

# Verteilte Systeme und Komponenten

Zusammenfassung Frühlingssemester 2018

Patrick Bucher

05.06.2018

## Inhaltsverzeichnis

<b>1</b>	<b>Komponenten</b>	<b>1</b>
1.1	Begriffe und Architekturen . . . . .	1
1.1.1	Der Komponentenbegriff . . . . .	1
1.1.2	Der Nutzen von Komponenten . . . . .	2
1.1.3	Der Entwurf mit Komponenten . . . . .	3
1.1.4	Komponenten in Java . . . . .	3
1.2	Schnittstellen . . . . .	4
1.2.1	Begriff und Konzept . . . . .	4
1.2.2	Dienstleistungsperspektive . . . . .	5
1.2.3	Spezifikation von Schnittstellen . . . . .	5
1.3	Modularisierung . . . . .	6
1.3.1	Modulkonzept . . . . .	6
1.3.2	Layers, Tiers & Packages . . . . .	8
<b>2</b>	<b>Entwicklungsprozess</b>	<b>9</b>
2.1	Projektplanung . . . . .	9
2.2	Source-Code-Management, Build und Dependency-Management . . . . .	9
2.3	Build-Server . . . . .	9
2.4	Integrations- und Systemtesting . . . . .	9
2.5	Entwurfsmuster . . . . .	9
2.6	Testing . . . . .	9
2.7	Continuous Integration . . . . .	9
2.8	Review . . . . .	9
2.9	Konfigurationsmanagement . . . . .	9
2.10	Deployment . . . . .	9
2.11	Code-Qualität . . . . .	9
2.11.1	Kommentare . . . . .	9
2.11.2	Namensgebung . . . . .	10
2.11.3	Weitere Massnahmen . . . . .	11

<b>3 Verteilte Systeme</b>	<b>11</b>
3.1 Socket-Kommunikation . . . . .	11
3.2 Serialisierung . . . . .	11
3.3 Message-Passing . . . . .	11
3.4 Verteilung & Kommunikation: RMI . . . . .	11
3.5 Uhrensynchronisation . . . . .	11
3.6 Verteilung: Data Grid . . . . .	11

## 1 Komponenten

Herkunft: *componere* (lat.) = zusammensetzen

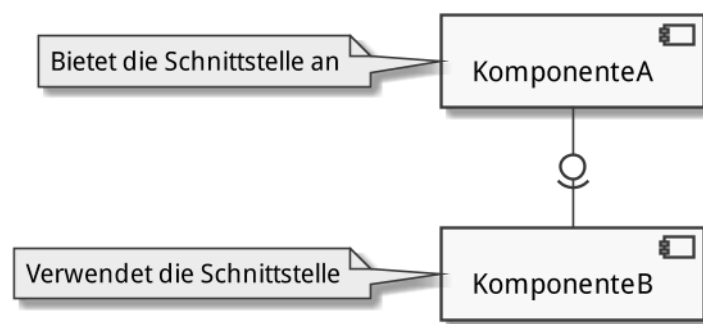


Abbildung 1: Komponentendiagramm (UML2)

### 1.1 Begriffe und Architekturen

#### 1.1.1 Der Komponentenbegrif

- Definition: Eine Software-Komponente
  1. ist ein Software-Element
  2. passt zu einem bestimmten Komponentenmodell
  3. folgt einem bestimmten *Composition Standard*
  4. kann ohne Änderungen mit anderen Komponenten verknüpft und ausgeführt werden
- Eigenschaften: Software-Komponenten
  1. sind eigenständig ausführbare Softwareeinheiten
  2. sind über ihre Schnittstellen austauschbar definiert
  3. lassen sich unabhängig voneinander entwickeln
  4. können kunden- und anwendungsspezifisch oder anwendungsneutral und wiederverwendbar sein
    - COTS (Commercial off-the-shelf): Software «von der Stange»
  5. können installiert und deployed werden

- 6. können hierarchisch verschachtelt sein
- Komponentenmodelle
  - sind konkrete Ausprägungen des Paradigmas der komponentenbasierten Entwicklung
  - definieren die genaue Form und Eigenschaften einer Komponente
  - definieren einen *Interaction Standard*
    - \* wie können die Komponenten miteinander über Schnittstellen kommunizieren (Schnittstellenstandard)
    - \* wie werden die Abhängigkeiten der Komponenten voneinander festgelegt
      - von der Komponente verlangte Abhängigkeiten: *Required Interfaces*
      - von der Komponente angebotene Abhängigkeiten: *Provided Interfaces*
  - definieren einen *Composition Standard*
    - \* wie werden die Komponenten zu grösseren Einheiten zusammengefügt
    - \* wie werden die Komponenten ausgeliefert (Deployment)
- Beispiele verbreiteter Komponentenmodelle:
  - Microsoft .NET
  - EJB (Enterprise Java Beans)
  - OSGi (Open Services Gateway Initiative)
  - CORBA (Common Object Request Broker Architecture)
  - DCOM (Distributed Component Object Model)

### 1.1.2 Der Nutzen von Komponenten

- Packaging: *Reuse Benefits*
  - Komplexität durch Aufteilung reduzieren (*Divide and Conquer*)
  - Wiederverwendung statt Eigenentwicklung spart Entwicklungszeit und Testaufwand
  - erhöhte Konsistenz durch Verwendung von Standardkomponenten
  - Möglichkeit zur Verwendung bestmöglicher Komponente auf dem Markt
- Service: *Interface Benefits*
  - erhöhte Produktivität durch Zusammenfügen bestehender Komponenten
  - erhöhte Qualität aufgrund präziser Spezifikationen und vorgetesteter Software
- Integrity: *Replacement Benefits*
  - erweiterbare Spezifikation durch inkrementelle Entwicklung und inkrementelles Testing
  - parallele und verteilte Entwicklung durch präzise Spezifizierung und Abhängigkeitsverwaltung
  - Kapselung begrenzt Auswirkungen von Änderungen und verbessert so wie Wartbarkeit

### 1.1.3 Der Entwurf mit Komponenten

- Komponentenbasierte Entwicklung
  - steigende Komplexität von Systemen, Protokollen und Anwendungsszenarien
  - Eigenentwicklung wegen Wirtschaftlichkeit und Sicherheit nicht ratsam

- Konstruktion von Software aus bestehenden Komponenten immer wichtiger
- Anforderungen (aufgrund mehrmaliger Anwendung) an Komponenten höher als an reguläre Software
- Praktische Eigenschaften
  - Einsatz einer Komponente erfordert nur Kenntnisse deren Schnittstelle
  - Komponenten mit gleicher Schnittstelle lassen sich gegeneinander austauschen
  - Komponententests sind Blackbox-Tests
  - Komponenten lassen sich unabhängig voneinander entwickeln
  - Komponenten fördern die Wiederverwendbarkeit
- Komponentenspezifikation
  - Export: angebotene/unterstützte Interfaces, die von anderen Komponenten genutzt werden können
  - Import: benötigte/verwendete Interfaces von anderen Komponenten
  - Kontext: Rahmenbedingungen für den Betrieb der Komponente
  - Verhalten der Komponente

#### 1.1.4 Komponenten in Java

- Komponenten in Java SE
  - Komponenten als normale Klassen implementiert
  - Komponenten können, müssen sich aber nicht an die *Java Beans Specification* halten
    - \* Default-Konstruktor
    - \* Setter/Getter
    - \* Serialisierbarkeit
    - \* PropertyChange
    - \* Vetoable
    - \* Introspection
  - Weitergehende Komponentenmodelle in Java EE
    - \* Servlets
    - \* Enterprise Java Beans
- Austauschbarkeit
  - Die Austauschbarkeit von Komponenten wird durch den Einsatz von Schnittstellen erleichtert.
  - Schnittstellen werden als Java-Interface definiert und dokumentiert (JavaDoc).
  - Eine Komponente implementieren eine Schnittstelle als Klasse.
    - \* mehrere, alternative Implementierungen möglich
    - \* Austauschbarkeit über Schnittstellenreferenz möglich
  - Beispiel: API von JDBC (Java Database Connectivity)
    - \* von Sun/Oracle als API definiert
    - \* von vielen Herstellern implementiert (JDBC-Treiber für spezifische Datenbanksysteme)
    - \* Datenbanktausch auf Basis von JDBC möglich
- Deployment

- über .jar-Dateien (Java Archive): gezippte Verzeichnisstrukturen bestehend aus
  - \* kompilierten Klassen und Interfaces als .class-Dateien
  - \* Metadaten in META-INF/manifest.mf
  - \* optional weitere Ressourcen (z.B. Grafiken, Textdateien)
- Deployment von Schnittstelle und Implementierung zum einfacheren Austausch häufig in getrennten .jar-Dateien mit Versionierung, Beispiel (fiktiv):
  - \* jdbc-api-4.2.1.jar enthält die Schnittstelle
  - \* jdbc-mysql-3.2.1.jar enthält die MySQL-Implementierung
  - \* jdbc-postgres-4.5.7.jar enthält die PostgreSQL-Implementierung
  - \* Versionierung idealerweise im Manifest und im Dateinamen (Konsistenz beachten!)

## 1.2 Schnittstellen

### 1.2.1 Begriff und Konzept

- Der Begriff *Schnittstelle* als Metapher
  - Beim Zerschneiden eines Apfels entstehen zwei spiegelsymmetrische Oberflächen.
  - Die Komponenten müssen so definiert werden, damit sie an der Schnittstelle zusammenpassen, als ob sie vorher auseinandergeschnitten worden wären.
  - Tatsächlich werden *Verbindungsstellen* erstellt, welche Kombinierbarkeit sicherstellen.
  - Eine Schnittstelle tut nichts und kann nichts.
  - Schnittstellen trennen nichts, sie verbinden etwas:
    - \* Komponenten untereinander (Programmschnittstellen)
    - \* Komponenten mit dem Benutzer
- Die Bedeutung von Schnittstellen (bei korrektem Gebrauch):
  1. machen Software leichter verständlich (man braucht nur die Schnittstelle und nicht die Implementierung zu kennen)
  2. helfen uns Abhängigkeiten zu reduzieren (Abhängigkeit nur von einer Schnittstelle, nicht von einer Implementierung)
  3. erleichtern die Wiederverwendbarkeit (bei der Verwendung bewährter Schnittstellen statt Eigenentwicklung)
- Die Beziehung zwischen Schnittstellen und Architektur:
  - System > Summe seiner Teile (Beziehungen zwischen den Teilen: durch Schnittstellen ermöglicht)
    - \* Schnittstellen & Beziehungen zwischen den Komponenten: wichtigste Architektur-aspekte!
    - \* Mehrwert des Systems gegenüber Einzelkomponenten liegt in den Schnittstellen & Beziehungen der Komponenten zueinander
  - Spezialisten für Teilsysteme konzentrieren sich auf ihr Teilproblem
    - \* Architekten halten das Gesamtsystem über Schnittstellen zusammen
    - \* Schnittstellen verbinden ein System mit der Aussenwelt und ermöglichen die Interaktion damit

- Kriterien für gute Schnittstellen
  1. Schnittstellen sollen *minimal* sein:
    - wenige Methoden (mit möglichst geringen Überschneidungen in ihren Aufgaben)
    - geringe Anzahl von Parameters
    - setzen möglichst keine oder nur wenige globale Daten voraus
  2. Schnittstellen sollen *einfach zu verstehen* sein
  3. Schnittstellen sollen *gut dokumentiert* sein

### 1.2.2 Dienstleistungsperspektive

- Die Schnittstelle als Vertrag:
  - Ein *Service Consumer* schliesst einen Vertrag mit einem *Service Provider* für eine *Dienstleistung* ab
- Design by Contract (DbC): Das Zusammenspiel zwischen den Komponenten wird mit einem Vertrag geregelt
  - *Preconditions*: Zusicherungen, die der Aufrufer einhalten muss
    - \* Nutzer: Prüfen der Vorbedingungen vor der Ausführung
    - \* Anbieter: Überprüfung mittels Assertions
  - *Postconditions*: Nachbedingungen, die der Aufgerufene garantiert
    - \* Nutzer: Überprüfung mittels Assertions
    - \* Anbieter: Prüfen der Nachbedingungen nach der Ausführung
  - *Invarianten*: Über alle Instanzen einer Klasse geltende Grundannahmen ab deren Erzeugung
    - \* Anbieter: Überprüfung mittels Assertions

### 1.2.3 Spezifikation von Schnittstellen

- Dokumentation von Schnittstellen
  - Umfang:
    - \* was ist wichtig für die Benutzung der Komponente
    - \* was muss der Programmierer verstehen und beachten
  - Eigenschaften der Methoden:
    - \* Syntax (Rückgabewerte, Argumente, Typen, call by value/reference)
    - \* Semantik (was bewirkt die Methode)
    - \* Protokoll (synchron/asynchron)
    - \* Nichtfunktionale Eigenschaften (Performance, Robustheit, Verfügbarkeit)
  - Schnittstellen an der Systemgrenze fliessen in die Systemspezifikation ein
- öffentliche Schnittstellen werden als *API* bezeichnet (Application Programming Interface)
  - objektorientierte API (sprachabhängig, z.B. API der JSE)
  - REST-API (Representational State Transfer, sprach- und plattformunabhängig, datenzentriert)

- Messaging-API (sprach- und plattformunabhängig, z.B. Push-Notifications für Mobile Apps)
- dateibasierte API (Informationsaustausch, Konfigurationsdateien)

### 1.3 Modularisierung

Modul: in sich abgeschlossener Teil des Programmcodes, bestehend aus Verarbeitungsschritten und Datenstrukturen

#### 1.3.1 Modulkonzept

- Kopplung und Kohäsion
  - Kopplung: Ausmass der Kommunikation zwischen Modulen
    - \* hohe Kopplung: grosse Abhängigkeit
    - \* Kopplung minimieren!
  - Kohäsion: Ausmass der Kommunikation innerhalb eines Moduls
    - \* geringe Kohäsion: geringer Zusammenhalt
    - \* Kohäsion maximieren!
  - Viele Module: Hohe Kopplung, geringe Kohäsion
  - Wenige Module: Geringe Kopplung, hohe Kohäsion
  - Idealer Kompromiss: Reduziert Gesamtkomplexität
- Arten von Modulen
  - Bibliothek: Sammlung oft verwendeter, thematisch zusammengehörender Funktionen (Datumsmodul, Mathematik-Modul, I/O-Modul)
  - Abstrakte Datentypen: Implementierung eines neuen Datentyps mit definierten Operationen (verkettete Liste, binärer Baum Hash-Tabelle)
  - Physische Systeme: Abgegrenztes Hardware-Modul (Ultraschallsensor, Anzeigemodul, Kommunikationsmodul)
  - Logisch-konzeptionelles System: Modellierung von Funktionalität auf hoher Abstraktionsstufe (Datenbankmodul, Bildverarbeitungsmodul, GUI-Framework)
- Entwurfskriterien
  - Zerlegbarkeit (*modular decomposability*): Teilprobleme können unabhängig voneinander gelöst werden
    - \* *Divide and Conquer*: Softwareproblem in weniger komplexe Teilprobleme zerlegen, sodass sie unabhängig voneinander bearbeitet werden können
    - \* Rekursive Zerlegung: Weitere Zerlegung von Teilproblemen
  - Kombinierbarkeit (*modular composability*): Module sind unabhängig voneinander wiederverwendbar
    - \* Module sollten möglichst frei kombinierbar sein und sich auch in anderen Umfeldern wieder einsetzen lassen
    - \* Zerlegbarkeit und Kombinierbarkeit sind unabhängig voneinander
  - Verständlichkeit: Module sind unabhängig voneinander verständlich

- \* Der Code eines Moduls soll ohne Kenntnis anderer Module verstehbar sein
  - \* Module müssen unabhängig voneinander versteh- und wartbar sein
- Stetigkeit: Änderungen der Spezifikation proportional zu Codeänderungen
  - \* Anforderungen können sich ändern, sollten sich aber nur auf ein Teilsystem auswirken
- Entwurfsprinzipien
  - lose Kopplung: schlanke Schnittstellen, Austausch nur des Nötigsten
  - starke Kohäsion: hoher Zusammenhalt innerhalb des Moduls
  - Geheimnisprinzip (*information hiding*): Modul nach aussen nur über dessen Schnittstellen bekannt
  - wenige Schnittstellen: zentrale Struktur mit minimaler Anzahl Schnittstellen
  - explizite Schnittstellen: Aufrufe und gemeinsam genutzte Daten sind im Code ersichtlich
- Vorgehen bei Modularisierung
  - Basiskonzepte: Kopplung & Kohäsion
  - Kriterien: Verständlichkeit, Kombinierbarkeit, Zerlegbarkeit, Stetigkeit
  - Modularten: Bibliotheken, abstrakte Datentypen, physische und logische Systeme
  - Prinzipien: geringe Kopplung, hohe Kohäsion, Geheimnisprinzip, wenige & explizite Schnittstellen
  - sinnvolle Modularisierung: eine der anspruchsvollsten Aufgaben der Informatik
- Parnas: *On the Criteria to be Used in Decomposing Systems into Modules* (1972)
  - Ziele der Modularisierung:
    1. Die Flexibilität und Verständlichkeit eines Systems verbessern
    2. Die Entwicklungszeit eines Systems reduzieren
  - Voraussetzung für modulares Programmieren:
    1. Ein Modul kann mit wenig Kenntnis des Codes eines anderen Moduls geschrieben werden.
    2. Module können neu zusammengesetzt und ersetzt werden, ohne dass das ganze System neu zusammengesetzt werden muss.
  - Nutzen der Modularisierung:
    - \* Verkürzung der Entwicklungszeit, da mehrere Teams gleichzeitig an je einem Modul arbeiten können und nur wenig Kommunikation zwischen ihnen nötig ist.
    - \* Erhöhte Flexibilität, da grössere Änderungen an einem Modul keine Änderungen in anderen Modulen zur Folge haben.
    - \* Bessere Verständlichkeit, da ein System nicht als ganzes, sondern Modul für Modul analysiert werden kann.
  - Ansätze der Modularisierung:
    1. *Flowchart-Analyse*: Jeder grosse Verarbeitungsschritt wird als Modul implementiert (konventionell).
    2. *Information Hiding*: Jede Design-Entscheidung wird in einem Modul versteckt (neuer Ansatz).
  - Interpretation:
    - \* Mit dem traditionellen Ansatz (*Flowchart-Analyse*) wird ein *Algorithmus* in einzelne Verarbeitungsschritte zerlegt.
    - \* Mit dem neuen Ansatz (*Information Hiding*) werden die *Datenstrukturen* herausge-



arbeitet. (Datenstruktur = Design-Entscheidung)

- \* Die einzelnen Schritte eines Algorithmus sind *nicht* beliebig austauschbar.
- \* Datenstrukturen können abstrahiert und über ein einfaches Interface angeboten werden.

### 1.3.2 Layers, Tiers & Packages

- Layer
  - öffentliche Methoden eines tieferstehenden Layers B dürfen vom höherstehenden Layer A genutzt werden
  - Beispiel (Layers von oben nach unten): A B C
    - \* richtig: A -> B, B -> C
    - \* zulässig: A -> C (gefährlich: Umgehung einer API)
    - \* falsch: C -> B, B -> A, C -> A (von unten nach oben)
    - \* falsch: A -> B -> C -> A (zyklische Abhängigkeit)
  - *call*-Beziehung: ein höherstehender Layer verwendet Funktionalität eines tieferstehenden Layers
  - *use*-Beziehung: korrektes Verhalten von Layer A hängt von der korrekten Implementierung des Layers B ab (initialisiertes Device, aufgenommene Netzwerkverbindung, erstellte Datei)
- Tier: oft mit Layern verwechselt
  - Presentation Tier
  - Business Logic (Tier)
  - Data Tier
- Packages: Implementierung des Layer-Konzepts
  - abstrakt: UML
  - konkret: Java-Package

## 2 Entwicklungsprozess

### 2.1 Projektplanung

### 2.2 Source-Code-Management, Build und Dependency-Management

### 2.3 Build-Server

### 2.4 Integrations- und Systemtesting

### 2.5 Entwurfsmuster

### 2.6 Testing

### 2.7 Continuous Integration

### 2.8 Review

### 2.9 Konfigurationsmanagement

### 2.10 Deployment

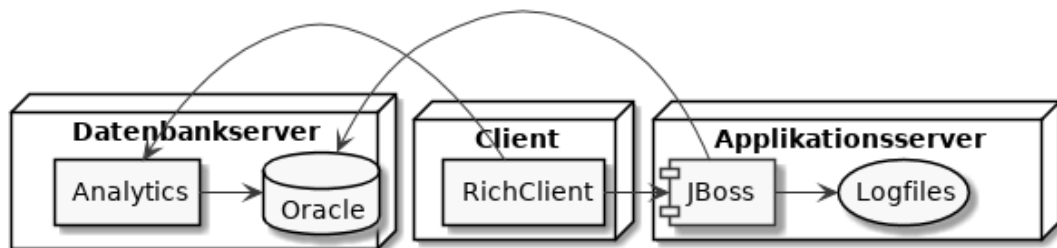


Abbildung 2: Deploymentdiagramm

### 2.11 Code-Qualität

#### 2.11.1 Kommentare

- Schlechter Code soll nicht kommentiert sondern umgeschrieben werden.
  - Kommentare dienen oft als Ausrede für schlechten Code.
  - Selbsterklärender Code ist besser als jeder Kommentar.
  - Gute Namen ersparen viele Kommentare.
- Kommentare sind kein Qualitätsmerkmal sondern ein (manchmal) notwendiges Übel.

- Code wird geändert, ohne dass die Kommentare nachgeführt werden.
- Die Wahrheit liegt im Code, Kommentare können «lügen».
- Akzeptable Kommentare
  - Copyright (Lizenzbedingungen)
  - temporäre TODO/FIXME-Einträge
  - Kommentare zur Hervorhebung unauffälliger aber wichtiger Dinge
  - zusätzliche Erklärungen zur Absicht des Codes
  - Warnungen
- Schlechte Kommentare
  - Redundante Kommentare: `file.save(); // save the file`
  - Irreführende Kommentare: `extension = ".xml"; // txt file`
  - erzwungene Kommentare: JavaDoc für Getter- und Setter-Methoden
  - Tagebuch- und Changelog-Kommentare: Versionskontrolle verwenden
  - Positionsbezeichner und Banner: `// 3) now save the changes`
  - Zuschreibungen und Nebenbemerkungen: `created by XYZ (xyz@foo.com)`
  - Auskommentierter Code: Kann dank Versionskontrolle gelöscht werden
  - HTML-formatierte Kommentare: Im Code schlecht lesbar
  - Zu viele und unnötige Informationen: `// written on my balcony at 5:00 am`

### 2.11.2 Namensgebung

- Namen von Klassen sollten gut überlegt sein, da sie ihn lange behalten.
  - Noch wichtiger bei Interfaces!
- Kriterien für gute Namen:
  - zweckbeschreibend (schlecht: `ThingDoer`, gut: `TreeWalker`)
  - korrekt und ohne Fehlinformationen (schlecht: `ConnectionSingleton` wenn gar kein Singleton implementiert wird)
  - differenzierend
    - \* schlecht: `int a; float b; char c;`
    - \* gut: `int divisor; float dividend; char decimalSeparator;`
  - gut aussprechbar und suchbar sein
    - \* schlecht: `int n_pâtés_mangées; String достопримечательность; float 101010101010fAc70r; double schnäderegägs;`
    - \* gut: `int pates_eaten; String attraction; float lolFactor; double babble;`
  - möglichst keine Codierungen enthalten
    - \* schlecht: `int range_0x00_0xff; int clr00ff00;`
    - \* gut: `int rangeFirstByte; int rgbGreen;`
- Heuristiken zur Namensgebung:
  1. Beschreibende Namen wählen
  2. Namen passend zur Abstraktionsebene wählen
  3. Standardnomenklatur verwenden
  4. Eindeutige Namen wählen

5. Namenlänge abhängig von Geltungsbereich
6. Codierungen vermeiden
7. Nebeneffekte in Namen miteinbeziehen

### **2.11.3 Weitere Massnahmen**

- Reviews
  - Effizienteste Methode zur Verbesserung der Code-Qualität
  - Zu Beginn alleine oder in kleinen Teams, später mit mehreren Teilnehmern
  - In vertrauensvoller Atmosphäre, nicht als QS-Massnahme oder in Verbindung mit KPIs
- Weitergabe von Erfahrungen im informellen Austausch
  - Mit Snacks (5 Minuten) als Anreiz, vor/nach Vormittags-/Nachmittagspause
- Bemühungen für sauberen Code als Motivation und zur Erinnerung optisch präsent halten
- Werkzeuge: Checkstyle, PMD, Findbugs, SonarQube etc.
- Pfadfinderregel: Den Platz (Code) sauberer verlassen als man ihn angetroffen hat.

## **3 Verteilte Systeme**

### **3.1 Socket-Kommunikation**

### **3.2 Serialisierung**

### **3.3 Message-Passing**

### **3.4 Verteilung & Kommunikation: RMI**

### **3.5 Uhrensynchronisation**

### **3.6 Verteilung: Data Grid**