Dissertation

Bestimmung der betriebswirtschaftlichen Effektivität von Informationssicherheitsmaßnahmen

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines Doktors der Sozial- und Wirtschaftswissenschaften unter der Leitung von

Univ.-Prof. Mag. Dr. Walter S.A. Schwaiger, MBA

E330
Institut für Managementwissenschaften
Bereich Finanzwirtschaft und Controlling

eingereicht an der Technischen Universität Wien
Fakultät für Informatik

von

Mag. Sascha Hönigsberger
Matrikelnummer 9603231
Heinrich-Collinstraße 3/8/20
A- 1140 Wien

Wien, im Mai 2009

Kurzfassung

In dieser Arbeit geht es darum ein Rahmenmodell zu erarbeiten, das es ermöglicht, alle Arten von Informationssicherheitsprozessen und -technologien in organisationsweite Risikomanagementmethoden und Risikomanagementprozesse zu integrieren. Dadurch soll Organisationen bzw. Unternehmen eine bessere Steuerung der Investitionen in diese Informationssicherheitsmaßnahmen ermöglicht werden.

Bei der Erstellung des Modells war es wichtig, die bereits existierenden Standards wie etwa die ISO 2700x Normenreihe, das Risikomanagementframework COSO2 und das IT-Governance Regelwerk COBIT zu berücksichtigen, um eine möglichst breite praktische Anwendbarkeit sicherzustellen.

Der Ausgangspunkt der Arbeit ist die Betrachtung einer Organisation bzw. eines Unternehmens als soziotechnisches System. Darauf aufbauend wird zuerst das System und dessen Kontext im Rahmen einer kurzen Systemumfeldanalyse näher betrachtet, und so die Basis für den Modellrahmen gelegt. Basierend auf diesem groben Modellrahmen wird ein Metamodell aufgespannt, in welchem das System, zur besseren Darstellung der Intrasystemabhängigkeiten, selbst in eine Prozessdomäne und in eine Infrastrukturdomäne aufgeteilt wird.
Das Metamodell beinhaltet bereits jene Instrumentarien, welche es ermöglichen den Weg, den ein Schadenereignis von seinem Eintritt in das System bis zu seinem Austritt aus dem System nehmen kann, darzustellen. Weiters können mit Hilfe des Metamodells bereits auch die jeweiligen Schadwirkungen näher analysiert werden. Zur Berechnung des mit einem Ereignis verbundenen Risikos wird die Methode des Probabilistic Risk Assessment (PRA) verwendet. Im erarbeiten Modell werden die für das Probabilistic Risk Assessment benötigten Schadenpotentiale aus der Prozessdomäne, und die Eintrittswahrscheinlichkeiten der jeweiligen Ereignisse aus der Infrastrukturdomäne abgeleitet. Die Schnittstellen zwischen diesen beiden Domänen bilden dabei die IT-Dienste, die den Prozessen von der Infrastruktur zur Verfügung gestellt werden.
Parallel zu der schrittweisen Vertiefung des Metamodells, die von den in den COSO2 Aktivitäten *Objective Setting*, *Event Identification*, *Risk Assessment* und *Risk Response* definierten Aufgaben eines Risikomanagementsystems geleitet wird, wird in der Arbeit auch immer eine einfache Ausprägung einer Implementierung des Metamodells mitgeführt, um die integrierten Konzepte besser darstellen zu können.

Nach dem vollständigen Aufspannen des Modellbogens ist es gegen Ende der Arbeit endlich möglich näher auf das eigentliche Ziel der Arbeit einzugehen. Es wird dort gezeigt, wie sich Informationssicherheitsmaßnahmen in das Modell integrieren lassen, und wie deren Wirkung sowohl auf die Eintrittswahrscheinlichkeit eines Ereignisses als auch auf die Schadenpotentiale eines Ereignisses bei der Berechnung des Risikos des Ereignisses berücksichtigt werden kann.

Nach der detaillierten Darstellung des Modells wird am Ende der Arbeit noch anhand eines praktischen Beispiels, in dem für einen IT-Dienst, wie etwa einen Webserver, eine Gegenüberstellung des inhärenten Risikos ohne implementierte Sicherheitsmaßnahmen für einen auf diesen Dienst angewiesenen Geschäftsprozess dem residualen Risiko nach Implementierung entsprechender Maßnahmen durchgeführt. Das Beispiel bedient sich dabei einer Monte Carlo Simulation des Dienstes und der von ihm benötigten Infrastruktur.

Vorwort

Diese Arbeit versucht sich an dem manchmal monströs erscheinenden Spagat, die tiefen Detailebenen der Softwarefehler, die daraus resultierenden Angriffspunkte von Softwaresystemen und die abstrakten Höhen der Unternehmenssteuerung auf eine Weise zu verbinden, die es ermöglicht ein in der Praxis anwendbares ganzheitliches organisationsweites Risikomanagement zu etablieren.

Stand am Beginn meiner Arbeit noch der schlichte Wunsch einfach nur einen Weg zu finden, wie der Einsatz von aufwändigen Informationssicherheitsinstrumenten abseits von Firewalls oder eines Virenschutzprogrammen besser gerechtfertigt werden kann, so bekam ich recht schnell eine Vorstellung davon, wie komplex und detailbehaftet dieses Vorhaben eigentlich ist.

Durch die Arbeit an dem nun vorliegenden Modell entwickelte sich bei mir ein tiefer Respekt all jenen gegenüber, die sich schon vor mir an einer theoretischen Lösung meines schlichten Wunsches versuchten, und ich bin dankbar, dass ich auf eine Fülle theoretischer Konzepte aus anderen Disziplinen zurückgreifen konnte, ohne die ich wohl keinen Weg durch den Informationssicherheitsdschungel gefunden hätte.
In diesem Zusammenhang möchte ich mich vor allem bei meinem Betreuer Prof. Walter Schwaiger bedanken. Ohne ihn wäre ich auf viele dieser theoretischen Konzepte nicht aufmerksam geworden und hätte sie auch bis heute wohl noch nicht verstanden. Ebenfalls bedanken möchte ich mich bei meinen Kollegen im Dissertantenseminar, deren kritisches Feedback während vieler Zwischenpräsentationen dazu beitrug das Modell von unzähligen Ecken und Kanten zu befreien.

Es gibt einen Spruch, nachdem der Weg das Ziel einer Reise ist. Auf diese Arbeit trifft er definitiv zu. Im Zuge der Erstellung des Modells hat sich mit der Lösung jedes Problems meine Sicht auf die Zusammenhänge innerhalb von Organisationen und deren Interaktionen mit ihrer Umwelt erweitert, sodass sich die investierte Zeit alleine vor diesem Hintergrund mehr als rentiert hat. In diesem Sinne hoffe ich, dass diese Arbeit für den Leser ebenso lohnend ist.

Inhaltsverzeichnis

1 Einleitung....2
2 Basiskonzepte....6
2.1 COSO2....6
2.2 Das Unternehmen als hierarchisches Multiebenen System....12
2.3 ANSI ISA-95 / Purdue Referenzmodell....17
2.4 ISO 27002....23
2.5 Probabilistic Risk Assessment....29
2.6 Das ISO OSI Modell....35
3 Ansatzpunkte....36
3.1 Zieldefinition / COSO2 Aktivität 2....36
3.2 Systemumwelt....38
3.3 Risikomodellierung....40
4 Integration in den Systemkontext....44
4.1 Relevante Teile von ISA 95....46
4.2 Basis der Identifizierung relevanter Ereignisse / COSO2 Aktivität 3....53
4.3 Ereignisidentifikation in der Prozessdomäne....55
4.4 Modellierung der Infrastrukturdomäne....58
4.5 Infrastrukturinteraktion exogener Ereignisse....65
5 Details der Risikomodellierung....68
5.1 Allgemeines....68
5.2 Details zur Schadenfunktion / Prozessdomäne....70
5.3 Details zum Zählprozess und zur Ereignisdauer der Infrastrukturdomäne....87
6 Beispiel zur Modellumsetzung....101
6.1 Das inhärente Risiko....103
6.2 Der Risk Response....104
6.3 Das residuale Risiko....105
6.4 Inhärentes vs. residuales Risiko....106
7 Zusammenfassung und Ausblick....109
7.1 Ausblick....112
Anhang A: Literaturverzeichnis....113
Anhang B: Abbildungsverzeichnis....115

1 Einleitung

Die Gefährdung der Sicherheit von Informationen und der diese verarbeitenden Systeme ist nach der Verschmelzung von Millionen von privaten Computern und von Tausenden von Firmencomputernetzen zu einem großen Netzwerk, dem Internet, mittlerweile nahezu jede Woche für eine Schlagzeile in den Print- bzw. Onlinemedien verantwortlich. Dies ist auch nicht verwunderlich; seit mehr als zwei Jahrzehnten treiben Viren ihr Unwesen auf den Festplatten der Welt, und Hacker entziehen sich mit ihren Fähigkeiten seit jeher dem Verständnis des gewöhnlichen Computerbenutzers. Zu diesen schon altbekannten Plagen gesellen sich seit einiger Zeit nun auch noch Spammer, die das virtuelle Postfach der Internetnutzer ständig mit unerwüschten Mails füllen, und Phisher, die sich nicht mit legalen Einkäufen begnügen sondern gleich über die Kreditkarten bzw. die Onlinebankingaccounts der Benutzer verfügen wollen.

Neben der medialen Präsenz haben wohl die meisten Organisationen in den letzten Jahren mehr oder weniger direkt die Auswirkungen der Bedrohungen für Informationen bzw. deren Sicherheit wahrgenommen. Dies ist wahrscheinlich mit ein Grund dafür dass es mittlerweile in allen größeren Unternehmen eigene Beauftragte oder sogar ganze Abteilungen gibt, die sich mit Informationssicherheit bzw. der Beseitigung der Informationsunsicherheit beschäftigen. Informationssicherheit ist dabei durch die recht allgemeine Formulierung der Ziele, die es zu erreichen gilt[1], ein sehr breites Betätigungsfeld, und die jeweilige Schwerpunktsetzung ist überdies noch abhängig vom Umfeld bzw. Geschäftsfeld, in dem eine Organisation tätig ist. Trotz dieser Heterogenität des Tätigkeitsfeldes gibt es dennoch ein Grundproblem, mit dem sich praktisch jeder Informationssicherheitsbeauftragte konfrontiert sieht, es ist dies die Frage wie das nötige bzw. adäquate Budget gesichert und argumentiert werden kann. Im Grunde ist dieses Budgetproblem nichts anderes als die Frage, welchem Risiko die jeweilige Organisation durch die Unsicherheit der Informationen ausgesetzt ist, und welchem sie ausgesetzt sein will.

Eine adäquate Erhebung des mit Informationssicherheitsereignissen verbundenen Risikos wird auch in den in diesem Bereich wichtigsten Standards, dem [ISO 27002] und [COBIT] (siehe auch Kapitel 2.4) thematisiert. Allerdings ist in diesen Werken und auch im [ISO 27001] keine Methodik definiert, wie exakt diese Risiken identifiziert oder bewertet / abgeleitet werden können.
Diese beiden Standards sind in diesem Zusammenhang „nur“ Sammlungen von Best Practices. Sie bilden damit aber eine gute Referenz- bzw. Vergleichsbasis für Audits, also zur Identifikation bzw. Bewertung von Informationssicherheitsrisiken.
Ein erster Leitfaden in Punkto Risikomanagement und Risikoerhebung im speziellen ist nun im [ISO 27005] enthalten, der im Juni 2008 veröffentlicht wurde. Bedingt durch den Umstand, dass bei der Erstellung eines Standards immer eine Vielzahl von Interessen berücksichtigt werden müssen und dieser auch eine möglichst breite Anwendbarkeit besitzen soll, wurde auch beim ISO 27005 nicht definiert wie die Risikobewertung exakt durchzuführen ist. Es wurde vielmehr auch hier der Fokus darauf gelegt die wichtigsten Komponenten eines Risikomanagementsystems für Informationssicherheitsereignisse darzustellen und zu zeigen wie diese Komponenten an einander gekoppelt werden müssen.

1 Sicherung der Vertraulichkeit, der Integrität und der Verfügbarkeit der Informationen

Zur Beantwortung der obigen Fragestellung nach dem Risiko gibt es nun zwei grundlegende methodische Ansätze: eine quantitative Berechnung des Risikos oder eine qualitative Bewertung des Risikos.[2]
Beide Ansätze haben ihre Wurzeln im sog. Probabalistic Risk Assessment (siehe auch Kapitel 2.5), in dem das Risiko, das mit einem Ereignis verbunden ist, über die Eintrittswahrscheinlichkeit und die potentiellen Auswirkungen des Ereignisses dargestellt wird. Der Unterschied zwischen den beiden Methoden liegt darin, daß das Ziel der quantitativen Methode die Berechnung von Zahlenwerten für die Eintrittswahrscheinlichkeit und die in den meisten Fällen finanziellen, Auswirkungen ist, während die qualitative Bewertung hingegen das Ziel verfolgt ein Risiko in beiden Dimensionen einer Kategorisierung zuzuführen.[3]

Zur quantitativen Berechnung des Risikos steht Informationssicherheitsbeauftragten derzeit nur eine weitgehend akzeptierte Methode, die sog. Annual Loss Expectancy (ALE), zur Verfügung, die aus dem Produkt aus erwartetem Verlust eines Einzelereignisses[4] und erwarteter Anzahl an Ereignissen pro Jahr errechnet wird. Diese Methode hat zwei Vorteile: sie ist einfach anzuwenden, und sie liefert reale Zahlenwerte. Der mit Abstand größte Nachteil ([Landoll, 2006] zeigt in Kapitel 13 eine detaillierte Gegenüberstellung beider Ansätze inkl. einer vollständigeren Auflistung der jeweiligen Vor- und Nachteile.) ist jedoch, dass diese Zahlen mangels statistischer Daten meist selbst nur Schätzungen sind und daher eine trügerische Sicherheit bezüglich der Ergebnisse entsteht.[5] Dies ist auch der Hauptkritikpunkt an dieser Methodik, der u.a. Jaquith ([Jaquith, 2007] S.28ff.) dazu veranlasst quantitative Methoden allgemein als untauglich für den Einsatz zur Bewertung von Informationssicherheitsrisiken zu klassifizieren.

Auf der anderen Seite gibt es schon eine recht beachtliche Anzahl von qualitativen Bewertungsmethoden wie etwa [FAA SRMP], [CMMI], [OCTAVE], [NSA IAM] oder [BSI 100-3], die auch zum Teil aus anderen Ingenieursdisziplinen stammen und für die Informationssicherheitssparte angepasst wurden. Eine weitere Best Practice Methode, die zu den qualitativen Methoden zu zählen ist, ist das Open Source Security Testing Methology Manual [OSSTMM] der Isecom, bei der die aktuelle Sicherheitskonfiguration einer Komponente oder eines Dienstes einem Ideal gegenübergestellt wird um das Risiko aus der Differenz zu erheben.

Da bei qualitativen Methoden von vornherein nicht der Anspruch erhoben wird konkrete Zahlen liefern zu können ist deren Unschärfe sofort offensichtlich und daher auch weitgehend akzeptiert. Bezüglich Einfachheit und Anwendbarkeit sind qualitative Methoden durch ihre Flexibilität den quantitativen überlegen, haben allerdings den Nachteil, dass die Grenzen der jeweiligen Kategorien von Anwendungsfall zu Anwendungsfall unterschiedlich sein können, was die Vergleichbarkeit verschiedener Anwendungsfälle reduziert.[6]

2 Siehe u.a. auch ISO 27005 Kapitel 8.2.2
3 wie etwa die Einordnung in die Kategorien gering, mittel, hoch o.ä.
4 dieser Verlust wird selbst wieder aus dem Produkt aus Wert eines Objekts und dem sog. exposure Factor errechnet.
5 Ein weiterer methodischer Kritikpunkt ist, daß dieser Ansatz sich auf Erwartungswerte stützt und dadurch Extremereignisse nicht ausreichend berücksichtigen kann.
6 Eine besondere Eigenschaft der qualitativen Ansätze im Allgemeinen, aber auch der meisten bis dato angewandten quantitativen Ansätze, ist die Subjektivität der Risikobewertungen, denn im Grunde basieren die Methoden beider Ansätze bisher auf Faktormodellen, in denen zumindest einige der eingesetzten Faktoren durch das professional Judgement von Experten gebildet werden, und damit eine subjektive Bewertung entsteht. Diese Eigenschaft wirkt sich in Hinblick auf die Nachvollziehbarkeit und Trennschärfe einer Risikobewertung wie zu erwarten ist negativ aus, weshalb zumindest in einem quantitativen Risikobewertungsansatz gänzlich auf ein professional Judgement verzichtet werden sollte.

Aus dieser kurzen Gegenüberstellung ist das Dilemma, in dem sich Informationssicherheits-beauftragte bei der Analyse ihrer Risiken befinden, schon sehr gut zu erkennen. Mit Hilfe qualitativer Methoden bekommt man sehr einfach und schnell einen Überblick über die wichtigsten Risiken und kann grobe Veränderungen derselbigen auch dokumentieren. Ohne Zahlenwerte ist es dennoch nicht möglich diese Risiken in das grundlegende Zahlengerüst eines Unternehmens einzugliedern, um für etwaig zu implementierende Gegenmaßnahmen entsprechende Wirschaftlichkeitsrechnungen, die ja die Grundlage für die meisten Managemententscheidungen sind, durchführen zu können. An diesem Umstand können auch adäquate Sicherheitsmetriken, wie sie etwa von Jaquith [Jaquith, 2007] vorgeschlagen werden, nicht viel ändern, da auch diese nur die Effektivität von Gegenmaßnahmen im Bezug auf die technische Seite der Informationssicherheit aufzeigen können, nicht aber deren wirtschaftliche.

Die Problemstellung der Informationssicherheitsbeauftragen, die Frage nach dem adäquaten Budget, ist also im Grunde ein Problem der Messung der wirtschaftlichen Effektivität der ihnen zur Verfügung stehenden Maßnahmen. Und dafür ist aus Sicht des Autors ein quantitativer Ansatz unumgänglich.

Aktuelle Arbeiten in den USA wie jene von Henry et al. [Henry et al., 2009] verfolgen auch bereits wieder den quantitativen Ansatz und versuchen durch die Simulation der Ausbreitung von gezielten (Hack-) Angriffen in Prozesskontrollnetzwerken (SCADA Systemen) zu Aussagen über das mit diesen Angriffen verbundene Risiko zu gelangen. Der von Henry et. al. verfolgte Ansatz nimmt allerdings zwei Einschränkungen in Kauf die bei der breiteren Anwendung des von ihm vorgeschlagenen Modells hinderlich sind. Zum Einen stellen die gezielten bzw. gerichteten Angriffe für eine durchschnittliche Organisation nur einen sehr geringen Anteil aller auftretenden Incidents dar[7], und zum Anderen verzichten Henry et. al. im Rahmen der Bestimmung der Schadenpotentiale auf eine detaillierte Integration der von einem Ereignis betroffenen Prozesse.[8] Ein weiterer Kritikpunkt am Modell von Henry et. al. ist der Umstand, daß immer nur ein Angriffsszenario für sich betrachtet bzw. bewertet werden kann, und es Teils auch der Einfachheit der Betrachtung wegen nicht möglich ist eine Risikoanalyse für einen Prozess oder für eine ganze Organisation durchzuführen.

Diese Arbeit versucht nun durch Zugrundelegung eines Unternehmensmodells, welches eine Vereinheitlichung der Risikoanalyse über verschiedenste Unternehmens- und Organisationsformen hinweg ermöglicht, einen Schritt zur Lösung des derzeitigen Hauptproblems der quantitativen Risikoanalysemethoden, den Mangel an statistischen Daten, beizutragen. Im Rahmen einer periodisch wiederkehrenden Reevaluation des Risikos erhält der Anwender des Modells auch in weiterer Folge die wichtigen Aussagen bezüglich der Effektivität der gewählten Gegenmaßnahmen durch den Vergleich der Ergebnisse der jeweiligen Risikoanalyse(n). Zur Analyse des Risikos von Ereignissen wird dabei wiederum auf die grundlegende Methodik des Probalistic Risk Assessment (siehe Seite 29) zurückgegriffen, wobei im Detail Konzepte der Sachversicherungsmathematik[9] die Basis zur genauen Modellierung der Risiken bilden.

Im nun folgenden zweiten Kapitel werden zunächst die wichtigsten Basiskonzepte, die das Fundament der Arbeit bilden, kurz vorgestellt um diese dem mit der Materie noch nicht so gut vertrauten Leser näher zu bringen.

7 Der Überwiegende Anteil der Incidents wird durch ungerichtete Angriffe wie etwa von Malware ausgelöst.

8 Dieser Verzicht auf Details ist u.a. auch als pragmatische Antwort auf die derzeit nicht vorhandenen statistischen Daten zu sehen, wodurch es ermöglicht wird das Modell in der Praxis einfacher anzuwenden.

9 Genauer gesagt die Konzepte des kollektiven Modells, in dem das Risiko eines Versicherungsbestands, also die Verteilung des Gesamtschadens, über eine Schadenhäufigkeit und eine Verteilung der Schadenhöhe hergeleitet wird.

Das darauf folgende Kapitel drei wird den Ansatzpunkt der Arbeit mit der Definition der relevanten Informationssicherheitsziele und der diese bedrohenden Ereignisse abhandeln. Ebenfalls Teil dieses Kapitels ist der Einstieg in das Modell, der über eine sehr vereinfachte Darstellung der Kernkonzepte des Modells deren grundlegendes Zusammenwirken darlegt.

Der vierte Abschnitt ist der detaillierteren Darstellung des Modells gewidmet. Hier wird gezeigt wie die Systemeigenschaften hinsichtlich der Unternehmensprozesse und des Unternehmensaufbaus in dem Modell abgebildet werden, und welche grundlegenden Annahmen zur Konstruktion des Modells getätigt wurden. Weiters wird in Kapitel vier gezeigt welche Schnittstellen des Unternehmens mit seiner Umgebung im Modell wichtig sind und daher berücksichtigt werden müssen.

Kapitel fünf bildet den Abschluss der Modellbeschreibung und beinhaltet vor allem die formelle mathematische Modellierung sowie die Einschränkungen und Detailannahmen des Modells. Es werden hier auch Lösungsansätze für relevante Einzelprobleme, die sich in der praktischen Anwendung des Modells ergeben können, skizziert.

Kapitel sechs beinhaltet ein Beispiel für die Anwendung des Modells in einem fiktiven Kleinunternehmen, das auf ein kleines IT-System angewiesen ist.

Im letzten Teil der Arbeit, dem Kapitel sieben, wird das Modell selbst noch einmal kurz zusammengefasst, bevor auf die noch offenen Punkte und die daraus resultierenden zukünftig sinnvollen Entwicklungsmöglichkeiten eingegangen wird.

2 Basiskonzepte

In diesem Kapitel sollen die wichtigsten dieser Arbeit zugrunde liegenden Konzepte und Standards einführend dargestellt werden. Der Leser soll dadurch einen grundlegenden Überblick über die von dieser Arbeit berührten Themengebiete erlangen. Da eine detaillierte Darstellung der jeweiligen Konzepte den Umfang dieser Arbeit sprengen würde, sei der interessierte Leser für die hier nicht enthaltenen Teile an die jeweilige Literatur verwiesen.

Bezüglich der Reihung der jeweiligen Konzepte sei darauf hingewiesen, daß diese nicht auf Basis deren Relevanz erfolgte, sondern vielmehr auf Grund von didaktischen Überlegungen vorgenommen wurde um einen möglichst einfachen Einstieg in die Materie zu gewährleisten. Da es das Ziel dieser Arbeit ist einen Beitrag zur Bewertung von Risiken und deren Gegenmaßnahmen zu leisten, wurde COSO2 [COSO2] für den Einstieg gewählt, da in diesem Rahmenwerk dargestellt wird in welchem organisatorischen Kontext die Risikobewertung zu sehen ist bzw. ablaufen sollte.

2.1 COSO2

COSO2 ist jener grundlegende Standard, der als primärer roter Faden dieses Dokumentes dienen soll. COSO2 wurde im September 2004 unter dem Namen „Enterprise Risk Management[10] – Integrated Framework“ vom Commitee of Sponsoring Organisations of the Treadway Commission veröffentlicht. Wie der Name schon andeutet, wurde mit COSO2 eine Vorgehensweise definiert wie Unternehmen unterschiedlichster Größe und Organisationsstruktur das Problem des unternehmensweiten Risikomanagements bewältigen können. Im Folgenden soll nun für jene Leser, die mit COSO2 noch nicht vertraut sind, eine kurze Einführung in die wichtigsten Elemente dieses de facto Standards gegeben werden.

2.1.1 Aufbau von COSO2

COSO2 definiert eine dreidimensionale Sicht auf das Unternehmen und seine Risikomanagementaktivitäten. Dadurch spannt sich ein Würfel (s.u.) auf, welcher die Dimension *Risikomanagement Komponenten* bzw. *Risikomanagementaktivitäten* die Dimension *Unternehmensziele* - die als Basis für diese Aktivitäten dienen- und die Dimension der *Unternehmensorganisationsebenen* - innerhalb derer die Aktivitäten durchgeführt werden sollen- beinhaltet.
Jede dieser Dimensionen und ihre Ausprägungen, wie sie im Rahmen von COSO2 definiert wurden, werden in den folgenden Unterkapiteln kurz beschrieben.

10 Enterprise Risk Management wird im weiteren Verlauf mit ERM abgekürzt.

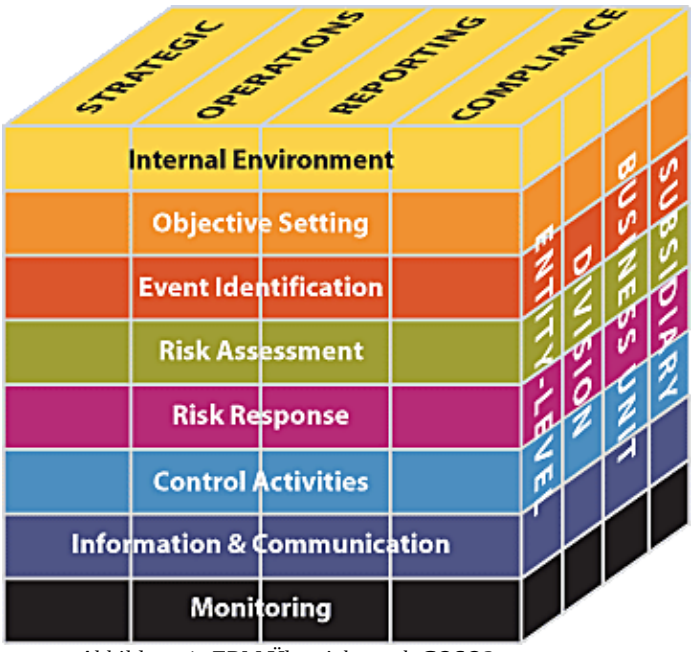

Abbildung 1: ERM Übersicht nach COSO2

2.1.2 Unternehmensorganisationsebenen

In COSO2 wurden vier Organisationsebenen definiert, wobei es von der Größe eines Unternehmens[11] abhängt wie viele davon wirklich benötigt werden um jeweils ein spezielles Unternehmen abbilden zu können.
Durch den Umstand, dass die Nutzung jeder Beschreibungsebene nicht zwingend vorgeschrieben ist, ist es möglich dieses Rahmenwerk auch für KMU's zu adaptieren.

- Entity-Level: Die Organisationsebene Entity-Level repräsentiert das Unternehmen als ganzes inklusive aller Geschäftsfelder und Beteiligungen.

- Division: Diese Ebene repräseniert einzelne Geschäftsfelder des Unternehmens.

- Business Unit: Die Business Unit repräsentiert einzelne von einander weitgehend unabhängige Geschäftszweige innerhalb eines Geschäftsfeldes.

- Subsidiary: Subsidiary sind all jene Beteiligungen eines Unternehmens, die von diesem gesteuert werden. Diese Ebene entspricht damit soweit der Division Ebene.

11 Unternehmen und eine Organisation im Allgemeinen sind im Zusammenhang mit COSO2 synonym zu verstehen.

2.1.3 Unternehmensziele

COSO2 definiert vier grundlegende Kategorien von Unternehmenszielen, die eine differenzierte Betrachtung einzelner Ziele erlaubt, die durchaus auch mehr als einer dieser Kategorien zugeordnet sein können. Diese differenzierte Betrachtung ist insofern hilfreich, da jede der Kategorien unterschiedliche Bedürfnisse des Unternehmens widerspiegelt und dementsprechend in unterschiedliche personelle Zuständigkeiten innerhalb des Vorstandes fallen kann.

- Strategic: Strategische Unternehmensziele sind jene, die direkt die Mission des Unternehmens unterstützen sollen und auch mit der Mission abgestimmt sein müssen. Die Aktivitäten (siehe Kapitel 2.1.4) im Rahmen von COSO2 sollen sicherstellen, dass die strategischen Unternehmensziele auch wirklich die Mission unterstützen.

- Operations: Operative Unternehmensziele stellen Vorgaben für den effektiven und effizienten operativen Einsatz von Unternehmensressourcen zur Erfüllung der strategischen Ziele dar. COSO2 Aktivitäten sollen prüfen ob die operativen Tätigkeiten effektiv und effizient im Umgang mit den Ressourcen des Unternehmens sind.

- Reporting: Diese Kategorie umfasst all jene Ziele und Vorgaben, die die Definition eines korrekten von vollständigen Berichtswesens darstellen. In Bezug auf das Berichtswesen soll die Anwendung der COSO2 Aktivitäten sicherstellen, dass die jeweiligen Berichte korrekt und vollständig, also zuverlässig sind.

- Compliance: Zu dieser Kategorie gehören alle - meist extern vorgegebenen - gesetzlichen Vorschrifen und Vorgaben die das Unternehmen gezwungen ist einzuhalten. Die COSO2 Aktivitäten sollen sicherstellen, dass diese Vorgaben auch eingehalten werden.

2.1.4 Risikomanagement Komponenten

Den eigentlichen Kern von COSO2 stellen die nun folgend beschriebenen acht Aktivitäten dar.[12]

2.1.4.1 Internal Environment

Internal Environment kann wohl am besten mit dem Begriff internes Umfeld / Rahmenbedingungen beschrieben werden. Das Internal Environment leitet sich zum Teil aus der Unternehmenskultur ab und beeinflusst das Risikobewusstsein der Mitarbeiter maßgeblich. Das interne Umfeld ist aber auch die unabdingbare Basis für alle anderen Komponenten von COSO2, da es unter anderem auch den Risikoappetit[13] des Managements, die ethischen Werte und die Integrität der Mitarbeiter und des Unternehmens mitbestimmt. Es ist also unbedingt notwendig, dass eine Unternehmensführung vor der Etablierung aller anderen Aktivitäten die internen Rahmenbedingungen festlegt, innerhalb derer das Risikomanagement und der Betrieb des Unternehmens stattfinden sollen.

2.1.4.2 Objective Setting

Diese Komponente von COSO2 beschreibt den Prozess der Zieldefinition. Hier sollen für alle Organisationsebenen und alle Zielkategorien die entsprechenden Ziele der jeweiligen Organisationsebene definiert werden. Bei der Durchführung dieser Aktivitäten sind ebenfalls für jedes Ziel der Risikoappetit und die Risikotoleranzen[14] zu definieren.

2.1.4.3 Event Identifikation

Im nächsten Schritt muss das Management jene Ereignisse identifizieren, die - soferne sie eintreten - einen massiven Einfluss auf das Unternehmen und seine Zielerreichung haben können. Negative Ereignisse werden in COSO2 als Risiken gesehen (z.B.: Naturkatastrophen) und positive Ereignisse als Chancen (z.B.: Konjunkturverbesserungen)[15]. Um ein möglichst vollständiges Bild des Unternehmens zu erhalten muss das Management eine Fülle von internen und externen Ereignissen einbeziehen.

2.1.4.4 Risk Assessment

Bei der Risikobewertung beschränkt man sich im Allgemeinen auf jene Ereignisse, die negative Wirkungen haben können. Hier muss das Management feststellen, wie wahrscheinlich ein Ereignis ist und welche potentiellen Auswirkungen (Schadenausmaß) mit seinem Eintreten verbunden sind.[16] COSO2 trifft hier die Unterscheidung zwischen inhärenten und residualen Risiken.
Inhärent sind Risiken, so lange sie „unbehandelt“, das heißt ohne aktivierte entsprechende Gegenmaßnahmen, sind. Residuale Risiken sind hingegen ursprünglich inhärente Risiken, die

12 Die folgenden Kurzbeschreibungen sind inhaltliche Übersetzungen der jeweiligen Kapitelbeschreibungen in COSO2.
13 Das ist jenes Risiko welches das Management des Unternehmens noch bereit ist zu tragen.
14 Jene Abweichungen von den Zielen, die gerade noch akzeptabel sind.
15 COSO2 weicht in diesem Punkt etwas von der gängigen Nomenklatur ab, in der Ereignisse die nur negative Wirkung haben können als reine Risiken, und jene Ereignisse die auch positive Wirkung haben können als spekulative Risiken bezeichnet werden.
16 Diese Methode der Risikobewertung wird auch Probabalistic Risk Assessment (PRA) genannt. Siehe auch Kapitel 2.5.

schon durch adäquate Gegenmaßnahmen in ihrer Eintrittswahrscheinlichkeit und / oder ihren Auswirkungen reduziert wurden. Teil der Aktivität Risikobewertung ist auch die Bestimmung des Risikos des Portfolios von Risiken, die einer Organisationseinheit zugeordnet sind.

2.1.4.5 Risk Response

Die Risk Response Aktivität definiert, dass das Management hier zu bestimmen hat, wie auf die identifizierten und bewerteten Risiken reagiert werden soll. Als Entscheidungsbasis für die zur Verfügung stehenden Reaktionen auf das Risiko (Risiko vermeiden, Risiko reduzieren, Risiko teilen und Risiko akzeptieren) soll einerseits der Risikoappetit bzw. die definierte Risikotoleranz, aber auch das aggregierte Risiko der jeweiligen Organisationseinheit, dessen residualer Anteil mit dem Risikoappetit (-toleranz) in Übereinstimmung gebracht werden sollte, dienen.
Teil dieser Aktivität ist auch das Monitoring bezüglich der Effektivität der gewählten Risk Responses. Beispiele für Risk Responses sind unter anderem Maßnahmen auf technischer / infrastruktureller Ebene wie etwa Redundanzen, Lastverteiler, das Anbringen von Feuerlöschern udgl.

2.1.4.6 Control Activities

Unter Control activities versteht COSO2 all jene Maßnahmen und Vorgaben (policies) die dabei helfen sicher zu stellen, dass die im Rahmen der Risk Response Aktivität gewählte Reaktion auf das Risiko auch umgesetzt wird.[17] Ein Beispiel für solche Kontrollaktivitäten wäre etwa die Prüfung, ob die Feuerlöscher auch wirklich auf den dafür vorgesehenen Lokationen aufgestellt wurden und ob sie auch gewartet werden. Ein weiterer Teil dieser Aktivität ist auch die Prüfung, wie effizient die der Risk Response gewählten Maßnahmen umgesetzt wurden.

2.1.4.7 Information & Communication

Diese essentielle Aktivität von COSO2, die effizientes Risikomanagement erst ermöglicht, verlangt, dass ein Unternehmen sicherstellen muss, dass sämtliche zum Erkennen und zur Bewertung von Risiken notwendigen Informationen auch erzeugt und gesammelt werden. Weiters ist es notwendig, dass diese Informationen über geeignete und effiziente Kommunikationskanäle zeitgerecht an jene Personen weitergeleitet werden, die diese Informationen benötigen. Das bedeutet, dass sämtliche am Risikomanagement beteiligten Personen[18] sich im Klaren sind, wie wichtig ihr Beitrag für ein funktionierendes Risikomanagement ist und dass es außer Zweifel steht, dass das Top-Management ein solches Risikomanagement will. Des weiteren ist es wichtig, dass nicht nur die Kommunikationskanäle von Management zu den Mitarbeitern, über die die Aufgabenverteilung im Rahmen des Risikomanagements abgewickelt wird, effektiv und effizient sind, sondern auch jene von den Mitarbeitern zum Management, über die ja das Berichtswesen abgewickelt werden sollte. Die zweite wichtige Kommunikationsachse im Unternehmen, die funktionieren sollte ist die horizontale Achse der Organisationseinheiten bzw. deren Mitarbeitern untereinander.

17 Im Kontext der Informationssicherheit sind die zwei Standardwerke Cobit v.4 und ISO 27002 zum Großteil den COSO2 Komponenten *Risk Response* und *Control activities* zuordenbar.

18 Im allgemeinen sollten dies in der einen oder andern Form alle Mitarbeiter des Unternehmens sein.

2.1.4.8 Monitoring

Die Aktivität, die COSO2 abschließt, ist die wiederkehrende Überprüfung ob die sieben obig genannten Aktivitäten wie gewünscht funktionieren, und festzustellen wo es Abweichungen bzw. Verbesserungspotential gibt. Hierbei unterteilt COSO2 die Monitoringaktivität noch in zwei eigenständige Teile, die auch unterschiedliche Zielsetzungen haben.
Der erste Teil ist ein ständiges Überwachen der einzelnen Aktivitäten von COSO2 hinsichtlich deren Effektivität und deren Effizienz um kontinuierliche Verbesserungen bei der Durchführung der einzelnen Aktivitäten zu erreichen.
Der zweite Teil ist eine eigenständige[19] Evaluation, die zum einen von Zeit zu Zeit überprüfen soll, ob das Enterprise Risk Management des Unternehmens als Ganzes effektiv ist, und zum anderen Verbesserungspotentiale ausmachen soll.

2.1.5 COSO2 ERM im Unternehmensregelkreis

Betrachtet man die in COSO2 definierten Aktivitäten aus einer Systemperspektive, lassen sich diese auch als Unternehmensregelkreis in der folgenden Form darstellen.

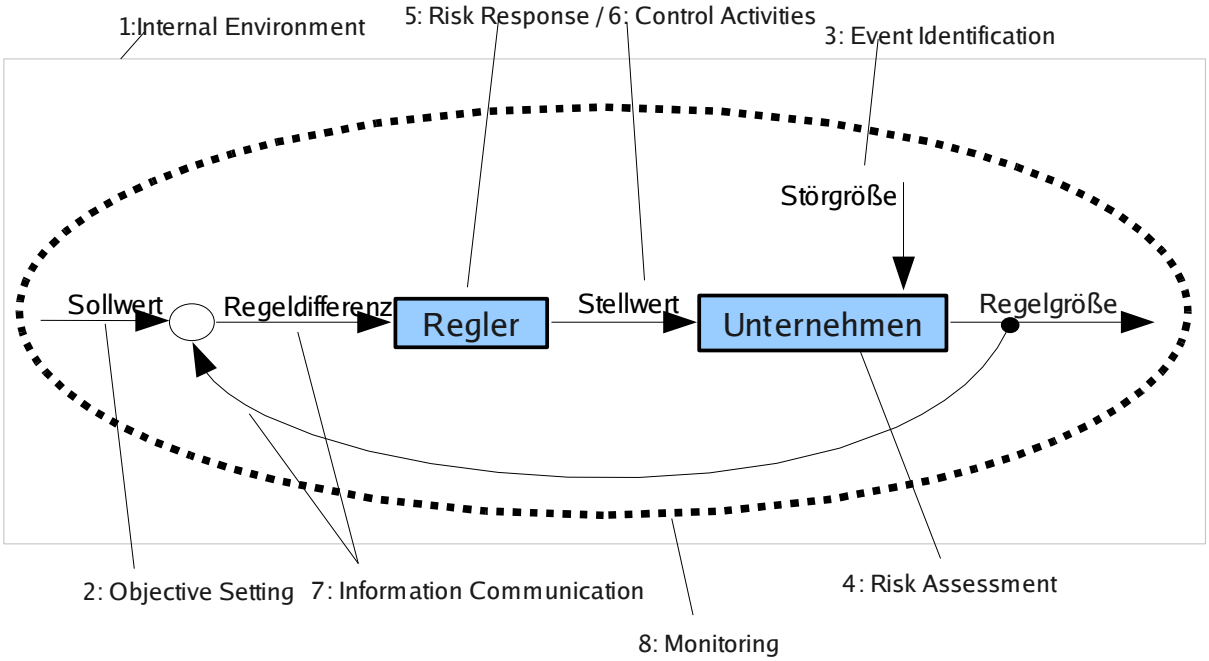

Abbildung 2: COSO 2 im Unternehmensregelkreis

In der obig dargestellten Interpretation von COSO2 kommt der Aktivität Internal Environment die Aufgabe zu, die Existenz der einzelnen Komponenten des Regelkreises und deren Zusammenspiel festzulegen. Durch die Aktivität Objective Setting werden die Sollwerte - also die Ziele, der sich die Regelgröße[20] möglichst annähern soll definiert.
Durch die Aktivität Event Identification sollen alle relevanten Störgrößen identifiziert werden. Basierend auf den identifizierten Störgrößen soll deren Einfluss auf das System Unternehmen und damit auch deren Einfluss auf die Regelgröße in der Aktivität Risk Assessment festgestellt werden. Durch die Aktivität Risk Response wird der Regler bzw. die

19 meist mit Hilfe von externen Fachkräften durchgeführte
20 Als betriebswirtschaftliche Interpretation der Regelgröße kann der Unternehmensgewinn oder eine ähnliche Kenngröße gesehen werden.

Regelungsstrategie definiert, durch die der Stellwert, der den Störgrößen entgegenwirken soll, erstellt wird. Die Control Activities müssen danach sicherstellen dass der Stellwert im Unternehmen ankommt bzw. umgesetzt wird.
Die Aktivität Information und Communication soll in diesem Regelkreis sicherstellen, dass die relevanten Informationen über die Regelgröße bzw. über die aus dieser und der Sollgröße ableitbaren Regeldifferenz in der geeigneten Form dem Regler zur Verfügung gestellt werden.
Im Zuge der Monitoring Aktivität muß hierbei sichergestellt werden dass der Regelkreis als Ganzes effektiv und effizient funktioniert.

Aus obiger Darstellung ist gut ersichtlich, dass der Regler nur dann seiner Aufgabe effizient gerecht werden kann, wenn er abschätzen kann, wie der von ihm festgelegte Stellwert auf das System Unternehmen wirkt und gemeinsam mit der Störgröße die Regelgröße beeinflusst. Das bedeutet, dass der Regler ein Modell vom System Unternehmen benötigt, in dem die wichtigsten Wirkungszusammenhänge akkurat abgebildet sind.

Im nun folgenden Kapitel wird ein systemtheortischer Ansatz vorgestellt, der eine sehr gute Grundlage zur Bildung dieses Unternehmensmodells bietet.

2.2 Das Unternehmen als hierarchisches Multiebenen System

Eine wichtige dieser Arbeit zugrunde liegende Betrachtungsweise ist jene, dass ein Unternehmen als hierarchisches soziotechnisches Multiebenensystem begriffen wird. Einen guten Ausgangspunkt für die Analyse solcher Systeme stellt die Arbeit von Mesarovic [Mesarovic et al., 1970] dar, deren grundlegende Konzepte, die auch die Basis für diese Arbeit bilden, nun folgend kurz beschrieben werden sollen.
Eine wichtige allgemeine Eigenschaft von Systemen und damit auch jener die von Mesarovic behandelt wurden ist, dass diese Systeme einen für sie definierten Input in einen definierten Output transformieren.
Hierarchische Systeme werden dabei als vertikale Zusammenstellung von Subsystemen betrachtet, in der das Gesamtsystem aus mehreren miteinander interagierenden Familien von Subsystemen besteht, wie dies in Abb. 3 Dargestellt ist.[21]

21 siehe auch Mesarovic et. al. [Mesarovic et. al, 1970] S.35 Abb. 2.1

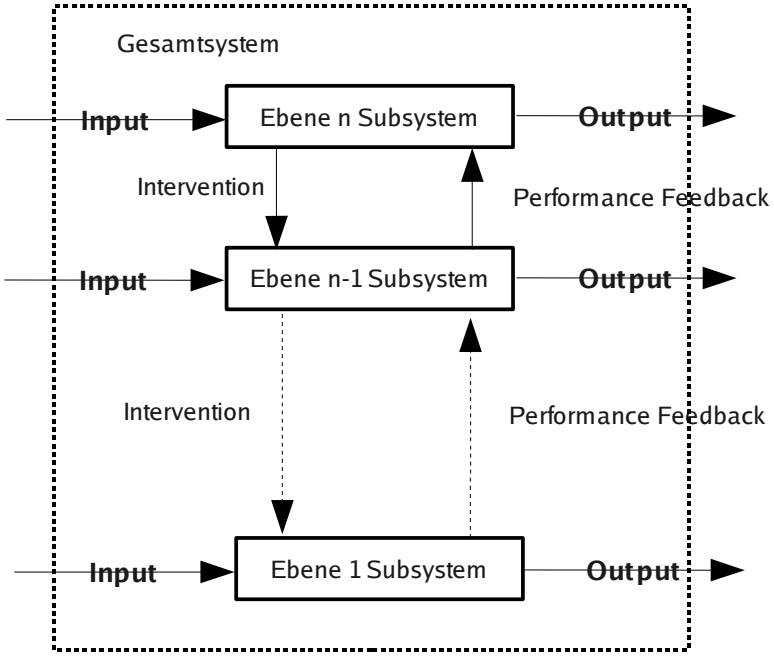

Abbildung 3: Multiebenensysteme nach Mesarovic

Neben dem Umstand, dass sämtliche Subsysteme über ihren Input und Output mit der Systemumgebung kommunizieren können, sind vor allem die Interaktionen der Subsysteme untereinander von Interesse.
Durch die Intervention sind die Abläufe in einem Subsystem - gleich welcher Ebene - direkt von den Subsystemen der höheren Ebenen beeinflusst. Mesarovic spricht im Zusammenhang mit der Intervention auch vom Recht auf Intervention, das bei den übergeordneten Subsystemen liegt, und drückt dies durch den Terminus „supremal unit“ und „infimal unit“ aus.
Das Performancefeedback zeigt die Abhängigkeit der übergeordneten Einheiten von den untergeordneten auf, da die übergeordneten Einheiten und das System als Ganzes nur dann ihre Ziele erreichen bzw. erfolgreich sind, wenn alle Subsysteme ihre Ziele erreichen bzw. erfolgreich sind. Das Performancefeedback zeigt den übergeordneten Einheiten an, wie erfolgreich die untergeordneten Einheiten sind, und kann nach Mesarovic als Antwort auf die Intervention der übergeordneten Einheiten gesehen werden.

Dieser grundlegende Aufbau einer Systemhierarchie kommt in allen drei, in den Folgekapiteln beschriebenen Hierarchietypen für Systeme, die von Mesarovic beschrieben werden, zur Anwendung. Die Typisierung der Hierarchien entspricht dabei einer Klassifikation und nicht einer Partition; das heißt, ein System kann durchaus mehr als einem Typus zugeordnet bzw. durch diesen beschrieben werden.

2.2.1 Strata

Das Stratum ist der von Mesarovic benutzte Terminus für die Beschreibungs- bzw. Abstraktionsebene eines Systems. In jedem Stratum gelten dabei eigene Regeln nach denen das System auf dieser Ebene funktioniert, und es existiert eine Funktionalität, die nur in dieser Ebene bereitgestellt werden kann. Die einzelnen Strata sollten dabei in ihrer Funktionalität soweit wie möglich unabhängig voneinander sein, damit eine vernünftige Abgrenzung der Strata untereinander möglich wird. Ein gutes Beispiel für eine Darstellung eines Systems anhand von Strata ist das ISO-OSI Modell [ISO/OSI] für Netzwerke (siehe auch Kap. 2.6) in dem jeder Ebene eine spezielle Funktionalität zugeordnet wird und die Kommunikation der Ebenen untereinander über definierte Schnittstellen die die jeweilige Ebene (Stratum) kapseln sollen abgewickelt wird. Der Zweck der Strata ist es also, die Gesamtkomplexität eines Systems durch geschickte Kapselung von Funktionalität bzw. Komplexität in verschiedenen Ebenen zu reduzieren bzw. übersichtlicher zu gestalten.

Die folgende, ebenfalls der Arbeit von Mesarovic entnommene Abbildung soll dieses Hierarchiekonzept noch besser darstellen.

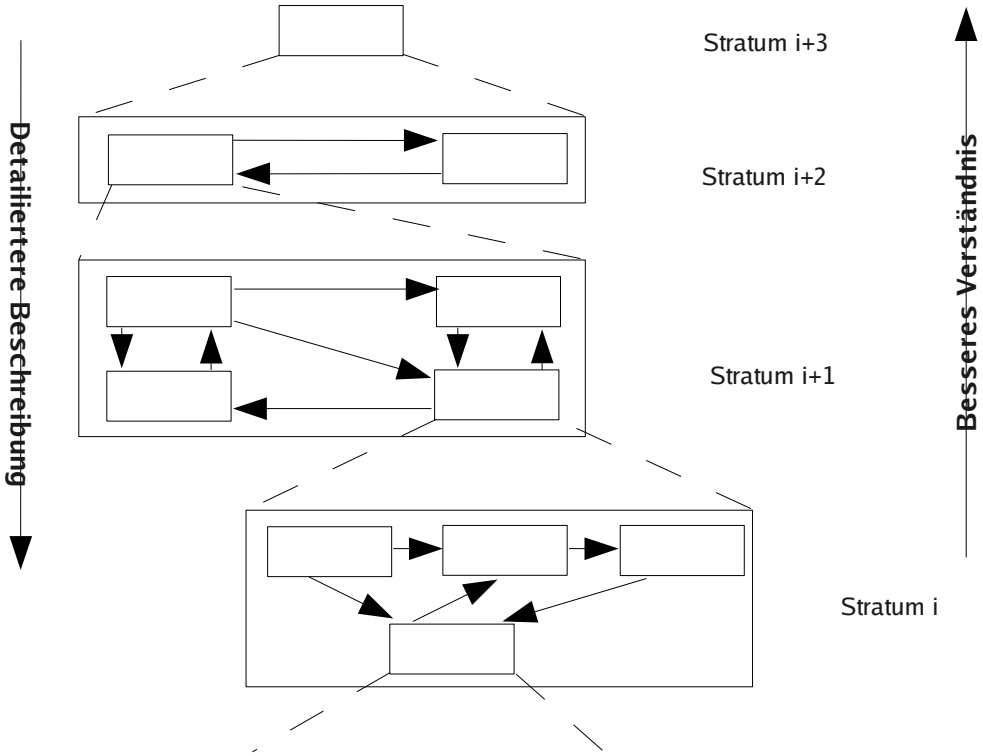

Abbildung 4: Systemgliederung durch Strata

Aus dieser Darstellung wird deutlich, dass ein System in einem Stratum zum Subsystem in einem höherliegenden Stratum wird bzw. werden kann. Dieser Zusammenhang zwischen den einzelnen Strata bedingt auch die Eigenschaft, dass im Laufe eines Abstiegs durch die Hierarchie mehr Details des Systems sichtbar werden und im Gegensatz dazu bei einem Hinaufsteigen der Hierarchie die Zusammenhänge im Gesamtsystem klarer werden.

2.2.2 Layer

Mit Layern werden von Mesarovic Ebenen mit unterschiedlicher Komplexität in Entscheidungssystemen modelliert. Dieser Ansatz basiert darauf, ein komplexes Entscheidungsproblem eines Systems soweit in simplere Subprobleme zu zerteilen, daß diese von jeweils einem Layer bzw. einer Einheit gelöst werden können. Die Lösung des Gesamtproblems entsteht dabei aus der Lösung aller Einzelprobleme. Eine Hierarchie von Entscheidungslayern hat nach Mesarovic den in Abb. 5 dargestellten Aufbau.

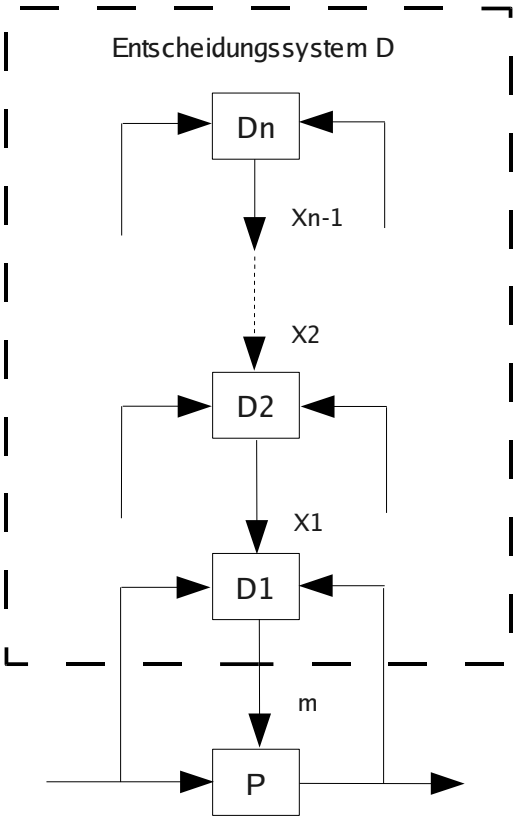

Abbildung 5: Entscheidungssystem nach Mesarovic

Die Aufgabe des hier dargestellten Systems ist dabei, den optimalen Steuervektor m für den Prozess P zu bestimmen. Dabei entspricht der Output jeder Entscheidungsebene D_i der Lösung eines Entscheidungsproblems bzw. der Konsequenz der Lösung, die auf den Lösungen der Entscheidungsprobleme der übergelagerten Ebenen bzw. Einheiten beruht. Geht man zum Beispiel davon aus, dass das gesamte System n Bedingungen bzw. Einschränkungen bestmöglich zu erfüllen hat um den optimalen Steuervektor m zu erhalten, so trifft ein Layer die Entscheidung, in welchem Grad die ihm zugeteilte Einschränkung berücksichtigt werden soll, bzw. liefert die Lösung für die jeweilige Einschränkung. Dadurch wird die Entscheidung des Gesamtsystems mit jeder Ebene, die in der Hierarchie nach unten gestiegen wird, komplettiert und die unterste Ebene muss basierend auf den n-1 Lösungen der darüberliegenden Ebenen, die im Vektor X1 enthalten sind, nur noch die Entscheidung über eine Einschränkung treffen.
Für dynamische bzw. regelnde Systeme versteht es sich von selbst, dass die getroffenen Entscheidungen auf Grundlage des Performancefeedback evaluiert und ggf. geändert werden.

2.2.3 Echelons

Die dritte von Mesarovic angeführte Hierarchie ist die organisatorische. Damit eine solche Hierarchie in einem System entsteht ist es notwendig, dass das System aus mehreren miteinander interagierenden Subsystemen besteht, von denen einige Entscheidungssysteme in einer entsprechenden Hierarchie sein müssen. Jede Ebene in einer solchen Hierarchie wird als Echelon bezeichnet. Die folgende Abbildung zeigt ein System mit einer solchen organisatorischen Hierarchie[22].

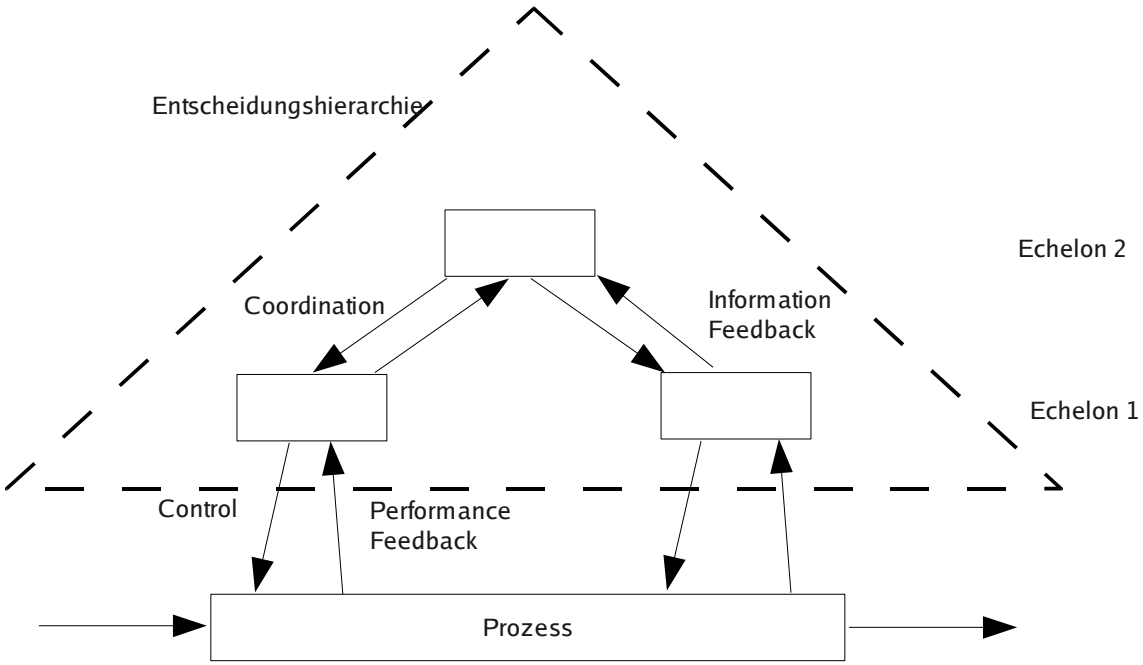

Abbildung 6: Entscheidungshierarchie / Echelons nach Mesarovic

Die obige Entscheidungshierarchie wird als multilevel multigoal System[23] bezeichnet, da das System aus mehreren Echelons mit Entscheidungseinheiten, die unterschiedliche Ziele haben können, besteht. An dieser Stelle ist es wichtig daran zu erinnern, dass auch die Darstellung eines Systems als organisatorische Hierarchie nur eine Sicht auf ein System darstellt, und dass diese den anderen beiden vorgestellten Sichten gleichgestellt ist. Dies kann man auch daran erkennen, dass eine Entscheidungseinheit wie sie in Kap.2.2.2 dargestellt wurde, selbst wieder eine multiechelon Hierarchie sein kann.
In Zusammenhang mit Entscheidungshierarchien werden von Mesarovic auch zwei grundlegende Arten von Interventionen aufgezeigt: die Koordination, die zwischen zwei Entscheidungseinheiten notwendig ist, und die Kontrolle, die gegenüber den geregelten Prozessen notwendig ist. Die Arbeit von [Mesarovic et.al., 1970] fokussiert sich nach der Darstellung dieser drei unterschiedlichen Hierarchietypen primär auf die Optimierung der Koordination zwischen den Entscheidungseinheiten der Entscheidungshierarchie, auf die aber hier nicht näher eingegangen werden soll.

22 Das hier gezeigte System ist nur ein Beispiel für eine Entscheidungshierarchie, weitere Beispiele finden sich bei [Mesarovic et. al., 1970]

23 Neben mulitlevel mulitgoal Systemen identifizierte Mesarovic auch noch singlelevel singlegoal und singlelevel multigoal Systeme, die aber eher Spezialfälle darstellen und in der Praxis selten vorkommen.

2.3 ANSI ISA-95 / Purdue Referenzmodell

Nach der Darstellung, wie ein Unternehmen als hierarchisches Multiebenensystem grundsätzlich modelliert werden kann, soll in diesem Kapitel mit ANSI ISA 95 ein Standard vorgestellt werden, der eine praktische Implementierung des obig vorgestellten theoretischen Ansatzes ermöglicht.
Dieses Kernkonzept ist im ersten Teil des Standards ANSI ISA-95 [ISA 95-1], der auch unter der internationalen Bezeichnung IEC 62264 veröffentlicht ist, enthalten. Veröffentlicht wurde er vom American National Standards Institute und verfasst von der Instrument Society of America. Das komplette Standardwerk umfasst aktuell fünf Teile, von denen zur Zeit aber nur die Teile eins bis drei offizielle Standards sind.
Der Standard selbst definiert ein Rahmenwerk, wie die Komponenten, Produkte und die Architektur / der Systemaufbau von Produktionsbetrieben in ein informationstechnologisch verarbeitbares Datenmodell abgebildet werden können.[24]
Die Hauptmotive, die zur Entstehung dieses Standards führten, sind die Probleme, die bei der Zusammenschaltung von produktionsnahen IT-Systemen - wie sie schon jahrzehntelang im Computer aided Manufacturing (CAM) oder dem Computer aided Design (CAD) zum Einsatz kommen - und Systemen zur Unterstützung der Produktionsplanung und anderer produktionsrelevanter Kontrollprozesse entstehen. Ein Großteil dieser Probleme entsteht dadurch, dass diese Systeme auf grund unterschiedlicher Datenmodelle und Softwarearchitekturen nur sehr eingeschränkt miteinander kommunizieren können. Dies wiederum schränkt die Optimierungspotentiale[25] der betroffenen Unternehmen ein.

Die fünf Teile von ANSI ISA-95 haben kurz zusammengefasst folgende Inhalte.

Teil1:

Teil 1 beinhaltet das Purdue Modell (siehe Kap. 2.3.5), welches eine funktionale Hierarchie für Unternehmensprozesse definiert und sich damit gut zur Darstellung der Prozesshierarchie und der Vernetzung der Prozesse bzw. der Organisationsstruktur eines Unternehmens eignet. Weiters enthält dieser Teil ein Modell, das die wichtigsten mit der Produktion assoziierten in Purdue Ebene 3 und teilweise Ebene 4 liegenden Funktionen[26] eines Produktionsbetriebes und deren Kommunikationspfade untereinander definiert.
Es werden auch die Datenobjekte inklusive deren Hierarchie definiert, die für die Kommunikation der Funktionen untereinander benötigt werden.
Im Grunde wird in diesem Teil die Schnittstelle zwischen den Prozessen zur Gesamtunternehmenssteuerung und jenen zur Steuerung der operativen Einheiten definiert.

Teil2:

In Teil 2 von ISA-95 werden die Attribute für die in Teil 1 definierten Datenobjekte festgelegt. Dieser Teil stellt somit eine Detaillierung von einzelnen Kapiteln aus Teil 1 dar, die für eine möglichst einheitliche Implementierung dieser Kapiteln notwendig ist.

Teil 3:

Teil 3 beschäftigt sich sehr detailliert mit den Funktionen in der Purdue Ebene 3, dem Management von Produktionsoperationen (Manufacturing Operations Management) und den

24 zur Beschreibung des in ISA-95 definierten Datenmodells wurde die Unified Modelling Lauguage [UML] verwendet.
25 etwa im supply chain management oder agile manufacturing
26 wie etwa Produktionskontrolle, Qualitätskontrolle oder Produktionsscheduling

Interaktionen der zu den Funktionen gehörenden Aktivitäten mit dem Datenmodell. Damit vertieft Teil 3 die Ansätze aus Teil 1.

Teil 4:

Teil 4 definiert das Datenmodell und die darin enthaltenen Datenobjekte inklusive ihrer Attribute, die für die Kommunikation zwischen den Aktivitäten innerhalb, der in Teil 3 definierten, Funktionen notwendig sind.

Teil 5:

Dieser Teil definiert jene Transaktionen, die für den Datenaustausch zwischen Geschäftssystemen und Produktionssystemen genutzt werden sollen. Damit regelt dieser Teil, auf welcher Basis die Kommunikation zwischen Purdue Ebene 4 und Ebene 3 (siehe Kap. 2.3.5) in den IT-Systemen durchgeführt werden soll.

Auf Basis der im ersten Teil von ISA-95 definierten Konzepte und dem Purdue Modell ist es möglich, die Prozessarchitektur bzw. -hierarchie eines Unternehmens strukturiert und in unterschiedlichen Detailtiefen zu beschreiben. In den nun folgenden Unterkapiteln werden die für das in diesem Dokument thematisierte Modell maßgeblichen in ISA-95 enthaltenen Konzepte näher dargestellt.[27]

2.3.1 Production definition information

Die Production definition information umfasst all jene Informationen die beschreiben, wie ein bestimmtes Produkt hergestellt wird. Im Detail werden in ISA-95 folgende Subkategorien angeführt:

- Product production rules: Dies sind die exakten Regeln, nach welchem „Rezept“ ein bestimmtes Produkt hergestellt wird. Es sind dies also Informationen über die anzuwendenden Arbeitsschritte bzw. Handlungsfolgen.
- Bill of material: Dies ist die Liste aller für die Herstellung eines bestimmten Produktes notwendigen Materialien, seien dies nun Rohmaterialien oder Hilfsstoffe.
- Bill of ressources: Dies ist die Liste aller Ressourcen, die für die Herstellung eines bestimmten Produktes notwendig sind. Die Bill of material ist eine Untermenge der Bill of ressources, die allerdings auch die benötigte Infrastruktur, Arbeitskraft, Verbrauchsgüter udgl. enthält.
- Product Segment: repräsentiert einen Arbeitsschritt und beinhaltet all jene Informationen über benötigte Ressourcen und durchzuführende Handlungen, die zur Durchführung des Arbeitsschritts und damit zur Bearbeitung eines Produktes in diesem Produktionsschritt notwendig sind.

2.3.2 Product Information

Dieser Informationstypus umfasst all jene Informationen die angeben, wie die Produktion

27 Jene in ANSI ISA 95 definierten Informationsobjekte die in der untenstehenden Auflistung nicht enthalten sind, sind für das Aufspannen des in der Folge beschriebenen Modells zwar nicht relevant, im wesentlichen sind diese aber als Input bzw. Output der jeweiligen Prozesse in einer Organisation zu sehen.

eines bestimmten Produktes bisher verlaufen ist und welche Produkte wann noch produziert werden sowie ein Inventar aller Materialien und Produkte die während der Produktion verbraucht und hergestellt wurden. Von besonderem Interesse in diesem Dokument ist allerdings nur die Production scheduling information, die genaue Daten über die zukünftig geplanten Produktionsschritte und deren Zuteilung zur benötigen Infrastruktur enthält.

2.3.3 Processsegments

Prozesssegmente repräsentieren einen Schritt im Produktionsprozess und beinhalten eine Liste an Anforderungen hinsichtlich der benötigten Tauglichkeit bzw. Leistungsfähigkeit von Material, Personen und Equipment (Infrastruktur), die zur Durchführung des Prozessegments notwendig sind. ISA-95 definiert diese drei Gruppen als Material capability, Personnel capability und Equipment capability.
In einem Unternehmen ist ein Processsegment immer mit einem oder mehreren Productsegment(s) assoziert. Hierbei ist das Processsegment als Abbild eines Teils des Produktionsprozesses, als eine Art Schablone für einen Produktionsschritt, zu sehen. Der Produktionsschritt für ein physisches Produkt wird dabei dann durch das Productsegment definiert.

2.3.3.1 Personnel capability

Der Begriff Personnel capability umschreibt all jene Fähigkeiten wie Qualifikation, Arbeitsleistung oder Anzahl der Arbeiter, die einem Processsegment bzw. einem Productsegment aus dem Pool der Personalressourcen eines Unternehmens zur Verfügung stehen müssen um dieses erfolgreich durchlaufen zu können. Der Begriff steht synonym für den Bedarf an Humanressourcen eines Process- bzw. Productsegments.

2.3.3.2 Equiment capability

Dieser Begriff umschreibt die Anforderungen wie etwa Durchsatz oder Verfügbarkeit, die die Produktionsinfrastruktur einem Processsegment bzw. einem Productionsegment zur Verfügung stellen können muß, damit dieses erfolgreich durchlaufen werden kann.

2.3.3.3 Material capability

Dies ist der Sammelbegriff für sämtliche Eigenschaften hinsichtlich Qualität, Menge, Mischverhältnisse oder ähnlichem, die die in einem Produktionsschritt zu verarbeitenden Materialien erfüllen müssen um eine erfolgreiche Durchführung des selbigen zu gewährleisten. Wichtig ist in diesem Zusammenhang, dass auch informationsverarbeitende Produktionsschritte Anforderungen hinsichtlich der Qualität und Verfügbarkeit der zu verarbeitenden Daten / Informationen haben können, die kritisch für den zu erzeugenden Output sind.

2.3.4 Abbildung in Produktionsprozesse

Basierend auf den obig angeführten Definitionen von ISA-95 lassen sich für reale Produktionsprozesse und allgemeine Geschäftsprozesse bzw. deren Segmente folgende Eigenschaften ableiten:

- Ein Prozess benötigt Steuerinformation, die aus der Product definition information und der Production scheduling information gebildet wird.
- Ein Prozess hat Anforderungen hinsichtlich der Verfügbarkeit bestimmter personeller Ressourcen.
- Ein Prozess hat Anforderungen hinsichtlich der Verfügbarkeit bestimmter infrastruktureller Ressourcen
- Ein Prozess hat Anforderungen hinsichtlich der Verfügbarkeit des Inputs, den er zu verarbeiten hat.

Ist eine dieser Eigenschaften, die zugleich Voraussetzung für eine erfolgreiche Prozessdurchführung ist, nicht erfüllt, so steht der / das betroffene Prozess(segment) still.

Die obig aufgelisteten Eigenschaften der Prozesssegmente sind nach exakter Auslegung der Definitionen in ANSI ISA95 eine Kombination aus Elementen des Productsegment und des Processsegment. Dies ist auch dadurch bedingt, dass diese beiden Datenobjekte eine sehr enge Verbindung haben, worauf auch in ISA-95[28] explizit hingewiesen wird.
In weiterer Folge wird aber aus Gründen der Einfachheit nur der Begriff des Prozesssegments synonym für beide ISA-95 Datentypen verwendet.

28 Siehe ISA-95 Part 1 Figure 13

2.3.5 Purdue Referenzmodell

Das Purdue Referenzmodell stellt eine Möglichkeit dar, das Prozessstratum einer Organisation hierarchisch zu strukturieren, wobei das Purduemodell die hierarchischen Grenzen zwischen den Subsystementeilen auf Basis von betriebswirtschaftlichen und unternehmens-organisatorischen Kriterien zieht. Ein derartiges Modell für ein Unternehmen mit nur einem Produktionsstandort sieht dabei wie in Abb. 7 dargestellt aus.

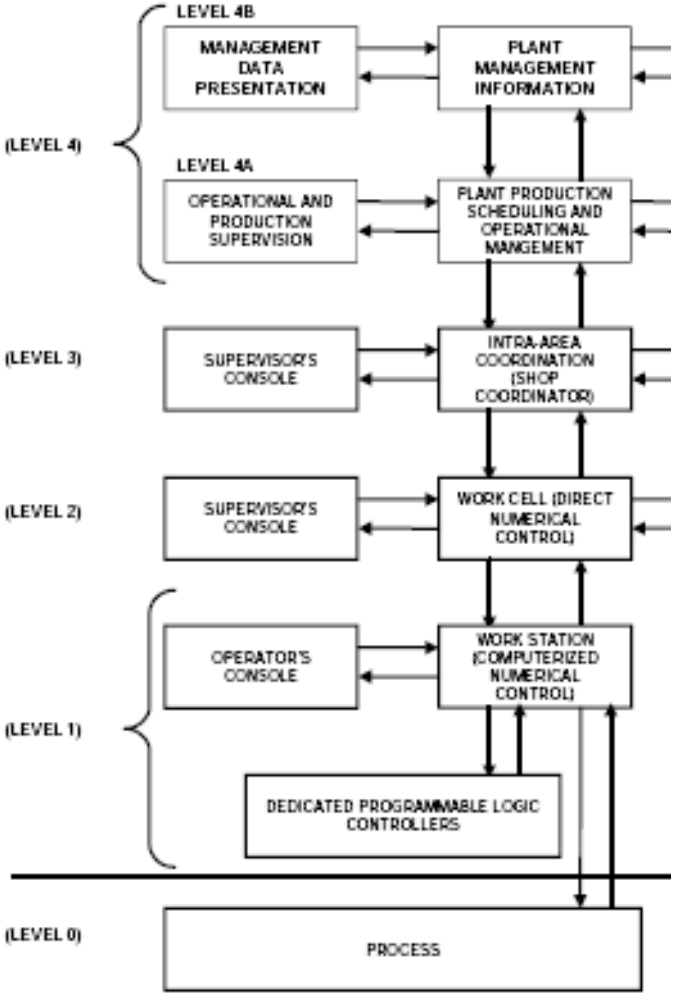

Abbildung 7: Prurdue Refernezmodell für ein Produktionsunternehmen

Auf den ersten Blick ist hierbei schon die Ähnlichkeit dieses Ansatzes zu den hierarchischen Multiebenensystemen, wie sie von Mesarovic beschrieben wurden, sichtbar. Die einzelnen Ebenen (Levels) sind dabei wie folgt zu interpretieren.

Level 4 (a & b):

Auf Level 4 sind jene Funktionen und Aufgaben angesiedelt, die der Koordination des gesamten Werkes bzw. des gesamten Unternehmens dienen. Das bedeutet, Level 4 repräsentiert die Gesamtunternehmenssteuerung.

Level 3:

Level 3 umfasst die Aufgaben, die mit der Koordination eines gesamten Produktionsbereichs

verbunden sind.

Level 2:

Level 2 repräsentiert einen Arbeitsbereich wie etwa eine Dreherei oder eine Fräserei.

Level 1:

Level 1 entspricht einem einzelnen Arbeitsplatz bzw. repräsentiert eine einzelne Maschine

Level 0:

Auf Level 0 befinden sich im Purduemodell die physischen Prozesse. Aus dieser Hierarchie ist auch ersichtlich, daß die physische Ebene nur durch Vorgänge auf Level 1 manipuliert werden kann.

2.4 ISO 27002

Der Standard ISO 27002 [ISO 27002] entspricht dem ISO 17799-2005 (Information technology – Code of practice for information security management) der International Organisation for Standardisation und ist der Nachfolger des ehemals rein britischen Standards BS 7799. Dieser Teil entspicht dabei - einem Referenzhandbuch gleich - einer Sammlung von Best Practices zum Thema Management von Informationssicherheit. Der Standard ISO 27002 ist eng mit dem ISO 27001 [ISO 27001] (ehemals ISO 17799-2) verbunden, in dem festgelegt ist welche der Best Practices eine Organisation, die sich nach ISO 27001 zertifizieren lassen möchte, umsetzen muß.
Betrachtet man nun die Aktivitäten, die ein Unternehmen im Rahmen eines Risikomanagements nach COSO2 durchführen sollte, so liefert der ISO 27002 durch die taxative und relativ vollständige Auflistung[29] von Maßnahmen zur Verbesserung von Informationssicherheit sehr wertvolle Vorschläge für Sicherungsmaßnahmen, die im Rahmen der Aktivität *Control Activities* etabliert werden können. Da viele Sicherungsmaßnahmen in der Infromationssicherheit direkt mögliche Ereignisse addressieren, ist der Standard aber auch eine gute Basis zur Identifizierung von möglichen schädlichen Ereignissen im Rahmen der Aktivität *Event Identification*. Der Umstand, dass der Standard in einem Dokument eine große Breite an Themen bis in einen relativ hohen Detaillierungsgrad abdeckt, reduziert seine Lesbarkeit massiv, weshalb nun folgend die wichtigsten Inhalte kurz, primär anhand der Ziele (control objectives) der einzelnen Kapitel, dargestellt werden sollen.

2.4.1 5- security policy

Die security policy oder Informationssicherheitsrichtline ist das hierarchisch oberste Dokument zum Umgang mit Informationssicherheit, dass für die Mitarbeiter verständlich darlegen sollte warum Informationssicherheit notwendig ist bzw. welche anderen strategischen Unternehmensziele sie unterstützt, und dass die Informationssicherheit ein Anliegen des obersten Managements ist. Weiters sollte sie kurz darlegen, wer im Unternehmen mit dem Management der Informationssicherheit betraut ist und wie das Reporting im Falle von Informationssicherheitsereignissen abzulaufen hat. Ebenfalls enhalten sein sollen Referenzen auf jene Dokumente, die die allgemeinen Vorgaben der Richtline im Detail regeln.
Die Informationssicherheitsrichtline hat drei relevante Anküpfungspunkte an COSO2:

- Der Wunsch des Managements dass Informationssicherheit im Unternehmen gelebt werden soll, und der Umstand dass dies durch die Richtline kundgetan wird, sind Teil der Aktivität *Internal Environment*.
- Die Definition der Informationssicherheitsziele des Unternehmens und die Anküpfung der Richtline an die anderen relevanten Unternehmensziele sind Teil der Aktivität *Objective Setting*.
- Die Ausarbeitung des Reportings von Informationssicherheitsereignissen ist Teil der Aktivität *Information and communication*.

2.4.2 6- Organizational security

Dieses Kapitel des ISO 27002 behandelt die organisatorischen Grundlagen, die in einem

29 vollständig bezieht sich hierbei auf die Adressierung aller relevanten Themen, und nicht auf eine exakte Defnition dessen wie im Detail eine spezielle Sicherungsmaßnahme ausgestaltet werden muss. Die Adaptierung des Standards muß von den jeweiligen Unternehmen selbst durchgeführt werden.

Unternehmen notwendig sind um ein effektives Informationssicherheitsmanagment zu ermöglichen. Zu diesen Grundlagen zählen u.a. die Etablierung eines Managementforums[30], dass sich den Informationssicherheitsthemen widmet, oder die Definition wie die Informationssicherheit zwischen den einzelnen Organisationseinheiten koordiniert werden soll. Ebenso Teil dieser Grundlagen ist die Festlegung, wer mit den einzelnen Aufgaben im Rahmen der Informationssicherheit betraut ist. Weiters bedarf es grundlegender Regelungen, wie die Autorisierung für den Zugriff auf die IT-Infrastruktur erfolgen sollte, oder in welcher Art und Weise auf externes Know-How zurückgegriffen werden kann, und wie die Schnittstelle zu den Behörden im Falle von legislativ relevanten Ereignissen ausgestaltet sein soll.
Zwei ebenfalls wichtige Punkte sind Regelungen zu Zugängen von Dritten[31] und grundlegende Vorgaben, wie mit outgesourcten Komponenten bzw. den Outsourcingpartnern hinsichtlich Informationssicherheit umgegangen werden soll.

2.4.3 7-Asset management

Dieses Kapitel zeigt Vorgehensweisen, die zum Schutz der für das Unternehmen wichtigen Werte angewandt werden können. Im wesentlichen geht es darum für jedes relevante Asset einen Verantwortlichen (owner) zu definieren, es in ein Inventar aufzunehmen und zu definieren, unter welchen Voraussetzungen und auf welche Weise dieses Asset verwendet werden darf (acceptable use policy / AUP). Ein weiterer Punkt in diesem Kapitel widmet sich der korrekten Klassifikation und der AUP von Daten.[32]

2.4.4 8- Personnel security

Dieses Kapitel behandelt die Aspekte der Informationssicherheit, die sich aus dem Umgang jedes einzelnen Mitarbeiters mit den jeweiligen Informationen ergeben. So wird auf die Notwendigkeit der Etablierung eines Rollenkonzeptes, das den Zugang zu den Informationen für jeden Einzelnen regelt, ebenso eingegangen wie auf grundlegende Dinge, die bei der Auswahl von Mitarbeitern vor allem in sensiblen Bereichen beachtet werden sollten.
Es sind hier auch die Punkte Schulung der Mitarbeiter, eventuell notwendige disziplinäre Maßnahmen und die für die korrekte Beendigung eines Dienstverhältnisses sinnvollen Prozessschritte enthalten.

30 und damit verbunden auch die Demonstration des Managements, dass Informationssicherheit ein wichtiges Ziel ist, dass vom Management mitgetragen wird.
31 dies inkludiert auch die Zugänge von Kunden zu den IT-Systemen
32 Daten sind zwar auch „nur“ Assets und unterliegen damit schon den allgemeinen Regelungen in diesem Kapitel, aber auf Grund ihrer besonderen Eigenschaften hinsichtlich Reproduktion und Weitergabe ist diese detailliertere Regelung im Allgemeinen sinnvoll.

2.4.5 9- Physical and environmental security

In diesem ISO 27002 Kapitel sind zwei große Themenbereiche enthalten.
Der erste befasst sich mit der Schaffung und Aufrechterhaltung von sicheren Zonen, die rund um die kritischen Komponenten errichtet werden sollen, um diese vor unbefugtem Zugriff oder Schaden zu bewahren. Hierein fallen Überlegungen zum Zutrittsschutz genauso wie der Schutz vor ungewollten Umwelteinflüssen wie Feuer oder Wasser.
Der zweite Themenbereich beschäftigt sich mit der Sicherheit der IT-Komponenten selbst. Dies umfasst wieder den Schutz vor schädlichen physikalischen Einflüssen oder Stromausfällen. Ebenfalls thematisiert wird die Wartung und die Regeln, nach denen die Wiederverwendung bzw. Entsorgung von IT-Komponenten ablaufen kann.

2.4.6 10- communications and operations management

Dieses Kapitel ist umfangreich und in manchen Punkten sehr detailliert. Im Kern geht es hier darum, einen aus Informationssicherheitssicht sicheren Betrieb der IT-Komponenten und einen sicheren Austausch von Informationen zu gewährleisten, wobei folgende Punkte bedacht werden sollten:

- Es ist sinnvoll eindeutige Prozeduren und Zuständigkeiten für den Betrieb der IT-Infrastruktur zu definieren und zu dokumentieren.
- Bei Leistungserbringungen durch Dritte sollten ebenfalls die grundlegenden IT-Sicherheitsvorkehrungen eingefordert und überwacht werden.
- Es ist sinnvoll Prozesse zur Kapazitätsplanung und Systemabnahme zu definieren um einen gleich bleibenden Qualitätslevel aufrechterhalten zu können.
- Es ist sinnvoll Regeln zu definieren, wie gegen Schadsoftware (z.B.: Viren,..) vorgegangen werden soll, bzw. wie im Falle eines Befalls mit Schadsoftware zu verfahren ist.
- Es sollten Regelungen bezüglich Backup von Daten und Regeln über die Auditierung von Zugriff auf diese Daten erstellt werden.
- Es sollten Regelungen erstellt werden, wie mit portablen Medien umgegangen werden soll, und es ist notwendig Regeln für den Austausch von Informationen und Software zu definieren.
- Es sollten Maßnahmen und Regeln definiert werden, wie das Netzwerk als grundlegende IT-Ressource geschützt werden soll.

2.4.7 11- Access control

Dieses Kapitel ist ähnlich umfangreich und detailliert wie das vorangegangene, und enthält Punkte, die bei der Erstellung von Vorschriften, die den Zugang zu Daten bzw. IT-Komponenten regeln, beachtet werden sollten. Dieses Kapitel umfasst dabei die folgenden Themenbereiche:

- Allgemeine geschäftliche Voraussetzungen, die den Zugriff auf Daten bzw. Infrastruktur rechtfertigen.
- Regelungen und Prozeduren bezüglich des Managements von Benutzerzugängen wie etwa Registrierung oder Prüfung von Benutzerrechten.
- Aufgaben und Sorgfaltspflichten der Benutzer wie etwa der Umgang mit deren Passwörtern.

- Zugangskontrollen und Authentifizierung im Netzwerk bzw. der Netzwerkinfrastruktur.
- Zugangskontrollen auf der Ebene der Betriebssysteme und der Applikationen
- Überwachung der Systemzugriffe und der Systemnutzung
- Punkte, die bei Mobile Computing oder Telearbeit berücksichtigt werden sollten.

2.4.8 12- IT-System Akquisition und Wartung

Kapitel 12 beinhaltet Best Practices, wie eine sichere Entwicklung und Wartung von IT-Systemen erreicht werden kann. Daß heißt, es werden auf der einen Seite Sicherheitsmerkmale angeführt, die in Applikationen enthalten sein sollten und schon beim Design der Applikationen berücksichtigt werden sollten. Andererseits beschäftigt sich dieses Kapitel mit Methoden, wie sichergestellt werden kann dass die Software während des Entwicklungsstadiums, oder zum späteren Zeitpunkt ihrer Weiterentwicklung, nicht beschädigt oder manipuliert wird.
Ebenfalls in diesem Kapitel enthalten sind Vorgaben, wie mit Systemschwächen (Vulnerabilities) umgegangen werden sollte.

2.4.9 13- Information security incident management

Dieses Kapitel kann direkt der COSO2 Aktivität *Information and Communication* zugeordnet werden. Hier sind im ISO 27002 Vorgaben enthalten, wie ein Berichtswesen für Informationssicherheitsereignisse bzw. erkannte Sicherheitsschwächen aussehen sollte. Weiters sind auch Vorgaben angeführt, wie diese Ereignisse abgehandelt werden sollten und wie Verbesserungen erreicht werden können um zukünftig ähnliche Ereignisse zu vermeiden.

2.4.10 14- business continuity management

In diesem Kapitel wird detailliert beschrieben, welche Schritte unternommen werden können um die Informationssicherheit in die Pläne zur Aufrechterhaltung der Geschäftsprozesse miteinzubeziehen. Es wird auch darauf eingegangen, welche Punkte solche Pläne im Allgemeinen berücksichtigen sollten.

2.4.11 15- complinace

Dieses Kapitel unterstützt im ersten Teil das Ziel eines Unternehmens, möglichst wenige relevante legislativen Vorgaben zu verletzen. In diesem Zusammenhang sollte es einen Prozess geben, der regelmäßig die relevanten legislativen Vorgaben für das Unternehmen identifiziert, wobei im speziellen die folgenden Vorgaben beachtet werden sollten:

- Urheberrechte
- Legislativ vorgeschriebene Geschäftsdaten
- Datenschutzvorgaben
- Vorgaben bezüglich Mißbrauch von Unternehmensinfrastruktur
- Einschränkungen bei der Anwendung von Verschlüsselungssoftware

Der zweite Teil dieses Kapitels gibt Anregungen, wie die Einhaltung von internen Vorschriften kontrolliert werden kann und welche Punkte im Zuge der Durchführung von solchen Kontrollen (Audits) beachtet werden sollten.

2.4.12 ISO 27002 vs. COBIT

Zum Abschluß dieses Kapitels soll noch oberflächlich[33] auf den zweiten weithin etablierten Standard in punkto Management von Informationssicherheit eingegangen werden, nämlich auf die Control Objectives for Information and related Technology [COBIT]. Cobit ist ähnlich zu COSO2 ein Framework, welches aber im Gegensatz zu COSO2 eine Detailebene tiefer liegt und als Basis für IT-Governance dienen kann. In der Tat ist auch die nahe Verwandtschaft[34] der beiden Frameworks in den darin angewandten Konzepten ersichtlich.
Cobit wird von den Autoren selbst als high level Framework beschrieben, dass primär für die Anwendung durch das Top- und das IT Management sowie für IT-Auditoren ausgelegt ist. Es kann daher für den IT-Bereich als Bindeglied zwischen COSO2 und der ISO 27002 gesehen werden. Dies ist insoferne begründet, da COBIT selbst eine Vielzahl von Prüfzielen (Control Objectives) auf Prozessebene definiert, die auch einige Messkriterien[35] enthalten die eine Beurteilung wie gut die Zielerreichung schon ist ermöglichen. Aber es sind keine konkreten Maßnahmen definiert, wie diese Ziele erreicht werden können. Die Definition dieser konkreten Maßnahmen wird größtenteils im ISO 27002 durchgeführt.
Die wichtigsten Inhalte und Konzepte von COBIT sind die folgenden:

COBIT ist ein Prozessframework, in welchem 34 generische IT-Managementprozesse definiert werden. Jeder dieser 34 Prozesse ist dabei einer der vier folgenden Prozessdomänen zugeordnet:

- Planung und Organisation: In dieser Domäne sind all jene Prozesse zusammengefasst, welche sich mit der zukünftigen Entwicklung der IT-Infrastruktur eines Unternehmens befassen. Darin inkludiert sind auch Vorgehensweisen wie die IT-Infrastruktur an die allgemeinen Geschäftsziele angepasst werden kann, bzw. welche IT-Ziele zur bestmöglichen Unterstützung der Geschäftsprozesse abgeleitet werden können. D.h., die Prozesse dieser Domäne definieren die IT-Strategie.
- Beschaffung und Umsetzung: Die Prozesse dieser Domäne sind für die Implementierung der IT-Strategie zuständig. Typischerweise werden durch diese Prozesse die IT-Lösungen die zur Umsetzung der IT-Strategie notwendig sind, identifiziert und aufgebaut. D.h. hier sind die meisten Prozesse, derer sich IT-Projekte bedienen, angesiedelt.
- Delivery und Support: Diese Domäne beinhaltet jene Prozesse, die für den Betrieb der IT-Dienste und für die Unterstützung der Kunden dieser Dienste notwendig sind.[36] In diesen Bereich fallen auch das Securtiy- und Continuitymanagment.
- Überwachung und Evaluierung: Die Prozesse dieser Domäne haben die Aufgabe die Prozesse der anderen Domänen hinsichtlich ihrer Effektivität und ihrer Konformität gegenüber legislativen und internen Vorgaben zu überwachen. Ebenso soll hier geprüft werden, ob die IT-Strategie und die daraus abgeleiteten Aktivitäten auch wirklich die übergeordneten Geschäftsziele unterstüzen.

Ein weiteres wichtiges Konzept von COBIT sind die sog. Maturity levels, die ein generisches sechsstufiges Konzept darstellen, das nahezu jedem Anwender von COBIT die Möglichkeit

33 Eine erschöpfende Einführung in COBIT würde leider den Rahmen dieser Arbeit sprengen und ist auch auf Grund der guten Aufarbeitung und Darstellung der Themen im Standard selbst nicht zielführend.
34 COBIT wurde ursprünglich aus den Ansätzen von COSO 1 abgeleitet.
35 Key Performance Indicators, Key Goal Indicators und Maturity levels
36 Im Detail wird diese Domäne sehr gut vom IT Infrastructure Library [ITIL], in dem eine Vielzahl von Best Practices für einen effektiven und effizienten IT-Betrieb aufgeführt werden, abgedeckt.

geben soll für jeden der definierten Prozesse selbst zu beurteilen, auf welcher Entwicklungsstufe die diesem Prozess äquivalenten unternehmensinternen Prozesse stehen. Die einzelnen Stufen sagen dabei etwas darüber aus, wie gut die Prozesse in der Lage sind die in sie gesetzten Erwartungen bzw. Ziele die sie realisieren sollen auch wirklich zu erreichen.

2.5 Probabilistic Risk Assessment

Probabilistic Risk Assessment (PRA), also die wahrscheinlichkeitstheoretische Bewertung von Risiken, ist eine Disziplin die Ihren Ursprung in den späten 1960-er Jahren in der US-amerikanischen Raumfahrt bzw. im Atomenergiesektor hatte. 1981 veröffentlichte die US-Atomenergiebehörde mit dem Fehlerbaum Handbuch [Vesely et al., 1981] ein Handbuch das als Basis eines modernen PRA gesehen werden kann. (vgl. [Bedford, 2001] Kap. 1.1) In weiterer Folge gab es eine Fülle von Veröffentlichungen, die die Konzepte der PRA auf andere Anwendungsgebiete übertrugen oder diese verfeinerten. Heute ist die wahrscheinlichkeitstheoretische Bewertung von Risiken die gängige Methode[37] um ein Risiko zu analysieren und zu beschreiben. Das PRA soll im Zuge seiner Durchführung folgende Fragen klären:

- Welche Ereignisse können eintreten?
- Wie wahrscheinlich ist es, dass solche Ereignisse eintreten?
- Gegeben ein solches Ereignis tritt ein, was sind die Folgen?

Wie aus den obigen Fragestellungen ersichtlich ist, liegt es in der Methodik des PRA zuerst zu untersuchen, welche Ereignisse eintreten können. Danach wird versucht das Risiko, das mit einem identifizierten Ereignis verbunden ist, durch folgende beiden Kennzahlen darzustellen:

- Die Auswirkungen, die der Eintritt eines Ereignisses bedingt.
- Die Wahrscheinlichkeit, dass das betrachtete Ereignis eintritt.

Das PRA ist dadurch ein Werkzeug, das auf eine nahezu unbegrenzte Fülle von Ereignistypen angewandt werden kann. So kann ein solches Ereignis genauso der Bruch einer Schweißnaht sein wie der Ausfall eines vergebenen Kredits. Welche Auswirkungen eines Ereignisses herangezogen werden um das Ereignis zu bewerten, ist von Anwendungsfall zu Anwendungsfall unterschiedlich. Im Zusammenhang mit der Bewertung von Betriebsrisiken sind dies meist die finanziellen Auswirkungen, die sich in Gewinnen bzw. Aufwänden niederschlagen.
Vor allem bei der Analyse von komplexen technischen Systemen ist es zur Bestimmung der Eintrittswahrscheinlichkeit eines Ereignisses auf grund mangelnder statistischer Daten meist notwendig die möglichen Ursachen für den Eintritt des untersuchten Ereignisses zu bestimmen, um in weiterer Folge die gesuchte Eintrittswahrscheinlichkeit aus der Eintrittswahrscheinlichkeit der ursächlichen Ereignisse herzuleiten. Zwei der im Rahmen des PRA definierten Methoden für eine nähere Systemanalyse sind die Fehlerbaumanalyse (FTA) und die Ereignisbaumanalyse (ETA), die in den folgenden beiden Kapiteln einführend beschrieben werden sollen.

37 Siehe [COSO 2], [COBIT], [ISO 27002],...

2.5.1 Event tree analysis

Die Ereignisbaumanalyse stellt einen Bottom UP Ansatz dar der, sich einer vorwärts gerichteten Logik bedient (siehe auch [Bedford, 2001] Kap. 6.1). Die Analyse startet dabei mit einem für das betrachtete System meist externen Ereignis, das einen abnormen Vorfall darstellt. Danach wird schrittweise der Einfluß dieses Ereignisses auf das System untersucht, indem es durch das System verbreitet wird. Dabei werden alle möglichen Einflüsse auf das System berücksichtigt und im Ereignisbaum festgehalten. Die Knoten in einem Ereignisbaum repräsentieren dabei entweder funktionierende oder ausgefallene Subsysteme oder das ganze System selbst.
Das folgende Beispiel untersucht ein elektrisches Leitungsnetz, in dem es einen Haupt- und einen Endverteiler gibt, in denen sich jeweils Leitungssicherungen befinden. Das zu untersuchende externe Ereignis, ist ein Kurzschluss in einem der an das Leitungsnetz angeschlossenen Geräte.

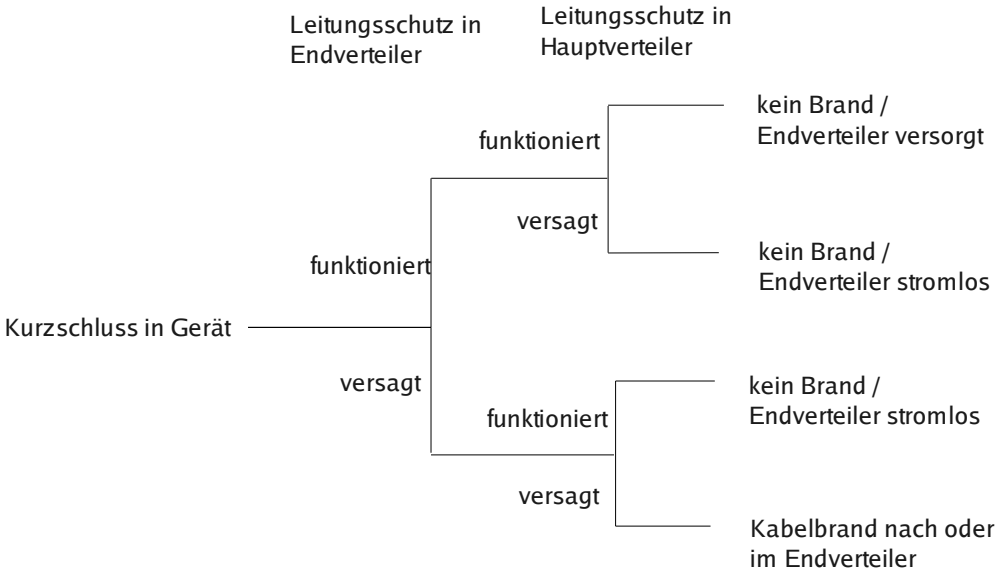

Abbildung 8: Ereignisbaum in einem elektrischen Leitungsnetz

Wie man gut erkennen kann, ist die gewünschte Funktionalität (alle nicht vom Kurzschluss betroffenen Stromkreise bleiben versorgt und der betroffene wird abgeschaltet) nur bei korrekter Funktionalität beider Sicherungsmaßnahmen gewährleistet. Sollte einer der Schutzschalter versagen oder zu empfindlich reagieren, so entsteht entweder ein Stromausfall, der größer als notwendig ist, oder im schlimmsten Fall sogar ein Kabelbrand im betroffenen Stromkreis.
Man kann ebenfalls erkennen, dass ein Ereignisbaum ebenso wie ein Fehlerbaum (siehe Kap. 2.5.2) eine bildhafte Darstellung einer Menge von Ereignissen ist, die durch eine Boolesche Logik verknüpft sind.

2.5.2 Fault tree analysis

Die Fehlerbaumanalyse ist im Gegensatz zur Ereignisbaumanalyse ein Top-Down Ansatz. Das bedeutet, dass von einem Topereignis ausgegangen wird, welches schon ein Schadereignis darstellt. Für dieses definierte Topereignis wird systematisch nach (Zwischen-) Ereignissen gesucht, die Ursache für dieses Topereignis sind; danach werden die möglichen Ursachen für diese nun gefundenen Ereignisse untersucht. Diese schrittweise Verfeinerung des Detaillierungsgrades wird so lange durchgeführt, bis die maximal gewünschte Auflösung erreicht ist, und die an dieser Grenze gefundenen Ereignisse nennt man Basisereignisse.
Beim Aufbau des Baumes kann es dabei zu Darstellungsproblemen kommen, wenn Systemkomponenten oder Subsysteme, die drei oder mehr Zustände annehmen können, in die binäre Logik des Baumes abgebildet werden müssen.
Das nun folgende Beispiel soll eine kurze Einführung in den Aufbau eines Fehlerbaumes darstellen, und berücksichtigt nur die grundlegendsten Schritte, die für den Aufbau und die Darstellung eines Fehlerbaumes notwendig sind. Der darüber hinaus interessierte Leser sei daher an dieser Stelle an [Bedford, 2001] Kap. 6 verwiesen.
In Fortführung des obigen Beispieles eines elektrischen Leitungsnetzes soll uns nun das Ereignis Stromausfall im Endverteiler als Topereignis interessieren, da wir annehmen dass dies auch ein für ein Unternehmen wichtiges System ausser Betrieb setzt. Ein möglicher Fehlerbaum könnte also wie folgt aussehen:

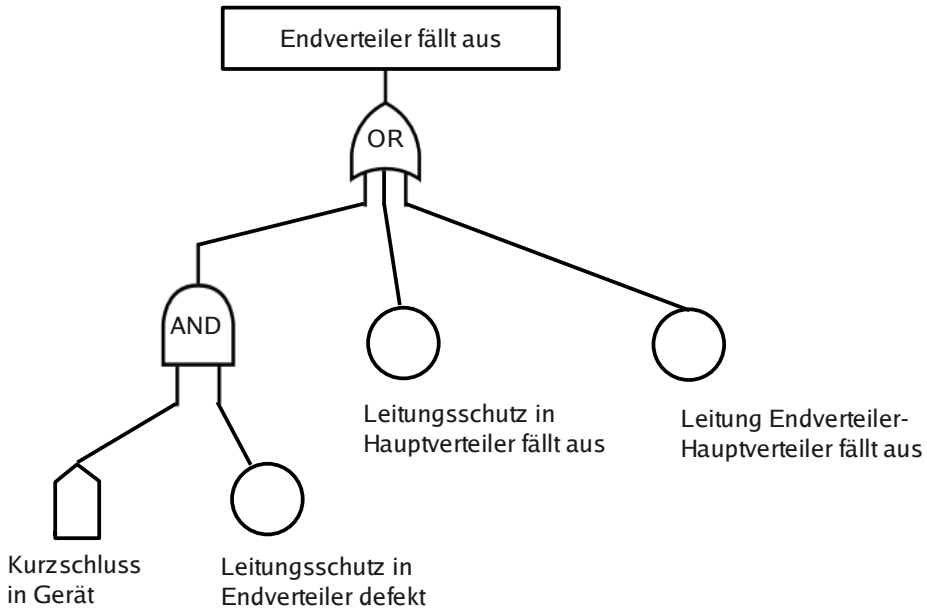

Abbildung 9: Fehlerbaum eines elektrischen Leitungsnetzes

Beginnend vom Endverteiler kann der Baum wie folgt gelesen werden: Der Endverteiler fällt aus, wenn sein Versorgungsstrang unterbrochen wird (dieses Ereignis wurde der Einfachheit halber nicht explizit eingetragen). Der Versorgungsstrang kann unterbrochen sein wenn entweder die Leitung physikalisch ausfällt oder der Leitungsschutz im Hauptverteiler auslöst. Dieses Auslösen des Leitungsschutzes im Hauptverteiler kann entweder durch einen Defekt desselbigen ausgelöst sein (dies ist ein Basisereignis welches nicht näher untersucht wird) oder durch eine Überlast am Versorgungsstrang zum Endverteiler. Diese Überlast selbst kann wiederum durch einen Kurzschluss in einem Gerät (als externes Ereignis) in Verbindung mit einem defekten (nicht auslösenden) Leitungsschutz im Endverteiler entstehen. Wenn man nun die einzelnen Ereignisse in boolesche Variablen transformiert, kann der Fehlerbaum in einen

booleschen Term übergeführt werden.

Sei:

- Das externe Ereignis Kurzschluss in Gerät die Variable A
- Das Basisereignis Leitungsschutz in Endverteiler defekt die Variable B
- Das Basisereignis Leitungsschutz in Hauptverteiler fällt aus die Variable C
- Das Basisereignis Leitung Endverteiler – Hauptverteiler fällt aus die Variable D
- Das Ereignis Endverteiler fällt aus die Variable E

so ergibt sich der folgende boolesche Term: $E=(A\cdot B)+C+D$

Für die Konstruktion des Baumes wurde im obigen Beispiel nur auf die folgenden grundlegenden Symbole zurückgegriffen.[38]

Zwischenereignis

ODER Verknüpfung

UND Verknüpfung

Basisereignis das bereits einen Fehler auf unterster Ebene darstellt.

Externes Ereignis, das von außen auf das betrachtete System einwirkt.

38 [Bedford] S. 107 enthält eine vollständige Liste.

2.5.3 Critical path analysis

Die Analyse des kritischen Pfades (CPA) ist eine PRA Methode, die sich vor allem im Projektmanagement etabliert hat. Die Grundlage der CPA[39] ist dabei der vollständige Netzplan aller in einem Projekt notwendigen Aktivitäten inklusive deren voraussichtlicher Dauer. Der kritische Pfad in einem Projekt ist dann jene Abfolge von Aktivitäten, die die längste kumulierte Dauer hat. Während der CPA wird für jede Aktivität jener Zeitpunkt, zu dem die Aktivität bedingt durch die Ausführung der vorhergehenden Aktivitäten frühestens abgeschlossen werden kann[40], ebenso wie der Zeitpunkt, zu dem die Aktivität abgeschlossen sein muß[41] um die Nachfolger nicht zu verzögern, erhoben und festgehalten. Das nachfolgende Beispiel (Abb. 10 bis 12) soll die Anwendung der CPA und die Berechnung der *earliest-* und *latest delivery time* (edt bzw ldt) anhand eines Netzplanes veranschaulichen.

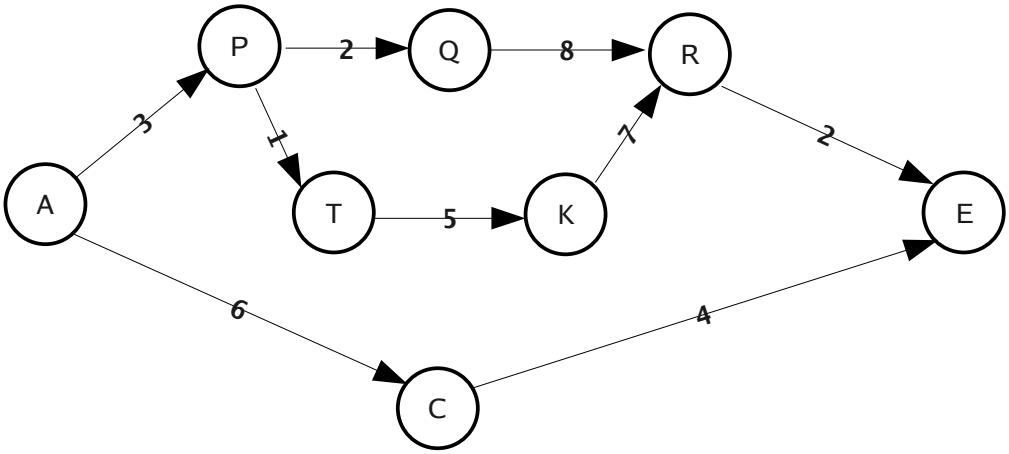

Abbildung 10: Kritischer Pfad, Schritt 1

Jeder Knoten im obigen Netzplan repräsentiert einen Meilenstein nach einer oder mehreren abgeschlossene(n) Tätigkeit(en). Die Ziffern auf den Kanten zwischen den Knoten geben die Dauer, die für die Durchführung der entsprechenden Tätigkeit notwendig ist, an. Auf Basis dieser Informationen kann mit Hilfe eines vorwärtsgerichteten Algorithmus der frühest

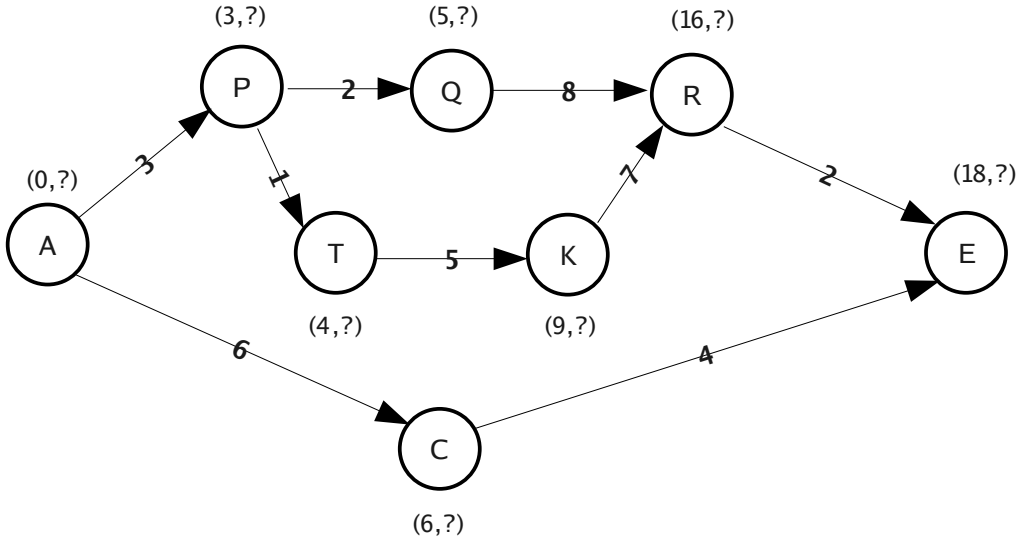

Abbildung 11: Kritischer Pfad Schritt 2

39 vgl. auch [Bedford, 2001] Kap. 15.2
40 earliest delivery time
41 latest delivery time

mögliche Zeitpunkt für den Eintritt des jeweiligen Meilensteines errechnet werden.

In Abb. 11 wurden die edt's der jeweiligen Knoten als erste Ziffer in den Klammern bei den Knoten eingetragen. Die zweiten Ziffern repräsentieren die ldt's, die allerdings zu diesem Zeitpunkt noch nicht bestimmt sind.
Der vorwärts Algorithmus ist - wie man sehen kann - sehr einfach aufgebaut, es wird in jedem Knoten jeweils zur edt des Vorgängerknotens die Dauer der Transittätigkeit zum aktuell untersuchten Knoten hinzugezählt. Für den Fall, daß ein Knoten mehr als eine Transittätigkeit benötigt (z.B. Knoten R oder E) wird das Maximum der berechneten edt's für diesen Knoten eingetragen.

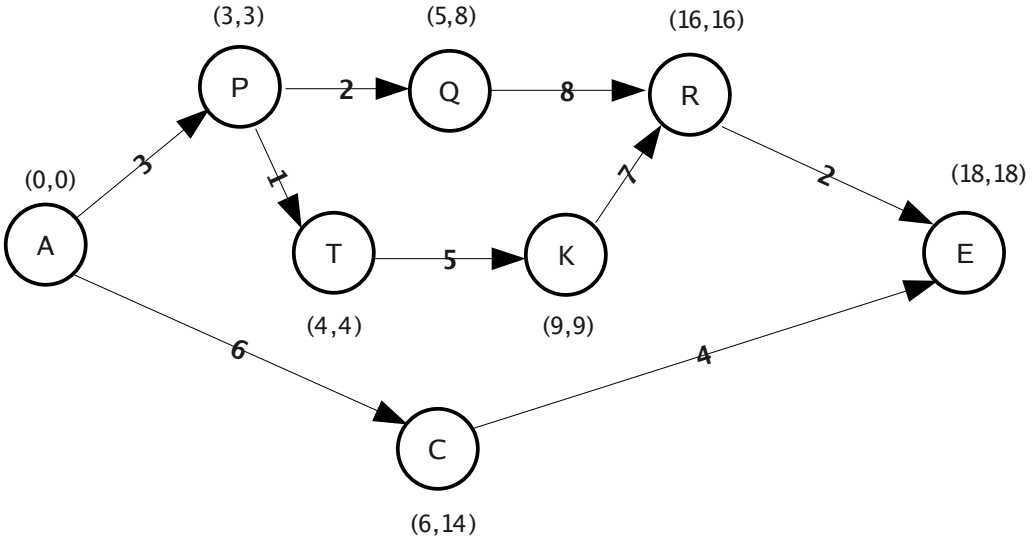

Abbildung 12: Kritischer Pfad Schritt 3

Nunmehr sind im Netzplan auch die ldt's der jeweiligen Knoten eingetragen, deren Berechnung wiederum relativ einfach ist. Es wird hierzu von der ldt des jeweiligen Nachfolgeknoten die Dauer der Transitaktivität zum aktuellen Knoten abgezogen, wobei im Endknoten mit einer ldt, die gleich der edt gesetzt ist, begonnen wird. Sollte ein Knoten mehrere Nachfolgeknoten besitzen (Knoten A bzw. P), so ist in diesem Falle das Minimum der berechneten ldts heranzuziehen. Der kritische Pfad ist nun jene Folge von Knoten im Netzplan, deren edt's und ldt's jeweils gleich sind. Dies ist also jene Folge von Aktivitäten, bei der jede Verzögerung einer der Aktivitäten, wie gering sie auch ausfallen möge, sofort eine Verzögerung des Projektabschlusses nach sich zieht. Der kritische Pfad bestimmt somit auch die Mindestdauer eines Projektes.

2.6 Das ISO OSI Modell

Das ISO Modell für Open Systems Interconnection [ISO/OSI] ist das Referenzmodell für den Aufbau von Kommunikationsnetzen zwischen Computersystemen. Das Modell hat folgende Struktur.

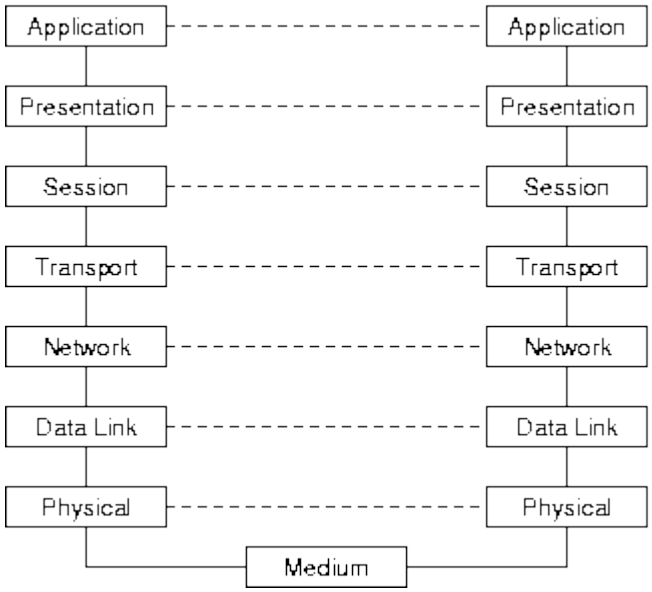

Abbildung 13: ISO / OSI Schichten

Für diese Arbeit relevant ist dabei weniger das exakte Zusammenspiel der sieben Protokollebenen oder deren genaue Funktionalität, vielmehr ist dieses Modell ein sehr gutes Beispiel für die Anwendung des Konzepts der Strata von Mesarovic. Das ISO OSI Modell ist auch eine gute Referenz für die Umsetzung des Konzepts der Kapselung von Funktionalität zwischen den Ebenen und der Abstraktion von Funktionalität durch die einzelnen Ebenen.[42] Diese beiden Konzepte haben sich allgemein im Design von komplexen IT-Systemen durchgesetzt, da sie erst die Komplexität beherrschbar machen.[43]

42 Die konzeptionelle Nähe des ISO OSI Modells zu den Strata von Mesarovic ist hauptsächlich dadurch bedingt, dass die beiden Konzepte Abstraktion und Kapselung ja wichtige Eigenschaften der einzelen Strata sind.

43 vgl. auch die Konzepte der objektorientierten Programmierung.

3 Ansatzpunkte

Orientiert sich der Leser bei der Etablierung eines Risikomanagement Systems an den Vorschlägen von COSO2, so ist einer der wichtigsten Schritte die Definition der Ziele eines Unternehmens bzw. einer Organisation. Da eine Risikoanalyse wesentlich einfacher und fundierter ist, wenn der Einfluss der jeweiligen Ereignisse auf die zu erreichenden Ziele betrachtet wird, ist die Zieldefinition auch bei der Behandlung von Informationssicherheits-risiken ein erster unabdingbarer Schritt. Deshalb beginnt dieses Kapitel mit eben jener Zieldefinition.

3.1 Zieldefinition / COSO2 Aktivität 2

Sämtliche Aktivitäten, die im Rahmen des Aufbaus der Informationssicherheit stattfinden oder deren Erhaltung dienen, haben die Aufgabe zumindest Teile der für die Abwicklung der Geschäftsprozesse notwendigen Infrastruktur zu sichern. Dies ist ein Grund, warum Informationssicherheitsziele einen großen Beitrag zur Sicherung der geschäftlichen Ziele eines Unternehmens leisten und somit implizit Teil dieser übergeordneten Ziele sind. Daher sollen an diesem Punkt diese Ziele, die sich aus Informationssicherheitsüberlegungen ergeben, näher betrachtet werden.
Als Basis für diese Betrachtung kann der ISO 27002 herangezogen werden, da in diesem Standard die Informationssicherheitsziele sehr griffig definiert sind. Es sind dies:

- Confidentiality: Die Sicherstellung der Vertraulichkeit von Informationen bzw. Daten.
- Integrity: Sicherstellen, dass die Informationen bzw. Daten nicht verfälscht oder anderweitig manipuliert werden können.
- Availability: Sicherstellen, dass die Informationen, Daten und Datenverarbeitungs-einrichtungen den berechtigten Personen (Prozessen) bei Bedarf auch in der benötigen Qualität zur Verfügung stehen.

Betrachtet man die gesamte IT- Infrastruktur eines Unternehmens näher, so kann man feststellten, dass sich die obig genannten Ziele auf die drei Teile dieser Infrastruktur nämlich Software, Hardware und Daten, applizieren lassen.
Das bedeutet: aus Infomationssicherheitssicht interessant sind die Software, die eine benötigte Funktionalität (Dienste / services) für einen Geschäftsprozess zur Verfügung stellt und die Daten, auf die diese Funktionalität angewendet wird. Die folgende in Tabelle 1 dargestellte Matrix gibt dabei an, welche der generischen Informationssicherheitsziele - auf die unterschiedlichen IT-Infrastrukturobjekte angewendet - aus Sicht des Autors Teil einer operativen Risikobetrachtung sein sollten.

Objekte / ***Informationssicherheitsziele nach ISO 27002***	***Hardware***	***Software***	***Daten***
Vertraulichkeit (confidentiality)	nein	nein[44]	ja
Integrität (integrity)	ja	ja[45]	ja
Verfügbarkeit (availability)	ja	ja	ja

Tabelle 1: Relevanz der Informationssicherheitsziele für ausgewählte IT Objektgruppen

Neben diesen sieben direkt aus den im ISO 27002 angeführten generischen Infomationssicherheitszielen ableitbaren Zielen, deren Erreichung im Detail für die jeweils betrachteten Objekte überprüft werden kann, ist es auch sinnvoll, die folgenden direkt von diesen Zielen unterstützten allgemeineren Unternehmensziele in eine Betriebsrisiko-betrachtung aufzunehmen.

- Erreichung der gewünschten Zuverlässigkeit bei der (Dienst-) Leistungserbringung
- Erhaltung des Vertrauens der Kunden in die erbrachte Leistung und die Art / Qualität der Leistungserbringung, bzw. Erhaltung eines guten Images des Unternehmens bei den Kunden.

Diese beiden Ziele sind insofern notwendig, als sie in der Außenwirkung, die das Unternehmen auf die Kunden hat, auf grund des Umstandes dass sie eine große Anzahl der obigen Informationssicherheitsziele subsumieren, teilweise besser zu messen sind als etwa die Vertraulichkeit einzelner Datenobjekte. Damit sind sie eine bessere Basis für die weiter unten dargestellte Risikoanalyse.

Wie vielleicht auch zu erkennen ist, sind die aus dem ISO 27002 abgeleiteten Ziele allesamt Ziele, die der Infrastruktur eines Unternehmens zugeordnet werden können; und die beiden allgemeinen Ziele sind Ziele aus der Prozessebene, die auf grund der Abhängigkeit der Prozesssegmente (siehe Kapitel 4.1 f.) von deren Infrastruktur auch von der Erreichung der Infrastrukturziele abhängen. In den folgenden Kapiteln wird nun eine Methodik vorgestellt, wie sich diese Abhängigkeiten auffinden und darstellen lassen.

44 Wenn Software als Asset / Endprodukt betrachtet wird, das nur einem eingeschränkten Kundenkreis zur Verfügung stehen soll, so ist sie als Datenobjekt zu behandeln, da Software im Kontext einer Betriebsrisikobetrachtung nur als Funktionalität bzw. Dienst gesehen werden sollte.

45 Integrität von Software ist insoferne wichtig, als dadurch sichergestellt ist, dass sie gemäß den Vorgaben, die zu ihrem Einsatz führten, funktioniert. Es ist daduch ebenfalls ein Faktor, der zu Vertaulichkeitsproblemen von Daten führen kann, ausgeschaltet (z.B. durch Trojaner,..)

3.2 Systemumwelt

Ruft man sich die Darstellung von COSO2 im Unternehmensregelkreis (siehe Kapitel 2.1.5) in Erinnerung, so wurde bereits festgehalten, dass es aus Sicht des Reglers sehr wichtig ist die grundlegenden Wirkungszusammenhänge zwischen den Störgrößen, den Stellwerten und den resultierenden Regelgrößen zu kennen, um entsprechend sinnvolle Stellwerte erzeugen zu können. Der erste Schritt zur Modellierung dieser Wirkungszusammenhänge ist die Definition der Systemgrenzen und der Interaktionspunkte des Systems mit seiner Umwelt, die nun Teil dieses Kapitels ist.

Betrachtet man ein Unternehmen bzw. eine Organisation von außen - wie dies etwa ein Versicherungsunternehmen, bei dem das Unternehmen seine Betriebsrisiken versichern will, machen müsste - so ergibt sich grundsätzlich folgendes vereinfachte Bild, das einer Black Box Betrachtung gleich kommt.

Abbildung 14: Ereignisse im Kontext des Systems Unternehmen

Dieses eher einfache Modell beinhaltet eine Menge von bekannten, für das Unternehmen nicht beeinflussbaren und somit exogenen Ereignissen, die auf das Unternehmen einwirken. Als Resultat der Interaktion der exogenen Ereignisse mit dem Unternehmen entstehen Ereignisse deren Wirkung über den Systemhorizont hinaus reicht.[46] Diese vom Unternehmen zumindest teilweise beeinflussbaren Ereignisse seien als externe[47] Ereignisse benannt.

Kommt man zurück zu dem Versicherungsunternehmen, das ein Anbot zur Versicherung der Betriebsrisiken legen soll, so ist in einem ersten Schritt für diesen Versicherer vor allem das Gesamtschadenaufkommen der Betriebsrisiken in einer Periode (meist ein Jahr) von Interesse. Dieses Gesamtschadenaufkommen ist prinzipiell ohne tiefgehende Überlegungen zur Unternehmensarchitektur meßbar, indem über mehrere Perioden alle Schäden, die durch Abweichungen von den Sollabläufen[48] entstanden sind festgehalten werden. Dadurch ist es prinzipiell auch für das Management des Unternehmens selbst möglich festzustellen, ob sich zum Beispiel der erwartete Schaden pro Periode innerhalb des Risikoappetits (siehe COSO2) des Managements befindet, bzw. ob sich der erwartete Schaden der dem Schadenaufkommen zugrundeliegenden Schadenverteilung auf einem vom Management definierten Konfidenz-niveau innerhalb des jeweiligen Risikoappetits befindet.

46 Diese Definition ist nicht sehr trennscharf, beinhaltet aber all jene Ereignisse, die finanzielle Auswirkungen auf das Unternehmen selbst oder auf von dem Unternehmen abhängige Stakeholder haben.

47 Die Begriffe exogen und extern sind aus einer Systemsicht grundsätzlich synonym, die gleichzeitige Anwendung beider Begriffe dient hier jedoch dem Zweck um zwischen den vom Unternehmen beeinflussbaren Ereignissen und jenen, die dies nicht sind, unterscheiden zu können. Die Tatsache, daß Ereignisse aus der externen Menge auch Einfluss auf die Menge der exogenen haben können wurde hier nicht dargestellt sie sei aber angenommen.

48 Diese Abweichungen von den Sollabläufen sind als eben jene externe, auf dritte (meist Kunden) wirksame Ereignisse zu verstehen.

In der Praxis ergeben sich bei dieser sehr oberflächlichen Betrachtung allerdings meist zwei schwerwiegende Probleme.

- Viele Ereignisse (vor allem jene mit hohen Schadenpotentialen) haben extrem geringe Eintrittswahrscheinlichkeiten und sind daher nicht oder nur durch sehr lang zurückreichende Messreihen[49] in der Gesamtschadenverteilung berücksichtigt. Aus diesem Grund ergibt sich eine Tendenz, dass der Erwartungsschaden systematisch unterschätzt wird.
- Das Unternehmen kennt die Hebeln die ihm die Umsetzung einer entsprechenden Reaktion auf die identifizierten Risiko- und Schadenpotentiale ermöglichen würden, nicht, da es über die internen Zusammenhänge bzw. Ursachen, die zu diesen Schäden führen, zu wenig Information besitzt.

Für einen Versicherer ist vor allem der erste Problempunkt sehr schwerwiegend, da es in den meisten Fällen auf Grund einer fehlenden Datenbasis nicht möglich ist mit Hilfe statistischer Methoden Aussagen über die Ursache – Wirkungszusammenhänge zwischen den relevanten exogenen Ereignissen (Faktoren) und den durch sie bedingten externen Ereignissen zu treffen. Dadurch ist für einen Versicherer aber auch der Weg zu einer adäquaten Tarifierung des Betriebsrisikos für das jeweilige Unternehmen versperrt.

Daher ist es sinnvoll, für ein Risikomanagment nach COSO 2 dieses Black Box Modell wie folgend dargestellt zu verfeinern und bis zu einem vom Management für sinnvoll erachteten Detaillevel die strukturellen und funktionalen Eigenschaften des Unternehmens zu berücksichtigen.

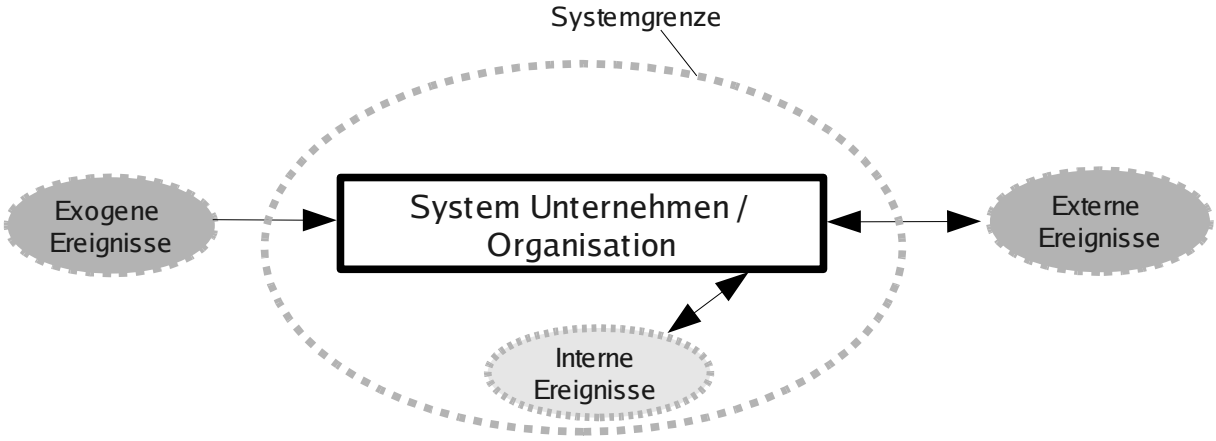

Abbildung 15: Erweiterte Systemumwelt

Die hier dargestellte Erweiterung des Black Box Ansatzes beinhaltet nun auch die innerhalb des Systems auftretenden internen Ereignisse. Interne Ereignisse können dabei sowohl Folgeereignisse von exogenen Ereignissen, als auch gewöhnliche interne Ereignisse, wie etwa ein Fehlverhalten eines Mitarbeiters sein, die keinen exogenen Ursprung haben.
Die hier gezeigte grobe Sicht auf das System dient als Ausgangsbasis für die weiteren Überlegungen.

49 die selten in der benötigten Qualität existieren

3.3 Risikomodellierung

Nach der Definition der Informationssicherheitsziele und der einführenden Darstellung der die Zielerreichung beeinträchtigenden Ereignisse, in den vorangegangenen Unterkapiteln, soll in diesem Unterkapitel nun der dieser Arbeit zugrunde liegende Risikomodellierungsansatz dargestellt werden.

Betriebsrisiken zählen zu jenen Risiken, die im Falle des Eintritts eines solchen Ereignisses keine Gewinnpotentiale, sondern nur Verlust- bzw. Schadenpotentiale haben. Diese Risiken werden allgemein als reine Risiken bezeichnet. Reine Risiken sind seit je her das Geschäft der Versicherungsbranche, daher kommen die gängigen mathematischen Ansätze zur Modellierung dieser Risiken auch aus diesem Bereich. Da nur selten genaue statistische Informationen zu den einzelnen Informationssicherheitsrisiken, die eine Untermenge der Betriebsrisiken darstellen, vorhanden sind, bietet sich zu deren Modellierung das kollektive Modell aus der Sachversicherungsmathematik an. Mit Hilfe des kollektiven Modells wird im Allgemeinen das Risikoportfolio des Versicherungsunternehmens als Ganzes, d.h. unter Vernachlässigung der Einflussfaktoren der einzelnen Risiken, betrachtet. Dieser Ansatz ist notwendig um festzustellen, ob der Versicherer unter Berücksichtigung all seiner Verträge (Risiken) auf Dauer wirtschaftlich überlebensfähig ist. Eine ähnlich Aussage ist im Rahmen von COSO 2 auch für ein normales Unternehmen von Interesse, da dieses ja langfristig den Erwartungsschaden all seiner Risiken unter seinen Risikoappetit senken will.
Im kollektiven Modell wird dabei der Gesamtschaden einer Periode als Zufallssumme der Schadenhöhen der jeweilig in dieser Periode aufgetretenen Schäden berechnet, wobei sowohl die jeweilige Schadenhöhe als auch die Anzahl an Schäden in einer Periode eine Zufallszahl darstellt. Die Verteilung des Gesamtschadens kann, wenn *n* die Anzahl der Schäden der Periode darstellt, aus der n-fachen Faltung der jeweiligen Schadenhöhenverteilungen errechnet werden (siehe hierzu u.a. [Mack, 2002]).
Im Gegensatz zum kollektiven Modell könnte bei entsprechender Informationslage der Gesamtschaden einer Periode bzw. dessen Verteilung auch über das sog. individuelle Modell errechnet werden. Im individuellen Modell wird davon ausgegangen, dass jedes Einzelrisiko im Portfolio des Versicherers bekannt ist und sich die Gesamtschadenverteilung aus der Summe der Verteilungen der Einzelrisiken zusammensetzt.
Der nun folgende Risikobewertungsansatz orientiert sich primär am kollektiven Modell, da nur die eingetretenen Ereignisse (Schäden) berücksichtigt werden und über eine Schadenanzahl und eine Schadenhöhe in das Modell integriert werden.[50]

Die im Zuge des PRA zu treffende Aussage bezüglich der Eintrittswahrscheinlichkeit eines Ereignisses und dessen Schadenhöhe ergibt, wenn diese beiden Faktoren quantitative Aussagen repräsentieren und entsprechend kombiniert werden, einen erwarteten Schaden pro Periode, also den Erwartungswert des Schadens. Diese Information ist in der Schadenverteilung des jeweiligen Ereignisses über die Zeit enthalten, deren Approximation durch das langfristige Erheben der Schäden, die durch das jeweilige Ereignis in jeder Periode verursacht wurden[51], durchgeführt werden kann.

50 Im Gegensatz dazu würde ein Ansatz, der dem individuellen Modell folgt, wahrscheinlich versuchen das mit einem Unternehmenswert (wie einer Infrastrukturkomponente oder einem Datensatz) verbundene Risiko zu messen.

51 hierbei sind allerdings in Abweichung vom kollektiven Modellansatz auch sog. Nullschäden zu berücksichtigen, also jene Perioden in denen das Ereignis nicht auftrat.

3.3.1 Definition des Begriffs Ereignis

Bevor in diesem Kapitel näher auf die Bewertung von Ereignissen eingegangen werden kann, ist es wichtig den Begriff Ereignis selbst näher zu definieren. Im Kontext dieses Modells wird ein Ereignis als Nachricht über eine Zustandsänderung in der Systemumwelt, einem Prozess bzw. einer Infrastrukturkomponente des untersuchten Systems Unternehmen verstanden. (siehe hierzu auch die Kapitel 4.1 und 4.4, wo die Details zu beiden Domänen abgehandelt werden) Ein Ereignis kann dabei zusätzlich zu der Information, welche Zustandsänderung erfolgte, auch noch die Informationen über verschiedenste Ausprägungen der Intensität dieser Zustandsänderung, wie etwa die voraussichtliche oder tatsächliche Dauer udgl, tragen.

Hinsichtlich der durch Ereignisse anzeigbaren Zustandsänderungen gibt es natürlich keine grundsätzlichen Einschränkungen; der Fokus dieses Modells liegt allerdings auf jenen Zuständen, die eine Verschlechterung im Vergleich zu den Regelbetriebszuständen der Prozesse und der Infrastrukturkomponenten darstellen. Das bedeutet, dass in der Folge nur auf jene Ereignisgruppen eingegangen wird, die derartige für das Unternehmen als negativ zu bewertende Zustandsänderungen anzeigen.

3.3.2 Anzahl der Ereignisse

Zur Modellierung der Anzahl von Ereignissen in einer Periode wird angenommen, dass den Ereignissen ein Zählprozess N(t) zugrundeliegt.
Der Einfachheit der Veranschaulichung wegen werden im Folgenden die Konzepte des hier vorgestellten Modells stellvertretend für alle Zählprozesse anhand eines homogenen Poissonprozesses[52] gezeigt.

3.3.3 Schadenhöhe

Eine wichtige Eigenschaft von Ereignissen, die Informationssicherheitsziele bedrohen ist jene, dass diese nur sehr selten schon die Information über die durch sie verursachten Schäden tragen. Vielmehr ist es so, dass die einzige in praktisch allen Fällen vorhandene Information jene über die Intensität (meist in Form der Dauer) des Ereignisses ist. Damit ist es naheliegend die mit einem Ereignis verbundenen Schäden aus der jeweiligen Ereignisintensität abzuleiten. Das bedeutet: im Rahmen des Modells werden sämtliche Schadenpotentiale, wie etwa Serviceunterbrechungen, negative Reputation, Verletzungen von Rechtsvorschriften, die Gefährdung der physischen Sicherheit von Personen uvm.[53] in eine allgemeine Schadenfunktion $s = h(I)$ integriert, mit deren Hilfe die Kopplung zwischen Ereignisintensität I und Schadenhöhe s des Ereignisses erreicht wird.
Ein Beispiel für eine solche Schadenfunktion zeigt Abb. 16 .

52 Bei dem ja bekanntlich N(]tj-1, tj] ~ Poi(λ*(tj -tj-1)), also die Anzahl an Ereignissen im Intervall tj-1 bis tj, einer Poissonverteilung mit Intensität λ folgt.

53 Diese Liste beinhaltet nur einen Auszug der im ISO / IEC 27005 auf den Seiten 36 u.37 enthaltenen Schadenpotentiale.

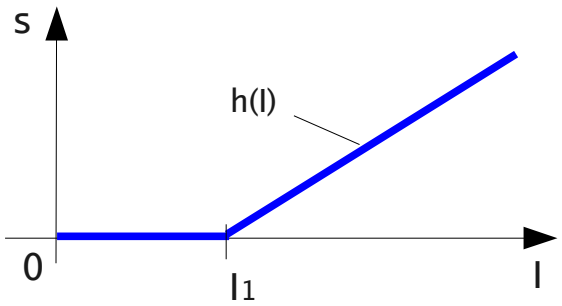

Abbildung 16: Beispiel für eine einfache Schadenfunktion

s... Schadenhöhe
I... Intensität eines Ereignisses
I1... Intensitätsschwelle ab der ein Schaden eintritt.

Wahrscheinlichkeitstheoretisch betrachtet findet durch die Anwendung der Schadenfunktion auf die Intensität eines Ereignisses eine Transformation einer messbaren Intensitätsverteilung, z.B.: G(I) zu der benötigten und direkt meist nicht messbaren Schadenhöhenverteilung F(s), statt.

Wählt man dabei eine Schadenfunktion, die im Intervall [0;∞] streng monoton ist, so hat dies den Vorteil, dass nur eine Transformation des Bildraumes von G zu F stattfindet, womit F auch analytisch bestimmt werden kann.[54]

Dieser für eine analytische Lösung wichtigen weiteren Eigenschaft der Schadenfunktion sollte bei der Wahl bzw. Approximation der Schadenfunktion großes Gewicht zukommen, da dies dem Ziel einer möglichst transparenten Berechnung des Risikos sehr zuträglich ist. Im Anwendungsfall sollte daher aus Sicht des Autors eher eine Funktion gewählt werden, die eine Umkehrfunktion besitzt, auch wenn dies zu Lasten der Anpassungsgüte an die empirisch erhobene Schadenfunktion gehen sollte.

Sollte es im Zuge der durchzuführenden Risikoanalyse nicht notwendig sein die Charakteristika der Schadenhöhenverteilung wie die Momente udgl. näher zu bestimmen - ist es also ausreichend die Schadenhöhen an einigen charakteristischen Quantilen wie etwa dem 5% oder 1% Quantil zu kennen, so genügt es die Umkehrfunktion G' der Ereignisintensitätsverteilung G(I) zu kennen - um mit deren Hilfe die Ereignisintensität der jeweiligen Quantile zu errechnen, die dann eingesetzt in die Schadenfunktion den Schaden am jeweiligen Quantil ergibt.
In einem solchen Fall ist es natürlich nicht notwendig eine streng monoton steigende Schadenfunktion zu benutzen.

3.3.4 Der Schadenprozess

Das Endziel einer Risikoanalyse ist es den Schadenprozess S(t) eines Ereignisses bzw. einer Gruppe von Ereignissen zu kennen und möglichst exakt nachbilden zu können.
Hierfür ist es notwendig die Informationen über den den Ereignissen zugrunde liegenden Zählprozess mit jenen Informationen über die Schadenhöhenverteilung der Ereignisse zu

54 Dies ist dadurch bedingt, dass mit:
G(I) ... Ereignisdauerverteilung des analysierten Infrastrukturereignisses
h(I) ... Schadenfunktion mit Umkehrfunktion h'(s)
F(s) ... Verteilungsfunktion des Schadens, sich aus G(h'(s)) für s > 0 ergibt und h'(s) nur dann für alle Schadenhöhen s eindeutig bestimmt ist wenn h(I) streng monoton ist. h(I) sollte überdies streng monoton steigend sein, da negative Schäden ja nicht Teil der Betrachtungen dieses Modells sind.

kombinieren.

In der Sachversicherungsmathematik werden zur Charakterisierung derartiger Schadenprozesse oftmals sog. Mixture Distributions (siehe dazu auch [Mack, 2002]), also Mischverteilungen, eingesetzt.
Folgt die Ereignisintensität (-dauer) zum Beispiel einer Gammaverteilung Ga(α, β)[55] und ist weiters die Schadenfunktion durch z.B.: s=2*I gegeben so folgt die Schadenverteilung ebenfalls einer Gammaverteilung mit den Parametern α_1 und β_1 .
Berücksichtigt man nun den homogenen Poissonprozess aus 3.3.2 so ist der resultierende Schadenprozess durch eine Compound Poissonverteilung $CP(\lambda; Ga(\alpha_1; \beta_1))$ charakterisiert. Ein möglicher Schadenprozess könnte dann etwa wie folgt aussehen:

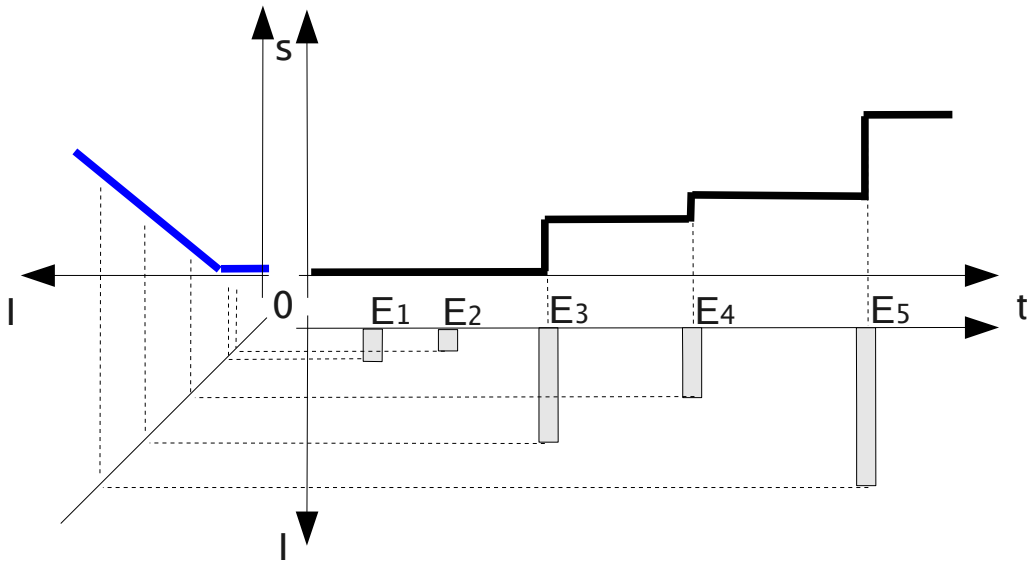

Abbildung 17: Zusammenhang zwischen Zählprozess, Schadenfunktion h(I) und Schadenprozess

Die obige Abbildung enthält die graphische Aufarbeitung, wie sich der gesuchte Schadenprozess aus dem Zählprozess der jeweiligen Ereignisintensität und der gewählten Schadenfunktion zusammensetzt.[56]
Der untere Zeitstrahl in Abbildung 17 beinhaltet dabei die Ereignisse E1 bis E5.
Die Ereignisse E1 und E2 haben aber eine zu geringe Intensität (kleiner als die Intensitätsschwelle) um einen Schaden verursachen zu können. Für die Ereignisse E3, E4 und E5 wurde die Schadenfunktion des Dienstes zur Bestimmung der jeweiligen Schadenhöhe benutzt.
Im oberen Teil ist der aus den Ereignissen resultierende Verlauf des Gesamtschadenprozesses abgebildet.

55 Diese Verteilung eignet sich i.a. gut für mittellange Ereignisdauern und hat den Vorteil, durch ihre zwei Parameter gut anpassbar zu sein. Des weiteren hat die Gammaverteilung die im Mischverteilungsmodell angenehme Eigenschaft, dass ihre Faltungen einfach zu bestimmen sind.

56 Hinweis: die in diesem Beispiel genutzte Schadenfunktion ist nicht streng monoton steigend, besitzt dafür aber eine Pufferzeit, wie sie in der Praxis oft vorkommen kann.

4 Integration in den Systemkontext

Nachdem im vorangegangen Kapitel die Grundlagen der in diesem Modell angewandten Risikomodellierung dargestellt wurden, soll dieses Kapitel dazu dienen das Risikomodell selbst in den Systemaufbau zu integrieren, bzw. soll nun auch gezeigt werden, welche Gründe zur obig gezeigten Ausgestaltung des Risikomodells geführt haben. Am Beginn dieses Kapitels wird das in Kapitel 3.2 eingeführte Systemmodell noch um eine Detailebene erweitert und in den Kontext der jeweils angewandten Konzepte gesetzt. Dadurch können die in diesem und dem nachfolgenden Kapitel näher beschriebenen Konzepte und deren Abhängigkeiten untereinander gut eingesehen werden.

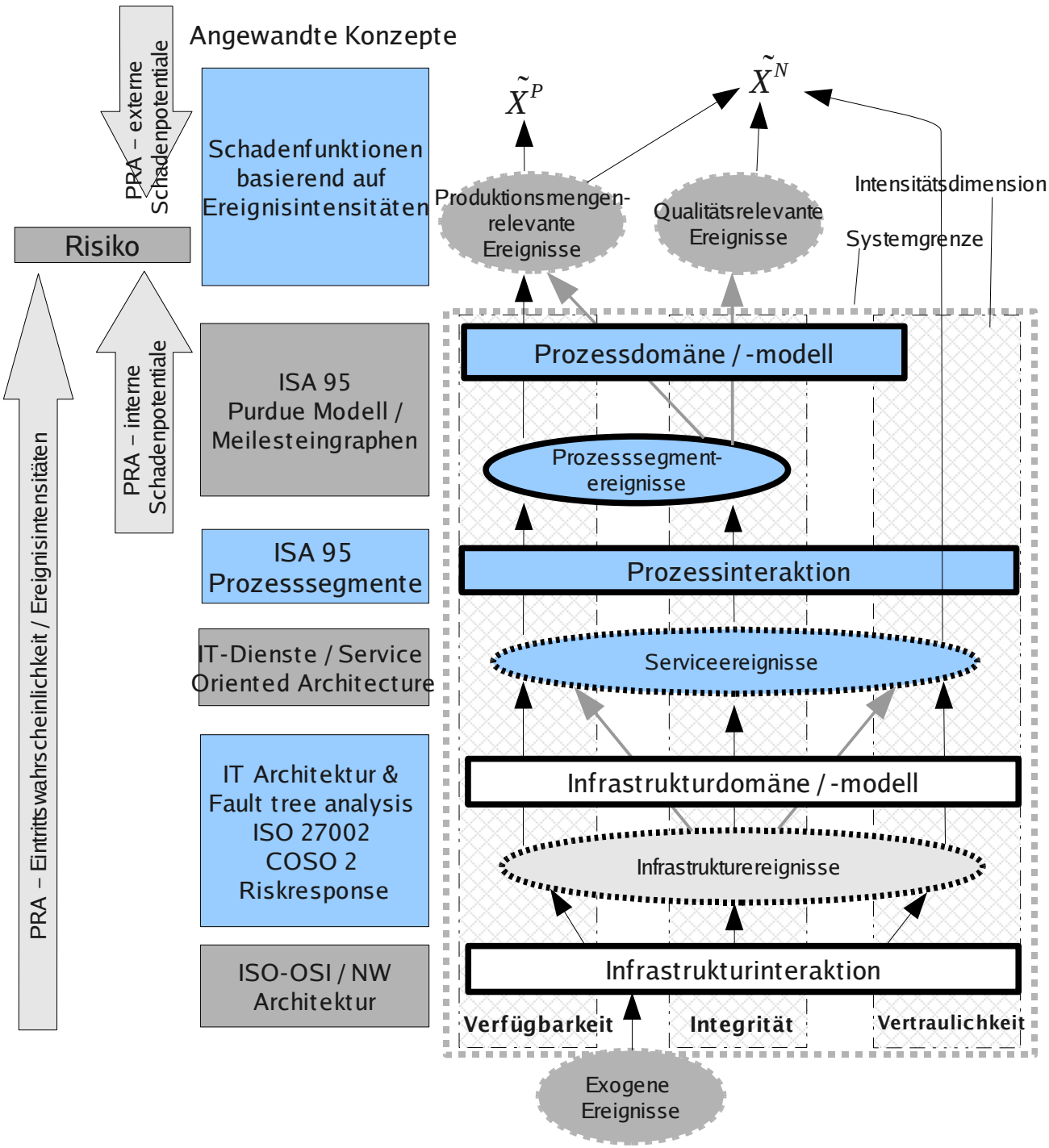

Abbildung 18: Rahmenmodell zur Integration von Informationssicherheitsereignissen in ein Risikomodell nach COSO2

Die Abbildung 18 stellt nun das dieser Arbeit zugrundeliegende Framework, dar wie sich exogene Ereignisse durch das System Unternehmen fortpflanzen und schließlich wieder zu externen Ereignissen werden, die dann auf die Zufallsvariablen $\tilde{X}^P$ und $\tilde{X}^N$ - also die Produktionsmenge und die Nachfrage des Unternehmens - wirken.
Der Eintrittspunkt der exogenen Ereignisse in das System wird dabei mit Hilfe eines Modells zur Interaktion der exogenen Ereignisse mit der Infrastruktur und der darauf folgenden, durch die Vernetzung der Komponenten bedingten Verbreitung der Ereignisse modelliert. Diese Infastrukturinteraktion wird in dieser Arbeit aber nicht näher detailliert. (Siehe dazu auch Kap.4.5 und 5.3.6.)
Der Einstiegspunkt dieser Arbeit ist die Annahme, dass das Modell zur Infrastrukturinteraktion exakte Daten bezüglich des Auftretens von Infrastrukturereignissen hinsichtlich der drei Intensitätsdimensionen Verfügbarkeit, Vertraulichkeit und Integrität auf jeder zu betrachtenden Infrastrukturkomponente liefert. Durch das weiter unten (Kapitel 4.4) dargestellte Modell für die Infrastrukturdomäne wird gezeigt, wie diese Infrastrukturereignisse zu Ereignissen die die Infrastrukturdienste (Services) betreffen aggregiert werden können. Basierend auf diesen Serviceereignissen der Infrastrukturdienste können mit Hilfe der in den Kapiteln 4.1.1 und 4.1.2.5 dargestellten Prozessinteraktion die Ereignisse in der Prozessdomäne hergeleitet werden. An diesem Punkt kommt auch eine weitere Annahme des Modells zum Tragen: es wird hier davon ausgegangen, dass Ereignisse, die die Vertraulichkeit von Informationen schädigen, keine direkten Auswirkungen auf die Prozesse haben, sondern direkt auf die Nachfrage wirken.
Mit Hilfe des primär in den Kapiteln 4.1.2, 4.3 und 5.2 dargestellten Prozessdomänenmodells können dann darauf folgend jene externen Ereignisse ermittelt, werden die direkt auf die Produktionsmenge und die Nachfrage des Unternehmens wirken.[57]
Bezüglich der Messung des Risikos wird aus Abbildung 18 ersichtlich, dass diese Arbeit die Ebene der externen Ereignisse als jenen Punkt heranzieht, an dem die Messung des Risikos mit Hilfe des Probabalistic Risk Assessment Ansatzes erfolgt. Das bedeutet, dass das jeweilige Prozess- bzw. Produkt(portfolio)risiko gemessen wird.[58]
Eine weitere Eigenschaft des Frameworks ist aus jenen Pfeilen ersichtlich, die nicht voll gezeichnet sind. Diese repräsentieren Abhängigkeiten zwischen den einzelnen Intensitätsdimensionen, die je nach Anwendungsfall des Modells unterschiedlich stark ausgeprägt sein können und für die im Rahmen dieser Arbeit zur Vereinfachung auch Annahmen getroffen wurden.

Zur weiteren Präzisierung des Modells wird nun in den nächsten Unterkapiteln gezeigt, wie der Aufbau des Systems Unternehmen (im speziellen die Infrastrukturdomäne und die Prozessdomäne) analysiert und dargestellt werden kann. Danach wird gezeigt, wie auf Basis der Systemanalyse die für die Konstruktion der Schadenfunktion relevanten Teile identifiziert werden können, und wie die Wirkung der externen Ereignisse auf die Produktionsmenge und die Nachfrage modelliert wird.
Gegen Ende dieses Kapitels wird dann der Fokus auf die Infrastruktur des Unternehmens gelegt und die wichtigsten Modellinhalte diesbezüglich eingearbeitet.

57 HINWEIS: Grundsätzlich haben auch externe Ereignisse unterschiedliche Intensitätsdimensionen (Nachfrage-, Produktions-, Qualitätsintensität,..). Diese konnten allerdings aus Gründen der Übersichtlichkeit nicht so wie die der internen Ereignisse visualisiert werden, weshalb in der Abbildung stellvertretend für die Dimensionen die zwei gezeigten externen Ereignistypen enthalten sind.

58 Im Rahmen dieser Arbeit wird davon ausgegangen, dass ein Produktionsprozess nur jeweils ein bestimmtes Produkt bzw. eine hinsichtlich ihrer Abhängigkeiten von den externen Ereignissen homogene Produktgruppe erzeugen kann. Diese Eigenschaft wird benötigt, damit die externe Schadenfunktion für einen Prozess einfacher bestimmt werden kann, und sie führt dazu, dass das Schadenpotential eines Produktes gleich dem externen Schadenpotential des zugeordneten Prozesses ist.

4.1 Relevante Teile von ISA 95

Betrachtet man nun das in Abb. 18 dargestellte Systemmodell (siehe auch Kap. 3.2) näher, so lässt sich aus Betriebsablaufsicht nahezu jedes Unternehmen in zwei große Teilbereiche gliedern, nämlich das Infrastrukturstratum[59] und das Prozessstratum. Die Prozessdomäne umfasst dabei alle betrieblichen Abläufe (Prozesse), die entweder direkt oder indirekt wertschöpfend sind. (Siehe hierzu auch das Wertkettenmodell von Porter [Porter, 1998]) Die Infrastrukturebene umfasst all jene betrieblichen Einrichtungen, die für die Abwicklung der Prozesse benötigt werden. Das Bindeglied zwischen diesen beiden Ebenen bilden die von den betrieblichen Einrichtungen den Prozessen zur Verfügung gestellten Dienste[60].

Worin sich diese Aufteilung, gründet wird u.a. im folgenden Unterkapitel ersichtlich, wo die Prozessebene und ihre Bestandteile näher betrachtet werden.

4.1.1 Das Prozesssegment

Das zentrale Element der Unternehmensabläufe sind seine Prozesse bzw. die Teile der Prozesse, die nach ANSI ISA 95 als Prozesssegmente bezeichnet werden. Ein Prozesssegment kann dabei den gesamten Prozess als auch einen beliebig kleinen Teil des Prozesses repräsentieren. In der folgenden Abbildung ist die direkte Umwelt, in der ein Prozesssegment abläuft, dargestellt.

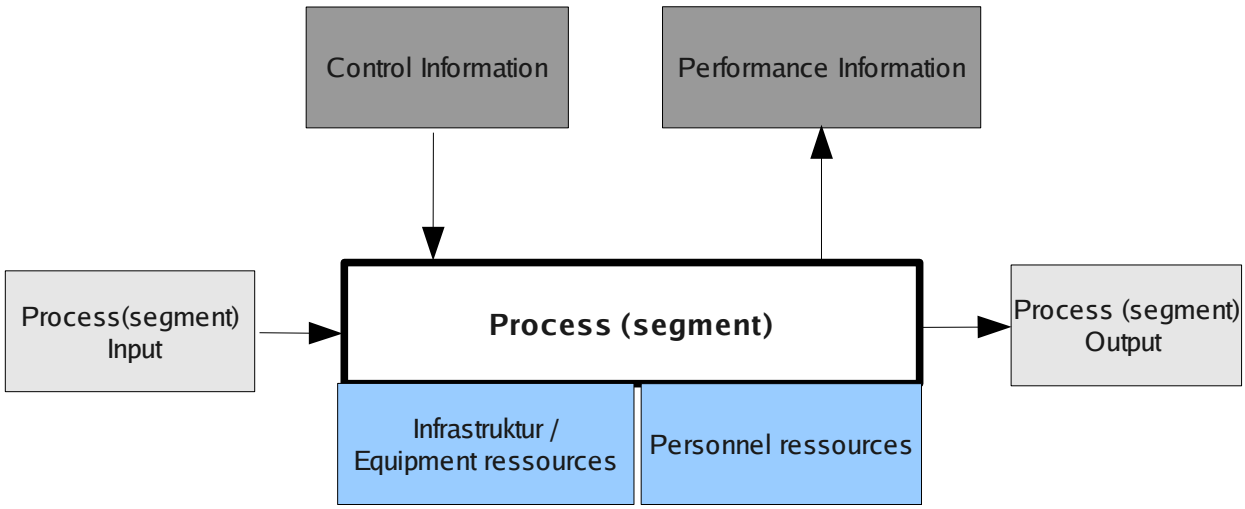

Abbildung 19: Modellrelevante Elemente der Prozesssegmentumwelt

Die Elemente der obigen Grafik sind in ISA 95 Notation. Ihre Bedeutung wird in den folgenden Unterkapiteln dargestellt.

4.1.1.1 Processsegment

Dieses Element ist repräsentativ für einen Prozess bzw. einzelne Teile (Segmente) dieses Prozesses.

59 Der Terminus Stratum stammt aus der in Kap. 2.2 vorgestellten Systemtheorie von Mesarovic et. al., in weiterer Folge wird im Dokument aber der Lesbarkeit und Verständlichkeit wegen auf den allgemeiner gefassten Begriff „Ebene“ bzw. „Domäne“ zurückgegriffen.

60 Nimmt man etwa eine Bohrmaschine als Infrastruktureinrichtung, so stellt diese z.B.: den Dienst Bohren von Löchern in Holz von 2 bis 38mm Durchmesser zur Verfügung.

4.1.1.2 Processinput

Der Prozess(segment)input ist die Menge aller Rohstoffe bzw. Rohdaten, die die Ausgangsbasis für die Erstellung des Endproduktes des Prozesses im Rahmen der Prozessdurchführung bilden.
Hinsichtlich der Typenzugehörigkeit des Inputs kann man die folgenden drei Fälle unterscheiden.

- Purdue Level 0 in einem Produktionsunternehmen: Ein Prozess bzw. Prozesssegment auf Purdue Level 0 ist in einem Produktionsunternehmen ein physikalischer Prozess; dementsprechend kann der Input eines solchen Prozesses nur ein physikalischer Stoff wie ein Rohstoff oder ein Zwischenprodukt sein.
- Purdue Level 0 in einem Dienstleistungsunternehmen: Ein solcher Prozess ist im weitesten Sinne auch ein physikalischer Prozess, da er die physikalische Repräsentation von Daten bzw. Informationen manipuliert. Der Input eines solchen Prozesses sind allerdings trotzdem die Rohdaten (Informationen), die von einem Kunden geliefert wurden bzw. die intern erzeugten Daten, die Teil eines Kundenproduktes sind.
- Purdue Level 1+ : Ab Purdue Level 1 und darüber entspricht der Input für einen Prozess bzw. für ein Prozesssegment der Performanceinformation (vgl. [Mesarovic et. al., 1970]) des von diesem Prozess geregelten (gesteuerten) Prozesses oder besteht aus dem Output eines diesen Prozess unterstützenden Prozesses[61]. Dies ist auch unabhängig davon, ob der untersuchte Prozess Teil eines Produktions- oder Dienstleistungsunternehmens ist.

4.1.1.3 Processoutput

Der Prozessoutput ist jene Menge an Zwischen- bzw. Endprodukten, die vom betrachteten Prozess an die Umwelt abgegeben werden. Analog zum Prozessinput gilt es auch hier die folgenden Fälle zu unterscheiden.

- Purdue Level 0 in einem Produktionsunternehmen: Hier stellt der Prozessoutput jene Zwischen- bzw. Endprodukte dar, die entweder an nachgelagerte Bearbeitungsprozesse oder den Kunden weitergegeben werden.
- Purdue Level 0 in einem Dienstleistungsunternehmen: Hier kann der Prozessoutput ähnlich wie in einem Produktionsunternehmen definiert werden. Der Output sind jene Daten- bzw. Informationsobjekte die entweder an einen nachgelagerten Bearbeitungsprozess oder an den Kunden abgegeben werden.
- Purdue Level 1+: Der Output eines Regelprozesses (bzw. Prozesssegments) ist unabhängig vom Unternehmenstypus entweder die Control Information für den von diesem Prozess geregelten Prozess bzw. der Input eines von diesem Prozess unterstützten Prozesses.

4.1.1.4 Control Information

Die in ANSI ISA definierte Controlinformation ist die Stellgröße eines Prozesses; sie ist jenes Mittel, mit deren Hilfe die von Mesarovic definierte Intervention von der höheren zur niedrigeren Hierarchieebene in der Prozessdomäne durchgeführt wird. In ihr enthalten sind alle Informationen die definieren, welchen Output der Prozess erzeugen soll und wie dieser Output erzeugt werden soll. ANSI ISA 95 definiert diese Informationen als Product Definition

61 mit unterstützenden Prozess sind jene Prozesse gemeint, die das jeweils betrachtete Prozesssegment insoferne unterstützen, als daß sie Informationen aufbereiten bzw. bereitstellen die das betrachtete Prozesssegment für einen erfolgreichen Durchlauf benötigt.

Information, Product Production Rules und Product Scheduling Information. Ohne diese drei Informationen kann ein Prozess den gewünschten Output entweder garnicht oder nur unkontrollierbar, und damit in einer minderen Qualität erzeugen.

4.1.1.5 Performance Information

Die PerformanceInformation ist nach Mesarovic das Feedback, das der geregelte Prozess dem regelnden Prozess zurück gibt. Sie beinhaltet jene Informationen, anhand derer der regelnde Prozess beurteilen kann wie groß die Abweichung des Outputs des geregelten Prozesses von den Sollgrößen ist. In ANSI ISA 95 wird diese Information unter dem Titel Production Information zusammengefasst. Erzeugt ein geregelter Prozess diese Informationen nicht oder können diese nicht zum regelnden Prozess kommuniziert werden, so findet keine Regelung statt, und auch in diesem Fall kann der regelnde Prozess den geregelten Prozess nicht kontrollieren und damit die Erzeugung des gewünschten Outputs nicht sicherstellen.

4.1.1.6 Infrastructure

Die Infrastruktur repräsentiert sämtliche physikalischen Ressourcen[62] eines Unternehmens, die ein Prozess(segment) benötigt um innerhalb der für ihn definierten Parameter ablaufen zu können. Dies können z.B.: Produktionsstätten, Büroeinrichtungen, Computernetze oder Energieträger sein. Die korrekte Funktion der Infrastruktur stellt die Basis für die korrekte Funktion eines Geschäftsprozesses bzw. seiner Segmente dar.

4.1.1.7 Personnel Ressources

Unter Personnel Ressources werden sämtliche Fertigkeiten und die Arbeitskraft der für die Durchführung des Prozesses benötigten Mitarbeiter des Unternehmens subsummiert. Sollten die eingesetzen Mitarbeiter nicht über die notwendigen Fähigkeiten, verfügen so wird der Prozessoutput nicht den Vorgaben entsprechen. Es wird ebenso nicht möglich sein die Vorgaben einzuhalten, wenn nur eine nicht ausreichende Anzahl von Mitarbeitern eingesetzt wird.
Da sich diese Arbeit auf Aspekte der Informationssicherheit und deren Beitrag zum Betriebsrisiko eines Unternehmens fokussiert, werden ausreichende personelle Ressourcen als gegeben angenommen und Ereignisse, die deren Bereitstellung beeinträchtigen, können nicht weiter untersucht.

62 die weder Prozessinput noch Prozessoutput sind.

4.1.2 Modellierung der Prozessdomäne

Fügt man die in Kapitel 4.1.1 dargestellte Sicht auf ein einzelnes Prozesssegment bzw. auf einen einzelnen Prozess in das Purdue Referenzmodell ein und berücksichtigt man dabei die Definitionen von Mesarovic [Mesarovic et al., 1970] bezüglich eines Systems, das Entscheidungsprobleme lösen muß, so kann die Struktur bzw. Hierarchie in der Prozessdomäne eines Unternehmens wie in Abbildung 20 dargestellt werden. Als Prozessdomäne ist dabei jene Sicht auf die Organisationsstruktur eines Unternehmens zu verstehen, die ausschließlich die Geschäftsprozesse beinhaltet.

Das unten visualisierte Unternehmen soll einen Produktionsbetrieb darstellen, der seine komplette Produktpalette in einem einzigen Produktionsprozess (Purdue Ebene 0), bestehend aus drei unterschiedlichen Prozesssegmenten, herstellen kann. Aus Gründen der Übersichtlichkeit[63] wurde auf Purdue Ebene 4 verzichtet. Ebenfalls der Übersichtlichkeit wegen wurde die Control- und Performance Information nicht mehr explizit dargestellt.

Da in diesem Kapitel primär die Abhängigkeiten der Prozesse eines Unternehmens (hier einer Gießerei) dargestellt werden sollen, wurde in der unten gewählten Darstellung vor allem im Produktionsprozess auf Purdue Ebene 0 bewusst darauf verzichtet, die genauen Relationen von Input und Output darzustellen.[64]

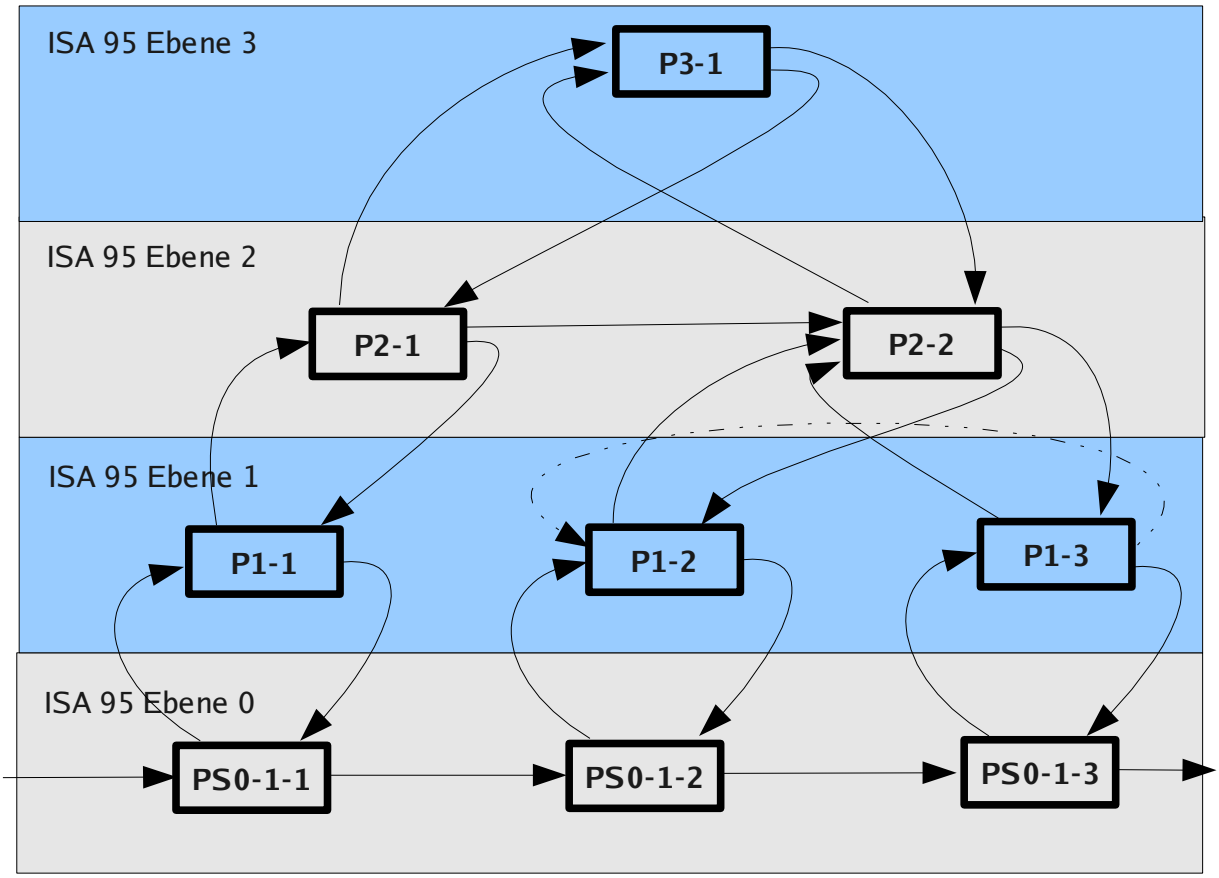

Abbildung 20: Prozesshierarchie basierend auf ISA 95 / Purduemodell

63 und weil ein Unternehmen dieser Komplexität die Funktionalität der Ebene 4 auch in Ebene 3 abbilden kann.

64 Dies ist insofern auch nicht notwendig, da eine exakte Modellierung dieser Zusammenhänge bereits mit der Theorie der Produktionsfunktionen ausreichend detailliert möglich ist. (siehe auch [Fandel, 2005])

Die Abkürzungen in obiger Darstellung[65] haben folgende Bedeutung.

P: *Prozess mit Beschriftung ISA Ebene – Prozessnummer*
PS: *Prozesssegment mit Beschriftung ISA Ebene – Prozessnummer – Segmentnummer*

Nachfolgend werden die angenommenen Eigenschaften des modellierten Unternehmens, die zur obig dargestellten Prozessdomäne führen, für jede ISA 95 Ebene dargestellt.

4.1.2.1 Ebene0

Ebene 0 stellt wie schon weiter oben angesprochen (Kap. 4.1.1) jene Ebene dar, in der die physikalischen Prozesse des Unternehmens ablaufen. Damit ist dies die Ebene, auf der in dem Produktionsunternehmen die Gütererzeugung stattfindet. Für diese Gütererzeugung wurde angenommen, dass jedes Gut auf seinem Weg vom Rohstoff zum fertigen Endprodukt jeden der drei Arbeitsschritte, die jeweils durch ein Prozesssegment dargestellt werden, durchlaufen muß. Durch die direkte Anbindung von Prozessoutput an Prozessinput soll dargestellt werden, dass der Output des vorangegangenen Segments dem Input des nachfolgenden Segments entspricht.

4.1.2.2 Ebene1

Durch die im Beispiel unterstellte Verschiedenartigkeit der einzelnen Arbeitsschritte im Produktionsprozess ist es notwendig, jeden dieser Schritte durch einen eigenen Prozess zu überwachen. Ein Beispiel für einen solchen Betrieb könnte etwa eine Gießerei mit den Arbeisschritten „aufschmelzen im Hochofen“, „gießen des Produktes“, und „Nachbearbeitung des Produktes“ sein. In einem solchen Betrieb kann es auch notwendig sein, dass der Prozess P1-2 in seinem Input Teile des Outputs von Prozess P1-3 - wie oben durch den strichlierten Pfeil dargestellt - berücksichtigt, z.B. um, durch kleine Modifikationen während des Gusses den Arbeitsschritt Nachbearbeitung zu erleichtern. Dieser Input für P1-2 ist allerdings optional, und sein Fehlen hat keine Auswirkungen auf den Ablauf von P1-2.

4.1.2.3 Ebene2

Ebene 2 besteht aus zwei Koordinationsprozessen, von denen der zweite Prozess die Ebene 1 Prozesse vom Typ 2 und 3 kontrolliert[66]. In dem hier dargestellten Fall muß Prozess P2-2 auch Teile des Outputs von Prozess P2-1 als Input berücksichtigen. Auch dieser Umstand kann durch die Gießerei gut veranschaulicht werden, da es keinen Sinn macht den Guß eines Produktes zu veranlassen solange das Einschmelzen der Erze noch nicht abgeschlossen ist.

4.1.2.4 Ebene3

Ebene 3 enthält nur einen Prozess[67], der die Steuerung bzw. Zieldefinition des gesamten Unternehmens ebenso wie die Kommunikation des Unternehmens mit seinen Lieferanten und

65 Hinweis: Es bestünde auch die Möglichkeit, die Prozesse bzw. Prozesssegmente ebenfalls als Komponenten mit UML zu modellieren und die Abhängigkeiten der Prozessebene derart darzustellen. Diese Vorgehensweise erscheint aber wenig sinnvoll, da Prozesse untereinander nicht durch Funktionalität verbunden sind, sondern durch Objekte. Daher ist es nicht zielführend Objektabhängigkeiten mit Methoden zur Darstellung funktionaler Abhängigkeiten zu visualisieren. Für eine detaillierte Darstellung der Prozessebene eignen sich daher eher die UML Aktivitäts-, Kollaborations-, bzw. Sequenzdiagramme.

66 dies ist u.Ust. mit Hilfe eines Prozesssegments pro geregeltem Prozess sinnvoll möglich, dies wurde hier aber der Übersichtlichkeit wegen nicht dargestellt.

67 bestehend aus mehreren hier nicht explizit dargestellten Segmenten, die die unterschiedlichen Aufgaben durchführen.

Kunden abwickelt. Da - wie schon weiter oben begründet - die Ebene 4 nicht abgebildet wurde, ist dieser Prozess in diesem Beispiel auch für die unternehmensweite Zielerreichung verantwortlich.

4.1.2.5 Allgemeine Eigenschaften der Prozessebene

Werden in einem Unternehmen die Geschäftsprozesse durchgehend einer der ISA 95 Ebenen zugeordnet, so lassen sich folgende teils triviale Eigenschaften hinsichtlich des Zusammenspiels bzw. der Voraussetzungen für selbiges zwischen den einzelnen Prozessen innerhalb der Ebenen und über die Ebenengrenzen hinweg festhalten:

- geregelte Prozesse sind ohne die Controlinformation der regelnden Prozesse mit hoher Wahrscheinlichkeit ineffektiv und ineffizient und damit de facto unkontrollierbar. Daher wird als Basis dieser Arbeit angenommen, dass ein Fehlen der Controlinformation für einen Prozess dessen Stillstand bewirkt.
- Regelnde Prozesse, die keine Performanceinformation[68] von den von ihnen geregelten Prozessen erhalten, werden zu steuernden Prozessen und sind mit hoher Wahrscheinlichkeit nicht in der Lage sinnvolle Controlinformation an die von ihnen geregelten Prozesse zu liefern. Daher wird angenommen, dass nach einer für einen regelnden Prozess jeweils zu erhebenden Zeitspanne nach dem Versiegen der Performanceinformation dieser Prozess auch keine Controlinformation mehr erzeugen kann.
- Die zeitliche Abhängigkeit zwischen einzelnen Prozessschritten und Prozessen innerhalb einer Ebene sinkt mit steigender Ebene, innerhalb der sich dieser Prozess befindet. D.h., ein Prozess auf Ebene 0 ist wesentlich abhängiger vom zeitgerechten Abschluss der Vorgängertätigkeit als beispielsweise ein Prozess auf Ebene 3. Dies ist ja auch insofern schlüssig, weil operative Prozesse (Ebene 0-2) meist deutlich geringere Pufferzeiten haben als taktische und strategische Prozesse (Ebene 3 bzw. 4 u. 5).
- die zeitliche Abhängigkeit zwischen Prozessen auf unterschiedlichen Ebenen sinkt mit steigender Ebene, in der sich die untersuchten Prozesse befinden. Das bedeutet, die Zeitspanne, nach der ein Fehlen der Controlinformation bzw. Performanceinformation zu einem Versagen des Prozesses führt steigt mit zunehmender ISA 95 Ebene.
- Je höher die Ebene, der ein Prozess angehört, desto größer ist die Anzahl der von ihm über die Controlinformation geregelten und damit abhängigen Prozesse.

4.1.3 Modellierung und Analyse der Prozessabläufe

Betrachtet man die Analyse- und Darstellungsmethodik aus Kap. 4.1.2 Näher, so kann man erkennen, dass diese auch schon die wichtigsten Informationen über die Prozessabläufe enthält, da darin schon sämtliche Input- und Outputbeziehungen für einen Prozess einer ANSI ISA Ebene inkl. aller seiner regelnden Prozesse enthalten sind. Damit fehlen für diesen Analyseschritt nur noch die Informationen über die Fertigstellungszeitpunkte der jeweiligen Outputs für die benötigte Ablaufdarstellung. Das Ziel der Ablaufanalyse für dieses Modell ist die Erstellung eines Graphen ähnlich jenem, der in Kapitel 2.5 zur Bestimmung des kritischen Pfades diente. Die dafür benötigte Information über die Fertigstellungszeitpunkte ist dabei in den Outputmeilensteinen der jeweiligen Prozesssegmente enthalten.

68 die ja den größten Teil des Inputs für diesen Prozess darstellt.

4.1.3.1 Prozessmeilensteine

Analysiert man einen Meilenstein[69] wie etwa die Lieferung einer bestimmten Menge der Ware X an den Kunden A[70], so kann dieser Meilenstein nur erreicht werden, wenn zuvor jeder der beteiligten Geschäftsprozesse erfolgreich beendet wurde. Aus dieser engen Verknüpfung des zu analysierenden Meilensteines mit seinem jeweiligen Prozess(segment) ergeben sich Abhängigkeiten zwischen den Meilensteinen, die jenen der Prozess(segment)e untereinander sehr ähneln. Jeder Prozess hat dabei folgende Eigenschaften bezüglich der Meilensteine, die er erreichen soll.

- Jedes in der Controlinformation eines Prozesssegments definierte Ziel dieses Prozesssegments stellt einen dar Meilenstein der zu erreichen ist. Daraus ergeben sich die zwei weiter unten aufgeführten Meilensteintypen.
- Ein Prozess kann seine Ziele / Meilensteine nur dann erreichen wenn er über sämtlichen dafür notwendigen Input, sämtliche dafür notwendige Controlinformation und alle benötigten Ressourcen (Infrastruktur und Humane) verfügen kann[71].
- Jeder zeitliche Verzug eines Meilensteines stellt eine Zustandsänderung des Prozesses dar, die von einem entsprechenden Ereignis kommuniziert wird.[72]
- Outputmeilensteine sind jene definierten Zeitpunkte für die Fertigstellung des Prozessoutputs in der definierten / vereinbarten Qualität.
- Performancemeilensteine zeigen die Erreichung all jener Ziele, an die nicht direkt den Output des Prozesssegments betreffen. Z.B. könnte ein solcher Meilenstein die erfolgreiche Fertigstellung von 5 Kaffeemaschinen sein, der ein Ziel für einen Prozess auf ISA Ebene 3 sein kann, jedoch nicht direkt durch den Output dieses Prozesses erreicht wird.
- Für Prozesse auf ISA Ebene 0 sind alle Outputmeilensteine auch Performance-meilensteine, da diese Prozesse keine Regelungsaufgaben besitzen.
- Durch die Prozesshierarchie setzen sich Performancemeilensteine eines Prozessegments auf Ebene n immer aus den Performancemeilensteinen und den Outputmeilensteinen der von ihm geregelten Prozesssegmente auf Ebene n-1 zusammen.

Eine wichtige Basis für die Ableitung der Meilensteine und ihrer Hierarchie aus der Prozesshierarchie ergibt sich aus dem letzten Punkt der Eigenschaften eines Prozesses: es muß um diese Ableitbarkeit gewährleisten zu können, für jedes Ziel eines regelnden Prozesses definiert sein, welches der Ziele, die dieser Prozess seinen geregelten Prozessen vorgibt, dieses eigene Ziel unterstützt. (siehe auch das Kapitel Koordination in [Mesarovic et. al.,

69 Unter einem Meilenstein ist dabei jener Zeitpunkt an dem ein Zwischen- oder ein Endziel einer bestimmten Handlung oder Tätigkeit erreicht ist, zu verstehen.

70 Hinweis: Um in weiterer Folge externe Ereignisse identifizieren zu können ist es ratsam, an dieser Stelle stets einen Meilenstein als Basis der Analyse zu wählen, der in einem Liefervertrag oder Service Level Agreement mit einem Kunden enthalten ist.

71 Dieses Prozessversagen bei Fehlen der Prozessvoraussetzungen könnte in weiterer Folge über eine Versagenswahrscheinlichkeit abgebildet werden, die für Prozessressourcen fast sicher immer den Wert eins annehmen wird. Für den Prozessinput oder die Controlinformation besteht aber auch die Möglichkeit die Versagenswahrscheinlichkeit zeitabhängig zu gestalten, da es durchaus möglich ist für kurze Zeit auf die aktuellste Version der jeweiligen Informationen zu verzichten. So kann ein chemischer Prozess wie etwa die Verbrennung von Holz in einem Ofen auch ohne Controlinformation (z.B.: ob Holz nachgelegt werden muß) für eine gewisse Zeit im gewünschten Toleranzbereich bleiben. In diesem Dokument wird dieser Ansatz aber aus Gründen der Abgrenzung und Einfachheit der Darstellung nicht weiter verfolgt.

72 Diese Ereignisse sind jene, die bei wichtigen Meilensteinen von Interesse und somit Ziel der Ereignisidentifikation sind.

1970]). Diese Zuordnung der Ziele kann bereits in der Controlinformation als Vorgabe enthalten sein, oder sie wird vom jeweiligen Prozess selbst erstellt. Wichtig ist jedenfalls, dass sie existiert, sonst kann die Prozessdomäne nicht hinreichend analysiert werden.
Um den Meilensteingraphen zur Analyse des kritischen Pfades zu erhalten, sind unter Berücksichtigung obiger Eigenschaften folgende Schritte durchzuführen:

1. Ordne den zu analysierenden Meilenstein einem Prozess der Domäne zu.
2. Identifiziere alle Performance- und Outputmeilensteine dieses Prozesses, die zur Erreichung des zu analysierenden Meilensteins notwendig sind.
3. Wiederhole Schritt 2 für alle jene Prozesse, die vom analysierten Prozess geregelt werden, die diesen regeln bzw. die Input für den analysierten Prozess bereitstellen müssen. Hierbei ist Schritt 2 iterativ so oft zu wiederholen, bis alle Performancemeilensteine auf Outputmeilensteine zurückgeführt wurden[73] oder eine Schätzung der jeweiligen Performancemeilensteine für ausreichend erachtet wird.[74]

4.2 Basis der Identifizierung relevanter Ereignisse / COSO2 Aktivität 3

Basierend auf der Definition für ein Ereignis (siehe Kapitel 3.3.1) soll nun folgend kurz diskutiert werden, welche Vor- und Nachteile jeder der drei durch das obig vorgestellte Unternehmensmodell definierten, grundlegenden Ereignistypen mit sich bringt, wenn er als Ausgangsbasis für das in Kapitel 3.3 eingeführte Risikomodell herangezogen wird.

4.2.1 Externe Ereignisse

Gemäß den obigen Definitionen (siehe hierzu auch die detaillierte Aufarbeitung in Kapitel 4.3.2) sind die externen Ereignisse jene Ereignisse, die auch die Kunden eines Unternehmens messen, d.h. wahrnehmen können. Da die Kundenbeziehungen eines Unternehmens meist hinsichtlich der Geldflüsse und der Vorbedingungen für diese sehr klar geregelt sind können bei einer Bewertung der externen Ereignisse - vor allem jener, die sich auf den Absatz des Unternehmens auswirken (siehe Kap. 4.3.2) - die Schadenhöhen bzw. die Schadenpotentiale und damit auch die Schadenfunktion in einem Großteil der Fälle recht einfach aus den Leistungs- bzw. Liefervertägen des Unternehmens bestimmt werden.
Im Gegensatz dazu ist es ohne genaue und langfristige statistische Aufzeichnungen solcher eher seltenen Ereignisse nur sehr schwer möglich, den zugrundeliegenden Zählprozess bzw. im Falle des Poissonprozesses dessen Intenstiät zu bestimmen. Will man die Intensität[75] möglichst exakt festmachen, so ist es notwendig, eine Top-Down Analyse durchzuführen, um mit Hilfe des dadurch entstehenden Fehlerbaumes die Eintrittswahrscheinlichkeit für das Topereignis aus den dafür ursächlichen Basisereignissen abzuleiten.
Ein Vorteil dieser Ereigniskategorie ist, dass es aufgrund der leichter zu bestimmenden Schadenpotentiale einfacher ist die relevanten Ereignisse dieser Gruppe zu identifizieren. Der Nachteil dieser Ereigniskategorie kann aber sehr gut aus der in Kapitel 4.3.2 gezeigten Ereigniskategorisierung abgelesen werden, sie ist nur eine Teilmenge aller Ereignisse die in

73 Ein Mapping aller Performancemeilensteine zu Outputmeilensteinen bewirkt daß die Zielerreichung aller ANSI ISA Ebenen über 0 auf Zielerreichungen in Ebene 0 zurückgeführt werden und umgekehrt. -> es ist dadurch für diesen Analyseschritt nicht so wichtig, welchem Prozesssegment (welcher Ebene) die Endverantwortung für die Leistungserbringung zugeordnet wird.
74 Dies ist wird in jenen Fällen zutreffen, wo eine Analyse bis zur ANSI ISA Ebene 0 nicht gewünscht wird.
75 bzw. die Eintrittswahrscheinlichkeit

einer vollständigen Risikoanalyse enthalten sein sollten. Das bedeutet vor allem, dass in einem Unternehmen, das bereits viele Sicherungsmaßnahmen bzw. Kontrollaktivitäten enthält viele relevante Ereignisse meist nicht in dieser Gruppe enthalten sind. (näheres zu den Kontrollaktivitäten in den Kapiteln 5.2.5 und 5.3.5)

4.2.2 Exogene Ereignisse

Nach den obigen Definitionen (siehe weiters auch Kap. 4.5) besteht die Gruppe der exogenen Ereignisse aus jenen Ereignissen, die ihren Ursprung in der Unternehmensumwelt haben und die auf das Unternehmen messbar einwirken. Fokussiert man sich bei der Erhebung des Betriebsrisikos eines Unternehmens auf diese Ereignisgruppe, so ist es bei einer geschickten Auswahl der zu bewertenden Ereignisse oftmals recht gut möglich deren Intensität / Eintrittswahrscheinlichkeit aus statistischen Daten zu berechnen. Betrachtet man etwa z.B. den Versuch einer Schadsoftware sich auf einem PC einzunisten, so ist dies ein Ereignis, das bei einem privat genutzten PC mit Internetverbindung durchaus mehrmals täglich auftreten kann.[76] Im Gegensatz zu den externen Ereignissen ist es aber de facto unmöglich den exogenen Ereignissen direkt Schadenpotentiale oder Schadenfunktionen zuzuorden. Für die Bestimmung des Schadenpotentials ist es notwendig eine Ereignisbaumanalyse (siehe Kap. 2.5.1) durchzuführen, deren Endziel es ist die in Folge des exogenen Ereignisses möglichen intern messbaren Ereignisse in der Prozessdomäne eines Unternehmens zu erheben. Das Schadenpotential des exogenen Ereignisses setzt sich dann aus dem Schadenpotential aller durch es bedingten externen Ereignisse in seinem Ereignisbaum zusammen.
Der große Nachteil der exogenen Ereignisse bei einer strukturierten Risikoanalyse ist die „geschickte Auswahl“, es ist auf Grund der Fülle an exogenen Ereignissen ohne Information über die dadurch bedingten externen Ereignisse und deren Schadenpotentiale nur sehr eingeschränkt möglich die relevanten exogenen Ereignisse zu identifizieren. Das bedeutet, um die relevanten exogenen Ereignisse identifizieren zu können, ist es notwendig diese bereits einer Risikoanalyse zu unterziehen.

4.2.3 Interne Ereignisse

Interne Ereignisse sind Nachrichten über Zustandsübergänge bei Infrastrukturkomponenten bzw. bei Prozesssegmenten. Für eine Risikobewertung sind dabei jene Zustandsübergänge, die bei den Diensten (services) am Übergang zwischen Infrastruktur und den Prozessen stattfinden bzw. jene in den Prozesssegmenten von besonderem Interesse, da hier die internen Schadenpotentiale im Fall des Ereigniseintritts auftreten. Diese Ereignisse haben den Vorteil, dass sie wesentlich häufiger auftreten als die externen und daher statistisch besser erfassbar sind. Auf der anderen Seite stellen diese Ereignisse nur die kleinen und mittleren Schadereignisse dar, und wenn nur diese in einer Risikobetrachtung berücksichtigt werden, gibt es keine Möglichkeit die möglichen Großschadenereignisse (die zu den externen Ereignissen zählen) bei den Riskresponses zu berücksichtigen.

Daher ist es sinnvoll einen kombinierten Ansatz, der sowohl die internen Schadenpotentiale als auch die externen berücksichtigt, zu wählen.

76 und ohne entsprechenden Schutz des PCs werden diese Versuche auch oftmals erfolgreich sein.

4.3 Ereignisidentifikation in der Prozessdomäne

In der Prozessdomäne wird eine der grundlegenden Herausforderungen der Aktivität Eventidentification sehr schnell offensichtlich: die Notwendigkeit zur Beschränkung auf die relevanten bzw. wesentlichen Ereignisse. Im Vergleich zur Infrastrukturebene, in der es ja meist durchwegs enge Definitionen für die erwünschten bzw. geplanten Betriebszustände gibt, herrscht in der Prozessebene eine weit losere Kopplung der beteiligten Komponenten (Prozesssegmente), und es existieren oft nicht so klare Formulierungen von Zielen bzw. Sollzuständen.

In diesem Unterkapitel werden deshalb vorab zwei Methoden zur Reduktion der Menge der identifizierten Ereignisse auf jene, die wesentlich sind, beschrieben.

4.3.1 Fokus auf Meilensteine

Ein erster zur Fokussierung auf die relevanten Ereignisse hilfreicher Ansatz[77] ist, sich auf die Zustandsübergänge von Meilensteinen zu beschränken, und zwar auf jene Übergänge vom Status „im Plan“ zum Status „in Verzug“ bzw. „überfällig“. Als Bezugspunkt für die Bewertung, ob ein Meilenstein nicht mehr im Plan ist, dient die latest delivery time, also jener Zeitpunkt zu dem der Meilenstein spätestens erreicht sein muß um den Endmeilenstein bzw. Liefertermin, zu dem sich das Unternehmen gegenüber seinen Kunden verpflichtet hat, noch fristgerecht erreichen zu können.[78]

4.3.2 Ereigniskategorisierung

Die zweite Möglichkeit der Beschränkung auf die relevanten Ereignisse ist es die Auswahl derselbigen basierend auf einer Kategorisierung wie der unten beschriebenen zu treffen. Denn wählt man den in den Kapiteln 4.1.1 und 5.2 beschriebenen Top-Down Ansatz zur Analyse der Ereignisse in einem Unternehmen und im speziellen in dessen Prozessebene, so stellt sich im Rahmen der Definition der zu analysierenden Meilensteine das Problem, welche Meilensteine als Ausgangspunkt der Betrachtungen gewählt werden sollen.[79] Der hier angebrachte logische Ansatz ist, jene Meilensteine bzw. Prozesssegmente bevorzugt auszuwählen, die geschäftskritisch sind bzw. die die größten Schadenpotentiale haben. Die systematische Beurteilung, welche Ereignisse dabei geschäftskritische Auswirkungen haben können, soll im Rahmen dieses Modells auf Basis der folgenden Ereigniskategorisierung für ein Unternehmen erfolgen.

Grundsätzlich lassen sich aus Sicht des Autors sämtliche Schadenpotentiale und damit auch die sie verursachenden Ereignisse in eine der folgenden vier Kategorien einteilen.[80] Diese Kategorisierung erlaubt aber neben einer Priorisierung gewisser Ereignistypen auch einen schrittweisen Aufbau der jeweils durch diese Ereignistypen definierten Teile der Schadenfunktion (siehe hierzu die Kapitel 5.2.3 f.).

77 Für all jene Fälle, in denen die Dauer eines Ereignisses als einziges Merkmal einer Intensität herangezogen wird.

78 Diese Fokussierung auf die Meilensteine entspricht einer Einschränkung der Risikomodellierung auf das Thema Verfügbarkeit im Prozess bzw. Analyse der realisierbaren Produktionsmenge.

79 im vorangegangenen Kapitel wurde implizit davon ausgegangen, dass diese Ausgangspunkte jene Meilensteine sein sollten, zu denen sich ein Unternehmen gegenüber seinen Kunden verpflichtet hat. Dies muss aber nicht notwendigerweise so sein und ist in einigen Fällen auch eine zu große Einschränkung.

80 Dies gilt auch für jene exemplarisch im ISO IEC 27005 Anhang B2 aufgelisteten Ereigniskategorien.

4.3.2.1 Interne Schadenpotentiale

Interne Schadenpotentiale beinhalten all jene Aufwände, die im Ereignisfall durch die interne Ereignisregulierung entstehen. Beispiele für interne Schadenpotentiale sind etwa Instandsetzungsarbeiten oder angefallene Überstunden während eines Ereignisses.
So sind zum Beispiel die im ISO 27005 [ISO 27005] Anhang B2 aufgelisteten Ereignistypen „disruption of internal operation“ oder „loss of goods / funds / assests“ primär dieser Kategorie von Schadenpotentialen zuzuordnen.
Für eine genaue, auch mit einem Beispiel hinterlegte Abgrenzung der internen Schadenpotentiale sei der Leser an dieser Stelle auf das Kapitel 5.2.4.5 verwiesen.

4.3.2.2 Pönalezahlungen

Dieses Schadenpotential ist recht klar abgegrenzt: es umfasst all jene Aufwände, die durch mit Strafzahlungen sanktionierte Vertrags- oder Gesetzesverletzungen entstehen. Der Vorteil dieser sehr klaren Abgrenzung ist auch die in weiterer Folge sehr präzise Abschätzbarkeit dieser Aufwände.

4.3.2.3 Schadenpotentiale auf Grund von Produktionsmengenänderungen

Diese Schadenpotentiale entstehen grob umrissen aus entgangenen Gewinnen auf grund der negativen Beeinflussung der Produktionsmenge $\tilde{X}^P$ durch die aufgetretenen Ereignisse. Für diese Kategorie von Schadenpotentialen sind im ISO 27005 unter anderem die folgenden ursächlichen externen Folgeereignistypen wie „interruption of service“ oder „disruption of a third party's operation“ (Siehe auch Kapitel 5.2.4.2) genannt.

4.3.2.4 Nachfragebedingte Schadenpotentiale

Dies sind all jene Schadenpotentiale, die aus entgangenen Gewinnen auf grund der negativen Beeinflussung der Nachfragemenge $\tilde{X}^N$ durch die aufgetretenen Ereignisse enstehen. Dieser Kategorie von Schadenpotentialen sind wohl die meisten externen Folgeereignistypen direkt zuordenbar, wenngleich dieses Schadenpotential am schwierigsten zu bestimmen ist. Aus dem ISO 27005 lassen sich u.a. die folgenden Ereignistypen direkt hier zuordnen: „loss of customer confidence“, „Danger to personnel / user safty“, „loss of technical lead“ oder „loss of customers“ (siehe auch Kapitel 5.2.4.3)

4.3.3 Eingliederung der Schadenfunktion

Die Schadenfunktion ist ein generisches Konstrukt, das dazu dient die jeweiligen Prozess- und Markteigenschaften bzw. auch deren Umfelder, die für jedes Unternehmen unterschiedlich sein können, auf eine Weise zusammen zu führen, daß einer Intensität eines Ereignisses direkt ein resultierender Schaden zugeordnet werden kann. Das bedeutet, dass die Schadenfunktion grundsätzlich für jeden Anwendungsfall empirisch zu erheben ist, wobei in Kapitel 5.2.4 die allgemeinen Bestandteile einer Schadenfunktion genauer dargestellt sind.

Die mit einem externen Ereignis verknüpften Schäden in den Prozessen entstehen dabei primär durch das Verfehlen der mit den Kunden vereinbarten Qualitätsziele bzw. Liefermengen und -termine. Bei der Erhebung der Schadenfunktion eines Prozesses bzw. eines Prozesssegmentes bildet daher die Schnittstelle des Unternehmens mit seinen Kunden den Ausgangspunkt. Die Hauptkomponenten der an dieser Schnittstelle festgestellten Schadenfunktion sind dann in weiterer Folge auch für alle an der Herstellung des Produktes beteiligten Prozesssegmente nahezu gleich. (die Details der Herleitung der Komponenten der Schadenfunktion und deren Applikation auf die jeweiligen Prozesssegmente sindt in Kapitel 5.2 näher ausgeführt).

Ein Beispiel für eine Schadenfunktion eines Prozesssegments, in der primär mögliche Ausfälle in der Produktion Eingang gefunden haben, ist in der folgenden bereits aus Kapitel 3.3.3 bekannten Abbildung dargestellt, in der die Intensität eines Ereignisses sich ausschließlich auf der Dauer des Ereignisses gründet.

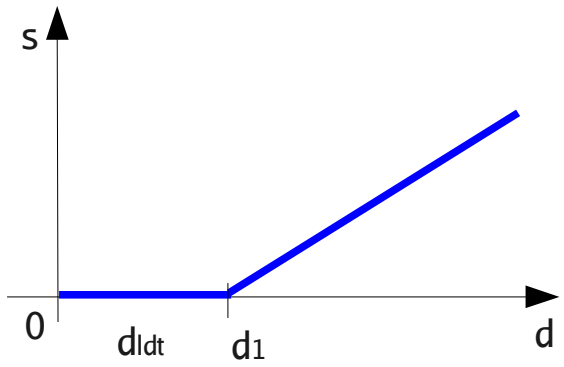

Abbildung 21: Beispiel für eine einfache Schadenfunktion

s	...	*Schadenhöhe*
d	...	*Ereignisdauer eines Prozessstillstandsereignisses*
d_{ldt}	...	*Pufferzeit zwischen geplanter Fertigstellung des Prozessoutputs und der spätest notwendigen Fertigstellung*[81].
d_1	...	*Ereignisdauer, ab der ein Schaden eintritt.*

Wie ersichtlich ist gründet sich die auch bereits in Kapitel 3.3.3 als Intensitätsschwelle visualisierte Zeitspanne bis ein Schaden eintreten kann darin, dass erst nachdem alle Pufferzeiten im Prozess aufgebraucht wurden auch der Meilenstein seinen Status verändert. Der Verlauf der Schadenfunktion ab dem Punkt d1 ist dabei in weiterer Folge von den Liefer- bzw. Leistungsvereinbarungen mit den jeweiligen Kunden des Unternehmens abhängig.
Für jenen einfachen Fall, in dem ein Infrastrukturdienst nur von einem Prozesssegment benötigt wird, ist es durch die in den Kapiteln 4.1.1 und 4.1.2.5 aufgeführten Eigenschaften der Prozesssegmente und deren Abhängigkeiten von der Infrastruktur bedingt, dass die Schadenfunktion eines Infrastrukturdienstes sogar gleich jener des Prozesssegmentes ist.

81 latest delivery time

4.4 Modellierung der Infrastrukturdomäne

Nachdem in den vorangegangen Unterkapiteln die Modellierung der Prozessdomäne und darauf aufbauend die Identifikation von Ereignissen sowie die Hintergründe für den Einsatz der Schadenfunktion erläutert wurden, soll dieses Unterkapitel eine Möglichkeit zur Modellierung der Infrastrukturdomäne zeigen. Basierend darauf wird im Anschluss gezeigt, wie der Zählprozess und die Dauer der jeweiligen Infrastrukturereignisse hergeleitet werden können.

4.4.1 Allgemeine Infrastrukturmodellierung

Zur Darstellung der Infrastruktur eines Unternehmens eignen sich jene Teile der Unified Modelling Language [UML], die zur Strukturmodellierung konzipiert wurden sehr gut. Daher wird im folgenden die Infrastruktur mit Hilfe von Komponenten bzw. Klassen,diagrammen dargestellt. Als Schnittstelle zu den Prozessen dient hierbei wie schon weiter oben angeführt der von einer Infrastrukturkomponente zur Verfügung gestellte Dienst (Service). Dieser Dienst dient auch in weiterer Folge als Ausgangspunkt für die Analyse der Infrastrukturabhängigkeiten.

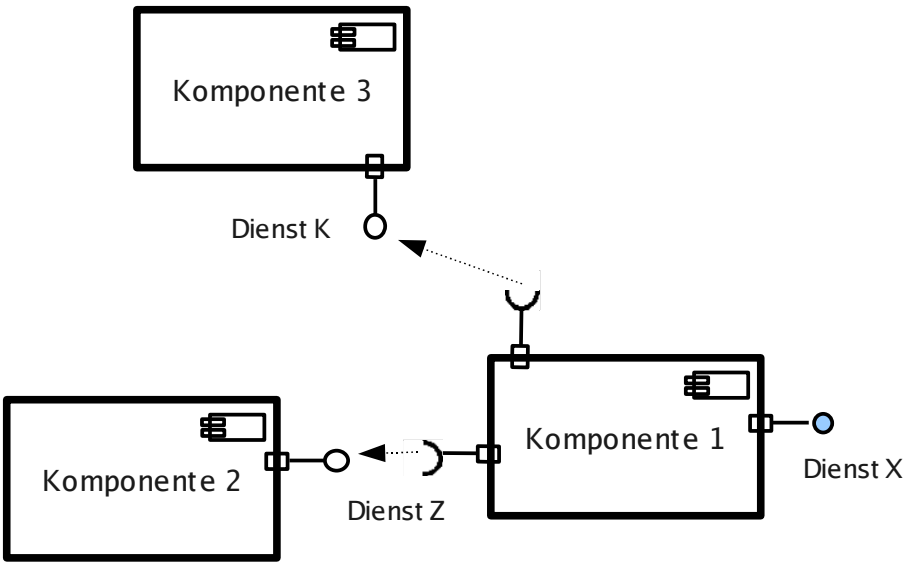

Abbildung 22: Einfache Infrastrukturabhängigkeiten

Das obige UML-Komponentendiagramm zeigt einen *Dienst X,* der von Komponente 1 den Geschäftsprozessen bzw. Prozesssegmenten zur Verfügung gestellt wird. Visualisiert wurde die Schnittstelle zur Prozessdomäne mit Hilfe des hellblau eingefärbten Schnittstellen--symbols. Zur Erbringung dieses Dienstes X muß Komponente 1 allerdings auf die Funktionalität in Dienst Z von Komponente 2 und in Dienst K von Komponente 3 zurückgreifen.
Die obige Darstellung mittels Komponenten ist eine sehr grobe Modellierung von Infrastrukturabhängigkeiten, die einen guten Überblick geben kann, aber dafür nicht genug Detail zur Abbildung eines Praxisproblems enthält. Wie ein sinnvoller Detaillierungsgrad erreicht werden kann, soll in den folgenden Unterkapiteln gezeigt werden.

4.4.2 IT / Infrastrukturebenen

Ähnlich wie in der Prozessdomäne ist auch die Infrastrukturdomäne ein Ausschnitt aus einem

hierarchischen System. Daher ist es vor allem für eine Analyse von IT-Systemen notwendig, die einzelnen Hierarchieebenen und deren Eigenschaften kurz darzustellen. Zur Beschreibung von IT-Architekturen gibt es eine Vielzahl von möglichen Ansätzen. Für den nun folgend gezeigten wurde dabei das in Kap. 2.6 bereits beschriebene ISO / OSI Referenzmodell zugrunde gelegt.

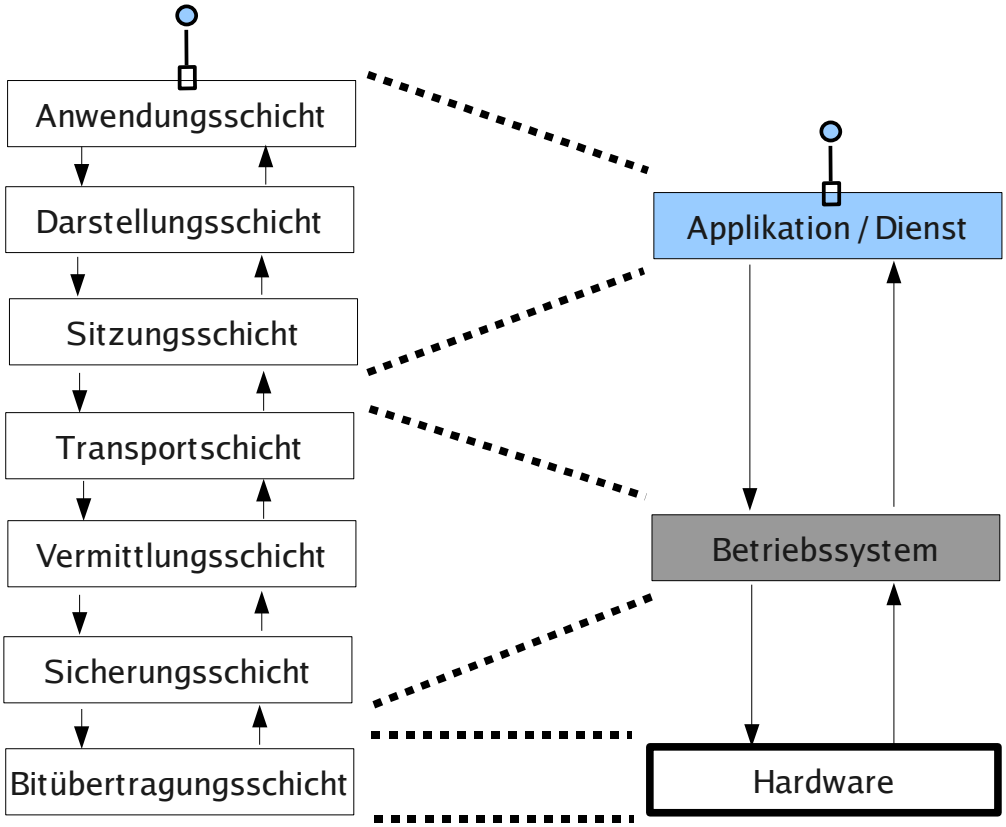

Abbildung 23: Gegenüberstellung ISO OSI Modell / IT- Systemebenen

Die obige Darstellung beinhaltet einerseits ein gängiges Schema, nach dem IT-Systeme aufgebaut sind, und es enthält ebenfalls eine Gegenüberstellung, welche IT-Systemebene in einem überwiegenden Teil der Fälle[82] für die Implementierung der jeweiligen ISO / OSI Ebene verantwortlich ist. Die Schnittstelle zur Prozessdomäne wurde auch hier wieder gesondert gekennzeichnet.
Für das korrekte Funktionieren einer Applikation kann es durchaus notwendig sein, dass diese mit mehreren anderen Applikationen bzw. Diensten auf dem selben oder anderen Systemen kommuniziert bzw. diese in Anspruch nimmt. Dadurch können sich wieder Abhängigkeiten ergeben, wie sie u.a. bereits in Abb. 22 dargestellt wurden. Eine triviale aber wichtige Eigenschaft der IT-Systeme kann in der obigen Darstellung direkt eingesehen werden: eine Komponente, die Funktionalität der Ebene n zur Verfügung stellt, kann nur funktionieren, wenn alle unter ihr liegenden Ebenen (1,...,n-1) ebenfalls korrekt funktionieren.

82 Vgl. die TCP/IP Stacks in MS Windows, SUN Solaris, Linux,...

4.4.3 Detailliertere Infrastrukturmodellierung

Eine korrekte Erhebung der Eintrittswahrscheinlichkeiten von Ereignissen bzw. eine korrekte Identifikation der Ereignisse selbst basiert auf einer akkuraten Aufarbeitung der Abhängigkeiten in der Infrastrukturebene. Aus diesem Grund wird an dieser Stelle noch etwas tiefer auf die bereits in Kapitel 4.4.1 erläuterte Darstellung der selbigen eingegangen.
Als Aufhänger dazu dient eine Erweiterung des Beispieles aus Abb. 22.

Als Erweiterung sei nun angenommen, dass Komponente 1 aus dem obigen Beispiel aus mehreren Subkomponenten besteht, die ebenfalls dargestellt werden sollen.

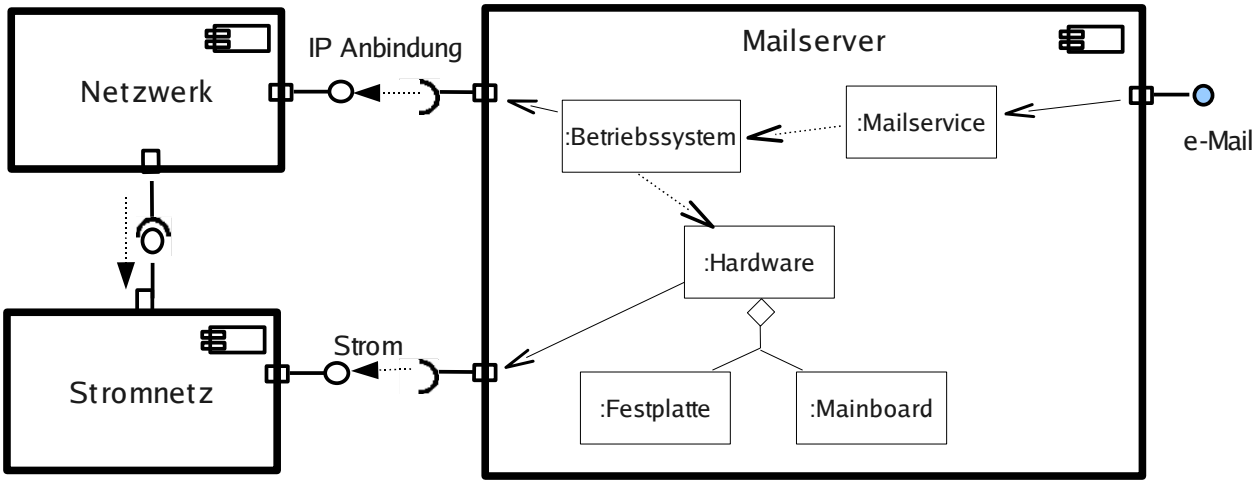

Abbildung 24: Infrastrukturabhängigkeiten unter Berücksichtigung ausgewählter Subkomponenten

Das obige Beispiel zeigt eine mögliche Analyse eines Teils der Infrastrukturdomäne eines Unternehmens, das diverse IT-Infrastrukturen nutzt. In diesem Falle beschränkt sich die Analyse auf das Service e-Mail. Darin enthalten sind alle Komponenten und Subkomponenten, die im Rahmen der Analyse für relevant erachtet wurden, wobei die Darstellung wie folgt zu lesen ist:

- Das Mailservice der Komponente Mailserver ist abhängig von der Subkomponente Betriebssystem.
- Die Subkomponente Betriebssystem ist abhängig von der Hardware des Mailservers und benötigt für die vom Mailservice verlangte Funktionalität den Dienst IP Anbindung von der Komponente Netzwerk.
- Die Subkomponente Hardware benötigt Strom von der Komponente Stromnetz und besteht selbst aus den beiden Subkomponenten Festplatte und Mainboard.

Die Funktionalität IP-Anbindung der Komponente Netzwerk soll im obigen Beispiel das korrekte und zeitgerechte Zustellen aller IP-Pakete in das Unternehmensnetzwerk und das Internet umfassen.

Mit UML ist es, wenn dem obig gezeigten Analyse- und Darstellungsschema gefolgt wird, möglich die funktionalen Abhängigkeiten[83] in der Infrastrukturdomäne bis in sehr große Detailtiefen darzustellen, es ist aber in der jeweiligen Anwendungssituation zu entscheiden ob dies sinnvoll ist, da hier oftmals nur Abhängigkeiten in der Hardware Ebene weiter präzisiert werden.

83 die für die Infrastrukturdomäne auf grund der engen Kopplung zwischen Funktion eines Teils und dessen Platz in der strukturellen Hierarchie der Infrastruktur auch meist den strukturellen Abhängigkeiten entsprechen.

4.4.4 Information als Teil der Infrastrukturdomäne

Controlinformation, Performanceinformation, bzw. Information allgemein in Form von Datenobjekten als Prozessinput und -output sind neben einer funktionierenden Infrastruktur wie schon weiter oben dargestellt eine Grundvoraussetzung für einen funktionierenden Prozess. Da Informationen bzw. Daten nicht für sich existieren können, sondern ein Trägermedium benötigen um transportiert oder gespeichert werden zu können, ist es nahe liegend diese Medien als Teil der Infrastruktur in das Modell zu integrieren.
Damit ein Geschäftsprozess auf die für ihn relevanten Daten zugreifen kann ist es notwendig, dass der Dienst, der die Daten bereitstellt, verfügbar ist. Dienste haben also primär die Aufgabe den Zugriff auf Daten bzw. deren Manipulation zu ermöglichen. Jeder Dienst kann sich dabei aber wieder anderer Dienste[84] bzw. Diensten anderer Komponenten, auf denen die Daten abgelegt / gespeichert sind, bedienen.
Um diese Abhängigkeiten zu integrieren wurde das obige Beispiel derart erweitert, dass der innere Aufbau der Komponente „Mailservice“ nun sichtbar ist. Dabei ist nun zu sehen, dass sich diese aus zwei Subkomponenten zusammensetzen kann: zum einen aus der Komponente „Mailverteilung“, die für das Senden und Empfangen von e-mails zuständig ist und die der Prozessdomäne den Dienst „email“ zur Verfügung stellt, und zum anderen aus der Komponente „Mailobjekte verwalten“, die den Zugriff auf sämtliche gespeicherten e-mails (also die darin enthaltenen Informationen) ermöglicht.
Diese Erweiterung kann wie folgt dargestellt werden.

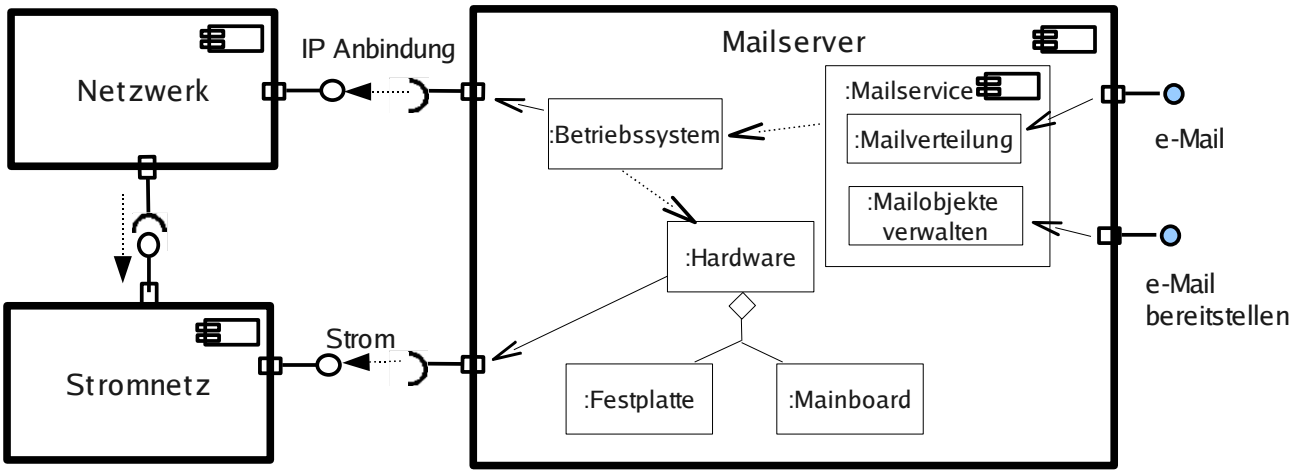

Abbildung 25: Beispiele für die Trennung in datenverarbeitende und datenbereitstellende Dienste

4.4.5 Ereignisidentifikation in der Infrastrukturdomäne

Nachdem in Kapitel 4.3 gezeigt wurde, wie in der Prozessdomäne die relevanten Ereignisse identifiziert werden können, soll dieses Kapitel dazu dienen zu zeigen, wie diese Ereignisidentifikation in der Infrastrukturdomäne stattfinden kann.

4.4.5.1 Verfügbarkeit in der Infrastrukturebene

Um die Methodik für den Übergang von der Strukturanalyse hin zur Ereignisanalyse zu veranschaulichen, eignet sich vor allem die Analyse der Verfügbarkeit in der Infrastrukturdomäne auf grund der Intuitivität der funktionalen Abhängigkeiten. Deshalb soll hier an die Beispielerweiterung aus Kapitel 4.4.3 angeknüpft werden.

84 die meist Aufgaben in einer niedigeren ISO OSI Ebene bzw. IT-Systemebene erfüllen

In einem ersten Schritt ist es notwendig, für jede der relevanten Komponenten und Subkomponenten alle jene Zustände zu identifizieren bzw. zu definieren, in denen sich diese Komponenten befinden können. Der Einfachheit halber nehmen wir an, dass sämtliche Komponenten des Beispiels nur zwei Zustände annehmen können, nämlich „in Betrieb“ und „ausgefallen“.
Der nächste Schritt besteht darin festzulegen bzw. zu erheben, welche funktionalen Mindeststandards eine Komponente auf den von ihr angebotenen Schnittstellen erfüllen muß um sich noch in dem Zustand „in Betrieb“ zu befinden[85]. An diesem Punkt kann z.B. definiert werden, dass der Dienst e-Mail mindestens drei e-Mails pro Sekunde versenden oder empfangen können muß um sich im Zustand „in Betrieb“ zu befinden. In der Praxis lassen sich diese Schwellwerte aus eventuell vorhandenen Service Level Agreements (SLA's) ablesen oder ableiten.
Der dritte Schritt besteht darin, wie bereits beschrieben Ereignisse als Nachrichten über Zustandsübergänge der Komponenten zu betrachten, wobei sich dieser Zustandsübergang jedes Mal dann vollzieht, wenn die definierten Schwellwerte entweder über- oder unterschritten werden. Diese Zustandsübergänge lassen sich sehr gut mit Hilfe von UML Zustandsdiagrammen (sog. statemachines) darstellen, die es ermöglichen sämtliche relevanten Zustände und Ereignisse einer (Sub)Komponente systematisch darzustellen.

Die folgende Abbildung stellt beispielhaft die Zustände und Zustandsübergänge für die Infrastrukturkomponente „Betriebssystem“ dar.

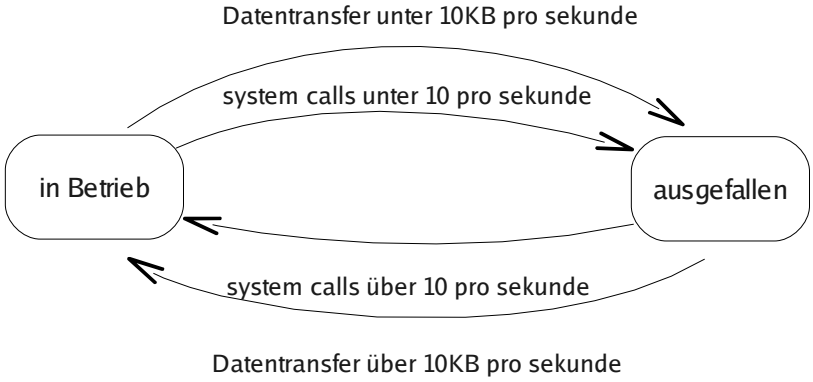

Abbildung 26: Beispiel für ein UML Zustandsdiagramm der Komponente Betriebssystem

Als Kriterien für das korrekte Funktionieren des Betriebssystems wurden die Messwerte

- Anzahl der bearbeiteten bzw. bearbeitbaren Systemaufrufe pro Sekunde und
- Anzahl der übertragenen bzw. übertragbaren Daten in kilo Byte pro Sekunde

herangezogen.
Ist man nun daran interessiert, aus Sicht des Betriebssystems die Gründe für das Unterschreiten der Schwellwerte festzustellen, so ist es notwendig jene Komponenten oder genauer die Schnittstellen der Komponenten, von denen das Betriebssystem abhängt, mit zu berücksichtigen. Abb. 27 zeigt eine solche Erweiterung des ursprünglichen Beispiels.

85 dies ist insofern notwendig, um in der Praxis auftretende Zustandsübergänge - also die Ereignisse - überhaupt messen zu können.

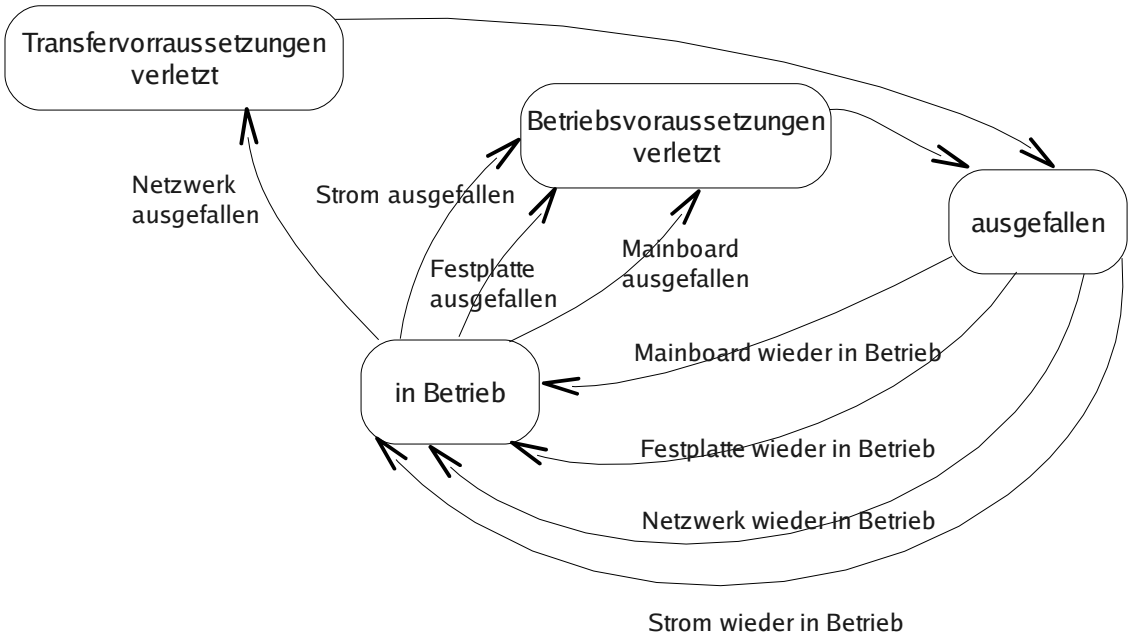

Abbildung 27: Erweiterung um die Zustände von aus OS Sicht benötigten Komponenten

Wie gut zu erkennen, ist führt der Ausfall einer jeden der vier vom Betriebssystem benötigten Schnittstellen zu den genutzten Komponenten auch zum Ausfall des Betriebssystems, der nur durch die Wiederinbetriebnahme der jeweiligen Schnittstelle zu beheben ist. In der obigen Darstellung ist zu beachten, dass jede Komponente nur eine Schnittstelle besitzt, und somit der Ausfall der Schnittstelle mit dem Ausfall der Komponente gleichgesetzt werden kann. Man kann ebenfalls erkennen, dass durch das Einbeziehen der durch den Systemaufbau vorgegebenen Ursachen für die Betriebssystemereignisse *system calls zu gering* bzw. *Datentransferrate zu gering* bereits der erste Schritt einer Top-Down Analyse der Komponentenabhängigkeiten durchgeführt wurde.

4.4.5.2 Top-Down Analyse

Zur Durchführung einer Top-Down Ereignisanalyse in der Infrastrukturebene ist es notwendig die folgenden Schritte unter Umständen in mehreren Iterationen durchzuführen.

1. Beginne bei jener Schnittstelle, die einen Dienst für die Prozessebene zur Verfügung stellt.
2. Definiere jene Zustände der Schnittstelle bzw. der diese Schnittstelle anbietenden Komponente, die analysiert werden sollen.[86]
3. Identifiziere jene Ereignisse, die das Eintreten der zu analysierenden Zustände verursachen (Der Einfachheit halber wird von nun an angenommen, dass nur der Zustand *ausgefallen* von Interesse ist). An dieser Stelle ist es notwendig die möglichen Ereignisse in zwei Gruppen zu teilen, nämlich die der endogenen[87] und die der exogenen[88]

86 Meist wird dies nur der Zustand *ausgefallen* sein, da dieser am leichtesten zu untersuchen ist.

87 endogene Ereignisse sind jene Ereignisse, die aus Sicht der analysierten Komponente Zustandsübergänge von Komponenten anzeigen, die im Systemaufbaumodell und damit im System selbst enthalten sind. z.B.: ist im obigen Beispiel aus Sicht des Betriebssystems die Festplatte Teil des Systemaufbaumodells.

88 exogene Ereignisse sind jene Ereignisse, die aus Sicht der analysierten Komponente Zustandsübergänge anzeigen, die nicht auf weitere im Systemaufbaumodell enthaltene Komponenten zurückgeführt werden können. So ist es zum Beispiel aus Sicht der Festplatte nicht möglich ihren eigenen Ausfall noch auf eine andere Komponente zurückzuführen, womit dieser Ausfall ein exogenes Ereignis ist.

Ereignisse. Zur Bestimmung der endogenen Ereignisse ist es lediglich notwendig, allen Abhängigkeitspfeilen der betrachteten Komponente im Systemaufbaumodell vom Schaft zur Spitze zu folgen und die in unserem Falle Ausfallsereignisse der an der Spitze liegenden Schnittstellen bzw. Komponenten heranzuziehen.
Ist eine Komponente von keiner anderen Schnittstelle des Systemaufbaumodells mehr abhängig, so sind ihre Zustandsübergänge exogen verursacht, und das dazugehörige Ereignis ist exogen.

4. führe die Schritte 2 und 3 so lange durch bis jedes endogene Ereignis auf exogene Ereignisse zurückgeführt wurde.

Die zentrale Fragestellung, die für jede untersuchte Komponente zu klären ist, lautet dabei folgendermaßen: *von welchen Komponenten ist diese abhängig, und welche Ursachen haben die definierten Zustandsänderungen in der untersuchten Komponente?* Der unter Ausführung der obigen Schritte entstehende Fehlerbaum des vorangegangenen Mailserverbeispiels sieht dann folgendermaßen aus.

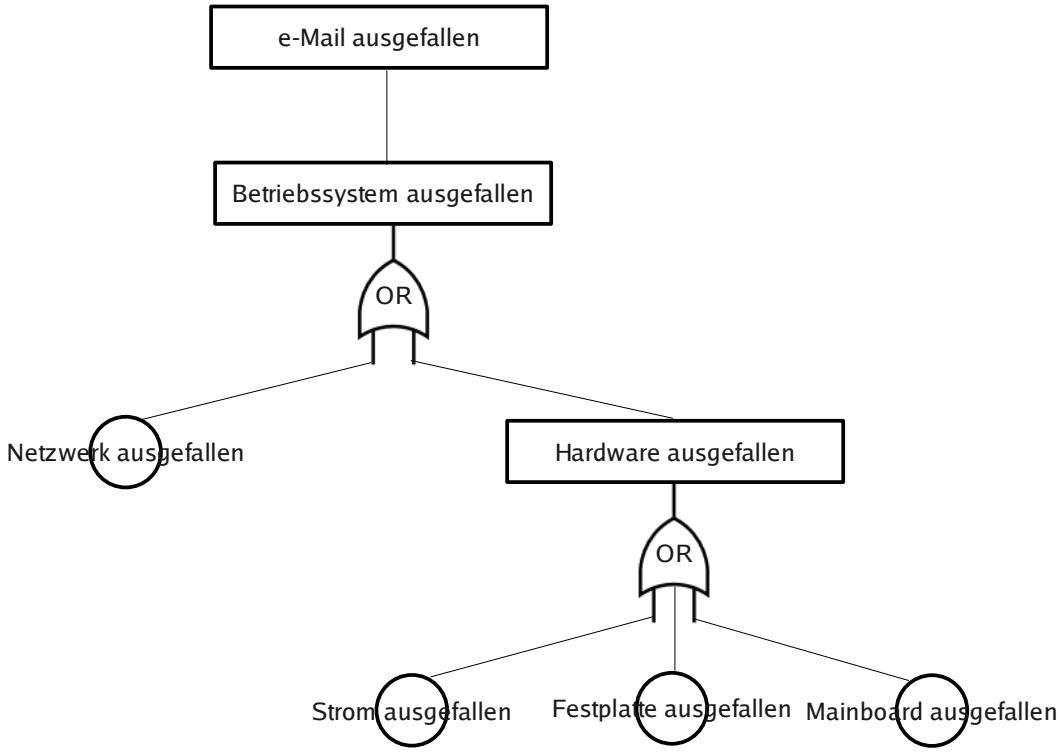

Abbildung 28: Möglicher Ausfallsfehlerbaum für einen Mailserver

Wie man erkennen kann, sind sämtliche Ursachen für den Ausfall des Betriebssystems und damit des e-Maildienstes auch im Zustandsdiagramm weiter oben im Kapitel enthalten. Die Darstellung im Fehlerbaum ist aber weit übersichtlicher und damit auch besser für die Analyse größerer Systeme geeignet.
Dies ist auch dadurch bedingt, dass im Zustandsdiagramm für jede Ebene von Subkomponenten eigene Substati eingefügt werden müssen, um die Zusammenhänge korrekt abbilden zu können.

4.4.5.3 Ausgangspunkt für Top-Down Analysen

Betrachtet man nochmals Schritt 2 der Methode zur Durchführung von Top Down Analysen

aus obigem Abschnitt etwas näher, so kann dieser Schritt der initialen Findung der relevanten Zustandsänderungen bzw. Ereignisse durch das Einbeziehen der in Kap. 3.1 definierten Informationssicherheitsziele erheblich vereinfacht werden.
Denn die Informationssicherheitsziele sind nichts anders als die Sollzustände jeder Komponente, und damit ist jede Abweichung von diesen Sollzuständen ein relevantes Ereignis. (siehe auch Kapitel 5.1.1)

4.5 Infrastrukturinteraktion exogener Ereignisse

Eine genaue Aufarbeitung der aus exogener Sicht vorhandenen Erstangriffspunkte - also jener Komponenten, auf denen ein exogenes Ereignis (also ein Angriff) zuerst auftreten kann - und wie sich diese Ereignisse auf angrenzende Komponenten verbreiten, ist, wie auch aus unten stehender Auflistung (siehe Kapitel 5.3.6) teilweise ersichtlich ist, keine triviale Aufgabe. Insbesondere die Modellierung der Ausbreitung eines Angriffs in einem Unternehmensnetzwerk ist eine schwierige Aufgabe, da keine Netzwerkimplementierung der anderen gleicht. Dies und der Umstand, dass eine solche Aufarbeitung den Rahmen dieser Arbeit sprengen würde, sind die Gründe warum auf diesen Teil des Frameworks verzichtet wird, und direkt die Messung der Ereignisse auf den jeweiligen Komponenten als Ausgangsbasis für die jeweiligen Ereigniszählprozesse und die Ereignisdauerverteilungen herangezogen wird. Vor allem die Arbeiten von Mc Queen [Mc Queen et al., 2006] und Henry [Henry et al., 2009] bieten bezüglich der Modellierung der Ausbreitung im Netzwerk allerdings interessante Ansätze, die sich auch in das hier vorgestellte Modell integrieren lassen.

Die in Kapitel 5.3.6 enthaltene Auflistung von aus Sicht dieses Modells exogenen Ereignissen ist dabei nur eine eher oberflächliche Kategorisierung von aus Informationssicherheitssicht relevanten Ereignissen; sie dient hauptsächlich dazu, die von diesen Ereignissen betroffenen Komponenten und damit die notwendige Analysetiefe in der Infrastrukturebene zu identifizieren.
Als eine Ausgangsbasis für die Analyse der Infrastrukturebene soll dieses und das Unterkapitel 5.3.6 einen Überblick darüber geben, welche grundlegenden Typen von Ereignissen existieren und auf welche Art und Weise diese mit den in Kap. 3.1 definierten Zielen in Konflikt stehen bzw. auf welche Komponenten sie wirken.

4.5.1 IT-Systemabhängigkeiten

Rekapituliert man noch einmal die in Kap. 4.4.2 aufgezeigten Zusammenhänge zwischen den Kommunikationsebenen im ISO OSI Modell und den Komponenten, die die für die Kommunikation erforderlichen Protokolle implementieren, so können die Angriffspunkte der exogenen Ereignisse auf recht einfache Art und Weise den jeweils betroffenen Komponenten zugeordnet werden (siehe wieder Kap. 5.3.6). In weiterer Folge ist es möglich, die auf grund des Angriffstypus und der damit verbundenen Ziele des Angreifers entstehenden Auswirkungen auf die Informationssicherheitsziele der jeweiligen Komponente zu bestimmen.

Um die im Rahmen der Risikoanalyse zu erstellenden Fehlerbäume vollständig aufbauen zu können bzw. damit es möglich ist die durch die exogenen Ereignisse verursachten Basisereignisse zu identifizieren und zuzuordnen, ist es notwendig die Abhängigkeiten der Informationssicherheitsziele über alle IT-Systemebenen hinweg zu berücksichtigen.[89]
Ziel dieses Kapitels ist es, eben diese Abhängigkeiten kurz mit Hilfe der folgenden beiden Diagramme zu skizzieren.

Das unten stehende UML-Zustandsdiagramm zeigt dabei nochmals die trivialen Zusammenhänge für die Verfügbarkeit der Systemkomponenten und der Informationen auf.

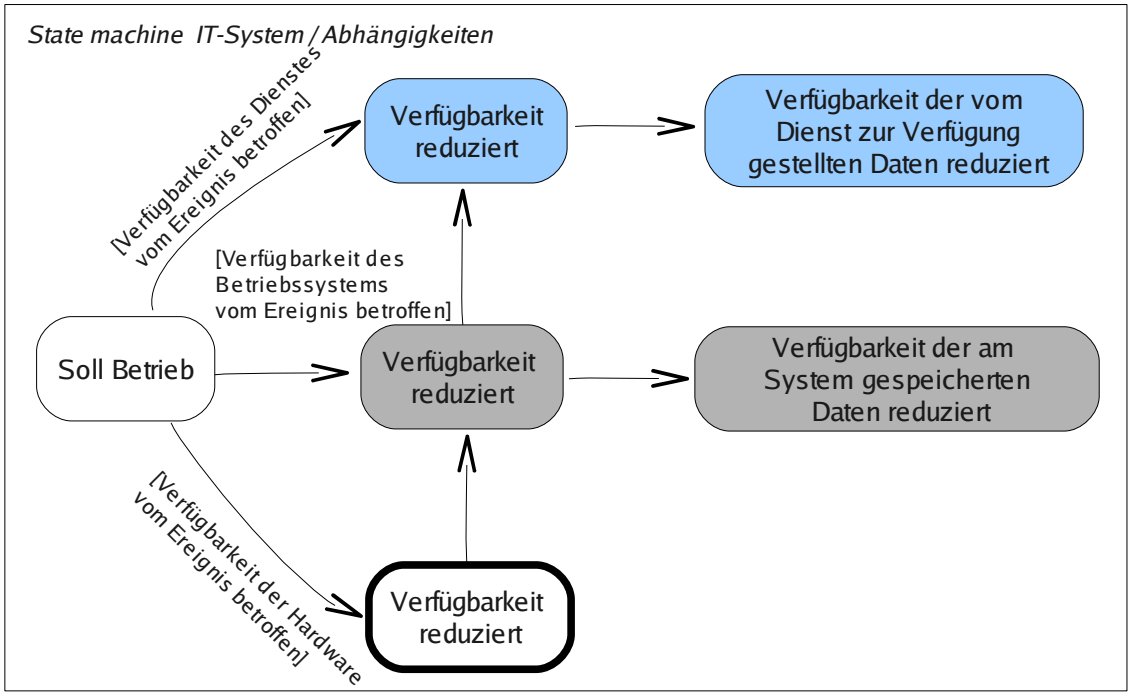

Abbildung 29: Verfügbarkeitsabhängigkeiten zwischen den Ebenen eines IT-Systems

89 Die Abhängigkeiten zwischen den jeweiligen IT-Systemebenen müssen insoferne berücksichtigt werden, um die Ereignisanzahl- und Ereignisdauerverteilungen der jeweiligen Infrastrukturereignisse korrekt bestimmen zu können.

Das folgende Zustandsdiagramm zeigt die Abhängigkeit der Informationsvertraulichkeit von der Integrität der Komponenten an, die auf die Informationen zugreifen können.

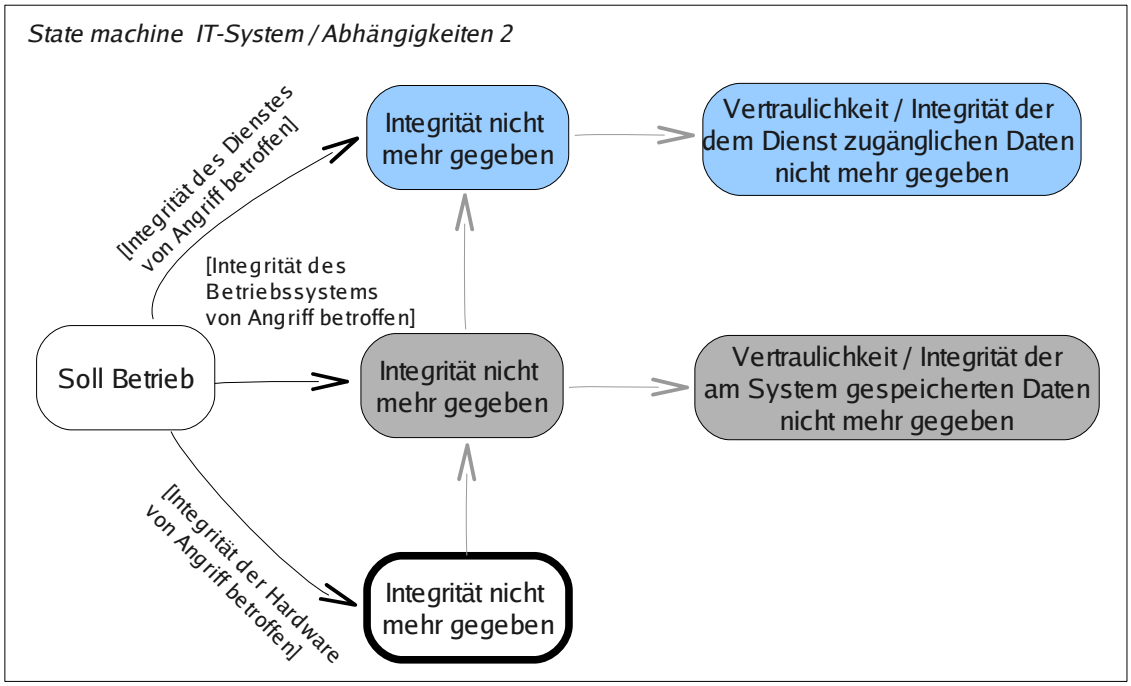

Abbildung 30: Integritätsabhängigkeiten zwischen den Ebenen eines IT-Systems

Die in Abbildung 30 enthaltenen grau dargestellten Abhängigkeiten zwischen den IT-Systemebenen bzw. zwischen Integrität und Vertraulichkeit entspechen dabei jenen hinsichtlich ihres Ausprägungsgrades nicht fix vorgebbaren Abhängigkeiten in der Infrastrukturdomäne, die schon in der Framworkübersicht auf Seite 44 durch graue Pfeile dargestellt wurden.
Der Umstand, dass es nicht möglich ist den Ausprägungsgrad dieser Abhängigkeiten zu fixieren, ist zum einen dadurch begründet, dass es in der Praxis oft eine Fülle von Einflussfaktoren gibt, die nicht in diesem Modell enthalten sind, die aber darüber entscheiden, ob diese Abhängigkeit nun schlagend wird oder nicht. So kann z.B. ein Management durchaus darauf beharren Daten, deren Integrität nicht mehr gesichert ist trotzdem zu verwenden.
Daher wird in diesem Modell der Ansatz verfolgt, all diese Abhängigkeiten mit einem Wert für ihre Aktivierungswahrscheinlichkeit zu hinterlegen, der im jeweiligen Anwendungsfall entsprechend der vorliegenden Gegebenheiten zu wählen ist.
Für die in dieser Arbeit angewandten Beispiele wird jedenfalls angenommen, dass diese Aktivierungswahrscheinlichkeit den Wert eins hat. Dies entspricht der Umsetzung des recht kompromisslos erscheinenden Standpunktes der Informationssicherheits Best Practices, nach dem jeglicher Verlust an Integrität in einer Systemebene sofort den Verlust der Integrität in den darüber liegenden Systemebenen nach sich zieht, und dass jegliche Information auf einem kompromittierten System nicht mehr vertrauenswürdig bzw. nicht mehr vertraulich ist. Dies führt zwar zu einer systematischen Überschätzung der erwarteten Schäden, ist aber - da es fast nie möglich ist festzustellen welche Aktionen ein Angreifer auf einem System wirklich gesetzt hat - der wohl seriöseste Weg, die Auswirkungen des jeweiligen Angriffs zu definieren.

5 Details der Risikomodellierung

In den vorangegangenen Kapiteln 3 und 4 wurde zuerst das Modell im Groben dargestellt und danach im nächsten Detaillierungsschritt dessen Integration in das System Unternehmen gezeigt. In diesem Kapitel sollen nun die noch fehlenden Modelldetails, wie etwa eine Formalisierung des Modells sowie die für dessen praktische Anwendung getroffenen Annahmen und die bestehenden Einschränkungen dargelegt werden.

5.1 Allgemeines

Bevor auf die Modelldetails in der Prozess- und Infrastrukturdomäne eingegangen werden kann, ist es noch notwendig einige allgemeine Modelleigenschaften abzuhandeln.

5.1.1 Ereignisauswahl näher betrachtet

Neben der Zuordnung der Ereignisse zu einer der in Kapitel 4.3.2 aufgeführten Ereigniskategorien ist es, wie schon erwähnt, oftmals notwendig sich zu Beginn einer Risikoanalyse auf die wichtigsten externen Ereignisse, also die Topereignisse zu fokussieren. Zur Bestimmung eben dieser Topereignisse, die ja der Ausgangspunkt der in Kapitel 5.2 dargestellten Top-Down Analyse der Prozessdomäne sind, ist u.a. folgende Vorgehensweise sinnvoll.[90]

- Auflistung aller Erzeugnisse bzw. Dienstleistungen, die aus Sicht des Managements zu den tragenden Säulen des Absatzes bzw. Gewinnes des Unternehmens zählen.
- Auflistung aller Unternehmenswerte, die zu den tragenden Säulen der Nachfrage des Unternehmens zählen. Hier sollten auch all jene gesetzlichen oder vertraglichen Verpflichtungen aufgelistet werden, die das Unternehmen einhalten muß.
- Auflistung aller Unternehmenswerte, deren Ersatz oder Wiederbeschaffung nur mit großen Aufwänden möglich wäre.

Hinsichtlich der Ereignisidentifikation und -auswahl gilt es außerdem zu beachten, dass eine systematische Ereignisidentifikation nur über die Top-Down Analyse der Prozess- bzw. Infrastrukturebene erfolgen kann, da nur auf diese Weise strukturiert jene Ereignisse entdeckt werden können, die bis dahin noch nicht berücksichtigt wurden.
Dabei können durchaus auch die im Zuge der Kapitel 4.2 und 4.3.2 bisher beschriebenen Methoden zur Vorselektion der im Detail zu analysierenden Ereignisse angewandt werden. An dieser Stelle ist anzumerken, dass diese an sich sinnvolle Vorselektion allerdings vor allem in großen und komplexen Unternehmen dazu führen kann, dass relevante Basisereignisse übersehen und durch die Analyse nicht aufgedeckt werden. Daher ist es oftmals sinnvoll auch jene Ereignisse, für die es schon Erfahrungswerte gibt und die auch einer größeren Gruppe von Menschen zugänglich sind, wie etwa die Elementarereignisse Feuer oder Flut, ebenfalls einer tieferen Risikobewertung[91] zu unterziehen.
Das bedeutet in weiterer Folge, dass es sinnvoll sein kann einen Schwellwert zu definieren, der als Entscheidungskriterium dafür dient, ob ein Ereignis ausgewählt werden soll. Dieser Schwellwert sollte sich in der Infrastrukturdomäne auf die Anzahl der von dem bereits

90 Siehe auch [ISO 27005]
91 welche sich dann allerdings der Ereignisbaumanalyse bedienen muß.

bekannten Basisereignis betroffenen Komponenten, und in der Prozessdomäne auf die Anzahl der betroffenen Prozesssegmente beziehen. Ob eine Komponente betroffen ist lässt sich daran beurteilen, ob durch das Ereignis eines der in Kap. 3.1 definierten Infrastrukturziele gefährdet ist oder nicht erreicht werden kann. In der Prozessdomäne ist das Beurteilungskriterium dabei, ob die Einhaltung des(r) Meilensteine(s) des Prozesssements gewährleistet werden kann bzw. ob eines der Prozessziele aus Kap. 3.1 nicht erreicht werden kann.[92]

5.1.2 Fehlerbaum vs. Ereignisbaum

In Kapitel 4.4.5.2 wurde der Fehlerbaum bzw. die mit seiner Erstellung verbundene Top-Down Analyse als einziges Mittel zur Identifikation von Ereignissen verwendet. Dies ist vorallem darin begründet, dass im Gegensatz zur Top-Down Analyse im Zuge einer Bottom-Up Analyse, wie in Kap 2.5.1 angeführt, ein Ereignisbaum für ein gegebenes (exogenes) Ereignis erstellt wird. Mit dem Ereignisbaum Ansatz können daher keine substantiell „neuen“, bisher nicht berücksichtigten Ereignisse gefunden, sondern nur bereits bekannte Ereignisse bewertet werden. Daher ist diese Analyse weniger eine Methode zur Identifikation von Ereignissen als vielmehr eine zur Bewertung.
Hinsichtlich der Bewertung von Ereignissen verfolgt der Ereignisbaum aber nun die gegenteilige Strategie zum Fehlerbaum, d.h. es müsste für eine akkurate Risikobewertung eine Fülle von bereits bekannten exogenen Ereignissen integriert und hinsichtlich ihrer Auswirkungen auf das System untersucht werden. Da diese Herangehensweise in der Praxis nur sehr schwer mit einer systematischen Identifikation und Auswahl der Ereignisse verbunden werden kann, spielen der Ereignisbaum und die Bottom-Up Analyse in diesem Modell nur eine untergeordnete Rolle.

5.1.3 COSO2 Risk Response und Control Activities

Nach einer erfolgreichen Risikoanalyse sieht COSO 2 für ein Risikomanagement die Schritte Risk Response und Control Activities vor. Wie bereits in Kap. 2.1.4 dargestellt, beinhaltet die Aktivität Risk Response jene Entscheidungen des Managements eines Unternehmens, in denen festgelegt wird, wie mit einem durch die Risikoanalyse bestimmten Risko umgegangen werden soll, das heißt welche Taktik (auslagern, verringern, tragen) im Umgang mit dem Risiko gewählt werden soll, und wie diese Taktik operativ durch entsprechende Maßnahmen umzusetzen ist.

Basierend auf der gewählten Taktik und den operativen Maßnahmen werden in der Aktivität Control Activities die meist prozessualen Kontrollmechanismen definiert, die die Umsetzung der gewählten Taktik und der daraus abgeleiteten operativen Maßnahmen überwachen sollen.
In COSO 2 wird im Zusammenhang mit der Aktivität Risk Response und den Control Activities von inhärenten und residualen Risiken[93] gesprochen.
Betrachtet man die von COSO 2 definierten Risk Response Taktiken, so ändert sich das Risiko nur wenn es verringert werden soll, da ja ein Auslagern des Risikos dieses nur verteilt, aber substantiell nicht verändert. Daher werden in den Kapiteln 5.2.5 und 5.3.5 nur jene Sicherungsmaßnahmen bzw. Kontrollaktivitäten betrachtet werden, die das Risiko verringern sollen.

92 Dieses Auswahlkriteruim entspricht wie auch die anderen Methoden zur Vorselektion bereits einer sehr oberflächlichen und in diesem Falle vorweggenommenen Risikoanalyse.

93 Inhärente Risiken sind die Risiken, so wie sie sich nach der Risikoanalyse darstellen, es sind dies also die rohen Risiken. Als residual werden jene Risiken bezeichnet, auf die der gewählte Risk Response mit den dazugehörigen Control Activities angewendet wurde.

Da Kontrollaktivitäten aus Gründen der Übersichtlichkeit in den bisherigen Kapiteln nicht betrachtet wurden, wurden bisher auch nur inhärente Risiken behandelt. Das Instrumentarium zur Berücksichtigung der Kontrollaktivitäten in der Risikoanalyse entspricht aber auch jenem zur Bewertung der Risiken. Deshalb werden die Kontrollaktivitäten in den beiden folgenden Unterkapiteln gemeinsam mit den Rahmenbedingungen und den noch nicht behandelten Details des Modells dargestellt bzw. in dieses integriert.

5.2 Details zur Schadenfunktion / Prozessdomäne

Dieses Kapitel soll das Konzept der Schadenfunktion und deren Ableitung aus dem Aufbau der Prozessdomäne und dem des Marktumfeldes näher vorstellen, wobei an dieser Stelle an die in Kapitel 5.1.1 bereits ausgeführte Beschränkung auf die Topereignisse angeknüpft wird.

Hat ein Unternehmen also durch die Fokussierung auf die Meilensteine bzw. durch die zusätzliche Berücksichtigung der Ereigniskategorien die relevanten obersten Ereignisse identifiziert, ist der Grundstein für eine Top-Down Analyse in der Prozessdomäne gelegt. Diese folgt dabei in weiten Teilen der Vorgehensweise aus der Infrastrukturebene (siehe auch Kapitel 4.4.5.2), mit dem einen Unterschied, dass die Auswirkungen einer Zustandsänderung eines Prozesssegments auf dessen Meilensteine bzw. auf dessen Outputqualität nun das Auswahlkriteruim dafür darstellen, ob diese Zustandsänderung relevant ist. Nur Ereignisse, die Zustandsänderungen anzeigen, welche die Meilensteine des Prozesses gefährden oder sogar verschieben, bzw. die die Qualität des Endproduktes messbar verändern, sind von Interesse.
Eine Top-Down Analyse in der Prozessebene sollte also folgende Schritte beinhalten.

1. Wähle einen für das Unternehmen wirtschaftlich bedeutsamen Meilenstein.
2. Erstelle für diesen Meilenstein den Meilensteingraphen

Hierzu ist es im Allgemeinen notwendig, den zweiten Schritt nach der bereits in Kapitel 4.1.3 beschriebenen Methodik zur Erstellung eines Ablaufgraphen durchzuführen. Dies soll anhand des folgenden Beispiels veranschaulicht werden.

5.2.1 Beispiel Gießerei

Zur Darstellung, wie die Meilensteine aus der Prozessebene gewonnen werden können, soll wieder auf das Beispiel der Gießerei aus Kapitel 4.1.2 zurückgegriffen werden. Es sei angenommen, die Gießerei hätte einen Auftrag zur Produktion von X Ambossen mit Liefertermin Y erhalten, und das Produkt Amboss sei ein Standardprodukt, für das schon alle Konstruktionszeichnungen existieren, sodass deren Produktion nur noch optimal in den Produktionsablauf der anderen hergestellten Produkte eingefügt werden muß (Also die exakte Scheduling information von den Prozessen P3-1 bzw. P2-1/ P2-2 erzeugt werden muß). In der Gießerei sei die Prozessebene weiters so organisiert, dass auf ISA 95 Ebene 3 neben der Erstellung der Konstruktionszeichnungen für die Gußformen auch die grobe Zuordnung von Aufträgen zu den Arbeitswochen und die Überwachung der fristgerechten Lieferung erfolgt. Die Zuteilung der Aufträge zu den Tagen bzw. Stunden wird auf Ebene 2 durchgeführt.

Definiert man nun den Liefertermin Y als zu untersuchenden Meilenstein, so ist das Ziel, das zu diesem Meilenstein erreicht sein muss, durch die vollständige Lieferung der gesamten Menge von Ambossen an den Kunden definiert.

Schritt 1: Zuordnung des Meilensteins zu einem der Prozesse.
Aus den Aufgaben auf ISA Ebene 3 folgt, dass Prozess P3-1 das Ziel hat, die fristgerechte Lieferung der Ambosse sicherzustellen. Daher kann dieser Meilenstein (Y) diesem Prozess 3-1 zugeordnet werden.

Schritt 2: Identifiziere jene Meilensteine des Prozesses, die Basis für den Meilenstein Y sind. Dabei kann festgestellt werden, dass die Erstellung des Prozessoutputs (z.B.: die grobe Zeitplanung, in welchen Wochen dieser Auftrag abgearbeitet werden soll) einen Basismeilenstein darstellt, da dieser ja die Controlinformation für die untergeordneten Prozesse darstellt. Der Performancemeilenstein wird durch die Performance- und Outputmeilensteine der beiden Prozesse P2-1 und P2-2 gebildet.
An dieser Stelle sieht die Darstellung des kritischen Pfades (bzw. Ablaufgraphes) folgendermaßen aus:

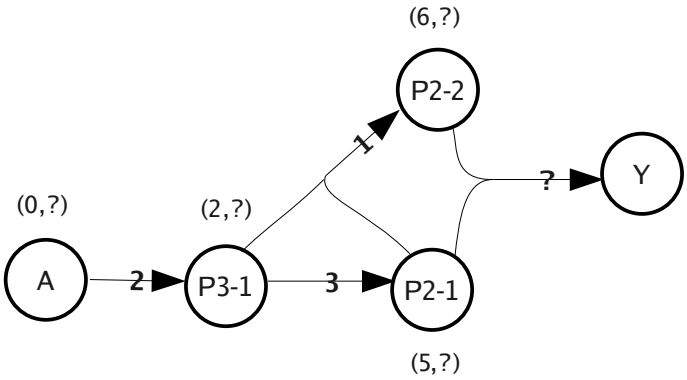

Abbildung 31: Ablaufgraph nach Berücksichtigung der ISA 95 Ebenen 3 und 2

Da die Performancemeilensteine der Prozesse P2-1 und P2-2 noch nicht evaluiert wurden, sind die Übergänge von diesen Meilensteinen zum Meilenstein Y noch nicht quantifizierbar. Es können also zum derzeitigen Zeitpunkt der Analyse nur die Zeitspannen, bis die Prozesse ihren Output erzeugt haben, angegeben werden.

Schritt 3: Eine Analyse der Performancemeilensteine von P2-1 und P2-2 führt zu folgendem Bild.

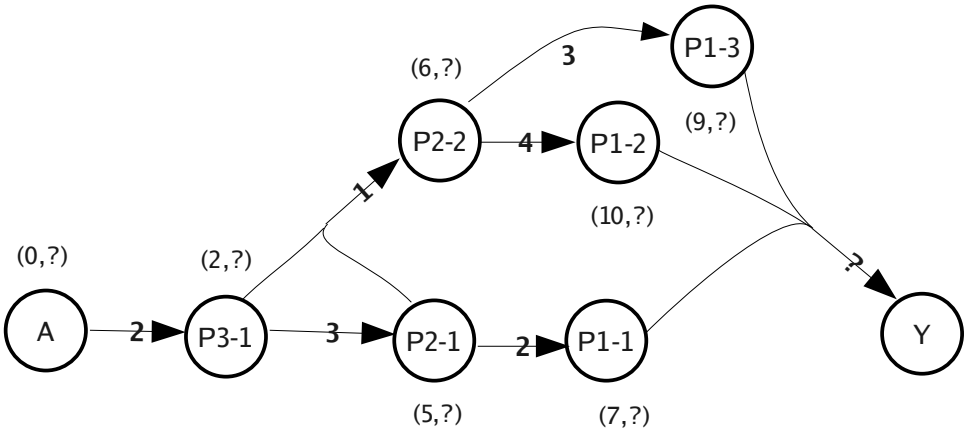

Abbildung 32: Ablaufgraph nach Berücksichtigung der ISA 95 Ebenen 3,2 und 1

Auch in diesem Schritt ist der Graph noch nicht determiniert, da aus der Prozesshierarchie bis

zu Ebene 1 nicht abgelesen werden kann, in welcher Reihenfolge die Prozesse angeordnet sein werden. Dies wird erst im nächsten Schritt offenkundig.

Schritt 4: Analyse der Performancemeilensteine der Ebene 1 Prozesse P1-1 bis P1-3. Da die Performancemeilensteine der Ebene 0 gleich den Outputmeilensteinen der Ebene 0 Prozesse sind, ist die Analyse der Prozessarchitektur nach diesem Schritt abgeschlossen und führt zu folgendem Meilensteingraphen.[94]

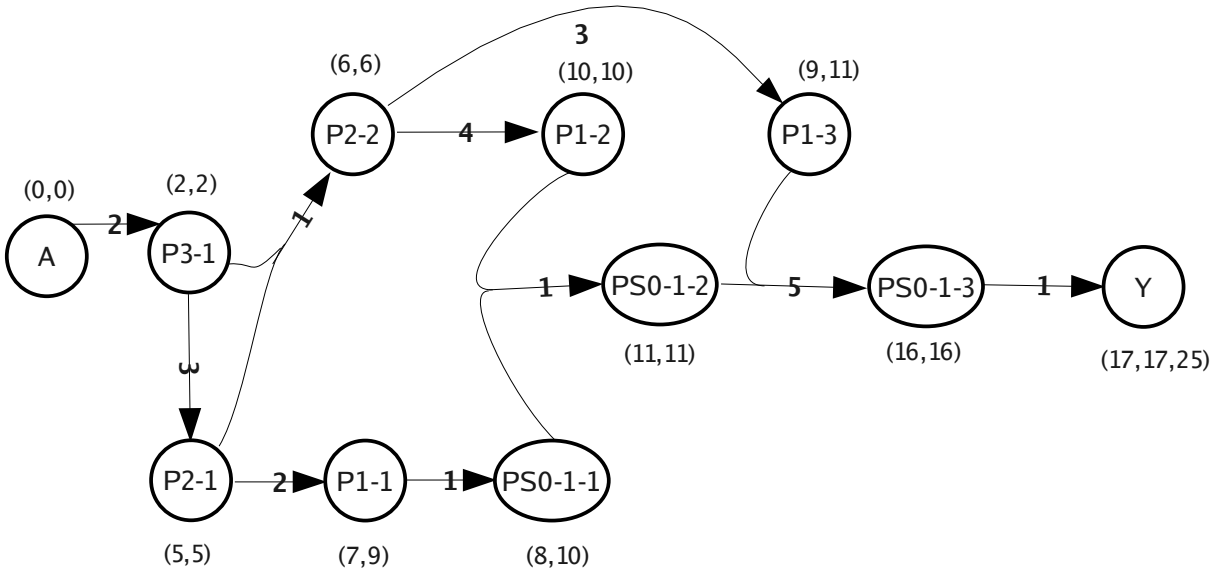

Abbildung 33: Ablaufgraph nach Berücksichtigung aller relevanten ISA 95 Ebenen

Der in der oben dargestellten Abbildung enthaltene kritische Pfad, kann nun auch zur Beurteilung, ob Meilenstein Y gefährdet ist zu verzögern, herangezogen werden. Im obigen Beispiel bildet sich dieser kritische Pfad aus der Prozesssegmentkette P3-1; P2-1; P2-2; P1-2; PS0-1-2; PS0-1-3; Y, und für den betrachteten Meilenstein Y existiert zwischen der frühest möglichen Fertigstellung, die durch den kritischen Pfad determiniert ist, und der spätest notwendigen Fertigstellung (dritter Wert in den Klammern des Meilensteins Y) acht Zeiteinheiten Puffer. Das bedeutet, die Prozesskette scheint relativ wenig anfällig gegen störende Einflüsse, da erst eine Verzögerung von mehr als acht Zeiteinheiten im kritischen Pfad zu einer Verzögerung des Meilensteines führt. Sollte durch Verzögerungen in den Prozesssegmenten P1-1 bzw. PS0-1-1 ein alternativer kritischer Pfad entstehen, so müssen diese sogar mehr als 10 Zeiteinheiten betragen, bevor der Meilenstein Y verzögert ist. (Selbst bei so großen Pufferzeiten ist allerdings zu beachten, dass die allgemeine Auslastungssituation einzelner benötigter Infrastrukturkomponenten durchaus auch bei geringfügigen Störereignissen zu langfristigen Verzögerungen im Prozesssegment führen kann. So ist es z.B.: möglich, dass ein sechsstündiger Ausfall des Schmelzofens genau in jenes Zeitfenster fällt, in dem der Ofen der Ambossproduktion zugeteilt war, und sich das nächste Zeitfenster dafür erst 5 Tage später ergibt. Dadurch würde sich sofort ein neuer kritischer Pfad ergeben, der - sofern eine Zeiteinheit einen Tag repräsentiert - erst frühestens nach 20 Tagen terminiert und somit die Pufferzeit auf 5 Tage reduziert.

94 Hinweis: Die benötigte zeitliche Auflösung bestimmt an diesem Punkt die Detailtiefe und Sichtbarkeit von Submeilensteinen.

5.2.1.1 Erweiterungen im Prozessablauf

Wie dem Leser wahrscheinlich schon aufgefallen ist, ist das Beispiel aus dem vorangegangenen Unterkapitel nur für jene Fälle repräsentativ, in denen ein einziger Amboss produziert werden muß bzw. in denen jeder Produktionsschritt (Prozesssegment) für alle Ambosse abgeschlossen wird bevor der nächste gestartet wird, weil nur dann der strikt sequentielle Ablauf, so wie er im Meilensteingraph dargestellt ist, eingehalten wird. Diese Art der Prozesssegmentorganisation ist aber meist ineffizient und daher in der Praxis eher die Ausnahme. Viel öfter kommt es vor, dass vorgelagerte Prozesssegmente bereits mit den Arbeiten für den Output n+1 beginnen während das nachgelagerte Prozesssegment gerade mit Output n beschäftigt ist. Im obigen Beispiel also wäre es effizient Amboss 9 aus der Gussform zu holen während Amboss 8 gleichzeitig gerade nachbearbeitet wird.
Betrachtet man in diesem Beispiel nun den Meilenstein Y (Fertigstellung und Lieferung aller Ambosse an den Kunden), so kann dieser Meilenstein erst erreicht werden, wenn jeder einzelne Amboss fertiggestellt wurde, also die Meilensteine Y1- Yn erreicht sind. Dadurch ist impliziert, dass Yn erst erreicht werden kann, wenn Yn-1 erreicht ist. Somit sind aber auch alle Meilensteine im Graphen, die zu Yn führen, abhängig von den Meilensteinen, die zu Yn-1 führen. Dies setzt sich rekursiv bis zu Y1 fort.
Da der Startzeitpunkt für jede Tätigkeit, die zu Yn führt, von allen Fertigstellungszeitpunkten (also den Meilensteinen) der entsprechenden Tätigkeiten für die vorangegangen Ambosse abhängen, ist es möglich etwaige durch vorangegangene Prozessdurchläufe verursachte Verzögerungen dem letzten Prozessdurchlauf zuzuordnen.

Erweitert man nun das Beispiel der Gießerei um folgende aufgezählte Annahmen, so kann dieser Prozessablauf durch den in Abblidung 34 gezeigten Meilensteingraphen für Yn dargestellt werden:

- Der Guß aller Ambosse erfolgt in einem Arbeitsschritt und muss unmittelbar, nachdem das Aufschmelzen beendet wurde, erfolgen. -> d.h. PS0-1-1 und PS0-1-2 werden nur einmal durchlaufen und müssen direkt nacheinander abfolgen.
- Mit der Nachbearbeitung wird begonnen, sobald der Guß abgeschlossen ist, und es kann immer nur ein Amboss nachbearbeitet werden.
- Die im Regelprozess P1-3 benötigte Zeit für die erneute Erstellung des Outputs sei vernachlässigbar gering.
- Die zu liefernde Anzahl von Ambossen n sei 10.

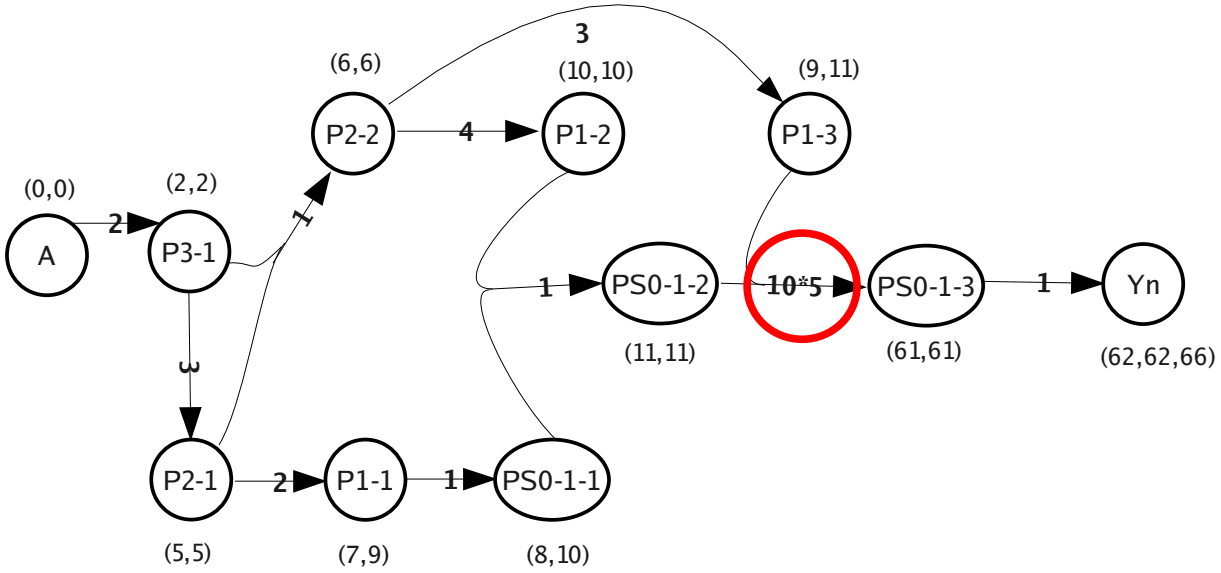

Abbildung 34: Beispiel für einen Ablaufgraph der parallelisierte Prozessschritte berücksichtigt

Im Zusammenhang mit der obigen Abbildung sein darauf hingewiesen, dass hier nur eine von einer Vielzahl von Möglichkeiten für eine vereinfachte Darstellung angewandt wurde. Für die Visualisierung der Eigenschaften und Abhängigkeiten in der Prozessebene der Gießerei erschien dem Autor aber diese als die schlüssigste.
Im Allgemeinen ist es für die Erstellung des Meilensteingraphen grundsätzlich vom Anwendungsfall abhängig, auf welche Art und Weise die Einschränkungen und Scheduling-probleme der jeweiligen Organisation berücksichtigt werden.
Wichtig für die Anwendung des in dieser Arbeit beschriebenen Modells ist hierbei nur, dass ein Meilensteingraph, der die wichtigsten zeitlichen Abhängigkeiten der Prozessebene korrekt widerspiegelt, erstellt werden kann.

5.2.1.2 *Weitere Annahmen zu Abläufen in der Infrastrukturdomäne*

Als Basis für die u.a. in Kap. 4.5 ausgeführte Zuordnung von Ereignissen zur jeweiligen Infrastruktur waren bezüglich der Abläufe in der Infrastukturdomäne die folgenden Annahmen unterstellt:

- Es soll kein Ablauf in der Infrastrukturebene existierten, der nicht durch einen Ablauf in der Prozessdomäne gestartet wird bzw. mit diesem assoziiert ist.[95]
- Die Durchführungsdauer eines Ablaufs in der Infrastrukturebene ist damit bereits in der Durchführungsdauer des assoziierten Ablaufs in der Prozessebene enthalten.
- Die funktionale Koppelung zwischen den Komponenten der Infrastrukturebene sei so eng, dass das Versagen einer Komponente sofort , d.h. ohne Verzug, auch zum Versagen der von dieser abhängigen Komponenten führt, soferne diese Komponente zum Zeitpunkt des Versagens gerade von einem Prozesssegment in Anspruch genommen wird.[96]

95 Diese Annahme ist insoferne auch dadurch bedingt, dass es in einem Unternehmen keine Vorgänge in der Infrastruktur geben sollte, die nicht der Erreichung eines Prozesszieles dienen, da diese ja die Ineffizienz des Unternehmens erhöhen.

96 Von anderer Seite betrachtet bedeutet dies, dass sämtliche vorhandenen Pufferzeiten in der Prozessdomäne enthalten sind, womit ein kritischer Pfad in der Prozessdomäne auch dem kritischen Pfad der gesamten Organisation entspricht.

Ist der Aufbau des Meilensteingraphen erfolgreich durchgeführt, so kann mit dem zweiten Teil der Top-Down Analyse begonnen werden, in der der Fehlerbaum für ein ausgewähltes Topereignis erstellt wird.

5.2.2 Aufbau des Fehlerbaumes

Unter Berücksichtigung der Annahme, dass ein Meilenstein nur dann nicht fristgerecht erreicht wird, wenn das ihm zugehörige Prozesssegment in seiner Ausführung gestört wird, kann der Fehlerbaum direkt aus dem Meilensteingraphen abgeleitet werden, da bei der Erstellung des Meilensteingraphen bereits in gewisser Weise eine Top-Down Analyse der Prozessebene durchgeführt wurde. Es ist nun nur noch notwendig, vom Anfangspunkt aus den Pfeilen vom Schaft zur Spitze folgend durch den Graphen zu gehen.
Die Fehlerbäume, die dabei Schritt für Schritt entstehen, sehen in Anlehnung an das Beispiel der Gießerei folgendermaßen aus:

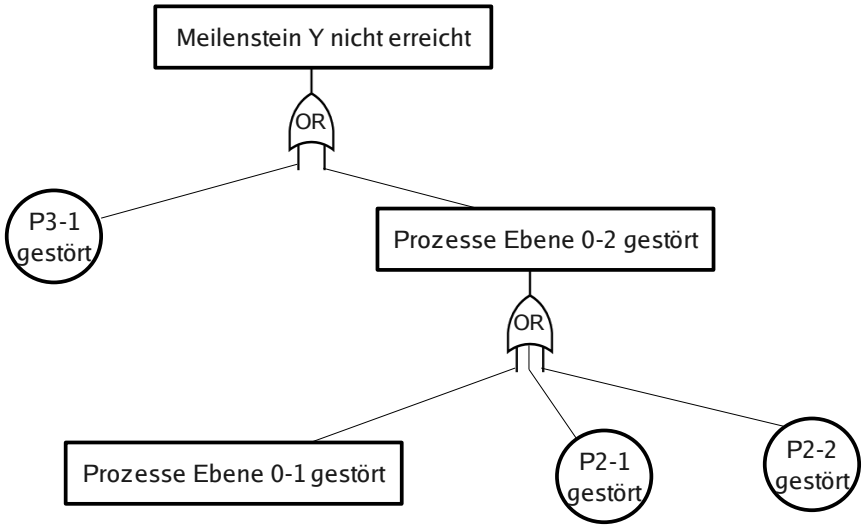

Abbildung 35: Meilensteinfehlerbaum zur Darstellung der Ausfallsabhängigkeiten in der Prozessdomäne

5.2.2.1 Prozessinteraktion

Eine Störung eines Prozesssegments stellt aus Sicht der Prozessebene ein Basisereignis dar. Um die Brücke von der Prozessebene zur Infrastrukturebene zu bauen, ist es notwendig die Ursachen für eine solche Prozessstörung zu identifizieren, wofür die Abhängigkeiten eines Prozesssegments, wie sie in Kapitel 4.1.2.5 bereits dargestellt wurden, noch einmal untersucht werden müssen. Betrachtet man die Voraussetzungen, die für eine korrekte Funktion des Prozesssegments erfüllt sein müssen, näher, so lassen sich die folgenden Ursachen für eine Prozesssegmentstörung bzw. einen Stillstand feststellen:

- Ein Prozesssegment steht still, wenn es keinen ausreichenden Input hat oder dieser nicht die geforderte Mindestqualtität hat.
- Ein Prozessegment steht still oder erbringt nicht die Sollleistung, wenn ihm nicht die benötigte Controlinformation zur Verfügung steht.
- Ein Prozesssegment steht still oder erbringt nicht die Solleistung, wenn ihm nicht die benötigte Infrastruktur zur Verfügung steht.

- Ein Prozesssegment steht still oder erbringt nicht die Solleistung, wenn ihm nicht die benötigten Humanressourcen zur Verfügung stehen.

Appliziert man nun diese Ursachen für einen Prozesssegmentstillstand auf die in Kapitel 3.1 aufgelisteten Infomationssicherheitsziele, so ergeben sich für Prozesse, die auf IT-Infrastruktur aufsetzen unter Berücksichtigung der in Kapitel 4.5.1 bereits für Integritätsereignisse in der Infrastruktur dargelegten Abhängigkeiten der Verfügbarkeit folgende grundlegende Annahmen für die praktische Anwendung dieses Modells.

- Ein Prozesssegment steht still mit Wahrscheinlichkeit p=1, wenn die Inputdaten nicht verfügbar sind, bzw. mit Wahrscheinlichkeit q [0..1], wenn deren Integrität[97] nicht gewährleistet ist.
- Ein Prozesssegment steht still bzw. erbringt nicht die Solleistung mit Wahrscheinlichkeit p=1, wenn seine Controlinformation nicht verfügbar ist, bzw. mit Wahrscheinlichkeit q [0..1] wenn deren Integrität nicht gewährleistet ist.
- Die Qualität des Outputs des Prozesssegments entspricht mit Wahrscheinlichkeit q [0..1] nicht der Sollqualität, wenn die Integrität der Inputdaten oder jene der Outputdaten nicht gewährleistet ist.
- Ein Prozesssegment steht still bzw. erbringt nicht die Solleistung, wenn die von ihm benötigten Infrastrukturdienste nicht verfügbar sind.
- Humanressourcen sind zwar eine wichtige Basis für die Prozesssegmentabläufe, werden aber nicht näher betrachtet.

Durch die Top-Down Analyse in der Prozessdomäne ist es also möglich sämtliche für ein Top Ereignis möglichen ursächlichen Ereignisse bis hinunter zum Übergang zwischen Prozessdomäne und Infrastrukturdomäne, also bis zu den Infrastrukturdiensten, aufzulisten. Als Resultat dieser Analyse erhält man, quasi als Nebenprodukt, auch eine Liste aller Topereignisse für die ein spezielles Ereignis auf einem Infrastrukturdienst der Auslöser sein kann[98]. Diese Liste bildet nun die Ausgangsbasis für die folgenden Überlegungen.[99]

5.2.3 Herleitung der externen Schadenfunktion für Ausfallsereignisse

Betrachtet man den Meilenstein Y aus dem Beispiel der Gießerei (siehe Kap. 5.2.1), so ist die Höhe des Schadens, der durch den Verzug dieses Meilensteines entsteht, abhängig von der Größe des Verzugs. Die Tatsache, dass der Meilenstein nicht eingehalten werden kann, ist im obigen Beispiel ein als relevant identifiziertes Ereignis. Damit ist die Größe des Verzugs des Meilensteines mit der Ereignisdauer, des Ereignisses gleichzusetzen.
Der funktionale Zusammenhang zwischen der Schadenhöhe und der Ereignisintensität(-dauer) wird nun für diesen Meilenstein durch die Schadenfunktion h_y(I) hergestellt.
Aus der Schadenfunktion für den Endmeilenstein und aus den jeweiligen internen Aufwänden zur Ereignisregulierung am jeweiligen Prozesssegment lassen sich nun, wie in Kap. 4.3.3 bereits erwähnt, alle Schadenfunktionen für die vorgelagerten Meilensteine im

97 Also deren Inhalt nicht auf Vollständigkeit und Korrektheit verifiziert werden kann.
98 Diese Liste ist äquivalent zur Liste aller von einem Infrastrukturdienst abhängigen Prozesssegmente.
99 Hinweis: Die Prozessinteraktion bzw. die Prozessdomäne hat eine allgemeine, in Bezug auf die Intensität der externen Ereignisse sehr wichtige Eigenschaft. Die Tatsache, dass oftmals eine Vielzahl von Prozesssegmenten in unterschiedlichen Prozessen von einem Infrastrukturdienst abhängig sind, führt zu einer Vervielfachung der Dimensionen der Ereignisintensität des jeweiligen externen Folgeereignisses, da ja die externe Intensität für jeden Prozess unterschiedlich sein kann.

Meilensteingraphen direkt ableiten, es ändern sich nur die jeweiligen Pufferzeiten d_{ldt}. Es findet also quasi nur ein Parallelverschieben der Schadenfunktion h_y(I) statt, womit dieser Teil der Schadenfunktionen der Meilensteine, die am kritischen Pfad liegen, exakt gleich ist. Das folgende Beispiel zeigt die Schadenfunktionen für ausgewählte Meilensteine aus Kap. 4.1.3 unter der Annahme, dass keine zusätzlichen Aufwände zur Ereignisregulierung am jeweils betroffenen Prozesssegment anfallen.

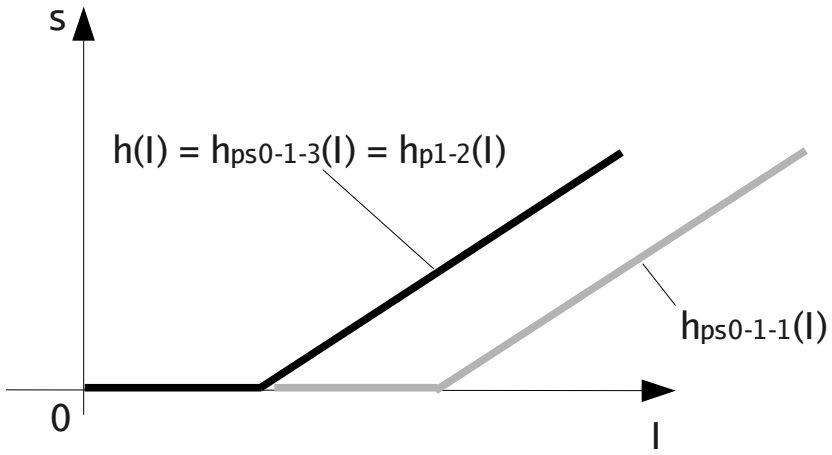

Abbildung 36: Beispiel für verwandte Schadenfunktionen innerhalb eines Ablaufgraphs

Die unterschiedlichen Pufferzeiten d_{ldt} beruhen in diesem Beispiel auf der Annahme, dass die geplanten Fertigstellungszeiten den frühest möglichen Fertigstellungszeitpunkten aus dem Meilensteingrapen entsprechen.
Da nun über diese Methodik[100] mit Hilfe des Meilensteingraphen für jedes Prozesssegment die externe Schadenfunktion determiniert ist, ist es umgekehrt auch möglich, die Ausfallsintensität der externen Folgeereignisse aus jener der Prozesssegmente zu bestimmen (siehe nächstes Kapitel).

5.2.4 Weitere Beiträge zur Schadenfunktion

Die in Kap. 4.3.3 gezeigte Herleitung der Schadenfunktion hat, wie bereits erwähnt, nur für jene Ereignisse Gültikeit, die Verfügbarkeitsziele gefährden. In der Infrastrukturdomäne sind aber auch jene Ereignisse relevant, die Integritätsziele und Vertraulichkeitsziele gefährden. Durch die in Kap. 5.2.2.1 dargestellten Eigenschaften der Prozesssegmente ist es in der Praxis möglich auch die Schadenfunktionen für jene Ereignisse die integritätsgefährdend[101] sind, nach der gleichen Methodik wie für jene, die verfügbarkeitsgefährdend sind herzuleiten. Ereignisse, die Vertraulichkeitsziele gefährden, interagieren hingegen operativ nicht mit der Prozessebene und werden daher als nicht produktionsmengenrelevant erachtet. Sie können aber sehr wohl extern von Kunden messbar sein und sind damit Teil der nachfragerelevanten externen Ereignisse. Daher muß die Schadenfunktion für diese Ereignisse direkt aus den Qualitäts- und Imagezielen gegenüber dem Kunden abgeleitet werden.

100 Hinweis: Diese Ableitung der Schadenfunktion eines Infrastrukturdienstes aus der Schadenfunktion des abhängigen Prozesssegments ist nur für jene Ereignisse möglich, die eine Degradierung der Dienstverfügbarkeit darstellen; für alle anderen Ereignistypen sei hier auf das anschließende Kapitel 5.2.4 verwiesen.

101Hinweis: Ereignisse, die die Integrität der Infrastruktur gefährden, haben aus sich selbst oft kein Schadenpotential, sie haben allerdings stets Verfügbarkeits- bzw. Vertraulichkeitsereignisse zur Folge. Siehe die Kapitel 4.5.1 und 5.2.2.1.

In den folgenden Unterkapiteln werden als Ausgangsbasis zur Bestimmung der jeweiligen Schadenfunktion basierend auf der in Kapitel 4.3.2 bereits vorgestellten Ereigniskategorisierung die allgemeinen Bestandteile der jeweiligen Schadenfunktionen näher beschrieben.

5.2.4.1 Stochastisches Gewinnmodell

Realisierte Schäden manifestieren sich am Ende des Jahres in jedem Unternehmen in einem reduzierten Gewinn. Diesem Umstand folgend werden auch in diesem Modell die Schadenfunktionen an den negativen Abweichungen des Gewinns in Abhängigkeit von den Ereignisintensitäten festgemacht. Das bedeutet, dass die Schadenfunktion für ein Ereignis grundlegend durch die Differenz aus Gewinn ohne Ereigniseinwirkung und Gewinn mit Ereigniseinwirkung, also durch $h(I)=G(I=0)-G(I)$ definiert ist.
Dabei gilt es allerdings zu berücksichtigen, dass der Gewinn keine deterministische Größe ist, sondern selbst eine von weiteren stochastischen Größen abhängige Variable. Diese Eigenschaft des Gewinns lässt sich aus folgendem stochastischen Gewinnmodell[102] für ein Produkt und eine Periode t abbilden.

$\tilde{G}_t = db_t * min(\tilde{X}_t^N ; \tilde{X}_t^P + X_{t-1}^{Lager}) - K_t^{fix} - k_t^{Lager} * max(\tilde{X}_t^P + X_{t-1}^L - \tilde{X}_t^N ; 0)$ Wobei

db_t ... Deckungsbeitrag des produzierten Gutes in Periode t.

$\tilde{X}_t^N ; \tilde{X}_t^P$ ... Nachfrage bzw. Produktionsmenge des Produktes in Periode t.

X_{t-1}^{Lager} ... Lagerstand des Produktes am Ende der Vorperiode.

K_t^{fix} ... Fixkosten in Periode t.

k_t^{Lager} ... Lagerkosten pro Produkt in Periode t.

5.2.4.2 Produktionsmengenrelevante Ereignisse

Betrachtet man vor dem Hintergrund des obig dargestellten stochastischen Gewinnmodells jene externen Ereignisse, die sich auf die Produktionsmenge auswirken, so lassen sich zwei Ursachen für eine verringerte Produktionsmenge ausmachen: zum einen sind dies Ausfallsereignisse in den Prozesssegmenten, die direkt zu Produktionsausfällen führen können, und zum anderen sind dies Probleme mit der Produktqualität auf grund von Integritätsproblemen in den Prozesssegmenten, die zu einem erhöhten Ausschuss führen. Das bedeutet, die Intensität der externen Ereignisse setzt sich aus der Ausfallsintensität und der Integritätsintensität zusammen und lässt sich folgendermaßen modellieren: $\tilde{I}_P^{ext} = a_P(\tilde{I}_a ; \tilde{I}_i)$.

Der mit einem externen Ereignis verbundene Produktionsausfall ist dabei durch die Funktion[103] $\tilde{X}^{P\,aus} = b_P(\tilde{I}_P^{ext})$ definiert[104], und die realisierte Produktionsmenge in einer Periode liegt entsprechend um die Summe der Produktionsausfälle unter der geplanten

102 Das hier dargestellte stochastische Gewinnmodell ist eine Subkomponente des von Schwaiger [Schwaiger, 2001] Kap. 2.1 aufgestellten Modells zur Bildung einer Cash Flow Bilanz. Bezüglich der mathematischen Fundierung dieses Modells sei der Leser auf [Bauer, 2002] verwiesen.

103 Die Funktion b ist hier nun jene Funktion, die (wie in Kapitel 4.3.3 bereits beschrieben) empirisch für jeden Anwendungsfall zu bestimmen ist.

104 Gemäß den in der Produktionstheorie (siehe dazu Fandel [Fandel, 2005] S. 120ff.) beschriebenen Produktionsfunktionen vom Typ C lässt sich der Produktionsausfall aus dem Produkt aus externer Ereignisintensität in der Intensitätsdimension Zeit und Produktionsintensität des betroffenen Prozesses bestimmen.

Produktionsmenge, also $\tilde{X}_t^P = X_t^{P\ plan} - \sum_{m=1}^{M} \tilde{X}_m^{P\ aus}$ mit M gleich der Anzahl externer Ereignisse in der Periode.

5.2.4.3 Nachfragerelevante Ereignisse

Wie bereits weiter oben angeführt haben im Grunde alle Ereignisse, die von Kunden eines Unternehmens wahrgenommen werden, eine Auswirkung auf die Nachfrage der Kunden nach Produkten oder Dienstleistungen des Unternehmens. Zu diesen externen Ereignissen sind neben jenen, die auch die Produktionsmenge beeinflussen, vor allem jene Ereignisse zu zählen, die zu einer Veröffentlichung von aus Sicht des Kunden vertraulichen Informationen führen oder die Qualität der Leistungserbringung auf eine andere Art und Weise beeinträchtigen. Das Schadenpotential dieser Ereignisse errechnet sich grundsätzlich aus dem durch das Ereignis verursachten Rückgang der Nachfrage multipliziert mit dem Preis der Produkte bzw. Dienstleistungen.
Da Kunden bzw. deren Nachfrage grundsätzlich sowohl auf Lieferverzögerungen, Qualitätsprobleme als auch auf die Offenlegung von vertraulichen Informationen und Betriebsgeheimnissen sensibel reagieren können setzt sich die für die Nachfrageänderungen verantwortliche Ereignisintensität I_N^{ext} aus allen drei internen Ereignisintensitätsdimensionen zusammen, folgt also der allgemeinen Funktion $I_N^{ext} = a_N(I_a; I_i; I_c)$, z.B. mit der Funktion a als einfaches Faktormodell: $I_N^{ext} = \alpha_N * I + \beta_N * I_i + \gamma_N * I_c$.

Im Zusammenhang mit externen Ereignissen, die die Nachfrage beeinflussen, ist es wichtig festzuhalten, dass Kunden sich diese Ereignisse meist eine gewisse Zeitdauer, die auch oft länger als eine Betrachtungsperiode andauert merken, und daher auch vergangene Ereignisse bei der Kaufentscheidung zu einem Zeitpunkt berücksichtigt werden können. Im Modell kann dieses Verhalten über ein sogenanntes Intensitätsgedächtnis für das j-te Ereignis abgebildet, werden, welches sich wie folgt formalisieren lässt:

$I_{Nj}^{ext} = c(I_{Nj-1}^{ext}; a_N(I_{a\,j}; I_{i\,j}; I_{c\,j}))$ mit

$I_{Nj}^{ext}; I_{Nj-1}^{ext}$ ... Nachfrageintensität des aktuellen / vorangegangenen Ereignisses.

$I_{a\,j}; I_{i\,j}; I_{c\,j}$ ... Intensitätsdimensionen des aktuellen Ereignisses.

Betrachtet man weiter Unternehmen mit mehr als einem Produkt, so ist auch die obig dargestellte Nachfrageintensität noch nicht vollständig, da an diesem Punkt nur der Einproduktfall berücksichtigt ist. In vielen Fällen ist es aber so, dass auch die Ereignisse in unterschiedlichen Produktlinien auf die jeweils andere Produktlinie einwirken können. Man denke hier nur etwa an einen Autohersteller, der eine bestimmte Charge eines Kleinwagens zurückrufen muß. Fertigt dieser Hersteller auch Limousinen, so kann sich dies auch auf die Nachfrage in dieser Produktgruppe negativ auswirken. Im Grunde sind in einem solchen Fall zwei Aspekte zu berücksichtigen: zum Einen wird die Dimensionalität der externen Intensität um all jene Intensitäten der betroffenen Produkte / Prozesse erweitert, und es muß damit auch die weiter unten beschriebene Funktion b_N dementsprechend multidimensional sein. Zum Anderen lässt sich das Intensitätsgedächtnis nicht auf nur ein Produkt beschränken, sondern ist vielmehr als Nachfrageintensitätsbias des gesamten Unternehmens zu betrachten.[105]

Die Nachfragemenge einer Periode für ein Produkt ergibt sich hier aber im Gegensatz zur

105 Da eine erfolgreiche Anwendung dieses Modellelements in der Praxis aber nicht absehbar ist, wird hier auf eine weitergehende Formalisierung verzichtet.

Produktionsmenge entsprechend der empirisch zu bestimmenden bedingten Verteilungsfunktion $\tilde{X}_t^N \sim F(X^N | I_{Nt}^{ext\ \tilde{kumul}})$ in Abhängigkeit von der kumulierten Nachfrageintensität für dieses Produkt, wobei bei ausreichend genauer Kenntnis der bedingten Verteilungsfunktion auch die Funktion $X^{\tilde{N}\ aus} = b_N(I_N^{\tilde{ext}})$, die ja den Zusammenhang zwischen Nachfrageintensität und Nachfrageausfall für ein einzelnes Ereignis herstellt, ausreichend genau bestimmbar sein sollte. Die Gesamtnachfrage einer Periode ist dann durch $\tilde{X}_t^N = X_t^{\tilde{N}\ null} - \sum_{m=1}^{M} X_m^{\tilde{N}\ aus}$ gegeben, wobei M der Anzahl der externen Ereignisse in der Periode entspricht. $X_t^{\tilde{N}\ null}$ ist dabei die Nachfrage des Marktes ohne negative Beeinflussung durch externe Ereignisse.

An dieser Stelle lässt sich mit Hilfe des stochastischen Gewinnmodells auch bereits die externe Schadenfunktion für ein Produkt wie folgt bestimmen:

$h^{ext\ G}(I_N^{ext}; I_P^{ext}) = G(I_N^{ext} = 0; I_P^{ext} = 0) - G(I_N^{\tilde{ext}}; I_P^{\tilde{ext}})$

5.2.4.4 Pönale Zahlungen

Das Schadenpotential von externen Ereignissen, die pönale Zahlungen nach sich ziehen, ist wie in Kapitel 4.3.2.2 bereits angesprochen meist sehr gut definiert und vertraglich bzw. legislativ geregelt.

Die generische Schadenfunktion einer Periode für diese Ereignisgruppe lautet daher trivialer Weise

$h^{poe} = ((K^{poe} | I_P^{ext}) + (K^{poe} | I_q^{ext}) + (K^{poe} | I_c^{ext}))$ mit

$K^{poe} | I_x^{\tilde{ext}}$ ... Pönalehöhe in Abhängigkeit von der jeweiligen externen Ereignisintensität des Ereignisses j.[106]

$I_{jP}^{\tilde{ext}}; I_{jc}^{\tilde{ext}}$ ... Produktions- und Vertraulichkeitsintensitätsdimensionen für externe Ereignisse j.[107]

$I_{jq}^{\tilde{ext}}$ als Intensitätdimension in Bezug auf Abweichungen bei der zu liefernden Produktqualität dient dabei als Hilfskonstrukt, da in Pönaleverträgen meist neben der Termintreue primär auf die Qualität eingegangen wird. Im Endeffekt ist I_q^{ext} aber das Resultat einer Funktion der internen Integritätsintensität des Ereignisses j, da diese in der Prozessdomäne zu einem Versagen oder im Rahmen einer Sicherungsmaßnahme zu einem bewussten Deaktivieren von internen Qualitätskontrollen führt, so ist $I_{jq}^{ext} = f(I_{ji})$.

Die externe Schadenfunktion für ein Produkt ist nach dem Einbeziehen der Pönalezahlungen soweit vollständig und sieht folgendermaßen aus:

$h^{ext} = h^{ext\ G} + h^{poe}$

106 Die Additivität der jeweiligen Pönalen ergibt sich aus dem Umstand, dass diese nach Ansicht des Autors de facto unabhängig sind.

107 Dies sind jene Intensitäten, die trotz aller internen Sicherungsmaßnahmen noch nach aussen dringen. Sobald diese > 0 sind, sind die Ereignisse auch extern messbar. Diese Intensitäten können im Sinne von COSO2 auch als residuale Intensitäten betrachtet werden.

5.2.4.5 Interne Schadenfunktion von Ereignissen

Die Schadenpotentiale von rein internen Ereignissen, also jenen Ereignissen die extern nicht messbar sind, entstehen primär durch Aufwände für Sicherungsmaßnahmen[108], die kurzfristig aktiviert werden um die Auswirkungen dieser Ereignisse zu kompensieren bzw. um zu verhindern, dass diese Ereignisse extern messbar werden.
Allgemein formuliert entstehen die Schadenpotentiale aus den Aufwänden für eine zeitlich längere oder alternative Prozessdurchführung, die nach Eintritt des Ereignisses die Einhaltung eines oder mehrerer Prozessmeilensteine(s) sicherstellen sollen, bzw. aus der durch den Eintritt des Ereignisses notwendigen Durchführung von Reparaturprozessen, die die ursprüngliche Prozessdurchführung wiederherstellen sollen.[109]
Fällt zum Beispiel in der obig beschriebenen Gießerei der Schmelzofen für fünf Tage aus, so kann es trotzdem nach dessen Reparatur möglich sein, durch Überstunden in den nachgelagerten Produktionsprozesssegmenten den Meilenstein Y - also die fristgerechte Lieferung beim Kunden - zu halten. Die Schadenpotentiale für das Ereignis *Ausfall des Schmelzofens* setzen sich dann also aus den Reparaturkosten für den Schmelzofen, den Mehraufwänden für ein erneutes Aufheizen und den Mehraufwänden für die Überstunden in den nachgelagerten Prozesssegmenten zusammen.
Modifiziert man das hier angeführte Beispiel dahingehend, dass der Ausfall des Schmelzofens durch einen Stromausfall bedingt ist, der zu keinen Schäden am Schmelzofen führt und nur eine dreistündige Unterbrechung darstellt, so führt ein solches Ereignis zu keiner Verzögerung des Meilensteines Y; es verursacht aber trotzdem Kosten, da der Schmelzofen erneut aufgeheizt werden muß.

Die Schadenfunktion für diese Schadenpotentiale ist im Allgemeinen einfach zu erstellen, da die oben beschriebenen Mehraufwände im Unternehmen meist bekannt sind oder sehr genau geschätzt werden können. Es ist in diesem Zusammenhang allerdings zu bedenken, dass die Pufferzeiten dieser Schadenfunktionen weit kleiner sein können als die der extern bedingten Schadenfunktionen. Dies wird dadurch hervorgerufen, dass die Sicherungsmaßnahmen, die diese Aufwände verursachen, ja schon möglichst früh nach dem Eintritt eines Serviceereignisses aktiviert werden müssen um effektiv und effizient sein zu können. In weiterer Folge bedeutet dies, dass die Anzahl der Kleinschäden durch die Berücksichtigung dieser Ereignisse stark steigen kann, auch wenn die Anzahl oder Häufigkeit von Großschäden durch den Einsatz der Sicherungsmaßnahmen verringert wird.

Die interne Schadenfunktion für ein einzelnes Ereignis hat nun basierend auf obiger Unterteilung folgende Bestandteile:

$$h^{intern}(I_a; I_i; I_c) = (K_r|I_a; I_i) + (K_s|I_a; I_i; I_c)$$

$K_r|I_a; I_i$ ... Aufwände für Instandsetzungs- bzw Reparaturarbeiten in Abhängigkeit von den Ereignisintensitäten.

$K_s|I_a; I_i; I_c$ ... Aufwände für Sicherungsmaßnahmen bzw. Kontrollaktivitäten in Abhängigkeit von den Ereignisintensitäten.

108 Auch der Mißbrauch von Zugriffsrechten und Entscheidungsbefugnis, der zu finanziellen Schäden führt, also Fraud ist prinzipiell Teil der operativen Ereignisse, ist aber durch die Fokussierung der Arbeit auf die in Kap.3.1 definierten Ziele nur enthalten, wenn er sich auch als externes Ereignis manifestiert, da der Mißbrauch selbst nicht zwangsweise eines der angeführten Ziele gefährdet. Die Integration dieses Schadenpotentials kann aber grundsätzlich auch nach dem in diesem Kapitel beschriebenen Schema erfolgen.

109 Die Durchführung von Reparaturprozessen führt im Allgemeinen zu einer Schadenfunktion, die ab Eintritt des Ereignisses größer Null ist und nicht erst nach Ablauf einer etwaigen Pufferzeit.

I_a ... Ausfallsintensitätsdimension eines Ereignisses.
I_i ... Integritätsintensitätsdimension eines Ereignisses.
I_c ... Confidentiality Intensitätsdimension eines Ereignisses.

Die hier angeführten und auch die in den vorangegangenen Kapiteln verwendeten Intensitäten können grundsätzlich auch selbst mehrdimensionale Größen sein, in weiterer Folge wird in dieser Arbeit vor allem in dem in Kapitel 6 angeführten Beispiel als einzige Ausprägung für die Dimension der drei obigen Intensitäten jeweils die Dauer des Ereignisses verwendet.

5.2.4.6 Schadenfunktion eines Prozesses

Auf dem Weg zur Bestimmung der Schadenfunktion für einen Prozess kommt die schon in den Kapiteln 4.3.3 und 5.2.3 angeführte Eigenschaft, dass die externe Schadenfunktion für alle Prozesssegmente eines Produktionsprozesses gleich ist, zum Tragen.
Dadurch müssen nur noch die internen Schadenfunktionen der einzelnen Prozesssegmente aggregiert und zur externen hinzu gezählt werden. Die Schadenfunktion für einen Produktionsprozess, der nur ein einzelnes Produkt erzeugt, sieht daher folgendermaßen aus.

$$h^{proz} = \sum_{j=1}^{n} h_j^{intern} + h^{ext} \quad ,$$

wobei n gleich der Anzahl der von einem Ereignis betroffenen Prozesssegmente ist.

Ein wichtiger Teil der Schadenfunktion der Prozesssegmente ist nun auch die Festlegung der Funktion a_p zur Berechnung der externen Ereignisintensität der Produktion.
Wendet man hier das Konzept des Meilensteingraphen an und beschränkt man sich weiter auf jene Fälle, in denen die Integritätsintensität gleich Null ist oder keine Auswirkungen auf die Produktionsprozesse hat, kann die externe Ereignisintensität wie folgt aus der internen Ausfallsintensität abgeleitet werden: $\tilde{I}_p^{ext} = max(0; \tilde{I}_a - min(I_{s1}; ...; I_{sn}))$, wobei I_{si} die aus dem Meilensteingraphen ableitbare Intensitätsschwelle(= Pufferzeit) des i-ten Prozesssegments ist.

Aus Sicht des Autors erscheint es an dieser Stelle sinnvoll auch eine einfache Methode zur Integration der Integritätsintensität eines Ereignisses vorzustellen.
Und zwar sei hierzu angenommen, daß der Intensität der Integritätsdimension die gleichen Messgrößen wie der Intensität der Ausfalldimension zu Grunde liegen, und dass diese in der Prozessdomäne immer (d.h. Aktivierungswahrscheinlichkeit gleich eins) auch zu Ausschuss führen, der vollständig von den internen Qualitätskontrollen erkannt wird.[110] Dadurch wird der durch Produktionsausfälle in Folge eines Integritätsereignisses verursachte Schaden zwar systematisch überschätzt, es ist aber möglich die Intensität des Ereignisses auf einfache Art und Weise in die obige Formel zu integrieren. Die externe Intensität des Ereignisses in der Produktion ergibt sich dann wie folgt.

$$\tilde{I}_p^{ext} = max(0; max(\tilde{I}_a; \tilde{I}_i) - min(I_{s1}; ...; I_{sn}))$$

Für den Einsatz dieses Modells in der Praxis sei an dieser Stelle auch darauf hingewiesen, dass es - bedingt durch den stochastischen Charakter der Nachfrage im stochastischen Gewinnmodell - grundsätzlich nicht möglich ist die externe Schadenfunktion zu ermitteln, diese muß vielmehr aus einer großen Anzahl von Perioden und deren Nachfragerealisationen gemittelt werden.[111]

110 Dies bedeutet in weiterer Folge auch, dass Integritätsereignisse keinen Einfluss auf die Nachfrage haben können.
111 In vielen Fällen ist auch die interne Schadenfunktion nicht exakt, und somit nur gemittelt bestimmbar.

5.2.4.7 Erweiterter Schadenprozess

Sämtliche bis zu diesem Punkt skizzierten Schadenfunktionen basierten implizit auf der Annahme, dass Ereignisse nur dann eintreten, wenn die jeweils von diesen Ereignissen betroffene Infrastruktur gerade von einem Prozesssegment benutzt wird. Dies ist dadurch begründet, dass als Pufferzeit / Intensitätsschwelle jeweils nur die Pufferzeit, die sich aus dem Meilensteingraphen herleiten lässt, verwendet wurde. Tatsächlich kommt es in der Realität allerdings oft vor, dass ein Ereignis ja auch durchaus schon lange bevor ein Prozesssegment auf eine Infrastruktur zugreifen möchte eintreten kann. In einem solchen Fall stellt die bisher verwendete Pufferzeit allerdings nur das Minimum der tatsächlichen Pufferzeit dar.
Das bedeutet, dass durch das stochastische Auftreten der Ereignisse grundsätzlich auch die Pufferzeit zwischen Ereigniseintritt und dem Beginn eines etwaigen Meilensteinverzugs eine Zufallsvariable[112] ist.
In Erweiterung des Beispiels aus Kap. 3.3.4 kann ein beispielhafter Schadenverlauf für mehrere Ereignisse, der auch diese Eigenschaft berücksichtigt, dann z.B. wie folgt aussehen:

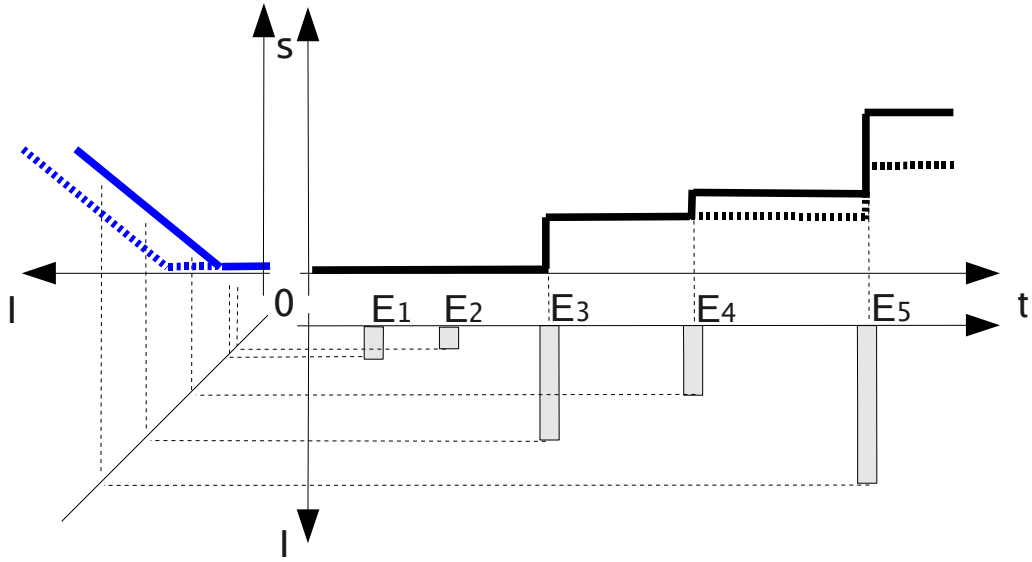

Abbildung 37: Schadenfunktionen und Schadenprozess bei stochastischen Pufferzeiten

Der obige Zeitstrahl in der Abbildung beinhaltet wieder die Ereignisse E1 bis E5. Die strichlierten Linien zeigen nun eine alternative Schadenfunktion mit längerer Pufferzeit und den resultierenden Schadenprozess, wenn die Ereignisse E4 und E5 nun mit dieser alternativen Schadenfunktion bewertet werden. Es zeigt sich, daß durch die längere Pufferzeit die Wirkung von E4 völlig ausgelöscht wird und auch der Schaden von E5 nun weit geringer ausfällt.
Eine wichtige Eigenschaft der Schadenfunktion ist also deren Abhängigkeit vom Zeitpunkt, zu dem ein Ereignis eintritt. Dies bedeutet, dass die Schadenfunktion selbst eine dynamische Komponente des Modells ist und sich im Zeitverlauf ändern kann.

112 Will man diese Eigenschaft der Pufferzeit ebenfalls berücksichtigen und stehen nicht genügend statistische Daten von Ereignissen zur Verfügung, so kann ein Durchschnittswert für den stochastischen Anteil an der Gesamtpufferzeit eines Prozess – Infrastruktur Paares aus dem Mittelwert aller Abstände zwischen den Zeitspannen, zu denen das jeweilige Prozesssegment die Infrastruktur in Anspruch nimmt, gebildet werden.

5.2.5 Risk Responses in der Prozessdomäne

Risk Responses haben generell zwei Ansatzpunkte, über die sie ein Risiko, dem sie entgegenwirken sollen, reduzieren können, und zwar über eine Reduktion der Eintrittswahrscheinlichkeit und/oder über eine Reduktion des Schadenpotentials.
Durch die Risikobetrachtung von Ereignissen, die die Infrastrukturdienste betreffen, und durch die Herleitung der Schadenfunktion für diese Dienste aus den Eigenschaften der Prozessdomäne wirken Sicherungsmaßnahmen in diesem Modell in der Prozessdomäne nur auf die Schadenfunktion und damit auf die Schadenpotentiale. Hierbei gibt es zwei unterschiedliche Ansätze, die eine Sicherungsmaßnahme verfolgen kann um die Schadenfunktion eines Prozesssegments zu modifizieren: die Erhöhung der Intensitätsschwelle und eine Reduktion der Steigung.

5.2.5.1 Erhöhung der Intensitätsschwelle

Die Erhöhung der Intensitätsschwelle der Schadenfunktion eines Prozesssegments ist wie unten abgebildet nichts anders als ein Parallelverschieben der Funktion entlang der Ordinate.

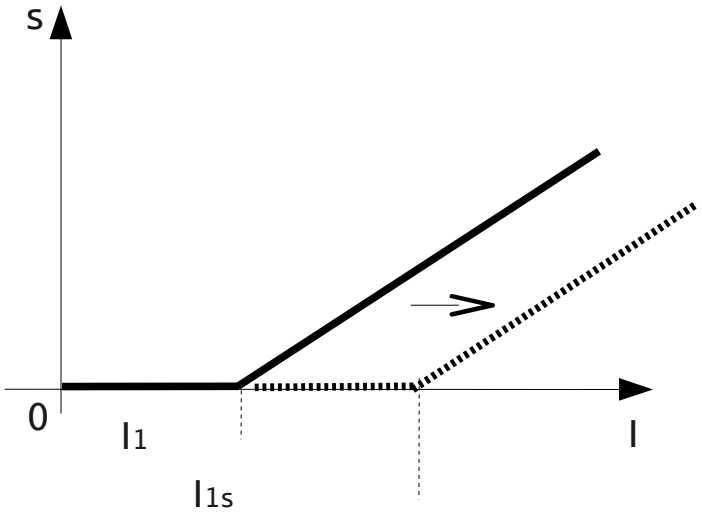

Abbildung 38: Integration von Risk Responses in die Schadenfunktion 1

Die Intensitätsschwelle I_1 wird dabei durch die Sicherungsmaßnahme auf I_{1s} vergrößert, und damit reduziert sich auch die zu erwartende Höhe eines Schadens, da $h_s(I) \leqslant h(I) \forall I$.

Eine Vergrößerung der Intensitätsschwelle kann im Falle der Ereignisdauer als einzige Dimension der Intensität grundsätzlich entweder durch eine Reduktion der Durchlaufzeit eines Prozesssegments im Rahmen eines Prozessreengineerings oder durch eine flexiblere Gestaltung der Liefertermine erreicht werden, wobei eine flexiblere Gestaltung der Liefertermine für die meisten Unternehmen auf grund des Konkurrenzdrucks allerdings leider keine wirkliche Alternative ist. Die Reduktion der Durchlaufzeit ist vor allem für jene Prozesssegmente die im kritischen Pfad des gesamten Prozesses liegen, eine effektive Maßnahme, da bei knapp kalkulierten Lieferterminen gerade im kritischen Pfad keine oder nur sehr geringe Pufferzeiten vorhanden sind. Neben der effizienteren Gestaltung des Prozesssegments selbst ist auch die parallele Abarbeitung der Aufgaben des Prozesssegments eine Möglichkeit, die Durchlaufzeit für das jeweilige Prozesssegment zu verringern.

5.2.5.2 Reduktion der Steigung

Eine Reduktion der Steigung entspricht einer Abflachung der Schadenfunktion und damit verbunden auch einer Reduktion der Schadenhöhe eines Ereignisses.

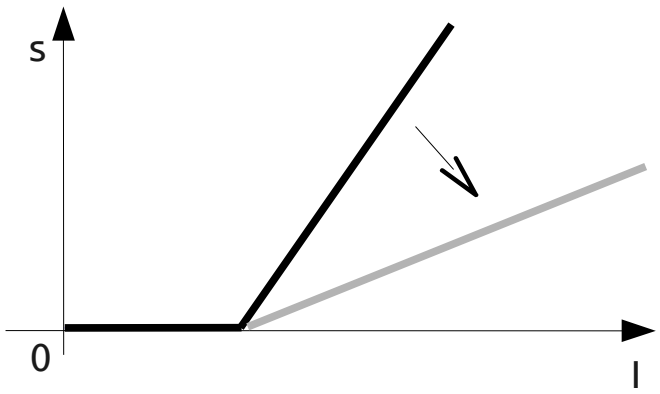

Abbildung 39: Integration von Risk Responses in die Schadenfunktion 2

Für die gewünschte Abflachung der Schadenkurve gibt es wieder zwei Ansatzpunkte.

- durch Parallelisierung des gesamten Prozesssegments, d.h. durch die parallele Ausführung mehrerer gleichartiger Prozesssegmente führt der Ausfall eines der Prozesssegmente nur zu einer Reduktion des Durchsatzes.[113]
- durch Verlagerung der betroffenen Prozesssegmente auf alternative Infrastruktur im Anlassfall.

5.2.5.3 Verlagerung

Die Bereitstellung alternativer Infrastruktur bzw. die Verlagerung eines Prozesssegments auf alternative Infrastruktur im Fall eines Infrastrukturereignisses ist die wohl am häufigsten angewandte Sicherungsmaßnahme, weshalb deren mögliche Wirkung auf die Schadenfunktion des Prozesssegments bzw. des Infrastrukturdienstes näher untersucht werden soll. Ausgangspunkt der Betrachtungen sind dabei die Schadenfunktionen der Infrastrukturdienste A und B, die der Einfachheit halber nur von jeweils einem Prozesssegment verursacht werden.

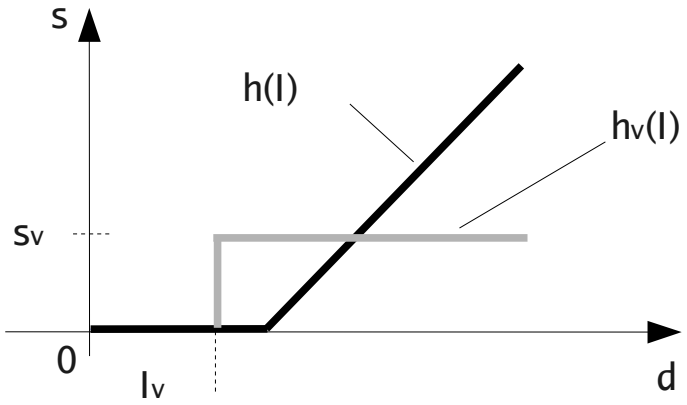

Abbildung 40: Integration der Verlagerung in die Schadenfunktion

113 Dies setzt allerdings voraus, dass die parallel geschalteten Prozesssegmente auf unterschiedlicher Infrastruktur aufsetzen.

$h(I)$ ... Schadenfunktion eines Infrastrukturdienstes

$h_v(I)$ ... Schadenfunktion eines Infrastrukturdienstes nach der Verlagerung des Prozsssegments auf einen alternativen Infrastrukturdienst.

s_v ... Aufwände, die durch die Verlagerung des Prozessegments auf den alternativen Infrastrukturdienst entstehen. Diese sind Teil der Kosten K_s für interne Ereignisse (siehe Kapitel 5.2.4.5)

I_v ... Intensität, die die Verlagerung initiiert.

Durch die Verlagerung des Prozesssegmentes auf den alternativen Infrastrukturdienst B wird auch die Schadenfunktion vom ursprünglichen Infrastrukturdienst auf den neuen verlagert und durch eine Schadenfunktion ersetzt, die nur die Migrationsaufwände[114] beinhaltet.
Von Seiten des aufnehmenden Infrastrukturdienstes gilt es dabei aber nun drei Fälle zu unterscheiden:

- Die Kapazitäten des Infrastrukturdienstes B reichen für die Bedienung beider Prozesssegmente aus. In diesem Fall wird die Schadenfunktion des verlagerten Prozesssegments nun einfach zur bisher bestehenden Schadenfunktion hinzuaddiert.
- Die Kapazitäten des Infrastrukturdienstes B reichen nur für eines der beiden Prozesssegmente, und das verlagerte soll priorisiert bedient werden. In diesem Fall findet ein kompletter Transfer der Schadenfunktionen der nicht mehr bedienbaren Prozesssegmente zum Infrastrukturdienst A statt.
- Die Kapazitäten des Infrastrukturdienstes B reichen nicht für beide Prozesssegmente, und beide Prozesssegmente werden als gleichwertig betrachtet. Dieser Fall stellt eine Kombination aus beiden obigen Fällen dar, wo nur jener Teil der Schadenfunktion zum Infrastrukturdienst B verlagert wird, der der verlagerten „Dienstkapazität[115]“ entspricht. Der restliche Teil der Schadenfunktion bleibt dem Prozesssegment A zugeordnet. Es muß hier auch jener Teil der Schadenfunktion, der der nicht mehr für das ursprüngliche Prozesssegment B zur Verfügung stehenden „Dienstkapazität“ von Infrastrukturdienst B entspricht, auf Infrastrukturdienst A verlagert werden (Siehe dazu Abbildung 41 auf Seite 87).

114 In diesen Migrationsaufwänden sind auch jene Aufwände enthalten, die für die Rückverlagerung des Prozesssegments auf den ursprünglichen Infrastrukturdienst anfallen.

115 Unter Dienstkapazität ist die vom Infrastrukturdienst zur Verfügung gestellte Funktionalität zu verstehen. Z.B.: bei einem Infrastrukturdienst e-mail, der 10000 Mails pro Stunde verarbeiten kann, ist dieser Wert die Dienstkapazität. Benötigt ein Prozesssegment nun eine Kapazität von 6000 Mails pro Stunde, und kann ein alternativer Mailserver nur 4000 zusätzliche Mails abarbeiten, dann kann nur die Abarbeitung dieser 4000 Mails auf den anderen Server verlagert werden. Die 2000 verbleibenden Mails können nun durch eine Schadenfunktion mit geringerer Steigung als jene für alle 6000 repräsentiert werden; diese Schadenfunktion bleibt dem ursprünglichen Mailserver zugeordnet.

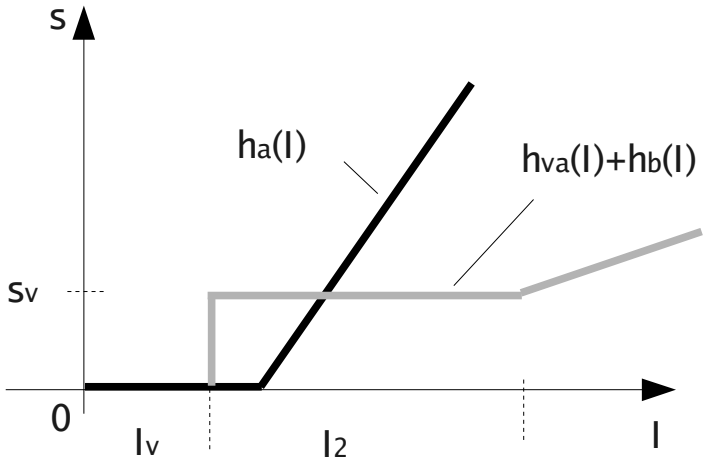

Abbildung 41: Integration der Verlagerung in die Schadenfunktion 2

In der obigen Abbildung ist gut ersichtlich, daß durch den - in Folge der Verlagerung von Prozesssegment A auf Infrastrukturdienst 2 - notwendigen Transfer der Schadenfunktion des Prozesssegments B auf den ersten Infrastrukturdienst nach Ablauf von I_2 - der Intensitätsschwelle von B - wieder ein Steigen der Schadenhöhen mit einer steigenden Intensität einhergeht. In diesem Fall ist die Entscheidung für den Transfer im Falle höherer Ereignisdauern trotzdem wirtschaftlich sinnvoll, da diese Schadenfunktion dort immer noch unter der des Prozesssegments A liegt und auch eine geringere Steigung aufweist.[116]

5.3 Details zum Zählprozess und zur Ereignisdauer der Infrastrukturdomäne

Dieses Kapitel soll im Anschluss an die Erweiterung des Infrastrukturmodells um jene Ereignisse, die die Integrität und die Vertraulichkeit von Informationen gefährden zeigen, wie exakt aus den Fehlerbäumen, die ja ein Resultat der Systemaufbauanalyse sind, die Ereignisraten und die Verteilungen für die Ereignisintensität der jeweiligen Infrastrukturdienste abgeleitet werden können.

5.3.1 Erweiterungen zur Infrastrukturmodellierung

Da in Kapitel 4.4.5 und den Folgekapiteln primär jene Ereignisse berücksichtigt wurden, die die Verfügbarkeit der Infrastruktur beeinflussen, wird in den beiden folgenden kurzen Kapiteln näher darauf eingegangen, wie Ereignisse, die die Integritäts- bzw. Vertraulichkeitsziele beeinträchtigen, in die Infrastrukturanalyse integriert werden können.

5.3.1.1 Integrität in der Infrastrukturebene

Auf grund der in Kapitel 4.4.2 dargestellten Infrastrukturebenen besteht bezüglich der Integrität des Systems als Ganzes der gleiche Zusammenhang zwischen den unteren und den übergelagerten Ebenen wie für die Verfügbarkeit. Dementsprechend ist auch der Fehlerbaum für die Integrität ein ähnlicher, wenngleich es oftmals nicht notwendig ist in die gleiche Detailtiefe einzudringen, wie dies für die korrekte Modellierung der Verfügbarkeit notwendig

116 Ein sinnvolles Kriterium ob ein Transfer durchgeführt werden soll oder nicht könnte hier z.B. eine Gegenüberstellung der Aufwände für den Transfer und der Schadenpotentiale ohne Transfer auf einem bestimmten Quantil der Ereignisdauer sein. Übersteigen die durchschnittlich zu erwartenden Kosten für den Transfer das durchschnittlich zu erwartende Schadenpotential so wird ein Transfer keine sinnvolle Alternative darstellen.

ist. Das nun folgende Beispiel repräsentiert den Integritätsfehlerbaum für das e-mail Beispiel.

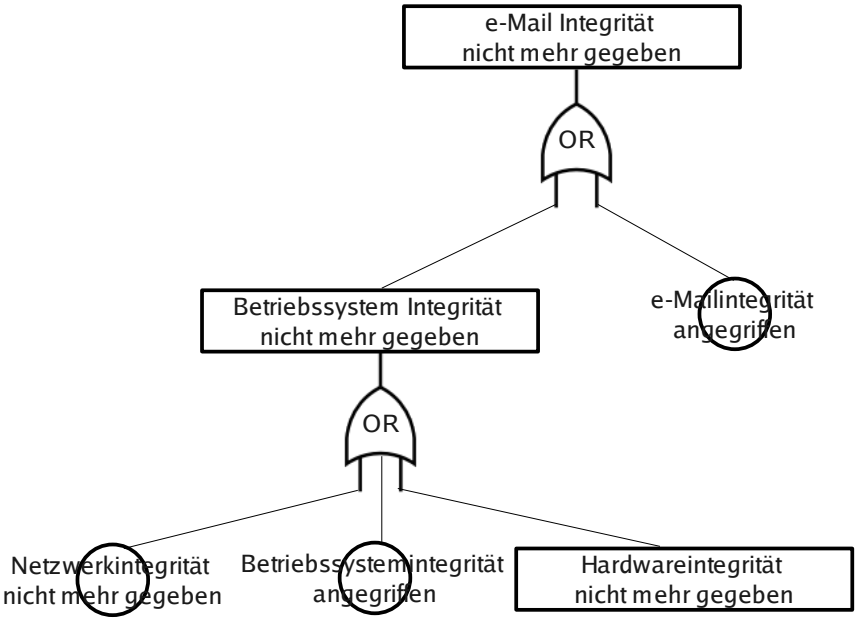

Abbildung 42: Möglicher Integritätsfehlerbaum für einen Mailserver

Im Gegensatz zu den Fehlerbäumen, die die Verfügbarkeitsabhängigkeiten zwischen den jeweiligen Komponenten abbilden, kommen bei der Analyse der Integrität jene Ereignisse, die die Integrität der Komponente direkt angreifen (siehe dazu Kapitel 5.3.6) hinzu. Hinsichtlich der Bedrohung der Integrität sind also auch die höheren Systemebenen Angriffspunkte für Ereignisse.

5.3.1.2 Vertraulichkeit in der Infrastrukturebene

Folgt man der in Kapitel 4.5.1 dargelegten Argumentation, so gibt es nur zwei mögliche Ursachen für den Verlust der Vertraulichkeit von Daten: sie wird entweder direkt angegriffen, oder die Integrität des Dienstes, der die jeweiligen Datenobjekte verwaltet, ist nicht mehr gegeben. Daher sind Vertraulichkeitsfehlerbäume schon fast vollständig durch die Integritätsfehlerbäume der jeweiligen Dienste definiert.

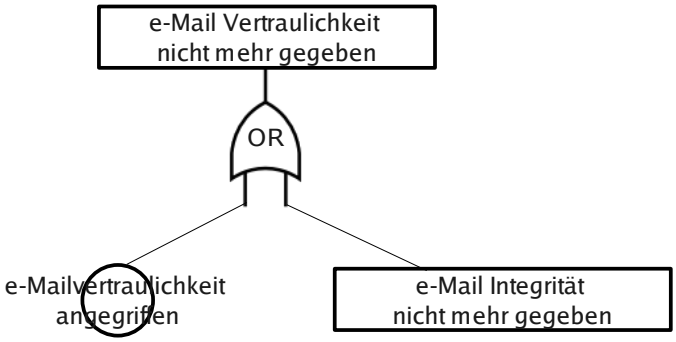

Abbildung 43: Möglicher Fehlerbaum für Vertraulichkeitsprobleme eines Mailservers

Dieses Beispiel enthält der Übersichtlichkeit halber nur noch den obersten Wurzelknoten des Integritätsfehlerbaumes.

5.3.2 Ereignisrate

Die Abschätzung bzw. Herleitung der Ereignisrate auf einem Infrastrukturdienst kann über die Aggregation der Ereignisraten der jeweiligen exogenen und endogenen Basisereignisse auf Grundlage der Systemaufbauanalysen der Infrastrukturebene erfolgen. Als Mittel zur Aggregation kann dabei entweder der Fehlerbaum oder der Ereignisbaum gewählt werden. Im Folgenden soll diese Aggregation anhand des Fehlerbaumes gezeigt werden.

Die Aggregation im Fehlerbaum ist an sich schon aus dem Aufbau des Baumes ersichtlich, es müssen hierzu nur iterativ die jeweiligen *und* bzw. *oder* Verknüpfungen zwischen den Ereignissen in einer Ebene des Baumes aufgelöst werden. *Oder* Verknüpfungen werden trivialer Weise aktiv, wenn eines ihrer Inputereignisse eintritt, und *und* Verknüpfungen können nur aktiv werden, wenn alle Inputereignisse gleichzeitig eintreten.
Im untenstehenden Beispiel wird anhand einer Erweiterung des Mailserverbeispiels aus Kap. 4.4.3 die den Verknüpfungen innewohnende Logik anhand der klassischen Aggregation von Eintrittswahrscheinlichkeiten für die jeweiligen (als statistisch unabhängig angenommenen) Ereignisse dargestellt.

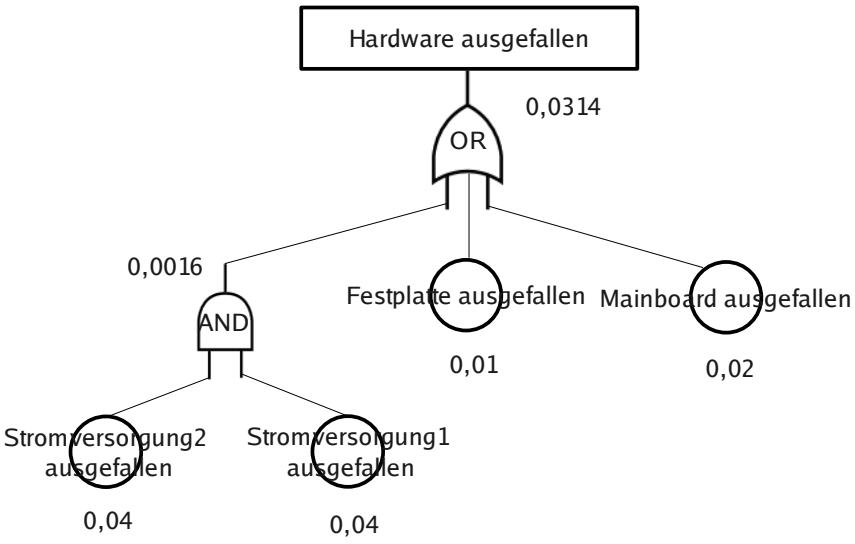

Abbildung 44: Kalkulation des Fehlerbaumes für eine Serverhardware

Die Eintrittswahrscheinlichkeit für das Ereignis *Hardware ausgefallen* errechnet sich also aus (1-(0,04*0,04)*(1-0,01)*(1-0,02))= 0,0314

Die Aggregation zur Ereignisrate des Wurzelereignisses selbst ist auf grund der Komplexität einer formalen mathematischen Aggregation der jeweiligen Ereignisraten dabei am einfachsten[117] über eine Simulation der Fehlerbäume auf Basis der Ereignisraten der jeweiligen Basisereignisse zu erreichen.

5.3.3 Verteilung der Intensität

Wie auch in Kap. 4.2 bereits angesprochen, können Zustandsänderungen ein und der(s) selben Komponente bzw. Prozesssegments eine unterschiedliche Intensität haben. Dieses Konzept geht weiter davon aus, dass ähnliche Ereignisse, die die selbe Komponente betreffen gruppiert und als Sammelereignis betrachtet werden können. Der Ausfall des Infrastrukturdienstes e-mail wird also dahingehend modelliert, dass dieser Ausfall ein Sammelereignis darstellt,

117 Vor allem, weil es dann auch gut möglich ist unterschiedliche Verteilungen für Basisereignisse mit stark unterschiedlichem Charakter zu integrieren.

dessen Ereignisintensität eine Zufallsvariable ist.[118] Durch die Zugrundelegung einer Wahrscheinlichkeitsverteilung für die Ereignisintensität ist es nun besser als durch die Schätzung bzw. Erhebung der durchschnittlichen Eintrittswahrscheinlichkeit für eine Fülle von Ereignissen möglich eine Risikoanalyse und Abschätzung mit Konfidenzniveaus für die Entscheidungsträger durchzuführen.

Bei der Auswahl einer geeigneten Verteilungsfunktion sowohl für die Intensität der Basisereignisse als auch für die Intensität der daraus resultierenden Ereignisse auf den Infrastrukturdiensten gilt es folgende zwei grundlegende Eigenschaften der Ereignisintensität im Zusammenspiel mit der Schadenfunktion und der benötigten Schadenhöhenverteilung zu berücksichtigen:

Setzt man voraus, dass jede Schadenfunktion eine Intensitätsschwelle $I_x \geq 0$ besitzt dann:

- führen alle Ereignisintensitätsverteilungen, deren Dichtefunktionen keine Punktmasse auf dem Wert Null, aber sehr wohl Masse im Intervall [0; I_x] besitzen zu Schadenverteilungen, die auch sog. „Nullschäden“ enthalten können[119]. Damit ist es möglich Verteilungen für die Ereignisintensität zu wählen, die günstige formale bzw. mathematische Eigenschaften besitzen[120], ohne dass etwaige Eigenschaften von in der Praxis vorkommenden Schadenverteilungen dadurch beeinflusst werden.
- führen nur jene Ereignisse, deren Intensität größer I_x ist zu Schäden. Damit ist in diesem Zusammenhang nur die Eintrittswahrscheinlichkeit $P(I \geq i)$ für eine zufallsverteilte Ereignisintensität I von Interesse. Die Verteilungsfunktion von F(I) für die Ereignisintensität hat daher $R^{+¿}$ als Träger, da die Ereignisintensität nicht negativ sein kann, und sie hat den Wert Eins im Nullpunkt des Trägers.

Die Familie aller Verteilungen, die die obigen Eigenschaften haben, sind also die Basis für die Anwendung dieses Modells.

Hinsichtlich der Aggregation der Ereignisintensitätsverteilungen der Basisereignisse ist es auch hier wieder am einfachsten die Ereignisintensitätsverteilung des Wurzelereignisses durch die Simulation des Fehlerbaumes zu bestimmen.

5.3.4 Anknüpfungspunkte zur Verfügbarkeitstheorie

Im Falle, dass die Ereignisdauer die einzige Dimension der Ereignisintensität ist, sind die oben ausgeführten Konzepte im Zusammenhang mit der Ereignisrate und der Dauer nicht grundlegend neu, und der Leser hat vielleicht schon Ähnlichkeiten mit den Konzepten der Verfügbarkeitstheorie festgestellt, die in der Tat wie folgend kurz ausgeführt auch vorhanden sind.

Eine wichtige Eigenschaft der Ereignisrate ist jene, dass in der Verfügbarkeitstheorie die erwartete Anzahl von Ereignissen pro Periode gleich der MTBF (Mean time between Failure) dividiert durch die Periodendauer ist. Dies eröffnet die Möglichkeit die Ereignisrate für eine Vielzahl von Komponenten bzw. Basisereignissen aus oftmals bereits bekannten Kennzahlen von Verfügbarkeitsmodellen abzuleiten.[121]

118 Entgegengesetzt dazu könnten ja auch die Ausfälle mit Intensität x und jene mit Intensität y als unterschiedliche Ereignisse betrachtet werden.

119 Siehe dazu auch die Kapitel 5.2.3 und 5.2.4

120 Verteilungen mit Punktmassen auf dem Wert Null können vor allem bei deren Aggregation aus formaler Sicht sehr unangenehm sein, da durch diese Eigenschaft oft eine analytische Lösung für die aggregierte Verteilung verhindert wird.

121 Siehe auch elementares Erneuerungstheorem

Jedoch ist eine tiefere Untersuchung über die genauen Charakteristika möglicher Zählprozesse und deren Ereignisraten nicht Teil dieser Arbeit und der interessierte Leser sei daher auf die weiterführende Literatur (u.a. [Bauer, 2002] Kap. VIII) verwiesen.

Beschäftigt man sich mit den Kennzahlen der Verfügbarkeitstheorie noch etwas näher, so kann man auch erkennen, daß der Erwartungswert der Ereignisdauerverteilung nicht anderes als die MTTR (Mean Time to Repair) ist.

5.3.5 Risk Response in der Infrastrukturdomäne

Risk Responses in der Infrastrukturdomäne haben vier grundlegende Ansatzpunkte zur Reduktion des Betriebsrisikos eines Unternehmens:

- Reduktion der Intensität des Zählprozesses auf einer Komponente bzw. am Infrastrukturdienst.
- Senkung der u.Ust. vorhandenen Variabilität der Intensität des Zählprozesses.
- Senkung des Erwartungswertes der Ereignisintensitätsverteilung.
- Reduktion der Varianz der Ereignisintensitätsverteilung.

In einigen Fällen haben Maßnahmen die in der Praxis angewandt werden mehr als einen der vier obig genannten positiven Effekte.
Neben dem Umstand, dass Risk Responses auf beide Eingangsgrößen des Risikomaßes (Intensität des Zählprozesses und Ereignisintensität) wirken können, ist auch eine Grenzziehung zwischen diesen beiden Größen, also die Definition, ab welcher Ereignisintensität ein Ereignis auch als solches zu werten ist, sehr wichtig.[122] Diese Definition ist insofern wichtig, da darauf basierend auch die Risk Responses kategorisiert werden können.
Im folgenden soll nun untersucht werden, wie die Risk Responses in das in Kap. 4.4.1 beschriebene Systemaufbaumodell eingefügt werden können, und wie sich ihr Einfluss auf die Ereignisintensitätsverteilung und die Ereignisrate beschreiben lässt.

5.3.5.1 Strukturelle / statische Sicherungsmaßnahmen

Das Systemaufbaumodell und im Speziellen der Fehlerbaum zeigen jene Ursachen und Zusammenhänge auf, die für den Eintritt eines Ereignisses auf einem Infrastrukturdienst verantwortlich sind. Das bedeutet, dass der Fehlerbaum eine Herleitung für die Ereignisrate des Zählprozesses einer Ereigniskategorie liefert. Dadurch ist aber bedingt, dass nur jene Sicherungsmaßnahmen sinnvoll im Fehlerbaum dargestellt werden können, die auf die Intensität des Zählprozesses wirken.[123] Jene Kontrollaktivitäten, die auf die Ereignisintensität wirken[124], können also hier grundsätzlich nicht explizit dargestellt werden und müssen bei der jeweiligen Komponente berücksichtigt werden.
Bei den strukturellen oder statischen Sicherungsmaßnahmen gibt es hinsichtlich deren Implementierung noch zwei zu unterscheidende Fälle, nämlich Absicherungen ohne Umschalteelemente und Absicherungen mit Umschalteelementen.

122 Die Definition der nötigen Ereignisintensität (in diesem Falle ist dies die Ereignisdauer), die ein Ereignis haben muß um als solches gewertet zu werden, kann als Festlegung der Empfindlichkeitschwelle der jeweiligen Messmethodik gesehen werden, die zur Kalibrierung des Modells notwendig ist und die auch erst die Vergleichbarkeit von Risikobewertungen ermöglicht.

123 Dies sind alle jene Maßnahmen, die dauerhaft aktiviert statisch in das System eingebaut sind oder als Teil eines Risk Responses eingebaut werden.

124 i.A. sind dies die ablauforganisatorischen Sicherungmaßnahmen, also die Sicherungsprozesse.

Absicherung ohne Umschaltelemente

Als Sicherungsmaßnahmen ohne Umschalter werden jene Maßnahmen betrachtet, deren risikoreduzierende Eigenschaften ohne zusätzliche Steuerungs- oder Koordinationsmaßnahmen wirksam werden. In Erweiterung des Beispiels aus Kap. 5.3.2 ist mit dem Einsatz eines RAID 1 Verbundes, in dem nach dem Ausfall der primären Festplatte die Ersatzplatte sofort die komplette Funktionalität übernehmen kann, eine solche Sicherungsmaßnahme nun explizit dargestellt. Diese Sicherungsmaßnahme kann dabei wie folgt im Fehlerbaum abgebildet werden:

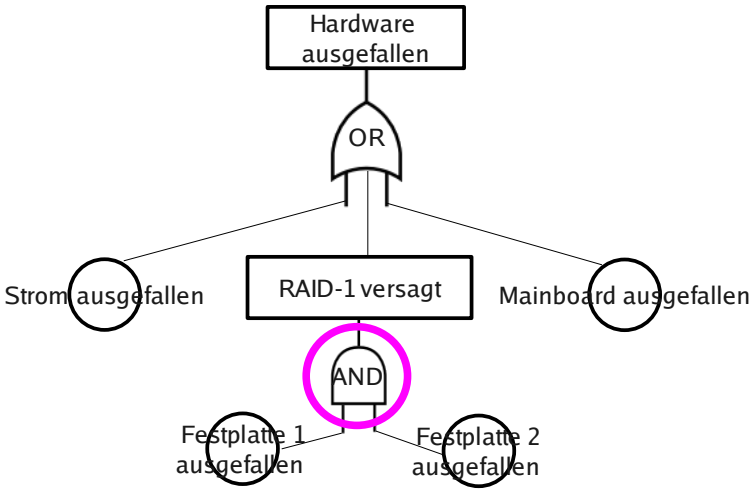

Abbildung 45: Einfache statische Sicherungsmaßnahmen in der Infrastrukturdomäne

Wie aus obiger Abbildung ersichtlich ist, entsteht durch die Sicherungsmaßnahme eine logische *und* Verknüpfung zwischen den jeweiligen Basisereignissen.
Bezüglich der Aggregation der Ereignisrate und der Ereignisdauer auf der nun entstandenen Komponente gelten die selben Regeln, die schon im vorangegangen Unterkapitel angeführt wurden.

Absicherung mit Umschaltelementen

Unter einer Absicherung mit Umschaltelementen sind all jene statischen Sicherungsmaßnahmen zu verstehen, bei denen Steuerungs- bzw. Koordinationsmaßnahmen in Form eines „Umschaltelementes“ notwendig sind um die risikoreduzierenden Eigenschaften entfalten zu können. Eine solche Maßnahme wäre etwa die Erstellung eines Clusters für einen Mailserver. Der Fehlerbaum, der sich durch die Integration einer solchen Maßnahme ergeben kann, ist in der folgenden Abbildung dargestellt.

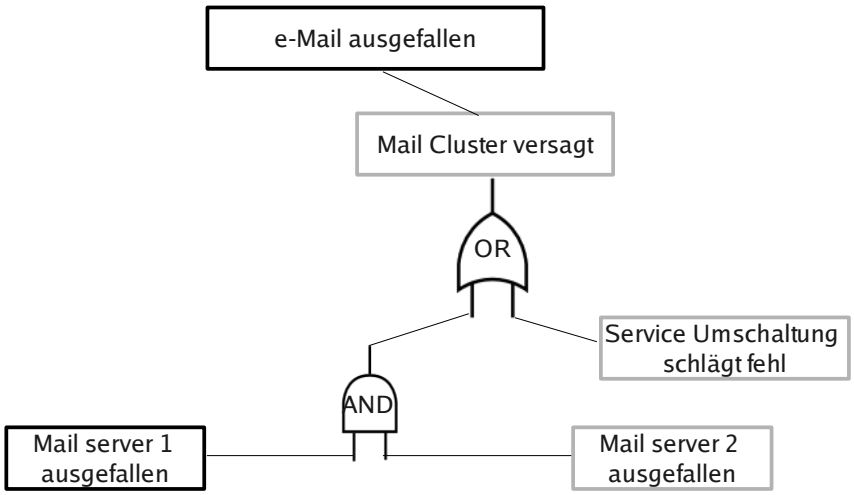

Abbildung 46: Komplexere strukturelle Sicherungsmaßnahmen in der Infrastrukturdomäne

Der Unterschied der Sicherungsmaßnahmen mit Umschaltern zu jenen ohne ist aus obiger Abbildung schon gut ersichtlich. Der Umschalter ist selbst eine Komponente, die ausfallen kann, und trägt somit zur Ereignisrate des Zwischenereignisses „Mail Cluster versagt“ bei. Dadurch verringert sich die positive Wirkung dieser Sicherungsmaßnahme auf die Ereignisanzahl in Abhängigkeit von der Zuverlässigkeit der Umschaltelemente.

Da Sicherungsmaßnahmen, die zusätzliche koordinative Tätigkeiten erfordern, die Komplexität der jeweiligen Komponente bzw. des jeweiligen Dienstes erhöhen, stellt eine im Vergleich zur Einzelkomponente gleichbleibende Reparaturzeit den Idealfall dar. Das bedeutet aber, dass Sicherungsmaßnahmen dieser Kategorie fast sicher schlechtere[125] Ereignisintensitätsverteilungen als die jeweiligen Einzelkomponenten haben.

125 Im Sinne von höherer erwarteter Ereignisdauer und auch höherer Varianz.

5.3.5.2 Kontrollaktivitäten

Als dynamische Kontrollaktivitäten oder Kontrollprozesse werden all jene manuellen, teilautomatisierten und vollautomatisierten ablauforganisatorischen Risk Responses verstanden, die entweder die Ereignisrate eines Ereignisses, oder die zu erwartende Ereignisintensität reduzieren.[126]

Ereignisratenwirksame Kontrollprozesse

Kontrollaktivitäten, die die Ereignisrate beeinflussen können, existieren, wenn man nur die exakte Aufgabe einer Kontrollaktivität - nämlich den Soll - Ist Vergleich des Zustands einer Komponente und das Aufzeigen von Abweichungen vom Sollzustand - betrachtet, nicht. Dies ist auch einleuchtend, da eine Zustandsänderung / ein Ereignis ja zuerst eintreten muß um gemessen werden zu können.
Allerdings können Kontrollaktivitäten bei Vorhandensein zumindest einer der folgenden beiden Eigenschaften dennoch Einfluss auf die Ereignisrate nehmen.

- Die Existenz der Kontrollaktivität hat im Zusammenspiel mit den normalen Geschäftsprozessen und den darin beteiligten Menschen abschreckende Wirkung (wie etwa durch eine Förderung des Unrechtsbewusstseins, oder die Bewußtmachung von Sanktionen bei Fehlverhalten) auf die Verursacher von Ereignissen (hauptsächlich Menschen), sodass weniger Basisereignisse stattfinden.
- Die Kontrollaktivität erkennt und beseitigt die Zustandsabweichung in so kurzer Zeit, dass deren Lebensdauer kürzer als die auf Seite 91 angeführte Zeit, bis ein Ereignis als solches zu werten ist, bleibt. Ein Beispiel für eine Kontrollaktivität, die solche Eigenschaften haben kann ist der bereits angeführte Cluster oder das Beispiel aus Abbildung 46.[127]

Jene prozessbasierenden Kontrollaktivitäten, die nun doch auf die Ereignisrate wirken, haben ähnliche Eigenschaften wie die statischen Sicherungsmaßnahmen ohne Umschalter und können daher auch äquvalent zu diesen in den Fehlerbaum integriert werden. Veranschaulicht soll dies anhand der folgenden Erweiterung des e-mail Beispiels werden.

126 Viele der dieser Kategorie zurechenbaren Kontrollaktivitäten sind prinzipiell in der Prozessebene beheimatet, werden aber - weil diese direkt auf die Infrastrukturkomponenten wirken - zur Infrastrukturebene gerechnet.

127 Im Allgemeinen können Kontrollaktivitäten, deren Latenzzeit nach dem Eintreten eines Ereignisses sehr viel kleiner als die Pufferzeit des von diesem Ereignis potentiell betroffenen Prozesssegments ist, die hier angeführten Eigenschaften zugesprochen werden.

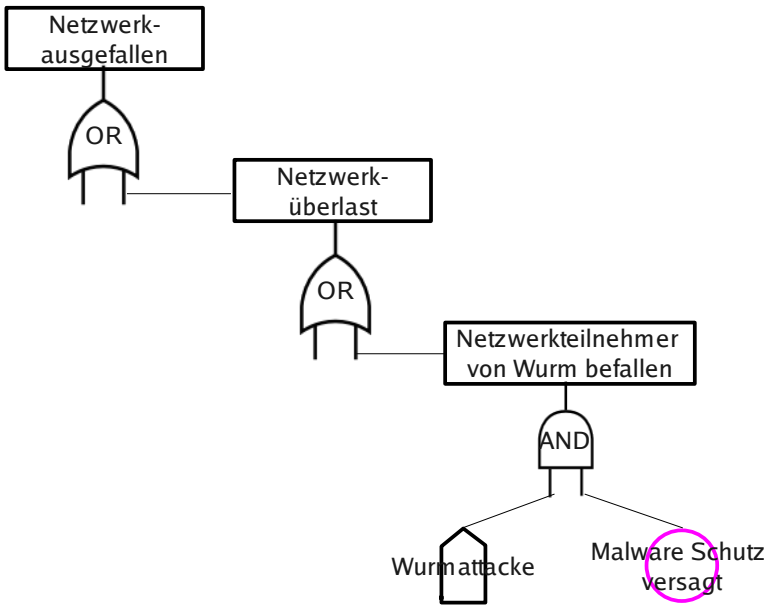

Abbildung 47: Einfaches Beispiel für die Integration einer vollautomatisierten Kontrollaktivität in die Infrastrukturdomäne

Im obigen Beispiel wurde nun das ehemalige Basisereignis „Netzwerk ausgefallen“ weiter analysiert und es wurde festgestellt, dass ein wurmverseuchter PC im Netzwerk einen Wurmangriff verursachen kann, welcher auf Grund der Verbreitungsversuche des Wurms zu einer Netzwerküberlast führen kann. Nun ist die Tatsache, dass ein Wurm versucht sich auf einem PC einzunisten ein exogenes Ereignis, dass vom Unternehmen nicht beeinflusst werden kann. Es besteht aber die Möglichkeit, zu versuchen den Wurmbefall durch einen entsprechenden Malwareschutz zu verhindern. Ist die Kontrollaktivität Malwareschutz aktiv, so kann sich der Wurm nur einnisten, wenn die Kontrollaktivität versagt[128] (daher die *und* Verknüpfung).
Die Kontrollaktivität Malwareschutz ist dabei ein gutes Beispiel für eine vollautomatisierte ablauforganisatorische Maßnahme, die bei jedem Zugriff auf ein Datenobjekt durchgeführt wird und prüft, ob sich dieses Objekt bzw. das gesamte System innerhalb der vordefinierten Grenzen für unbedenkliches Verhalten befindet.

Ereignisintensitätswirksame Kontrollaktivitäten

Ereignisintensitätswirksame Kontrollaktivitäten im Allgemeinen und im Speziellen solche deren Latenzzeit zu groß ist um sie als auf die Ereignisrate wirkend klassifizieren zu können, erzeugen durch ihre Aktivierung prinzipiell neue, d.h limitierte Ereignisintensitäten auf den Komponenten, auf denen sie angewandt werden. Das bedeutet, es ist für eine exakte mathematische Beschreibung des durch die Kontrollaktivität modifizierten Systems prinzipiell notwendig diese neue Ereignisintensität zu bestimmen.
Für die praktische Anwendung dieses Modells ist es daher meist notwendig, eine Anpassung der ursprünglichen Ereignisintensität bzw. deren Verteilungsfunktion mit Hilfe der durch die

128 Im hier vorgestellten Modell wird davon ausgegangen, dass die Sicherungsprozesse selbst auch fehlerhaft sind und auch ausfallen bzw. versagen können. Dies und die Tatsache, dass Sicherungsprozesse ständig parallel zu den Komponenten ablaufen, die sie kontrollieren, bedingt die *und* Verknüpfung der jeweiligen Ereignisse im Fehlerbaum. Im Gegensatz dazu bestünde auch die Möglichkeit von perfekten Prozessen auszugehen, die selbst nicht versagen und Ereignisse nur auf grund des Umstandes, dass sie ihnen diese nicht bekannt sind, nicht erkennen und abschwächen können. Ein solcher Ansatz entspräche einer „cold standby“ Konfiguration.

Anwendung der Kontrollaktivität neu zu bestimmenden statistischen Daten durchzuführen.[129]

5.3.5.3 Aufwände für Kontrollaktivitäten

Einfach ausgedrückt lässt sich festhalten, dass Kontrollaktivitäten praktisch immer Aufwände verursachen, wobei dies auf zweierlei Art geschehen kann:

- in Form von Fixkosten, welche durch dauerhaft aktivierte Kontrollaktivitäten entstehen.
- in Form von variablen Kosten, welche durch Kontrollaktivitäten verursacht werden, die im Anlassfall aktiviert. Zu dieser Kategorie gehören zum Beispiel die Reparaturmaßnahmen, die zur Wiederherstellung des Sollzustandes einer Komponente bzw. eines Prozesses dienen.

Dauerhaft aktivierte bzw. installierte Kontrollaktivitäten sind meist in der Infrastrukturdomäne vorzufinden, und ihre Kosten sind, da sie oft über lange Zeiträume gleich bleiben, recht einfach zu bestimmen.
Im Gegensatz dazu sind die Aufwände für Kontrollaktivitäten, die nur im Anlassfall benötigt bzw. aktiviert werden, schwer abschätzbar, da deren Aktivierung in einem Prozess oft zu Einschränkungen in einem anderen Prozess führt (siehe Kap. 5.2.5.3).

Es stellt also vor allem das Aktivieren einer anlassfallbezogenen Kontrollaktivität selbst ein Ereignis dar, das in anderen Infrastrukturkomponenten bzw. Prozessen als Basisereignis im Fehlerbaum aufscheinen kann.
Dementsprechend ist es sinnvoll, und aus wirtschaftlichen Überlegungen auch notwendig, vor der Installation bzw. vor der Aktivierung einer anlassfallbezogenen Kontrollaktivität deren Schadenpotential zu erheben, um nicht durch die Aktivierung mehr Aufwand als Nutzen zu generieren.

5.3.6 Einstiegspunkte in die Infrastrukturdomäne

Am Ende der Ausführungen über die Modelldetails in der Infrastrukturdomäne soll auf den folgenden Seiten noch kurz auf die aus Sicht des Modells wichtigen Typen von exogenen Ereignissen und deren mögliche Interaktionen mit den Komponenten der Infrastrukturdomäne eingegangen werden. Ziel dieser Darstellung ist es die Ursachen für die Basisereignisse des Fehlerbaumes aufzuzeigen, um deren Messung bzw. Ableitung aus in der Praxis bereits vorhandenen Kennzahlen zu ermöglichen.

Um ein Computersystem also zu infiltrieren, gibt es nur drei grundlegende Möglichkeiten:

- Das Ausnutzen einer Schwäche des Computersystems[130] (Vulnerability)
- Das Ausnutzen einer Schwäche im Verhalten der Benutzer des Computersystems.[131]
- Direkte Veränderung des Systems mit Hilfe eines physikalischen Systemzuganges.

129 In einem sehr oberflächlichen Ansatz könnte man hier z.B. einen Faktor bestimmen, um den die Kontrollaktivität die erwartete Ereignisintensität reduziert, oder es könnte ein Maximalwert für die Ereignisintensität definiert werden.

130 genauer: Die von externen Personen ausnutzbaren Schwächen befinden sich überwiegend in Diensten, die die ISO OSI Ebenen 5-7 implementieren. Die Zahl der erfolgreichen Angriffe ist dabei durchwegs nur vom Systemaufbau, also dem internen Aufbau der Dienste und deren exakter Konfiguration abhängig. Unter diesen Punkt fällt zum Beispiel auch die technische Absicherung von Passwörtern.

131 die Erfolgsrate bei diesem Einstiegspunkt ist durchwegs von der Sensibilität des angegriffenen Users gegenüber den Informationssicherheits Best Practices abhängig. In diese Kategorie fällt z.B. auch die Qualität von Passwörtern.

Vulnerabilities in Computersystemen sind immer nur nutzbar, wenn das angreifende System direkt mit einem Dienst auf dem angegriffenen System kommunizieren kann. Dies ist der Grund, warum oftmals nur öffentlich zugängliche Serversysteme bzw. Clientsysteme ohne Firewall über diesen Weg infiltriert werden können.
Um ungünstiges Benutzerverhalten mittels sog. Social Engeneering auszunutzen ist es aus Sicht eines Angreifers nur notwendig, mit dem jeweiligen Benutzer in Kontakt zu kommen, wofür es bereits ausreichend ist dem Benutzer ein vielversprechendes Dokument oder Angebot (Stichwort Phishing) zukommen zu lassen. Dies ist auch mit ein Grund, warum der überwiegende Teil der erfolgreichen Angriffe am Benutzer ansetzt und nicht an den Systemen selbst.
Ein physikalischer Systemzugang bietet zwar im Allgemeinen die beste Möglichkeit, unentdeckt am System die gewünschten Änderungen durchzuführen, er ist aber auch meist in Relation zu den anderen beiden Methoden am besten abgesichert.

5.3.6.1 Malware

Malware ist der Sammelbegriff für jedwede Schadsoftware, die teil- oder vollautomatisiert Datenverarbeitungssysteme attackiert um Zugriff auf diese zu erlangen. Die jeweilige Ausprägung von Malware kann dabei hinsichtlich ihrer Verbreitungsmethode und ihres Schadteils unterschieden werden.

Verbreitungsmethoden

Grundsätzlich gibt es zwei unterschiedliche Verbreitungsmethoden für Malware:

- Verbreitung als Virus: ein Virus verbreitet sich, indem er sich selbst in Dateien und Programme kopiert. Der Virus ist in den meisten Fällen für eine Verbreitung über die Grenzen des aktuell befallenen Computersystems hinaus auf die Distribution der infizierten Dateien durch den Benutzer angewiesen.
  Anfällig für eine Infektion, und somit Angriffspunkt für eine Infektion von außen sind primär die Schnittstellen zur Prozessdomäne (bzw. zum Benutzer), also meist Clientsysteme, wobei als Einstiegspunkt nur ungünstiges Benutzerverhalten ausgenutzt werden kann.
  Dieser Verbreitungsmechanismus beschädigt immer die Integrität der infizierten Dateien, Programme und des Betriebssystems, auf dem eine infizierte Datei geöffnet bzw. gestartet wird. In vielen Fällen werden auch wichtige Teile der infizierten Dateien überschrieben oder modifiziert, wodurch auch deren Verfügbarkeit beeinträchtigt ist.
- Verbreitung als Wurm: ein Wurm ist in der Lage sich vollautomatisiert über Computernetze zu verbreiten. Das bedeutet, ein Wurm sucht von sich aus von einem befallenen System aus nach neuen Systemen, die er attackieren kann.
  Angriffspunkt für eine Infektion von außen sind alle Dienste, die die ISO OSI Ebenen 5-7 implementieren, also Dienste, die entweder direkt mit dem Internet oder anderen WAN's kommunizieren bzw. verbunden sind und damit direkt von einem Wurm angesprochen werden können, oder alle Systeme, auf denen Daten die über diese Netzwerkanbindung übertragen wurden, verarbeitet[132] werden. Das bedeutet, dass Würmer grundsätzlich in der Lage sind, auch die Schnittstelle zur Prozessdomäne als Angriffspunkt zu nutzen, wenn der Benutzer sich entsprechend ungünstig verhält.
  Dieser Verbreitungsmechanismus beschädigt immer die Integrität der befallenen

132 genauer: interpretiert

Betriebssysteme. Bedingt durch die vollautomatische Verbreitung führt er oft zu Verfügbarkeitsproblemen bei den attackierten Systemen bzw. im Netzwerk, das das angreifende System mit den Zielsystemen verbindet.

Schadteile

Bezeichnungen für die Schadteile von Malware gibt es unterschiedlichste, wie etwa Trojaner, Spyware oder Backdoors. Im Grunde gibt es aber nur die folgenden beiden Typen von Schadteilen:

- Aufhebung der Vertraulichkeit von Daten: Dieser Schadteiltypus zielt darauf ab Informationen weiterzuleiten, die vertraulich sind.
- Reduktion der Verfügbarkeit: Malware mit diesem Ziel versucht Daten bzw. ganze Computersysteme unbrauchbar zu machen.

Zielkonflikte

Das folgende Zustandsdiagramm eines IT-Systems soll aufzeigen, welche Informationssicherheitsziele auf den unterschiedlichen IT-Systemebenen von Malware bedroht werden:

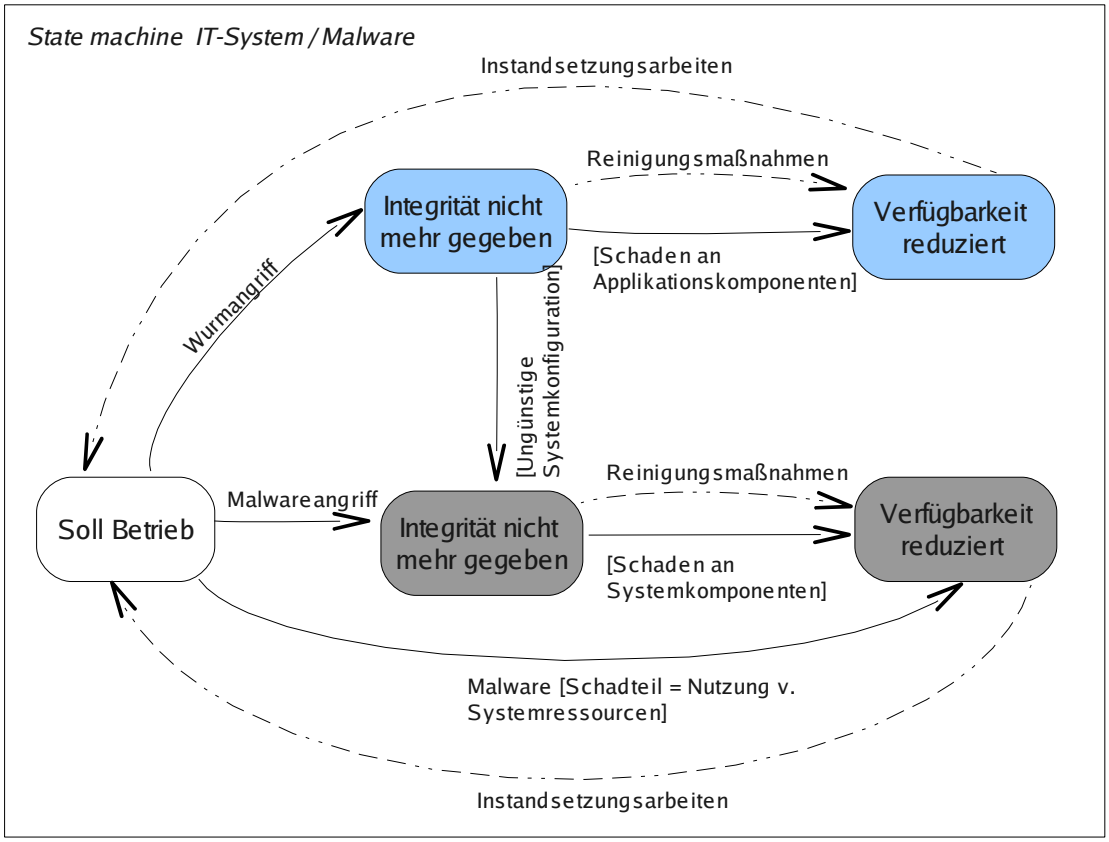

Abbildung 48: UML Zustandsdiagramm eines IT-Systems im Kontext eines Malwareangriffes

Es gilt hier zu bemerken, dass vor allem der Integritätsverlust des Betriebssystems (Türkis) zu massiven Zielbedrohungen bei den am System gespeicherten Daten führt (dies ist auch insoferne logisch, da ein Zugriff auf die Daten nur über das OS abgewickelt werden kann).

5.3.6.2 Hackangriff

Hacker bedienen sich prinzipiell der gleichen Zugangswege und -methoden wie Malware, in vielen Fällen wird auch Malware genutzt um Zugang zum Zielsystem zu bekommen. Die Angriffspunkte befinden dabei nahezu immer auf den Diensten die die ISO OSI Ebenen 5-7 implementieren, bzw. auf der Schnittstelle zur Prozessdomäne[133]. Auch der Schadteil ist prinzipiell gleich. In diesem Zusammenhang ist es sinnvoll zwischen automatisierten Hackangriffen, die in erster Linie Malware einsetzen um ein System zu infiltrieren, und manuellen zu unterscheiden. Die manuellen führen dabei in den seltensten Fällen zu Verfügbarkeitsproblemen in den Netzwerken, weil der Angreifer u.a. auch unentdeckt bleiben möchte. Hackangriffe unterscheiden sich hinsichtlich der Bedrohungen für die Informationssicherheitsziele und der betroffenen Infrastrukturkomponenten nicht von Malwareangriffen, und da automatisierte Hackangriffe in der überwiegenden Zahl der Fälle über Malware abgewickelt werden, wird im folgenden unter einem Hackangriff ein manuell ausgeführter Angriff verstanden.

Zielkonflikte

Auch für diesen Typus exogener Ereignisse soll das untenstehende Zustandsdiagramm wieder eine Auflistung jener Komponenten bieten, deren Informationssicherheitsziele durch diese Ereignisse bedroht sind.

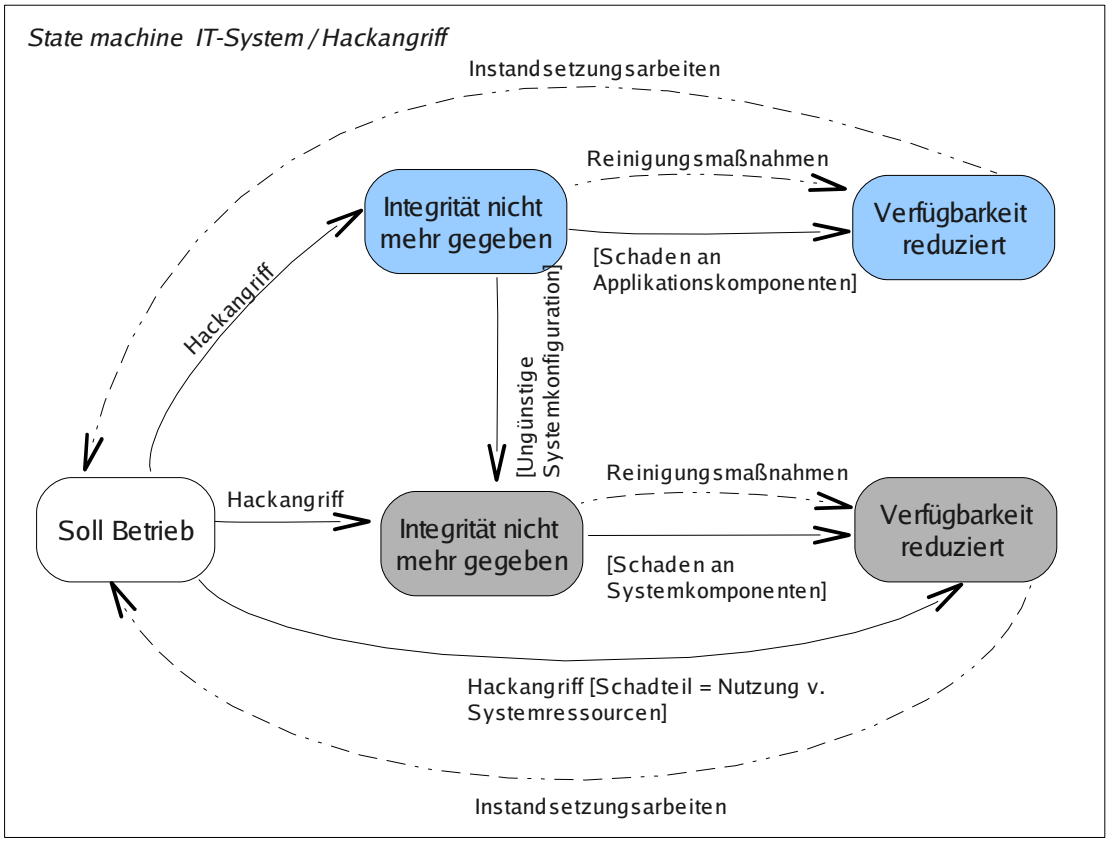

Abbildung 49: UML Zustandsdiagramm eines IT-Systems im Kontext eines Hackangriffes

Auch hier führt vor allem der Integritätsverlust des Betriebssystems zu massiven Zielbedrohungen bei den am System gespeicherten Daten.

133 Diese Schnittstelle wird dann genutzt wenn der reguläre Benutzer ungenügende Authentifizierungsmerkmale (Passwörter) benutzt.

5.3.6.3 Unautorisierte Datennutzung

Unautorisierte Datennutzung ist ein Ereignis, dass von Benutzern verursacht wird, die zwar Zugriffsrechte auf die jeweiligen Daten besitzen, diese Zugriffsrechte aber außerhalb des Rahmens eines gerade zu bearbeitenden Geschäftsfalls nutzen[134]. Als Angriffspunkt kommt dabei nur die Schnittstelle der Anwendungsschicht im ISO OSI Modell zur Prozessdomäne in Frage. Das bedeutet, aus Sicht des Unternehmens sind dies interne Vorfälle, die aber aus Sicht des Modells exogene stochastische Ereignisse sind, die auf grund ihrer Definition nur die Vertraulichkeit von Datenobjekten beeinträchtigen.

5.3.6.4 DOS Angriffe

Denial of Service Angriffe sind entweder ein Nebenprodukt von Malware oder Hackangriffen, oder aber der gezielte Versuch die Verfügbarkeit einzelner Komponenten zu beeinträchtigen. Angriffspunkte sind dabei sämtliche Dienste, die die ISO OSI Protokollebenen 3-7 implementieren und vom Angreifer aus erreichbar sind. Gezielte DOS Angriffe können dabei von extern nur auf jene Komponenten ausgeführt werden, die über WAN Netze erreichbar sind.

Zielkonflikte

DOS Attacken haben nur Zielkonflikte mit Verfügbarkeitszielen, und sie interagieren wie folgt mit den jeweiligen IT-Systemebenen:

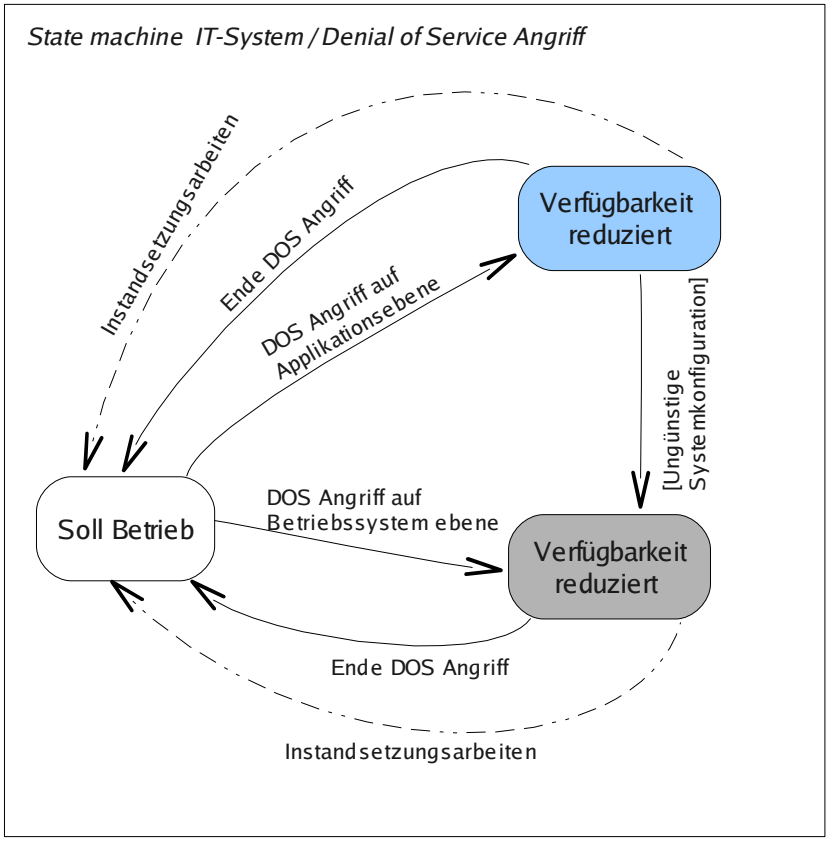

Abbildung 50: UML Zustandsdiagramm eines IT-Systems im Kontext eines DOS-Angriffes

134 also für den jeweiligen Zugriff nicht autorisiert sind bzw. waren.

6 Beispiel zur Modellumsetzung

In diesem Kapitel soll nun gezeigt werden, wie das in den Kapiteln 3 bis 5 beschriebene Modell in der Praxis angewendet werden kann. Der Einfachheit halber werden im Rahmen dieses Beispiels nur Ausfallereignisse betrachtet und deren jeweilige Dauer als Intensitätsmaß verwendet.
Die Grundlage für die folgenen Ausführungen bildet wieder das Infrastrukturmodell eines Mailservers, das schon in Kapitel 4.4.3 vorgestellt wurde. In der untenstehenden Abbildung sind zusätzlich nun auch jene Komponenten, die als Angriffspunkte für exogene Ereignisse gewählt wurden, Grau hinterlegt.

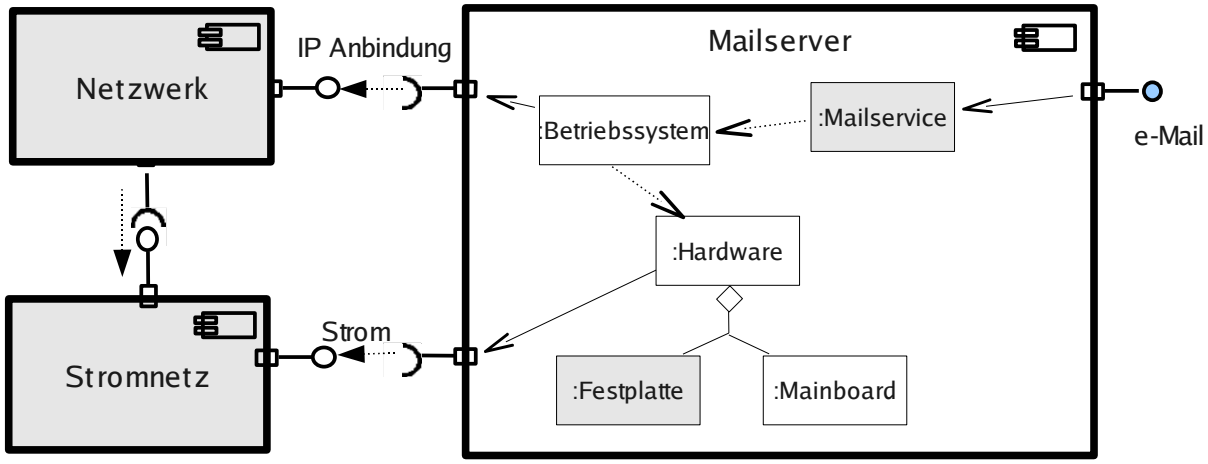

Abbildung 51: Beispielhaftes Infrastrukturmodell eines Mailservers

Wie ersichtlich ist, können im Rahmen des Beispiels das Stromnetz, das Netzwerk, die Festplatte des Servers und der Dienst jeweils unabhängig voneinander in Folge von exogenen Ereignissen ausfallen. Der aus obigem Modell ableitbare Fehlerbaum stellt sich wie auf der nächsten Seite gezeigt dar.

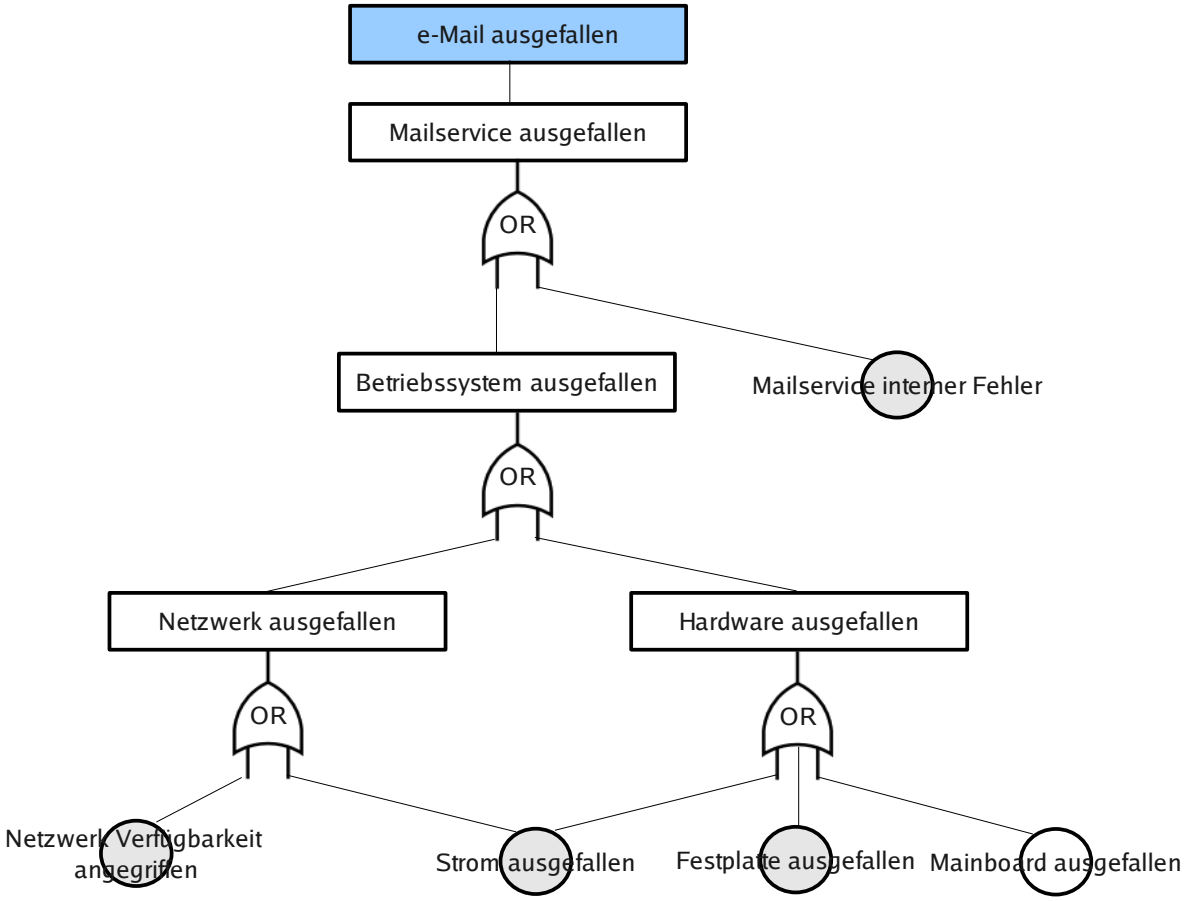

Abbildung 52: Beispielhafter Fehlerbaum für einen Mailserver

Hinsichtlich der zu erwartenden Ereignisraten der Zählprozesse und der Verteilungen der Ereignisdauern jedes der oben angeführten Basisereignisse wurden für das Beispiel die folgenden Annahmen getroffen[135]:

- Mailservice interner Fehler: Es wurde angenommen, dass dieser Fehler auf grund von Stabilitätsproblemen der eingesetzten Software verursacht wird. Als Zählprozess wurde ein Poissonprozess mit einem lambda von drei Ereignissen pro Monat angenommen. Die Ereignisdauer wurde als gammaverteilt mit Erwartungswert 20 Minuten und einer Varianz von 10 Minuten angenommen.
- Stromausfall: Auch hier wurde wieder ein Poissonprozess angenommen, der ein lambda von vier Ereignissen pro Jahr besitzt. Die Ereignissdauer ist gammaverteilt mit einem Erwartungswert von 6 Stunden und einer Varianz von 2 Stunden.
- Netzwerk Verfügbarkeit angegriffen: Dieses Ereignis soll eine Denial of Service Attacke repräsentieren, welche auch einem Poissonprozess mit einem lamda von zwei Ereignissen pro Jahr unterliegt und eine gammaverteilte Ereignisdauer mit einem Erwartungswert von 12 Stunden und einer Varianz von vier Stunden besitzt.
- Festplatte ausgefallen: Für dieses Ereignis wurde angenommen, dass der Poissonprozess durch ein lambda von einem Ereignis pro Monat definiert ist und die Ereignisdauerverteilung wieder eine Gammaverteilung ist, mit einem Erwartungswert von einem Tag und einer Varianz von einem Tag.

135 Hinweis: die hier angeführten Annahmen wurden zur besseren Veranschaulichung der Modellumsetzung gewählt. Die jeweiligen Ereignisse kommen in der Praxis nur selten in dieser Häufigkeit vor.

6.1 Das inhärente Risiko

Zur Berechnung der Schäden X jedes Ereignisses wurde die folgende Schadenfunktion zu Grunde gelegt: $x = h(d) = d^2$ [136]wobei die Ereignisdauer d auf Stunden normiert und der Schaden X in Euro gemessen wurde.

Basierend auf den obigen Spezifikationen wurden 5000 Simulationsdurchläufe mit einer Simulationsperiode von jeweils einem Jahr durchgeführt, um das inhärente Risiko für die vom Infrastrukturdienst e-mail abhängigen Prozesse zu bestimmen. Die Erhebung der Zustände der jeweiligen Komponenten und die darauf folgende Auswertung des Zustandes des gesamten Dienstes erfolgte in der Simulation nach jeder simulierten Minute.

Das inhärente Risko stellt sich nun wie folgt dar:

- Erwartete Anzahl von Ausfällen des Infrastrukturdiensts pro Jahr: E(N)= 53,4 / Varianz der Ereigniszahl: V(N)= 51,9.
- Erwarteter Schaden eines einzelnen Ereignisses: E(X)= 258,2 Euro / Varianz des Schadens eines einzelnen Ereignisses: V(X)~ 1,6*10^6 Euro.
- Erwarteter kumulierter Schaden einer Periode:[137] E(S)= 13797,8 Euro / Varianz des kumulierten Schadens pro Periode: V(S)~ 8,3*10^7 Euro.
- 90% Quantil des kumulierten Schadens einer Periode: 25883 Euro.

136 Hinweis: die Schadenfunktion ist rein willkürlich gewählt, und hat u.a. auch den Zweck die Wirkung der Risk Responses hervorzuheben. Der Einfachheit halber wurde im Rahmen des Beispiels auch auf eine detaillierte Analyse einer möglichen Prozessdomäne verzichtet, und statt dessen diese primitive Schadenfunktion als - für die Prozessdomäne - repräsentativ angenommen.

137 Nach dem Kollektiven Modell in der Schadenversicherungsmathematik ist E(S)=E(N)*E(X). Der erwartete kumulierte Schaden einer Periode ist also gleich dem Produkt aus erwartetem Einzelschaden, und erwarteter Schadenanzahl.

6.2 Der Risk Response

Betrachtet man die Annahmen für die Basisereignisse nocheinmal genauer, so lässt sich feststellen, dass es zwei primäre Treiber für das inhärente Risiko des Dienstes e-mail gibt. Es sind dies der monatliche Ausfall der Festplatte und der quartalsweise Ausfall des Stromnetzes. Aus diesem Grund wurden für diese beiden Komponenten als Risk Response jeweils statische Sicherungsmaßnahmen eingefügt. Die Festplatte wurde dabei durch einen RAID 1 Verbund abgelöst, und die Stromversorgung durch eine USV Anlage abgesichert. Das Infrastrukturmodell wurde dadurch wie folgt modifiziert.

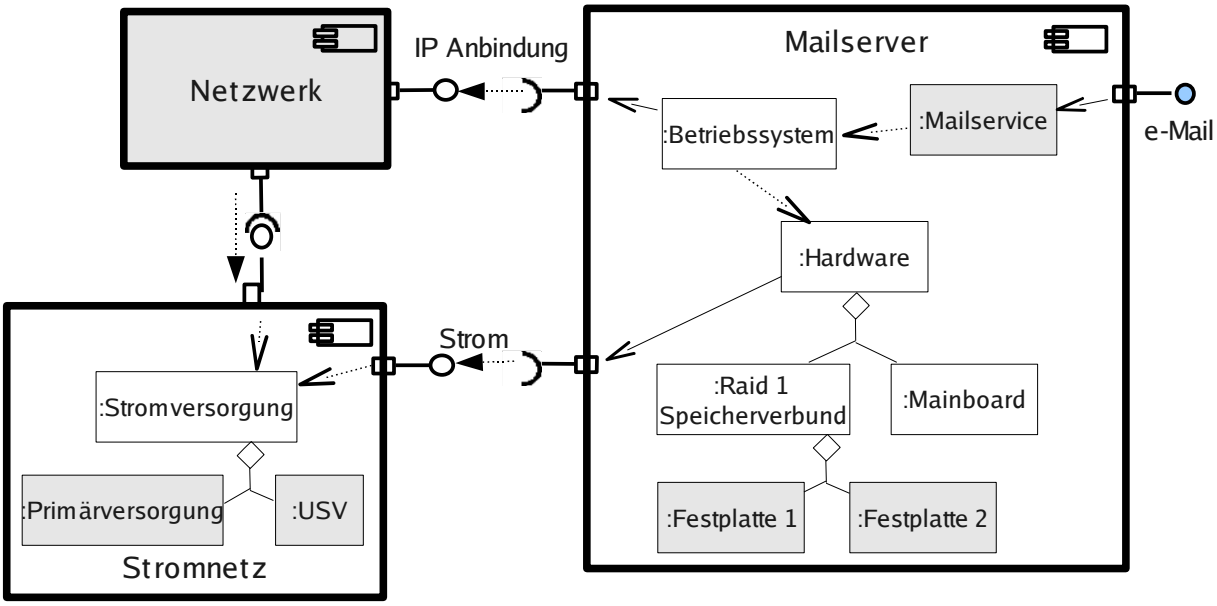

Abbildung 53: Infrastrukturmodell für den e-mail Dienst mit eingebauten Risk Responses

Der aus dem obigen Infrastrukturmodell ableitbare Fehlerbaum sieht wie folgt aus.

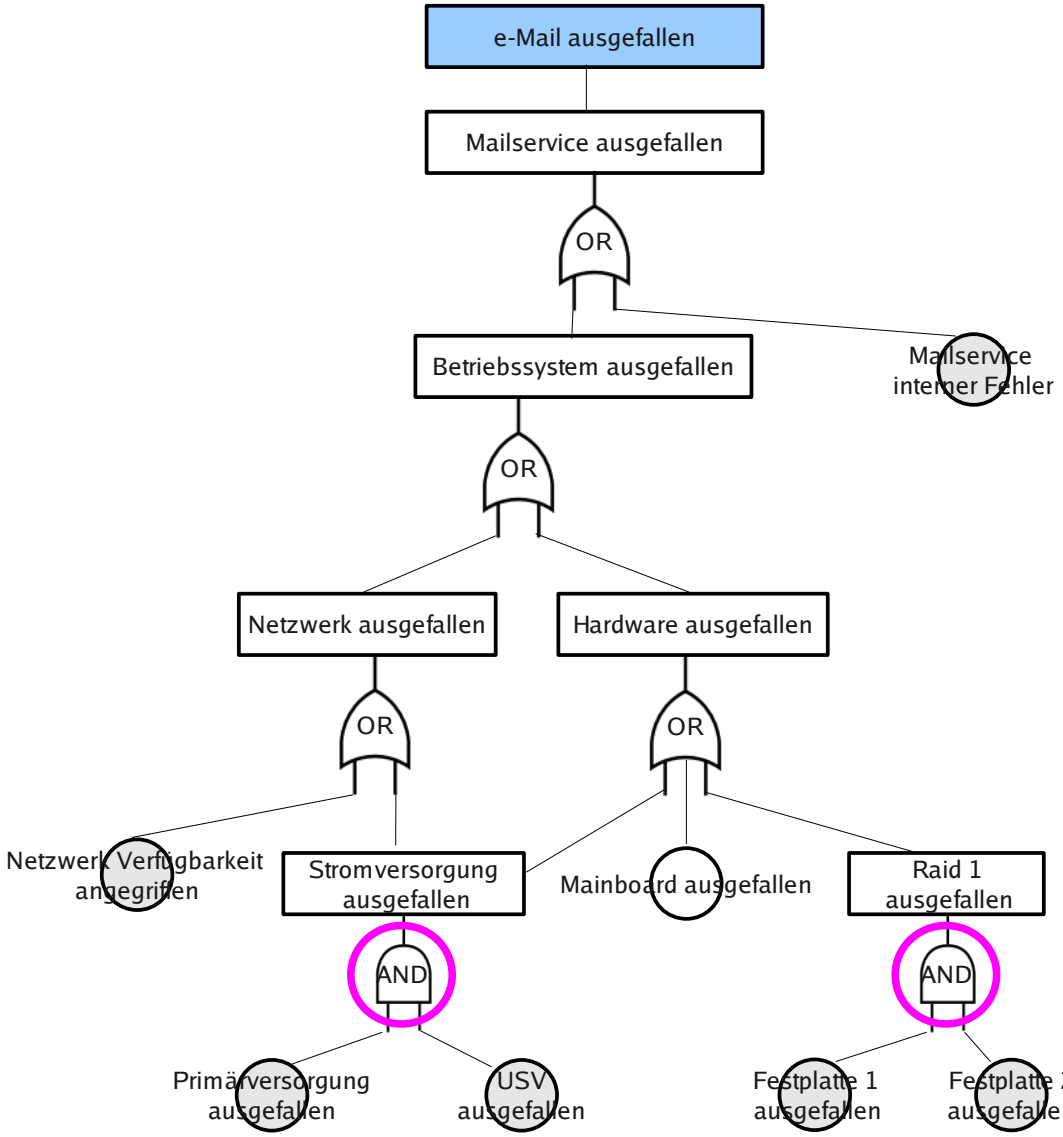

Abbildung 54: Fehlerbaum für einen e-mail Dienst inklusive der Risk Responses

Bezüglich der Basisereignisse dieses Fehlerbaumes wurden die selben Annahmen getroffen wie für den inhärenten Fall. Dies gilt auch für die neue Komponente USV, der die selben Zahlen wie für die Primärversorgung zu Grunde gelegt wurden.

6.3 Das residuale Risiko

Zur Berechnung des Risikos nach der Anwendung des Risk Responses, also des residualen Risikos, wurden wieder 5000 Jahre simuliert und die Schäden X ebenfalls wieder über die Schadenfunktion $x=h(d)=d^2$ aus der jeweiligen Ereignisdauer errechnet. Das residuale Risiko, der vom Dienstes e-mail abhängigen Prozesse, stellt sich dabei nun wie folgt dar.

- Erwartete Anzahl von Ereignissen pro Jahr: E(N)= 39,75 / Varianz der Ereignisanzahl: V(N)= 41,1.
- Erwarteter Schaden eines einzelnen Ereignisses: E(X)= 12,83 Euro / Varianz des Schadens eines einzelnen Ereignisses: V(X)~ 9400 Euro.
- Erwarteter kumulierter Schaden einer Periode: E(S)= 514 Euro / Varianz des kumulierten Schadens pro Periode: V(S)~ 3,8*10^5 Euro.

- 90% Quantil des kumulierten Schadens einer Periode: 1032 Geldeinheiten.

Wenn man diese Zahlen mit jenen für das inhärente Risiko vergleicht, kann man daraus schon sehr gut die Effektivität der eingesetzten Sicherungsmaßnahmen ableiten. Viel anschaulicher fällt noch die folgende Gegenüberstellung der beiden Fälle aus.

6.4 Inhärentes vs. residuales Risiko

Hier folgt nun noch eine graphische Aufarbeitung der beiden Risiken, die sehr gut veranschaulicht warum die Risiken so unterschiedlich groß sind. Den Anfang machen dabei die Verteilungen für die Anzahl der Ereignisse am Infrastrukturdienst innerhalb eines Jahres.

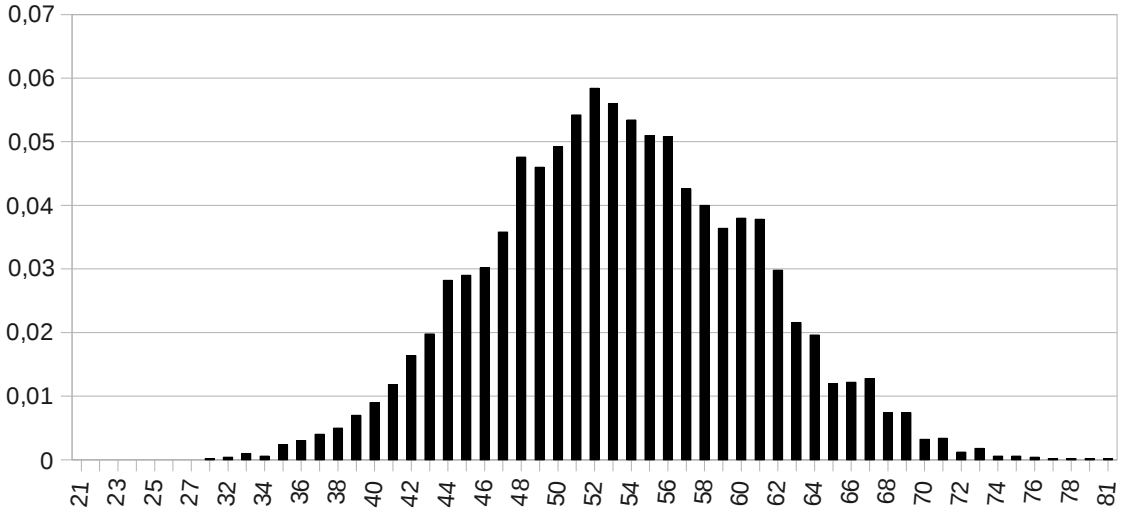

Abbildung 55: Dichte der Anzahl von Ereignissen pro Jahr im inhärenten Fall

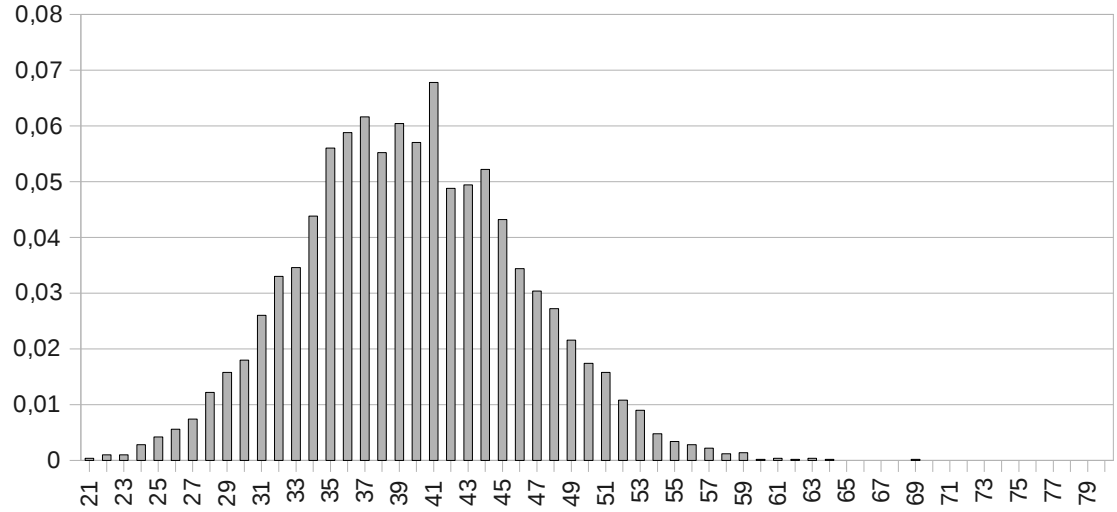

Abbildung 56: Dichte der Anzahl von Ereignissen pro Jahr im residualen Fall

In beiden obigen Diagrammen wird auf der Ordinate die Anzahl von Ereignissen in einer Periode dargestellt und auf der Abszisse deren relative Häufigkeit. Wie sich gut aus beiden Abbildungen ersehen lässt, verschiebt sich die ganze Verteilung für die Ereignisanzahl einer Periode durch die Anwendung des Risk Responses.

Das folgende Diagramm ist eine Gegenüberstellung der jeweiligen Dauer eines einzelnen Ereignisses für beide Fälle.

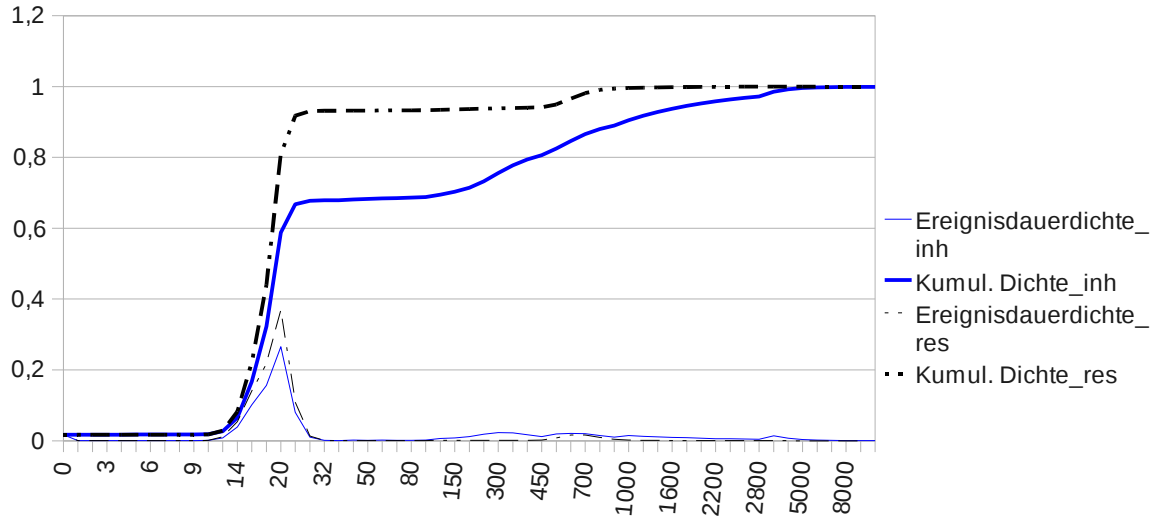

Abbildung 57: Gegenüberstellung der Ereignisdauern bei inhärentem und residualem Risiko

Die dünnen Linien repräsentieren dabei jeweils die Dichtefunktion und die dicken Linien die dazugehörige Verteilungsfunktion für die Dauer eines einzelnen Ereignisses am Infrastrukturdienst. wobei auf der Ordinate die Dauer eines Ereignisses in Minuten aufgetragen ist. Der inhärente Fall wird dabei blau mit durchgehenden Linien, und der residuale schwarz mit unterbrochenen Linien dargestellt.
Gut ersichtlich ist, dass im inhärenten Fall ein beträchtlicher Teil der Masse der Verteilungsfunktion bei sehr hohen Ausfalldauern zu finden ist, was auch letztlich den erheblichen Unterschied zwischen den beiden Risiken ausmacht. Explizit dargestellt ist dies in der folgenden Abbildung.

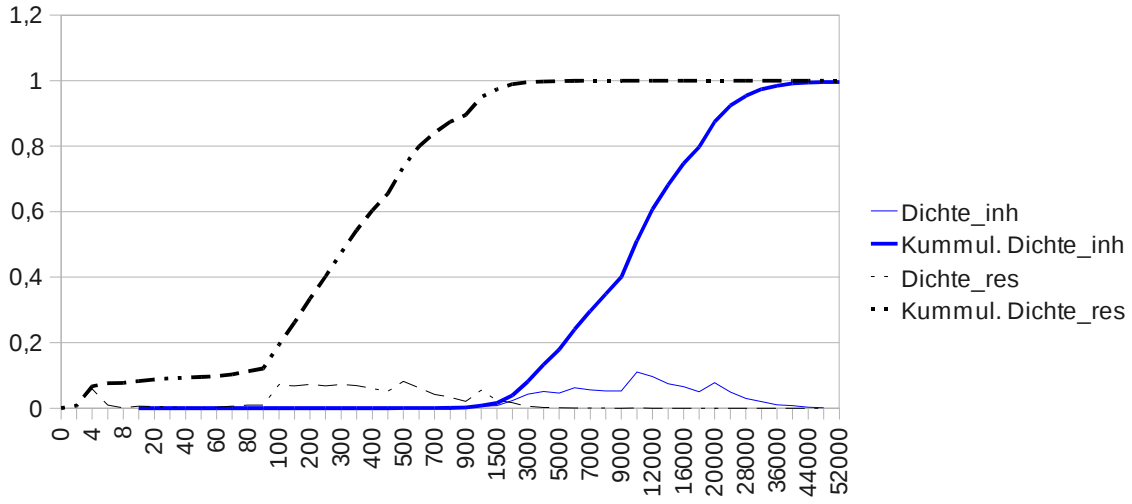

Abbildung 58: Dichte und Verteilungsfunktion der kumulierten Schadensummen eines Jahres

In Abbildung 58 ist auf der Ordinate die jeweilige Schadensumme eines Jahres abgebildet und auf der Abszisse die jeweilige Wahrscheinlichkeit. Was gut zu erkennen ist, ist die diskriminierende Wirkung der Schadenfunktion auf die jeweils realisierten Schäden. So gibt es im inhärenten Fall zumindest einen Großschaden pro Jahr, der bewirkt dass die Wahrscheinlichkeit für ein Jahr mit einem Gesamtschaden kleiner oder gleich 1500 praktisch null ist. Im residualen Fall hingegen gibt es praktisch keine Großschäden, mehr da auch fast keine Ereignisse mit mehr als 10 Stunden Dauer existieren, und dadurch ist die Wahrscheinlichkeit für einen Gesamtschaden größer als 1500 praktisch null.[138]

138 Hinweis diese derartig großen Unterschiede vor allem bei den Gesamtschäden eines Jahres sind hauptsächlich durch die quadratische Schadenfunktion bedingt. Würde hier eine Schadenfunktion gewählt die etwa auf Grund der Lieferverträge limitiert wäre (was ja durchaus oft der Realität entspricht) so wären diese Unterschiede weit geringer aber immer noch markant.

7 Zusammenfassung und Ausblick

Ziel dieser Arbeit war die Erstellung eines Modellgerüsts, innerhalb dessen eine ganzheitliche Modellierung aller aus Sicht der Informationssicherheit relevanten Vorgänge und Ereignisse einer Organisation möglich ist. Dabei war es wichtig Anknüpfungspunkte für die wichtigsten aktuell angewandten Informationssicherheits- und Risikomanagementstandards zu schaffen, ohne das Modell durch deren Grenzen zu beschränken. Aus dieser Vorgabe entwickelte sich auch eine sehr wichtige Eigenschaft des nun vorliegenden Modells: es ermöglicht es, die im jeweiligen Implementierungsfall angewandten Standards und Best Practices in Beziehung zu einander zu setzen. Dadurch wird es auch möglich, noch bestehende Schwächen im jeweiligen Risiko- bzw. Informationssicherheitsmanagementsystem (ISMS) strukturiert zu identifizieren und systematisch zu beseitigen.

Ausgangspunkt für die Etablierung und Anwendung des in den Kapiteln drei bis fünf beschriebenen Modells ist die Betrachtung eines Unternehmens bzw. einer Organisation als soziotechnisches System mit einem mehrschichtigen hierarchischen Aufbau. Das bedeutet, es müssen für eine Analyse sowohl die technischen als auch die organisatorischen Strukturen und Abläufe berücksichtigt werden.
Das Rüstzeug für die Analyse und Beschreibung eines derartigen Systems wurde durch die Arbeit von Mesarovic et. al. [Mesarovic et. al., 1970] (siehe Kap. 2.2) definiert. Zur Darstellung der für diese Arbeit wichtigen Aspekte organisatorischer Strukturen und Abläufe kommt der Standard ANSI ISA 95 zur Anwendung (siehe Kap. 2.3 und 4.1), der es ermöglicht sowohl die kleinsten Elemente eines Prozesses (Prozesssegmente) und deren Umwelt, als auch das hierarchische und funktionale Zusammenspiel dieser Prozesssegmente zu beschreiben.
Zu Beginn von Kapitel 4 wurde das dieser Arbeit zugrundeliegende Metamodell vorgestellt. In diesem wird davon ausgegangen, dass die internen Infrastrukturereignisse mit Hilfe eines Modells zur Infrastrukturinteraktion aus den exogenen Ereignissen abgeleitet werden können. Als eine weitere Annahme wird davon ausgegangen, dass diese internen Infrastrukturereignisse durch ein entsprechendes Infrastrukturmodell zu Ereignissen auf den Infrastrukturdiensten (-services) aggregiert werden können. In weiterer Folge führen diese als Serviceereignisse bezeichneten Ereignisse auf den Infrastrukturdiensten durch die Prozessinteraktion zu Prozesssegmentereignissen. Diese Prozesssegmentereignisse selbst können nun externe Ereignisse zur Folge haben, wobei deren Zusammenhang über das Prozessdomänenmodell hergestellt wird. Jene Ereignisse der Prozessdomäne, die keine externen Ereignisse zur Folge haben, besitzen dabei nur interne Schadenpotentiale wie etwa Instandsetzungskosten (siehe Kapitel 5.2.4.5), für jene Ereignisse, die auch externe Ereignisse zur Folge haben, müssen zusätzlich auch die externen Schadenpotentiale berücksichtigt werden (siehe Kapitel 5.2.4.2 und 5.2.4.3). Die Höhe der jeweiligen Schäden wird dabei aus der Intensität des jeweiligen Ereignisses bzw. des Ereignistyps abgeleitet.

Die in dieser Arbeit vorgestellte Detailausprägung des Metamodells der Prozessdomäne berücksichtigt bedingt durch die Anwendung des Konzeptes des Meilensteingraphen als Ereignisintensität ausschließlich die Dimension Zeit bzw. die Ausfalldauer[139] und die damit verbundenen Einschränkungen (constraints) bezüglich der Fertigstellungszeitpunkte und der Abarbeitungsdauer für jedes Prozesssegment (Prozessschritt). Darauf basierend kann man so zusammen mit den hierarchischen und funktionalen Abhängigkeiten Ablaufgraphen und die

139 Hinweis: Zur Herleitung der Auswirkungen eines Ereignisses können sich in der Praxis verschiedenste Ereigniseigenschaften anbieten, die Aufschluss über die Intensität des Ereignisses geben. Bei der detailierten Darstellung des Modells fokussiert sich diese Arbeit aber der Einfachheit und Nachvollziehbarkeit halber nur auf die Ereignisdauer als Intensitätsmaß.

darin enthaltenen kritischen Pfade ableiten (siehe Kap. 4.1.3 und 5.2). Vor allem auf Basis der kritischen Pfade ist es möglich einen Schadenverlauf (Schadenfunktion) in Abhängigkeit von der Dauer eines Stillstands oder eines Ausfalls eines der analysierten Prozesssegmente zu bestimmen.
Steigt man nun eine Detailebene tiefer in das Modell der Prozessinteraktion und analysiert die möglichen Ursachen für ein Ereignis in einem Prozesssegment, so kann man diese Information aus der Umwelt der jeweiligen Prozesssegmente erhalten. Im wesentlichen entstehen Probleme für die Prozesssegmente dann, wenn entweder der Input (sei es der zu verarbeitende oder die Steuerinformationen) nicht (in der benötigten Qualität) vorhanden ist, oder die für die Abarbeitung der Aufgaben benötigte Infrastruktur nicht verfügbar ist. Da diese Arbeit, einfach dargestellt, die Erarbeitung eines Modells zur Beschreibung der Auswirkungen von Informationssicherheitsereignissen zum Ziel hat, wurden in weiterer Folge nur jene Teile der Prozesssegmentumwelt betrachtet, die der Informations- und Kommunikationstechnologie zugeordnet werden können. Als Schnittstelle zwischen Prozess- und Infrastrukturdomäne wurden dabei die von der IT-basierenden Infrastruktur angebotenen Dienste / Services gewählt.
An jenem Punkt wurde die in COSO 2 (siehe Kap. 2.1) beschriebene Methodologie angewandt und die Ziele der Informationstechnologie oder genauer der Informationssicherheit, wie sie im ISO 27002 (siehe Kap. 2.4) definiert sind, als Grundlage für die weiteren Analysen herangezogen. Führt man nun bei der Anwendung des Modells die COSO 2 Aktivität *Event Identification* durch, so erhält man durch eine Top-Down Analyse der funktionalen Abhängigkeiten in der IT-Infrastruktur mittels Fehlerbaumanalyse ein Infrastrukturmodell und jene ursächlichen Ereignisse, die für ein Verfehlen der Informationssicherheitsziele in der Ebene der Infrastrukturdienste (-services) verantwortlich sind. Um in weiterer Folge die COSO 2 Aktivität *Risk Assessment* für ein Prozessereignis durchführen zu können ist es sinnvoll die Top-Down Analyse in der Infrastrukturdomäne so weit voran zu treiben, bis jene ursächlichen Ereignisse identifiziert werden konnten, die auch meßbar und damit statistisch erfassbar sind.(siehe auch Kap. 5.3.6)
Für die Durchführung des *Risk Assessment* bei den identifizierten Ereignissen wird im Rahmen dieses Modells auf die Methode des Probabilistic Risk Assessment (PRA / Kap. 2.5) zurück gegriffen. Dabei wird das mit einem Ereignis verbundene Risiko durch dessen Eintrittswahrscheinlichkeit und die Auswirkungen des Ereignisses beschrieben. Bezüglich der Intensität (Dauer) eines Ereignisses wird davon ausgegangen, dass ein im Rahmen der Top-Down Analyse identifiziertes Ereignis eine Gruppe von gleichartigen Ereignissen repräsentiert, die sich nur hinsichtlich deren Intensität unterscheiden. Dadurch ist es möglich die Intensität eines Ereignisses über eine Ereignisintensitätsverteilung der jeweiligen Gruppe zu modellieren. Für die Modellierung der Eintrittswahrscheinlichkeit wird unterstellt, dass die Ereignisse selbst das Resultat eines stochastischen Zählprozesses sind.
In weiterer Folge können die Ereignisraten der Zählprozesse und die Ereignisdauern der ursächlichen (Basis-)Ereignisse mit Hilfe des Fehlerbaumes im Infrastrukturmodell und mit Hilfe des Prozessinteraktionsmodells so weit aggregiert werden, um die Ereignisrate und die Ereignisdauerverteilung der jeweiligen Service- bzw. Prozesssegmentereignisse zu erhalten. Mit Hilfe der internen und externen Schadenfunktionen aus der Prozessdomäne kann nun aus der Verteilungsfunktion für die Ereignisintensität eine Verteilungsfunktion für den Schaden bestimmt werden, womit beide Variablen, die im Rahmen des PRA benötigt werden, vorliegen.
Bezüglich der im Rahmen der COSO 2 Aktivität *Risk Response* definierten Sicherungsmaßnahmen und der im Rahmen der Aktivität *Control Activities* durchzuführenden Kontrollaktivitäten ist festzuhalten, dass deren Wirkung im vorliegenden Modell in der

Prozessebene über deren Effekt auf die Schadenfunktion, und in der Infrastrukturebene über deren Effekt auf die Ereignisrate und die Ereignisdauerverteilung des jeweiligen Ereignisses abgebildet wird. Es gilt hier auch zu beachten, dass die Aktivierung einer Sicherungsmaßnahme selbst auch wieder ein mit einem Risiko oder zumindest mit einem Aufwand behaftetes Ereignis darstellen kann. Zu diesen Sicherungsmaßnahmen werden z.B. auch die Reparaturmaßnahmen, die in Folge eines Ereignisses aktiviert werden müssen, gezählt.

Mit den obig beschriebenen Methoden ist es möglich für ausgewählte Ereignisse in einem Produktions- bzw. Dienstleistungsprozess eine Schadenverteilung mit $R^{+\dot{c}}$ als Träger zu bestimmen.[140]
Basierend auf den Schadenverteilungen der Prozesse ist es auch möglich die Gesamtschadenverteilung der jeweiligen Risikostellen bzw. Risikoträger im Unternehmen (für nähere Informationen zum Konzept der Risikostellen sei der interessierte Leser an Eller et al. [Eller et al., 2002] und Schwaiger [Schwaiger, 2001] verwiesen) und in weiterer Folge auch die Verteilungsfunktionen für das gesamte Unternehmen zu errechnen.

Ziel dieser Risikobewertung auf einer Risikostelle oder für das gesamte Unternehmen jedenfalls ist es eine Aussage zu erhalten, wie hoch der Erwartungsschaden in einer Periode ist bzw. wie hoch der Schaden auf einem ausgewählten Quantil der Schadenverteilung ist, um so feststellen zu können, ob sich dieser innerhalb des Riskoappetits des Mangements bzw. des Verantwortlichen für die Risikostelle befindet. Erst diese Gegenüberstellung kann als Argumentationsbasis für etwaige weitere Sicherungsmaßnahmen und Kontrollaktivitäten dienen.

140 Durch die Eigenschaften der Schadenfunktion kann es dabei vorkommen, dass die Schadenverteilung auch Masse auf dem Punkt 0, also einen diskreten Anteil hat.

7.1 Ausblick

Im Grunde gibt es zwei große Themenblöcke die aus Sicht des Autors im Rahmen dieser Arbeit nicht genug Platz gefunden haben, die aber für eine bessere Anwendung des Modells in der Praxis auch noch große Relevanz haben.
Der erste Themenblock der auch bereits in Kapitel 4.5 angesprochen wurde ist ein detailliertes Modell zur Beschreibung der Infrastrukturinteraktion von exogenen Ereignissen. Im Rahmen dieser Arbeit wurde auf eine exakte Abbildung der Interaktion der exogenen Ereignisse mit der Infrastruktur und der darauf folgenden Verbreitung der Ereignisse auf die Komponenten der Infrastrukturdomäne nicht näher eingegangen. Es wird im Modell aber von der Existenz von Daten über das Auftreten dieser Ereignisse auf den jeweils im Rahmen der Top-Down Analyse identifizierten relevanten Komponenten ausgegangen. Für eine praxisnahe Anwendung dieses Modells liegt der nächste Entwicklungsschritt damit klar auf der Hand, und zwar ist dies die Modellierung der Verbindungen der Komponenten untereinander. Basierend auf den Eigenschaften der exogenen Ereignisse und der Konfiguration der Komponenten mit denen diese direkt interagieren können, sollten so Prognosen über das vordringen dieser Ereignisse zu den im Zuge der Top-Down Analyse als relevant identifizierten Komponenten ermöglicht werden. Das bedeutet es ist notwendig die Vernetzung der Komponenten untereinander in das Modell der Infrastrukturinteraktion zu integrieren.
Vor allem die Arbeiten von Henry [Henry et al., 2009] und Mc. Queen [Mc Queen et al., 2006] beinhalten sehr viel versprechende Ansätze wie diese Vernetzung der Komponenten modelliert werden kann.

Der zweite Themenblock der einen sinnvollen Erweiterungsschritt darstellt ist die Herleitung von Schadenfunktionen für jene externen Ereignisse die nicht von Ereignissen in der Prozessdomäne abstammen, sondern die direkt von Ereignissen auf Infrastrukturdiensten verursacht werden. Meist sind dies Probleme mit der Vertaulichkeit von Informationen die die Qualität der Leistungserbringung eines Unternehmens beeinträchtigen oder sich direkt auf die Nachfragemenge auswirken. Ein Beispiel hierfür wären etwa unbefugte Datenzugriffe, die die Abläufe der Prozessdomäne nicht direkt negativ beeinflussen, die aber einmal publik (d.h. extern meßbar) sehr wohl negativ auf die Nachfrage wirken können.

Anhang A: Literaturverzeichnis

[Bauer, 2002] Bauer H., Wahrscheinlichkeitstheorie, De Gruyter ISBN 3-11-017236-4.

[Bedford, 2001] Bedford T., Cooke R., Probabalistic Risk Analysis, Foundations and Methods Cambridge Universitiy Press ISBN 0-521-77320-2.

[Eller et al., 2002] Eller R., Schwaiger W., Federa R., Bankbezogene Risiko- und Erfolgsrechnung, Schäffer-Poeschl, Stuttgart, ISBN 3-7910-1652-0.

[Fandel, 2005] Fandel G., Produktion I Produktions- und Kostentheorie, Springer ISBN 3-540-25023-9.

[Henry et al., 2009] Henry H., Yacov Y., A Comprehensive Network Security Risk Model for Process Control Networks, Risk Analysis, Vol 29, No. 2, 2009.

[Jaquith, 2007] Jaquith A., Security Metrics, Addison-Wesley ISBN: 978-0-321-34998-9.

[Johannsen et al., 2007] Johannsen W., Goeken M., Referenzmodelle f. IT-Governance, dpunkt.verlag, ISBN 978-3-89864-397-9.

[Landoll, 2006] Landoll D., The Security Risk Assessment Handbook, Auerbach Publications, ISBN: 0-8493-2998-1.

[Mack, 2002] Mack T., Schadenversicherungsmathematik, Verl. Versicherungswirtschaft, Karlsruhe, 2002.

[Mc Queen et al., 2006] McQueen M., Boyer W., Flynn M., Beitel G., Quantitative cyber risk reduction estimation for a small SCADA control system. In Proceedings of the 39'th Hawaii International Conference on System Sciences. IEEE, 2006.

[Mc Queen et al., 2005] McQueen M., Boyer W., Flynn M., Beitel G., Time-to-compromise model for cyber risk reduction estimation. In Proceedings of the 1st Quality Protection Workshop at the University of Trento, Milan. Springer 2005.

[Mesarovic et al., 1970] Mesarovic M., Macko D., Takahara Y., Theory of Hierarchical, Multilevel, Systems, Academic Press, New York, 1970.

[Porter, 1998] Porter M., Competitive Advantage, Free Press, New York [u.a.], 1998.

[Schwaiger, 2001] Schwaiger W., Finanzwirtschaftlich basierte Unternehmenssteuerung, Gabler ISBN 3-8244-6798-4.

[Vesely et al., 1981] Vesely W., Goldberg F., Roberts N., Haasl D., Fault Tree Handbook of the U.S. Nuclear Regulatory Commission.

[BSI 100-3] Bundesamt für Sicherheit in der Informationstechnik, Risikoanalyse auf der Basis von IT-Grundschutz Version 2.5, www.bsi.bund.de.

[CMMI] Capability Maturity Model Integration der Carnegie Mellon University, www.sei.cmu.edu/cmmi/.

[COBIT] Control Objectives for Information and related Technology Version 4.1, www.itgi.org, www.isaca.org.

[COSO2] Enterprise Risk Management – Integrated Framework, by the Commitee of Sponsoring Organizations or the Treadway Commision 2004, www.coso.org.

[FAA SRMP] Federal Avivation Administration, Security Risk Management Process, http://fast.faa.gov/Riskmgmt/Secriskmgmt/secriskprocess.htm.

[ISA 95-1] ANSI ISA 95 Part 1 / ISO IEC 62264-1:2003 Enterprise-control system integration -- Part 1: Models and terminology, www.iso.org.

[ISO 27001] ISO IEC 27001:2005 Information technology — Security techniques — Information security management systems — Requirements , www.iso.org.

[ISO 27002] ISO IEC 27002:2005 Information technology -- Security techniques -- Code of practice for information security management, www.iso.org.

[ISO 27005] ISO IEC 27005:2008 Information technology -- Security techniques -- Information security risk management.

[ISO/OSI] ISO IEC 7498-1:1994 Information Technology - Open Systems Interconnection - Basic Reference Model, www.iso.org.

[ITIL] IT Infrastructure Library Version 3, www.itil.org.

[NSA IAM] National Security Agency, Infosec Assessment Methodology, http://www.iatrp.com/iam.php.

[OCTAVE] Operationally Critical Threat, Asset, and Vulnerability Evaluation, www.cert.org/octave/.

[OSSTMM] Open Source Security Testing Methodology Manual Version 3, www.isecom.org.

[UML] The Unified Modelling Language Version 2, www.uml.org.

Anhang B: Abbildungsverzeichnis

Abbildung 1: ERM Übersicht nach COSO2....7
Abbildung 2: COSO 2 im Unternehmensregelkreis....11
Abbildung 3: Multiebenensysteme nach Mesarovic....13
Abbildung 4: Systemgliederung durch Strata....14
Abbildung 5: Entscheidungssystem nach Mesarovic....15
Abbildung 6: Entscheidungshierarchie / Echelons nach Mesarovic....16
Abbildung 7: Prurdue Refernezmodell für ein Produktionsunternehmen....21
Abbildung 8: Ereignisbaum in einem elektrischen Leitungsnetz....30
Abbildung 9: Fehlerbaum eines elektrischen Leitungsnetzes....31
Abbildung 10: Kritischer Pfad, Schritt 1....33
Abbildung 11: Kritischer Pfad Schritt2....33
Abbildung 12: Kritischer Pfad Schritt 3....34
Abbildung 13: ISO / OSI Schichten....35
Abbildung 14: Ereignisse im Kontext des Systems Unternehmen....38
Abbildung 15: Erweiterte Systemumwelt....39
Abbildung 16: Beispiel für eine einfache Schadenfunktion....42
Abbildung 17: Zusammenhang zwischen Zählprozess, Schadenfunktion h(I) und Schadenprozess....43
Abbildung 18: Rahmenmodell zur Integration von Informationssicherheitsereignissen in ein Risikomodell nach COSO2....44
Abbildung 19: Modellrelevante Elemente der Prozesssegmentumwelt....46
Abbildung 20: Prozesshierarchie basierend auf ISA 95 / Purduemodell....49
Abbildung 21: Beispiel für eine einfache Schadenfunktion....57
Abbildung 22: Einfache Infrastrukturabhängigkeiten....58
Abbildung 23: Gegenüberstellung ISO OSI Modell / IT- Systemebenen....59
Abbildung 24: Infrastrukturabhängigkeiten unter Berücksichtigung ausgewählter Subkomponenten....60
Abbildung 25: Beispiele für die Trennung in datenverarbeitende und datenbereitstellende Dienste....61
Abbildung 26: Beispiel für ein UML Zustandsdiagramm der Komponente Betriebssystem....62
Abbildung 27: Erweiterung um die Zustände von aus OS Sicht benötigten Komponenten....63
Abbildung 28: Möglicher Ausfallsfehlerbaum für einen Mailserver....64
Abbildung 29: Verfügbarkeitsabhängigkeiten zwischen den Ebenen eines IT-Systems....66
Abbildung 30: Integritätsabhängigkeiten zwischen den Ebenen eines IT-Systems....67
Abbildung 31: Ablaufgraph nach Berücksichtigung der ISA 95 Ebenen 3 und 2....71
Abbildung 32: Ablaufgraph nach Berücksichtigung der ISA 95 Ebenen 3,2 und 1....71
Abbildung 33: Ablaufgraph nach Berücksichtigung aller relevanten ISA 95 Ebenen....72
Abbildung 34: Beispiel für einen Ablaufgraph der parallelisierte Prozessschritte berücksichtigt....74
Abbildung 35: Meilensteinfehlerbaum zur Darstellung der Ausfallsabhängigkeiten in der Prozessdomäne…75
Abbildung 36: Beispiel für verwandte Schadenfunktionen innerhalb eines Ablaufgraphs....77
Abbildung 37: Schadenfunktionen und Schadenprozess bei stochastischen Pufferzeiten....83
Abbildung 38: Integration von Risk Responses in die Schadenfunktion 1....84
Abbildung 39: Integration von Risk Responses in die Schadenfunktion 2....85
Abbildung 40: Integration der Verlagerung in die Schadenfunktion....85
Abbildung 41: Integration der Verlagerung in die Schadenfunktion 2....87
Abbildung 42: Möglicher Integritätsfehlerbaum für einen Mailserver....88
Abbildung 43: Möglicher Fehlerbaum für Vertraulichkeitsprobleme eines Mailservers....88
Abbildung 44: Kalkulation des Fehlerbaumes für eine Serverhardware....89
Abbildung 45: Einfache statische Sicherungsmaßnahmen in der Infrastrukturdomäne....92
Abbildung 46: Komplexere strukturelle Sicherungsmaßnahmen in der Infrastrukturdomäne....93
Abbildung 47: Einfaches Beispiel für die Integration einer vollautomatisierten Kontrollaktivität in die Infrastrukturdomäne....95
Abbildung 48: UML Zustandsdiagramm eines IT-Systems im Kontext eines Malwareangriffes....98
Abbildung 49: UML Zustandsdiagramm eines IT-Systems im Kontext eines Hackangriffes....99

Abbildung 50: UML Zustandsdiagramm eines IT-Systems im Kontext eines DOS-Angriffes.....100
Abbildung 51: Beispielhaftes Infrastrukturmodell eines Mailservers.....101
Abbildung 52: Beispielhafter Fehlerbaum für einen Mailserver.....102
Abbildung 53: Infrastrukturmodell für den e-mail Dienst mit eingebauten Risk Responses.....104
Abbildung 54: Fehlerbaum für einen e-mail Dienst inklusive der Risk Responses.....105
Abbildung 55: Dichte der Anzahl von Ereignissen pro Jahr im inhärenten Fall.....106
Abbildung 56: Dichte der Anzahl von Ereignissen pro Jahr im residualen Fall.....106
Abbildung 57: Gegenüberstellung der Ereignisdauern bei inhärentem und residualem Risiko.....107
Abbildung 58: Dichte und Verteilungsfunktion der kumulierten Schadensummen eines Jahres.....107