Power Quality Application Guide

Resilience, Reliability and Redundancy

Copper Development Association
IEE Endorsed Provider

Resilience
Resilience, Reliability and Redundancy

Gary Marshall, WSP Communications Ltd &
David Chapman, Copper Development Association
May 2002

European Copper Institute (ECI)

The European Copper Institute is a joint venture between ICA (International Copper Association) and IWCC (International Wrought Copper Council) contributing members. Through its membership, ECI acts on behalf of the world's largest copper producers and Europe's leading fabricators in promoting copper in Europe. Formed in January 1996, ECI is supported by a network of ten Copper Development Associations ('CDAs') in Benelux, France, Germany, Greece, Hungary, Italy, Poland, Scandinavia, Spain and the UK. It furthers the efforts initially undertaken by the Copper Products Development Association, formed in 1959, and INCRA (International Copper Research Association) formed in 1961.

Copper Development Association (CDA)

Copper Development Association is a non-trading organisation sponsored by the copper producers and fabricators to encourage the use of copper and copper alloys and to promote their correct and efficient application. Its services, which include the provision of technical advice and information, are available to those interested in the utilisation of copper in all its aspects. The Association also provides a link between research and the user industries and maintains close contact with the other copper development organisations throughout the world.

CDA is an IEE endorsed provider of training seminars and learning resources.

Acknowledgements

This project has been carried out with the support of the European Community and International Copper Association, Ltd.

Disclaimer

Copper Development Association

Resilience, Reliability and Redundancy
– the three R's of electrical design

Introduction

Availability - the proportion of time that the system is actually available to do real work - is one of the key parameters in determining the quality of a power supply. The absolute values of availability required or expected are very high - a typical domestic user in Western Europe would expect an availability of 99.98 % while a commercial user may require better than 99.996 %. Such high levels of availability do not happen naturally or by chance: they are the result of excellent design and maintenance procedure and attention to detail.

Unless the designer takes steps to avoid it, availability potentially reduces at every connection, protection device, cable, etc throughout the installation so that it is highest at the point of common coupling (PCC) and lowest at the equipment terminals.

Availability at the PCC is typically around 99.98%, largely because the supply network is highly redundant, i.e. there are many available routes between the consumer and the generation points, each route being continuously monitored and maintained to ensure the highest (economically justifiable) availability. If a fault develops there is a very high probability that an alternative route is available so that there is no interruption, or at the worst a short interruption, of power to the load. Installation design lags far behind in this respect, still using single string connections, individual transformers with no standby or cross-switching connections. Putting it another way, the network can survive the failure of one or more individual components or connections - it is resilient - while the installation has many individual single points where failure will cause interruption of supply.

Since the middle of the last century great emphasis has been placed on the reliability of components and sub-systems. The interest has been in reducing failure rate, thus increasing operational lifetime.

Over the last decade the focus has changed; designers and systems analysts are now much more concerned about the availability of their systems. High availability requires much more than high reliability. Depending on the situation and the importance of the equipment to the operation, it may also require short repair times, a degree of on-line redundancy, 'hot swap' replacement of critical sub-systems, good diagnostic strategies and appropriate spares stocks. Whatever the situation, maintaining high availability requires careful forward planning and good maintenance management. In fact, it is much more about management and planning than engineering.

High reliability, though not enough in itself, is still a fundamental requirement; maintaining high availability of unreliable equipment is unrealistic because it requires too much effort and a large stock of spares and redundant equipment. Resilient design requires the intelligent disposition of inherently reliable equipment in a redundant topology, coupled with well-defined and rigorously implemented maintenance procedures.

The basic requirement of high reliability of individual units is discussed in the next section.

Background

Reliability

Reliability is measured in terms of the Mean Time To Failure (MTTF). The MTTF is quoted in hours and is the average time taken for each individual member of a large population of standard production parts to fail. Reliability is difficult and time consuming to measure; it is a statistical number derived from a combination of historic experience with similar components and relatively short-term tests on large numbers of the components in question. As products mature, longer-term data accumulates and confidence in the MTTF figure increases. But it is not a guarantee!

Failure may occur as a catastrophic failure, i.e. it is complete and sudden, like a light bulb failing, or as degradation i.e. it is gradual or partial, like an electronic unit moving outside specification. In the case of an electrical supply, a complete loss of power would be a catastrophic failure, while voltage or frequency deviation would be regarded as degradation. A failure is primary if it is not caused by failure in another part of the system, and secondary if it is the result of the failure of another part of the system. As an example, the normal failure of a light bulb would be a primary failure, but failure following a severe overvoltage would be a secondary failure because it is the result of a failure elsewhere. Reliability data can only give information about primary failures.

Some failures may be obvious, such as the blowing of a light bulb, but in other cases, it may be much harder to determine that a failure has occurred. For example, a complex electronic circuit that must meet a detailed specification has failed if any parameter has moved outside the specified limits although this may not be apparent to the user. An electronic interface circuit may be required to have a certain immunity to noise voltage; failure to maintain this immunity would not be noticed under noise free conditions, but the symptoms would be apparent under more extreme conditions. In such cases failures and errors may appear to be random and it is likely to be difficult to localise the cause and effect a repair rapidly.

Figure 1 shows the failure rate $\lambda(t)$, plotted against time. It is often referred to as a 'bathtub curve' and is typical of that obtained for many electronic components. It has three distinct sections:
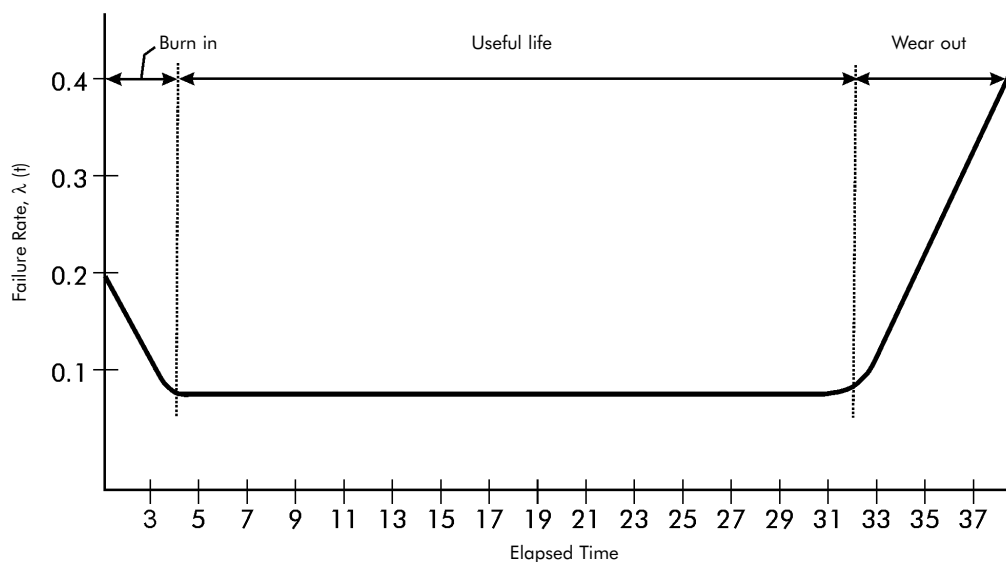

Figure 1 - Failure rate against time

Early life or burn-in

This is the period (up to time unit 4 in this illustration), when $\lambda(t)$ is decreasing as weak or substandard components fail. It is known as the early life, early failure, infant mortality or burn-in period. The reasons for the first three terms should be obvious. Burn-in is a process sometimes used in the final stages of manufacture of components in order to weed out early failures. It involves the components being run under normal conditions (or controlled conditions somewhat more severe than normal to accelerate the process) for sufficiently long to survive the early life period.

Useful life or normal operating period

This is a period of effectively constant, relatively low $\lambda(t)$ (from time unit 5 to time unit 31 in Figure 1), known as the useful life or normal operating period. During this time, the failure rate is independent of the time for which the component has been run. In other words, the probability of failure of a component is the same throughout this period.

Wear-out or old age

This is the period, when the failure rate increases steeply with time (beyond time unit 31 in this illustration), known as the wear-out or old age period.

The bathtub curve describes the behaviour that might well be expected for many types of component, or even for complex systems such as a UPS. Taking the familiar example of light bulbs, failures during early life might be due to filaments that have been badly attached to their supports, that have been locally drawn too thin or that have nicks in them. Another cause of early failure could be leaky glass envelopes, leading to filament oxidisation. During useful life the filaments gradually evaporate and become thinner until they break, usually under the thermal stress induced by current surges at switch on. If all bulbs were identical and were operated under identical conditions, they would all fail at the same time. However, since no manufacturing process will produce identical components, some will have serious faults leading to failure during early life, and the failure rate decreases as the weaker components are weeded out. Similarly, it is reasonable to expect a range of failure times during the wear-out period. In the case of light bulbs, filament composition, thickness, length and shape will vary slightly from bulb to bulb, thus leading to a spread in the time at which they finally break. As the bulbs get older, their probability of failure increases, giving a steep rise on the right-hand side of the bathtub curve.

It is perhaps harder to see why there should be any failures at all during the normal operating period, once early failures have been weeded out and there has not yet been any significant wear. A detailed analysis of the numerous ways in which a light bulb could fail would be extremely complex because of the variety of mechanical, thermal and chemical processes that can take place. The wide range of possibilities, each with its own time dependence, averages out to produce a failure probability that is effectively independent of time. The process is referred to as 'memory-less', because the probability of failure is assumed to be independent of previous history. These failures are also sometimes called catastrophic because they are unexpected. In fact, in many cases, the wear-out processes are already taking place at a very low level so the process is not truly 'memory-less'.

The manufacture of components and assemblies often involves a burn-in process so that those subject to early life failure can be removed from the supply chain. In many cases, electronic components do not reach their wear-out period during operational life - they may have useful lives that are much longer than the operational life of the system in which they are used. Routine maintenance procedures should be designed to ensure that components that can wear out – rechargeable batteries, for example - are replaced well before the onset of wear-out. Because of this, it is usually possible to assume that components are run only during their period of useful life and that they have a constant failure rate.

Mean time to failure (MTTF)

The mean time to failure is a term that is applied to non-repairable parts such as light bulbs and is a measurement of the average time to failure of a large number of similar parts which operate under specified conditions. The conditions of test are important, for example an increase in the operating temperature or voltage of most components will reduce the MTTF.

In practice, the MTTF is often calculated from data taken over a period of time in which not all the components fail. In this case

$$\text{MTTF} = \frac{\text{Total operating time for all components}}{\text{No of failures at that time}}$$

Mean time between failures (MTBF)

In the case of components or system elements that can be repaired, failure rates are often expressed in terms of the mean time *between* failures (MTBF) rather than mean time to failure (MTTF). This is a measure of the average time that a piece of equipment performs its function without requiring repair (although it may require routine scheduled maintenance).

Because MTTF and MTBF are statistical quantities, a large number of failures must be recorded in order to establish confidence in the result. Testing one piece of equipment for a very long time is impracticable so it is usual to test a large number of samples simultaneously for a shorter period, and to determine the total number of faults in the total operating time for all of the samples. This method assumes that burn-in and wear-out failure modes are not involved.

Availability and mean time to repair (MTTR)

System users are now more concerned with availability rather than reliability. The availability of a system or of a component is the proportion of time for which it is operating correctly and available for use. It is the ratio of operational time to total time, which is the sum of the operational and repair times. The average time needed for repairs is called the mean time to repair (MTTR).

$$\text{Availability} = \frac{\text{MTBF}}{\text{MTBF} + \text{MTTR}}$$

MTTR times are specified by manufacturers under ideal conditions, assuming, for example, that the site of the failure is obvious, that the required skill level to identify it is available, that suitable spares and repair procedures are on hand and that, once repaired, the system can be tested and recommissioned. In real life failures occur at the worst possible time. A quoted MTTR of twenty minutes is of little advantage if an engineer has to travel to site to identify the problem and then procure parts from a depot or, worse, a distributor. Following repair, the system must be tested and recommissioned before real work can begin again. Each of these steps may take many times longer than the repair itself. Some delays can be avoided - critical spares can be held locally, in-house staff can be trained to provide the necessary technical skills – but these steps have to be implemented well in advance and are not without their own cost implications. Taking the simple example of a television repair that takes, say 30 minutes to replace a circuit board, the manufacturer would quote a MTTR of 0.5 hours. But the user sees things differently. The set breaks down while the repair shop is closed. It has to be collected, the fault identified, and the replacement parts identified. After an estimate is prepared and approved, the part is procured, and - now the 30 minutes start - fitted and tested. The next day the set is returned to the user, who has experienced a MTTR of perhaps one week. If the MTBF were long enough the user may be quite content, but if the MTBF is short the user may well want to have a redundant television set available!

Often, the repair will take the form of replacement by a new or previously repaired component or sub-system. When a repaired sub-system is fitted it must be remembered that only a few of the components are new while the rest of the components are already aged. The MTTF must be less than that of a new unit.

To be completely fair, the MTTR should also include the downtime required for normal in-service preventative maintenance, such as servicing of a generator engine, replacement of the batteries or cleaning the ventilation fan filters of a UPS, etc. Fortunately, downtime to allow for preventative maintenance can be planned so that disruption is minimised but, often, cannot be eliminated.

Maximising the proportion of time a system is available involves trade-offs between component reliability and repair time. For instance, plug-in components are usually much less reliable than hard-wired ones, because of the relatively high failure rates of connections. On the other hand, the repair times of plug-in components will be much shorter than those for hard wired ones because they can simply be replaced by relatively less skilled operators. The optimum balance depends on the absolute values of MTBF and the true MTTR, taking full account of the facilities and expertise available on-site.

System availability is simply the MTBF divided by MTBF plus the true MTTR – including the extra items just discussed. Bald availability figures often look very impressive but must be used with care. An apparently excellent availability of 0.9999 translates to an *unavailability* of about one hour per year – a completely unacceptable level for many operations. Achieving better than this in practice - simply by increasing reliability - is quite difficult; MTBF figures, being statistical, cannot be relied upon in any particular instance and MTTR figures are usually optimistic, being based on the assumptions made in the calm atmosphere of a working system. Furthermore, the complexity of modern systems, such as commercial computer networks and factory process control, means that a large number of sub-systems have to work perfectly

together to keep the operation running. Achieving higher availability requires a different design approach, one that focuses on tolerating and surviving the inevitable failures rather than trying to reduce them - in other words, resilient design.

Resilience

A resilient system is one that can withstand a number of sub-system and components failures while continuing to operate normally. This can be achieved by installing additional equipment – redundancy – together with careful design – eliminating single points of failure - and well planned maintenance.

A balanced approach is essential. Adding redundant equipment is straightforward but it is expensive and so must be done after proper consideration of the costs and benefits. Introducing proper maintenance procedures is relatively cheap, but ensuring that they are properly observed is never easy! A careful balance of technical and managerial or procedural solutions will give the best cost/benefit ratio.

The essential elements of a resilient installation are:

- ◆ redundancy
- ◆ elimination of single points of failure
- ◆ good maintenance procedures

Redundancy

Redundancy is a useful method of increasing availability and optimising the balance between operation effectiveness and expenditure. Alternative circuits, equipment or components are installed so that, in the event of one or a few failures, functionality is preserved. The level and type of redundancy provided determines the level of functionality retained and the number and types of faults permitted.

Standby redundancy

Standby redundancy means that an alternative means of performing the function is provided but is inoperative until needed. It is switched on upon failure of the primary means of performing the function. An IT department might keep several spare monitors and use them to replace those that fail in service – this is an example of standby redundancy.

The disadvantage of standby redundancy is that there is an inevitable period of disruption between the failure occurring and the redundant unit being brought into use. Such schemes are rarely satisfactory for critical systems in modern commercial and industrial situations.

Active or parallel redundancy

In active or parallel redundancy all redundant units are operating simultaneously rather than being switched on when needed (Figure 2). The most obvious approach is to use two components, each capable of carrying the full load so that, if one should fail, the other will take over - this is referred to as 1+1 redundancy. An alternative approach is to split the load among a number of units, each capable of carrying only a fraction of the load, and provide just one additional redundant unit - this is referred to as N+1 redundancy. For very critical loads, more than one fully rated redundant unit may be provided. For example, a 1+2 redundancy scheme would have three fully rated units supporting the load and would require all three units to fail before the system fails.

Because there is no interruption, active redundancy is suitable for computer installations. The use of mirrored disks in a server computer is an example of 1+1 redundancy.

N+1 redundancy

A lower cost alternative is to use a larger number of lower rated units to support the load and to provide one extra redundant unit. For example, a 800 kVA load could be connected to five 200 kVA UPS units, any four of which would be sufficient to support the load. This would be referred to as 4+1 redundancy because four
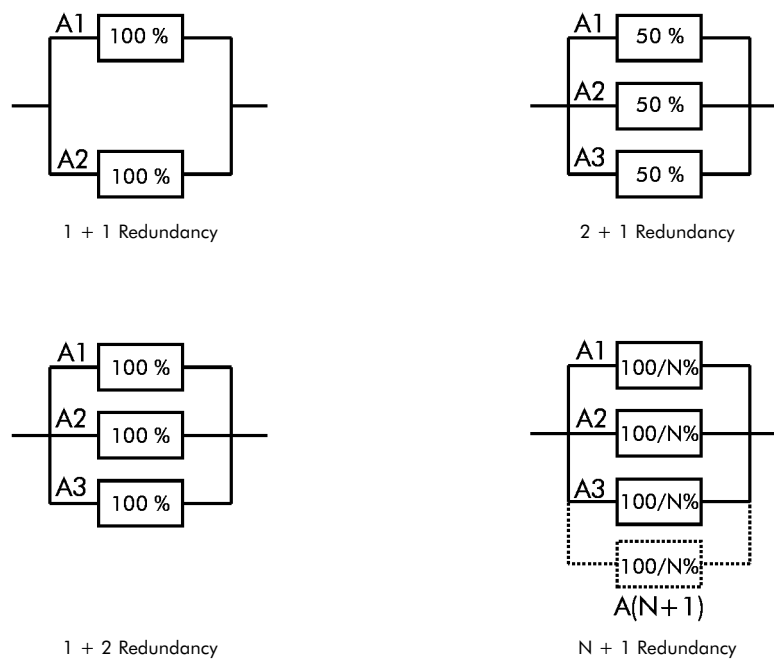
Figure 2 - Examples of active redundancy

units are required and one is redundant. N+1 redundancy can be cheaper to implement than 1+1, and is more flexible; if the load grows it is simple to add another unit so that a 2+1 system becomes a 3+1 system. On the other hand, extremes should be avoided; every additional unit is another potential failure so there is a balance to be struck between the additional *system* resilience versus additional cost and increased *unit* (not system) failure potential.

A RAID 5 disk system in a server computer is an example of N+1 redundancy. Additional error correction data is added to the original data and then logically dispersed over a number of disks. If one disk fails, all the data can be recovered in real time from the remaining disks and the failed disk contents reconstructed.

It is worth noting that the provision of a UPS in the first place is an acceptance of the need for a redundant power supply and adding redundancy to the UPS itself is a logical extension of that concept.

Eliminating single points of failure

Adding redundancy is costly so how much and what to provide must be carefully considered.

A risk assessment must be carried out, in which the risk of a particular failure occurring and the effect of that failure on the system is judged. Priority should be given to those failures where the combination of probability and system effect is highest - possibly including a very rare event of catastrophic effect and a frequent event that causes minor disruption. This study should identify all single points of failure – those simple single failures that can bring down a whole operation. The study should be exhaustive and pedantic, looking not only at major pieces of equipment but also the control boxes and signal wiring that connects them. If the control system or its interconnections are the cause of the failure, the expensive standby equipment will not help!

Figure 3 shows an example of a resilient power supply system. This design has two independent supplies from separate grid points backed up by two independent standby generators. Power from either input rail can be routed to either or both output rails directly or via either of two UPS units. Static transfer switches (STS) are used to connect and isolate equipment and paths as required. This example is symmetrical and assumes that the loads on each output are equally critical; in other circumstances, one output may be used to supply highly critical loads while the other supplies normal loads. In normal service power from one
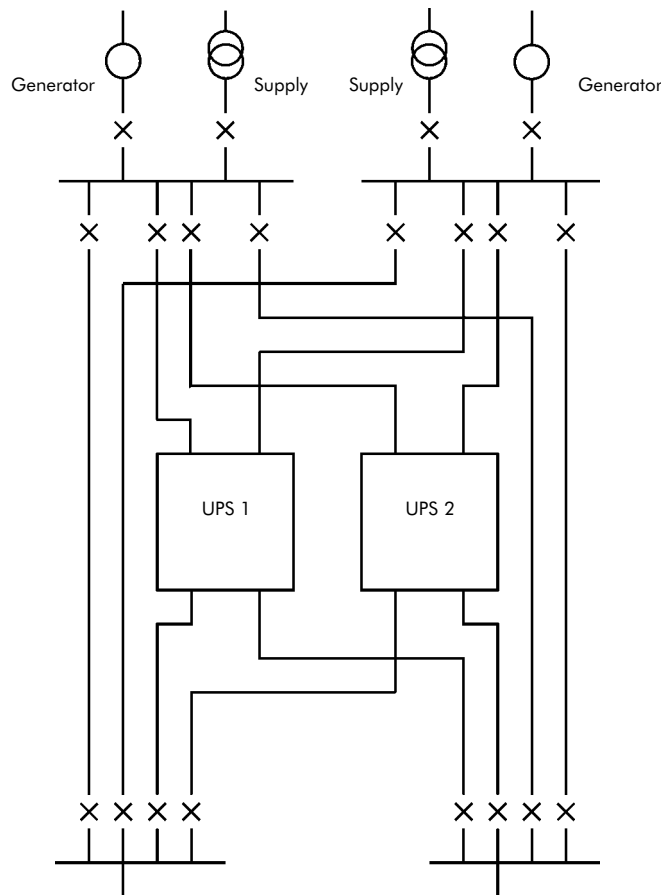
Figure 3 - An example of a resilient system

supply would be routed to both outputs via a UPS. In the event of a supply failure, connection would be made to the alternative supply with the UPS providing power during the short changeover period. If the alternative supply were unavailable, then a standby generator would be started and once up to speed, connected to the system; again a UPS would provide power during the start-up delay. If both supplies and both generators should fail, or be out of service for maintenance, then the UPS units would be used to supply the output rails independently. Load shedding would then be applied to extend the life of the supply.

This is an extreme example. It would be rather expensive to implement but may be justifiable in situations where loss of life or serious financial loss may result from failure. It is used here to illustrate what can be achieved but in most circumstances a subset of this system would be appropriate based on a critical examination of the risks involved and the cost of protection.

Maintaining a resilient system

Maintenance procedures are the key to preserving the resilience of a system and these procedures must be taken into account during the risk appraisal. If, for example, one of four UPS units has failed, and the remaining three are carrying the load, how is this situation recognised so that a repair can be effected? What is the risk of a second failure occurring while the remaining units are operating at the higher load? Could the load have grown after commissioning beyond the capacity of only three units without the maintenance staff being aware? Detailed procedures must be developed to cover these and many other situations during the risk analysis phase so that they can be rigorously tested.

The fact that an installation has been designed to be resilient can encourage lax maintenance standards.

After all, it doesn't matter if a particular unit fails - the site can survive without it. This attitude is a recipe for disaster. Firstly, there are more than enough reasons why a failure should occur and there is no need to multiply the possibilities. Secondly, if a unit breaks down due to lack of maintenance, it is likely that its counterparts are in a similar state and are likely to fail as well when faced with additional load.

Testing the system is also important. Maintenance staff need to be confident that the system will cope with failures and must test regularly by simulating the loss of individual items of equipment. This is a brave step, best first taken early in the commissioning of the site. Having established that the site is resilient, regular testing can be undertaken with high confidence of success. Naturally, testing should always be undertaken when the consequences of failure are at a minimum, but must also represent real site conditions.

Conclusion

Availability is the primary concern of installation designers and maintainers. High availability cannot be practically achieved by high reliability alone - good resilient design and proper maintenance procedures are also required. Designing and implementing a resilient system is only the first step in a long journey.

A hidden single point of failure

Figure 4 shows an installation in which a significant part of the load is made up of motors with power factor correction provided by a switched capacitor bank. Control circuits switch in the correct capacitance to correct the power factor of the instantaneous motor load, ensuring that reactive or 'wattless' current is minimised. This reduces the load on transformers, switchgear and cables as well as the tariff charged by the supplier.

Because the power factor of the load is compensated, the magnitude of the current is reduced from the normal, uncompensated value. Now the transformer has spare capacity and is capable of carrying additional load - but only while the PFC system is functional. If extra load is added and the PFC system fails, the load current will exceed the capability of the transformer resulting in an overload trip or ultimate failure.

The PFC system, its measurement and control system, even the PFC circuit breaker are all now critical to the survival of the transformer and the site. The reliability of the whole installation is reduced to the least reliable of these components.
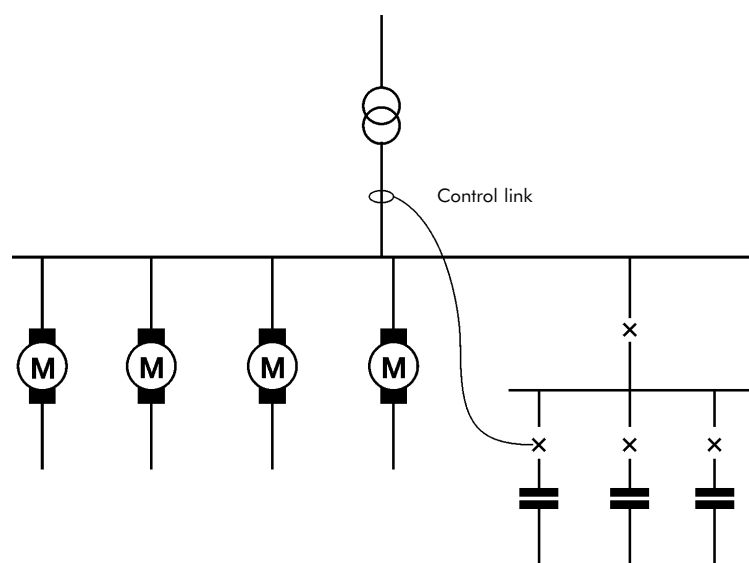

Figure 4 - Power factor correction using a controlled capacitor bank

Network Partners

Copper Benelux
168 Avenue de Tervueren
B-1150 Brussels
Belgium

Tel: 00 32 2 777 7090
Fax: 00 32 2 777 7099
Email: mail@copperbenelux.org
Web: www.copperbenelux.org

Contact: Mr B Dôme

Copper Development Association
5 Grovelands Business Centre
Boundary Way
Hemel Hempstead HP2 7TE
United Kingdom

Tel: 00 44 1442 275700
Fax: 00 44 1442 275716
Email: helpline@copperdev.co.uk
Webs: www.cda.org.uk & www.brass.org

Contact: Mrs A Vessey

Deutsches Kupferinstitut e.V
Am Bonneshof 5
D-40474 Duesseldorf
Germany

Tel: 00 49 211 4796 323
Fax: 00 49 211 4796 310
Email: sfassbinder@kupferinstitut.de
Web: www.kupferinstitut.de

Contact: Mr S Fassbinder

ECD Services
Via Cardinal Maffi 21
I-27100 Pavia
Italy

Tel: 00 39 0382 538934
Fax: 00 39 0382 308028
Email: info@ecd.it
Web: www.ecd.it

Contact: Dr A Baggini

European Copper Institute
168 Avenue de Tervueren
B-1150 Brussels
Belgium

Tel: 00 32 2 777 70 70
Fax: 00 32 2 777 70 79
Email: eci@eurocopper.org
Web: www.eurocopper.org

Contact: Mr H De Keulenaer

Hevrox
Schoebroeckstraat 62
B-3583 Beringen
Belgium

Tel: 00 32 11 454 420
Fax: 00 32 11 454 423
Email: info@hevrox.be

Contact: Mr I Hendrikx

HTW
Goebenstrasse 40
D-66117 Saarbruecken
Germany

Tel: 00 49 681 5867 279
Fax: 00 49 681 5867 302
Email: wlang@htw-saarland.de

Contact: Prof Dr W Langguth

Istituto Italiano del Rame
Via Corradino d'Ascanio 4
I-20142 Milano
Italy

Tel: 00 39 02 89301330
Fax: 00 39 02 89301513
Email: info@iir.it
Web: www.iir.it

Contact: Mr V Loconsolo

KU Leuven
Kasteelpark Arenberg 10
B-3001 Leuven-Heverlee
Belgium

Tel: 00 32 16 32 10 20
Fax: 00 32 16 32 19 85
Email: ronnie.belmans@esat.kuleuven.ac.be

Contact: Prof Dr R Belmans

Polish Copper Promotion Centre SA
Pl.1 Maja 1-2
PL-50-136 Wroclaw
Poland

Tel: 00 48 71 78 12 502
Fax: 00 48 71 78 12 504
Email: pcpm@miedz.org.pl

Contact: Mr P Jurasz

TU Bergamo
Viale G Marconi 5
I-24044 Dalmine (BG)
Italy

Tel: 00 39 035 27 73 07
Fax: 00 39 035 56 27 79
Email: graziana@unibg.it

Contact: Prof R Colombi

TU Wroclaw
Wybrzeze Wyspianskiego 27
PL-50-370 Wroclaw
Poland

Tel: 00 48 71 32 80 192
Fax: 00 48 71 32 03 596
Email: i8@elektryk.ie.pwr.wroc.pl

Contact: Prof Dr H Markiewicz

Gary Marshall

WSP

WSP Communications Ltd
Buchanan House
24-30 Holborn
London EC1N 2HS
United Kingdom

Tel: 00 44 20 7314 5000
Fax: 00 44 20 7314 5700
Email: gary.marshall@wspgroup.com
Website: www.wspgroup.com

David Chapman

Copper Development Association

Copper Development Association
5 Grovelands Business Centre
Boundary Way
Hemel Hempstead HP2 7TE
United Kingdom

Tel: 00 44 1442 275700
Fax: 00 44 1442 275716
Email: helpline@copperdev.co.uk
Websites: www.cda.org.uk & www.brass.org

Copper Development Association