Re-engineering a signal control system

Roy Heaton & Richard Kirby
Alcatel Canada Inc. Transport Automation

Abstract

Mass Transit systems are making increased use of technology to bring more information to the desk of the Operators. System performance is therefore directly impacted by the large amount of data being managed by the system both coming to the operators from various sensors or subsystems, and command and control information going in the opposite direction. In order to achieve optimum performance, the system architecture must be adapted to accommodate this increased level of information and associated data traffic. The Jubilee Line Extension Mainline Control System in London is a good representation of a railway control system that has to handle the increased information flow.

This paper presents an overview of the architectural and design changes that were introduced in the Jubilee Line Extension control system to handle the challenge of increased data. The goal of this system architecture change was to improve the reliability, maintainability, recoverability and performance characteristics of the overall signal control system. The major changes to the architecture included the introduction of IBM’s System Network Architecture (SNA) for use as the system’s communications backbone and a distributed catalogue mechanism to manage data distribution throughout the system. At the application level, a redesign of the Alarm subsystem and Access Control subsystem were performed. The new design of the alarm subsystem resulted in the addition of a flow control mechanism as well as superior failure and recovery characteristics. The Access Control subsystem was changed to eliminate the need for a centralised server.

For each change presented, an overview of the design prior to and after the change is given. The technical considerations are discussed from both a system and operational perspective.
1 Introduction

The London Underground Limited Jubilee Line Extension project started in 1993 with the goal of developing a state of the art railway for the new Millennium through the rejuvenated Docklands area. The Jubilee Line has been extended by 11 stations (16 km) for a total of 28 stations (38km). The line services the Millennium Dome at North Greenwich and provides a shorter travel time between the Docklands area and Central London. The extended line entered into end to end revenue service on November 20, 1999 and was supporting 225,000 journeys daily by late December 1999. The current revenue signaling system is a fixed block system capable of 24 trains per hour. The Mainline Control System (MCS) controls the signaling on the extension and the Metropolitan Line Control Centre at Baker Street controls the existing line. The two systems exchange control of trains using a “Seamless Handover” process. The MCS located at Neasden integrates the following services:

- Automatic Routing
- Manual Routing
- Fixed Block Train Tracking
- Service logs
- Alarm management
- Train Prediction
- Communication (Public Announcements, Closed Circuit TV, Radio, etc.)
- Overview display of the extended line

The control system is unusual in that it integrates traditional signal control system functionality with communication functions. The operators have access to all functions through an integrated operator’s desktop.

Although the MCS is not a safety critical system, it is mission critical to the smooth operation of the railway, so a great deal of design effort was directed at system performance, fault tolerance, and failure and recovery characteristics. This paper highlights four key areas and describes the corresponding solutions that were developed and implemented.

2 System Overview

This section details the objectives, constraints and system architecture for the mainline control system. It was in the context of these elements that the re-engineering was performed.
2.1 Mainline Control System Objectives

Although the current implementation of the MCS controls the extension of the Jubilee Line, the system was architected to control both the existing line and extension. The MCS for the extended Jubilee Line was designed to support a 36tph, moving block service. In addition, the system was designed to monitor 10,000 SCADA points and over 1000 CCTV cameras.

2.2 Reliability Constraints and Design Goals

The MCS was specified to meet the following reliability constraints and design goals:

- Minimise reliance on a single processor or single communications link.
- Minimise the loss of functionality during a single point of failure scenario.
- Minimise the loss of data due to a failure and recovery.
- Minimise the time to recover from a failure.
- 99.97% system availability

2.3 System Architecture

The MCS architecture is shown in figure 1. All processors in the system are dual redundant. This represents the current physical architecture of the system. The design changes described later did not require any significant changes to be made to the underlying physical architecture.

The control centre processors communicate via a redundant 10Mbps LAN. The following is a list of the control centre processors together with an overview of their functionality:

- Server - data logging, alarms management, and timetable management.
- Local Interface Controller - prediction services, master clock and interface to Metropolitan Line Control Centre.
- Two LNP Interface Controllers - network routing services.
- Eight Universal Workstations - operator interface for signal control and communications functions, and overview of the train service.
- CIMS Interface Controller - interface to the Communication System that supplies Public Announcement, Passenger information and Closed Circuit TV services.
- Three Overview Display Controllers - rear projection of the entire Jubilee Line status.

Distributed along the line are six Local Non-vital Processors (LNPs) situated at Stratford, West Ham, North Greenwich, Canary Wharf, London Bridge and...
Waterloo. They are responsible for issuing signalling commands to the interlocking processors within an area of control. An LNP interfaces with up to 3 interlocking processors. The LNPs communicate with the central processors using dedicated 64 kbps redundant channels and with neighbouring LNPs using dedicated 19.6kbps redundant channels. The communications channels are provided by an Optical Fibre Backbone. For clarity, the individual communications channels are not shown in figure 1.

Each LNP is responsible for making routing decisions, calling routes, tracking trains, and monitoring the interlocking processors within an area of control. LNPs are required to continue operation should communication with the central processors or neighbouring LNPs fail.

Figure 1. System Architecture (Not all processors shown)
3 Design Improvements

In consideration of the design goals, reliability constraints and the system architecture presented in the previous section, Alcatel’s project team reviewed the initial architectural design. The review and implementation of recommendations took place over a period of twelve months, prior to the start of revenue service.

Four of the areas identified in the review are presented in this section. For each area, an overview of the design prior to and after the change is given. Technical considerations are discussed from both a system and operational perspective.

3.1 Network Architecture

The original network design was limited by the lack of Commercial of the Shelf (COTS) software available for the selected hardware platform at the time of the original proposal. A decision was made to develop the networking software in-house, as had previously been the case on Alcatel’s centralised SELNET™ control systems. The design used two different protocols: one protocol to communicate between the remote processors (LNPs) and the central interface processors (LNPIFs); and another protocol to communicate between all the processors on the central LAN. The interface processors acted as a communications gateway and protocol translator. A life pulse mechanism was used to detect communications and processor failures. A management module located on one of the central processors managed the recovery from such failures. A higher-level session protocol was put in place to preserve message ordering and to guarantee message delivery. The result was a complex piece of software that had throughput constraints.

At the time of the architectural phase, fundamental assumptions were made regarding the data communications bandwidth required for proper signal control of a typical moving block system. Initial prototype testing demonstrated that the overhead associated with the messaging traffic was excessive and that a redesign was necessary.

A computer simulation of the network was created to analyse the network throughput and capacity constraints using TCP/IP and SNA. A model of the in-house solution was used as a benchmark. Using the results of the simulation, a decision was made to use the SNA technology. The advantages of using this technology compared to the in-house design were:

- Automatic communications link failure detection and recovery/failover.
- Link aggregation that provides a higher bandwidth capacity for peak message traffic.
- A single communications protocol used throughout the system.
- Vendor’s reputation for reliability and support.
The SNA technology is implemented in software so no major hardware modifications were required. The alternative solution using an IP network provided similar advantages but required extensive hardware changes and was rejected for that reason.

An additional advantage of using a standard network protocol is the availability of utilities such as TELNET, PING and FTP. These types of utilities are available with SNA and proved to be of enormous value in trouble-shooting, error correction and distribution of new software releases.

In addition, a new software task was introduced called a Proxy Server. The purpose of the Proxy Server is to reduce the amount of message traffic on the communication links between the LNPs and the central interface processors. This is achieved by programming the Proxy Server to act as a message forwarder for data originating in the LNPs that is required on the UWSs. In this way it performs the function of a fan-out message gateway.

3.2 Distributed Catalogue

One of the fundamental system design decisions was to make the system data driven. The software is able to register interest in a piece of data at runtime and define whether it produces or consumes that piece of data. A software architecture to support this capability was developed using producer-consumer relationships and a centralised catalogue containing the relationship definitions together with the interfaces to allow the application software to register for data at runtime. There were some problems with this approach:

- A centralised catalogue presents a single point of failure for the system
- The centralisation of the registration software implies that the processor where the catalogue resides must be started before all other processors. This makes system recovery after a processor failure much more difficult.
- Any change to a registration entry means that data must be sent to a central processor and this can lead to race conditions and latency.
- At start-up, an enormous amount of data is sent over the network to define all the producer-consumer relationships.

Upon further analysis, it was determined that nearly all the producer-consumer relationships are static - they do not change during the operation of the system. In fact, only the identity of the producers and consumers for data associated with trains changes dynamically and the rate of change is relatively slow. A change will occur when a new train enters or leaves the system, and when a train exits from the area of control of one LNP and enters into the area of control of the next LNP. A modified software architecture based on the concept of pre-defined data registration definition files that reside on each processor together with the ability to modify the producer-consumer relationships for train data at run-time was developed.
The following advantages are available with this approach:

- Processors can start independently, in any order.
- System start-up is quicker because data is local to each processor.
- The system has no single point of failure.
- The race condition and latency have been eliminated as the data catalogue resides on each processor.

3.3 Alarm Subsystem

In any control system, alarms play a very important role. Alarms keep operators informed of unusual events and prompt operators when manual intervention may be required. Operators of the control system need to be confident in the alarm subsystem as it is integral to their job function. In the MCS each processor in the system is responsible for raising alarms detected by the software running on that processor. Some processors will also raise alarms on behalf of external systems when requested e.g. the CIMS interface computer raises alarms within the MCS corresponding to alarms received from the communications system.

The original design was for a centralised thread that ran on the server processor. The clients sent a alarm message directly to the centralised thread, which then determined which operator(s) should receive the alarm and routed it accordingly. This thread also managed the lifecycle of the alarms.

The design was not resilient to failures of the server processor or to failures of other processors. Alarms could become lost or out of date.

The new design concentrated on improving the resilience of the alarm subsystem while retaining as much of the original design as possible. A flow control mechanism was added to ensure that alarms being raised did not overwhelm the system. The alarm lifecycle management was transferred to the processor raising the alarm. Alarms are buffered locally until acknowledged.

The new design of the alarms subsystem has:

- improved the fault tolerance of the alarms subsystem.
- increased system robustness under heavy load conditions.
- eliminated the possibility of losing alarms should the server processor fail.
3.4 Access Control

The Access Control subsystem provides the following services:

- operator authentication.
- ensure each operator is logged into only one active workstation at any given time.
- operator identification to assist in post incident analysis.

The original design comprised a single task that ran on the server processor. Operator authentication requests were sent to the access control task to be processed. The same task was also responsible for identifying that a failure had occurred so that the operator could be allowed to log into a second workstation and re-assume his duties. In the event of a server failure, no operators could log into the system. Upon recovery of the server, all operators would have to log in again. This lack of fault tolerance in the access control system led to a re-design.

The new design distributes the operator authentication functionality to each workstation, thus removing altogether the need for a centralised task. The workstations distribute the access control data amongst themselves to ensure that operators are only logged in once. Processor failures are detected by each workstation so that in the event of a failure of a single workstation, the operator simply moves to another workstation and logs in. The access control system is now resilient to single failures within the MCS.

4 Conclusions

Signal Control Systems are making more use of technology to integrate previously isolated signalling and communication functionality into a single operator interface. This paper presented the Jubilee Line Extension Mainline Control System as an example of an integrated control system. The paper described some of the problems identified in the system architecture during the design phase of the project together with the solutions that were identified and the resultant system characteristics in terms of system performance, fault tolerance and robustness.

By enhancing and adapting the system architecture, the MCS has been able to accommodate an increased amount of data typical of an integrated control system. This required a greater amount of effort to be invested in the design of the system architecture than is normally the case with smaller, non-integrated control systems.