

Sicherer Betrieb von IT-Systemen - Infrastruktur, das ungeliebte Kind

Service Level Agreements (SLA) lassen sich nur einhalten, wenn die Infrastruktur stimmt, von der Ausstattung des Gebäudes über den Rechneraum in den Serverschrank bis zum Aufbau der Systeme selber. Darstellung (und hoffentlich Diskussion) anhand zahlreicher Bilder aus verschiedenen Rechenzentren. Dabei möchten wir diskutieren, ob K-Fall Vorsorge (Vorsorge für den Katastrophenfall) tatsächlich nur eine Frage des Geldes ist und was die Konsequenzen aus dem 11.9. sind.

Christoph Maethner, Maethner Consulting GmbH
www.maethner-consulting.de

11.9.2001 – Der Maßstab ?

- Die Vorsorge für ein Desaster wie am 11.9. im World Trade Center muss nicht der Maßstab sein, denn ein solches Ereignis ist extrem unwahrscheinlich.
- Als Maßstab für den Standard-Anwender untauglich, hat das Ereignis aber verdeutlicht, dass gegen Totalverlust eines Rechenzentrums Vorsorge möglich ist
- Die Deutsche Bank war bereits wenige Stunden nach dem Zusammensturz der Twin Towers wieder voll operabel - im Ausweich-RZ in New Jersey ohne Datenverlust

Risiko Betrachtung

Verteilung der Ursachen von Katastrophen im RZ:

- Feuer 40 % => Infrastruktur
- Software/Viren 26 %
- Strom / Blitz 20 % => Infrastruktur
- Wasser 10 % => Infrastruktur
- Bauliche Mängel 2 % => Infrastruktur
- Sonstige 2 %

Der Totalausfall eines RZ ist also kein unrealistisches Risiko, auch ohne Krieg oder Terror. Dabei gilt oft:
Kleine Ursache, grosse Wirkung

Betrachtungsweise

- Es werden fehlertolerante Systemtypen gekauft, Fail-Over-Systeme oder (Disaster Tolerant) Cluster implementiert
- Es werden Journaling File Systeme eingesetzt und Transaktions-Monitore
- Platten werden gespiegelt
- In der Anwendungssoftware werden Wieder-Anlaufpunkte (Breakpoints) implementiert
- Der Infrastruktur wird selten genug Aufmerksamkeit geschenkt. Sie hat den Geruch von Beton und die Bedeutung der Toilettenreinigung: man braucht sie, aber man liebt sie nicht.

Beispiel 1: Infrastrukturbedingte Ausfälle

- Durch Wasserrohrbruch Ausfall der Wasserversorgung: die Klima-Anlage fällt aus (Fr. 23.30h) , das RZ hat am Mo. 8.00 h ca. 34 °C – zum Glück ohne Schaden.
- Durch Ausfall der (nicht redundanten) Klima-Anlage überhitzen die (redundanten) Datenbunker eines Unternehmens (42°C über 48 Stunden) ca. 3 Tbyte Daten in beiden (redundanten) Tape-Libs sind nicht mehr lesbar und endgültig verloren.
- Durch Verstopfung eines Abwasserrohrs dringt Regenwasser in die Unterverteilung des RZ ein, der Strom und damit alle Systeme müssen (ungeregelt) abgeschaltet werden. Bis alles wieder läuft: Schaden für den Anwender: > 2,5 Mio €

Beispiel 2 : Infrastrukturbedingte Ausfälle

- Arbeiten an einer USV-Anlage führen zu einem Kurzschluss in der USV – alles ausser den Servern läuft weiter
- Kurzschluss in einem schadhaften Kabel mit Personenschaden: während der Ermittlungen der Staatsanwaltschaft ist der Strom abgeschaltet und das Betreten des RZ verboten
- Kabelbrand in Unterverteilung: durch Verschmoren von ca. 3 m Kabel werden Systeme beschädigt (HCl-haltiges Brandgas) und kontaminiert (Furane und Dioxine) : Alle Oberflächen müssen feucht gereinigt werden.

SLA und Infrastruktur

- SLA ist das Buzzword der letzten Jahre. Ein SLA regelt die Verfügbarkeit von IT-Leistungen (Applikationen, Daten, Netzverbindungen)
- Ein SLA berücksichtigt die erforderlichen Auszeiten für Systeme, zu oft aber werden infrastruktur bedingte (meist massive) Ausfälle nicht bedacht und berücksichtigt.
- Deshalb verliert die Infrastruktur in der Regel beim Kampf um Auszeiten.
- Ordentliche SLA sehen die Möglichkeit dedizierter Auszeiten für Infrastrukturarbeiten vor. Das muss je nach technischer Gebäudeausstattung eine Auszeit für eine Totalabschaltung des RZ beinhalten.

Verfügbarkeit

- Die Gesamtverfügbarkeit ist bekanntermassen das Produkt (nicht die Summe) der Einzelverfügbarkeiten. Das schliesst **alle** Komponenten mit ein.
- An 5 Tagen je Woche Service von 8.00 – 18.00 h mit 98% Verfügbarkeit heisst: maximal 1 Stunde Ausfall pro Woche, 4 pro Monat oder 48 im Jahr.
- Ausserhalb dieses Zeitfensters finden Offline-Backups, OS- und Software-Updates, DB-Reorgs etc. statt.
- Auch alle Infrastruktur-Arbeiten müssen ausserhalb des Zeitfensters mit den o.a. Aktivitäten koordiniert werden.
- In der Regel bekommt die Infrastruktur die geringste Priorität, daher – definierte Auszeiten für Infrastrukturmassnahmen einfordern.

Infrastrukturbedingte Ausfälle

Es kommt zu Ausfällen im RZ, die wenn Infrastruktur

- versagt (ungeplant)
- gewartet werden muss (geplant)
- erweitert werden muss !!!!

Dabei gilt: infrastrukturbedingte Ausfälle sind in der Regel massiv, oft sogar Totalausfälle: Erweiterungen der Spannungsversorgung, Anheben des Doppelbodens, etc sind oft nur um den Preis einer Totalabschaltung zu haben. Daher muss mit entsprechender Weitsicht, Erweiterungsreserve geplant werden. (siehe Ende des Vortrags)

Theorie und Praxis

- Jeder Praktiker weiss, dass er bei guter Organisation wesentlich besser sein kann, als die theoretische Verfügbarkeit.
- Aber kein Praktiker kann diese bessere Verfügbarkeit garantieren – und nur darauf kommt es in Bezug auf den SLA an.
- Man kann unnötige Auszeiten vermeiden, indem man Um- und Erweiterungsbauten durch entsprechende Massnahmen ohne Störung des Betriebs vorbereitet und soweit wie möglich Erweiterungsreserven vorhält.

Ausfall-Folgen: ein Vergleich

- An einem Enterprise-Server dauert ein hardware bedingter Ausfall beispielsweise: 2 Stunden warten auf Servicetechniker, 1 Stunde Reparatur (er hat immer das passende Ersatzteil dabei ;-)) Hardwarecheck nach Einbau und Boot 1 Stunde, Reparatur der File-Systeme 1 Stunde = 5 Stunden Ausfall **eines** Systems, ca. 1 PT
- Der Ausfall der (nicht redundanten) USV-Anlage führt zum Crash aller angeschlossenen (ca. 350) Server. Nach Reparatur der USV, Hardware-Prüfung und -Reparatur, Boot der Betriebssysteme und Check und Reparatur der File-Systeme, teilweise Restore aus Backup. Auszeit für **alle** Systeme von ca. 12 Stunden, massiver Personaleinsatz (ca. 60 PT)

Elemente der Infrastruktur

- Gebäude
- Wasser Zu- und -ableitung
- Klimatisierung (Feuchte und Temperatur)
- Brandschutz und -erkennung, Löschmittelbevoratung
- Sicherheitseinrichtungen (Zutrittskontrolle)
- Transportwege / Durchgangsmaße, Tragfähigkeiten des Lifts, Beschaffenheit des Bodens
- Spannungsversorgung (des Gebäudes, NEA, USV und Verteilungen)
- Doppelboden: Klimatisierung, Netz- und Spannungsversorgung, Layout
- Passive Netzverkabelung

Gebäude (1)

- Beispiel: EMC3930 , 80x1,70x1,90 cm, $P = 22 \text{ kW}$, 1.890 kg auf 4 Rollen, Flächenlast ca. $12 \text{ kN} / \text{m}^2$, Punktlast ca. 5 kN
- Die gesamte Hauselektrik (Mittelspannungsanlage, GHV, NHV, Trafoanlage und das RZ selber gehören nicht in den Keller – Hochwasser ist wahrscheinlicher als Terror
- Das Gebäude und seine Infrastruktur müssen über ausreichende Ausbau-Reserven verfügen, wenn man nicht andauernd umziehen oder abschalten will.
- Lage sicherheitsrelevant: Hochwassergebiet, Einflugschneise, Risikogebiet, Autobahn, Öffentlicher Parkraum, Öffentlicher Zugang ...

Gebäude (2)

Neben dem Rechnerraum selber werden benötigt:

- Gebäudeautomatisierung und Überwachung
- Kommunikationsinfrastruktur (redundant) wie Telefonanlage
- Lagerräume für Neu- und Alt-Hardware
- Sichere Aufbewahrungsmöglichkeiten für Datenträger (Tapes), möglichst gebäudlich getrennt
- Büroräume für Administratoren
- Büro- und Lagerräume für Haustechnik, Netzwerk,
- Bevorratung Infrastruktur-Materialien: Kabel, Verlängerungskabel, Stecker, Kupplungen, Transportmittel, Ersatzplatten f. Doppelboden, Handlampen, etc.

Wasser Zu- und Ableitung

- Wasser läuft immer nach unten. Daher sollten kritische Installationen wie TGA, Spannungsversorgung oder das RZ selber nicht im Keller platziert werden.
- Wenn die Klima-Anlage für die Kälteübertragung auf die Wasserzufuhr angewiesen ist, wird der längerfristige Ausfall der Wasserversorgung zum Ausfall eines kompletten RZ führen
- Der Bruch von Wasser- oder Abwasserleitungen mit Wassereintritt in den Doppelboden ist üblicherweise katastrophal. Trotzdem werden immer wieder Heizungs-, Wasser- oder Abwasserleitungen in oder direkt neben RZ-Räumen geduldet.

Klimaanlagen

- Neben der Temperatur ist die Luftfeuchtigkeit für den zuverlässigen Betrieb von IT-Systemen entscheidend.
- Zu trockene Luft ($< 30\%$) führt zu elektrostatischer Aufladung und zu Rissbildung in Boards
- Zu feuchte Luft kann zu Kondensation und damit Pfützenbildung besonders im Doppelboden führen.
- Zu hohe Temperaturen ($>25^{\circ}\text{C}$) erhöhen signifikant die Ausfallrate von Systemen.
- Zu steile Gradienten beim Abkühlen von Systemen führen zu Problemen.
- Redundanzen in der Klimatisierung müssen gut geplant sein, sonst wird nach Ausfall eines Klimagerätes ggf. nur ein Teil des RZ kalt.

Brandschutz

- ▣ Vorbeugender Brandschutz: Minimierung der Brandlasten (Kartons, Doku, verpackte Ersatzteile – alles raus aus dem RZ) Wöchentliche Begehung nötig !!
- ▣ Schaffung von Brandabschnitten (über die LBO hinaus)
- ▣ Rauchdichtigkeit herstellen. Rauch ist ätzend, giftig
- ▣ Systeme zur Branderkennung und Früherkennung nutzen
- ▣ Kopplung Branderkennung / RLT (Brandschutzklappen)
- ▣ Ausreichend Gaslöscher bevorzugen, die Feuerwehr hat meist reichlich Wasser, aber wenig CO₂
- ▣ Bekämpfung, Eindämmungs- und Ausweichstrategie festlegen, kommunizieren und üben, üben, üben
- ▣ Identifizieren der Brandquelle üben

Zutrittskontrolle (1)

Wirksame Zutrittskontrolle umfasst:

- Einfriedung des Grundstücks (Zaun) mit Rolltor
- Sicherung und Überwachung der Aussenhaut (Einbruchssicherheit)
- Zutrittskontrolle zum Gebäude
- Definition von Sicherheitszonen / -bereichen
- Zutrittskontrolle zu Sicherheitsbereichen
- Sicherheitsregeln und deren Umsetzung
- Wachpersonal vor Ort
- Zurücksetzen der Zutrittsberechtigung für nicht mehr dem Unternehmen angehörige Personen

Zutrittskontrolle (2)

- Fremdfirmen müssen begleitet und beobachtet werden.
- Alles was reingebracht wird, muss überprüft und dokumentiert werden
- Alles, was raus soll, muss ebenfalls geprüft und dokumentiert werden. Auf einem einzigen Magnetband können sich genug Informationen befinden, ein Unternehmen zu kompromittieren oder zu ruinieren.
- Hier kommt man aber schnell an die Grenze des Machbaren.

Transportwege

- Transport von Systemen muss auch bei Regen möglich sein => Überdachte Transportwege inkl. Anlieferung
- Türen und Transportwege müssen den Durchgangsmassen der Systemen entsprechen (2,42 m hohe Racks passen in keinen Lift mit 2 m hoher Kabine)
- Der Boden muss glatt sein (kein Teppichboden)
- Der Transportweg darf keine Treppenstufen umfassen, auch steile Rampen können unüberwindlich sein.
- Der Transportweg muss das Gewicht der Systeme tragen können
- Der Transportweg muss (immer) frei sein.

Spannungsversorgung: Gebäude

Elektrischer Strom hält das Rechenzentrum am Leben:

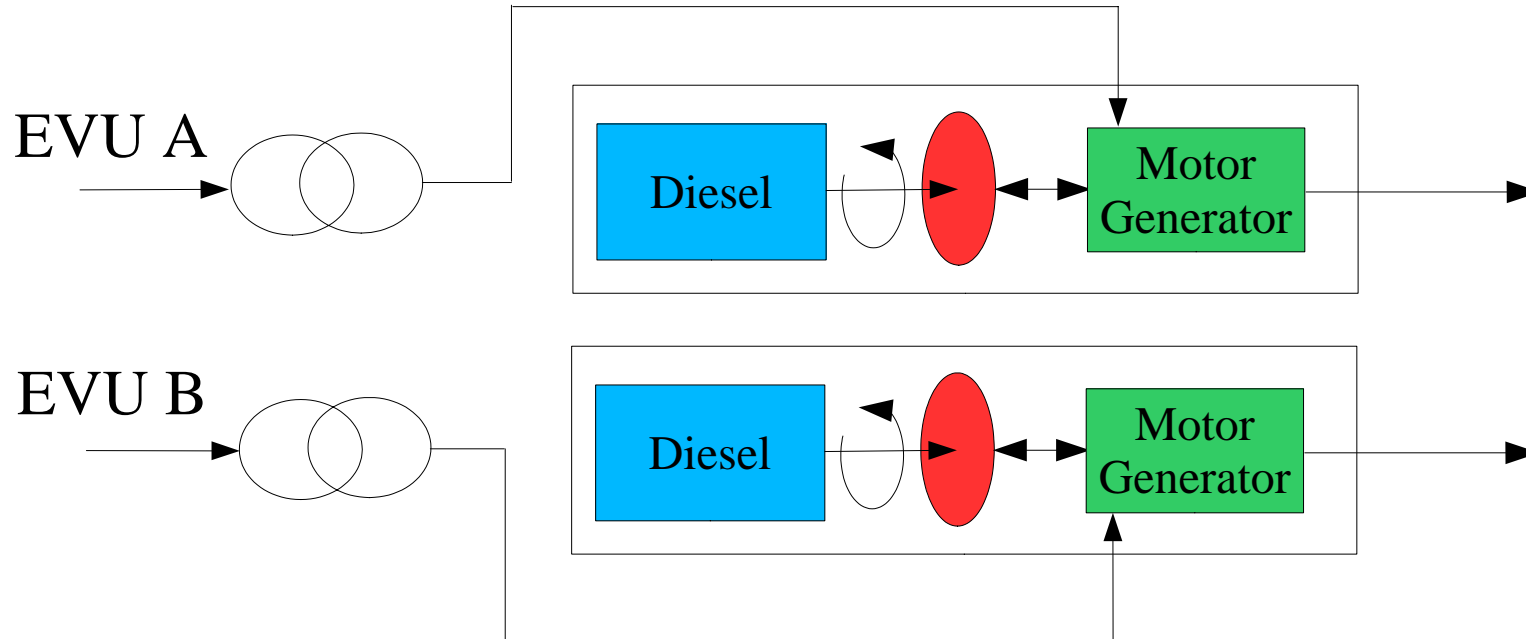
- Wer auf sein RZ bei Ausfall der Spannungsversorgung (Ausfall der Versorgung, Erweiterung der Anschlusswerte, Wartung der Mittelspannungsanlage, GHV, NHV) nicht verzichten will, braucht eine **NetzErsatzAnlage**.
- Eine NEA muss ausreichend dimensioniert sein, um neben den Systemen auch die **Klimaanlagen** bei jedem Lastzustand so zu betreiben.
- USV-Anlagen überbrücken nur kurzfristige Ausfälle oder Instabilitäten bei der Versorgung
- Bei Installation redundanter USV und NEA-Anlagen (egal ob N+M oder 2*N) muss der Ausfall einer Anlage jederzeit vollständig kompensiert werden können.

Spannungsversorgung: Redundanz

Redundanz bringt nur dann wirklich etwas, wenn sie konsequent realisiert wird:

- Redundante NEA verhindert Ausfall, wenn ein Diesel nicht anspringt
- Redundante USV sichert die Versorgung, wenn eine andere USV ausfällt
- Redundante Stromschienen verhindern Ausfall bei Beschädigung einer Schiene
- Redundante Spannungsversorgung im Systemschrank sichert den Betrieb bei Ausfall einer Spannungsquelle
- Redundante Netzteile erlauben den Ausfall einer Zuleitung

Redundante Gebäudeversorgung

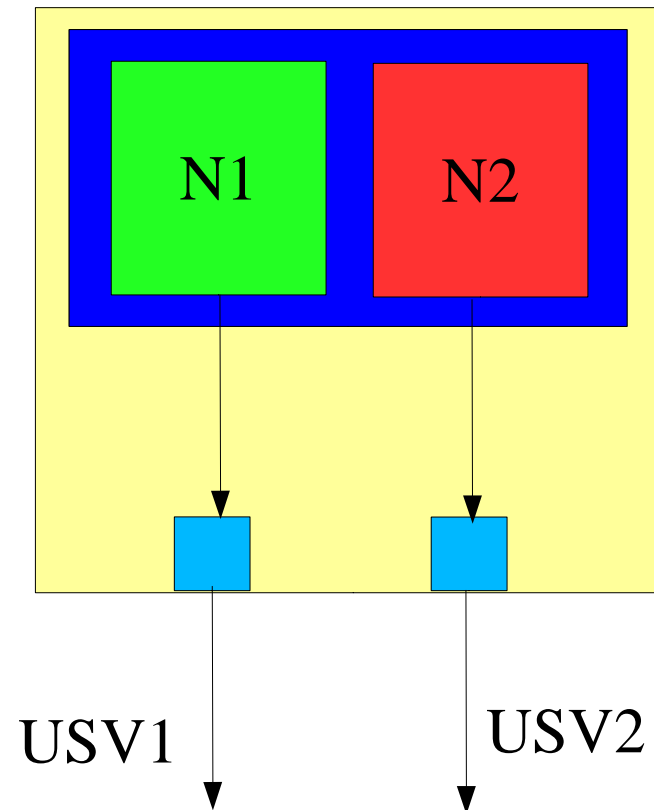


Stromversorgung: Rechnerraum

- Wartungen oder Änderungen in Unterverteilungen oder der Versorgung sollten nicht zu Unterbrechungen des Betriebs führen
- Stromschienensysteme vermeiden Probleme und Risiken mit der Bereitstellung zusätzlicher elektrischer Anschlüsse im laufenden Betrieb
- Die betriebenen Systemtypen sollten für eine voll redundante Spannungsversorgung ausgelegt sein, N+1 (und im Prinzip auch N+M) ist für Dauerbetrieb in RZ ungeeignet.
- Bei allen Systemtypen, besonders bei N+1 Systemen ist die Verkabelung des Systemschranks genau zu überprüfen und zu dokumentieren

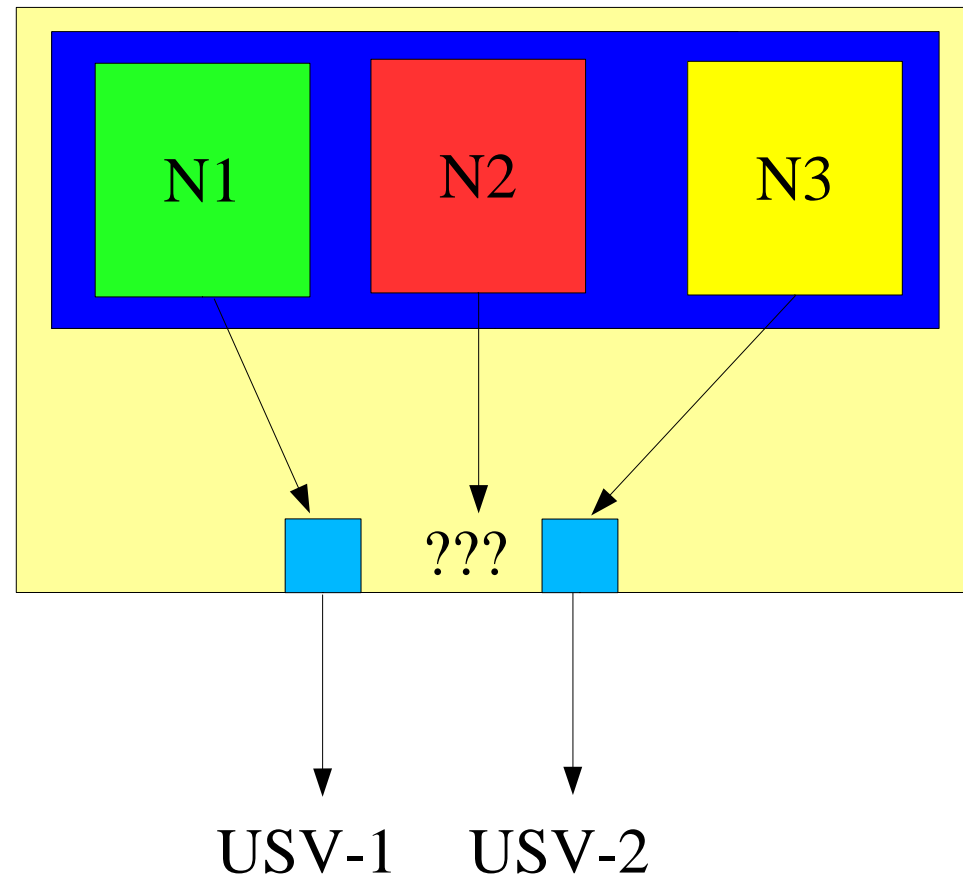
Vollredundante Systeme [2*N]

- Vollredundante Systeme sind für den Betrieb im RZ ideal geeignet
- Bei Ausfall einer Spannungsquelle bleibt das entsprechende System online
- Es gibt keinen „Single Point of Failure“



Teilredundante (asymmetrische) Systeme [N+M]

- Welches System bei Ausfall einer Spannungsquelle ausfällt, ist von der internen Verkabelung des Systemschrankes abhängig
- Es gibt keine eindeutige Zuordnung der Netzteile teilredundanter Systeme zu Spannungsquellen

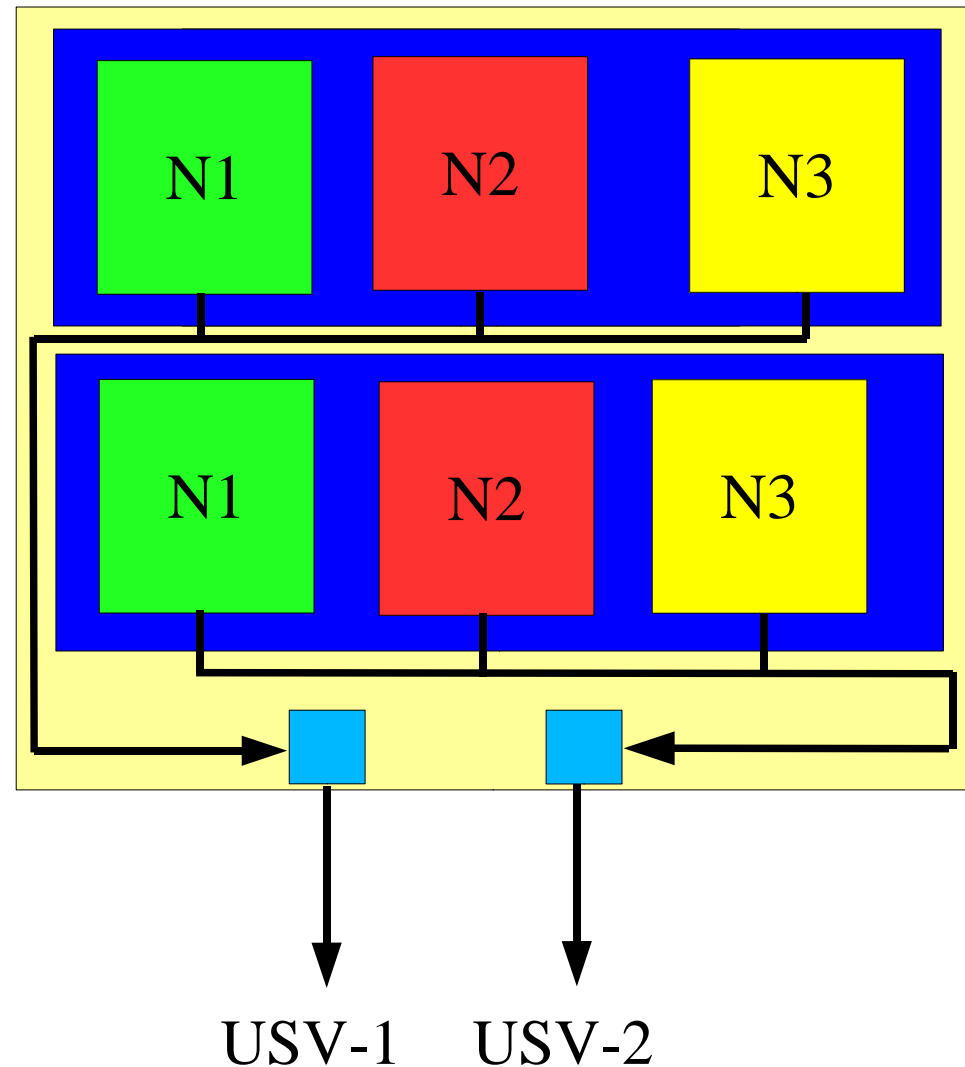


Asymmetrisch redundante Systeme

- Asymmetrisch redundante Systeme benötigen mehr als die Hälfte der elektrischen Anschlüsse, um online zu bleiben. Sie sind vor allem gegen interne Fehler (z.B. Usfall eines Netzteils) unempfindlich.
- Bei redundanter Spannungsversorgung des RZ führt der Ausfall einer Versorgung unter Umständen zum (vollständigen) Ausfall des betroffenen Systems.
- Für alle Infrastruktur-Wartungsarbeiten, die asymmetrisch redundante Systeme betreffen, müssen Auszeiten vereinbart werden. Der Aufwand ist daher besonders hoch

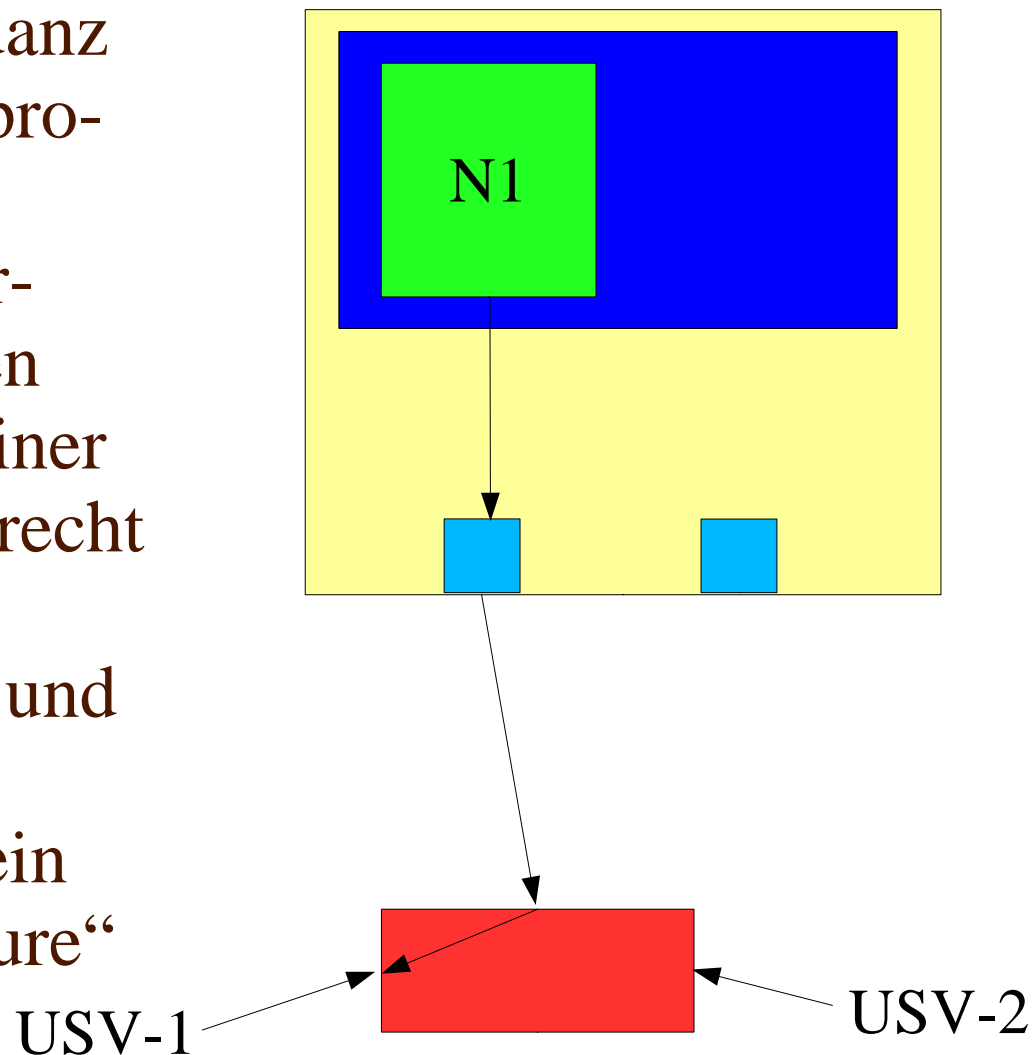
HA (High-Availability)-Cluster von teilredundanten Systemen

- Zumindest 1 Knoten eines HA-Clusters übersteht den Ausfall einer Spannungsquelle. Bei guter Doku weiss man auch vorher, welcher ;-)
- Es gibt eine eindeutige Zuordnung von Systemen zu Spannungsquellen
- Aber auch (geplante) Auszeiten erfordern vorherige Konfigurationsarbeiten



Systeme ohne Redundanz

- System ohne Redundanz sind für RZ-Betrieb problematisch
- Durch einen Transfer-schalter kann man den Betrieb bei Ausfall einer Spannungsquelle aufrecht erhalten
- Der Transferschalter und die dazu gehörigen Stromschienen sind ein „Single Point of Failure“



Layout des RZ

Bestimmende Faktoren sind:

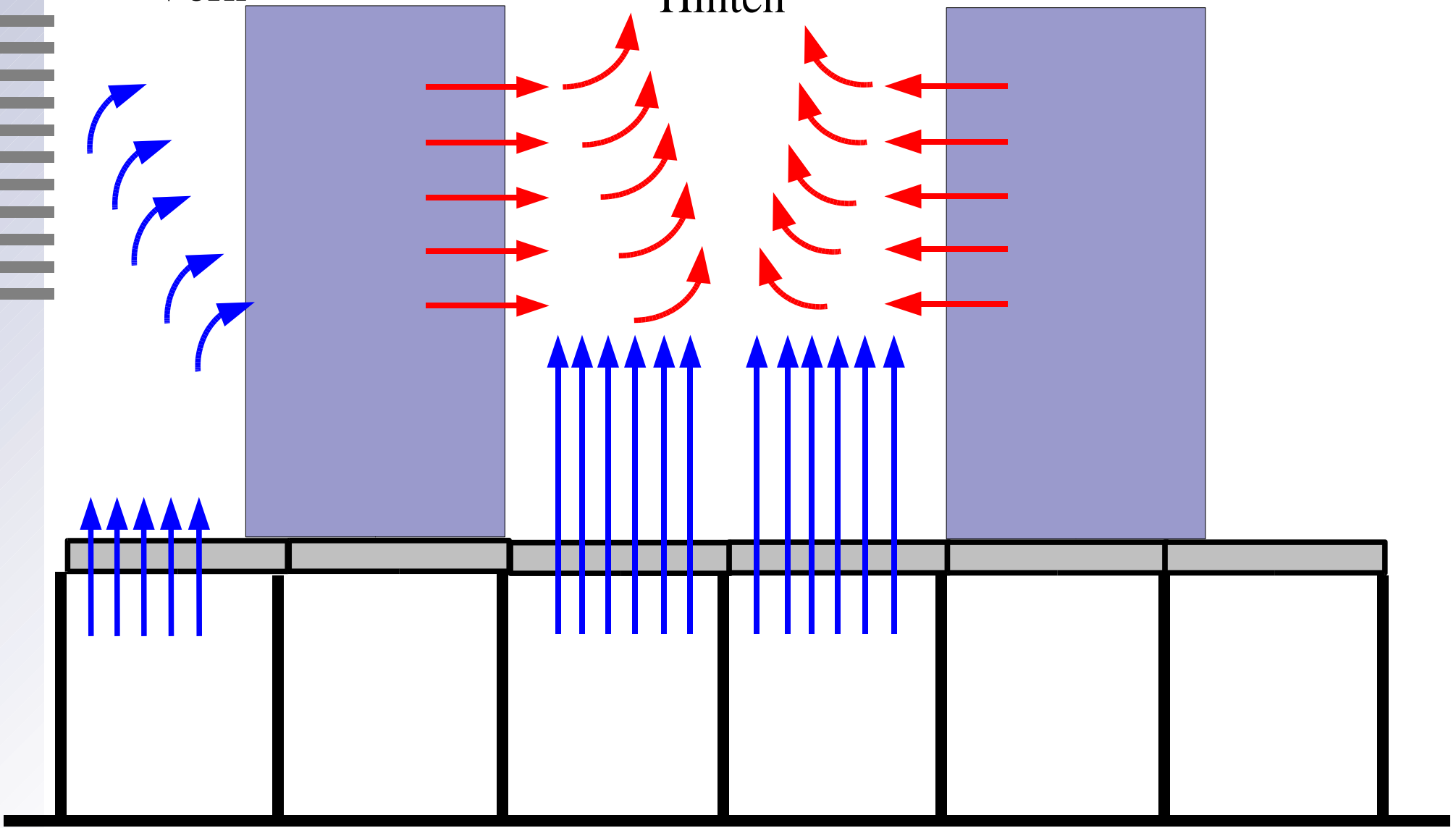
- Masse der installierten Systeme (z.B. BxT 80x90 cm)
- Erforderliche Servicefläche vor und hinter den Systemen (z. B. Je mindestens 100 cm)
- Transport- und Fluchwege, Ein- und Ausgänge
- Eventuell erforderliche Anschlussmasse an Säulen, Durchgängen etc.
- Bereiche mit veringierter Tragfähigkeit
- Rastermass der Bodenplatten (von 40x40 – 80x80 cm)
- Anforderungen an die Kühlung: je dichter die Systeme stehen, desto höher ist die erforderliche Kühlleistung und der Druck im Doppelboden

Doppelboden (Schema)

Vorn

Hinten

Vorn



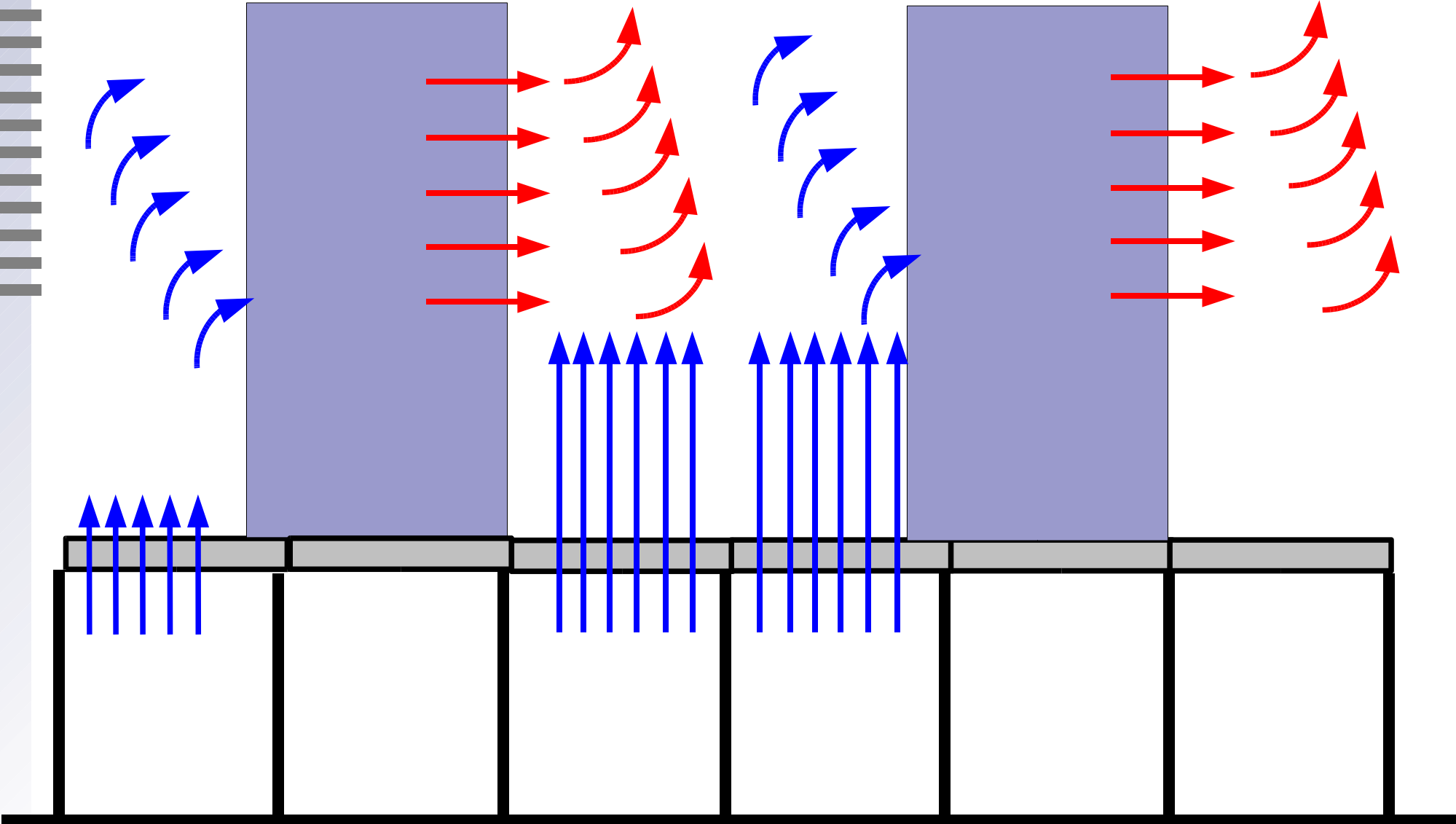
Doppelboden

Vorn

Hinten

Vorn

Hinten



Doppelboden – Funktion

Der Doppelboden hat mehrere Funktionen:

- Transport der Kühlungsluft
- Unterbringung der Kabel wie Spannungsversorgung, SCSI, Netzwerk, SAN etc.,
- Transportweg
- Stellfläche für die Systeme

Ein Doppelboden in einem in Betrieb befindlichen RZ kann nicht oder nur mit enormem Aufwand und Betriebsrisiko angehoben oder saniert werden. Er muss mit entsprechendem Weitblick geplant werden: das gilt für den Boden selber, wie die Kabletrassen und Stromschielen, die in ihm installiert werden.

Doppelboden - Planung

Für High-End-Systeme (EMC 3939, 1,78x0,90x1,95m³, 1.880 kg, Leistungsaufnahme ca. 22 kW) sollte ein Doppelboden nach dem Stand der Technik etwa wie folgt geplant werden:

- Traglast von $> 10 \text{ kN/m}^2$
- Bruchlast der Bodenplatten $> 5 \text{ kN}$
- Zul. Punktlast $> 5 \text{ kN/m}^2$
- Lichte Höhe $> 65 \text{ cm}$
- Breite der Transportwege $> 1,20 \text{ m}$
- Breite der Serviceflächen $> 1,00 \text{ m}$
- Keine Bodentanks für die Netzanbindung
- Ausreichend einstellbare Lüftungsplatten (ca. 50 %)
- Brand (frühest-) erkennung

Doppelboden Planung (2)

Die Zuführung der Daten-/ Powerleitungen ist ein ewiges Diskussionsthema:

- Zuführung von oben bringt die entsprechende Zuleitung bei jeder Bewegung in Gefahr
- Zuführung von oben erfordert immer eine Leiter für die Arbeiten
- Zuführung von unten ist „unbequem“, aber sicher
- Bei falsch geplanter Bodenhöhe blockieren Kabel den Luftstrom
- Praktisch alle Systemschränke sind für Zuführung aus dem Boden ausgelegt.

Doppelboden Ist

- Kabelvorhänge behindern den Luftstrom
- Kabelberge bilden eine Brandlast
- Altlasten belasten den Boden
- Stromschienen erlauben Anschlüsse bei Bedarf
- Netzanbindung unter dem Boden vermeidet Bodentanks / Schwachstellen



Doppelboden – Pflege

- Auf- und Abbau von Systemen darf nur in Abstimmung mit den Infrastrukturverantwortlichen erfolgen (z.B. wegen Grenzen in der USV-Belastung)
- Alle Kabel, die zu einem System gehören, **müssen** bei Abbau des Systems sofort zurückgebaut werden (Kontrolle erforderlich)
- Alle Netzanschlüsse müssen zurückgebaut werden.
- Kabeldurchführungen durch den Doppelboden müssen verschlossen werden, um Druckverlust zu vermeiden.
- Position, Anzahl und Öffnung der Lüftungsplatten muss permanent den sich ggf. ändernden Anforderungen angepasst werden.

Kennzahlen

- Ein gut geplantes RZ benötigt je Systemschrank (ca. 80X90cm) mit Transport und Service-Flächen ca. 2 m² je System, d.h bei 2 kW Leistung und einer Tapelib mit praktisch keiner Abwärme daneben entspricht das etwa 500 W/ m² . Das war bis vor wenigen Jahren ausreichend.
- Heute benötigt man ca. 900 W/ m² wobei die Reserven den Ausbau auf 1200 - 1500 W/ m² erlauben sollten.
- Mit den Werten kommt man an die Grenze der Abwärme, die sich mit einem herkömmlichen Doppelboden abführen lässt.

Ausblick (1)

- Tape-Libraries (ca. 1 kW) können durch SAN-Technik aus dem RZ entfernt werden. Die Flächen werden anschliessen für High-End Systeme (ca. 6 kW) genutzt
- EMC-Systeme (Sym VI) bieten im Vergleich zu Vorgänger-Modellen eine Reduktion des Flächenbedarfs **auf** ca. 10 % - entsprechend 10 facher Leistungsabgabe bezogen auf die Kapazität
- Blade-Server steigern die Systemdichte und damit die spezifische Leistungsaufnahme (48 HE = ca. 6 kW) auf einer Stellfläche, wo früher nur Intel 6-8 Server Platz fanden
- SUN Fire 15K , HP-Wildfire haben Leistungsaufnahmen um die 20 k

Ausblick (2)

- Um diese Wärme ab zu führen, muss der Druck im Doppelboden erhöht werden (Senken der Einblas-Temperaturen geht nicht wg. Kondensationsgefahr)
- Damit steigt der Druckverlust durch Undichtigkeiten
- Brandfrüh- und -frühesterkennung ist bei der entsprechenden Durchströmung nicht mehr möglich
- Feuer im Doppelboden wird sich durch die Strömungsgeschwindigkeit „explosionsartig“ ausbreiten
- Arbeiten im Doppelboden ohne Schutzausrüstung wird praktisch unmöglich bzw. unzumutbar

Lösung

- Doppelbodenlayout mit durchstömten und nicht durchströmten Zonen
- Der Luftaustritt je Stellplatz muss an den individuellen Bedarf anpassbar (regelbar sein)
- Dichtere Installation von Brandfrühesterkennung
- Tragfähigkeit $> 20 \text{ kN/m}^2$
- Das Hauptproblem wird die Regelung der Kühlleistung
- Bei verdichteter Stellweise wird auch das nicht helfen, ein zu warmes RZ ist ein unzuverlässiges RZ