Redundanz bei Prozeßleitsystemen

Dr.-Ing. Weinert, Albrecht, Siemens AG, Karlsruhe

1 Einleitung

In einigen Anwendungen werden an die leittechnischen Einrichtungen zum Führen von Fertigungs- oder verfahrenstechnischen Prozessen – insgesamt oder auch nur bei Teilprozessen – besonders hohe Sicherheits- oder Verfügbarkeitsanforderungen erhoben. Der vorliegende Beitrag stellt, mit dem Hauptgewicht auf dem Verfügbarkeitsaspekt, einige Redundanztechniken zur Erfüllung dieser besonderen Anforderungen an ein Prozeßleitsystem dar (ausführlicher u.a. in [13]).

2 Prozeßleitsysteme

Prozeßleitsystem – abgekürzt PLS – ist der Überbegriff für die leittechnischen Einrichtungen zum Regeln, Steuern, Bedienen und Beobachten eines Fertigungs- oder verfahrenstechnischen Prozesses, die mit einem einheitlichen und durchgängigen Systemansatz arbeiten. Insbesondere bei größeren Anlagen und Prozessen, wie beispielsweise einer chemischen Produktionsanlage oder einem Kraftwerk, ergibt sich ganz natürlich eine Unterteilung in Teilanlagen und Teilprozesse sowie eine Hierarchisierung der Leitebenen.
Diese Unterteilung und Hierarchisierung spiegelt sich auch in der Leittechnik, d.h. dem Prozeßleitsystem. Die Bezeichnungen der Ebenen sind je nach Hersteller, System und Branche unterschiedlich, aber das Prinzip ist weitgehend einheitlich. Für das PLS ergibt sich eine Baumstruktur bzw. eine Pyramide, in der Verarbeitungs- und Kommunikationsebenen einander abwechseln, vgl. **Bild 1**. In der Spitze der Pyramide liegen die prozeßfernen und zur Basis hin die prozeßnahen Komponenten. In den Kommunikationsebenen liegen jeweils Busse, die mehrere darunterliegende Verarbeitungseinheiten und eine darüberliegende Verarbeitungseinheit verbinden. In ausgeführten Systemen kann es durchaus auch Abweichungen von der „idealen" Pyramiden- oder Baumeinteilung geben, indem z.B. mehreren Ebenen (physikalisch) derselbe Bus unterlagert ist. Bei den meisten Prozeßleitsystemen ist es möglich, mehrere – dann nur logisch unterscheidbare – Ebenen in einem Verarbeitungsgerät zu vereinigen z.B. für Kleinanlagen. Von einer Pyramide ist in der Hardwarestruktur dann nur noch wenig zu sehen.
Die Prozeßsteuerung für eine Anlage kann man sich natürlich auch aus unterschiedlichsten Einzelkomponenten zusammensetzen, wie verschiedenen Rechnern, speicherprogrammierbaren Steuerungen (SPSen), PCs. Ein Vorteil eines solchen Vorgehens wäre, daß jeweils maßgeschneiderten Lösungen für alle Teilbereiche weitgehende Optimierungsmöglichkeiten eröffnen. Der Hauptnachteil eines derart uneinheitlichen Ansatzes ist die aufwendige Projektierung und Programmierung. Insbesondere bei Änderungen während der Inbetriebnahme kann es schwierig werden, zu einem konsistenten lauffähigen System zu kommen. Man lebt bei den meisten Projekten auf die Dauer besser, wenn man ein Prozeßleitsystem (Betonung liegt auf System) einsetzt, das in Hardware und Software bereits gewisse durchgängige Merkmale und Systemdienste bietet – vgl. **Bild 2**. Dies sind u.a. eine systemweit einheitliche Uhrzeit, zeitliche Auflösung von Ereignissen und Meldun-

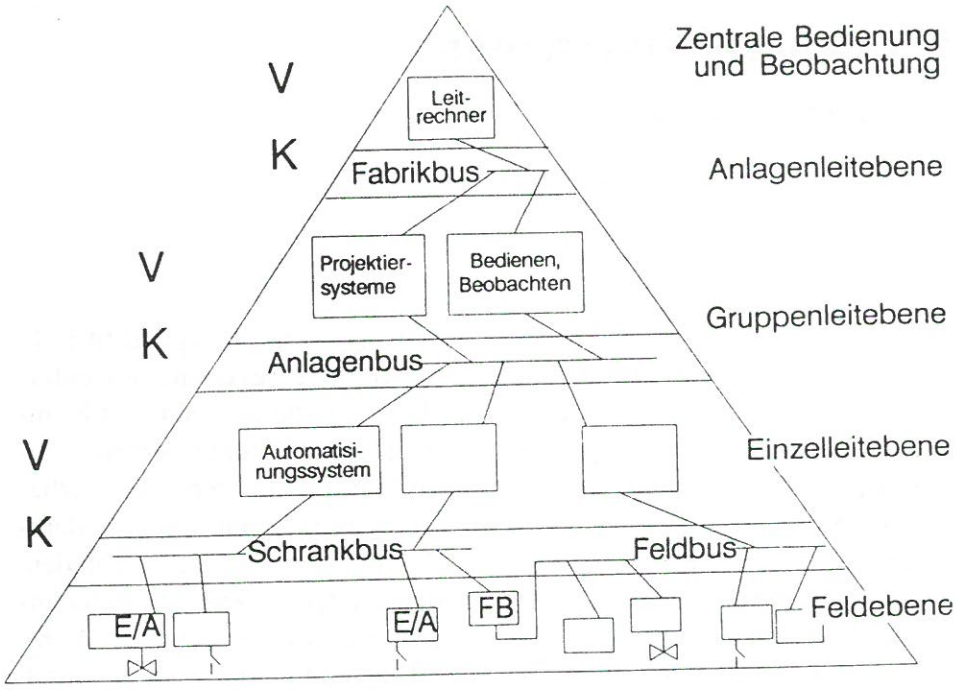


Bild 1 Die Struktur eines Prozeßleitsystems (PLS).

gen, überhaupt ein Leittechnikmeldesystem, systemweite Namen, leistungsfähige Projektierungshilfen, implizite Projektierung der Kommunikation u.a.m. sowie – und hierauf wollen wir uns im folgenden konzentrieren – durchgängige geeignete Redundierungsmöglichkeiten.

- Offenheit
- Objektorientierung
- Großer Fundus an Standardbausteinen
- Graphische Projektierung
- Meldekonzept
- Diagnosefunktionen
- Einbeziehung von Standards
- Bausteintechnik
- Branchenpakete
- Systemweite Zeit
- Zyklische Bearbeitung (garantierte Frequenz)
- Leittechniküberwachung
- Redundanz

Bild 2 Durchgängige Merkmale eines Prozeßleitsystems

3 Redundanz

In vielen Anwendungen sind die Folgen eines Versagens einer in einfacher Weise rein funktions- oder aufgabenbezogen eingesetzten Prozeßleittechnik unkritisch. Für Anwendungen aber, bei denen die sicherheitskritische oder verfügbarkeitsbestimmende Rolle einer so eingesetzten Leittechnik oder von Teilen davon nicht hingenommen werden kann, sind Architekturänderungen erforderlich, die letztlich auf den Einsatz von Redundanz hinauslaufen. Redundanz ist der – geeignete – Einsatz von mehr Mitteln, als zum Erfüllen der Funktion erforderlich wäre. Mehr Mittel ist in dem Sinne gemeint, daß das nicht-redundierte System für sich allein die geforderten Funktionen voll erbringen kann – solange es in Ordnung ist. Das mit dem Einsatz dieser zusätzlichen Mittel zu lösende Problem ist es, auch im Falle von Fehlern und Defekten ein gewünschtes Verhalten darzustellen. Das ist mit dem Wort „geeignet" gemeint. (Ungeeignete Redundanz wäre Verschwendung.)

Ein (in diesem Sinne) erhöhter Aufwand an Hardware oder Software wird bei Prozeßleitsystemen mit zwei Zielrichtungen eingesetzt, nämlich

- dem sicherheitsgerichteten oder "**F**ail-safe"-Verhalten ("F")
- und/oder der er**H**öhten Verfügbarkeit ("H"),

die im folgenden kurz definiert werden:

F : Beim sicherheitsgerichteten oder F-Verhalten darf ein Fehler in der Leittechnik, also in den genannten Verarbeitungs- und Kommunikationseinrichtungen (wie der Ausfall einer Baugruppe, Fehlfunktion eines Prozessors o.ä.) nicht zu einer Fehlansteuerung des Prozesses führen. Durch geeignete Maßnahmen ist bei Fehlern das Überführen des Prozesses in einen sicheren Zustand zu garantieren. Die minimale und typischerweise angewandte Redundanzstruktur ist 2-von-2. Die im nicht-redundanten Falle einzusetzende Verarbeitungseinheit wird zweifach eingesetzt, und die Prozeßführung wird nur aufrechterhalten, wenn beide (also zwei von zweien) arbeiten. Die doppelte Verarbeitung gestattet das Erkennen jedes Einzelfehlers durch Vergleiche. Und bei Erkennen eines Fehlers wird der Prozeß sicherheitsgerichtet stillgesetzt.

H : Beim verfügbarkeitsgerichteten Verhalten sollen die Prozeßführungsaufgaben trotz eines Fehlers aufrechterhalten werden. Die minimale und häufig verwendete Redundanzstruktur ist 1-von-2. Hier werden ebenfalls verdoppelte Verarbeitungseinheiten eingesetzt. Fehler können durch Vergleiche erkannt werden. Im Falle eines Fehlers führt die fehlerfreie Einheit die Aufgabe stoßfrei weiter. Das Gesamtsystem fällt nur aus, wenn auch diese Einheit während der Reparaturzeit einen Fehler erleidet. Die Wahrscheinlichkeit hierfür kann durch die Eigenschaft der sog. On-line-Reparierbarkeit und eine mittlere Reparaturzeit, die um Größenordungen unter der mittleren Ausfallzeit der Einheiten liegt, extrem klein gemacht werden.

Kombination HF : Für bestimmte Anwendungen werden auch Systeme eingesetzt, die sowohl H- als auch F-Eigenschaften haben. Die minimale Redundanzstruktur hierfür ist 2-von-3. Ein Teilsystemausfall wird verfügbarkeitsgerichtet toleriert, und man ist dann während der Reparaturzeit dieses Teilsystems auf ein sicherheitsgerichtetes 2-von-2-System reduziert (Beispielsystem siehe [4]).

Neben den gezeigten Zielen Verfügbarkeit und/oder Sicherheit – wir werden uns im folgenden auf die Verfügbarkeit konzentrieren – sind für den PLS-Anwender noch weitere Merkmale der eingesetzten Redundanzlösungen wichtig, die man oft unter dem Sammelbegriff Redundanzqualität zusammenfaßt:

- **1** Hoher Fehlererkennungsgrad.
- **2** Kleine Fehlerlatenzzeit.
- **3** Einteilung des Gesamtsystems in mehrere fehlerausbreitungsmäßiggutabgeschottete Fehlerabgrenzungsregionen.
- **4** Geringe oder keine Leistungsminderung bei tolerierten Fehlern.
- **5** On-line-Reparierbarkeit, d.h. Reparieren bei weiterlaufendem Prozeß.
- **6** Transparenz der Redundanz.
- **7** Durchgängigkeit der Redundierbarkeit.

Die ersten beiden Punkte sind wesentlich, da Fehlererkennung die Grundlage jeder Fehlertoleranz ist. Fehlerausbreitungsgrenzen und Fehlerabgrenzungsregionen müssen die Auswirkung von Fehlern lokal auf ein Teilsystem begrenzen. Ein Fehler darf sich nicht auf das Gesamtsystem und vor allem nicht auf das redundante Teilsystem ausbreiten; dies ist die Aussage des dritten Punktes. Mit dem Punkt 5, der Transparenz, ist gemeint, daß der Anwender und Projektierer wenig bis nichts von der Redundierung einer Komponente merken soll. Eine redundierte Komponente verhält sich an ihren Schnittstellen idealerweise genauso wie die nicht-redundierte Standardkomponente. Da die nicht-redundierten Glieder einer Verarbeitungskette deren Verfügbarkeit praktisch allein bestimmen (vgl. Beispiel in Kapitel 4, unten), lassen sich hohe Ge-

samtsystemverfügbarkeitsanforderungen nur bei durchgängiger Redundierung (Punkt 6) erfüllen. Dieser Aufwand wird i.a. dann getrieben, wenn die Kosten eines mehrstündigen Produktionsausfalls den Mehrkosten in der Leittechnik nahekommen, wie u.a. bei Großkraftwerksblöcken.

4 Verfügbarkeit

Die Verfügbarkeit ist die Eigenschaft eines Systems, seine Aufgabe zu erfüllen. I.a. wird sie als eine Wahrscheinlichkeit angegeben, die sich (unter der bei elektronischen Komponenten i.a. gegebenen Voraussetzung konstanter Ausfallraten) wie folgt berechnen läßt:

$$V = \frac{MTBF}{MTBF + MDT} \qquad (1)$$

Dabei bedeutet MTBF (mean time between failures) den Mittelwert der Betriebsdauer zwischen der (Wieder-) Inbetriebnahme eines intakten (reparierten) Systems und dem Systemversagen im Sinne von Nichtmehrerfüllen seiner Aufgabe. Eine MTBF einer Baugruppe von 5 Jahren bedeutet, daß von vielen solcher Baugruppen nach 5 Jahren über 60% ausgefallen sein werden, und nicht etwa eine garantierte Lebensdauer von 5 Jahren. Daß oft die zweite Annahme zuzutreffen scheint, liegt an den i.a. konservativen MTBF-Angaben seriöser Hersteller.
MDT ist die mittlere Zeitdauer zwischen einem Fehler, der zu einem Systemversagen führt, und der Wiederinbetriebnahme des Systems. Diese Zeit umfaßt u.a. Fehlerlokalisierung, Reparatur und Wiederanfahren des Systems. Die Mittelwerte sind über ein großes Ensemble gleicher Systeme oder viele Fehler eines Systems zu nehmen. Der Ausdruck (1) zeigt auch, daß die seltener betrachtete MDT mindestens so verfügbarkeitsbestimmend ist wie die vielbeachtete MTBF.

Die Verfügbarkeit eines n-von-m-Systems läßt sich bei gegebenen Verfügbarkeiten der Teilsysteme mit den Methoden der Wahrscheinlichkeitsrechnung ermitteln. Für ein 1-von-2-System aus zwei gleichen Teilsystemen ergibt sich:

$$V_{1v2} = 2 * V - V^2 \qquad (2)$$

und

$$MTBF_{1v2} = 1/2 * MTBF^2 / MDT \qquad (3)$$

Besteht ein System oder eine Verarbeitungskette aus 6 Einheiten, von denen jede eine mittlere Ausfallzeit MTBF von 2 Jahren habe, so ergibt sich für das nicht-redundierte Gesamtsystem eine mittlere Ausfallzeit von 4 Monaten. Bei einer durchgängigen Redundierung und einer mittleren Reparaturzeit MDT von 72 Stunden, ergibt sich für die einzelne redundierte Einheit eine MTBF von 242 Jahren und für das Gesamtsystem von 40 Jahren, wobei eine Reduzierung der MDT noch drastische Verbesserungen bringt. Werden nur 5 der 6 Einheiten redundiert, so verbessert sich die Gesamt-MTBF lediglich auf etwa 2 Jahre. Dieses Beispiel zeigt (trotz seiner bewußten Vereinfachung) eindrucksvoll die mit keinen anderen Maßnahmen als Redundanz und online-Reparierbarkeit möglichen Verfügbarkeitsverbesserungen und auch die Bedeutung einer durchgängigen Redundierungsmöglichkeit.

5 Redundierung der prozeßfernen Komponenten

Die Strenge der Redundanzanforderungen ist je nach Leittechnikebene durchaus unterschiedlich, und so unterscheiden sich auch die Lösungsprinzipien bei den verschiedenen Verarbeitungs- und Kommunikationseinheiten. So sind bei den prozeßfernen Komponenten, wie beispielsweise bei Bedien- und Pro-

jektierungsplätzen, die Echtzeitanforderungen häufig nicht so streng, daß nicht auch Ersatzgerätestrategien und eine Totzeit beim Umsetzen des Bedieners an einen überzähligen Bedienplatz tragbar wären. Allerdings dürfen durch einen Defekt keine Prozeß- oder Projektierungsdaten verloren gehen. Geeignete Prinzipien der redundanten Datenhaltung müssen hier angeboten werden, wie Spiegelplatten, Diskarrays o.ä.

Bei den Kommunikationswegen, den sogenannten Bussen, sind die Redundanzanforderung schon strenger: Ein Fehler muß stoßfrei toleriert werden. Ferner soll i.a. die Redundanz für den Anwender transparent sein, d.h. er muß dafür nichts programmieren oder projektieren. Außerdem muß man hier Redundanz gegen Leitungsfehler und gegen Anschaltungsfehler unterscheiden. Eine Leitung kann durch Kurzschlüsse, Unterbrechungen und durch Störungen eines angeschlossenen Teilnehmers (ein Teil der Teilnehmerelektronik ist fehlermäßig als zur Leitung gehörig zu rechnen!) beeinträchtigt werden. Will man trotz solcher Beeinträchtigungen die Kommunikation aufrechterhalten, so sind die Leitungswege zu redundieren, d.h. i.a. zu verdoppeln.

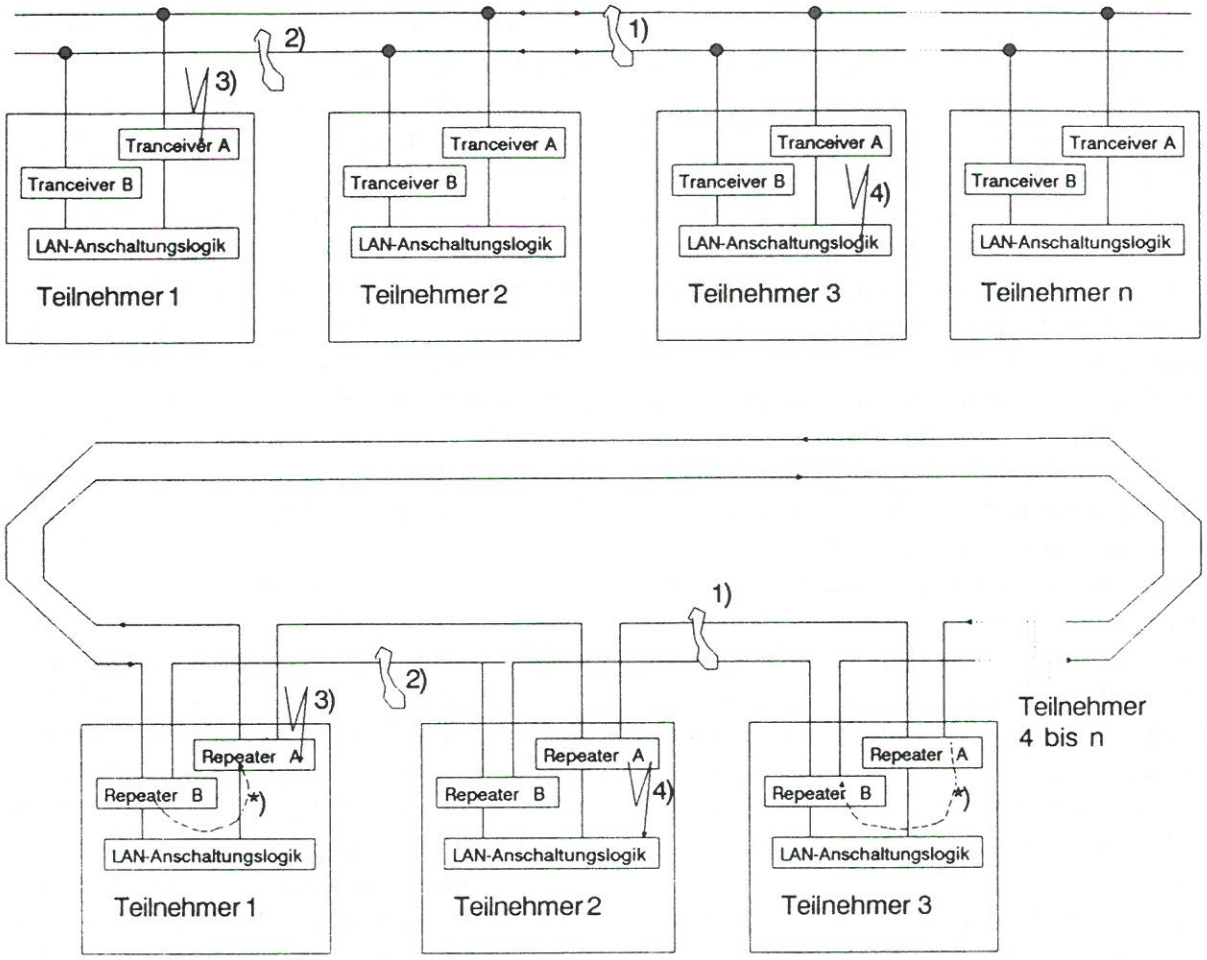
Während es das Ziel der Leitungsredundanz ist, die Kommunikation aller Busteilnehmer untereinander trotz Anwesenheit von oben angeführten Leitungsfehlern aufrechtzuerhalten, ist es das Ziel der Anschaltungsredundanz, die Kommunikation eines redundanten Teilnehmers (mit allen anderen) trotz eines Fehlers seiner Busanschaltung aufrechtzuerhalten. Hierzu erhalten redundante Systeme auch redundante Busanschaltungen. Der Einsatz von Anschaltungsredundanz ist vom Einsatz der Leitungsredundanz unabhängig.

Bei den Ausführungsdetails von Leitungs- und Anschaltungsredundanz sind noch Bus- und Ringstrukturen zu unterscheiden. Bei einer Busstruktur hängen sozusagen alle Teilnehmer an dem selben Draht, während bei einer Ringstruktur die Teilnehmeranschaltungen auch noch als sog. Repeater in einer bei jedem Teilnehmer unterbrochenen (i.a. Lichtwellen-) Leitung fungieren. Einfache, in **Bild 3** angedeutete Überlegungen zeigen, daß die Auswirkungen von bestimmten anzunehmenden Leitungs- und Anschaltungsfehlern bei Bus- und Ringstrukturen durchaus unterschiedlich sind und demzufolge auch die einzusetzenden Redundanzmaßnahmen.

Der in Bild 3 dargestellte Fehler 1, die Unterbrechung beider redundanter Leitungen, unterbricht beim Bus die Kommunikation; es gehen einige Teilnehmer verloren. Die Unterbrechung einer Leitung, der Fehler 2, wird durch die Leitungsredundanz toleriert. Der Ring kann wegen der Ringtopologie auch den Fehler 1 tolerieren, falls die Kommunikation in beide Richungen laufen kann. Sinn eines Lichtwellenleiter- (LWL-) Doppelrings ist die Realisierung beider Richtungen; die einer Unterbrechung benachbarten Stationen gehen in den sog. "Loopback"-Modus. Der Fehler 3, die Störung eines redundanten Tranceivers wirkt bei Bus und Ring wie der Ausfall einer Leitung und wird genauso toleriert. Der Ausfall oder das Ausschalten eines oder mehrerer Teilnehmer, Fehler 4, wird beim Bus ohne weiteres toleriert. Nicht tolerierbar ist hingegen eine Störung im nicht-redundanten Teil der Busanschaltung, die auf beide Leitungen wirkt; dies muß durch entsprechende Redundanzmaßnahmen in der Busanschaltung ausgeschlossen werden. Das Ausschalten einer Station, Fehler 4, wird beim Ring genau wie ein Fehler 1 durch Loopback der Nachbarstationen toleriert ebenso wie das Fehlverhalten einer Station (Vorteil Ring). Hingegen kann das Ausschalten mehrerer Stationen ohne Zusatzmaßnahmen nicht toleriert werden (Nachteil Ring).

6 Redundierung der prozeßnahen Komponenten

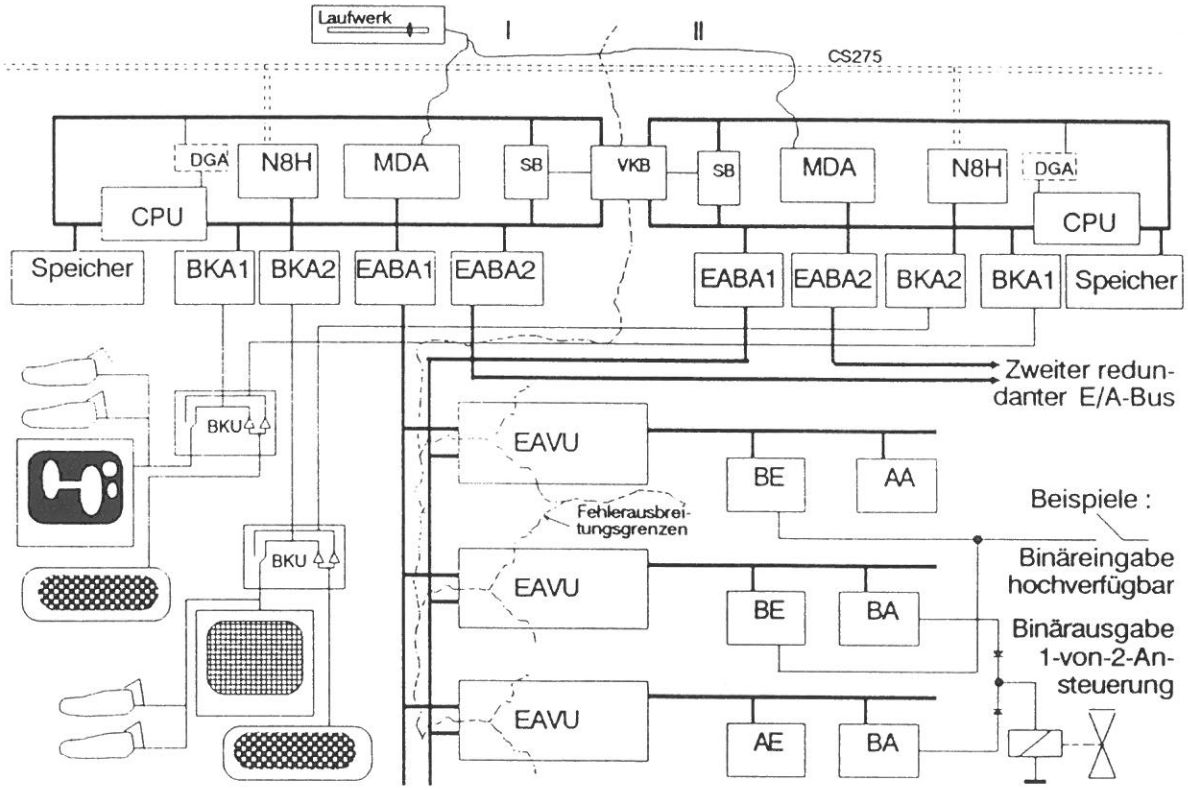
Bei den prozeßnahen Komponenten lassen sich die strengen Echtzeitanforderungen nach einem stoßfreien und weitgehend „totzeitlo-



*) "Loopback" der Teilnehmer 1 und 3 als Reaktion auf den beispielhaften Fehler 4).

Bild 3 Vergleich der Fehlerauswirkungen bei Bus- und Ringstruktur.

sen" Weiterarbeiten nur mit der synchronen Bearbeitung der Aufgabe in zueinander redundierten Teilsystemen erfüllen. Eine besonders attraktive Lösungsmöglichkeit hierfür ist die taktsynchrone Arbeitsweise, wie sie u.a. beim Zentralteil des Automatisierungssystems TELEPERM M AS 235 H realisiert wurde, vgl. **Bild 4** sowie [4], [11] und [12].
Die verdoppelte Standard-Hardware bearbeitet stets taktsynchron die gleiche Software. Die Taktgeber sind gegenseitig synchronisiert, aber voneinander unabhängig. Das Erkennen von Fehlern, sowie alle Synchronisierungsaufgaben übernimmt die einzige spezielle (Nicht-Standard-) Hardware in dem modularen System. Die Anwender- und weitestgehend auch die Systemsoftware „sieht" eine logische und mit dem entsprechenden einkanaligen Standardsystem identische Maschine.
Die zusätzlichen Synchronisierungs-, Vergleicher- und andere Schaltungen sind selbst redundant und vom System überprüfbar ausgeführt und bezüglich Fehlerausbreitung rückwirkungsfrei gekoppelt. Die Erfüllung der redundanzspezifischen Aufgaben kostet bei solchen hardwareorientierten Lösungen im Gegensatz zu Softwarelösungen keine zusätzliche Laufzeit, was auch zu der wünschenswerten Kompatibilität zum entsprechenden nichtredundanten Standardsystem beiträgt.
Die Redundanz im Prozeß-E/A-Teil ist bei den beispielhaft aufgeführten Systemen in



CS275 : LAN mit Leitungs- und Anschaltungsredundanz, N8H: redundierbare LAN-Anschaltung, DGA : Diagnosegeräteanschaltung, MDA : Minidiskettenlaufwerksanschaltung, SB, VKB : Synchronisier-, Vergleicher-/Koppelbaugruppe, BKA/BKU : Bedienkanalan-/umschaltung, EABA : E/A-Busanschaltung, EAVU : E/A-Bus-Vergleicher und -umschaltbaugruoppe, BE, BA, AE, AA : Binär/ Analog - E/A

Bild 4 Die Zentralteilstruktur des taktsynchronen 1-von-2-H-Systems TELEPERM M AS 235 H.

weiten Grenzen vom Anwender frei projektierbar.

7 Zusammenfassung

Neben anderen Systemanforderungen muß ein Prozeßleitsystem auch die Möglichkeit einer Redundierung auf allen Ebenen bieten. Die eingesetzten Lösungswege unterscheiden sich entsprechend den Redundanzqualitäts- und Echtzeitanforderungen der verschiedenen Anwendungen und Verarbeitungsebenen. Auch wenn die Grundprinzipien einfach erscheinen, stellt die Verwirklichung eines durchgängigen und für den Anwender transparenten Redundanzkonzepts eine technisch anspruchsvolle Aufgabe dar.

8 Literatur

[1] Delft University of Technology: Delft Progress Report, Functional Safety of Programmable Electronic Systems (PES), Workshop April 7, 1986, Delft (NL) 1986

[2] DIN V 19250, Messen-Steuern-Regeln: Grundlegende Betrachtungen für MSR-Schutzeinrichtungen, Beuth-Verlag, Januar 1989

[3] DIN VDE 31000, Teil 2, Allgemeine Leitsätze für das sicherheitsgerechte Gestalten technischer Erzeugnisse; Grundbegriffe, Beuth-Verlag, 1987

[4] Euringer,M; Reichert,W: Hochverfügbares und fehlersicheres Automatisierungssystem AS 220 EHF in 2-von-3-Technik, Siemens-Energietechnik 6 (1984), Heft 5 , September/Oktober 1984, S. 245 bis 249, Karlsruhe 1984

[5] Knight,J.C.; Leveson,N.G.: Correlated Failures in Multi-Version Software, S. 159-165 in [8]

[6] - , A Reply to the Criticisms of the Knight & Leveson Experiment, S. 24-35 in ACM SigSoft, Software Engineering Notes, vol 15 no 1, Jan 1990

[7] NTG: Architektur und Betrieb von Rechensystemen, Vorträge der NTG/GI-Fachtagung vom 10. bis 12. März 1986 in Stuttgart, NTG-Fachberichte Nr. 92, Berlin 1986

[8] Quirk,W.J. (ed.): Safety of Computer Control Systems 1985 (SAFECOMP'85), Achieving Safe Real Time Computer Systems, Proceedings of the Fourth IFAC Workshop, Como 1.-3. Oct. 1985, Oxford 1985

[9] Redmill,F.J. (ed.): Dependability of Critical Computer Systems - 1, Guidelines produced by the European Workshop on Industrial Computer Systems, Technical Committee 7 (EWICS TC7), Elsevier Applied Science Publishers, London 1988

[10]- (ed.): Dependability of Critical Computer Systems - 2, Guidelines produced by the European Workshop on Industrial Computer Systems, Technical Committee 7 (EWICS TC7), Elsevier Applied Science Publishers, London 1989

[11] Weinert,A.: Über hardwareimplementierbare Fehlertoleranz bei industriellen Automatisierungssystemen mit sehr hochintegrierten Prozessoren, S. 268-278 in [7]

[12]- : Hardware Implemented Fault-Tolerance and Safety for Programmable Automation Systems, S. 255-265 in [1]

[13]- : Automatisierung kritischer Prozesse, atp-Seminar (mehrere Fortsetzungen, gelbe Sonderseiten Seitennummern beginnen mit S), atp-Automatisierungstechnische Praxis 33 (1991), H.6 S1-S4, H.7 S5-S8, H.8 S9-S12, H.9 S13-S15