# Zusammenfassung Vss - Grundlagen TEIL SLM

## Programm

Definierter Start und Ende, nur Offline Update möglich, Deterministisch (Liefert nur reproduzierte Werte).

## VSS (Gruppe von Systemen und Programmen)

Wird möglicherweise nie angehalten, Akzeptiert die ganze Zeit Inputs, Konfigurierbar im laufenden Betrieb, Möglicherweise nicht Deterministisch, Event-Gesteuert

**Vorteile**: Verteilte Standorte, Parallele Verarbeitungen möglich, Skalierbarkeit, Fehlertoleranz, Verfügbarkeit

**Nachteile**: Erhöhte Komplexität, Performanz Probleme wegen Netzwerkverzögerungen je nach Design möglich, Zuverlässigkeit -> jedes Teilsystem kann einen Ausfall verursachen, Transaktionssicherheit

## Middleware

Für die Vereinfachte Implementierung von Systemen, Transparenz, Events, Security, Kommunikation, Abstraktion, Beispiel: Java RMI, Web Services

## ACID (Atomicity Consistency Isolation Durability)

Atomare Operation, alles oder nichts ausführen – Garantie, Hinterlässt immer konsistente Daten, Durch Abgrenzung stören Transaktionen einander nicht. Dauerhaft garantiert nach einem Abschluss einer Transaktion, dass die Daten dauerhaft gespeichert sind, auch bei Auftreten von Fehlern.

## CORBA (Common Object, Request Broker Architecture)

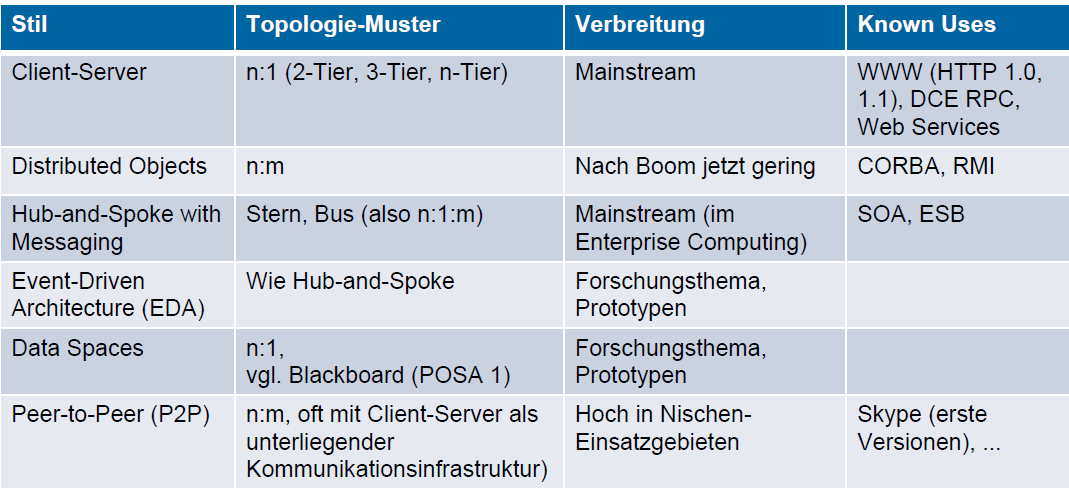
Spezifikation für eine Middleware für die Verbindung von verschiedene Softwareprogrammen zu einem System. CORBA-Middleware kann anstatt von Web Services eingesetzt werden. Ist Programmiersprachen unabhängig. Orbix ist eine Software die CORBA Standard implementiert. Mittels IDL (Interface Definition Language) wird eine Schnittstelle spezifiziert und später dann auch direkt Code daraus generiert. CORBA Einsatz heutzutage gering.

## Idempotent

Idempotente Methoden/Service usw. sind Methoden die beliebig oft aufgerufen werden können ohne einen Einfluss auf den Datenbestand im Server zu haben. Beispiel: «GET User» Request -> Idempotent, «POST new User» ist nicht idempotent.

# Enterprise and integration Patterns

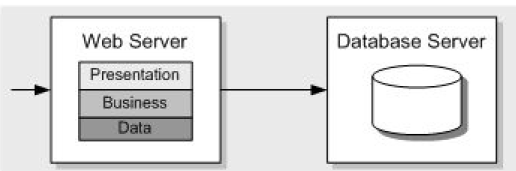
## Architekturstile (Architecture Patterns)



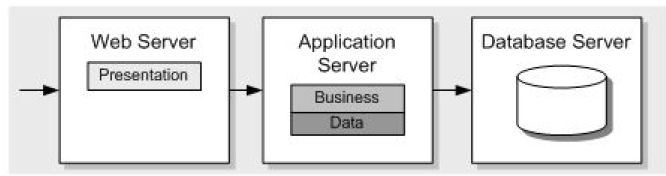
**Hub-And-Spoke:** Mainstream im Enterprise Computing. Benutzt einen Message Broker für das Verteilen von Nachrichten. Client bzw. Server verbinden zum Hub/Message Broker und müssen somit nicht für jeden Client/Server eine Verbindung aufbauen. Der Message Broker verwaltet die Verbindungen zwischen den Parteien.

## Deployment Patterns

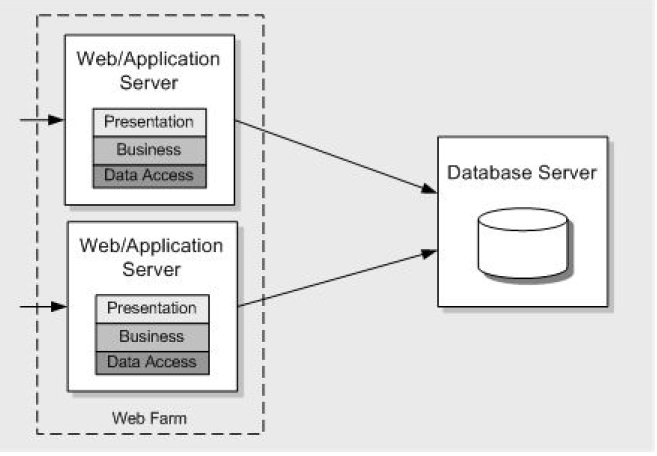
**Non-Distributed Deployment**

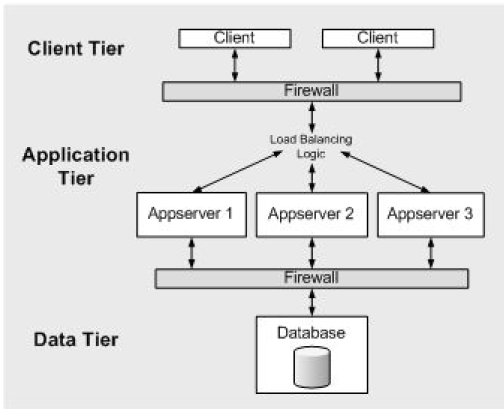


**Distributed Deployment** -> Vorteile: mehr Skalierbarkeit (Flexibilität), in der Regel bessere Performance, Kommunikationsaufwand darf aber nicht unterschätzt werden.

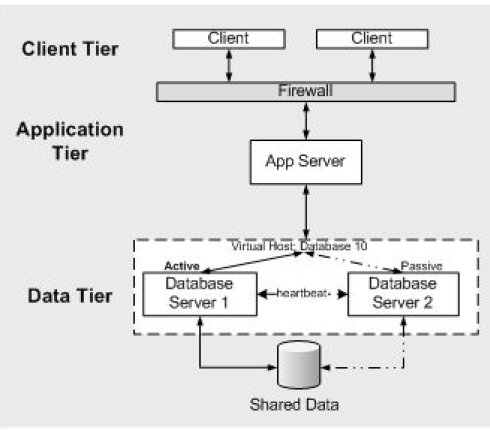


**Webfarm**



**Load-balancing Cluster** 

**Failover Cluster**



## Messaging Patterns

**Document Messages**: Transferieren von Daten zwischen Applikationen mittels Messages.

**Command Messages:** Führt Befehle anhand von Messages aus, sind grundsätzlich keine Spezielle Nachrichten, sie führen im Zielsystem einfach einen Befehl aus.

**Message Expiration:** Setzt ein Timeout auf einer Message um nicht mehr gültige Messages, welche zu lange Zustell-Zeiten haben, als ungültig zu erklären.

**Request-and-Reply:** Einfacher Einsatz eines Message Patterns, der Sender sendet eine Message und der Empfänger antwortet auch mit einer Message. 1<->1 Kommunikation.

**Push-pull:** Die Implementation von Publish und Subscribe Channel

**Fan-Out:** Verteilen von Messages auf unterschiedlichen Empfängern parallel. Der Sender wartet nicht auf eine Antwort.

## Messaging Channels

**Publish-Subscribe:** Anmelden von Pushmeldungen mit einem gewissen Topic oder gewissen Attributwerten um Messages zu erhalten. Typisch für mehrere Empfänger. Es gibt zwei Typen:

- Topic-based: Eine Message wird an einen bestimmten Channel geschickt, auf welchem er zu allen Subscriber des Channels kopiert wird. Der Sender ist verantwortlich für die Topics (= Channels).

- Content-based: Eine Message wird aufgrund von Attributen oder Content nur an bestimmte Subscriber geschickt. Die Subscriber spezifizieren die Attribute, welche die für sie interessanten Messages charakterisieren.

**Routing:** Auswahl des Consumers anhand Attribut.

**Point-to-Point Msg. Pattern** nur ein Empfänger, wichtig wenn Reihenfolge entscheidend ist. Sobald Parallelisierung ist die Reihenfolge nicht mehr gewährleistet.

## Messanging End Points

**Competing Consumer**: Mehrere Empfänger arbeiten Messages ab, eine Message wird nur einmal abgearbeitet (Parallelisierung der Consumers).

**Selective Consumer:** Ein Consumer verarbeitet nur Messages, welche seiner Filterkriterien entsprechen.

## System Management Patterns

Für die Überwachung von Queue basierten Systemen

**Wire Tap**: Punkt zu Punkt versenden von Messages inkl. Kopie auf einem Inspektionskanal.

**Detour:** Umleiten von Messages für die Analyse, Testen und für Debugging. Basierend auf dem Context wird die Message über einen speziellen Channel der Destination zu gestellt.

**Test Message:** Hilft mittels Test Messages rauszufinden, welche Komponenten Messages zwar korrekt entgegennehmen, sie aber wieder korrupt/falsch weiterleiten.

**Smart Proxy:** Adressen werden von ausgehenden Nachrichten überschrieben und bei einer Antwort wieder rekonstruiert. Das Zielsystem kennt die ursprüngliche Adresse nicht. Erstellt pro Service Message mit eigener integrierter Return Adresse einen ReplyChannel, auf welchen der Smart Proxy nachher die Antwort Message unbearbeitet weiterleitet.

**Message Store**: Speichert eine Kopie von allen Nachrichten in einem zentralen Ort.

**Message History**: Speichert den Fluss von Nachrichten

**Channel Purger:** Entfernt alte Messages um Spuren von alten Tests zu bereinigen.

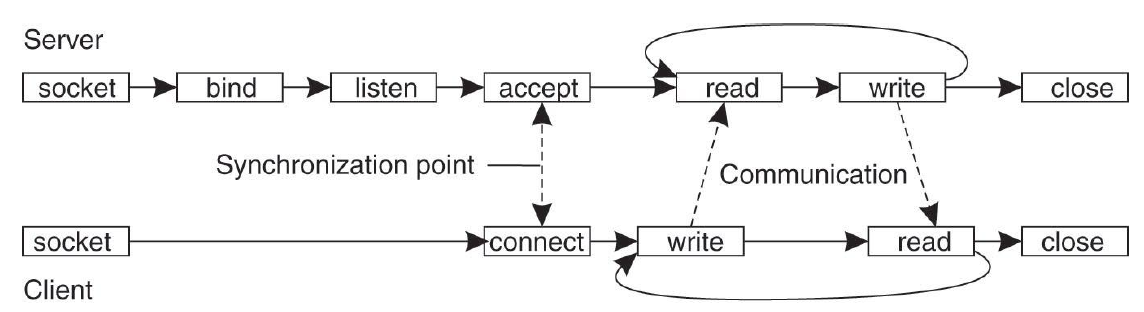
## Technologien und Ihre Messanging Patterns

**WCF**: Datagram (nur einseitig), Request-Reply, Duplex

**ZeroMQ:** Request-Reply (RPC, mehrere Clients auf mehreren Servern), Publish-subscribe (mehrere Clients auf mehreren Servern für Datenaustausch), Push-pull (verschiedene Nodes mit möglicher Parallelisierung, oder auch mehrere Empfänger für eine Message (Fan-Out Pattern)), Exclusive pair (1 Client, 1 Server)

# Low-Level Remoting TCP/IP Socket

## Socket

Ein Socket ist ein Verbindungsendpunkt der vom Programm wie eine gewöhnliche Datei beschreiben und gelesen wird. Es wird beim Server ein fester Port gewählt um den Socket anzusprechen. Der Client kann ein dynamischer Port wählen. 

**Nachteile von Sockets**: Keine Objekte, daher keine Typisierung. Message Exchange Pattern (MEP) muss selber implementiert werden. **Verbindung verweigern:** Hört kein Prozess auf dem Ziel Host, wird bei TCP ein ACK RST zurückgegeben, bei UDP ein ICMP Paket, welches aber nicht an die Java API weitergereicht wird. **Timeouts:** Der Server benötigt keine Timeouts da beliebig auf Anfragen gewartet wird. **Teilen von Datenübertragungen:** Kann Sinn machen da UDP sehr schlank ist und die Steuerung sollte persistent sein-> TCP. Abbruch einer UDP Verbindung wird vom Server nicht sofort erkannt.

## MPI (Berkeley Sockets and Msg Passing Interface)

Gleiches Protokoll für alle Plattformen, Standard für den Nachrichtenaustausch, Sammlung von Operationen, kein konkretes Protokoll, Für Datenaustausch innerhalb von Systemen zwischen Prozessen aber auch Austausch zwischen verteilten Systemen.

## Netcat

Tool um UDP und TCP Protokolle zu testen. UDP und TCP Socket. Kann auch gut mit Java Socket erstellt werden.

## Java Socket API

**TCP:** Erstellen von client Socket Socket(host, port); empfangen von Bytestream getImputStream(); (Blockierend) Server Socket der auf Requests wartet. ServerSocket.accept().

**UDP:** DatagramSocket(port) den Server Socket erstellen und der InetSocketAddress(host, port) für den Client. Der Client musst zuerst mal Daten Senden damit der Server den Port des Clients kennt. Darum socket.send(new DataPackat(bytesObj, bytesObj.lenght, InetAddress.getByName(‘localhost’), port)); danach mit while(true){ socket.recive(pkt); auf Lieferung von Server warten (Blockierend!). Warten auf mehrere Clients kann mit einer while(true){…} Schlaufe und jeweils starten eines neuen Threads Serverseitig erfolgen.

## WebSockets

Neue Entwicklung von Sockets auf Basis von HTML 5 -> Realtime Web Applikationen, hat nichts mit den 'alten' Sockets zu tun. HTTP Verbindungen werden vom Client geöffnet, danach kann Server an den Client Daten schicken ohne dass dieser eine Anfrage sendet. (Gegensatz zu reiner HTTP Verbindung)

## Service Activator

Der Service Activator kann eine Drittsoftware sein. Macht nicht anders als eine Message entgegennehmen und gibt sie weiter, Möglich als One Way oder auch als Two Way

# Asynchronous Messaging

**Message Descriptor:** Identifier und Steuerinfos

**Message Data:** Inhalt der Meldung in der Struktur des Zielprogrammes. Queue-basiertes Messaging gestattet die flexible und lose Koppelung unterschiedlicher Systeme.

## MQ Manager Vorteil Persistent vs non Persistent

Persistent ist teurer als non Persistent und muss ein Log nebenbei schreiben. Bei persistenten Nachrichten sind nach Stromausfall/Absturz, keine Daten verloren gegangen.

## Aufruf von Queue

Consuming -> nach dem Lesen löschen

Non-Consuming -> nach dem Lesen noch verfügbar

## Queue von Java: Java Message Service (JMS)

Dokumentation notwendig um JMS anzuwenden

JMS ist erst später zur Projektorientierten Programmierung "hinzugefügt" worden. JMS kann Point-to-Point Channel -> Queue API und Publish-Subscribe Channel -> Topic API. **Vorteil** Es können von Java <-> Java ganze Objekte ausgetauscht werden

**JMS Message Reliability Levels: Best effort nonpersistent**: Messages werden einfach verworfen, wenn ein System oder Netzwerk abstürzt oder überlastet ist. **Express nonpersistent**: Messages werden einfach verworfen, wenn ein System oder Netzwerk abstürzt. **Reliable nonpersistent**: Messages werden einfach verworfen, wenn ein System abstürzt. **Reliable persistent:** Messages könnten verworfen werden, wenn ein System abstürzt.

**Assured persistent:** Messages werden nicht verworfen.

## RabbitMQ

* OpenSource MQ
* Basiert auf AMQP (Standard protocol for the interoperability between all messaging middleware)

Work queues

Publish/Subscribe: Senden von Nachrichten zu verschiedenen Empfängern gleichzeitig

Routing: Erhalten von Nachrichten nach Kriterien

Topics: Erhalten von Nachrichten anhand von Muster;

# RMI & Web Services

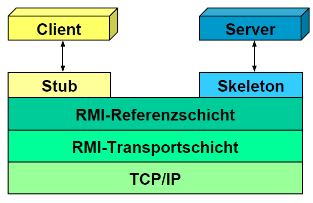
RMI ist RPC in Java, RMI kann ganze Java Objekte übermitteln. Interfaces und Klassen können vom Server zu Client und vom Client zum Server transportiert werden. Eine Aktualisierung einer Klasse kann nur auf dem Server stattfinden. Der Client kann per RMI die neuste Definition automatisch holen, es muss somit kein Code auf dem Client aktualisiert werden. Die Kommunikation ist für den Entwickler verborgen.

## RPC (Remote Procedure Calls)

Aufruf einer Prozedur auf einem anderen System. Dabei wird der Prozedurname bzw. eine ID und die Parameter übermittelt. Es sind keine spezifischen Messages, es sind einfach Messages mit Kommandos. Vereinfacht die Client/Server Programmierung. **RPC Aufruf:** 1. Client ruft eine Prozedur auf inkl. Parameter 2. Der *Client Stub* packt die Parameter in eine Message (**marshalling**) 3. OS verschickt die Nachricht 4. Das OS auf dem Server leitet die einkommende Messages an den richtigen Server Stub weiter 5. **Unmarshalling** auf dem Server 6. Stub ruft die Serverprozedur auf und verpackt die Antwort gleich. Während dieser Prozedur bleibt der Client Code meistens stehen und wartet auf die Antwort vom Server.

Asynchronous RPC: versenden/empfangen und verarbeiten von Nachrichten zu jeder Zeit.

## RMI Begriffe

Ein **Stub** Ist ein Stellvertreterobjekt (Remote Proxy), das Clientaufruf an Server weiterreicht. Ein **Skeleton** nimmt aufrufe des Stubs entgegen und leitet sie an Serverobjekt weiter. **Referenzschicht** = Namensdienst (Registry). **Transportschicht** = Verbindungen verwalten.

## Details von RMI

Wenn ein Objekt eine Instance von Remote ist wird die Referenz übertragen sonst der Wert (serialisiert). Server kann remote Wert ändern, Zugriff ist teuer.

**Wichtig im RMI ist/Nachteile:** Timeout Management, gleiche JVM Version notwendig, Firewalls blockieren im Standard die Ports. **Blockierender Stub** hat den Vorteil, dass es einfacher implementiert ist, jedoch ist asynchrone Programmierung nicht mehr möglich.

## Implementation von RMI

Serverseitig:

Public interface hello extends **Remote**{

String sayHello () throws **RemoteException**;

}

Public class HelloImpl extends **UnicastRemoteObject** implemnts Hello {

Public HelloImpl() threows Remote Exception{ super(); }

Public String sayHello() throws RemoteException{

Return "hello VSS";

} //create sever

Public class **HelloServer**{

Public static void main(String args[]){

**HelloImpl obj = new HelloImpl();**

**Naming.rebind("rmi://[hn]/remoteHello", obj);**

} //try and catch auch notwendig!

Clientseitig (Interface analog von Server):

Public class RmiClient {

Public static void main(String[] args){

**Hello obj** = (Hello)Naming.lookup("rmi://[hn]/remoteHello");

String message = **obj.sayHello();**

Sysout(message);

} //exeption handling mit try & catch notwenidg!

## IDL (Interface Definition Language)

IDL ist eine Spezifikationssprache welche Interfaces zwischen verschiedene Plattformen, Programmiersprachen und OS beschreiben. Mit IDL werden RPCs Programmiersprachen unabhängig beschrieben und dienen den Stubs für die Erstellung des Codes. **Beispiele** CORBA.

## WSDL

RPC Style mit WSDL wird mit wie folgt definiert:

<soap:envelope><soap:body>

<methodenname>

<param1>value</param1>

<param2>value</param2></…></…></..>

Unterschied RPC Style und Dokument Style (Standard für Internetkommunikation): Applikation macht das Mapping der XML Werte, Wie die WSDL Datei aussieht entscheidet die Applikation. Die Applikation muss *marshalling* und *unmarshalling* handeln. (Marshalling => Umwandeln der Daten in brauchbare Objekte)

# Operationale Modelle, Design for Performance

## Nicht funktionale Anforderungen (NFRs)

Verfügbarkeit, Compliance, Wartbarkeit, Performance (Reaktionszeit, Latenz, Durchsatz, Service Zeit), Privatsphäre, Wiederherstellung, Elastizität, Ausfalltoleranz, Robustheit, Skalierbarkeit, Usability

## Performance

Ein System ist nie schnell genug, es ist immer ein Gleichgewicht zwischen Kosten, Funktionalität und Performance.

## Zusammenhängende Aspekte von Performance

**Response Time** Steigt ab ca. 60% Auslastung exponentiell -> eingehende Anfragen müssen zuerst warten

**Throughput:** Druchsatz damit alle Konten einer Bank über Nacht abgeschlossen werden können. **Capacity**: Richtige Komponentengrösse um Auslastung zu Tragen. **Contingency**: Kann auch mit Ausnahmesituation noch umgehen -> Beispielsweise auch Jahresende Abschluss möglich.

# Design for Scalability and Availability

**Skalierbare Grösse:** Schnellere und mehr Komponenten (CPU, RAM, HDD, Switches usw.)

**Skalierung der Generation**: Technologie Evaluation

**Vertikale Skalierung**: Grössere/Schnellere Server

**Horizontale Skalierung**: Anzahl Server erhöhen.

## Load balancing

**IP Spraying**: eine IP steht für die Clients zur Verfügung und der HTTP Verkehr wird vom Balancer einfach an verschiedene Webserver weitergeleitet. Das Verteilen von Anfragen erfolgt nach 3 Prinzipen: **Round Robin**, **Least-recently-used**, **Workload-based** (nur mit Monitoring)

## HTTP Load Balancing

Behalten von Session trotz Load Balancing:

* Speichern der Session Daten auf dem Server
* IP source affinity: Anhand der Soruce IP Adresse wird wieder gleicher Server zugewiesen: Problem, was wenn der User IP Adresse wechselt?
* Application layer persistence: Balancer benutzt den Session Cookie um den User zu identifizieren

Identifizierung einer Session:

**Affinity**: Verwendet Layer unterhalb von Appl. Layer

**Persistence**: verwenden von Applikation Layer

**Session affinity:** verwendet Cookies

## Availability

System verfügbar und beantwortet Anfragen korrekt.

**MTTF**: durchschnittliche Zeit bis ein System ausfällt

**MTTR**: durchschnittliche Zeit bis ein System nach einem Ausfall wieder online ist.

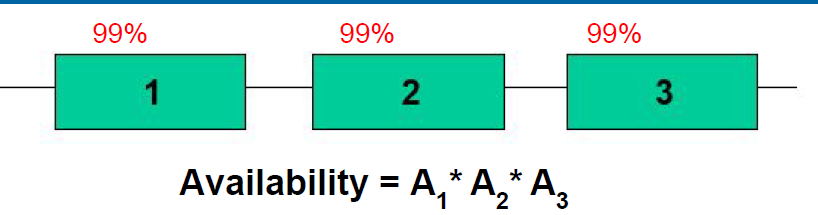
**MTBF**: Zeit zwischen 2 Ausfällen

**RPO**: Recovery Point Objective. Maximum tolerierte Zeit nach einem Fehler, in welcher Daten verloren gehen können.

**RTO**: Recovery Time Objective. Zeit in welcher der Service wiederhergestellt werden muss. Non Functionl Requiment (NFR): RTO + RPO

Kosten für die Verfügbarkeit steigen Exponentiell. Der Kritische Punkt ist bei ca. 95% Verfügbarkeit.

## Verfügbarkeit berechnen





**Verfügbarkeit verbessern**: Vermeiden von Single Point of Fail (SPoF)

**Cold Standby:** Das Backupsystem läuft nicht und hat keine aktuellen Daten. Wird bei einem Ausfall von den primären Komponenten hochgefahren und das Backup der Daten wird restored. Ausfallzeit von ein paar Stunden.

**Warm Standby:** Ein Watchdog überwacht das produktive System, antwortet dieses nicht, werden die Dienste auf dem Backupserver aktiviert. Verbessert MTTR. Bei Überlastung vom Produktiven Server kann es passieren, dass der Backupserver fälschlicherweise hochgefahren wird, Konflikte können die Folge sein.

**Hot Standby**: Min. 2 Nodes laufen immer. Daten werden immer wieder auf dem Backupserver kopiert. Recovery innerhalb von Sekunden.

**Active-Active (Load Balancing):** Ausfälle sollten nicht bemerkbar sein.

**Fail Fast Principle (Circuit Breaker Pattern):** Anfragen auf einem Service werden überwacht, wenn es Ausfälle/Fehler gibt, wird der Service gestoppt um nicht Folgefehler zu verursachen. Oft kann der Fehler auch automatisch behoben werden und dementsprechend den Dienst wieder starten.

## Scalability

**Size Scalability:** Bessere und mehr Komponenten **Generation Scalability**: Technology Evolution, **Vertical Scalability, Scale Up**: Grössere Server/Bandbreite/Storage **Horizontal Scalability Scale Out:** Anzahl Server/Nodes erhöhen.

# VSS Management

Richtige Verwaltung von Systemen ist ein grosser Bestandteil für die Verfügbarkeit von Systemen.

Monitoring, Trend Analyse, Configuration Mgmt. Usage, Performance Monitoring, Security (überwachen von Login)

## ITIL

**Service Strategy:** Financial, Service Portfolio, Strategy

**Service Design:** SLM, Availability, Capacity, Service Catalog, Information Security, Requirements Engineering

**Service Transition:** CM, CMDB, Release, KM, Validation and Testing

**Service Operation:** IM, PM, RF, Access, Event

**Continual Service Improvement**: Service Improvement, Service Measurement, Service Reporting

**Configuration Management:** Beispiele: Dokumentation der Build-beteiligten System, Unautorisierten Zugriff auf Files verhindern.

**CI:** Configuration Item. ITIL Bezeichnung: «Any Component that needs to be managed in order to

deliver an IT Service. […] CIs typically include IT Services, hardware, software, buildings, people and formal documentation such as Process documentation and SLAs.»

## FCAPS

**F:** Fault. Faults erkennen, isolieren, korrigieren und loggen. **C:** Configuration. System Konfiguration überwachen und informieren, im Falle von Änderungen **A:** Accounting/Administration. Usage Statistiken sammeln, Usage Informationen tracken, Quotas anwenden **P:** Performance. Durchsatz, Antwortzeiten, Error/Failure Raten, Utilization **S:** Security Authentifizierung & Autorisierung, Audit

## JMX (Java Management Extensions)

**Prinzip:** Verwaltete Ressourcen (z.B. Objekte aber auch SW oder HW) in Objekte verpacken - sogenannte MBeans. Für die Verwaltung und Überwachung von Java (JVM, Java Virtual Machine). MBeans ist die API der JMX. JMX wird in 3 Ebenen unterteilt **Distributed Service Level:** Schnittstelle für die Kommunikation gegen aussen (ausserhalb der JVM). **Agent Level**: Agent die die Kommunikation mit den Ressourcen übernehmen**. Instrumentation Level**: Auf diesem Level befinden sich die Ressourcen, die überwacht und konfiguriert werden.

JMX kann auch verwendet werden um Konfiguration in einer laufenden Software zu ändern. Achtung das Ändern der Konfiguration kann zu inkonsistenten Zuständen führen: **Complexity** wird erhöht, **Availability** kann erhöht werden, da kein Neustart der Applikation notwendig ist um neue Konfiguration zu laden, **Stability**: kann negative Auswirkungen haben (erhöhte Komplexität).

## Logging

Easy to use, Zentralisiert, kleiner Effekt auf die Performance, verwenden von Loglevels, Standard Formate und Tools, Vermeiden von W Fragen in Logs (Wer, Wann, Wo..usw) **Verschiedene Zielgruppen: Entwickler:** Technische Details, welche Methode, Variablenwerte usw. **Systemadministrator:** Netzwerkauslastungen, Start/Stop Protokoll, Konfigurationsfehler. **Supporter**: Vorgehen von User nachvollziehen.

**Loganforderungen für VSS:** Lognachrichten müssen an Instanzen gebunden werden können. Bei Netzwerkproblemen dürfen keine Nachrichten verloren gehen. Verteilte Aufrufe müssen ersichtlich sein.

# CFIA (Component Failure Impact Analysis)

Erkennen von kritische Verfügbar- und Performancerisiken, wie auch spezifizieren von Messgrössen. Prozess der beim Design eines Systems durchgeführt werden muss. Massnahmen welche getroffen werden können um Ausfälle zu vermeiden oder Ausfälle zu verkürzen. **Phase 1**: Tätigkeiten, welche sofort erreicht und gemacht werden können (in ein paar Wochen) **Phase 2**. Hoch priorisierte Tätigkeiten, welche in mittlerer Zeit erreicht und gemacht werden können (ein paar Monate) **Phase 3**. Niedrig priorisierte Tätigkeiten und solche, welche fundamentale Änderungen mit sich ziehen.

**Expanded CFIA**: für jeden CI folgende Fragen beantworten: Was ist es? Was kann schiefgehen? Wie kann ich es wiederherstellen? Was passiert, wenn dieses CI ausfällt? Wer ist zuständig? Usw.

**Risk Rating Tabelle:** Beantworten folgenden Fragen mit Zahlen von 1-10 = 1 sehr gut, kleines Risiko/ 10 sehr schlecht, grosses Risiko: Businessimpact, Häufigkeit des Ausfalls, Wie einfach ist es den Fehler zu erkennen, vorhandenes Wissen für die Wiederherstellung, Wiederstellungsmassnahmen, Automation im Wiederherstellungsprozess, Sicherheitsschutz (wie exponiert), Patchintervall/-status

**Allgemeine Massnahmen:** Eliminieren von SPOF, aktuelle und ausreichende Dokumentation, Monitoring, automatisierte Wiederherstellung von Systemen, angemessene Backups, eliminieren von «Key Personen» in kritischen Systemen.

## CFIA Matrix

|  |  |  |
| --- | --- | --- |
| **CI** | **Service 1** | **Service n** |
| **Pc1** | X | A |
| **App1** | M |  |

**X**-> S1 nicht verfügbar ohne PC1; **A** -> Redundanter PC1 verfügbar, kein Ausfall wenn nur PC1 ausfällt **M**-> Wie A nur manuell eingreifen nötig. Leer App1 hat keinen Einfluss auf Service n. CIs mit den meisten X sind am Kritischen.

# Verschiedenes

## CloudSum

Java Anwendung die als Client oder Server gestartet werden kann. Der Server ist nicht idempotent. Setzt für die Kommunikation ZeroMQ ein.

## ZeroMQ

Kommunikations-Library für die Kommunikation zwischen Prozessen und Threads, Neben TCP Kommunikation wird auch **inproc** als Kommunikationsprotokoll innerhalb eines Prozesses angewendet.

**Intraprozesskommunikation:** Kommunikation innerhalb eines Prozesses beispielsweise mit **inproc.** Kein Verbindungsaufbau nötig. Vorteil von Intraprozesskomm. In vergleich mit gemeinsamem Speicherzugriff: keine Race conditions, Synchronisation nicht notwendig.

**Interprozesskommunikation:** Datenaustausch über bsp. TCP. Je grösser die Nachricht desto grösser der Kommunikationsaufwand verglichen mit Intraprozesskomm.

## Weitere Tools

**Redis**: Persistent Datastore, Frei erhältliche KeyValue Data Store der auch als Verteilter Cache benutzt werden kann. Redis verwendet für die Kommunikation **RESP** (Redis Serialization Protocol). RESP kann verschiedene Datentypen serialisieren, wie Integers, Strings, Arrays. Befehle und Argumente werden als StringArray am Server versendet.

**Chirp:** Twitter ähnliche Applikation,Verwendet Redis, Chirp verwendet Request/Replay als Message Exchange Pattern, Aufrufe sind Synchron,

**Gatling:** Stress Tool für Blackbox Tests

**Slf4j:** Logging für API Logging, eingesetzt von Chirp.

**Logback**: Loggingframework von Chirp eingesetzt.

# Zusammenfassung Vss - TEIL MEF

# Bitcoin

Ein Bitcoin ist nichts anders als ein durch Transaktionen gesichertes Guthaben, welches in einem Peer-2-Peer Netzwerk in einer Block-Chain gespeichert ist. Geschützt durch einen privaten Key, welcher benötigt wird, den dir zugeteilten berechneten Betrag weiter zu überweisen.

## Transaktion

Eine Transaktion wird mit deinem privaten Schlüssel signiert und muss durch das Peer2Peer Netzwerk bestätigt werden. Eine Bestätigung dauert min. 10 min bis hin zu x Stunden. Dies hängt auch vom freiwilligen bezahlten Transferbetrag ab, je grösser desto schneller wird ein Transfer bestätigt

## Blocks

**Block Chain:** Da die verschiedenen Blocks alle zum jeweiligen Vorgängerblock referenziert werden, entsteht eine Kette – die Block Chain. Es wird immer die längste Chain verwendet, was sicherstellt, dass dort drin nur valide Transaktionen verarbeitet wurden.

**Block:** Beinhaltet: Random Hash Value (SHA256(SHA256(x))) vom Vorgängerblock, eine Liste von Transaktionen welche in der Zeit seit dem letzten gültigen Block stattgefunden haben und die Nonce. Ein Block ist gültig, wenn die Random Hash Funktion, über den Inhalt des Blockes angewandt, einen Hash Value rausbekommt, welcher mit mindestens X Nullen beginnt. Wird ein gültiger Block gefunden, erhält der Finder einen Block Reward von aktuell 12,5 Bitcoins.

**Mining:** Das Erzeugen von Blocks nennt sich Mining. Dafür sind grosse Rechenleistungen notwendig. Je mehr Leute die Hashes erzeugen desto grösser wird die Schwierigkeit den nächsten Hash bzw. Block zu finden. Die Währung reguliert sich selber, so dass ca. pro 10 min ein Block generiert wird und jeden 14. Tag wird die Schwierigkeit angepasst. Die momentane Grösse der Blocks von 2 MB lässt 3 Transaktionen pro Sekunde zu.

## Confirmation

Je mehr Bestätigungen eingetroffen sind, desto sicher ist es, dass die getätigte Transaktion gültig ist. Für grössere Beträge macht es Sinn bis zu 6 Bestätigungen abzuwarten.

## Schwächen

Wenn ein Minor, mehr als 50% der auf der Welt für minen zur Verfügung gestellte Rechenpower verfügt kann er seine ‘eigene’ block-Chain machen und so bestimmen welche Transaktionen akzeptiert werden.

## Double Spending

Will man versuchen den gleichen Bitcoin oder den gleichen Bitcoin Anteil (z.B. 0.1 Bitcoin) an zwei verschiedene Empfänger schicken, wird dies als Double Spending bezeichnet. Um diesem Problem entgegen zu wirken, wird jede Transaktion mit der längsten verfügbaren Blockchain abgeglichen. Bei diesem Abgleichen wird sichergestellt, dass dieser explizite Bitcoin überhaupt im Besitz des Senders ist und dass dieser nicht schon versendet wurde.

# Naming

Naming Systeme werden benötigt um Entitäten im Netzwerk zu benennen und so ‘erreichbar’ zu machen.

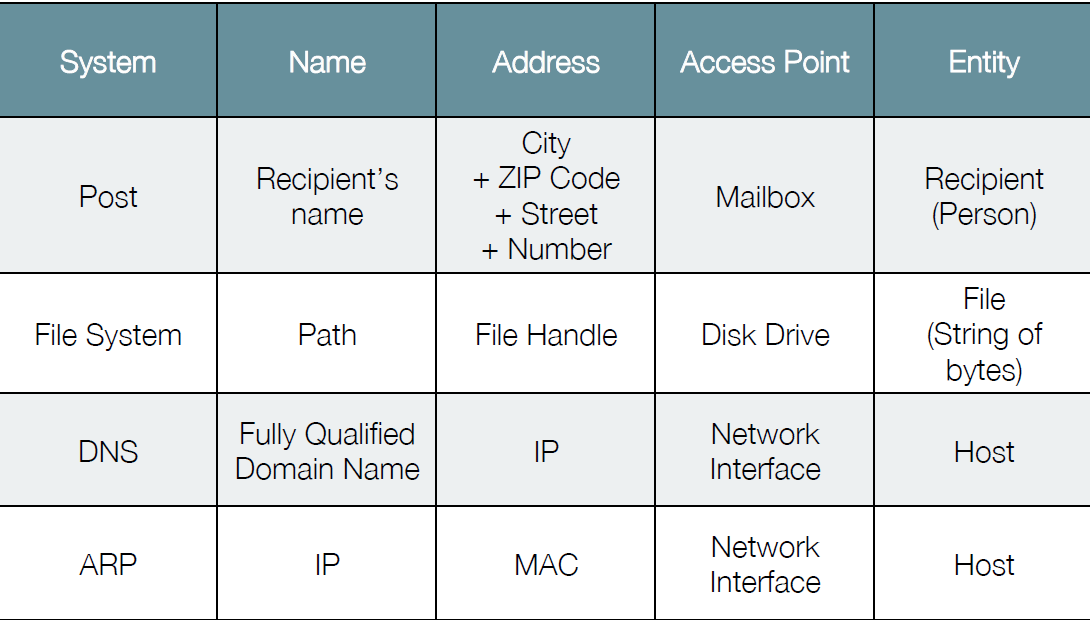
**Entity**: Ein Objekt das adressiert werden möchte.

**Name**: Bytestring als Referenz zu einer Entität

**Access Point**: Spezielle Entität die benötigt wird, um auf andere Identitäten zuzugreifen.

**Address**: Name von einem Access Point

**Identifier**: Ein Name der auf höchstens eine Entität referenziert.



## Flat Naming

Ein Name der aus dem Name keine Informationen zum Objekt schliessen lässt. Nicht Strukturiert. Eine Broadcast Domäne ist ein gutes Beispiel für eine Flat Naming Umgebung. Jede Entität hat eine Adresse, und fragt: «Was ist die Adresse zum Namen xyz?»

## Structural naming

Namen sind strukturiert wie z.B. ein DNS Name. Hohe Ansprüche an einen Namensserver. Kann gut gecached werden.

Interative resolution: Wenn der DNS Server den angefragten Host nicht kennt, setzt er interactive Request beginnend zuhinterst vom DNS Name (Bsp. .com) ab um die Anfrage vom Client zu beantworten

Recursive Resolution: Anfrage vom Client an den DNS Server

## Attribute-based Naming

LDAP (mix aus Strukturbased und Attribute Based naming)

## Bojour/Zeroconf

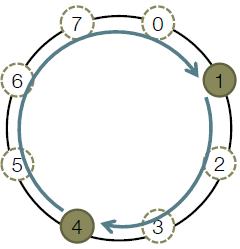
Protokolle, die mehrere Naming Standards verwenden z.B.: DHCP. Funktioniert nur Lokal, alle müssen dem DHCP Server vertrauen.

# Distributed Hash Table (DHT)

## Chord

P2P Lookup Service, Verbindet P2P Standorte im Netzwerk. Implementiert Consistent Hashing. Mappt Keys in einem P2P Netzwerk zu Locations. Ermöglicht das Suchen im P2P Netzwerk.

Aufbau: p ist der Hash der Host IP Adresse, jeder Node hat die Adresse für seinen Nachfolger und Vorgänger Node:

**Chord Lookup**

Geschwindigkeit O(log n), n = Anzahl Nodes im Netzwerk. O(log^2 n) Messages sind maximal notwendig um einen key zu finden.

**Node Adresse:** SHA1(IP + Port) -> die ersten m Zeichen davon. m wird definiert.

**Finger Table**: Beinhaltet alle direkten Nodes für jeden Client. Nämlich seine id + 2^(i – 1) (i = Laufvariable 1 bis m). Die Finger Table hat m Einträge.

**Files Stored:** Mit dem Key vom File (Bsp: Filename) wird durch hashen: SHA1(key) einen Schlüssel generiert, welcher beim Node, der >= sein Schlüssel ist, abgelegt.

**Ausfallsicherheit**: Die Ausfallsicherheit von einem chord P2P Netz wird dadurch gewährleistet, dass alle Dateien nicht nur auf dem dafür vorgesehenen Host abgelegt sind, sondern auch noch auf dem nächsten vorherigen Host mit 1 – (1/n^2) nahezu Unendlich.

**Stabilisierungsprotokoll:** Stabilisierungsprotokoll aktualisiert seinen Vorgänger und seinen Nachfolger. Jeder Knoten stabilisiert sich periodisch unabhängig und prüft ob sein Nachfolger und Vorgänger noch verfügbar sind.

## Lookup

Lookup, also ein P2P Netzwerk durchsuchen, ist eine grosse Herausforderung für ein P2P Netzwerk. Es gibt verschiedene Varianten:

**Broadcast eine Suche an alle Peers** (immer an die bekannten Peers weitergegeben): Problem hohe Last mit Suchanfragen.

**Zentraler Server, der ein Verzeichnis mit allen Standorten führt**: Bottleneck, 1. Angriffspunkt.

Suchbefehle wie DNS organisieren: Hohe Last auf «Root Server» Ausfall von «Root Server» führt zu einem Zusammenbruch

**Symmetrische Suche**: Suche von Node zu Node bis ein Suchergebnis gefunden worden ist: Anonymität nicht gewährleistet da Standort der Daten bekannt.

**DHT als Suche:** Verschiedene Algorithmen implementieren die Distributed Hash Table Suche. Das Grundprinzip ist jedoch immer das Gleiche. Es wird eine lokale Tabelle mit ausgesuchten Nodes geführt.

# Synchronization

Wir können die Reale Zeit nicht "genau" messen. Millisekunden sind schon schwierig, Nanosekunden unmöglich.

## beat

Anstatt von Stunden und Minuten wurde ein Tag in 1000 Stücke aufgeteilt. Ein Beat dauert 86.4 Sekunden. Es gibt keine Zeitzonen BMT (Biel Meantime) basiert auf dem Hauptstandort von Swatch.

## Logical vs. Physical clock

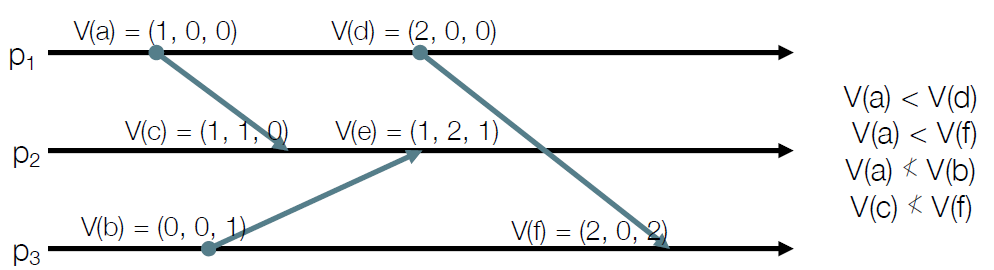
Die Physische Zeit ist die für uns interessante Zeit. Die Logische Zeit kann eine aufzählende Nummerierung sein, welche für eine gewisse Applikation die Reihenfolge von Transaktionen bestimmt. Die Logische Zeit ist nicht an die Physikalische Zeit gebunden.

## Lamport’s logical clock

Als Uhr wird ein Counter eingesetzt, welcher bei jedem Event um 1 erhöht wird. Der Counter muss immer > sein als der erhaltene Counter, wenn das nicht der Fall ist, muss sein Counter überschrieben werden. Pro Prozess oder Host wird noch eine eindeutige ID mitgeschickt damit auch zeitgleich und mit dem gleichen Counter ereignete Events unterschieden werden können.

## Vector clocks

Erweiterung von Lamport’s Uhr. Jeder Prozess/Host hat seine eigene Uhr und schickt jeweils einen Vektor mit allen Uhren an die anderen Prozesse/Hosts weiter. Inkrementiert wird nur die eigene Uhr, jedoch jeweils die anderen Uhren werden aktualisiert. Es können nebenläufige Events erkannt werden.



## Kausalität

Kausalität bezeichnet die Beziehung zwischen Ursache und Wirkung: Aktion und Reaktion.

**Transitive**: (a 🡪 b ^ b 🡪 c) => a 🡪 c [wenn a vor b, b vor c dann ist auch a vor c]

**Irreflexiv**: a ↛ a [kein Element in der Relation steht zu sich selbst]

**Asymmetrisch**: a 🡪 b => b ↛ a [wenn a vor b gilt aber nicht b vor a]

**Concurrency** (gleichzeitig/nebenläufig/Parallelität): a || b -> Heisst 2 Berechnungen können parallel ablaufen, möglicherweise völlig unabhängig, können sich jedoch integrieren, sprich Resultate austauschen. Es kann, muss jedoch nicht parallel ausgeführt werden.

# Map-Reduce / Spark

**Prinzip**: Map-Reduce ist ein Algorithmus, welcher speziell für die Auswertung von grossen und vor allem auch verteilten Datenmengen geeignet ist.

Map-Reduce bzw. Spark kann eingesetzt werden um über einen Cluster Logfiles zu analysieren und die Werte daraus zu entnehmen. Ein Node um grosse Logs zu verarbeiten sollte folgende Eigenschafen haben: angemessene Reaktion, Resistent (bsp. Bei Ausfall von einem Node) Skalierbar. **Nachteile**: Einfachere Jobs können teilweise bis zu 1 Seite Code beanspruchen, erhebliche Disk I/O

## RDD

Ein RDD ist eine Collection von Items des Types A welche folgende Eigenschaften hat:

**Immutable**: Änderungen führen zwingend in neue RDDs

**Lazy**: Operationen werden on-demand ausgeführt.

**Distributed**: Verteilt über einem Set von Nodes. Jeder RDD besteht aus mehreren Partitionen.

**Cacheable**: Zwischenresultate können im RAM oder auf Disk cached werden.

**Resilient**: Eine Partition kann neu erstellt werden, falls die Daten verloren gehen, denn jeder RDD enthält Informationen für die Wiederherstellung von verlorenen Partitionen. Eine Partition ist immer nur auf einem Node gespeichert.

**Sonstiges:** Deklarative API zum Beschreiben von Schritten auf einer Collection. Code wird auf irgendeinem Node ausgeführt. Cache und Wiederholung muss beachtet werden. Vorteile zu Parallel Stream: Schneller, Abbrechbar, elastisch

**Shuffling**: Nodes auf der Mapper Seite, speichern Values in Key Buckets. Nodes auf der Reducer Seite holen zugewiesene Buckets ab. Nach Daumenregel kann gesagt werden, dass wenn immer ein Resultat einer Transformation mehr als ein Record vom RDD beinhaltet geshuffled wird.

**Ausführung von RDDs**: 1. Aus RDDs ein DAG machen, 2. Logischen Ausführungsplan erstellen, 3. Stages in geplante Tasks aufteilen, 4. Tasks ausführen

# Peer-to-Peer Network – (bitTorrent / torrent)

Peer to Peer Netzwerk wächst kontinuierlich ca. 50% pro Jahr.

**Ziele:** Mehr Performance und Verfügbarkeit

**Tracker:** Hilft den Client einander zu finden. Koordiniert das Netzwerk über Port 80 UDP/http

**Client**: Kommuniziert mit Trackers. Wird «Leecher» genannt solange noch Teile des Files fehlen. Sobald das File vollständig ist, wird der Client «Seeder» genannt. Läuft normalerweise auf einem TCP Port im Range 6881 – 6889.

**Web Server:** Publizieren von .torrent Files

**Torrent File:** Beinhaltet: Tracker URL, Infos (als Directory): Name, Piece Länge (in Bytes), Hashlist mit Hashes aller Files, Länge (fall nur ein einzelnes File enthalten ist) oder ein Files Directory mit Pfad und Länge pro File.

**Fairness/ Choking**: Wer mehr hoch lädt wird nicht gedrosselt. Für neue Peers gibt es den Unchoked Algorithmus um den ersten Download nicht rauszuzögern. Englischer Begriff: **tit-for-tat**

**Seeders:** vollständige Downloaders (Bezogen auf einem File) **Leechers**: unvollständige Downloaders.

**Announce**: URL des Trackers **Path**: Eine Liste mit Teilen eines Pfades, Teil vom Files-Feld.

**Chunk**: Die Files werden in Chunks/Pieces aufgeteilt. Muss eine Grösse von 2x Bytes haben. Meist ist dies 218, also 256k. Der allerletzte Chunk kann von der Grösse der anderen Chunks abweichen.

Es wird ein SHA-1 Hash (20 Bytes) für jeden Chunk errechnet und dem Chunk hinzugefügt.

**Piece Selection**: Es gibt mehrere Möglichkeiten um Pieces auszuwählen:

**- Strict Priority**: «Unterstücke» vom eigentlichen Piece werden priorisiert.

**- Rarest First**: Seltenste Pieces zuerst, damit diese schnell wieder distributed werden können.

**- Random First Piece**: Hole ein erstes Piece schnell, damit man schnell bereits ist es wieder hochzuladen.

**- Endgame Mode**: Die noch verbleibenden Pieces von mehreren Peers anfordern.

## File

Eine Datei wird in 256KB grosse Blöcke aufgeteilt, welche jeweils mit einer SHA-1 Hash verifiziert werden können. Es werden jeweils 16Kb Grösse Paketstücke priorisiert: 1. Selten vorhandene Stücke, 2. Erste Stücke zuerst damit auch Upload möglich ist. 3. Ausstehende Teile werden von mehreren Clients angefordert.

## Ablauf

1. .torrent File von einem lokalen File erstellen.

2. ... selber im BitTorrent Client öffnen (lokales File im Zugriff)

3. ... auf einem Web Server veröffentlichen.

4. ... von einem Web Server runterladen.

5. ... im BitTorrent Client öffnen.

6. Download von den Peers starten (Upload startet zu gleichen Zeit automatisch)

# Stichwortverzeichnis

|  |  |
| --- | --- |
| ACID 1.4  AMQP 4.4  Asymmetrisch 15.5  Asynchronous Messaging 4  Availability 7.3, 7  beat 15.1  Bitcoin 12.1  bitTorrent 17  Bojour 13.4  Capacity 6.3  CFIA 9  CFIA Matrix 9.1  Channel Purger 2.6  Chirp 10.3  Chord 14.1  CloudSum 10.1  Concurrency 15.5  Contingency 6.3  CORBA 1.5  Deployment Patterns 2.2  Deterministisch 1.1  Detour 2.6  DHT 14  Enterprise and integration Patterns 2  Fan-Out 2.7, 2.3  FCAPS 8.2  Flat Naming 13.1  Gatling 10.3  Hub-And-Spoke 2.1  Idempotent 1.6  IDL 5.5, 1.5  inproc 10.2  Interprozesskommunikation 10  Intraprozesskommunikation 10  IP Spraying 7.1  Irreflexiv 15.5  ITIL 8.1  Java Socket API 3.4  JMS 4.3  JMX 8.3  Kausalität 15.5  Lamport’s logical clock 15.3  Load balancing 7.1  Logback 10.3  Logging 8.4  Logical vs. Physical clock 15.2 | Map-Reduce 16  marshalling 5.1  Marshalling 5.6  Message Store 2.6  Messaging Patterns 2.3  Middleware 1.3  MPI 3.2  MTBF 7.3  MTTF 7.3  MTTR 7.3  Naming 13  Netcat 3.3  NFR 7.3  NFRs 6.1  Peer-to-Peer Network 17  Performance 6.2  Presisten vs non Presistent 4.1  RabbitMQ 4.4  RDD 16.1  Redis 10.3  RESP 10.3  Response Time 6.3  Risk Rating Tabelle 9  RMI 5  RPC 5.1  RPO 7.3  RTO 7.3  Scalability 7.5, 7  Service Activator 3.6  Slf4j 10.3  Smart Proxy 2.6  Socket 3.1  Spark 16  Structural naming 13.2  Stub 5.2  Synchronization 15  Throughput 6.3  torrent 17  Torrent File 17  Transitive 15.5  Unmarshalling 5.1  Vector clocks 15.4  Verfügbarkeit berechnen 7.4  WCF 2.7  WebSockets 3.5  Wire Tap 2.6  WSDL 5.6  Zeroconf 13.4  ZeroMQ 10.2, 2.7 |