1. **State Data Centers (SDC)** are multiple data centers set up in various states of India to provide fundamental IT infrastructure for various eGovernance programs being run as part of National eGovernance Plan of India.

a. Accessing data

The State would follow the best practices in Data Security while sharing the Data from the SDC. To ensure that security is implemented and maintained within the State Data Center, a security policy would be developed and enforced. The security policy must include the following:

- The overall security goals.
- An outline of the overall level of security required.
- The security standards, including auditing and monitoring strategies.
- Definitions of training and processes to maintain security.

State would deploy Defense-in-depth strategy for securing the State Data Center architecture and enhance security level. This would comprise of Perimeter Defenses, Network Defenses, Host Defenses, Application Defenses and Data and Resources Defenses.

- Formulate and implement Trust and Identity Management Policy
  - Authenticate users prior to accessing services from the SDC, which would provide accountability for the transactions/activities performed within the system.
  - Authenticate users prior to accessing services from the SDC, which would provide accountability for the transactions/activities performed within the system.
  - State would use Public Key/Private Key infrastructure for AAA access mechanism to the users for providing access to the sensitive transactions.
  - State would use digital Signature, Digital certificates/biometrics for authentication of users performing critical transactions in the system.
  - In case of less sensitive data, State would use token based or strong password based authentication mechanism for services/transactions where public key certificates are not feasible.

- Security Posture Assessment Report to identify Security and Risks
  - State would mandate to define security zones at the SDC and set security levels for each zone: These separate the data center into areas that are logically separated from one another to contain an attack at minimal impact.
  - Communication between applications can be limited to specific traffic required for application integration, data warehousing, and Web services. Zones created at the Storage Area Network (SAN) can provide logical separation of each application's storage environment across a scalable, consolidated storage network. To achieve this efficiently, firewalls can be integrated and virtualized to provide secure connectivity between application and server environments.
  - State would also deploy control access between zones with firewalls and routers. Firewalls provide perimeter control for state-full inspection of connections to and from the data center while blocking access to nonpublic services and hosts through ingress and egress filtering. Routers provide Layer 3 segmentation between zones, inter-VLAN routing, bandwidth rate limiting, and traffic analysis.
  - State would need to implement Perimeter Firewall (Separating Internet from DMZ) and Internal Firewall (Separating DMZ from internal network) to increase the defense against vulnerabilities.
  - State would need to uses advanced stateful packet and application-layer inspection firewall, virtual private network (VPN), and Web cache solution that enables internet clients to retrieve the static content from the cache by improving network security and performance for both the Perimeter Firewall and Internal Firewall.
  - State would deploy endpoint protection for critical servers and hosts by deploying Host based IPS. This functionality discovers attacks in progress, protects operating systems and applications, and sends alarms to the management console when an exploit is detected.

- State would implement network IPS for critical network segments. Network IPS is used for analyzing traffic streams to identify and thwart attacks such as DoS and hacker activity. The system alerts the management console and/or invokes an automated response within the network infrastructure to "*shun*" or block attacks as they are identified. IDS can also dynamically command firewalls or routers to block packets from identified malicious sources, reducing the effort needed to mitigate the attack.

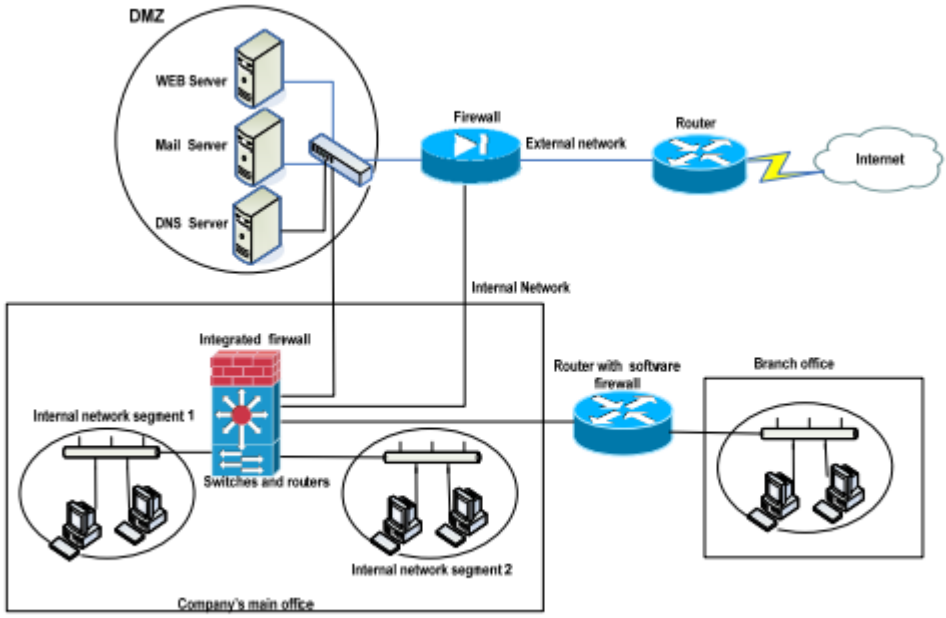


Fig.1. Firewall architecture

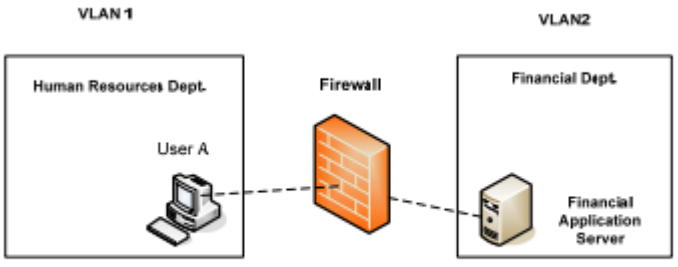


Fig.2 Traffic control between VLANS

- **SAN Security**

  - Secure the SAN from external threats, such as hackers and people with malicious intent.

  - Secure the SAN from internal threats, such as unauthorized staff and compromised devices.

  - Secure the SAN from unintentional threats by authorized users, such as misconfigurations and human error.

  - Secure and isolate each storage environment from other storage environments even if they share the same physical network.

  - SAN should support cloning (creating copy of production disks) onto less expensive disks from which the backup would be performed without affecting the performance of the production disks/LUNs.

  - In a SAN fabric, LUN storage is essential to the configuration of the environment and its performance. A LUN is a unique identifier given to separate devices, or logical units, so they can be accessed by a SCSI, iSCSI or Fiber Channel protocol.

b. Data Privacy policy for accessing data

- Obtaining consent, when appropriate, from individuals for any personal data collection activities that the State declares in its privacy policy. Consent can be obtained by using online forms containing checkboxes or by asking individuals to sign and return a written consent form.
- State would mandate access to the database/production servers and thus access to the data must be in control of system administrator. The root or administrator password must be known to both the nominated representative of user group and system group so that both should agree before making any major changes in the database.

- Each activity related to delete or update operation on the database even if the nominated authorized person does it must be logged for the purpose of audit trial and the logs must be protected via proper security mechanism.
- Console operator would also be given captive accounts for performing routine and repetitive jobs such as taking backup, doing recovery and generating the accounting reports. They must not be allowed to come on the OS prompt.
- State would need to use enterprise backup software to perform backups onto Automated Tape Library and these tapes should be transferred to a safe place away from the Datacenter to avoid loss of data in an event of disaster.
- State would mandate to have hierarchical layered structure defined for different types of users falling between super users (root user, account holder) and console operator with different access rights for the proper safety of the data.

c. Data Confidentiality

- The Operator shall ensure that all its employees, agents and sub-contractors execute individual non-disclosure agreements, which have been duly approved by the State, with respect to services provided from the SDC.
- The stakeholder of the data/applications and the party using the same should sign a Non-Disclosure-Agreement (NDA) with the State.
- State would formulate the policy of Intellectual Property rights with the concerned line departments while hosting/keeping their data into the SDC with overall control being with the State Government.

d. Data Protection mechanism from loss

- State would implement proper RAID mechanism to avoid loss of critical and comparatively less sensitive data.
- State would use Enterprise Storage Area Network based storage system for critical applications running on different hardware server machines. The SAN Storage system should be capable of Selective Storage Presentation (A feature by which storage volumes designated for access by a specific server would be fire-walled from all other devices on the SAN) and should support heterogeneous environments. The Storage system should also support hot add, hot removal and disk layout reconfiguration without the need to restart the system.

e. Data Recovery

- State would formulate a backup policy to periodically backup the data from online machine (hard disk) to offline. Database consistency check utilities must be run to verify that the data backup is consistent and can be used confidentially to recover data at the time of crisis. Periodic checks should be conducted on the backup tapes by way of restoration.
- State would advise the SDC operators to apply patches/upgrades regularly on the IT infrastructure including Servers, Operating systems, databases, application related, network equipment and on the storage system protecting the resources from known issues.
- State would form proper database recovery policy for different kind of failures to avoid even the slightest piece of data being getting lost. To reduce the recovery time of database, the database size should be kept under control by regularly purging the data and archiving it on the offline media, which is not required for online operation.
- State would be ready with a list of contacts of key personal as and when required in case of emergency.
- State would encourage keeping a vital system/software documentation at the backup site.

- State would keep a copy of complete Recovery Plan and steps involved at the off-site (backup site) with authority defined to use this documentation.

f. Security Audit

The State shall get the security audited by third party expert periodically (once in six months) to ensure and guarantee security of the Data Centre. The audit shall bring out any security lapses in the system and establish that the system is working as desired by the State.

2. Google Data Center

Google has data centers across the world running 24X7 in Oregon, Georgia, Oklahoma, South Carolina, Chile, Hong Kong, Singapore, Taiwan, Finland, Belgium, Ireland.

a. Physical Barriers and Perimeter Fencing

- Access to data center is highly controlled. It's in Google's Policy not to allow public tourists or visitors. Even Google's employee access is restricted.
- All vehicle access to the data center is controlled by a restricted barrier checkpoint.
- Google data centers are watched by a comprehensive set of Video Monitoring Camera System. Some of the Cameras are equipped with sophisticated thermal imaging that can identify potential intruders of the perimeter or within the grounds of the facility using heat signatures.
- Google also utilizes video analytics that are designed to automatically detect anomalies in the video and alert security staff to investigate further.
- Security personnel are on duty 24 hours a day 7 days a week both on patrolling cars and on foot with well trained dogs.
- Security fencing around the perimeter has a optical wiring system that detects anything near to the fence and indicates security team. It ensures that all authorized access occurs through checkpoint.

b. Access control with Badging and Biometric Identification

- Once granted access to the facility authorized personnel must check in at an access control desk.
- Access throughout the facility is controlled by badges that use a special lenticular printing mechanism that makes it especially difficult to replicate.
- Access logs are automatically scanned to make sure that anyone who entered an area also left it.
- Some of Google Data Centers utilize biometric devices, such as Iris Cameras to verify identity by scanning employees eyeballs.

c. Data Protection and Hard Drive Life Cycle Management

One of the most important machines that Google builds are the Hard Drives, especially for Google Apps Customers, as this is where their business data is stored.

- Google's customer data is stored in multiple locations to help ensure reliability.
- The files that store the data are not humanly readable. They are given random file names and are not stored in clear text.
- For each hard drive that is received in one of the data centers. Google rigorously tracks its location and status.
- When a hard drive fails or begins to exhibit performance problems. It's brought to a special area where it's brought to a special area where it's reformatted and retested.
- If the hard drive does not pass these tests, it is removed from the rotation.
- The data on the hard drive is then overwritten to help ensure that there is no trace of customer data remaining on the hard drive.

The drives are destroyed in a Multi-Step Process

- One device that is used to destroy old hard drive is known as 'the crusher'. A steel piston is pushed through the center of the hard drive and the platters are deformed making them unreadable.
- Another step in the process is the 'drive shredder'. This completely shreds the drives so that no one is able to read them.
- Remains are sent to recycling centers.
- Google maintains an extra backup of the data that is stored on special tapes. This provides a level of redundancy that helps to protect its customers' data.

d. Fire Detection and Suppression

In an event of a fire or any other destruction, data access for Google's customers is designed to automatically and seamlessly shift to another data center so that they can keep working and their business can continue uninterrupted.

e. Reliability of Operations

- Google's data centers are equipped with emergency backup generators that are capable of powering the data center operation in the event of a power failure.
- Google data centers are connected to internet via high-spped fiber optic cable.
- In each data center there are multiple redundant connections to protect against the possibility of a failure from a single connection.

Fig 3

3. National Security Agency (NSA) – The Country's Biggest Spy Center

- Electricity comes from the center's own substation built by Rocky Mountain Power to satisfy the 65-megawatt power demand.
- Private emails, cell phone calls, Google searches, parking receipts, travel itineraries, bookstore purchases, etc, are all stored in NSA's database.
- Costly antiterrorism protection program including a fence designed to stop a 15000 pound vehicle travelling 50 mph, closed-circuit cameras, a biometric identification system, a vehicle inspection facility, and a visitor-control center.
- The entire site is self-sustaining, with fuel tanks large enough to power the backup generators for 3 days in an emergency, water storage with the capability of pumping 1.7 million gallons of liquid per day, as well as about 60000 tons of cooling equipment to keep servers from overheating.

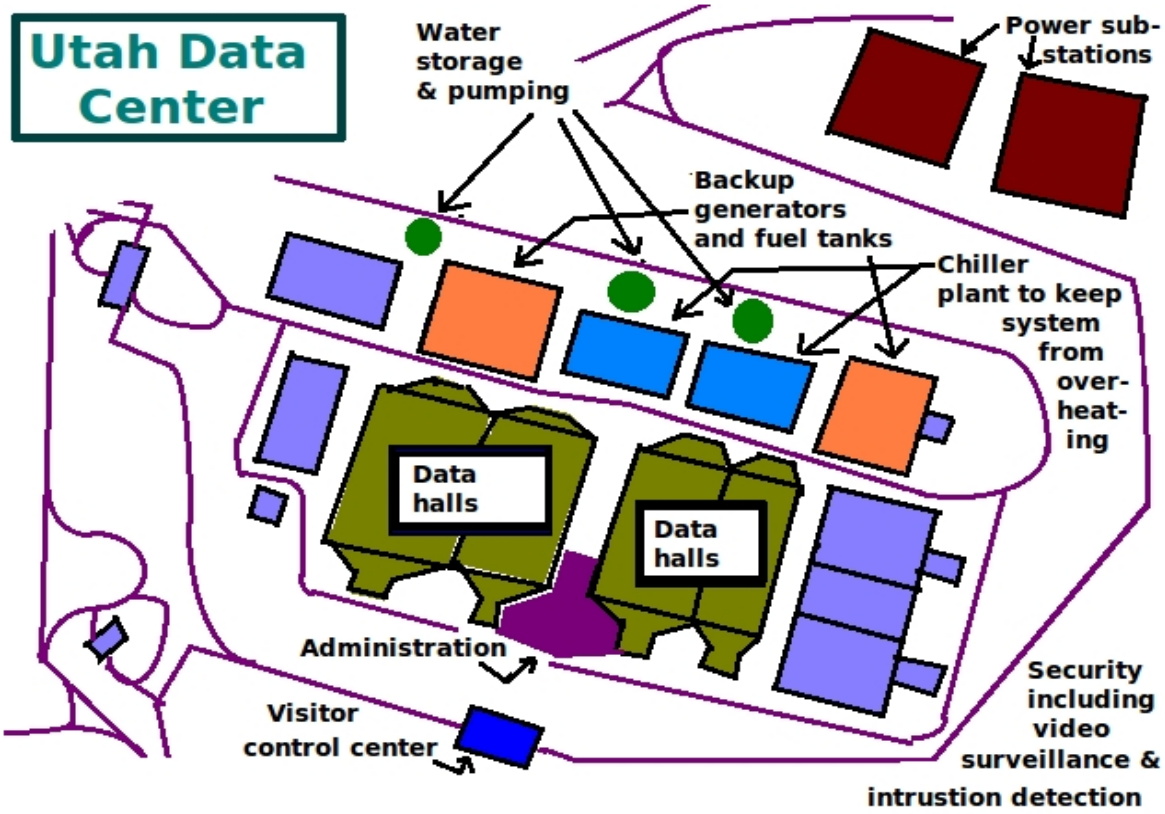


Fig 4

4. **Pionen Data Center**

Located in White Mountains in Stockholm, Sweden

- The data center is housed underneath the mountain, in what was originally a military bunker and nuclear shelter during the Cold War era.
- Data Center can withstand a hit from a Hydrogen Bomb.
- Entrance doors are almost 16 inches thick.
- The facility has 11950 sq ft of space and is located below 30 meters (almost 100 ft) of solid granite.
- The network has full redundancy with both fiber optics and extra copper lines with 3 different physical ways into the mountain.
- It is one of the best-connected places in northern Europe.
- The data center has simulated daylight, greenhouses, waterfalls, and a huge 2600 lt salt water fish tank.
- Cooling is handled by Baltimore Aircoil fans producing a cooling effect of 1.5 megawatt.
- Backup power is handled by 2 Mayback MTU diesel engines producing 1.5 Megawatt of power. The engines were originally designed for submarines.

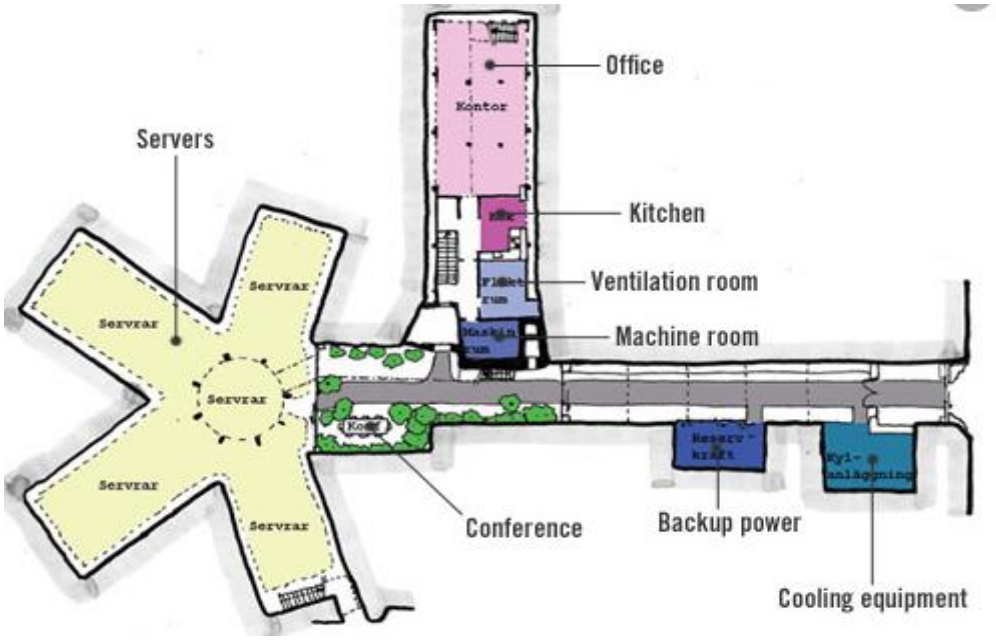


Fig 5

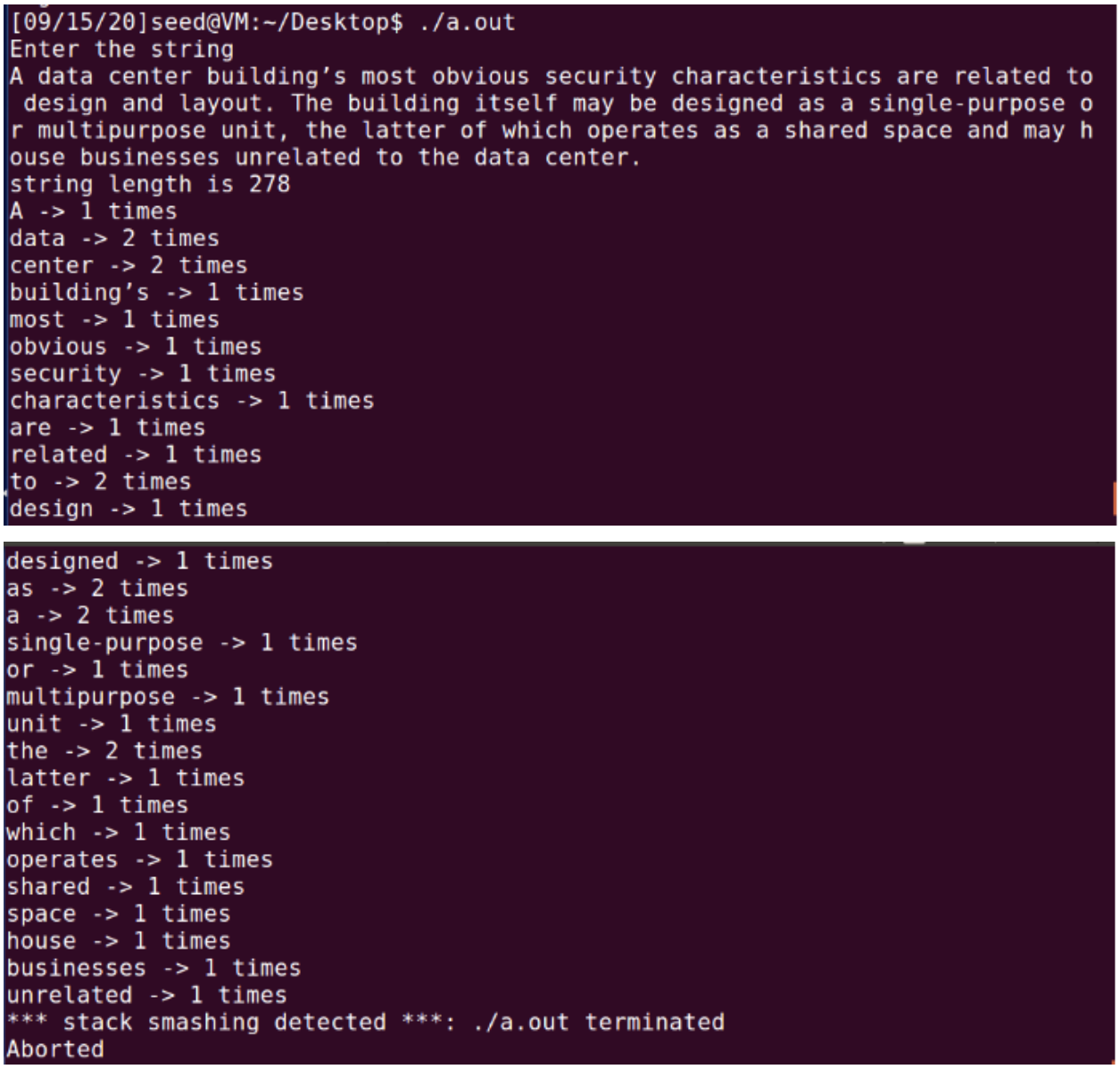
FREQUENCY ANALYSIS OF A PARAGRAPH

CODE:

```
#include <stdio.h>
#include <string.h>
void main()
{
   int count = 0, c = 0, i, j = 0, k, space = 0;
   char str[100], p[50][100], str1[20], ptr1[50][100];
   char *ptr;
   printf("Enter the string\n");
   scanf(" %[^\n]s", str);
   int n=strlen(str);
   printf("string length is %d\n",n);
   for(i=0;i<n;i++)
   {
      if((str[i]==' ')||(str[i]==','&&str[i+1]==' ')||(str[i] == '.'))
      {
         space++;
      }
   }
   for(i=0,j=0,k=0;j<n;j++)
   {
      if((str[j]==' ')||(str[j]==44)||(str[j]==46))
      {
         p[i][k]='\0';
         i++;
         k=0;
      }
      else
          p[i][k++]=str[j];
   }
```

```
    k=0;
    for(i=0;i<=space;i++)
    {
       for(j=0;j<=space;j++)
       {
          if(i==j)
          {
             strcpy(ptr1[k],p[i]);
             k++;
             count++;
             break;
          }
          else
          {
             if(strcmp(ptr1[j],p[i])!=0)
                continue;
             else
                break;
          }
       }
    }
    for(i=0;i<count;i++)
    {
       for(j=0;j<=space;j++)
           if (strcmp(ptr1[i],p[j]) == 0)
             c++;
        printf("%s -> %d times\n", ptr1[i], c);
       c=0;
    }
}
```

OUTPUT:


[09/15/20]seed@VM:~/Desktop$ ./a.out
Enter the string
A data center building's most obvious security characteristics are related to
design and layout. The building itself may be designed as a single-purpose o
r multipurpose unit, the latter of which operates as a shared space and may h
ouse businesses unrelated to the data center.
string length is 278
A -> 1 times
data -> 2 times
center -> 2 times
building's -> 1 times
most -> 1 times
obvious -> 1 times
security -> 1 times
characteristics -> 1 times
are -> 1 times
related -> 1 times
to -> 2 times
design -> 1 times
designed -> 1 times
as -> 2 times
a -> 2 times
single-purpose -> 1 times
or -> 1 times
multipurpose -> 1 times
unit -> 1 times
the -> 2 times
latter -> 1 times
of -> 1 times
which -> 1 times
operates -> 1 times
shared -> 1 times
space -> 1 times
house -> 1 times
businesses -> 1 times
unrelated -> 1 times
*** stack smashing detected ***: ./a.out terminated
Aborted