2015 41st Euromicro Conference on Software Engineering and Advanced Applications

A Maturity Model for Scaling Agile Development

Igor Stojanov
Dept. of Mathematics and Computer Science
Eindhoven University of Technology
Eindhoven, Netherlands
i.stojanov@tue.nl

Oktay Turetken, Jos J.M. Trienekens
Dept. of Industrial Engineering and Innovation Sciences
Eindhoven University of Technology
Eindhoven, Netherlands
o.turetken@tue.nl, j.j.m.trienekens@tue.nl

***Abstract*—Although the agile software development approaches have gained wide acceptance in practice, the concerns regarding the scalability and integration of agile approaches in traditional system development organizations have remained. The difficulty of adopting agile practices increases when there is a need to scale these practices. Scaled Agile Framework (SAFe) has emerged as a solution to address some of these concerns. Despite few encouraging results of SAFe adoption, case studies indicate several challenges of SAFe adoption. Currently, there is a lack of a well-structured gradual approach for implementing and establishing SAFe. Before and during SAFe adoption, organizations require a uniform model for assessing the current state and progress, and for establishing a roadmap for the initiative. To address this need, we developed a maturity model for adopting agile and SAFe practices. Taking an existing agile maturity model as the basis, we extended the model with practices that are key to scaling agile practices for the SAFe. The model is developed and refined using a Delphi study. Subsequently, a case study was conducted in a large organization where the model was applied to assess the maturity level of the organization in adopting SAFe.**

Keywords-agile software development; scaled agile framework; SAFe; maturity model; Delphi study

I. INTRODUCTION

The software development methodologies have evolved from predictive (such as waterfall) to iterative and incremental (such as spiral, RAD and RUP), and to agile approaches (such as Scrum, and XP). Although a large number of organizations have adopted agile practices, these approaches are often criticized of being applicable primarily to small teams and organizations rather than large enterprises with several hundreds of development teams [1]. In an attempt to scale the advantages of these new methodologies, a new model -commonly known as the Scaled Agile Framework (SAFe), has been developed ([2,3]). The SAFe has gained a rapid attention in the practice and has become an important choice for organizations that are in need of approaches for scaling agile development across the enterprise.

As the popularity of agile adoption increases, the questions organizations ask themselves shifts from why to adopt agile practices to how to adopt and scale these practices. SAFe tends to focus merely on describing the best practices, roles and artifacts of agile and lean principles but makes no attempt to describe any implementation strategy or method [2]. It claims multiple comprehensive and sustainable benefits such as accelerated time-to-market, increased productivity and quality and reduced risks and project costs. Although there are examples of successful SAFe adoptions that argue to have verified some of these benefits, these stories are typically narrowly focused and self-reported. The emerging growth of SAFe in industry and practice requires an academic attention.

Our literature review suggests a lack of a structured roadmap that can guide enterprises on the necessary preparation and adoption for agile and SAFe, highlighting the key components of an implementation strategy. Several studies ([4], [5], [6], [7]) have outlined the risks involved in agile transformations, and others (such as [8], [9], and [10]) have proposed some critical factors and practices, as well as maturity models for agile transformation. But these studies provide very limited contribution in offering a guideline for scaling the agile practices to the enterprise level. There is a need for a more structured approach for adopting SAFe and the scaling of agile in large organizations.

In this study, we introduce the SAFe Maturity Model (SAFe MM) that allows assessing the level of SAFe adoption and provides a roadmap for the implementation of SAFe practices in an enterprise. The SAFe MM was developed as a joint effort of industry experts, using the Delphi technique. To observe the applicability of the model, we applied it in an enterprise to assess the maturity level of the organization in adopting SAFe.

The remainder of the paper is organized as follows: Section II presents a brief overview of the SAFe and a discussion of maturity models for agile development. In Section III, we introduce the SAFe MM together with the elaboration of the method we followed in its construction. Section IV presents the case study and findings regarding the application of SAFe MM in an organization, which is followed by our conclusions.

II. BACKGROUND AND RELATED WORK

A. Scaled Agile Framework (SAFe)

The SAFe is primarily developed for organizing and managing agile practices in large enterprises. It is built upon the current set of practices of agile/lean principles that have been applied in practice by the creators and contributors of the SAFe. Currently, SAFe is being documented in version 3 of the Big Picture Framework and is publicly available.

978-1-4673-7585-6/15 $31.00 © 2015 IEEE
DOI 10.1109/SEAA.2015.29

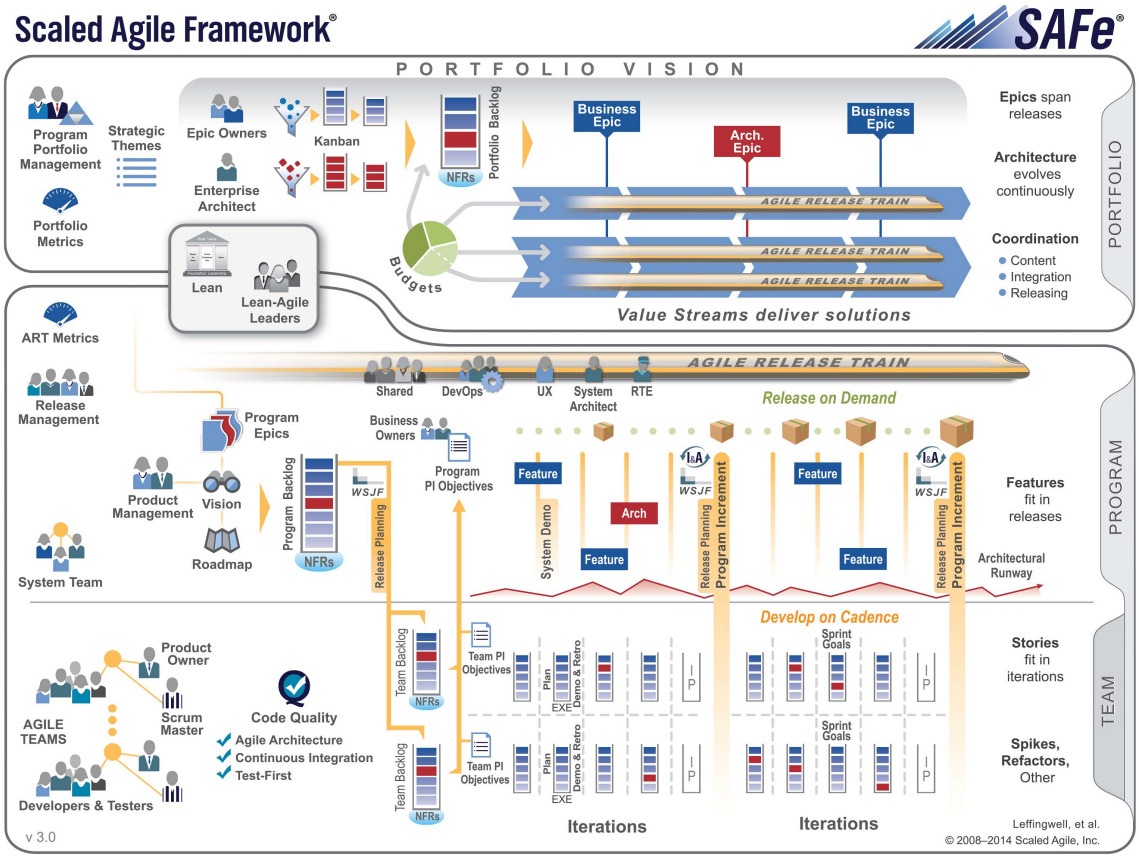

Figure 1. Enterprise Big Picture v3 (Copyright Scaled Agile, Inc. 2014. http://www.scaledagileframework.com. Used with permission. SAFe and Scaled Agile Framework are registered trademarks of Scaled Agile Inc.)

Figure 1. presents the SAFe Enterprise Big Picture, which is a visual representation of the framework to serve as both organization and process model for agile requirements practices [1]. SAFe uses already existing bodies of work in terms of Scrum, XP, Lean, and Product Development Flow. In a nutshell, the framework is separated into 3 levels, namely team, program, and portfolio levels. The boundaries between these three levels are arbitrary and serve as a model for abstraction of scope and scale between levels.

The *team level* of the framework consists of agile teams, which are collectively responsible for defining, building and testing software in fixed-length iteration and releases. The SAFe framework on this level contains a blend of agile project management practices (Scrum) and agile technical practices (XP). For instance, the concept of user stories is borrowed from XP, while sprint planning, daily stand-ups are typical Scrum components. 'Definition of done' and retrospectives are adopted at each iterations. Teams operate on an identical cadence and iteration lengths in order to provide better integration among teams. These agile teams typically consist of 7±2 team members [2].

The primary goal at the *program level* is to organize the agile teams at scale in order to optimize the value delivery of requirements. Furthermore, the program level also aligns the teams with a strategic vision and roadmap for each investment theme. At this level, business and architectural features are defined and prioritized in the program backlog. A major concept introduced at this level is the agile release train (ART), which provides cadence and synchronization. The ART produces releases or potentially shippable increments (PSI) at fixed time boundaries, typically 60 to 120 days [2]. The releases are planned during a two-days release planning event, which involves all relevant stakeholders. Furthermore, a system team is formed in order to establish an initial infrastructure and to support continuous integration and end-to-end testing efforts.

The *portfolio level* is needed for large enterprises, which require governance and management models. The essence lies in achieving a balance between four potentially conflicting goals [11]:

- maximizing the financial value of the portfolio by identifying value streams using Kanban systems,
- linking the portfolio to the strategy of an organization through investment themes,
- ensuring that the scope of activities is feasible by measuring appropriate metrics, and
- balancing the portfolio on relevant dimensions by defining and managing business and architectural epics, which run across value streams.

Since the introduction of the SAFe, a number of companies have applied the framework and published their experiences as technical reports [12], [13], [14]. These reports state improvements in several directions such as higher ROI, 20-30% faster time to market, 40-50% decrease in post release defects, better alignment with customer needs, and increase in productivity of 20-50%, within short periods of time, i.e. in the order of months. The reports also outline the challenges such as staying releasable throughout the development lifecycle due to the late discovery of defects, and defining the right level of requirement detail at the right time during the lifecycle. The studies claim that a proper preparation, orchestration and facilitation of distributed program events are essential for successful release planning. The studies also report with confirmation that geographically distributed teams experience lower productivity due to lack of alignment and solid program execution. However, the fact that these case studies are self-reported by SAFe affiliated practitioners raises questions about the validity and accuracy of the findings.

B. *Maturity Models for Agile Development*

In practice, organizations are unable to fully adopt agile development practices immediately or over a short period of time [15]. Maturity models can guide organizations in providing the directions concerning the practices and the manner that they can be introduced and established in the organization. A maturity model is a conceptual framework that comprises a collection of best practices that help organizations to improve their processes in a particular area of interest [16].

The basic purpose of maturity models is to outline the stages of maturation paths. Based on the assumptions of predictable patterns of organizational evolution, maturity models typically represent theories about how an organization's capabilities evolve in a step-by-step manner along an anticipated, desired and logical path [17]. Maturity models are characterized by a limited and ordered number of maturity levels, and each maturity level defines characteristics or practices which have to be achieved at each level [18].

In a systematic review of literature, Schweigert et al. [19] identifies around forty maturity models for agile development. The analysis concluded that, despite an urgent need for agile maturity models, currently none of the models proposed in the literature is widely accepted in practice and academia. In addition, there is no model that addresses the practices, principles and challenges of scaling agile in large settings.

Ozcan Top & Demirors [20] presents a case study based comparison of nine agile maturity models in terms of their fitness of purpose, completeness, definition of agile levels, objectivity, correctness, and consistency. In their study, they applied each model in a software organization that follows agile practices. Accordingly, the model proposed by Sidky (SAMI - Sidky Agile Measurement Index) scored the highest.

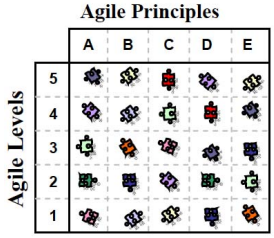

Figure 2. Structure of the SAMI Model [9]. Agile levels are populated with agile practices, which are categorized within agile principles.

The SAMI model is structured into four components: agile levels, agile principles, agile practices and concepts, and indicators [9]. Driven from the four values and twelve principles of the Agile Manifesto, the SAMI model defines five *agile levels*. Collaboration is one of the essential values and qualities of agile and thus it is enumerated as Level 1.

Developing software through an evolutionary approach is the next key principle in agile and the objective at Level 2 in SAMI. Effectiveness and efficiency in developing high quality, working software is the objective at Level 3. The key objective at Level 4 is to gain the capability to respond to change through multiple levels of feedback. Establishing a vibrant and all-encompassing environment to sustain agility is the objective at Level 5.

The SAMI model clusters twelve *agile principles* into five categories that group the agile practices. These principles are: (i) embracing change to deliver customer value, (ii) plan and deliver software frequently, (iii) human-centricity, (iv) technical excellence, and (v) customer collaboration.

Agile practices are a set of techniques or methods that are used for developing software in a manner that is consistent with the agile principles. Each agile level contains practices, which -when adopted- collectively leads to significant improvements in agility. These practices are also categorized under the agile principles. Figure 2. shows the relationship between these components. The model incorporates 38 agile practices in total. Organization should adopt agile practices on lower levels first, since the agile practices on a higher level are dependent on the practices introduced at the lower levels.

Indicators are used to assess certain characteristics of an organization or project, such as its people, culture and environment, in order to ascertain the readiness of the organization or project to adopt an agile practice.

III. THE SAFE MATURITY MODEL AND ITS DEVELOPMENT

After an extensive literature review on the scalability of the agile approaches, SAFe and relevant case studies, and software process oriented maturity models with emphasis on the agile approaches, we took the SAMI agile maturity model [9] as the starting point (mainly due to its comprehensive and well-organized structure, which is also confirmed by the literature [20]). Next, we adapted and extended the SAMI model in accordance with the SAFe principles and practices defined in the main sources of SAFe,

namely [1] and [2]. As a result, 24 SAFe practices were incorporated in the initial design of the model. Three of the original SAMI practices were altered in order these practices to be applicable in the large scale.

The initial model was evaluated and refined through a Delphi study with two rounds of feedback. A panel of industry experts was gathered for the Delphi study in order to elicit domain expertise and facilitate consensus on the domain knowledge and eventually on the maturity model. Online questionnaires were used for collecting data during the two Delphi rounds.

A. Model Development/Refinement using a Delphi Study

The initial maturity model for SAFe was constructed through a synthesis of various concepts and best practices aligned with main SAFe sources (i.e. [2,3]). However, to increase the relevancy and applicability of the proposed model, we performed a Delphi study. The key benefit of this method is that it uses group decision-making techniques while involving experts from the field, which increases the validity of the research [21]. Furthermore, the anonymity of the participants and the discussions resolves the difficulties commonly associated with group interviews (such as the effects of dominant individuals [22]).

The Delphi study that we organized consisted of two rounds, with a timespan between the two rounds of about thirty days. Seven SAFe and agile experts participated in both the first and second round of the study.

An important aspect of the Delphi study is the selection of the panel of experts. The panel typically consists of academic and industrial experts in order to balance the views from both theoretical and practical perspectives. However, due to the high practical relevance of the topic and to the fact that SAFe is a recent development that has not yet gained sufficient attention in the academia, the panel of experts comprised of industry professionals. The experts are agile coaches with multi-year experience in practical implementation of agile practices. Six experts (out of 7) are also SAFe Program Consultants (SPC). SPC are internal change agents in an enterprise, who have domain expertise on SAFe implementation and are qualified to launch such initiatives including training management and practitioners in the organization on SAFe practices. Due to the homogeneity and internal nature of the group, two rounds were deemed appropriate to reach a consensus. The profile of the experts is shown in Figure 3.

1) Delphi Study Round 1

The aim of the first round of the study was to elicit broad comments from the panel of experts. The initial maturity model was presented and a draft report was sent to the participants of the panel. An online questionnaire was used to gather responses to the multiple-choice questions (with 5-points Likert scale) and open-ended questions. The panel members were subject to detailed questions regarding whether the five agile levels are sufficient and if the SAFe and agile practices are complete and aligned appropriately. In cases when experts disagreed, they were asked to elaborate their responses and rationale based on their practical experience and expertise.

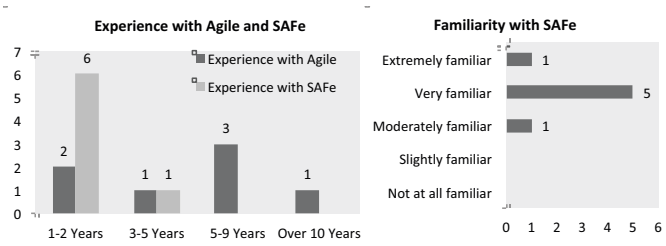

Figure 3. Profile of the experts participated in the Delphi study

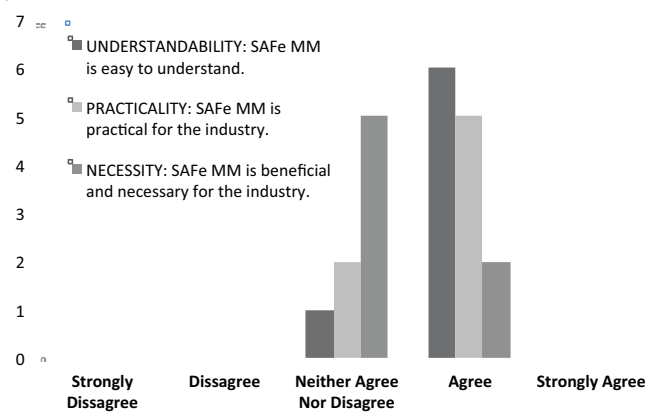

Figure 4. Results of the expert evaluation of the SAFe maturity model

2) Delphi Study Round 2

The primary benefits of multi-round Delphi technique is the fact that the selected Delphi participants are able to reassess their initial views after seeing the results from the previous round and to reach a consensus. Also, since the results are presented anonymously the effects of dominant individuals typically associated with group-based interviewing are diminished. During the second round the expert panel was confronted with results from the first round and requested to reach consensus on the proposed improvements. The improvement revisions were grouped into three categories: fundamental, additive and corrective revisions. Fundamental revisions addressed the essential changes proposed by the experts and used for a further clarification of the existing agile and SAFe practices. Additive revisions covered the additions of SAFe practices, which were proposed during Round 1. The corrective revisions included the corrections of the existing elements, which mainly focused on aligning practices under the correct maturity levels and principles.

Based on the feedback gathered in these rounds, five new SAFe practices were added to the initial model, six were altered and one practice was removed. At the end of the round, the experts were asked to reach consensus on these revisions and evaluate the final version of the maturity model in terms of its necessity, practicality, and understandability. Figure 4. presents the results of the responses gathered from the experts.

Most of the experts agreed that the model has a practical merit and can be used in the industry. The model is also considered as easy to understand and use. However, the

experts were reserved regarding the necessity of the model since five out of the seven experts stated that they neither agree nor disagree that the maturity model is beneficial and necessary for the industry. After the completion of the Delphi study, the descriptions of practices were further refined.

The final version of the SAFe Maturity Model (MM) obtained after the development, refinement and evaluation through the Delphi study is given in TABLE I. The practices from the SAMI that remained unchanged in the SAFe maturity model are displayed in black color (35 practices). The SAFe practices that scale to enterprise level and have been introduced and evaluated with the Delphi study are shown in blue color (in bold) (24 practices). The SAMI practices that were altered in the current model are displayed in red color (in bold-italic) (3 practices). The alteration involved not only content-wise changes for scalability, but also changes regarding the maturity level where the practice is positioned (e.g., practices L1P9 and L2P13 were re-positioned to the lower levels in order to provide a better foundation for the SAFe practices at higher maturity levels.

IV. Case Study

Although our SAFe Maturity Model has been developed as a joint work of expert practitioners, its applicability –as a design artifact- should be evaluated in a business environment.

We applied the SAFe MM in a large international company, which is in transition from a traditional to agile way of developing software. The *objective* of the case study is to employ SAFe MM as an assessment tool to evaluate the current maturity level of the case organization in terms of the agile and SAFe practices, and thereby to investigate the applicability of the model in real-life settings.

TABLE I. SAFe Maturity Model

	Principles				
	Embrace Change to Deliver Customer Value	***Plan and Deliver Software Frequently***	***Human Centricity***	***Technical Excellence***	***Customer Collaboration***
Level 5 Encompassing	• L5P1: Low process ceremony • **L5P2: Continuous SAFe Capability Improvement**	• L5P3: Agile project estimation	• L5P4: Ideal Agile physical setup • **L5P5: Changing organization**	• L5P6: Test driven development • L5P7: No/minimal number of level -1 or 1b people on team • **L5P8: Concurrent testing**	• L5P9: Frequent face-to-face interaction between developers and users (collocated)
Level 4 Adaptive	• L4P1: Client driven iterations • L4P2: Continuous customer satisfaction • **L4P3: Lean requirements at scale**	• L4P4: Smaller and more frequent releases • L4P5: Adaptive planning • **L4P6: Measuring business performance**	• **L4P7: Managing highly distributed teams**	• **L4P8: Intentional architecture** • L4P9: Daily progress tracking meetings	• L4P10: CRACK customer immediately accessible • L4P11: Customer contract revolves around commitment of collaboration
Level 3 Effective	• **L3P1: Regular reflection and adaptation**	• L3P2: Risk driven iterations • L3P3: Plan features not tasks • **L2P4: Roadmap** • **L3P5: Mastering the iteration** • **L3P6: Software Kanban Systems** • **L3P7: PSI/Release** • **L3P8: Agile Release Train**	• L3P9: Self-organizing teams • L3P10: Frequent face to face communication • **L3P11: Scrum of Scrum**	• L3P12: Continuous integration • L3P13: Continuous improvement (refactoring) • L3P14: Unit tests • L3L15: 30% of level 2 and level 3 people	• **L3P16: DevOps (Integrated Development and Operations)** • **L3P17: Vision, features** • **L3P18: Impact on customers and operations**
Level 2 Evolutionary	• L2P1: Evolutionary requirements • **L2P2: Smaller, more frequent releases** • **L2P3: Requirements discovery**	• L2P4: Continuous delivery • ***L2P5: Two-level planning and tracking*** • **L2P6: Agile Estimating and Velocity** • **L2P7: Release planning**	• **L2P8: Define/Build /Test team**	• L2P9: Software configuration management • **L2P10: Automated testing** • L2P11: Tracking iteration progress • L2P12: No big design up front (BDUF) • ***L2P13: Product Backlog***	• L2P14: Customer contract reflective of evolutionary development
Level 1 Collaborative	• L1P1: Reflect and tune process	• L1P2: Collaborative planning	• L1P3: Empowered and motivated teams • L1P4: Collaborative teams	• L1P5: Coding standards • L1P6: Knowledge sharing • L1P7: Task volunteering • **L1P8: Acceptance testing**	• ***L1P9: User stories*** • L1P10: Customer commitment to work with development team

[a] Due to the space limitation, the description of each practice listed in the SAFe Maturity Model could not be provided in this paper. However, a supplementary online document with more information about each practice can be found at: https://drive.google.com/file/d/0B0i4ZY_RKOZ0Z3JFWVpEZE5TTG8)

A. The Case Company

The company chosen for the case study is a corporation headquartered in Europe. It employs more than 115,000 employees and operates in over 100 countries worldwide.

In 2011, the company launched a change program, which aimed at redefining the way the company conducts business through making fundamental changes to its products and processes. As a part of this corporate change program, an initiative in the IT landscape was put in place to introduce new practices in the organization, which laid the foundation for the transition of the company to an agile way of working. The company introduced multi-disciplinary teams, which adopted incremental and iterative approach to software development instead of the traditional waterfall model. Furthermore, the company created output-based partnerships with several partners contributing to the development process. This led to the creation of around 150 Scrum teams, which consisted of internal employees as well as members from the partner companies. The teams are grouped into *programs* (each with varying number of teams; 5 to 15) with respect to the projects and the particular line of business they serve for.

A typical Scrum team consisted of company employees and external (output-based) partners. The partner typically provides the Scrum master and developers, while the company provides the product owner, business analysts and tech leads. The product owner is a non-IT business employee -very close to the company's business problem to be solved.

Given the size of the organization and projects, scaling agile practices across the teams emerged as an immediate need. The majority of these Scrum teams were geographically dispersed due to the location of several output-based partners. In 2013, the company decided to adopt SAFe in attempt to better organize and streamline these Scrum teams. Initially, program managers took the initiative to implement varying aspects of SAFe, which was later taken over by SAFe program consultants who were formally trained on SAFe.

B. Case Conduct: Application of the SAFe MM

For the purpose of assessing the current SAFe maturity, we randomly selected a single program within the company and conducted an assessment among the teams within this particular program. The authors performed the assessment meetings with a Scrum master and a RTE (Release Train Engineer). (RTE can be considered as a 'chief scrum master' who facilitates the program level processes in SAFe.) The RTE's experience and considerations provided input for the existing SAFe practices on the program and portfolio levels of the SAFe, while at the SAFe team level, the Scrum master's experience and reflections were consulted to evaluate the practices employed within the teams in that particular program.

The SAFe practices employed at these levels are assessed using SAFe MM *indicators*. For each practice in the SAFe MM, we developed a set of indicators to appraise certain characteristics of that particular practice. For the SAMI practices that remained unchanged in the SAFe MM, we adopted the original indicators. For instance, one of the indicators used for assessing the 'Roadmap' practice in the SAFe MM Level 3, is as follows:

L3P4-In1: The organization has a program roadmap, which provides a view of the intended deliverables over a time horizon of three to six months.

Based on the current practices applied in the projects, each indicator was rated using a 5-point agreement Likert scale (strongly agree to strongly disagree). Each rating was confirmed either by the RTE or the Scrum master. On the practice level, the results from the indicators were aggregated into an achievement scale: fully-achieved, largely-, partially-, not-achieved (aligned with [23]).

C. Assessment Results and Discussions

The results of the assessment are summarized in TABLE II. The level of achievement tends to decrease towards higher maturity levels. In line with this result, the practices that are '*fully achieved'* are at the lower maturity levels. This result is expected as the SAFe maturity model is developed in such a way that each practice contributes to the foundation required for the practices that are at higher maturity levels. In level 1, for instance, the majority (7 out of 10) of the practices are either fully or largely achieved.

However, the results for the practices that are '*not achieved'* hardly confirm this inclination; that is, the practices that are 'not achieved' are spread over all levels. In order to have a better understanding of the underlying reasons for this situation, we analyzed all the practices that are 'not achieved' in more detail. The practice that was 'not achieved' at level 1 (L1P2) relates to the *collaborative planning*, where all stakeholders are expected to come together during the planning phase. Given that the teams are globally distributed, issues related to the achievement of this practice came as no surprise.

We identified this issue (of a globally dispersed structure) also as an influential factor in many of the upper level practices that failed to achieve their goals. For example, the two practices at level 2 (that the company was not able to achieve) are both related to the planning of the releases. The release planning in SAFe requires all members of the program to come together for synchronization and planning, but having teams (and sometimes team members) located in different time-zones made it challenging for organizing such critical gatherings.

At level 3, this situation manifested itself as a communication barrier between the user representatives (operations) and the development team members, which prevented them to establish a tight integration of development and operations. These themes of weaknesses in the lower maturity levels were indicated as the foremost points on which the company should direct its attention in adopting SAFe.

TABLE II. RESULTS OF THE ASSESSMENT

Level 5 Encompassing	L5P3 (L)	L5P5* (L)	L5P7 (L)	L5P2* (P)	L5P4 (P)	L5P6 (P)	L5P8* (P)	L5P9 (P)	L5P1 (N)
Level 4 Adaptive	L4P2 (L)	L4P5 (L)	L4P6* (L)	L4P7* (L)	L4P8* (L)	L4P9 (L)	L4P10 (L)	L4P1 (P)	L4P3* (P)	L4P4 (P)	L4P11 (N)
Level 3 Effective	L3P7* (F)	L3P1* (L)	L3P5* (L)	L3P15 (L)	L3P2 (P)	L3P3 (P)	L3P4* (P)	L3P8* (P)	L3P9 (P)	L3P10 (P)	L3P11* (P)	L3P12 (P)	L3P13 (P)	L3P14 (P)	L3P17* (P)	L3P18* (P)	L3P6* (N)	L3P16* (N)
Level 2 Evolutionary	L2P6* (F)	L2P8* (F)	L2P9 (F)	L2P4 (L)	L2P11 (L)	L2P1 (P)	L2P3* (P)	L2P5** (P)	L2P10* (P)	L2P12 (P)	L2P13** (P)	L2P14 (P)	L2P2* (N)	L2P7* (N)
Level 1 Collaborative	L1P4 (F)	L1P6 (F)	L1P9** (F)	L1P10 (F)	L1P3 (L)	L1P5 (L)	L1P8* (L)	L1P1 (P)	L1P7 (P)	L1P2 (N)

** SAFe practices that scale to enterprise level and that have been introduced with the Delphi study (pure SAFE Practices)*
*** SAMI practices that were changed and adopted*

	(F) → Fully achieved	**(L) → Largely achieved**	**(P) → Partially Achieved**	**(N) → Not Achieved**	***Total # of practices***
Original SAMI practices	11%	31%	49%	9%	35
*(Pure) SAFe Practices**	13%	29%	42%	17%	24
*Adopted SAMI practices***	33%	0%	67%	0%	3
Total	***13%***	***29%***	***47%***	***11%***	**62**

When the achievement in the original agile practices (that were adopted from the SAMI model as-is) and the pure SAFe practices (that were introduced through the Delphi study) are considered, we cannot distinguish a significant difference. The percentages of practices that were fully, largely, partially and not achieved for these two groups of practices are similar.

The overall assessment results were discussed not only with the company employees that were involved in the assessment but also with the delivery and program managers to confirm the findings. The confirmations validated the ability of the SAFe MM to reveal and pinpoint company's strong and weak points in achieving agile and SAFe practices, and those aspects that require immediate attention.

V. CONCLUSIONS

The challenge in adopting agile practices in developing software increases further with the need to scale these practices to large settings. The SAFe is rapidly emerging as a key industry approach to address these challenges. However, the current wholesale adoption approach of the SAFe is considered risky and complex. Moreover, driven by practitioners' efforts in the industry, there are very few studies about SAFe adoption reported in the academic literature.

In attempt to provide a structured approach to increase the chance of success in scaling agile practices through SAFe, we developed a SAFe maturity model. SAFe MM serves as an evolutionary path that increases organization's SAFe maturity in stages. It can be considered as a *descriptive* model (as opposed to prescriptive) since it describes only the essential practices that organizations should possess at a particular level of maturity.

The model was developed through a Delphi study with the participation of agile/SAFe experts in two rounds of feedback. The panel's central opinion confirms the understandability, practicality and necessity of the maturity model. The Delphi study contributes significantly to the completeness and validity of the model. To address the concerns regarding the applicability of the SAFe MM, we applied it in an international company that is in the process of adopting SAFe. We assessed a single program of the company with respect to the SAFe MM and confirmed the points where company's initial effort should be directed.

This exploratory research addressed the initial challenges of adopting SAFe and measuring SAFe maturity, and should pave the way for future research in this area. Due to the current early state of SAFe itself, there needs to be further exploration of approaches and risks behind SAFe adoption possibly through case studies in several organizations. The constant changes and improvements in the SAFe framework should also be reflected continuously in the SAFe MM.

Our study has several limitations. One limitation concerns the application of the model in the case organization. Due to the organizational constraints, we were able to involve only two members of the company in assessing their level of maturity in applying SAFe. Although these members were highly experienced and were closely and actively involved in the organizational initiative of SAFe adoption, involving more members to gather additional viewpoints on the application of the SAFe MM would be valuable. This might have also allowed us to assess the practices in multiple programs to get a better and more accurate picture of the overall status of the initiative.

There are also limitations pertaining to the use of the Delphi method in constructing the SAFe MM. First limitation concerns the number of Delphi participants. Although the literature does not reach a consensus on the optimal number of subjects in a Delphi study [24], seven members can be considered as limited. We aimed at addressing this issue by bringing together experts that are highly trained and competent within this specialized area of

knowledge. Second, the rounds of feedback were conducted anonymously, which allowed for the panel members to rethink their initial opinions without the influence of the other members. However, the anonymity also can lead to exclusion of group interaction, which in some cases may reduce the accuracy of group judgments [25]. To insure soundness of the gathered data we paid particular attention to panel selection and motivation, questionnaire construction, and for aggregating expert opinion.

Given that the research was done and relates to a single case organization, further studies are required to better assess the applicability, completeness and effectiveness of the maturity model. A complete evaluation of the model's efficacy requires longitudinal studies that relate increasing SAFe maturity with the process and product quality, as well as business performance in quantitative forms.

REFERENCES

[1] D.J. Reifer, F. Maurer, H. Erdogmus, Scaling agile methods, IEEE Softw. 20 (2003) 12–14. doi:10.1109/MS.2003.1207448.

[2] D. Leffingwell, Agile software requirements: lean requirements practices for teams, programs, and the enterprise, Addison-Wesley Professional, 2011.

[3] D. Leffingwell, Scaling Software Agility: Best Practices for Large Enterprises, Addison-Wesley Professional, 2007.

[4] J. Highsmith, Agile: From Rogue Teams to Enterprise Acceptance, The Cutter Business Technology Council, 2006.

[5] B. Boehm, Get ready for agile methods, with care, Computer (Long. Beach. Calif). 35 (2002) 64–69. doi:10.1109/2.976920.

[6] S. Nerur, R. Mahapatra, G. Mangalaraj, Challenges of migrating to agile methodologies, Commun. ACM. 48 (2005) 72–78. doi:10.1145/1060710.1060712.

[7] B. Boehm, R. Turner, Management Challenges to Implementing Agile Processes in Traditional Development Organizations, IEEE Softw. 22 (2005) 30–39. doi:10.1109/MS.2005.129.

[8] M. Pikkarainen, O. Salo, J. Still, Deploying Agile Practices in Organizations: A Case Study, in: I. Richardson, P. Abrahamsson, R. Messnarz (Eds.), EuroSPI 2005, LNCS 3792, Springer Berlin Heidelberg, 2005: pp. 16–27. doi:10.1007/11586012.

[9] A.S. Sidky, A Structured Approach to Adopting Agile Practices: The Agile Adoption Framework. PhD Thesis, Virginia Polytechnic Institute, Virginia Polytechnic Institute, 2007.

[10] Anna Rohunen, P. Rodriguez, P. Kuvaja, L. Krzanik, J. Markkula, Approaches to Agile Adoption in Large Settings: A Comparison of the Results from a Literature Analysis and an Industrial Inventory, in: PROFES 2010, LNCS 6156, Springer Berlin Heidelberg, 2010: pp. 77–91. doi:10.1007/978-3-642-13792-1.

[11] K. Rautiainen, J. von Schantz, J. Va, Supporting Scaling Agile with Portfolio Management: Case Paf.com, in: HICSS, IEEE, 2011: pp. 1–10. doi:10.1109/HICSS.2011.390.

[12] Scaled Agile Framework: Case Studies. http://www.scaledagileframework.com/case-studies/ [5 May 2015], (n.d.). http://www.scaledagileframework.com/case-studies/.

[13] Rally Software Development, Case Study on SAFe Adoption: BMC. https://www.rallydev.com/sites/default/files/Case Study-BMC.pdf [5 May 2015], 2011.

[14] Scaled Agile Inc., Applying the Scaled Agile Framework in Outsorcing Context: Infogain Case Study. http://www.scaledagileframework.com/infogain-case-study [5 May 2015], 2012.

[15] A. Qumer, B. Henderson-Sellers, A framework to support the evaluation, adoption and improvement of agile methods in practice, J. Syst. Softw. 81 (2008) 1899–1919. doi:10.1016/j.jss.2007.12.806.

[16] SEI, CMMI for Development (CMMI-DEV), Version 1.3, CMU/SEI-2010-TR-033. Carnegie Mellon University, Software Engineering Institute (SEI), 2010.

[17] M. Röglinger, J. Pöppelbuss, J. Becker, Maturity models in business process management, Bus. Process Manag. J. 18 (2012) 328–346. doi:10.1108/14637151211225225.

[18] J. Becker, R. Knackstedt, J. Pöppelbuß, Developing Maturity Models for IT Management, Bus. Inf. Syst. Eng. 1 (2009) 213–222. doi:10.1007/s12599-009-0044-5.

[19] T. Schweigert, D. Vohwinkel, M. Korsaa, R. Nevalainen, M. Biro, Agile Maturity Model: A Synopsis as a First Step to Synthesis EuroSPI 2013, CCIS 364, Springer Berlin Heidelberg, 2013: pp. 214–227. doi:10.1007/978-3-642-39179-8.

[20] O. Ozcan Top, O. Demirors, Assessment of Agile Maturity Models: A Multiple Case Study, SPICE 2013, CCIS 349, Springer Berlin Heidelberg, 2013: pp. 130–141. doi:10.1007/978-3-642-38833-0.

[21] J. Osborne, S. Collins, M. Ratcliffe, R. Millar, R. Duschl, What "ideas-about-science" should be taught in school science? A Delphi study of the expert community, J. Res. Sci. Teach. 40 (2003) 692–720. doi:10.1002/tea.10105.

[22] I.I. Mitroff, M. Turoff, Philosophical and Methodological Foundations of Delphi, in: Delphi Method Tech. Appl., 2002: pp. 17–34.

[23] ISO/IEC 15504–2: 2003, Software Engineering – Process Assesment – Part 2: Performing An Assessment, 2003.

[24] C.-C. Hsu, B.A. Sandford, The Delphi Technique: Making Sense Of Consensus, Pract. Assessment, Res. Eval. 12 (2007) 1–8.

[25] F. Woudenberg, An evaluation of Delphi, Technol. Forecast. Soc. Change. 40 (1991) 131–150. doi:10.1016/0040-1625(91)90002-W.