Comments on
Implementation of Additional Export Controls: Certain Advanced Computing Items; Supercomputer and Semiconductor End Use
RIN 0694–AI94
Bureau of Industry and Security, Department of Commerce

January 17, 2024

CSI is the leading U.S. industry association devoted to promoting the trade policy priorities of the U.S. services sector on services and digital issues. Our members include companies that provide financial services, information and communication technology services, telecom services, express delivery and logistics, media and entertainment, distribution, and professional services to all sectors of the economy, including manufacturing and agriculture. CSI members include manufacturers of consumer electronics, telecommunications equipment and health and nutrition products. Our members operate in all 50 states and in nearly 200 countries, representing both large and small firms. CSI also works with counterpart organizations around the globe as co-chair of the Global Services Coalition and the Asia Pacific Services Coalition.

- ***Introduction***

Thank you for the opportunity to provide comment. CSI and its members support the Bureau of Industry and Security's (BIS) objective to protect U.S. national security and appreciate that BIS is carefully considering approaches to expanding AI-related export controls to Infrastructure as a Service (IaaS) due to the complexities of this topic. We hope these comments will aid the agency in achieving its objective.

IaaS is the backbone of the digital economy and the world's technological security infrastructure, and an engine of economic growth for the United States, supporting U.S. jobs and contributing to the U.S. export base. U.S. IaaS providers are global leaders in the IaaS industry, providing the most secure, reliable, and innovative products and services to hundreds of millions of customers around the world. U.S. leadership on IaaS is critical to U.S. national security, ensuring that the United States remains at the forefront of the development of technologies that are central to the world's security infrastructure and global economic growth. To ensure export controls do not impede U.S. economic growth and technological leadership, we urge BIS to craft narrowly tailored controls to address essential and clearly defined U.S. national security interests.

Narrow controls are necessary to avoid a scenario in which new rules undermine the market standing of U.S. providers and undercut the global adoption of U.S. technology. Overly restrictive requirements will only drive IaaS customers to choose non-US providers, which is not in the United States' economic or national security interests. While considering the proposed controls, the US government should simultaneously seek to build strong support for similar actions among allies to ensure effective, broad-based global implementation.

Although the United States is currently a global leader in IaaS, other countries are advancing in the industry, particularly China, which has the largest IaaS providers outside of the United States (e.g., Huawei, Alibaba Cloud, Tencent). Many European and Middle Eastern countries are also seeking to build their own IaaS champions and are seeking to leverage extraterritoriality and continuity of service concerns around U.S. laws—including concerns about loss of access to U.S. technologies due to sanctions—to their own benefit. Overly broad and discretionary export controls would reinforce fears about dependence on U.S. technology companies and put U.S. IaaS providers at a competitive disadvantage.

For the United States to remain a technology leader, it is critical that customers around the world be able to easily access and utilize U.S. IaaS technologies to facilitate their adoption worldwide.

- ***Avoiding duplicate IaaS-related regulations***

Before imposing any new controls related to IaaS, BIS should take into account the Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence ("AI EO"), which establishes new requirements around reporting and identity verification of foreign IaaS customers accessing significant compute capacity and, as such, appears to address the risks BIS seeks to target.

To avoid duplication of effort and establishment of multiple, overlapping regulations, BIS should focus first on the EO and refining and implementing the regulations identified therein related to the use of IaaS products to develop large dual-use AI foundation models, before imposing any new IaaS-related export controls. It will be important, in particular, to observe and assess the implementation and operation of the EO's regulations, both in terms of any difficulties in implementation and in their effectiveness in addressing the national security concerns, before imposing additional regulatory requirements with respect to the same products.

Although IaaS-related export controls would be premature, if BIS nonetheless chooses to pursue them, then any controls should be narrowly tailored, including by applying:

- only to compute power from AI chips that are already subject to export controls (i.e., any forthcoming controls should be consistent with October 7 IFR controls);

- only to compute capacity at the threshold necessary to develop large, highly capable foundation models (e.g. 10^{26} FLOPs within a specified time period);

- interoperably with the AI EO and its forthcoming regulations;

- only to specific entities that pose national security risks;

- any applicable export license or notification requirement should apply only to the IaaS user and not the provider; and

- to both U.S. and non-U.S. companies alike.

We have elaborated further on some of these points below.

- ***Controls on underlying semiconductors***

The scope of any IaaS-related controls should be tied directly to the semiconductors that are subject to the Oct. 7, 2023 controls, namely the provision of computational capacity from integrated circuits subject to the ECCN 3A090 controls. This approach supports BIS's objective of ensuring that IaaS solutions do not undermine the effectiveness of the October 7 IFR controls, without restricting the ability of U.S. IaaS providers to provide global customers with access to less powerful chips that do not present national security concerns.

- ***Relevance of compute capacity as threshold for controls***

Any IaaS-related controls should apply to use of controlled AI chips at a compute capacity equal to the threshold the Secretary of Commerce sets pursuant to the AI EO (currently greater than 10^26 FLOPS) within a limited period of time. This is the best measure currently available because model capabilities scale in relation to the amount of compute (and data) used for training and the 10^26 FLOPs threshold, which has already been adopted by the AI EO, focuses on next generation models most likely to raise national security concerns. Access to the relevant compute capacity must be contained to a limited period of time because otherwise the control would inadvertently capture use cases where a customer crosses the 10^26 FLOPS threshold after using compute power for an extended period for purposes other than training an AI foundation model.

As noted below, any applicable notification requirement for access to compute capacity should be on the IaaS user and not the IaaS provider. However, to the extent IaaS providers have a role in any notification or licensing regime -- i.e., BIS issues best practice guidance that IaaS providers should confirm that a customer has notified BIS -- then compute capacity is again the most effective threshold for IaaS-related controls, because IaaS providers can determine how much compute capacity a customer is using.

Controls based on the amount of training, parameter count, or capabilities of a model will not be effective because they require knowledge of the specifics of the model development and evaluation process, which an IaaS provider will not have. IaaS providers do not have visibility into the models that customers are developing on their infrastructure, as the IaaS provider has no access to customer workloads and customers content is encrypted to protect customer security and data privacy. Moreover, IaaS customers consider information about the amounts or types of training data being used, the number of parameters, and the methods of training AI models to be sensitive, proprietary information. IaaS customers also have a strong commercial interest in limiting the ability of IaaS providers to access their customer data. This is due to a general interest in protecting sensitive corporate data as well as to potentially comply with other privacy laws and regulations.

Any regulations should also avoid relying on characteristics that cannot be determined until after a model has been trained. IaaS providers—indeed, even the customers themselves—cannot be certain what levels of performance a model will exhibit at particular tasks until after training is complete. IaaS providers cannot know in advance of providing services how many operations will be used in training the customer's model. The number of chips used by the customer in training has some impact on this, but key determinative factors like training time and training efficiency are not visible to IaaS providers until after training.

- ***Value of list of restricted entities***

To the extent that BIS applies any restrictions on IaaS providers, they should apply equally to U.S.-based or non-U.S. based IaaS providers. If export controls apply only to U.S. IaaS providers, it will put U.S. industry at a critical disadvantage versus foreign IaaS providers, which are primarily Chinese. If future IaaS-related controls applied to a similarly broad set of countries and entities as the AI chip controls, the global chilling effect on using U.S. IaaS providers would be extensive, and accelerate other countries' industrial policies to build their own IaaS champions. It would damage U.S. IaaS providers by potentially restricting their ability to work with a wide variety of customers globally and impeding delivery of services while providers conduct extensive due diligence on customers' ownership structure that is unlikely to identify restricted entities. If potential customers face delays or hurdles in accessing U.S. technology, they will simply look elsewhere, including to Chinese or other foreign providers. Overly broad controls will also drive away customers due to concerns that their access to U.S. IaaS providers will be restricted. Even the perception that global access may be restricted will result in customers moving away from U.S. IaaS providers, rather than take the risk that they will lose access to critical IaaS technologies. To maintain U.S. technological leadership and ensure the IaaS industry remains an engine of economic growth in the United States, U.S. IaaS providers must be able to offer their innovative and cutting-edge technologies to customers worldwide with minimal friction in order to facilitate the global adoption of those technologies.

Rather than maintain broad, country-wide prohibitions on remote access, we recommend that BIS develop more specific list of entities subject to controls, with a focus on organizations located in D:5 countries such as those on the Entity List or military end users. A list would also facilitate U.S. IaaS providers' compliance with controls, as providers could use standard screening and other compliance measures to quickly identify whether a customer is on the list and subject to controls. In order for BIS to achieve its national security objectives, it is crucial to issue regulations with which IaaS providers can develop mechanisms to comply, rather than imposing overly broad controls that providers cannot effectively implement.

- ***Any controls should apply to IaaS customers, not providers***

With regard to BIS question 1, IaaS providers typically do not and should not monitor their customers. Any regulations in this area should take into account the significant challenges that IaaS providers would face if asked to identify whether a customer has developed a "dual-use AI foundation model."

While IaaS providers are aware of the identity of their customers, the IaaS providers do not have visibility into customers' activities or into the details of the model itself, including how it is trained, what its capabilities are, or how the customer intends to use it, which will be critical for BIS's ability to determine whether the model presents a threat. Only the customer has and can submit this information to BIS.

If BIS imposes export licensing or notification requirements, the IaaS provider's customer – and not the IaaS provider - should be responsible for notifying/obtaining a license from BIS prior to accessing restricted computational capacity. This position is consistent with the Export Administration Regulations (EAR) and BIS's Cloud Advisory Opinions, which establish that provision of computational capacity is not

an export and an IaaS provider is not the exporter when providing computational capacity.[1] This framework recognizes the practical realities of the provision of computational capacity and will be the most effective way to achieve the U.S. government's goal of identifying dual-use models that raise national security concerns.

- ***Equal applicability of controls to U.S. and foreign entities***

As explained above, and consistent with BIS's longstanding interpretations, the IaaS provider is not the exporter and therefore any export restriction should apply to the customer of the IaaS provider. However, to the extent that BIS applies any restrictions on IaaS providers, they should apply equally to U.S.-based or non-U.S. based IaaS providers. If export controls apply only to U.S. IaaS providers, it will put U.S. industry at a critical disadvantage versus foreign IaaS providers, which are primarily Chinese. This will undermine both U.S. industry and the U.S. Government's national security goals, as customers will simply move to Chinese and foreign IaaS providers in order to develop models without any constraints.

- ***Applicability of deemed exports and deemed reexports***

Requiring a license for deemed exports and deemed reexports would significantly reduce U.S. technology companies' ability to recruit and develop both U.S. and non-U.S. talent. In many instances, the internal security and trade-secret protection procedures of companies working with the restricted semiconductor technology may be as, if not more, effective than deemed export licensing to protect against the risk of diversion. Companies invest significant resources in the development of technologies, and it is in the companies' economic interest to ensure proprietary technologies remain protected from third parties.

The need to obtain and renew deemed export licenses also creates uncertainty for medium- and long-term planning for projects subject to heightened controls. Given the lengthy processing times for the US government to review deemed export license applications and, in the case of approved licenses, determine the conditions that should attach to the license, employers must decide to either push back start dates or allow the employee to start work without being able to perform their intended responsibilities. Individuals who face the uncertainty of whether or when they will be able to perform the work they were hired to do may choose to forgo working on export-controlled projects that contribute to US leadership in important fields.

Top U.S. scientists and engineers want to collaborate with the best scientists and engineers from around the world, regardless of nationality or location. Placing restrictions on their ability to do so would discourage top U.S. talent from accepting positions in the United States or with U.S.-based companies, sending that talent offshore where these restrictions do not apply to foreign companies. There is already an existing shortage of talent that can develop cutting edge semiconductors. Deemed export or deemed reexport restrictions would further exacerbate the issue, increase the cost of hiring talent, and delay projects that are already underway. In short, such restrictions damage companies' ability to compete in the open global marketplace for top talent, undermine U.S. leadership in the semiconductor technology, and negatively impact the U.S. economy.

- ***Designed or marketed for datacenters***

[1] Application of EAR to Grid and Cloud Computing Services at 2-3 (Jan. 13, 2009); *see* 15 C.F.R. § 734.13.

In order to use License Exception NAC for ECCN 3A090.a, an entity must first determine whether the chip is "designed or marketed for use in datacenters." For companies reselling chips made by other companies, it is not possible or practical to know whether a chip is "designed or marketed for use in datacenters," as these companies only market the chips. To address this concern, we suggest that BIS modify the following 3A090 subparagraph by adding the language in ***bold*** to explicitly assign an ECCN for items that are designed or marketed for use in datacenters. This approach will allow manufacturers to use the ECCN to communicate the correct level of control and identify the appropriate compliance requirements within automated systems. For example, if a manufacturer used proposed ECCNs 3A090.a.1.a and 3A090.a.1.b and communicated those to a reseller, the reseller would be able to easily determine if the item requires a license or is NAC eligible, respectively.

a. Integrated circuits having one or more digital processing units having either of the following:

a.1. a 'total processing performance' of 4800 or more ***and meeting the following, or****:*

a.1.a designed or marketed for use in datacenters

a.1.b designed or marketed for use for any application other than those identified in a.1.a

a.2. a 'total processing performance' of 1600 or more and a 'performance density' of 5.92 or more.

- ***Clarifying definition of "headquartered in"***

"Headquartered in" is a vague and overly broad concept. For example, some companies claim to have multiple headquarters or no headquarters. Other companies may be joint ventures and have two independent parent company structures with different headquarters. Still other companies are ultimately owned by holding companies incorporated in a jurisdiction in which they don't have real operations. Such ambiguities create difficulties when determining whether a license is required.

Applying controls so broadly risks undermining U.S. companies as reliable suppliers of strategic technology globally and driving customers to foreign suppliers, many of which are Chinese. It is also ineffective because detailed ownership information on non-public companies is generally not publicly available, particularly if an entity is set up for the purpose of circumventing export controls. It will be extremely difficult for U.S. companies to determine whether a customer is "headquartered in" a D:5 country.

To facilitate compliance with restrictions, BIS should also provide and continually update a list of entities that meet this criterion, which will allow companies to use standard screening processes to quickly identify customers subject to restrictions. For example, BIS could narrowly define the term "headquartered in" in the EAR to target only entities wholly or majority-owned by a Chinese entity.