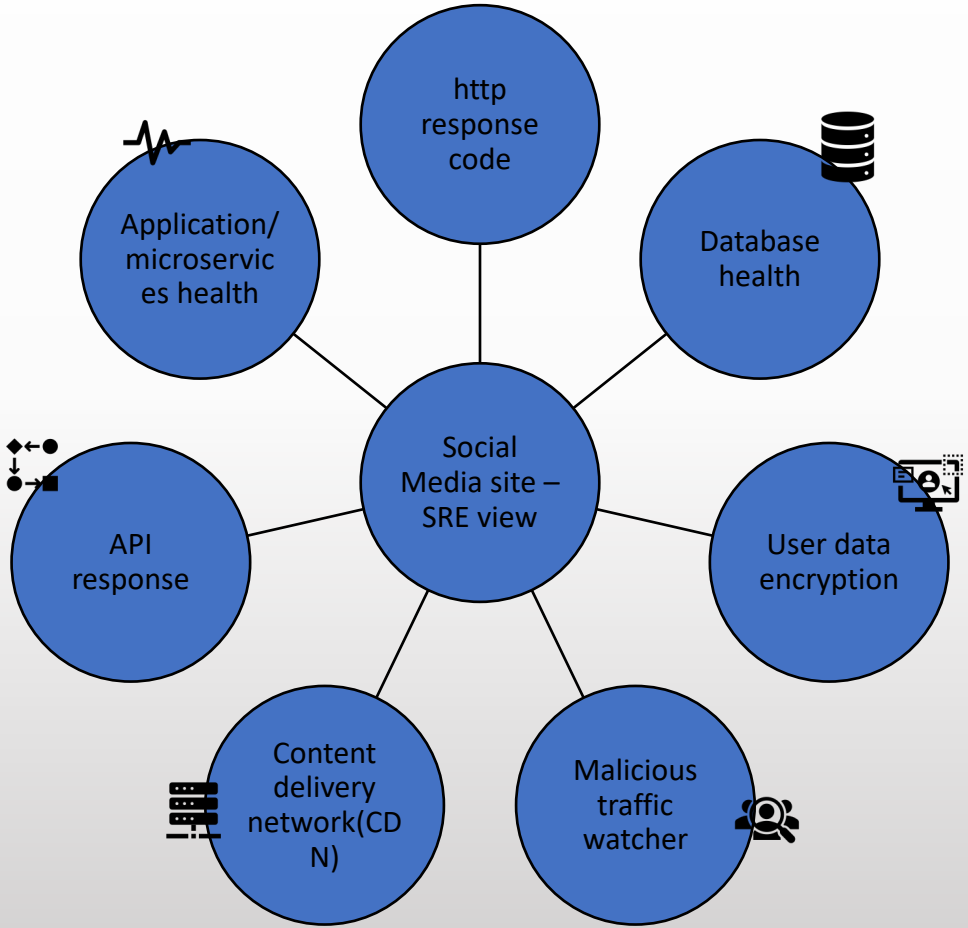
Examples of application

Social media site – user view



Examples of application

Social media site – SRE view



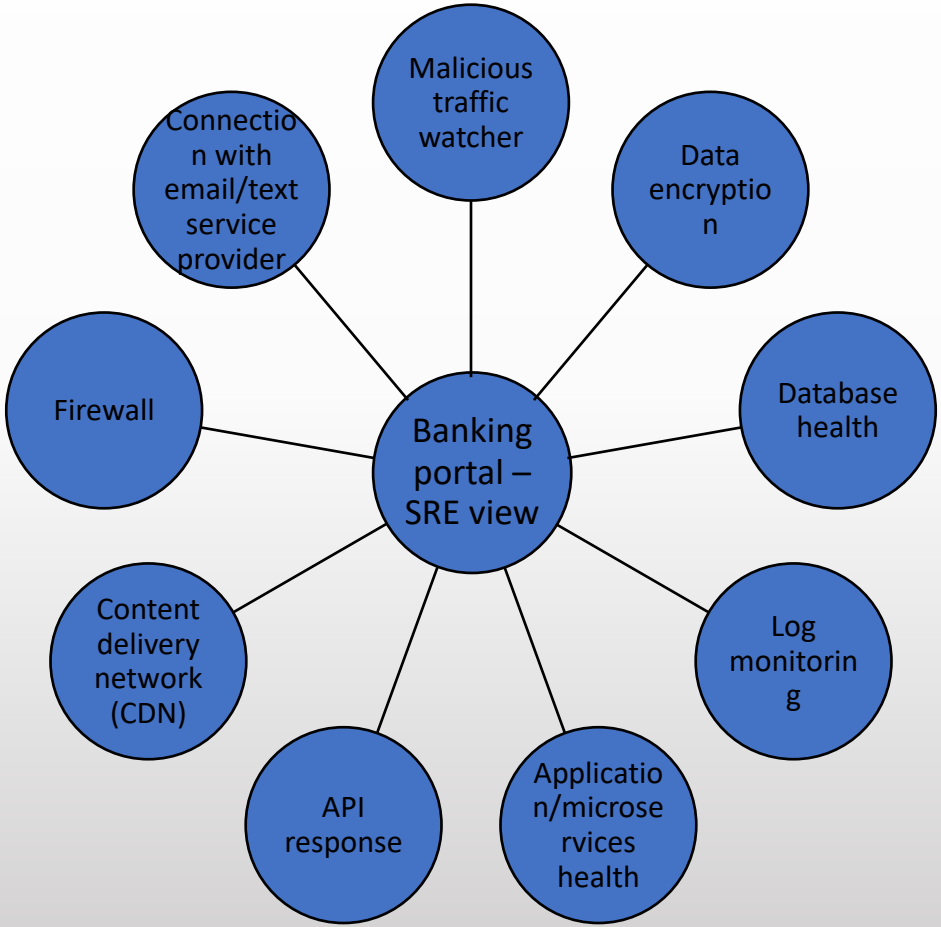
Examples of application

Banking portal – user view

Account	Credit card	Banking transaction	Loan	Insurance	Saving/investment plan
• Saving account • Checking account	• Travel card • No annual fee • Low rate Card	• Account Summary • 3rd party transaction	• Home loan • car loan	• Home insurance • Car insurance	• Tax saving fund • Retirement fund

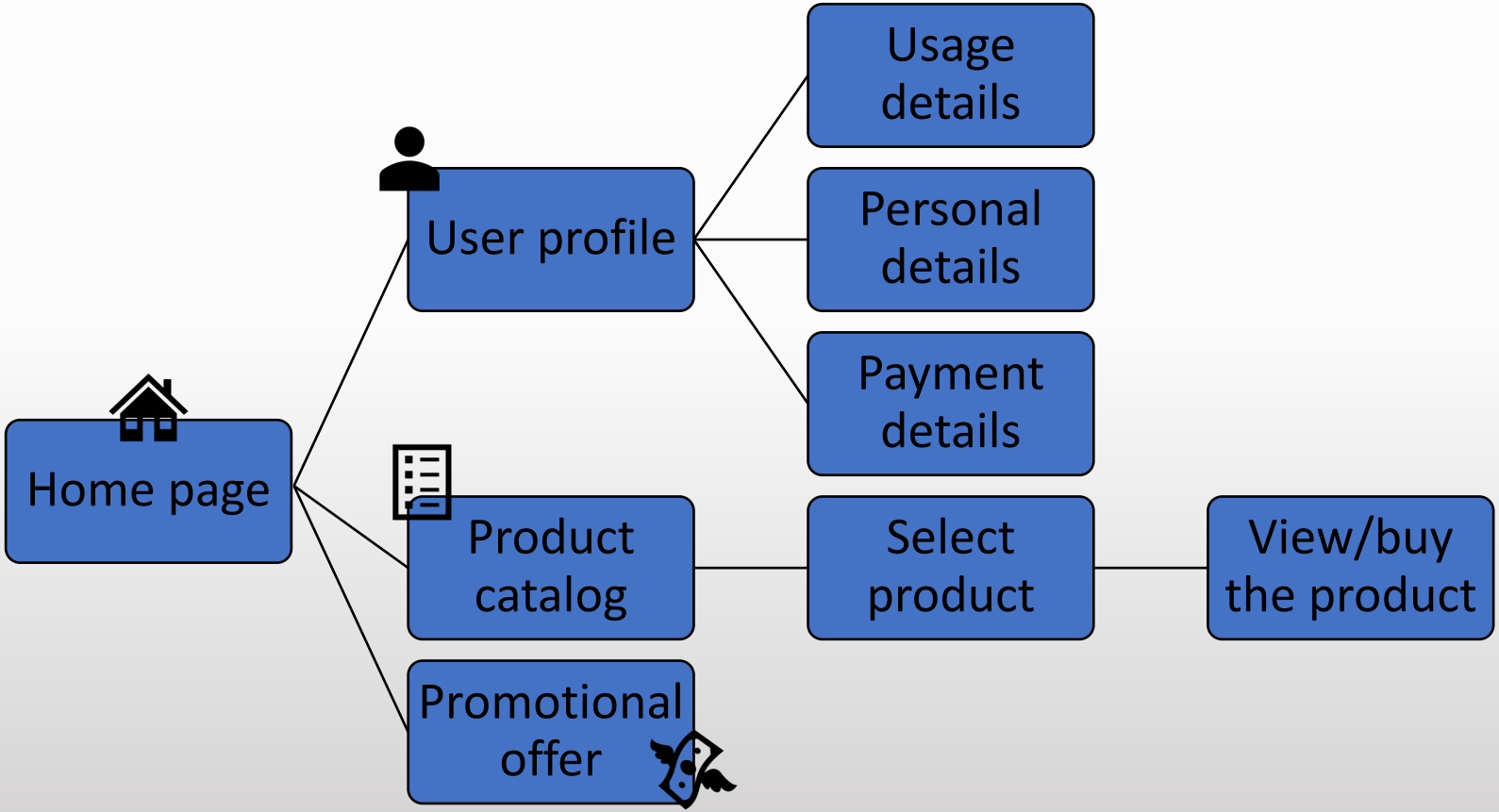
Examples of application

Banking portal – SRE view



Examples of application

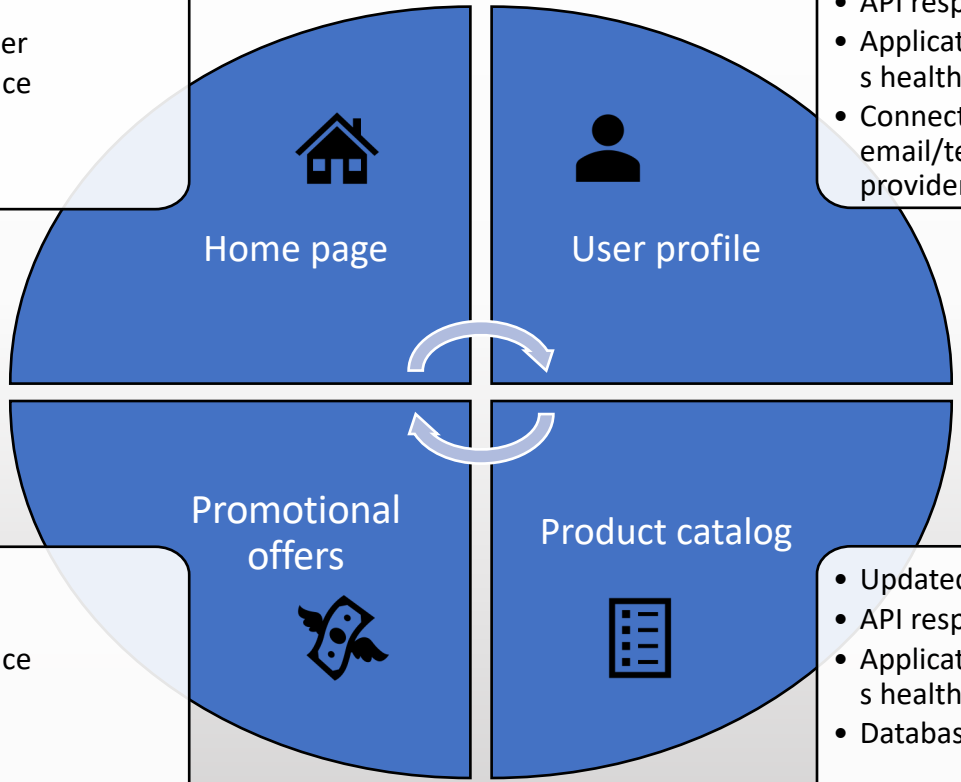
Utility app – user view



Examples of application

Utility app – SRE view

- http response code
- API response
- Malicious traffic watcher
- Application/microservice s health
- Certificate

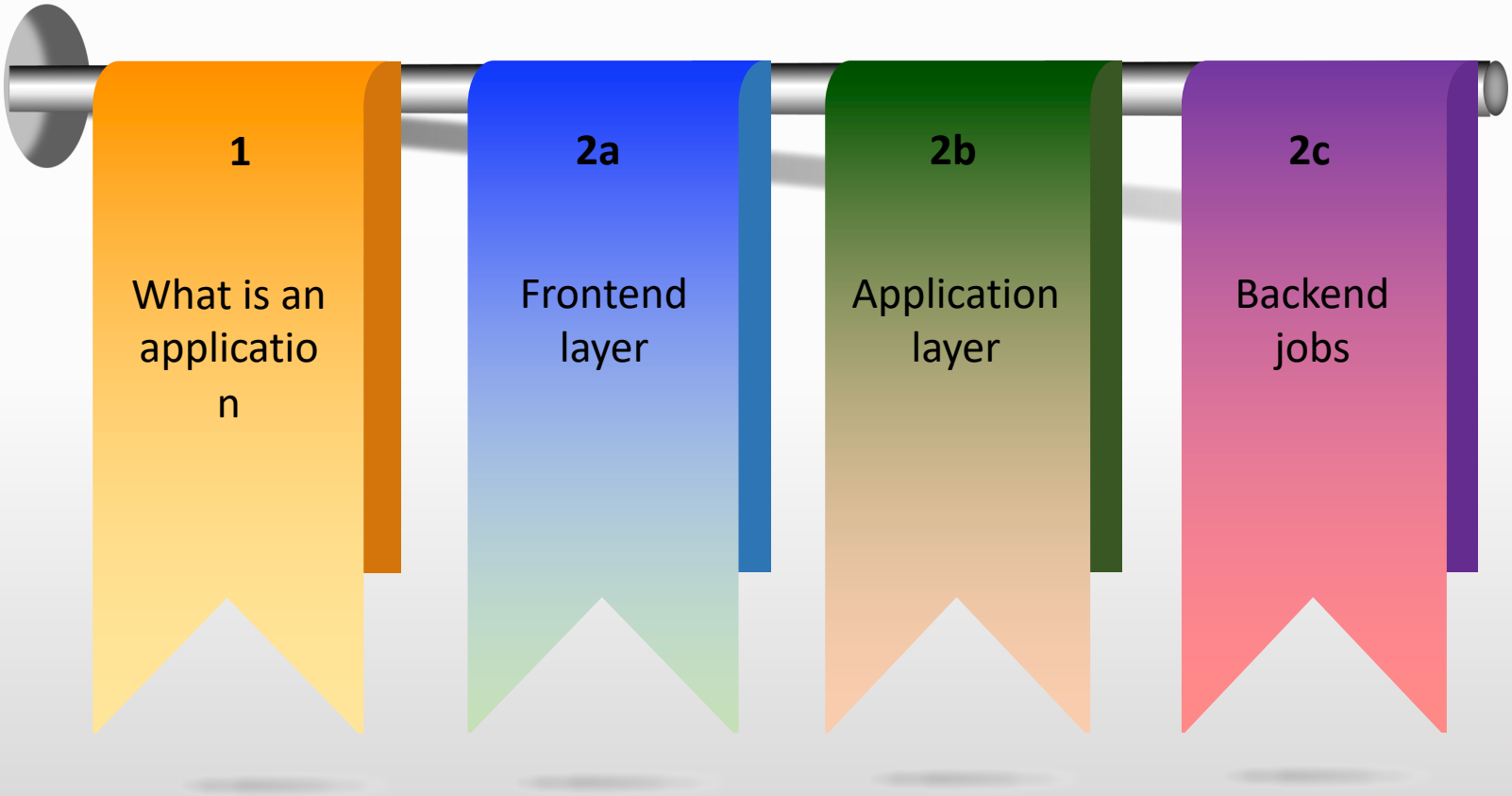


- Data encryption
- API response
- Application/microservice s health
- Connection with email/text service provider

- Updated list of offers
- Database health
- Application/microservice s health

- Updated catalog
- API response
- Application/microservice s health
- Database health

Section -2 Summary



Section - 3

Users types

Users

Who are these users

Application users/Consumer

- Directly interact with application
  - *Example: banking portal*
- Consumers internal to organization(direct employee)
  - *Example: help desk ticketing portal*
- Federated users – applications used by third party vendors
  - *Example: Client's timesheet*

Application integrations

- Interfacing application
  - *Example: Schedular job*
- IoT devices
  - *Example: Scanner app*

Application users – direct consumer

Characteristics :

- Direct customer interface
- Fast GUI response
- Such applications are very critical due to business impact
- Example: Netflix, HBOMax, Facebook, etc

What it means to SRE :

- Customer login success/failure rate
- User experience monitoring
  - API response
  - HTTP status code
- Pattern of failure responses
- Certificate validity
- Usage volume monitoring
- Network watcher

Application users – consumer internal to organization

Characteristics :

- These applications are for employee internal to organization
- Criticality can be identified based on business usage
- Example: HR portal, internal helpdesk

What it means to SRE :

- User experience monitoring
  - API response
  - HTTP status code
- Pattern of failure responses
- Certificate validity
- Intranet and internet connections
- Connection origin/source(geography)

Application users – Federated users

Characteristics :

- Users of such applications are third party vendors but they are authorized to use client applications
- Criticality can be identified based on business usage
- Example: VPN application, client timesheet portal

What it means to SRE :

- Connection origin/source(geography)
- Monitoring user access vs activity
- User experience monitoring
  - API response
  - HTTP status code
- Pattern of failure responses
- Certificate validity
- Intranet and internet connections

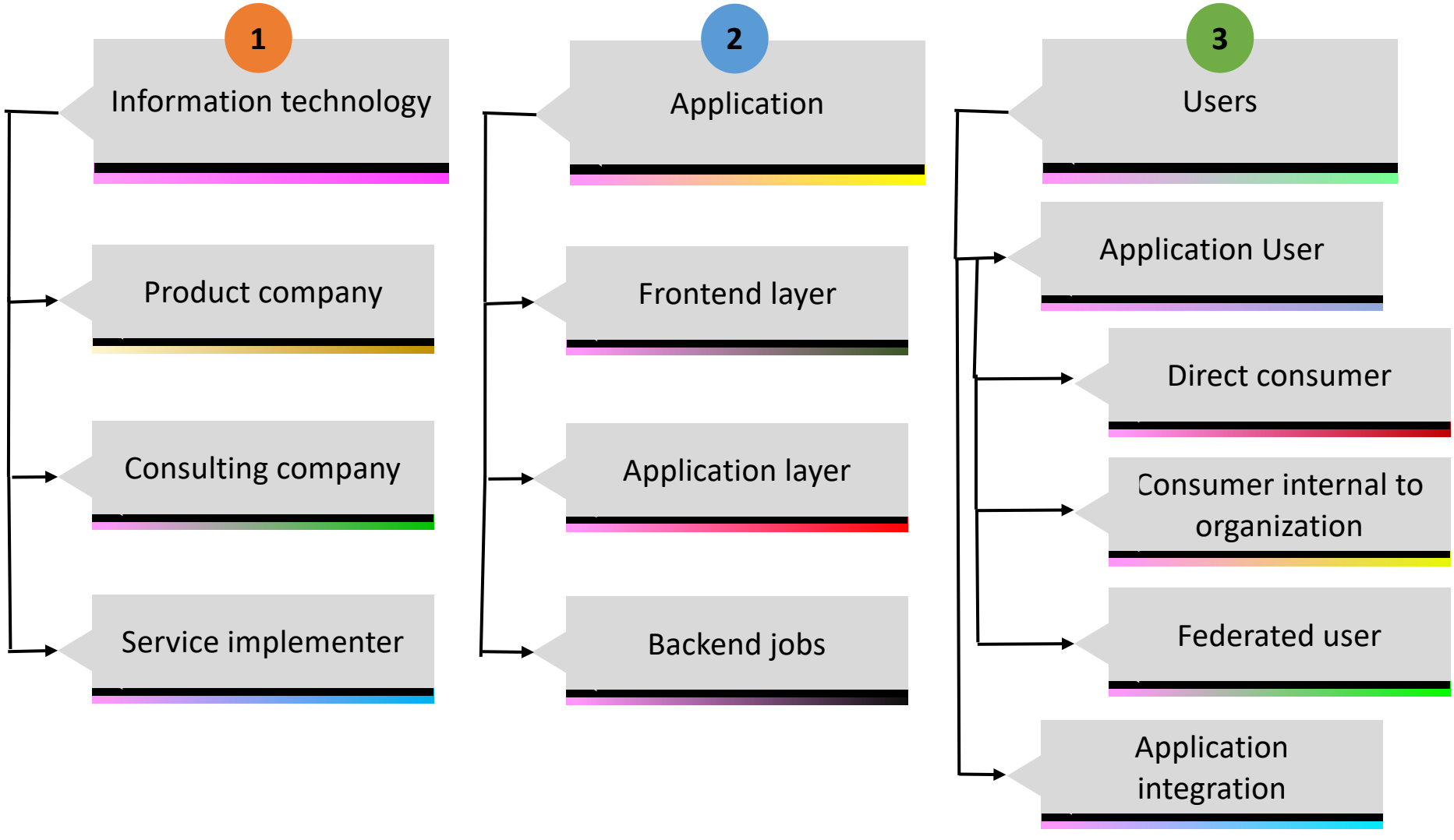
Application integrations

Characteristics :

- Under this category the users by itself are applications
- These can be interfacing applications which takes input from frontend application and passes the data to backend application for processing
- These can be IoT apps which collects data and sends to another application
- Example: Scanner application, WhatsApp

What it means to SRE :

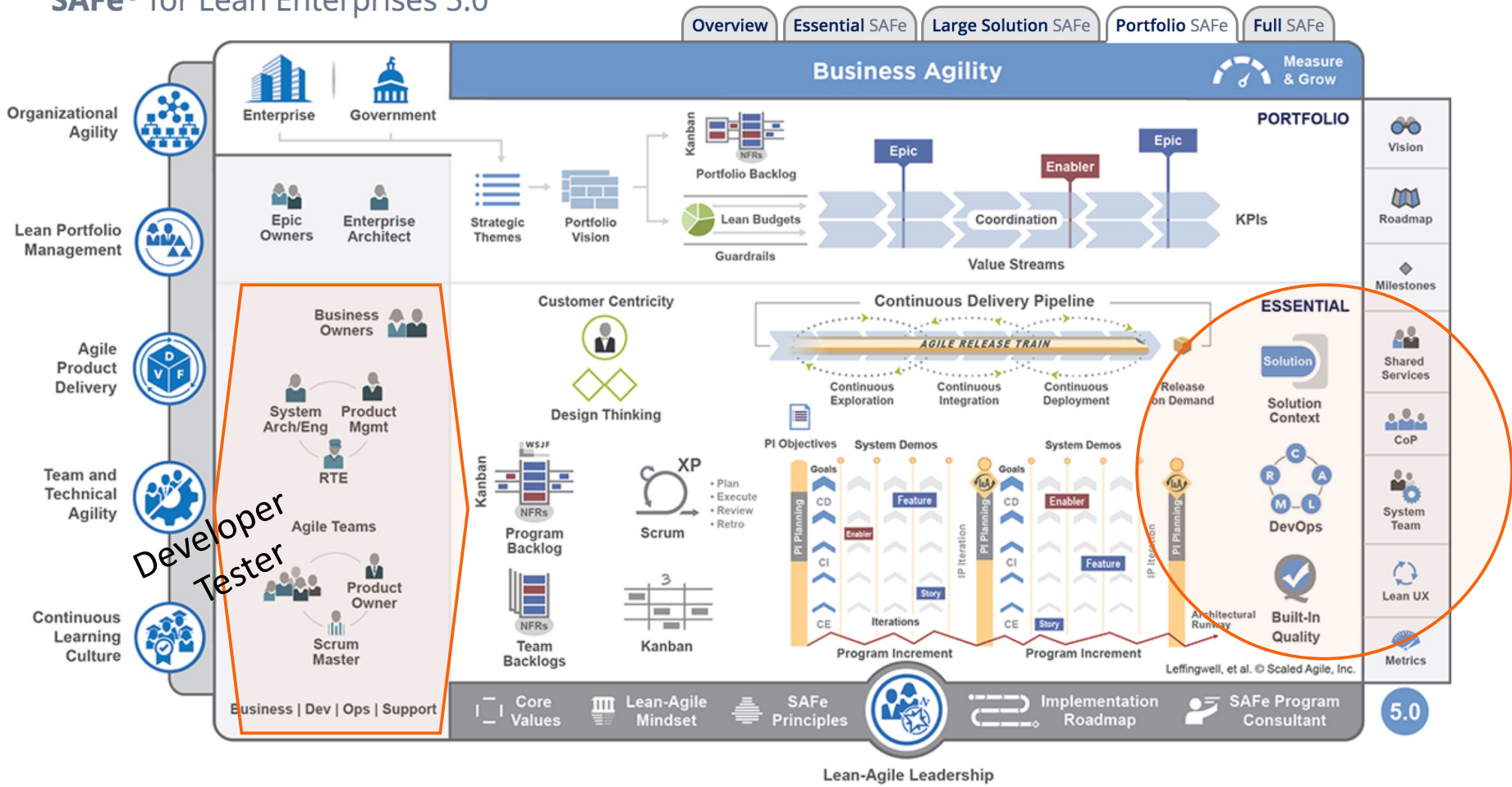
- User traffic volume monitoring
- Token expiration monitoring
- API response monitoring
- API error pattern monitoring
- Interface application connectivity


1
Information technology
Product company
Consulting company
Service implementer
2
Application
Frontend layer
Application layer
Backend jobs
3
Users
Application User
Direct consumer
Consumer internal to organization
Federated user
Application integration

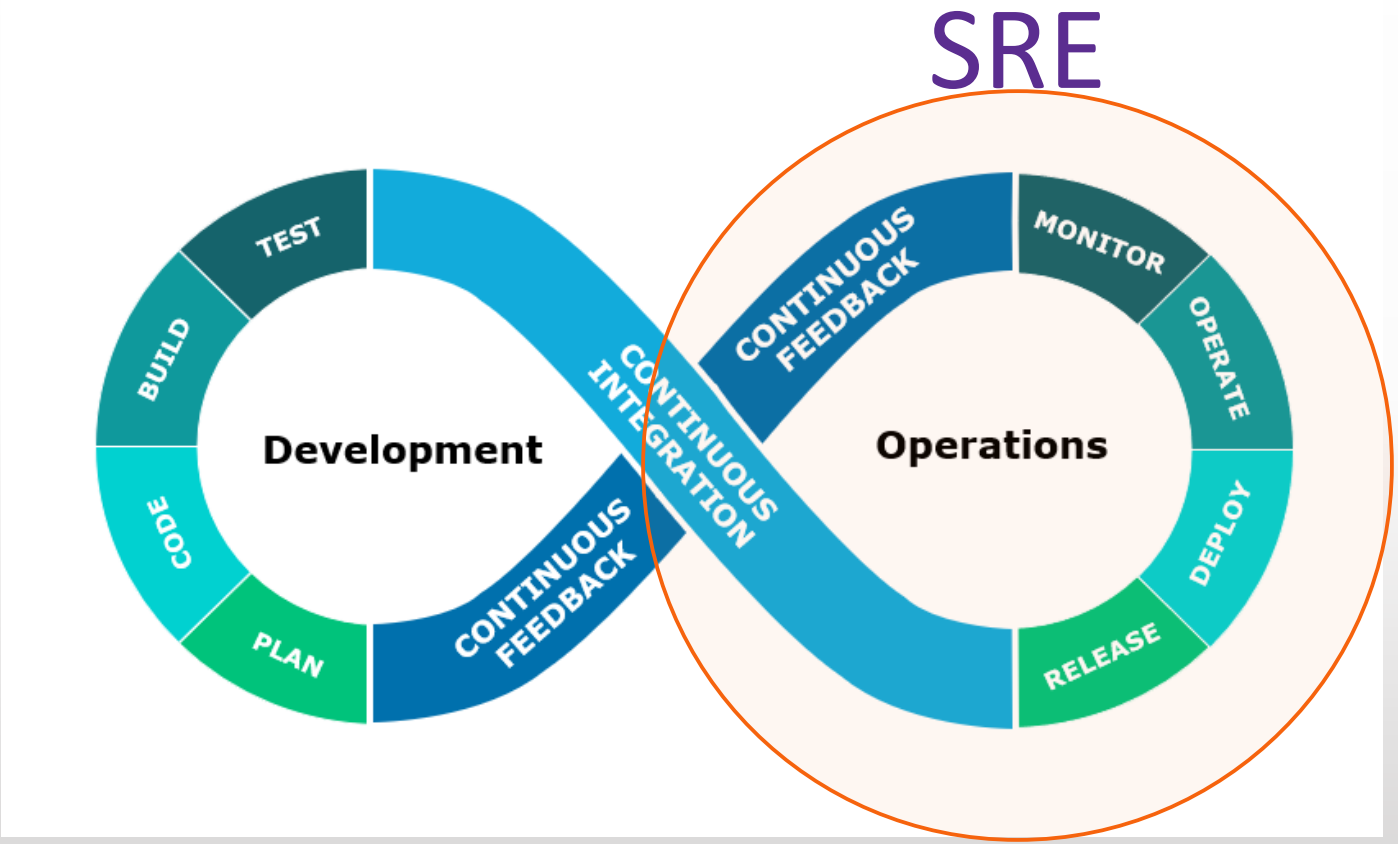
Section - 4

How different teams are structured for an application

https://www.scaledagileframework.com/#Select Configuration



DevOps model



SRE Team

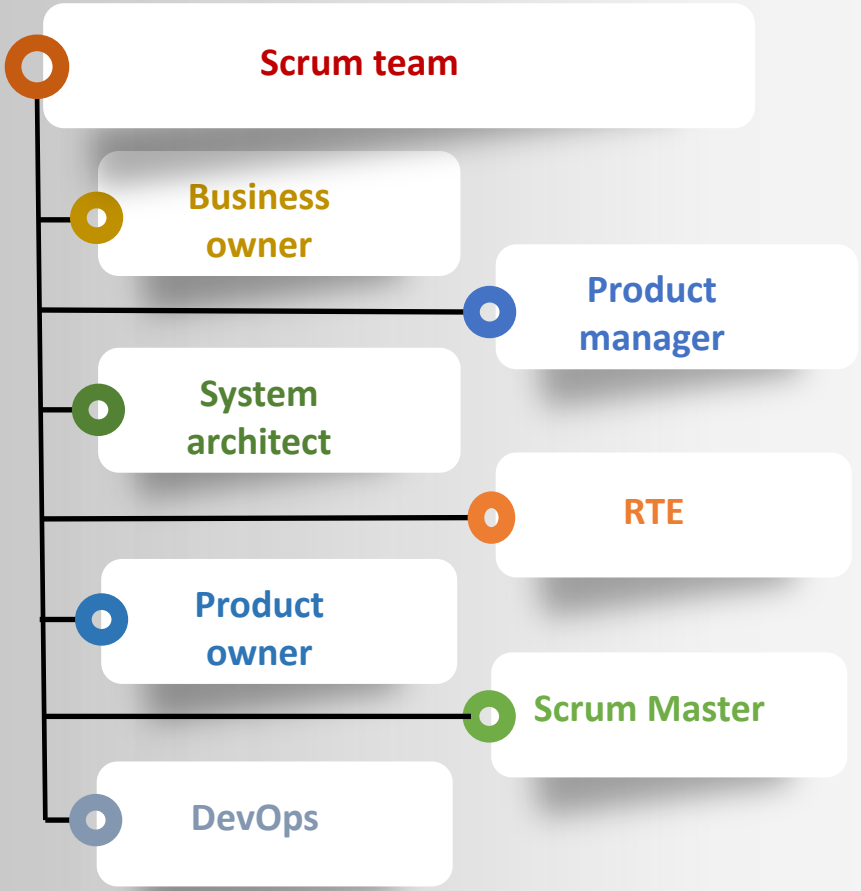
Application Support

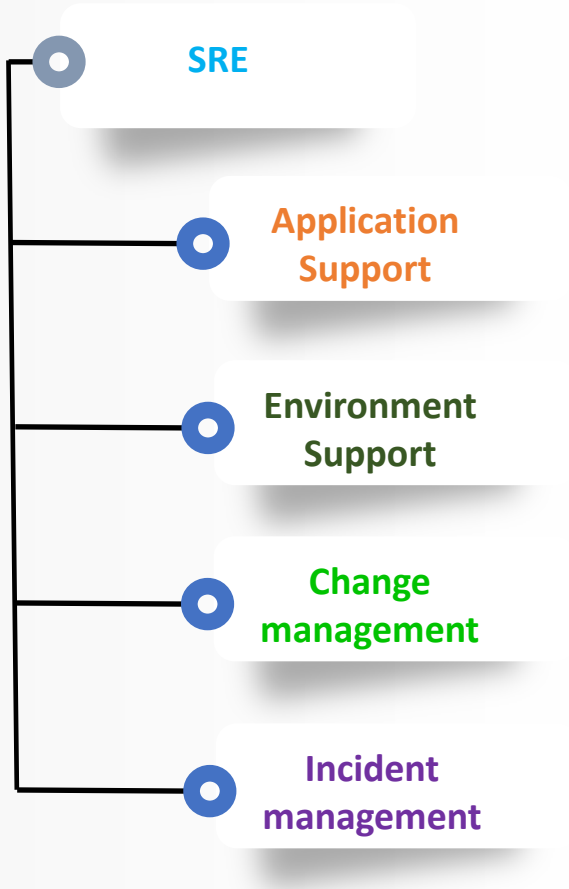
Environment Support

Change Management

Incident Management

Section - 4 Summary

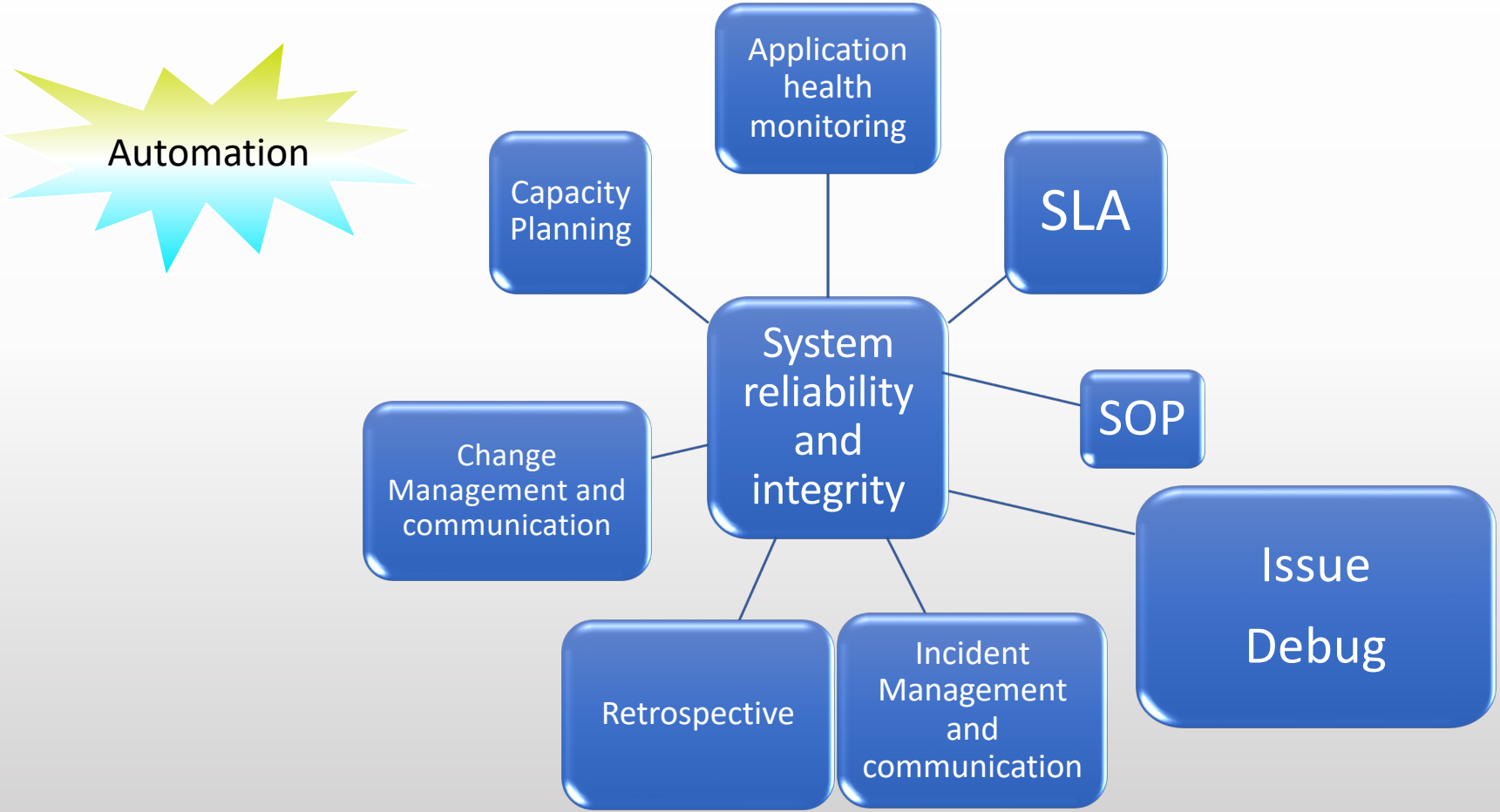




Section - 5

Role of SRE in Production Environment

SRE - Responsibility



System reliability

Online system providing services to the customers needs to monitored, ensuring product developed for customer is working as expected

Detected issues should be reported immediately

How do you know something is not working

- Email alert
- Monitoring dashboard

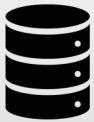
Verticals of monitoring

Security

System health

Application health

Application functionality



Database

Verticals of monitoring – Security

KPIs of security monitoring

- Accounts & credentials
- Infrastructure access
- Phishing attack pattern
- Access from unusual location
- Unauthorized deployment
- Pattern of APIs access incidents
- Intrusion attempts

How to monitor

- SPI(Sensitive Personal Information) data encryption
- Failed login attempts threshold monitoring
- Geolocation mapping from IP
- Error codes
- Vulnerability scanning
- Antivirus update
- Latest patch installation

Tools

Verticals of monitoring – System health

Structural system health monitoring

CPU utilization

Server ping test

Server load average

File system usage

Memory utilization

Network

Firewall

How to monitor

Server performance

- Percentage of CPU utilization
- Load average
- Percentage of memory utilization

Server response

- Ping test – data packet loss
- Percentage of file system used
- Network – data loss between source and destination
- Firewall access log

Tools

Nagios

Azure Monitor

Grafana

Amazon Cloudwatch

AppDynamics

Verticals of monitoring – Application health

Real time application monitoring

Application performance metrics

Application health

How to monitor

Performance

- Response time
- End user experience
- Time consuming transactions
- Cross-application tracing

Health

- Port/URL monitoring
- GC(Garbage Collection) monitoring

Tools

Azure Monitor

Amazon Cloudwatch

Grafana

dynatrace

AppDynamics

ca wily

Verticals of monitoring – Application functionality

Metrics of functional monitoring

Application flow

API response

How to monitor

Application flow

- Business logic
- Error code

API response

- Response time
- Failure rate
- Anomaly detection

Tools

Amazon Cloudwatch

kibana

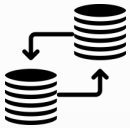
dynatrace

AppDynamics

splunk>

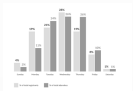
Verticals of monitoring – Database

Critical database monitoring elements



Replication lag

Open connection



IOPS

Expensive query

Throughput

Space usage

Error

How to monitor

Performance

- Data replication lag
- Number of open connections
- IOPS
- Long running DB queries
- DB throughput
- DB storage usage percentage

Health

- DB connection error
- DB VM monitoring

Tools

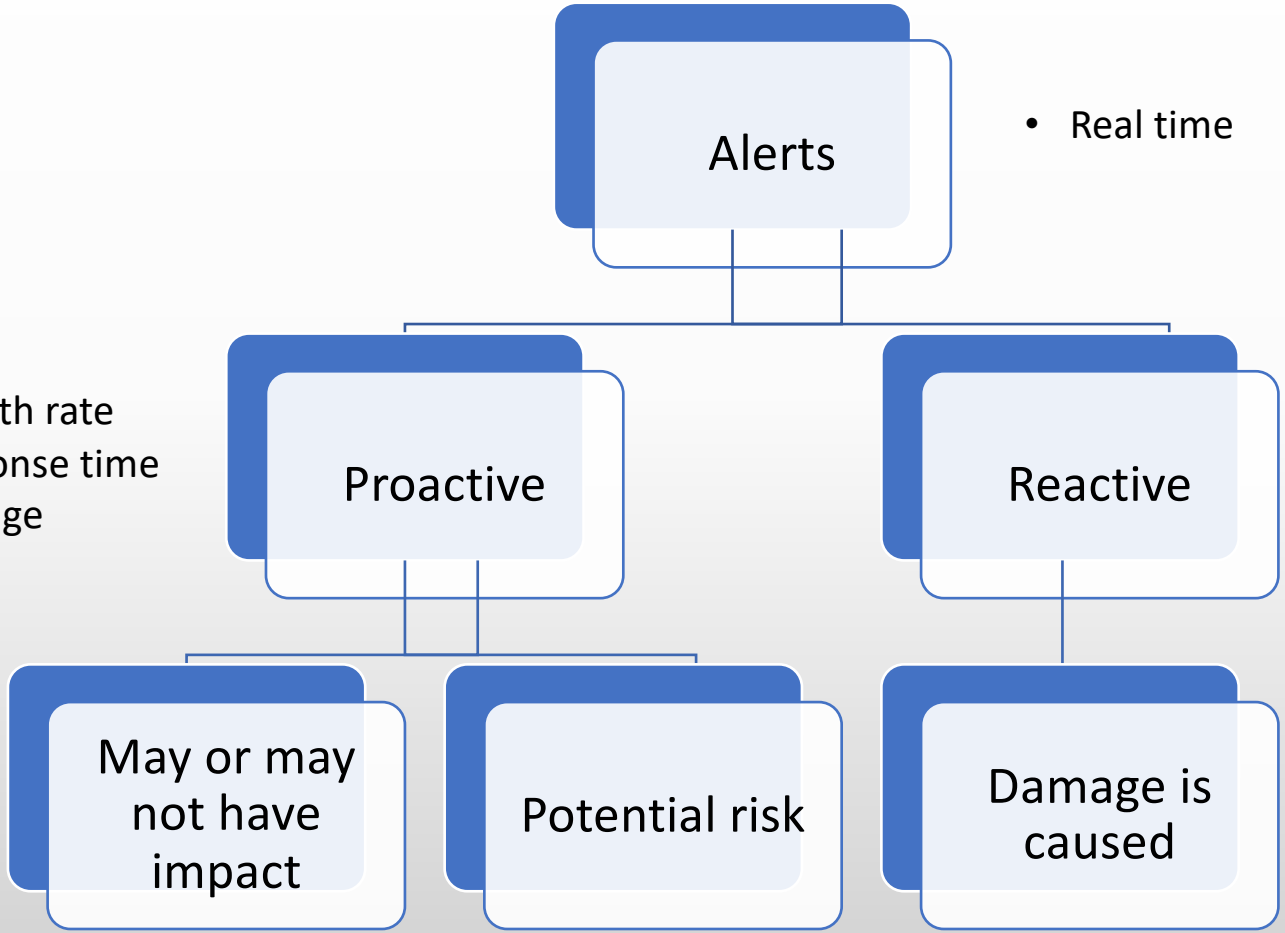
Amazon Cloudwatch

ORACLE ENTERPRISE MANAGER

splunk>

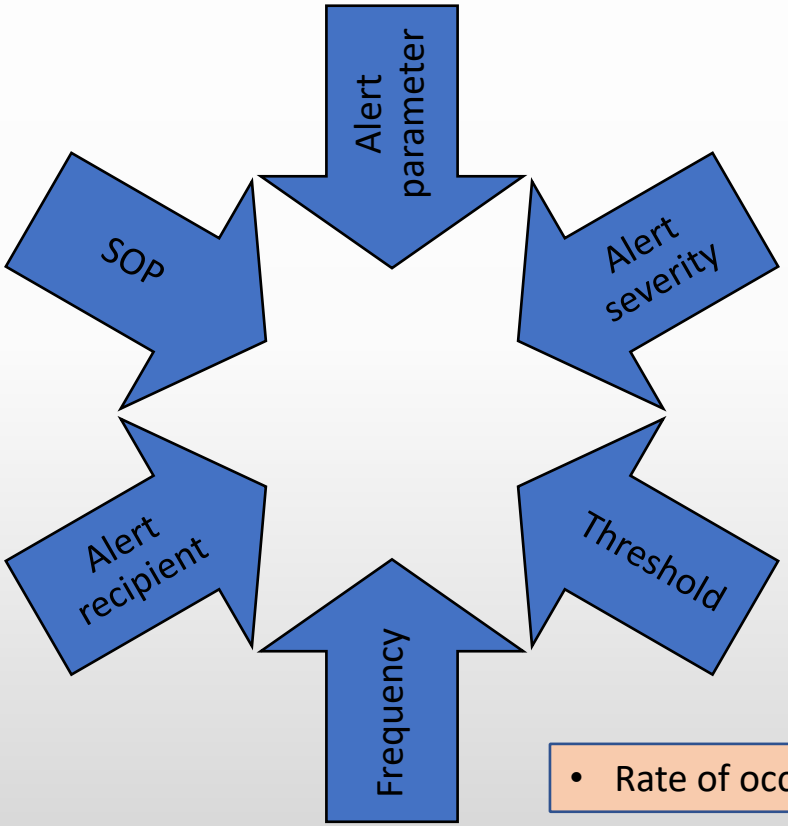
Nagios

Effective alerting



Powerful alert management

- Step by step action plan
- Store at common location
- Accessible to team
- Limited edit access



- Error code
- M
- C
- Fi

- Critical
- High
- Medium
- Low

- Number of events in a time interval

- Whom to notify when alert triggers

- Rate of occurrence

SRE Fundamentals: SLO, SLA, SLI and Error Budget

SLO 

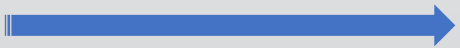
Service Level Objective

Predefined objective SRE Team or its application needs to meet

SLA 

Service Level Agreement

Agreement between Client and SRE Team

SLI 

Service Level Indicator

Actual performance metric

SRE Fundamentals: SLO, SLA, SLI and Error Budget

SLI – SLO = Error Budget

99.95 % 90 % 9.95 %

Severity and SLA

	Sev-1	Sev-2	Sev-3	Sev-4 and below
Business Impact	Critical incident with very high impact	Major incident with significant impact	Minor to low impact	Low or no impact
MTTA	minutes	minutes	minutes	minutes/hours
MTTR	minutes	hours	hours	days
Uptime	minutes	hours	hours	days

SLA – Service Level Agreement
MTTA – Mean Time to Acknowledge
MTTR – Mean Time to Resolve

Incident Management

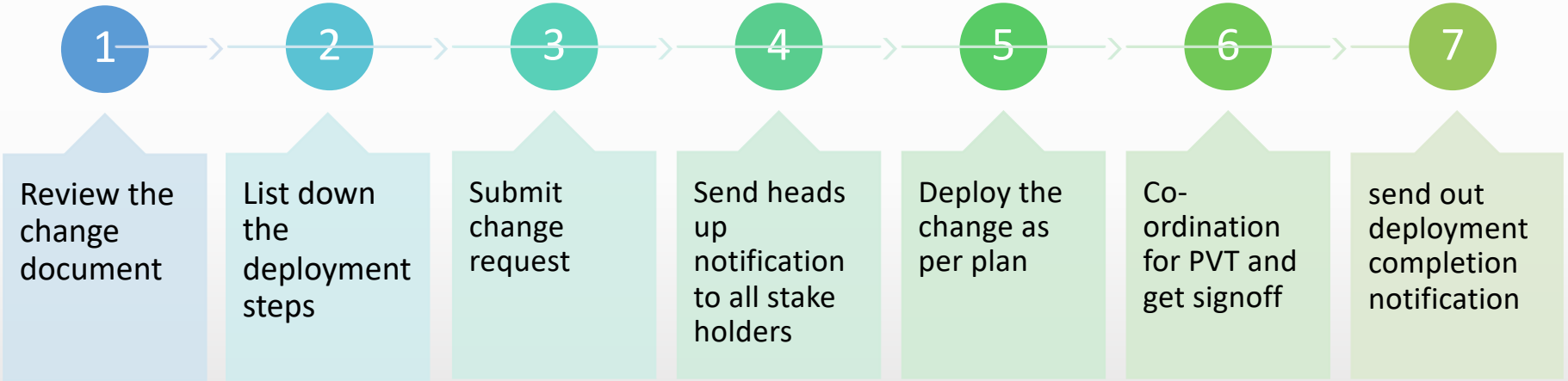
The **purpose** of the **Incident Management** process is to restore normal service operation as quickly as possible and minimize the adverse impact on business operations, ensuring that agreed levels of service quality are maintained.

How to achieve:

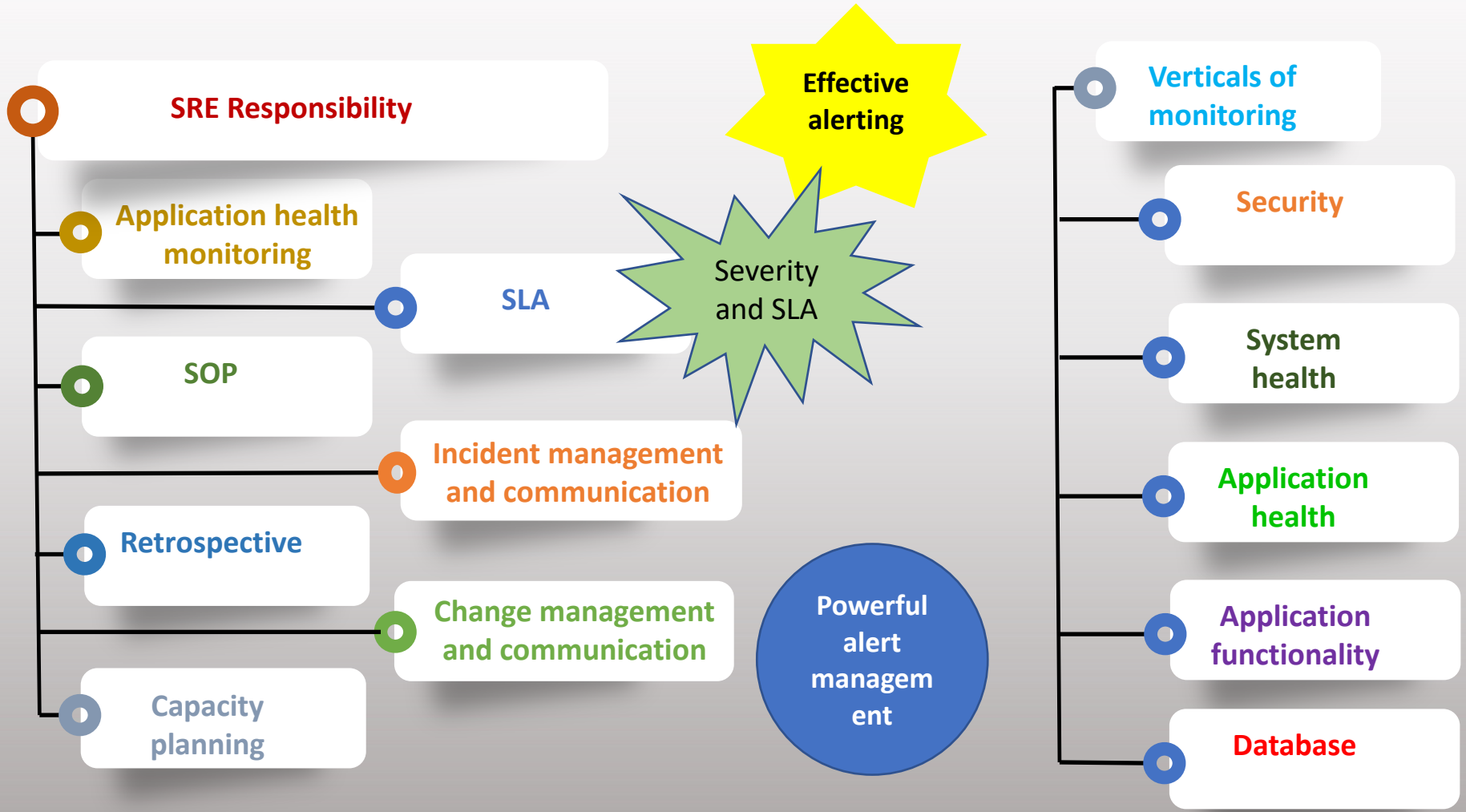
- Alert acknowledge
- Identify severity
- Follow SOP
- Business impact assessment
- Escalate issue to appropriate team/person
- Notification of issue to stake holders

❖ SOP – Standard Operating Procedures

Change Management



Section - 5 Summary



Section - 6

Levels of production support

Levels of production support

Tier-1

- Monitor issue
- Report/Acknowledge issue

Tier-2

- Analyze the issue reported
- Follow SOP
- If issue is not resolved by following SOP then escalate it to next level

Tier-3

- Analyze the issue, replicate it in lower env and debug
- Deliver the fix, if needed engage Development and Architect for analysis

Section - 7

Important elements for SRE

Checkpoint for SRE

Monitoring Tool

Purpose:

- Capture metrics from application
- Compare it with alerting condition
- Trigger email/ticket if condition is meet
- Build dashboard on these tools for graphical representation of data

Log monitoring tool

- Ingest log and then search through to extract information
- Splunk forwarder -> Indexer -> Respond to search queries
- Splunk search processing language
- Licensed

kibana

- Search fields defined before hand
- Elastic search -> Logstash -> Kibana
- Lucene query syntax
- Open source

Security monitoring tool

solarwinds

intruder

vulnerability scanner tool that finds cyber security weaknesses in digital infrastructure, to avoid data breaches

SAAS	Category	SAAS
Access control/firewall/ Internet usage/ Vulnerability scanning	Features	Vulnerability scanning

System health monitoring tool

visualization tool that allows you to see and analyze all of your metrics in one unified dashboard

Usage

- beautiful, simple, annotated graphs.
- multiple sources of metrics or logs at one place
- share your dashboards across the organization.
- reorganize information based on specific team needs.
- easy to use query builder.
- alerting for events.

proprietary software for server, network and log monitoring

Usage

- monitor servers.
- monitor networks.
- monitor applications.
- Windows or Linux monitoring.

Database monitoring tool

ORACLE ENTERPRISE MANAGER

- web-based tools aimed at managing software and hardware produced by Oracle Corporation as well as by some non-Oracle entities.
- set up preferences that help you manage the database
- Home page lists all alerts and for email alert requires additional setup

solarwinds

- The tool's multidimensional root cause analysis looks at:
  SQL statements
  Wait-based analytics
  System health
  Response time analysis
- Popular tool for database monitoring by network admin
- Dynamic UI to show to show most critical information

Email alert/ticket

- After alert condition has meet, based on frequency and threshold email will trigger or ticket will be created
- SRE will receive those emails/tickets and take action as mentioned in SOP

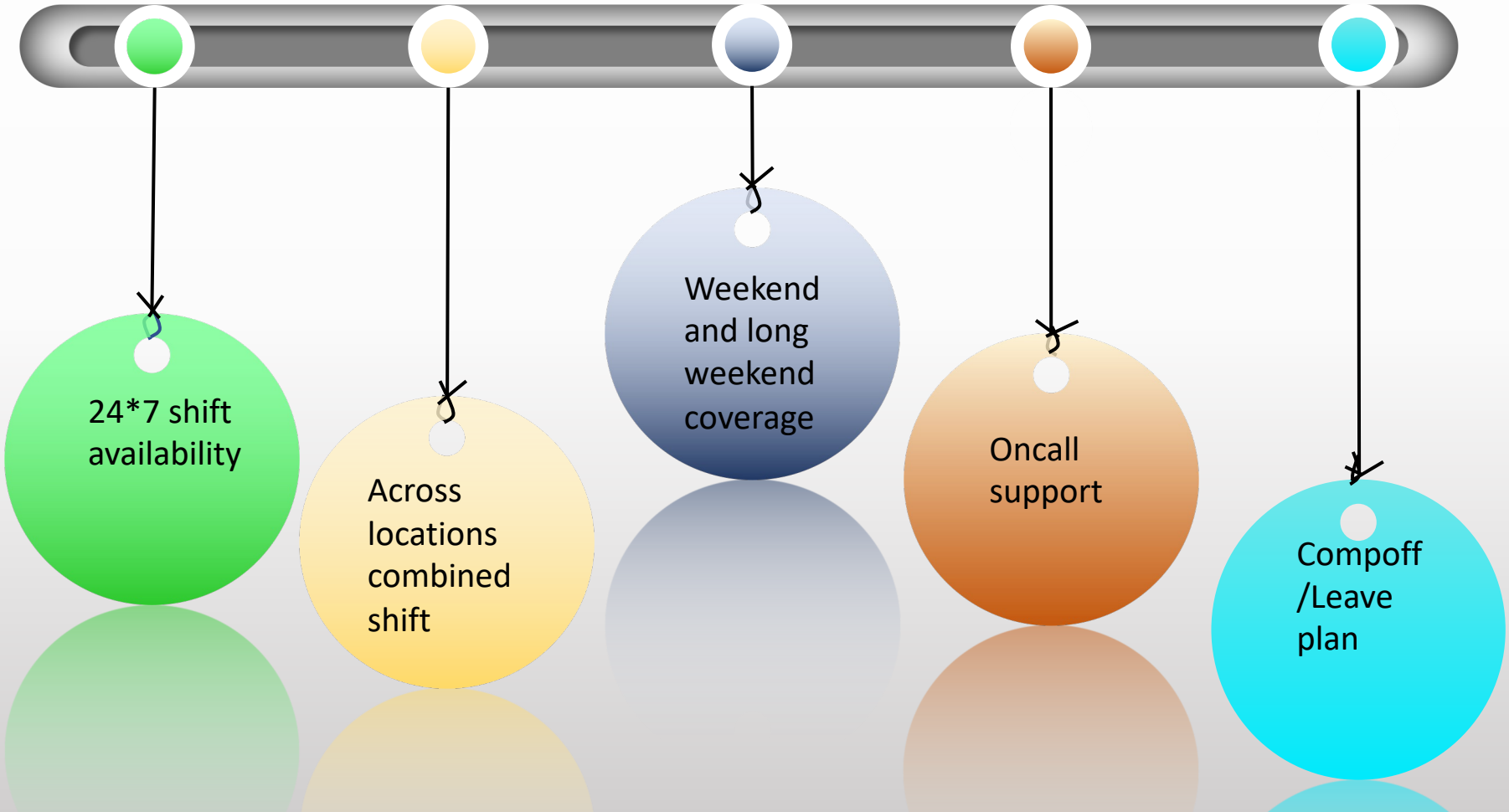
Important details in SOP

What is SOP – document with details

What type of details ?

Issue	Mitigation step	Impact	Severity	Notification
• Issue must be elaborated in detail • It helps in identifying what steps needs to be followed	• Known issue/ unknown issue	• Criticality of the issue is directly proportional to impact to the customer	• It depends upon how many consumers are impacted	• Who all needs to be engaged or notified when issue has occurred

Shift Roster



Section - 8

Issue Debugging

Elastic Search


```
GET IndicesName/_search
{
  "query": {
    “query_string”: {
      “field”: {
        "value": ""
      }
    }
  }
}
```

1. **Automatic indentation**

2. **Elements in search query:**
   a. Index under which the query needs to search
   b. What type if search it is, example: term, match, etc
   c. Under which field the string needs to be checked
   d. Value of the string which needs to be checked

Example 1:

Index = IndexName source = SourceName sourcetype = SourceTypeName “string”

Example 2:

Index = IndexName source = SourceName sourcetype = SourceTypeName string*

1. **Always provide the index name**

2. **Limit the search by providing time range for quick search result**

3. **Elements in search query:**
   a. Index under which the query needs to search
   b. For effective search include source and source type in search query
   c. For exact match use double quotes
   d. Use wild character based on need

Example 3:

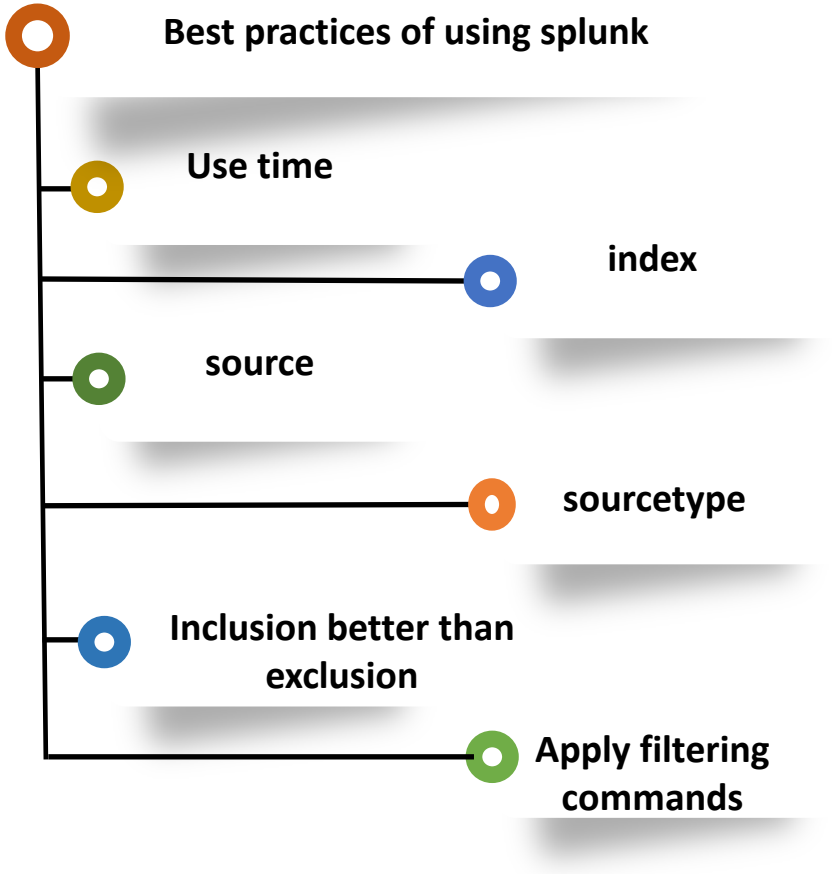
```
Index = IndexName source = SourceName sourcetype = SourceTypeName “string1” OR “string2”
```

Example 4:

```
Index = IndexName source = SourceName sourcetype = SourceTypeName | timechart count by host
```

Example 5:

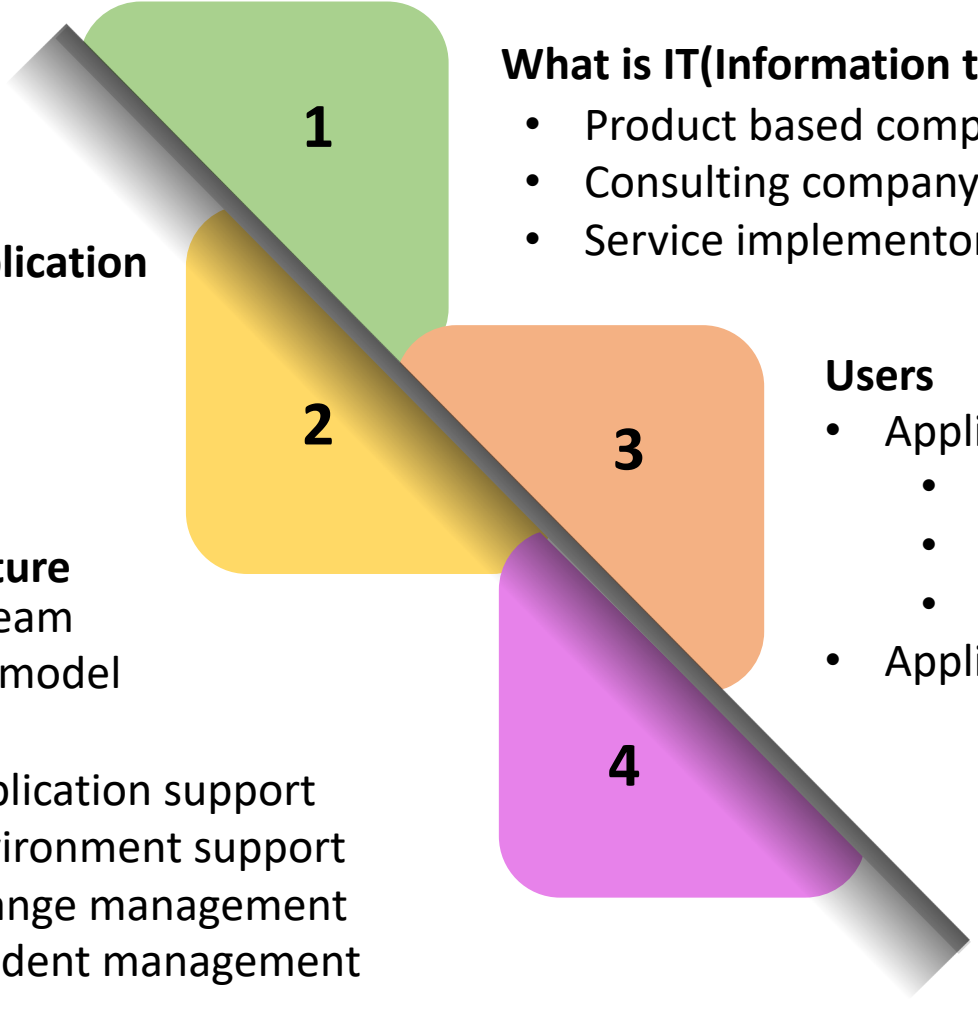
```
Index = IndexName source = SourceName sourcetype = SourceTypeName | sort ip, -url
```



"authentication denied" is better than using NOT "access granted"

1. Data metrics – Prometheus
2. Visualization – Grafana
3. Build dashboard over metrics

Summary



What is IT(Information technology) system

- Product based company
- Consulting company
- Service implementor

Understanding an application

- Frontend layer
- Application layer
- Backend jobs

Users

- Application user
  - Direct consumer
  - Consumer internal to org
  - Federated user
- Application integration

Team structure

- Scrum Team
- DevOps model
- SRE
  - Application support
  - Environment support
  - Change management
  - Incident management

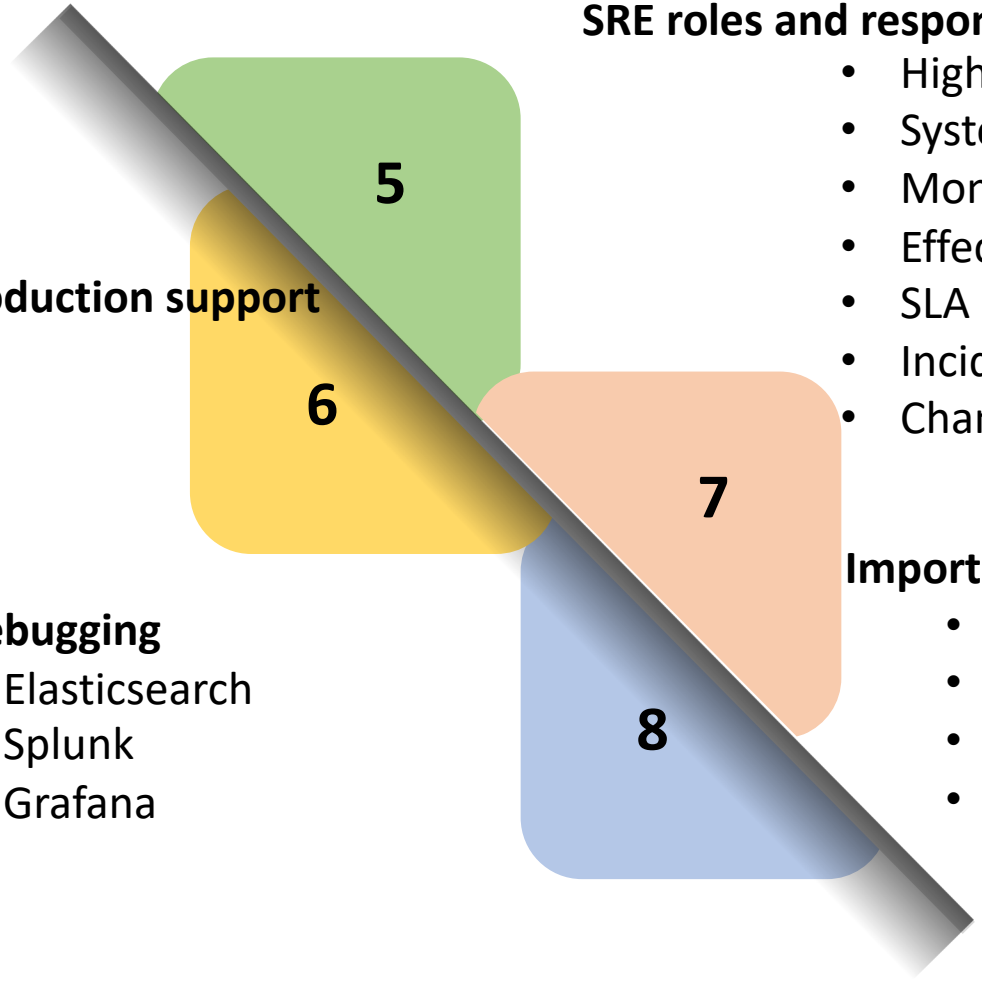
Summary

SRE roles and responsibilities

- Highly available application
- System reliability
- Monitoring
- Effective alerting
- SLA adherence
- Incident management
- Change management

Levels of production support

- Tier-1
- Tier-2
- Tier-3



Issue debugging

- Elasticsearch
- Splunk
- Grafana

Important elements for SRE

- Monitoring tool
- Email alert/ticket
- SOP
- Shift roster

Production Support – SRE in real world

Nidhi Singh

Thank You