Processo e Sviluppo del Software

UniShare

Davide Cozzi
@dlcgold

Indice

Capitolo 1

Introduzione

Questi appunti sono presi a lezione. Per quanto sia stata fatta una revisione è altamente probabile (praticamente certo) che possano contenere errori, sia di stampa che di vero e proprio contenuto. Per eventuali proposte di correzione effettuare una pull request. Link: `https://github.com/dlcgold/Appunti`.
Le immagini presenti in questi appunti sono tratte dalle slides del corso e tutti i diritti delle stesse sono da destinarsi ai docenti del corso stesso.

Capitolo 2

Metodi Agili

I *metodi agili* sono stati definiti per rispondere all'esigenza di dover affrontare lo sviluppo di software in continuo cambiamento. Durante lo sviluppo si hanno vari passaggi:

- comprensione dei prerequisiti
- scoperta di nuovi requisiti o cambiamento dei vecchi

Questa situazione rendeva difficile lo sviluppo secondo il vecchio metodo *waterfall* (a cascata) portando al fallimento di diversi progetti.
I *metodi agili* ammettono che i requisiti cambino in modo "naturale" durante il processo di sviluppo software e per questo assumono un modello di processo *circolare*, con iterazioni della durata di un paio di settimane (figura 2.1). Potenzialmente dopo un'iterazione si può arrivare ad un prodotto che può essere messo "in produzione". Dopo ogni rilascio si raccolgono *feedback* per poter rivalutare i requisiti e migliorare il progetto.
Si hanno quindi aspetti comuni nei metodi agili e nel loro processo:

- enfasi sul team, sulla sua qualità e sulla sua selezione
- il team è *self organizing*, si da importanza ai vari membri del team dato che non esiste un *manager* ma è il team stesso a gestire lo sviluppo
- enfasi al pragmatismo, focalizzandosi su una documentazione efficace evitando di produrre documenti inutili e difficili da mantenere
- enfasi sulla comunicazione diretta, sostituendo i documenti suddetti con meeting e riunioni periodiche

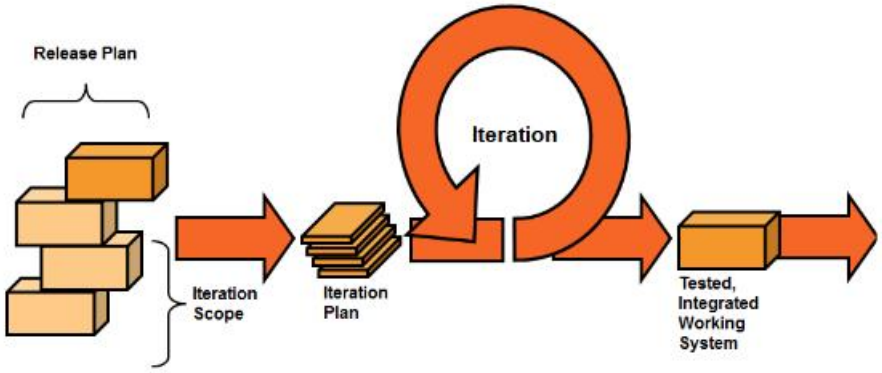

Figura 2.1: Rappresentazione grafica del modello agile. I blocchi a sinistra rappresentano i requisiti (da sviluppare secondo una certa priorità). Lo sviluppo si articola tramite varie iterazioni che porta ogni volta ad aggiungere una parte al prodotto finale (ogni iterazione produce, a partire da un certo requisito, un parte di prodotto di qualità già definitiva, con testing, documentazione etc...)

- enfasi sull'idea che nulla sia definitivo: la perfezione non deve essere seguita fin da subito ma saranno gli step a portare al raggiungimento di una perfezione finale (anche dal punto di vista del design)
- enfasi sul controllo costante della qualità del prodotto, anche tramite
  - *continuous testing* grazie al quale un insieme di test viene eseguito in modo automatico dopo ogni modifica
  - *analisi statica e dinamica* del codice al fine di trovare difetti nello stesso
  - *refactoring*

I metodi agili sono molto "elastici" e permettono la facile definizione di nuovi metodi facilmente adattabili al singolo progetto.

2.1 Scrum

Uno dei più famosi, tra i vari *metodi agili*, è **scrum** (figura 2.2).
In questo caso la parte di sviluppo e iterazione prende il nome di *sprint* ed

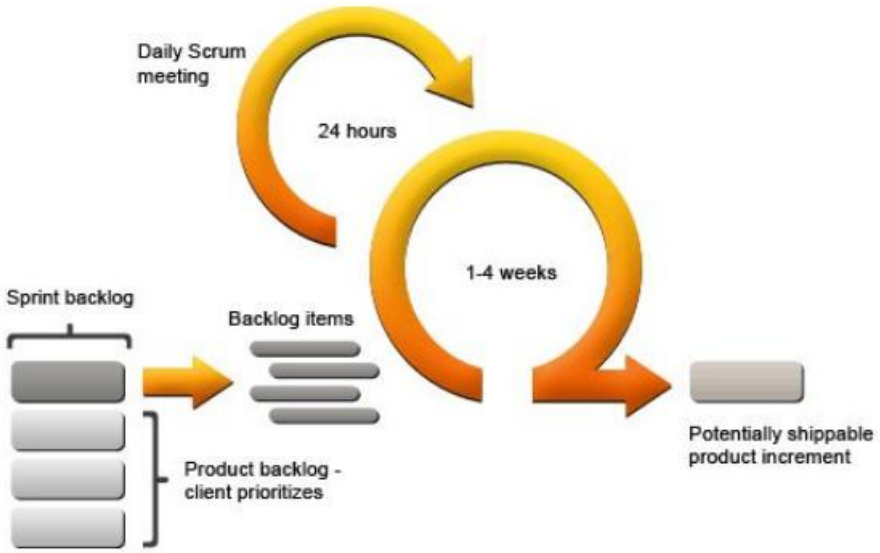

Figura 2.2: Rappresentazione grafica del processo scrum

ha una durata variabile tra una e quattro settimane, per avere un rilascio frequente e una veloce raccolta di feedback. I requisiti sono raccolti nel cosiddetto *product backlog*, con priorità basata sulla base delle indicazioni del committente. Ad ogni *sprint* si estrae dal *product backlog* lo *sprint backlog*, ovvero il requisito (o i requisiti) da implementare nello *sprint*. Lo *sprint backlog* viene analizzato nel dettaglio producendo i vari *backlog items*, ovvero le singole funzionalità che verranno implementate nello *sprint*. Si ottiene quindi di volta in volta un pezzo di prodotto finale, testato e documentato. Durante le settimane di *sprint* si effettua anche un meeting giornaliero utile per mantenere alti i livelli di comunicazione e visibilità dello sviluppo. Durante il meeting ogni sviluppatore risponde a tre domande:

1. Cosa è stato fatto dall'ultimo meeting?
2. Cosa farai fino al prossimo meeting?
3. Quali sono le difficoltà incontrate?

L'ultimo punto permette la cooperazione tra *team members*, consci di cosa ciascuno stia facendo.
Durante il processo *scrum* si hanno quindi tre ruoli:

1. il **product owner**, il committente che partecipa tramite feedback e definizione dei requisiti

2. il **team** che sviluppa
3. lo **scrum master** che controlla la correttezza di svolgimento del processo scrum

Essi collaborano nelle varie fasi:

- product owner e team collaborano nella definzione dei backlog e nella loro selezione ad inizio *sprint*
- durante lo *sprint* lavora solo il team
- nello studio del risultato collaborano tutti coloro che hanno un interesse diretto nel progetto (team, product owner e stakeholders)

Lo scrum master interagisce in ogni fase, fase che viene comunque guidata tramite meeting:

- **sprint planning meeting**, ad inizio *sprint*
- **daily scrum meeting**, il meeting giornaliero
- **sprint review meeting**, in uscita dallo *sprint* per lo studio dei risultati
- **sprint retrospective meeting**, in uscita dallo *sprint* per lo studio tra i membri del team di eventuali migliorie al processo e allo sviluppo del prodotto (anche dal punto di vista delle tecniche e delle tecnologie)

2.2 Extreme Programming

Un altro tipo di metodo agile è l'*extreme programming* (figura 2.3), ormai poco usato. I requisiti prendono i nomi di *stories*, delle narrazioni in cui l'attore (futuro utente del sistema) cerca di svolgere un compito. Vengono scelti quindi *stories* per la prossima iterazione, dove si hanno testing e revisione continua. Le release di ogni iterazione vengono catalogate per importanza (con anche la solita collezione di feedback). In *extreme programming* si hanno davvero molte pratiche:

- The Planning Game
- Short Releases
- Metaphor

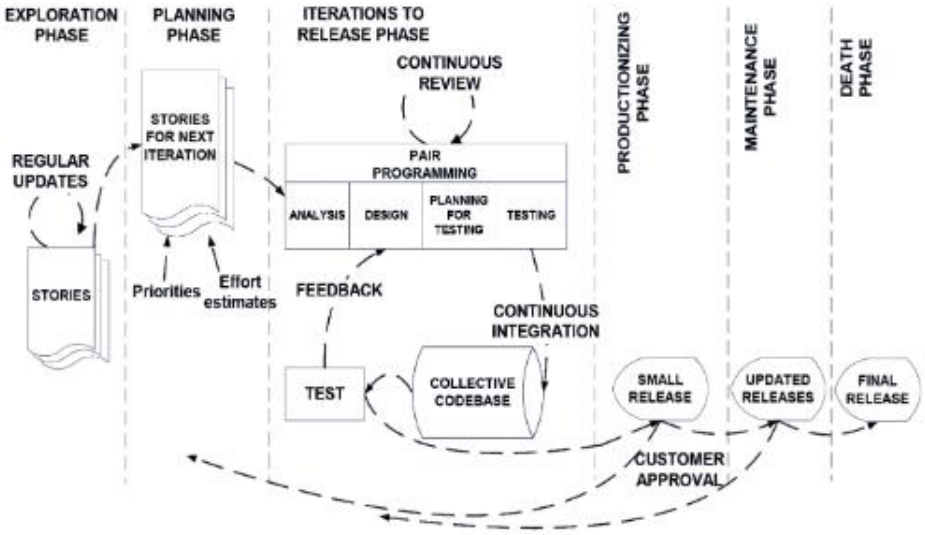

Figura 2.3: Rappresentazione grafica dell'extreme programming

- Simple Design
- Refactoring
- Test-First Design
- Pair Programming
- Collective Ownership (codice condiviso da tutti senza alcun owner, con tutti in grado di effettuare modifiche)
- Continuous Integration
- 40-Hour Week
- On-Site Customer
- Coding Standard (per avere il collective ownership avendo un solo standard di codifica comune a tutti)
- Open workspace

Capitolo 3

DevOps

Nel tempo si è passato dal *modello waterfall* (dove tutto era definito ad inizio progetto) al *modello agile*. Negli ultimi tempi si è sviluppato un altro metodo, chiamato **DevOps**, dove anche la parte di *operation* deve essere agile: il rilascio in produzione e il *deployment* devono essere agili quanto lo sviluppo. Amazon, Netflix, Facebook e molte altre società già adottano le pratiche *DevOps.*
Nel *DevOps* i team di sviluppo e operation sono indipendenti tra loro, diminuendo il costo di impegno necessario al team di sviluppo per la parte di deployment (che viene resa anche più sicura grazie alla diminuzione dell'intervento umano in favore di automazioni). Inoltre, avendo i due team dei tempi di lavoro diversi, si riesce a prevenire ritardi causati dalla non organicità delle operazioni.
DevOps promuove la collaborazione tra di due team al fine di ottenere una sorta di team unico che curi sia sviluppo che operation.
DevOps include quindi diversi processi che vengono automatizzati:

- Continuous Development
- Continuous Integration
- Continuous Testing
- Continuous Deployment (anche con tecnologie di virtualizzazione e con l'uso di *container* in *ambiente cloud*)
- Continuous Monitoring

Con DevOps il feedback arriva in primis dal software (i tool di *monitor*) inoltre il focus viene spostato sui processi automatici.
Nel DevOps si introducono nuovi ruoli:

- il **DevOps Evangelist**, simile allo *scrum master*, supervisiona l'intero processo di DevOps
- l'**automation expert**, dedicato a curare gli aspetti di automatismo
- un **Security Engineer**
- un **Software Developer**, nonché **Tester**
- un **Quality Assurance** che verifica la qualità del prodotto rispetto ai requisiti
- un **Code Release Manager** che si occupa sull'infrastruttura e sul deploy della release

Il DevOps (figura 3.1) si basa su sei principi base:

1. **Customer-Centric Action**, ovvero il committente è al centro dell'azione
2. **End-To-End Responsibility**, ovvero il team gestisce interamente il prodotto, avendone responsabilità totale
3. **Continuous Improvement**, ovvero cercare continuamente di migliorare senza sprechi il prodotto finale e i servizi
4. **Automate everything**, ovvero cercare di automatizzare l'intera infrastruttura di processo, dalle attività di testing e integrazione fino ad arrivare alla costruzione della release e del deployment
5. **Work as one team**, ovvero unificare tutti gli aspetti sotto un unico team o comunque con due team che collaborano fortemente come se fossero uno
6. **Monitor and test everything**, ovvero testare e monitorare costantemente il prodotto

Il quarto e il sesto punto sono i due punti tecnici principali.

Figura 3.1: Rappresentazione famosa del *lifecycle* di DevOps, con le due parti, con le relative parti fondamentali, di sviluppo e operation distinte dai colori ma unite in un singolo metodo unico

3.1 Build, Test e Release

Partiamo dalla parte "dev" di DevOps.
Bisogna pensare a questi step in ottica di automatismo vicina al DevOps.
Innanzitutto bisogna introdurre i sistemi di **version control**, in primis **Git**, un sistema di version control **distribuito**, dove ogni utente ha una copia della repository (con la storia dei cambiamenti), con la quale interagisce tramite *commit* e *update*. Esiste poi una repository lato server per permettere di condividere i vari cambiamenti tramite sincronizzazione.
Git permette anche lo sviluppo *multi-branch* per permettere di lavorare su più branch, pensando allo sviluppo separato dove ogni branch alla fine può essere unito (*merge*) con quello principale (solitamente chiamato *master*). Un altro branch standard è quello di sviluppo, detto *develop*. Un altro ancora è quello detto *hotfix*, per le modifiche più urgenti, che vive in uno stadio intermedio tra i due sopracitati, prima ancora del branch *develop*. Volendo ciascuna feature può essere creata in un branch dedicato. Si procede con le versioni "merge-ndo" quando necessario, "step by step" fino ad un branch *release* per poi andare dopo il testing su *master*. Sul branch *release* possono essere effettuati fix che poi verranno riportati anche in *develop*.
Lo sviluppo su multi-branch si collega alle operazioni di verifica che possono essere attivate automaticamente a seconda dell'evoluzione del codice su ogni branch. Avendo ogni branch una precisa semantica possiamo definire precise attività di verifica, corrispondenti a *pipelines* precise, solitamente innescate da un *push* di codice su un certo branch, in modo sia automatico che manuale. Le pipeline vengono attivate in fase di test di un componente, in fase di creazione di un sottosistema, di assembramento di un sistema intero o di

deployment in produzione. Si hanno quindi quattro fasi:

1. component phase
2. subsystem phase
3. system phase
4. production phase

Spesso le pipelines sono usate come *quality gates* per valutare se un push può essere accettato in un certo branch. Una pipeline può essere anche regolata temporalmente, in modo che avvenga solo ad un certo momento della giornata.

Component phase e Subsystem phase

Dove il fuoco è sulla più piccola unità testabile che viene aggiornata (una classe, un metodo etc...) che non può essere eseguita senza l'intero sistema. In tal caso si può fare:

- code review
- unit testing
- static code analysis

Un cambiamento può anche essere testato nell'ambito del sottosistema di cui fa parte, in tal caso si hanno anche check di prestazioni e sicurezza. Il servizio però potrebbe essere da testare in isolamento rispetto ad altri servizi, usando quindi dei *mocks* o degli *stubs*, ovvero creando degli alter ego dei servizi mancanti in modo che il servizio da testare possa funzionare.

System phase

In questo caso si testa l'intero sistema che viene "deployato" in ambiente di test. Si hanno:

- integration tests
- performance tests
- security tests

Tutti test che richiedono l'interezza del sistema e sono spesso molto dispendiosi e quindi bisogna regolare la frequenza di tali test in molti casi (sfruttando ad esempio la notte).

Production phase

Questa fase è legata alla necessità di creare gli artefatti che andranno direttamente "sul campo", ovvero il deployment in produzione. In tale fase potrebbe essere necessario creare container o macchine virtuali. Si hanno dei check molto veloci sugli artefatti finali (che non siano per esempio corretti), dando per assodato che la qualità del codice sia già stata testata. Si hanno quindi strategie anche di *deployment incrementale*, per cui esistono più versioni del software contemporaneamente con diversa accessibilità per gli utenti finali (accessibilità che viene man mano scalata). In tal caso si usano anche vari tool di monitor. Si hanno anche eventualmente tecniche di *zero downtime* (dove il software non è mai in uno stato *unstable*).

Fasi diverse corrispondono a branch diversi

3.2 Deploy, Operate e Monitor

Studiamo ora la parte "Ops" di DevOps.
Si studia l'evoluzione automatica del software da una versione all'altra in produzione. Avanzare di versione in modo *naive* e istantaneo è troppo rischioso (qualora la nuova versione fosse corrotta non si avrebbe controllo sull'update) e quindi spesso non attuabile (spesso anche per ragioni tecniche). Si ha quindi un insieme di tecniche che si basano in primis sull'*evoluzione incrementale*. Tali tecniche si distinguono in base alla dimensione su cui sono incrementati:

- **Incremental wrt users:** *Dark launching, Canary releases (and User Experimentation)*, ovvero legata agli utenti esposti alla nuova release
- **Incremental wrt requests:** *Gradual upgrades/rollout*, ovvero legata alle richieste per la nuova release
- **Incremental wrt components/replicas:** *Rolling upgrade*, incentrata sulle componenti che vengono aggiornate
- **Non-incremental with backups:** *Green/blue deployment, Rainbow deployment*, non incrementali ma che offrono comunque un backup di sicurezza

Tali schemi possono essere usati in un contesto DevOps.
Per studiare la prima tipologia (*Incremental wrt users*) abbiamo:

- **Dark launching** che arriva dal mondo di Facebook dal 2011. In tale schema l'update è esposto solo ad una parte della popolazione, per la quale viene effettuato il deployment per studiare gli effetti (tramite continuous monitoring) ed eventuali modifiche e migliorie al software, che infine verrà deployato per il resto della popolazione in modo comunque incrementale fino a che l'intera popolazione godrà della feature. Spesso usata per front-end
- **Canary releases**, che studia l'impatto di update relativi al back-end

Tali schemi spesso sono usati di pari passo per le varie sezioni del software, nonché possono essere usati in modo intercambiabile.
Collegato a questi schemi si ha l'approccio basato sull'**user experimentation**, che non è un reale schema di gestione dell'evoluzione del software ma è comunque correlato agli schemi sopra descritti. In questo approccio si studiano diverse varianti del sistema e il loro impatto esponendole agli utenti (perlomeno ad una sottoparte degli stessi in modo incrementale), cercando di capire per l'utente cosa sia meglio e come (si ispira alla *sperimentazione scientifico*). Si hanno quindi più release diverse, per parti di popolazione comparabili, tra le quali si sceglierà la migliore.
Per la seconda tipologia (*Incremental wrt requests*) si ha una divisione a seconda delle richieste fatte dagli utenti (esempio un momento d'uso diverso porta all'uso di componenti diverse), detto *gradual rollout*. Si ha quindi un *load balancer* che permette la coesistenza di due versioni, una nuova e una vecchia, dello stesso servizio. In modo graduale, partendo da pochissime, si passano le richieste alla versione nuova per poter studiare e testare la nuova versione (in caso di problemi il load balancer dirotterà tutte le richieste alla vecchia versione). Alla fine tutto il traffico sarà diretto verso la nuova versione, mentre la vecchia verrà dismessa.
Per la terza tipologia (*Incremental wrt components/replicas*), si ha lo schema del *rolling upgrade*, dove l'upgrade non riguarda un singolo upgrade ma tanti componenti di un sistema distribuito, verificando efficacia di ogni singolo update tramite il continuous monitoring prima di effettuare l'upgrade di un'altra componente. La stessa idea su applica anche a diverse versioni dello stesso prodotto, aggiornandone una prima e poi le altre progressivamente. Le nuove versioni delle componenti "upgradate" devono essere compatibili con quelle ancora prive di upgrade.
Per la quarta tipologia (*Non-incremental with backups*) si ha il **blue/green deployment**, dove vengono isolate due copie della stessa infrastruttura (anche hardware), dove una ospita la versione nuove l'altra la vecchia. Un router ridireziona le richieste degli utenti verso le due unità e quella che ospita la

nuova versione subirà le solite operazioni di test che, se superate, porteranno il router a direzionare verso quella unità, ignorando la vecchia. Se ci sono problemi si fa rollback alla vecchia unità che rimane come backup. Questo schema può essere generalizzato nel **rainbow deployment** dove il momento di coesistenza tra le due versioni (se non più versioni) viene prolungato al fine che vecchie richieste che richiedono una lunga elaborazione vengano elaborate dall'unità vecchia mentre le nuove dall'unità nuova.
In ogni caso le applicazioni devono essere costruite per supportare tutti questi schemi di deployment (a causa di stati delle applicazioni, backward compatibility etc...)

3.2.1 Deployable units

Il caso più tipico in merito alle unità dove fare deployment è il mondo del **cloud**, con **unità virtualizzate e virtual machine (VM)**, dove magari ogni servizio vive in una diversa VM. Si hanno diversi casi in merito a questo tipo di deployment:

- **cloud** basato su **VMs**, dove si ha un'infrastruttura gestita dal cloud provider che gestisce l'hardware e l'hypervisor. Ogni VM, che sono le nostre unità di deployment, ha un sistema operativo arbitrario che lavora con l'hardware mostrato dall'hypervisor. Ogni VM avrà una o più applicazioni e fare deployment porterà all'update di una o più VM. In alcuni casi si fa deployment di intere VM e in altri si modifica il software di una VM già in esecuzione. L'ambiente cloud solitamente è **multi-tenant** (ovvero su una piattaforma unica di un provider si hanno più VM di diverse organizzazioni). Una VM è grossa in quanto contiene un sistema operativo intero e la loro gestione può quindi essere difficoltosa

- **cloud** basato su **containers** che risolvono il problema della grandezza delle VM. In questo caso lo schema è il medesimo ma si ha un **container engine** al posto dell'hypervisor e ogni container non contiene l'intero sistema operativo ma solo il minimo necessario al funzionamento dell'applicazione (il sistema operativo viene condiviso dalla macchina sottostante, riducendo il volume dei singoli containers). In questo caso lo schema di update spesso consiste nel distruggere e ricreare i singoli containers. Anche qui si ha un contesto *multi-tenant*

- **bare metal**, dove i provider offrono direttamente risorse hardware, guadagnando prestazioni ma aumentano anche i costi eco-

nomici , che vengono comunque gestite dal cloud provider. Non si ha virtualizzazione ma accesso diretto alle risorse su cui fare deployment. Questa è una soluzione tipicamente **single-tenant** (sulla macchina gira il software di una sola organizzazione)

- **server dedicati**, un metodo ormai superato con difficoltà causate dall'uso di script, shell e connessione *ftp* completamente autogestiti dall'organizzazione e non da un provider

Il deployment "stile cloud" non è comunque l'unico possibile. Un esempio quotidiano è il deployment di app mobile sui vari store, dove il back-end probabilmente sarà gestito come sopra spiegato mentre l'app in sé viene rilasciata negli store e sarà l'utente finale a fare il deployment installando l'app sul proprio device. Le forme di monitoraggio e di feedback sono spesso diverse e provengono dagli utenti finali stessi.

3.2.2 Monitor

In ambiente cloud ci sono tante soluzioni per il **monitoring**, ad esempio lo **stack di ELK**, formato da:

- **Elasticsearch**
- **Logstash**
- **Kibana**

I dati, ad esempio log o metriche d'uso hardware, vengono raccolti e passano da **Logstash**, finendo in un database, per la memorizzazione di *time series* (serie temporali) di dati (questo in primis per le metriche d'uso che per i log), gestito da **Elasticsearch** e venendo visualizzati da una dashboard grafica, gestita da **Kibana**. Si ha quindi un ambiente di **continuous monitoring**.

3.2.3 DevOps tools

Ogni step del DevOps è gestito tramite moderne tecnologie e tools , con varie alternative per ogni fase (per questo servono figure esperte per ogni step). Vediamo qualche esempio (in ottica più spinta ad un progetto in Java):

- code: Git, Svn, Jira, Eclipse
- build: Apache Ant, Maven, Gradle
- test: JUnit

- release: Jenkins, Bamboo
- deploy: Pupper, Chef, Ansible, SaltStack
- monitor: New Relic, Sensu, Splunk, Nagios

Capitolo 4

Risk management

Partiamo da un semplice esempio:

Esempio 1. *Durante lo sviluppo di un progetto software l'unico dev, insoddisfatto del salario, che conosceva un modulo di importanza critica lascia la società, rallentando in modo serio lo sviluppo del progetto (nonché aumentandone i costi, nel tentativo di cambiare il modulo conosciuto solo dal dev) fino a farlo uscire troppo in ritardo rispetto, ad esempio, alle opportunità di marketing a cui puntava.*

Un esempio del genere non è così raro e il **risk management (*gestione dei rischi*)** si occupa di prevenire questo tipo di complicazioni e fallimenti.

Definizione 1. *Il **risk management** è la disciplina che si occupa di identificare, gestire e potenzialmente eliminate i rischi prima che questi diventino una "minaccia" per il successo del progetto (o anche per eventuali situazioni di revisione del progetto stesso).*

Dobbiamo però dare qualche definizione:

Definizione 2. *Definiamo **rischio** come la possibilità che ci sia un danno.*

Bisogna cercare di prevenire n evento che può portare ad un danno serio (e tanto più è serio il danno tanto è alto il rischio)

Definizione 3. *Definiamo **risk exposure**, che è una grandezza (calcolabile), per calcolare quanto un progetto sia esposto ad un rischio. Viene calcolato come:*

$$RE = P(UO) \cdot L(UO)$$

dove:

- $P(UO)$ *è la probabilità di un* unsatisfactory outcome*, ovvero la probabilità che effettivamente un danno (o comunque un risultato non soddisfacente) sia prodotto*
- $L(UO)$ *è l'entità del danno stesso, ovvero è la perdita per le parti interessate se il risultato non è soddisfacente*

Tanto più un rischio è probabile e tanto più il rischio crea un danno tanto cresce il ***risk exposure****.*

Definizione 4. *Definiamo* ***outcome unsatisfactory (risultato non soddisfacente)*** *come un risultato non positivo che riguarda diverse aree:*

- *l'area riguardante l'esperienza degli utenti, con un progetto che presenta le funzionalità sbagliate, una UI carente, problemi di prestazioni o di affidabilità etc.... In questo caso se i problemi sono gravi si hanno alti rischi, come l'utenza che smette di usare il prodotto, portando al fallimento del prodotto, o anche a conseguenze legali*
- *l'area riguardante i dev, con rischi che per esempio si ritrovano superamento del budget e prolungamenti delle deadlines*
- *l'area riguardante i manutentori, con rischi che per esempio si ritrovano nella qualità bassa di software e hardware*

Se abbiamo un rischio che produce un *outcome unsatisfactory* il primo elemento su cui soffermarsi è lo studio degli eventi che abilitano il rischio, detti **risk triggers**, per evitare che avvengano (comportando di conseguenza che il rischio diventi realtà comportando un *outcome unsatisfactory*, come nell'esempio 1). Sempre in base all'esempio 1 si potrebbe pensare di non sussumere un solo dev con una certa conoscenza o comunque di alzare la paga, per evitare di attivare i *risk triggers*.
Abbiamo quindi due principali **classi di rischio**:

1. **process-related risks**, rappresenti rischi con impatto negativo sul processo e sugli obbiettivi di sviluppo, come ritardi o superamento di costi
2. **product-related risks**, rappresenti rischi con impatto sul prodotto e su obbiettivi del sistema funzionali o meno, come fallimenti riguardanti la qualità del prodotto (sicurezza, prestazioni etc...) o la distribuzione dello stesso

Entrambe le classi possono portare al fallimento del progetto e quindi vanno gestite entrambe.
Bisogna quindi imparare a gestire i rischi. Si hanno principalmente due fasi:

1. una prima fase riguardante il **risk assessment (*valutazione del rischio*)**. In questa fase si hanno:
   - **risk identification**, ovvero l'identificazione dei rischi
   - **risk analysis**, ovvero l'analisi dei rischi identificati (tramite calcolo del *risk exposure* e studio dei triggers)
   - **risk prioritization**, ovvero la prioritizzazione dei rischi analizzati, per focalizzarsi sui più pericolosi per poi scalare ai meno pericolosi

   Alla fine si produce una lista ordinata sulla pericolosità dei rischi
2. una seconda fase riguardante il **risk control** e consiste, in primis, sul **risk management planning**, producendo piani di controllo di due tipi:
   (a) **piani di management**, per la gestione del rischio prima che si verifichino
   (b) **piani di contingency**, per il contenimento di rischi divenuti realtà qualora il *piano di management* fallisca, sapendo cosa fare a priori in caso di emergenza

   Si hanno quindi due sotto-fasi per i due tipi di piani:
   (a) **risk monitoring**
   (b) **risk resolution**

Queste due fasi vengono ciclicamente ripetute durante il ciclo di vita dello sviluppo di un software.

4.1 Risk identification

Si studia come identificare i rischi.
È un'operazione complessa legata alla competenza degli analisti. Un modo comune di farlo è usando delle **check-list**, liste che includono un insieme di rischi plausibili comuni a molti progetti. L'analista scorre tale lista cercando rischi che possono essere applicati al progetto in analisi.

Esempio 2. *Una lista di 10 punti comoda nell'ambito dello sviluppo include:*

1. *problemi con il personale*
2. *budget e deadlines irrealistici*
3. *sviluppo delle funzionalità sbagliate*
4. *sviluppo della UI sbagliata*
5. *gold-plating, ovvero aggiungere più funzionalità del necessario o usare tecnologie troppo avanzate*
6. *continue modifiche ai requisiti*
7. *problemi in componenti software fornite da esterni*
8. *problemi con task svolti da esterni*
9. *problemi con le prestazioni real-time*
10. *voler superare i limiti delle tecnologie attuali oltre le regole e le capacità della computer science*

Alcuni rischi, in astratto, possono verificarsi sempre ma va preso in considerazione solo per **motivi specifici identificabili** nel mio progetto.
Si hanno altri metodi per identificare i rischi:

- riunioni di confronto, *brainstorming* e *workshop*
- confronto con altre organizzazioni e con altri prodotti

4.2 Risk analysis

Per analizzare i rischi si sfrutta esperienza delle reali probabilità che un rischio diventi realtà. Anche in questo si hanno degli schemi su cui basarsi, come *modelli di stima dei costi*, *modelli delle prestazioni*, etc... basati su *simulazioni*, *prototipi*, *analogie con altri progetti* e *check-list* (con stime di probabilità e danno).
Le stime sono molto specifiche sul singolo progetto.
L'analisi dei rischi può anche comportare lo studio delle decisioni da prendere al fine di minimizzare il **risk exposure**, scegliendo o meno tra varie opzioni, scegliendo in modo "guidato" dai rischi. A tal fine si usano i *decision tree* con la radice che rappresenta il problema (un esempio si ha in figura

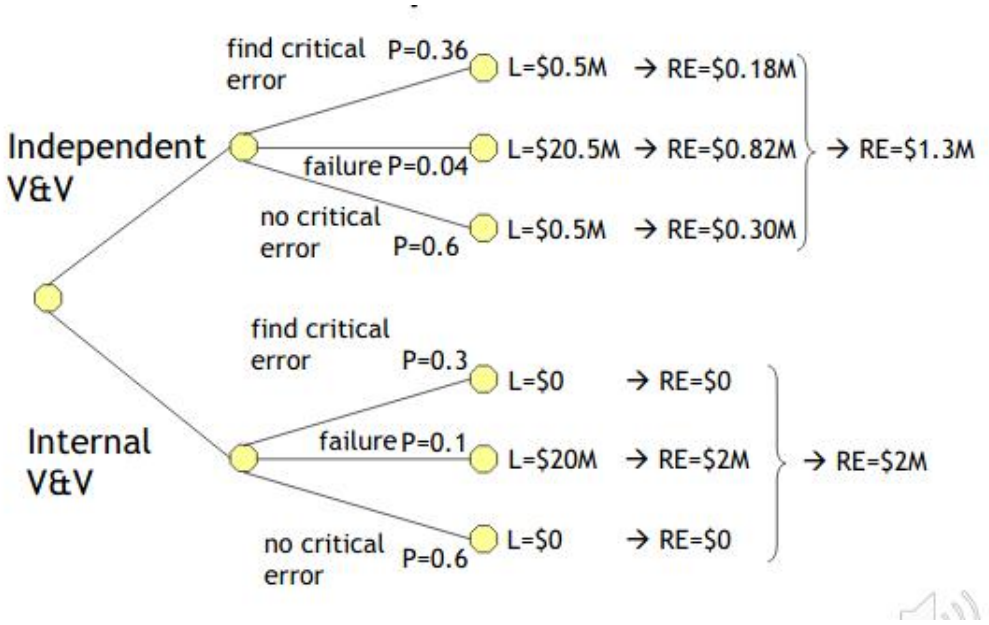

Figura 4.1: Esempio di decision tree

4.1). Si hanno di volta in volta i vari scenari, con le stime di probabilità di trovare un errore critico, di fallimento, di non trovare errori critici etc.... Tali probabilità verranno usate per il calcolo del *risk exposure* insieme ad un quantificatore di $L(UO)$ spesso pari all'effettivo costo che conseguirebbe al risultato ottenuto. Infine i vari *risk exposure* di ogni caso vengono sommati per ottenere il *risk exposure* finale. Si può fare un'**analisi di sensitività** cambiando le percentuali o i costi al fine di ottenere un certo risultato, in modo da capire come si dovrebbe comportare.
Ragionando sulle cause dei rischi usiamo il cosiddetto **risk tree** (esempio in figura 4.2). Questo albero ha come radice il rischio. Ogni nodo, detto **failure node**, è un evento che si può scomporre "via via" in altri eventi, fino alle foglie. La scomposizione è guidata da due tipi di **nodi link**:

1. **and-node**, dove i figli di tali nodi sono eventi legati dal un *and*
2. **or-node**, dove i figli di tali nodi sono eventi legati dal un *or*

Nodi And/Or vengono rappresentati tramite i simboli delle porte logiche.
Dato un *risk tree* cerco le combinazioni di eventi atomici che possono portare al rischio. Per farlo si esegue la **cut-set tree derivation**, ovvero, partendo dalla radice, si riporta in ogni nodo la combinazione di eventi che possono produrre il fallimento e si vanno a calcolare le varie combinazioni degli eventi

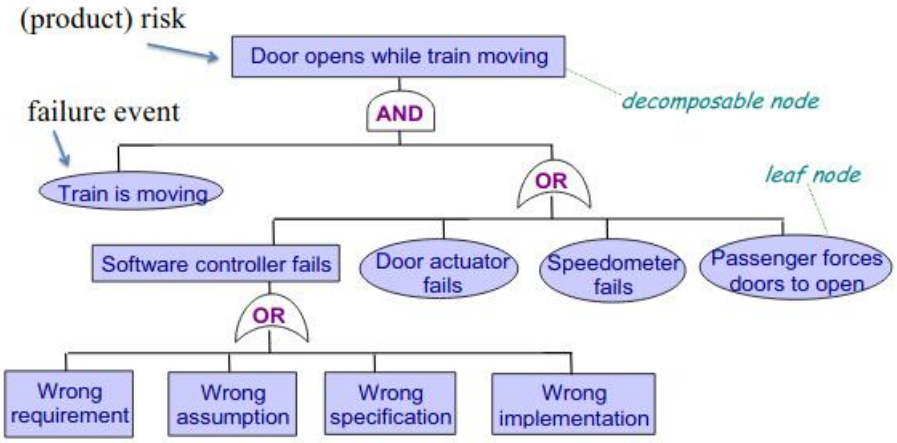

Figura 4.2: Esempio di risk tree

foglia. Praticamente si deriva un insieme di eventi non scomponibili sulle combinazioni dell'*and*.

4.3 Risk prioritization

Bisogna capire quali rischi sono più "rischiosi" degli altri. Per farlo si pongono i valori di $P(UO)$ e $L(UO)$ in un range, per esempio, da 1 a 10, ricalcolando il *risk exposure*. Una volta fatto si lavora in base al *risk-exposure* (che può anche essere un intervallo se le probabilità o i costi sono in un certo intervallo). Si procede plottando i dati su un piano con $P(UO)$ sull'asse delle y e $L(UO)$ su quello delle x e facendo lo *scatterplot* degli eventi (ponendoli quindi come punti in base a $P(UO)$ e $L(UO)$). Qualora i valori siano in un range si rappresentano con un segmento tra i due valori limite di *risk exposure*. Con delle curve posso identificare zone di rischio diverse in base al *risk exposure* per poter catalogare gli eventi. Solitamente alla fase di *risk control* passano una decina di eventi.

4.4 Risk control

Bisogna quindi capire come gestire i rischi.
Per ogni rischio bisogna definire e documentare un piano specifico indicante:

- cosa si sta gestendo
- come mitigare il rischio e quando farlo

- di chi è la responsabilità
- come approcciarsi al rischio
- il costo dell'approccio al rischio

Anche in questo caso ci vengono incontro liste e *check-list* con le tecniche di *risk management* più comuni in base al rischio specifico.
Ci sono comunque strategie generali:

- lavorare sulla probabilità che il rischio avvenga, sulla probabilità dei triggers. Bisogna capire come diminuire la probabilità
- lavorare, nel limite del possibile, sull'eliminazione stessa del rischio
- lavorare sulla riduzione della probabilità di avere conseguenze al danno, non viene quindi ridotto il rischio
- lavorare, nel limite del possibile, sull'eliminazione stessa del danno conseguente al rischio
- lavorare sul mitigare le conseguenze di un rischio, diminuendo l'entità del danno

Bisogna anche studiare le contromisure, da scegliere e attivare in base alla situazione. Si hanno due metodi quantitativi principali per ragionare quantitativamente sulle contromisure:

1. **risk-reduction leverage**, dove si calcola quanto una certa contromisura può ridurre un certo rischio, secondo la seguente formula:
$$RRL(r, cm) = \frac{RE(r) - RE\left(\frac{r}{cm}\right)}{cost(cm)}$$
dove r rappresenta il rischio, cm la contromisura e $\frac{r}{cm}$ la contromisura cm applicata al rischio r. Calcolo quindi la differenza di *risk exposure* avendo e non avendo la contromisura e la divido per il costo della contromisura.
La miglior contromisura è quella con il RRL maggiore, avendo minor costo e maggior efficacia dal punto di vista del *risk exposure*

2. **defect detection prevention**, più elaborato del primo è stato sviluppato dalla NASA. Questo metodo confronta le varie contromisure, confrontando anche gli obiettivi del progetto, in modo

quantitativo facendo un confronto indiretto, producendo matrici in cui si ragiona in modo indipendente sulle singole contromisure e sui singoli rischi ma confrontando anche in modo multiplo.

Si ha un ciclo a tre step:

(a) elaborare la matrice di impatto dei rischi, detta **risk impact matrix**. Questa matrice calcola l'impatto dei rischi sugli obiettivi del progetto. I valori della matrice, ovvero $impact(r, obj)$ (che ha per colonne i rischi r e righe gli obiettivi del progetto obj) variano da 0, nessun impatto, a 1, completa perdita di soddisfazione (e totale non raggiungimento dell'obiettivo indicato). Ogni rischio viene accompagnato dalla probabilità P che accada. Ogni obiettivo è accompagnato dal **peso** W che ha nel progetto (la somma di tutti i pesi è pari a 1). Si possono calcolare altri valori di sintesi. In primis la **criticità** di un rischio rispetto a tutti gli obiettivi indicati:

$$criticality(r) = P(r) \cdot \sum_{obj}(impact(r, obj) \cdot W(obj))$$

La criticità sale se sale l'impatto e se sale la probabilità del rischio.
Un altro dato è la **perdita di raggiungimento** di un obiettivo qualora tutti i rischi si verificassero:

$$loss(obj) = W(obj) \cdot \sum_{r}(impact(r, obj) \cdot P(r))$$

(b) elaborare contromisure efficaci per la matrice. In questa fase si usa il fattore di criticità del rischio. Viene prodotta una nuova matrice con colonne pari ai rischi (con probabilità e criticità) e righe pari alle contromisure. I valori saranno le riduzioni di rischio di una contromisura cm sul rischio r ($reduction(cm, r)$). La riduzione va da 0, nessuna riduzione, a 1, rischio eliminato. Si possono calcolare altri valori di sintesi. Possiamo calcolare la **combineReduction**, che ci dice quanto un rischio viene ridotto se tutte le contromisure sono attivate:

$$(combineReduction(r) = 1 - \prod_{cm}(1 - reduction(cm, r))$$

Un altro valore è l'**overallEffect**, ovvero l'effetto di ogni contromisura sull'insieme dei rischi considerato:

$$overalleffect(cm) = \sum_{r}(reduction(cm, r) \cdot criticality(r))$$

si avrà effetto maggior riducendo rischi molto critici

(c) determinare il bilanciamento migliore tra riduzione dei rischi e costo delle contromisure. Bisogna considerare anche il costo di ogni contromisure e quindi si fa il rapporto tra effetto di ciascuna contromisura e il suo costo e scegliendo il migliore

Analizzando il *contingency plan* viene attuato qualora il rischio si traduca in realtà.
I passa quindi al **risk monitoring/resolution**. Queste due parti sono tra loro integrate. I rischi vanno monitorati e occorrenza vanno risolti il prima possibile. Tutte queste attività sono costose e si lavora su un insieme limitato di rischi, una decina.