OPTIMAL COMMUNICATION BACKBONE DESIGN FOR A TYPICAL LRT URBAN TRANSPORTATION CASE

Obrad Aleksic, Systems Specialist
Craig Smith, Sr. Systems Specialist
Djoko Corovic, P. Eng
Hatch Mott MacDonald Ltd.
Toronto, Canada

Introduction

In conjunction with the construction of a LRT line, a new communication infrastructure becomes essential to support safe, reliable and continuous operations. In addition to remote monitoring and control of LRT operations, such an infrastructure facilitates interactions and exchanges between various City services. Typical communication infrastructure covers the entire alignment including operational control center and offices close or remote to the alignment itself. In any case, communication backbone infrastructure is an integrated solution that brings together various aspects of communication services and the entire user's scope of requirements. Therefore, industry best practices include the following:

- Sufficient coverage and capacity to guarantee current and future requirements based on scalable configurations and throughputs.
- Measures for continued operations including redundancy for all major communication aspects
- The backbone infrastructure established along the entire corridor with connectivity to each Traction Power Substation (TPSS), wayside cabinet, station, stop and data/control centre including specific needs along bridges, elevated sections and/or tunnels.

Traditional networking design in these applications employs the use of ring-based network architectures. In these scenarios a ring network is susceptible to a service disruption should there be an event that occurs at more than one point of failure. As a result of advances in technologies, lower media costs, and a distributed device design, we are able to construct a communications backbone that can operate despite multiple points of failure and will allow higher speed applications as compared to conventional designs.

A typical communications backbone infrastructure supports interoperability and communication centric applications with preferably IP based field devices to ensure true distributed networking approach.

Communication Backbone Infrastructure

A communication backbone typically interconnects all major components and its configuration and layout may take different modes and layouts. Traditionally, a ring configuration has been used with a benefit of reducing the amount of expensive optical cabling and terminations. The latest networking developments and significant cost reductions in optical installations have open new opportunities to build safe, high capacity communication backbones.

Segmenting a communication backbone into three separate layers: Core, Distribution and Access provides interconnected hierarchical infrastructure straightforward to construct, expand and operate. An increased optical cabling plant than interconnects each Access node in the field to Core nodes in the control room via a number of Distributed nodes (at strategic locations) to ensure aggregation and optimization services. Figure 1 represents such a typical communication backbone infrastructure with Access nodes spread out in the field and Distribution nodes strategically located within each TPSS.

The Access layer is comprised of access switches placed within stop/station perimeter or in the field, connecting all IP devices that exist around the platform plus providing Power over Ethernet (PoE) wherever required. Typical devices that connect to an access switch would be CCTV cameras, TVMs, PA/VMS, IP phones, SCADA and other field equipment.

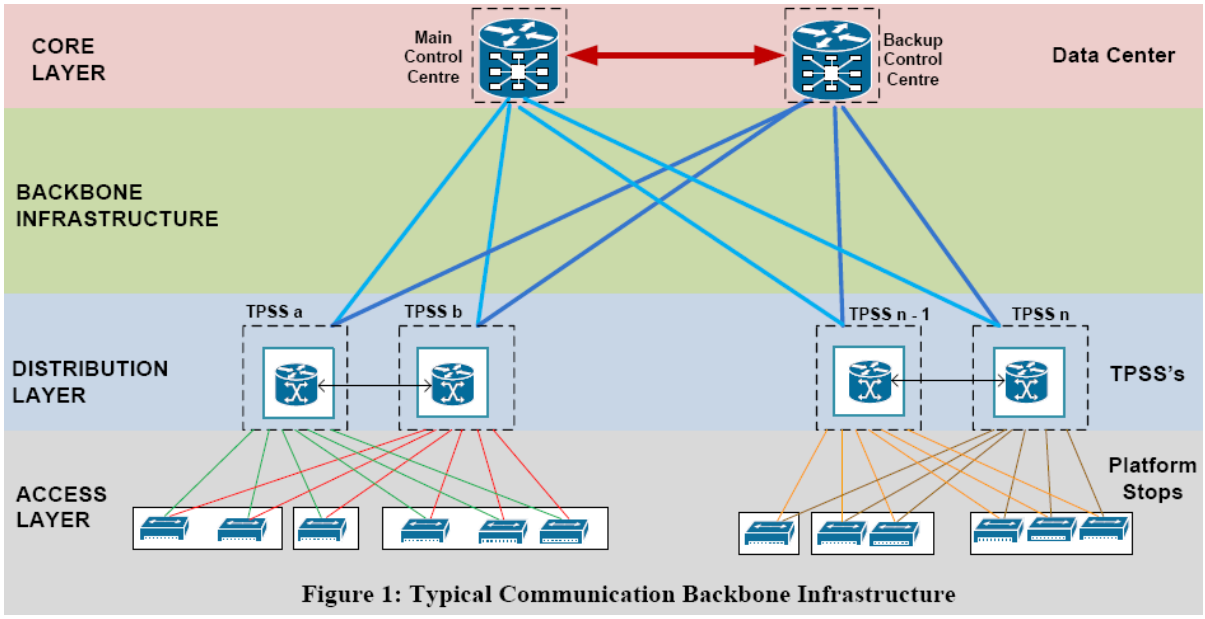

Figure 1: Typical Communication Backbone Infrastructure

Each access switch includes uplinks to the corresponding distribution node to preserve redundancy, in the case of an access link or distribution switch failure.

The required bandwidth for Access node uplinks should be minimum 1 Gbps with Port Channel protocol utilization becoming mandatory to eliminate Spanning Tree issues and get logical paired uplink speed doubled to 2 Gbps.

The Distribution layer consists of distribution switches placed inside a dedicated communication room at each TPSS location. Its role is to aggregate data from connected access switches. Such data can be CCTV video files, PA/VMS data, TVM file and other stop (field)-typical data. After aggregation, distribution layer switch separates various types of data into dedicated data-streams, where each type of data conforms into a separate VLAN with such classified traffic routed to the Core layer.

The Distribution layer should be designed to be fully redundant due to its role and importance. Uplinks to the next-layer network (Core) are also recommended to be redundant – from each distribution switch there are two uplinks, one to each Core. The required bandwidth for distribution switch uplinks should be minimum 10 Gbps. Port channel protocol utilization is mandatory in order to get logical paired uplinks with speed doubled to 20 Gbps.

The Core layer is at the top of the hierarchy. It is responsible for transporting large amounts of traffic efficiently while providing backbone connectivity. The distribution layer devices aggregate to the core layer, and the traffic transported across the Core is common to all users and/or devices connected.

Data Centers are typically built as redundant. The Core switches will service each data centre. They collect and terminate all links from distribution switches, managing data flow to/from application servers and control room operators, including external access.

Due to the large numbers of uplinks between the distribution layer and other connections that terminate on the core switch itself (e.g. firewalls, DC local switches), the core switch technology needs to be carefully selected to ensure a high port density.

A dominant factor for a designer's choice of a core switch is the required minimum throughput speed between the core-to-core links, especially at distances greater than 10 km. Dedicated fibre pairs are recommended for interconnection of the two core switches together through high-speed transceivers. The protocols that maintain redundancy should provide operationally seamless transition, in the event of a core failure. VPC (virtual Port Channel) or VSS (Virtual Switch System) are protocols that enable core communications redundancy.

Such operational requirements are currently confirmed to be available for the Cisco Nexus series.

At Data Center, all of the necessary components for operation and redundancy are provided. Components inside Data Center includes all servers, storage components, Core switch,

firewalls and other networking equipment, as chosen by the designer.

Two data centers should be built to allow for full redundancy including dedicated core-to-core links. Each data center is mirrored for both application servers and archival to act as a network hub with the Core switch and firewalls at the top. Data backup and replication procedures are then established to ensure data retention and replication for general analysis and incident reviews.

Access Switches Concept

Access switches are field devices typically not redundant. In the event of a failure, some localized downtime is experienced until the affected switch is repaired or replaced. Specific devices affected by such a failure may include an IP camera, TVM, or any other IP devices connected to that particular switch. Typically field devices are spread between more than a single access switch and as such any single switch failure does not affect the entire location. Therefore, some level of functionality is preserved at all significant field locations in the case of a single switch failure. Each access switch in the field connects directly to the corresponding redundant Distribution node pair.

Implementing additional redundancy at the Access level is not recommended due to high accommodation costs, limitation of redundant protocols for access switch interconnection, and increased cabling requirements.

Distribution Switches Redundancy Concept

Distribution switches perform data aggregation linking field to the Data Centers. By doing so, Distribution switches become critical networking components requiring solid platform and redundancy. Those are strategically located along the alignment to include coverage and optimal hardware/cabling requirements (***Ref. Figure 2***).

Switch interconnection and redundancy is done through proprietary network protocols where two separate distribution switches act as a single logical unit. The direct link between such two adjacent distribution switches would be via transceivers that support 10 Gbps throughput for redundant operations. Similar to the access-to-distribution links, the Port Channel protocol would be required for those links to double their bandwidth capacity to 20 Gbps (2x10Gbps) and eliminate Spanning Tree issues.

Distribution switch redundancy can be within a single location (usually TPSS) but also pairing of distribution switches can be across two adjacent TPSS locations which would reduce a number of required switches, reduce costs but still improving reliability. At least one distribution link remains

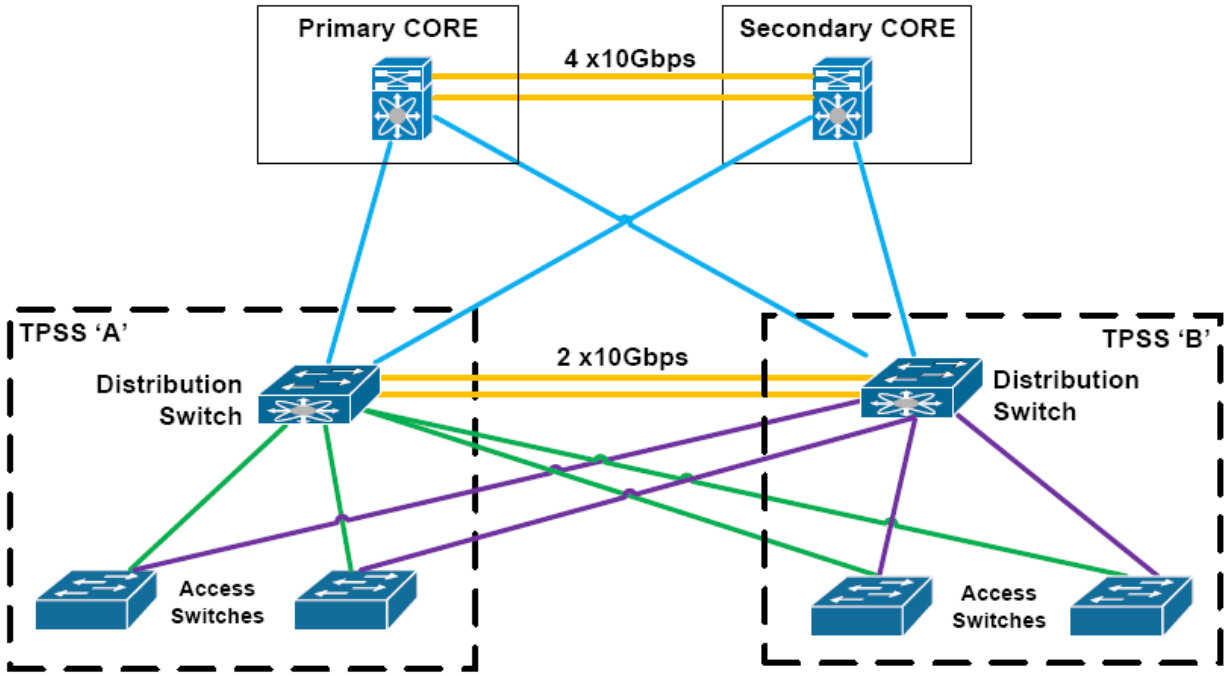

Figure 2: Paired Distribution Switches Concept

active even if one of the paired TPSS nodes fails. As a result, the availability on the backbone links at the Distribution Layer is consistent with systems such as traction power, so that operations remain unaffected in the event of a single catastrophic TPSS failure.

A connection to the Core should be than established with a throughput of not less than 10 Gbps, with transceivers capable of supporting such speed at distances greater than 10 km. For redundancy purposes, two of these transceivers and associated fibre links would be required per distribution switch, one going to each Core. The Port Channel protocol for Distribution-to-Core uplinks would be necessary for the same reasons as stated above.

Network Resilience

Failure cases and solutions shown below illustrate the resilience and robustness of the proposed networking design:

Catastrophic TPSS failure

In the event of a catastrophic TPSS failure, the distribution switch may become inaccessible. In such a case, the second distribution switch, at the adjacent TPSS, becomes aware of the failure and continues to operate. Access switches in the field continue to operate and are not affected due to redundant uplinks and protocols enforced at the Distribution layer. The only effect that the Access layer encounters is a reduction of bandwidth for the affected services. The core switch will lose the uplink from the failed TPSS, continuing operations via the second uplink from the adjacent TPSS. This would not require any manual intervention.

Uplink failure

In the case of the failure of one uplink from the Distribution layer, the Core will notice that an uplink has failed, but will continue receiving data from another uplink. Access-to-Distribution traffic remains unaffected, as that traffic would be automatically forwarded to the switch that has the active uplink. No manual intervention is required to maintain operations, as the entire distribution switch assembly across two TPSS nodes is considered to be one logical unit.

Core Switches

Core switches are typically placed in two separate data centre (DC) facilities. Each component of the Core switch is redundant and paired with its counterpart in the back up DC. As mentioned above, interconnection between Core switches should be at minimum 40 Gbps, with network protocols that provide seamless and immediate transitions in a case of the main DC failure.

Network protocols examples that provide such operations and redundancy include: VSS (Virtual Switching System), vPC (virtual Port Channel) from Cisco, MLAG (Multi-Chassis Link Aggregation) from Arista, Virtual Chassis from Juniper.

Not all vendors support recommended speeds and bandwidth and one should be careful when choosing networking equipment.

Fiber Optic Backbone Infrastructure Design

Typically two segregated conduits for backbone cable infrastructure run along the entire alignment connecting all data centres, TPSS's, stations and stops while ensuring separate paths and redundancy.

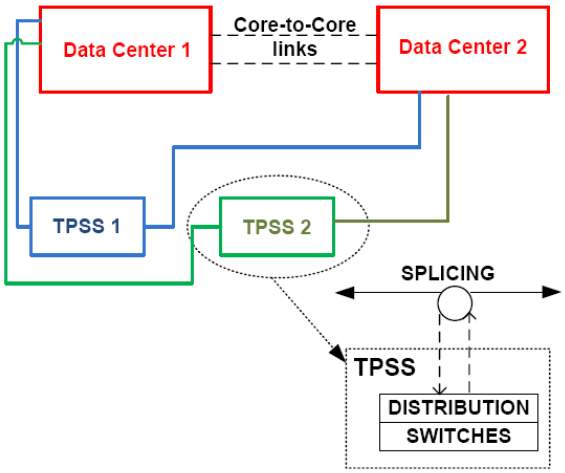

Figure 3: Distribution-to-Core Links

In order to have sustainable redundancy on physical fibre level, designer's goal is to have a dedicated physical links (strings) between each TPSS and each DC. Another dedicated links are than provided for direct DC-to-DC links (***Ref Figure 3***). The pairing of distribution switches across two adjacent TPSS locations reduces the number of required switches, which reduces costs, while improving reliability as at least one distribution link remains active even if one of the paired TPSS nodes fail. As a result, the availability for the backbone links at the Distribution layer is consistent with that of traction power, so that operations remain unaffected in the event of a single catastrophic

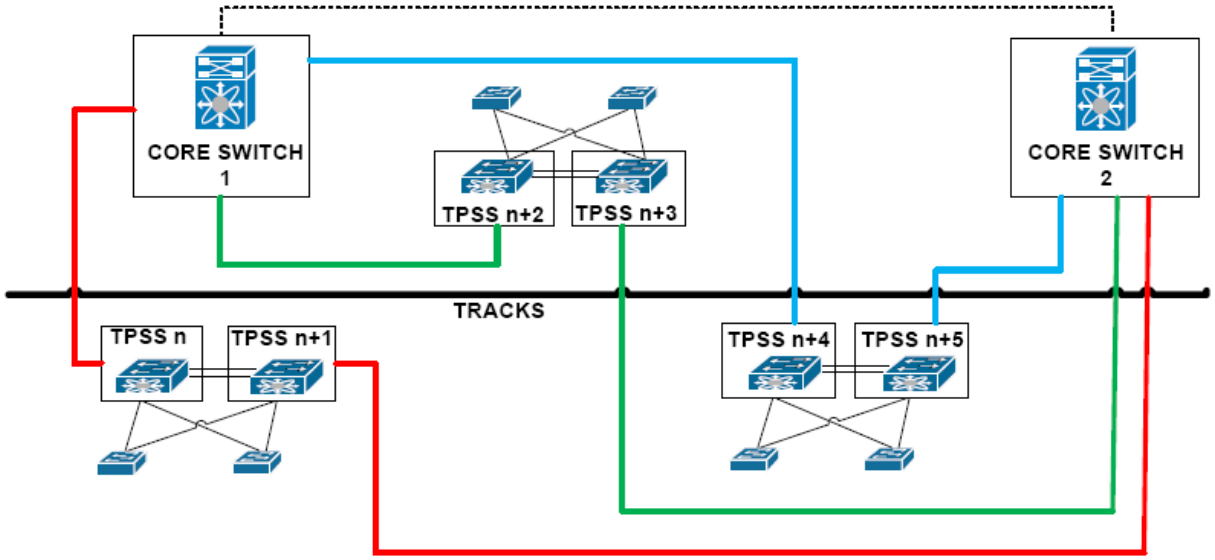

Figure 4: Alignment-driven pairing of distribution switches

TPSS failure. In such cases additional links are required for direct connections between paired distribution switches. (See **Ref Figure 4** for details.)

If the design follows that direction, distribution segments can be identified along the alignment. Each distribution segment comprises of paired distribution switches (TPSSs) and their corresponding and connected access switches (Stops/Stations).

The whole LRT line can then be considered as a fibre-connected series of distribution segments (plus their associated Access nodes) with their ultimate connection to the Core (Data Center).

The entire fibre infrastructure is dictated by alignment configuration and available space for cable runs. Typically, fibre cables are segregated in the field based on ownership, services or/and utilization, Figure 5 represent one way of segregation by which

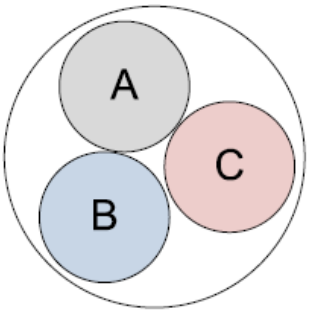

Figure 5: Communication Cable Conduit Utilization

the entire conduit is subdivided for use by different parties.

The example of co-existence within the same cable run shown on ***Figure 5*** includes:

A. Access layer sub-conduit
   - Fiber cable for connection of all Access switches and corresponding Distribution switch pairs (nodes) within a segment

B. Distribution layer sub-conduit
   - Fiber cables for links between TPSS Distribution nodes and data centre cores (and core-to-core)

C. 3rd party sub-conduit - may be empty and available to use by 3rd parties (e.g., City Services).

Other Components Affecting Communication Backbone Design

Data Center

Data Centers are main repository and management points for data and communication links. All data processing and communication links termination takes place inside DC. Main and backup data centres must ensure continuous operations under all conditions. Data center can be co-located within Control Center, for easier access for LRT operators.

Data centres should comply with ANSI/BICSI 002-2011 standard with Class rating sufficient to meet the operational availability requirements. The

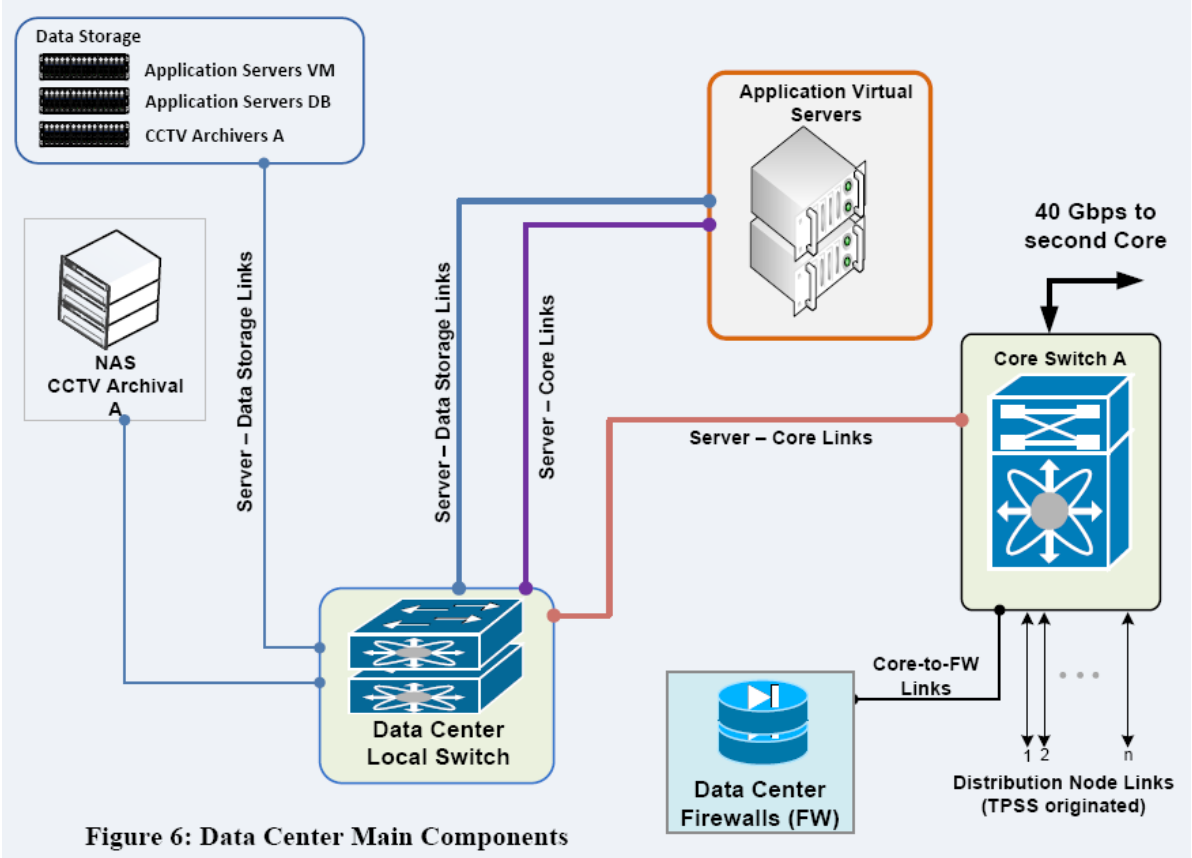

Figure 6: Data Center Main Components

data centre architecture should be modular and easily expandable with no or minimal operational downtime. Virtualization has been recognized as the best approach to building a data centre.

Typical components and their purpose within the DC should be as follows (***Ref. Figure 6***):

1. Core switch – As mentioned above, the core is used to collect links from distribution and forward such data to designated servers.

2. Storage – Storage is used to keep all data for installed application servers, as well as virtualized machines
   a. CCTV storage – Due to the nature of CCTV files (size and growth) the data should be separated from other critical services. Special types of storage with easily expandable drawers are recommended.

3. The local DC switch is used to interconnect all components within DC, with high-speed links (at least 10Gbps). Due to various natures of components implemented inside the DC, and keeping in mind the current trend of reducing a number of devices, it is recommended to implement a redundant device with universal ports that can accept any type of fibre or Ethernet connection.

4. Firewalls are a standard component for protection from outside mistrusted environment. A careful design and setup must be made in order to not disrupt operational and production traffic and protect inside network from malicious access.

5. Servers within DC should be virtualized, except those which are not (yet) supported for virtualization, and/or are requested to stay physical.

Applications

Typical systems applications generally expected for LRT services are as follows:

NMS (Network Management System)
A dedicated simple network management protocol (SNMP) server for network monitoring, maintains an internal database for all system configuration data and monitoring computing equipment. Access to this server and data security are critical.

CCTV (Close Circuit TV)
A number of dedicated servers based on the number of on-line cameras and archival requirements.

TVM (Ticket Vending Machines)
Fare equipment utilizes some sort of database for storing and processing; hence this application shall be backed up and replicated through the vendors' storage mechanism.

VoIP (Telephone)
Includes a pair of redundant servers (or hardware) for central telephone switching and distribution. Voice recording is included with accurate time synchronization for the entire infrastructure. Voice recording and storage (mp3 or wav files) are easily transferred to any repository type.

PA/VMS (Public Address/Visual Message Signs)
A pair of redundant servers that synchronize each other. Synchronization with signaling application is also required to provide train arrival time and track information for announcement and display at stations and stops.

BMS (Building Management System)
A pair of redundant servers dedicated to monitoring various field devices, including status of alarms, measurements and status controls. An automatic switchover mechanism and embedded self-synchronization are implemented. Data historian, dedicated to BMS or shared with other applications (e.g., SCADA) provides storage and access to all historical data for all users.

TPSS SCADA (Traction Power Substation SCADA)
A pair of redundant servers dedicated to traction power substations infrastructure monitoring and controls with an automatic switchover mechanism and embedded self-synchronization. Dedicated historian provides storage and access to all historical data.

Radio
Radio communication for emergency, operations and general services may be shared and provided by other City emergency services. Depending on the strategy undertaken by City, there may be a need to have a dedicated radio infrastructure for LRT operations. In such a case, dedicated servers shall be available for managing and recording radio traffic.

Virtualization in Data Center

The term virtualization describes the ability to run multiple operating systems on a single physical system and share the underlying hardware resources. Virtual machines are a representation of a real machine using software that provides an operating environment which can run or host a guest operating system.

Each VM (virtual machine) uses shared resources (CPU, memory, network card) from underlying hardware machines. The usage of virtual resources and monitoring of operations and control is done by the "Hypervisor" which is brain of the virtualization system. The hypervisor is installed on physical servers which provide these resources. There are few of those hypervisor programs for virtualization platforms on today's market that include; VMware ESXi, Microsoft Windows Server with HyperV and Citrix Zen Server. See **Figure 7** for more details.

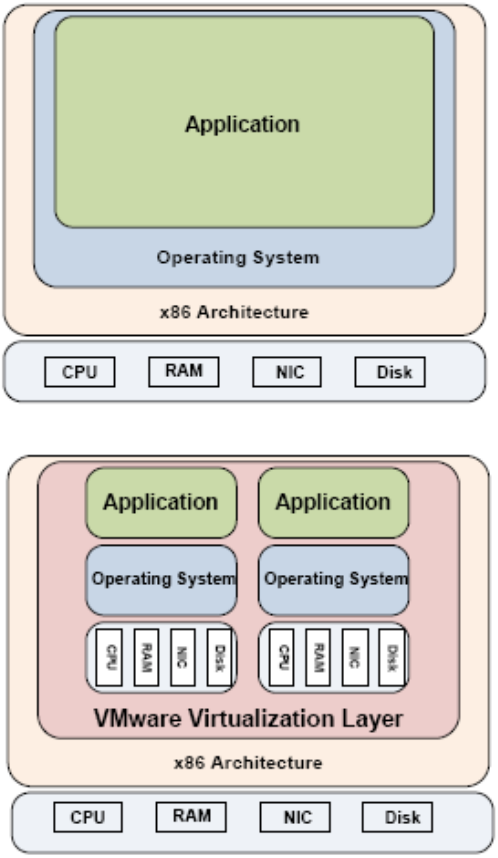

Figure 7: Virtualization concept

There are multiple benefits from using virtual systems:

1. Business continuity and system availability:
   - Minimizing downtime by reducing the cost and complexity to ensure high availability and simpler disaster recovery
2. Minimized infrastructure cost:
   - The number of servers and related IT hardware is reduced.
3. Portability:
   - Combined virtual machines are easy to move from one environment to another for maintenance efficiency and for resource utilization.

Usage of VMs is closely connected with storage devices since the VMs themselves reside in storage array disks.

Storage device and redundancy in Data Center

Storage is used in Data Center for storing VMs and other data types including databases, video files, etc. There are several options for the storage technology implementation:

- Storage Area Network (SAN), which is referred to as block-level storage. This type of storage is considered and recommended for storing relational database files that requires high-speed block-level access.

- Network Attached Storage (NAS) is referred as file-level storage. This type of storage is recommended for storing typical data files, and is not designed for database reposition.

- Direct Attached Storage (DAS) represents pure extension of the server's hard disk.

A combination of file-level NAS and block-level SAN devices into one unified storage device is recommended for the entire data storage needs. This type of storage combines the best features of both types (SAN and NAS) into a single device thus making it best for keeping every type of file and for every possible future scenario.

Storage, as any other device, is not protected in case of incidents in data centre, such as fire, flood or earthquakes. A secondary data centre is proposed including data replication and backups for full redundancy.

Various failover policies needs to be defined based on systems downtime and recovery requirements.

CCTV system in LRT

CCTV is a network demanding critical service and as such mentioned here as an application that significantly drives networking configuration requirements.

CCTV systems comprise of multiple (ideally) IP-based surveillance cameras (in a case of analog cameras utilisation of encoders is assumed). As with any other devices on platforms, CCTV cameras are connected to local power-over-Ethernet (PoE) switch and their video signal is sent to the dedicated CCTV servers that controls and monitors the whole system. These servers are usually a point where controllers and viewers are connected in order to view recordings and change parameters as required.

There are two types of data for CCTV video transmissions; video files and indexing file. The video file is raw video record. The indexing file includes metadata that shows camera, time/date and other file properties. Those files are transferred together and the CCTV controlling software manages those two types of data and places an index file into the database (usually SQL) and raw video in NAS.

As with any other application in the data center, the CCTV system requires redundancy and failover. Current regulatory requirements may determine a need for the full archival redundancy. A method to accomplish full availability of the recordings is to have cameras and network devices that will send their data to two data center locations simultaneously (multicast traffic) which needs to be supported by the networking concept.

The simultaneous transfer of video to both data centers (dual stream) is the best practice for using multicast transport on the distribution level (See ***Figure 8***). The distribution network capacity and multicast enabled switching and routing must be able to support the potential needs for full archival redundancy. These requirements need to be developed based on the final number of cameras, number of viewers, and their locations.

The following ***Figure 8*** presents an example of the overall CCTV infrastructure and provides insight into multicast and dual stream concept of CCTV video transfer.

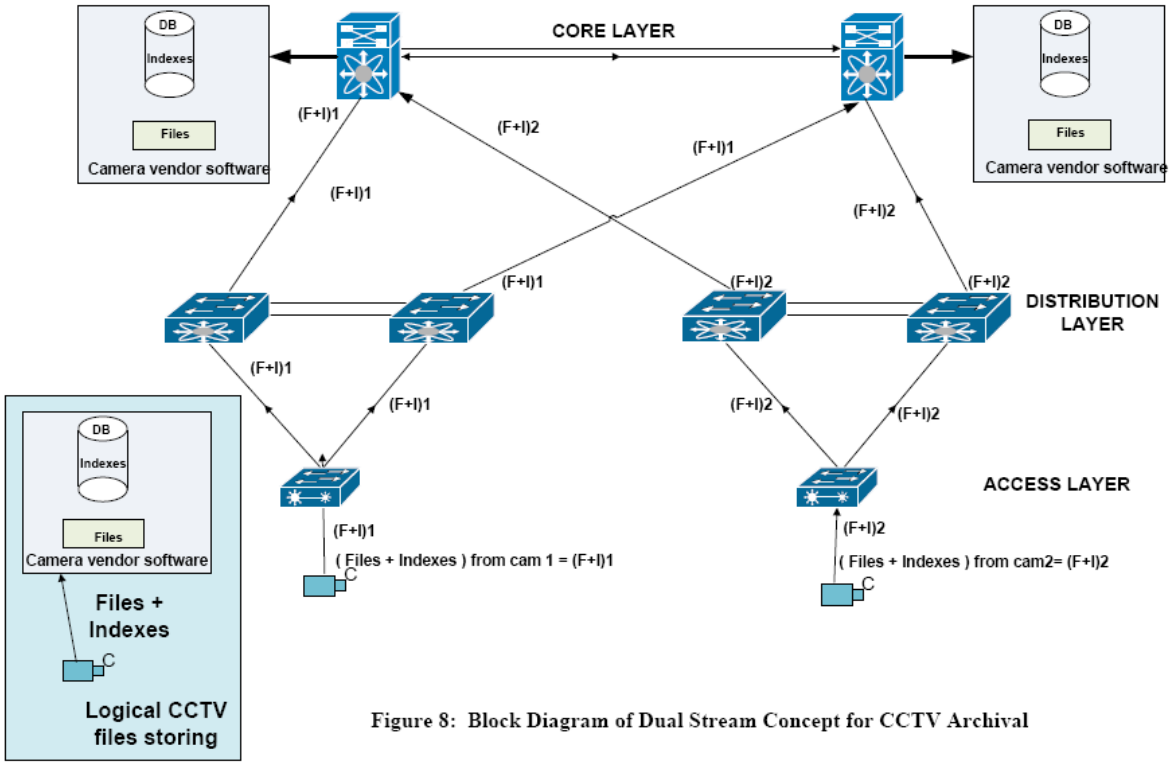

Figure 8: Block Diagram of Dual Stream Concept for CCTV Archival

Conclusion

The optimal backbone infrastructure has to be able to serve a great number of devices and users covering a large area. Presented concept of splitting the entire networking infrastructure to Access, Distribution and Core layers enables clear distinctions between different networking requirements at various project locations. Development of the new and advanced switching technology enables increased functionality and availability operations.

Price reductions for fiber optical cables and their common use make it possible to develop such an extensive fibre plant covering the entire alignment and connecting to different user's levels.

Advances in network switching technologies have made it possible to design solutions that include;

- Communications links are bundled together to allow for cumulative bandwidth between multiple locations. This technology will also allow for no network downtime when there are failures to communications cables.
- New optical transceivers allow for greater communications speed between locations allowing for higher performance of networked applications.
- Technologies are now able to encapsulate networks so that multiple systems can share higher bandwidth connections by means of virtual networking.

Redundancy, throughput, scalability, and maintainability are some of the reasons that the modern switching and data processing IT technologies are gaining acceptance for the LRT communication infrastructure implementations.